ISSN: 1696-2281
SORT 29 (1) January - June 1-138 (2005)

SORT

Statistics and Operations Research Transactions

Sponsoring institutions

Universitat Politècnica de Catalunya
Universitat de Barcelona
Universitat de Girona
Universitat Autònoma de Barcelona
Institut d'Estadística de Catalunya

Supporting institution

Spanish Region of the International Biometric Society

Generalitat de Catalunya
Institut d'Estadística de Catalunya

SORT

Volume 29 (1), January-June 2005

Formerly Qüestiió

Contents

SORT 29 (1) January-June 2005, 1-10

Graphical display in outlier diagnostics; adequacy and robustness

Nethal K. Jajo*

University of Western Sydney

Abstract

Outlier robust diagnostics (graphically) using Robustly Studentized Robust Residuals (RSRR) and Partial Robustly Studentized Robust Residuals (PRSRR) are established. One problem with some robust residual plots is that the residuals retain information from certain predicated values (Velilla, 1998). The RSRR and PRSRR techniques are unaffected by this complication and as a result they provide more interpretable results.

MSC: 62-09, 62G35, 62J05, 62J20

Keywords: Masking; outlier; robust diagnostics; robust residuals; swamping.

1 Introduction

Graphical methods play an important role in fitting linear models in general and in outlier detection in particular. The informal graphical display and the formal testing procedures (numerical display) are complementary in outlier detection, and we emphasize that it should be implemented together. The informal graphical display is more useful than formal testing procedures since: first, the patterns of the residuals, the graphical displays dependant on them are often more informative than their magnitudes (Atkinson, 1985). Second, truth in the saying "one picture is worth thousand words". Third, a heavy computational requirement for numerical display is required, and in some cases the numerical display may fail to identify the potentiality of outlier observations.

* *Address for correspondence:* Dr. Jajo, N.K. University of Western Sydney, Blacktown Campus, Locked Bag 1797, Penrith South DC NSW 1797, Australia. Email: n.jajo@uws.edu.au

Received: October 2003
Accepted: June 2004

Several books devoting large portions to graphical presentations have been published, among them: Belsley, *et al.* (1980), Cook and Weisberg (1982), Atkinson (1985), Chatterjee and Hadi (1988), Barnett and Lewis (1994) and Hocking (1996).

Associating with others, we think robust residuals' plots are among the important graphical techniques in outlier detection. Many authors advocated robust residuals plots, among those are Rousseeuw and Leroy (1987). Different approaches are proposed by Velilla (1998), based on M-estimators using Huber's ψ function (Huber, 1973), GMM: three-step estimator of Simpson, *et al.* (1992) and GMS one-step estimator of Coakley and Hattmansperger (1993), who notices; "for high breakdown robust estimators, the residuals retain information on the regressor (here are X_i's) and this might complicates the interpretation of residuals plots". We are in agreement with others to a degree, but that is not the case when we are using RSRR or PRSRR plots in outlier detection.

The unaffected performance of using RSRR in outlier detection by the problem mentioned above will be shown through sections 2 – 4. Section 2 introduces RSRR and PRSRR, while Section 3 illustrates how to use RSRR /PRSRR plots and PRSRR probability plots in outlier detection. Real data examples will be given in Section 4, and Section 5 ends with a short conclusion.

2 Robust residuals

Consider the general linear regression model

$$Y_{n\times 1} = X_{n\times k}\beta_{k\times 1} + \epsilon_{n\times 1}, \tag{1}$$

where ϵ's have zero mean and variance matrix $\sigma^2 \times I_n$ (σ is unknown). The ordinary least squares (OLS) estimators $\hat{\beta}$ of β could be obtained by the minimization of

$$R(\hat{\beta}) = \sum \hat{r}_i^2 \quad \text{where} \quad \hat{r}_i = y_i - x_i^T\hat{\beta}. \tag{2}$$

While the robust estimators $\tilde{\beta}$ could be obtained by the minimization of

$$D(\tilde{\beta}) = \sum \rho(\tilde{r}_i/\sigma), \quad \tilde{r}_i = y_i - x_i^T\tilde{\beta}, \tag{3}$$

where ρ is a propriety function.

It's clear from equations 2 and 3, the robust estimators depend on the value of σ (while OLS does not). For this reason and other reasons given by Jajo (1999), we suggest using RSRR (robust residuals studentized by their robust scale estimators) and PRSRR for outlier detection. Moreover, Velilla's note (1998) is of relevance when we replace plotting robust residuals against regressors by: plotting RSRR against probability of each residuals (RSRR probability plots), plotting PRSRR against regressors (PRSRR plots) or plotting PRSRR against probability and regressors in three dimensions. This

mostly avoids (in fact is unaffected by) the problems of masking, swamping, or complicating the interpretation of residuals' plots, when residuals from high breakdown robust estimators retain information on the regressor variable, as mentioned by Velilla (1998) and Mckean *et al*. (1993, 1994).

2.1 RSRR

To obtain robust estimators; for the simple case of model 1, we use Theil-type method NMK; originally, the MK method was developed by Hussain and Sprent (1983) and Jajo (1999) proposed a modification termed NMK. Jajo's estimators are robust with high breakdown point (= 0.5) and can be defined as follows (Jajo, 1999):

$$\tilde{\beta} = \text{med}\{a_{i,i+m}\} \quad i = 1, 2, \ldots, m,\ m = n/2 \text{ if } n \text{ is even and } (n-1)/2 \text{ if } n \text{ is odd,}$$

where $a_{i,j} = (Y_j - Y_i)/(x_j - x_i)$, for $1 \le i < j \le n$, and $x_i \ne x_j$. The intercept parameter α can be estimated as $\tilde{\alpha} = \text{med}\{y_i - \tilde{\beta} x_i\}$.
In case of multiple linear regression (model 1) and to obtain robust estimators, we apply orthogonal modified Theil method based on NMK set of $a_{i,i+m}$ elements, for more details see Jajo and Wu (1998) and Jajo (1999) . For both cases, scaling robust residuals $\tilde{r}_i = y_i - x_i^T \tilde{\beta}$ by dividing them by corresponding s_i, we obtain RSRR as:

$$e_i = \tilde{r}_i / s_i, \tag{4}$$

where $s_i = s\sqrt{1 - p_{ii}}$. We recommend using s equal to the interquartile range (IR) since it achieves better results in robustness, as mentioned by Kianifard and Swallow (1989, 1996) in their comparison study of using variety of s as a robust estimate of the dispersion. For simple linear regression $p_{ii} = \frac{1}{n} + (x_i - \bar{x})^2 \Big/ \sum_{i=1}^{n} (x_i - \bar{x})^2$ and $p_{ii} = x_i^T x_i \sum \lambda_r^{-1} \cos^2 \theta_{ir}$ for multiple case where λ_r is the r^{th} eigenvalue of $X^T X$ and θ_{ir} is the angle between x_i and the r^{th} normalized eigenvector of $X^T X$ (Chatterjee and Hadi 1988).

2.2 PRSRR

Suppose that

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik}$$

and

$$\tilde{y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 x_{i1} + \tilde{\beta}_2 x_{i2} + \cdots + \tilde{\beta}_k x_{ik},\ i = 1, 2, \ldots, n$$

are robust estimates for y_i. It follow that $\tilde{y}_i - \tilde{\beta}_\ell x_{i\ell}$ is a robust estimates of the i^{th} response

when all the predictors except the ℓ^{th} one ($x_{i\ell}$) are used. Hence the partial residuals at trial i can be defined as (du Toit *et al.* 1986):

$$\begin{aligned} p\tilde{r}_{i\ell} &= y_i - (\tilde{y}_i - \tilde{\beta}_{i\ell} x_{i\ell}), \qquad \ell = 1, 2, ..., k. \\ &= \tilde{r}_i + \tilde{\beta}_\ell x_{i\ell}. \end{aligned} \tag{5}$$

To obtain PRSRR $d_{i\ell}$ we scaled $p\tilde{r}_{i\ell}$ by the same way given in the preceding definition of RSRR.

3 Graphical display

An advantage of graphical display is that it can exhibits the effect of each observation, but some disadvantages occur in high breakdown robust residuals' plots as mentioned here by Velilla (1998). Suitability of a plot for a specified purpose could be the key to overcome the problem. Hadi (1993) states "the focus here is not on how the graph is constructed but rather on (a) what to graph, (b) what information can be extracted from a graph". Two-and three-dimensional graphs will be the main category of our graphical display. We used 2-D graphs as a way to compare the graphs of Velilla (1998) and Rousseeuw and Leroy (1987). We use 3-D graphs for explanation and confidence because software has made it possible to rotate 3-D plots and the choice of rotating position could be of interest.

3.1 RSRR Probability plot

In this plot, the probability (x-axis) axis is with $(i - .5)/n$ scale, where i (in 2-D) is the index of RSRR e_i, after ordering them in ascending magnitude. The y-axis is the RSRR e_i. In 3-D the i index is for x_i after ordering them also in ascending magnitude and represented as the x-axis; y-axis is the probability, and the z-axis is the e_i.

3.2 PRSRR and PRSRR Probability Plots

Plotting PRSRR as y-axis against corresponding x_i as x-axis will obtain PRSRR plot, while PRSRR probability plot is obtained by plotting PRSRR against corresponding x_i after ordering them in ascending magnitude. Probability axis is in the same scale as that in RSRR probability plot in 3-D.

For the two kind of plots, outliers will be recognized as the points that deviate markedly from the pattern of the whole observations. We emphasize that the graphical display is not enough for outlier detection, but it must be accompanied by the numerical display, which we do not present here to go along with Velilla's paper that is confined to the graphical display. For more details of using the two displays in outlier detection see Jajo (1999).

4 Real data examples

To illustrate the fact that the high ability of using robust residuals (RSRR or PRSRR) in outlier detection is unaffected by the retaining information on the regressor variable, possibly caused by the robust residuals, as Velilla (1998) notices. And to be nearer to Velilla's note, we make use of the two examples he used in his paper, and we add another well-known example for more explanation and confidence since it contains multiple outliers.

4.1 Gesell adaptive score data

These data are given by Mickey, Dunn and Clark (1967), and have been analyzed extensively in the statistical literature. These data contain 21 observations, with y regressed on x simply. Observation number 19 was regard as outlier. Figure 1: RSRR probability plot in 2-D marks observation 19 as outlier easily and so does RSRR probability in 3-D. Figure 3 of Velilla (1998) shows masking and swamping problems in RGMM-X plot and RGMM-INDEX plot, which confirms the improved achievement of our graphs.

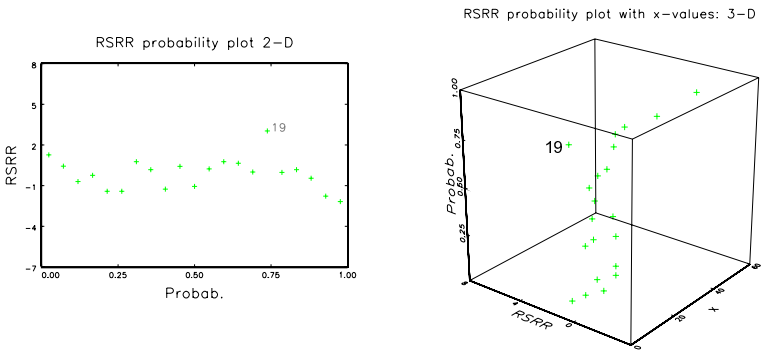

Figure 1: *RSRR probability plots for Gesell data.*

4.2 Salinity data

Ruppert and Carroll (1980) give this set of data, which contains 28 observations, three regressor variables and certainly two real outliers (observations number 5 and 16). Carroll and Ruppert (1985) and Atkinson (1985) perform residuals analysis for different models to fit salinity data and confirm certainly observations 5 and 16 as outliers. Rousseeuw and Leroy (1987) plot standardized LMS (Least Median of Squares) against estimated response and they mentioned that the horizontal band of points that must contain the data but not outliers is structureless. Velilla (1998) associated observation 16 as outlier, and he only recognized that case 5 has a large positive residual.

We agree that the salinity data is extremely complex, but still one of the better examples for masking, and there are a variety models that can be used to fit these data. Among other researchers, and for diagnostic purpose, we suggest our model to regress the water flow on bi-weekly average salinity, the salinity lagged two weeks and the trend. Our chosen model may not be the best possible one, but it is chosen because of its simplicity, going smoothly with our diagnostic and association with others in their results.

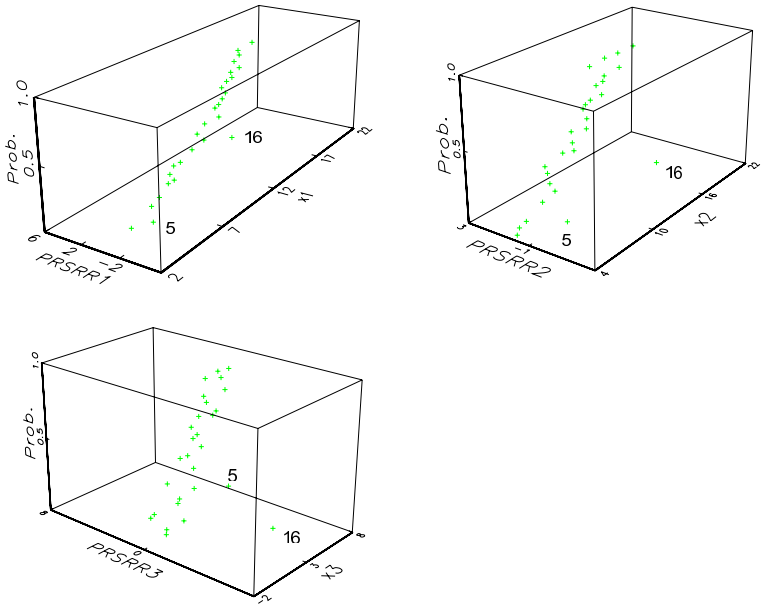

Figure 2: *PRSRR probability plots for Salinity data.*

From Figure 2, we can distinguish observations number 5 and 16 from others since it locate far from the main body of the observations, so regarded as outliers. The same could be mentioned for these observations in Figure 3. Both figures achieve better than Rousseeuw and Leroy (1987): Figure 4, page 84 and Velilla (1998): Figure 4 in detecting outliers without any problem of masking and swamping. Rousseeuw and Leroy (1987): Figure 4, shows swamping while Velilla (1998): figure 4, shows masking and swamping at plots of RGMM-INDEX, RGMM-X_1 and RGMM-X_2

4.3 Esoteric example (Dilemma data)

This set of data was used by Hocking and Pendleton (1983) and others. It contains 26 observations with 3 regressor variables, a constant term and 3 outliers (observations number 11, 17 and 18). Figure 4 uses PRSRR against corresponding regressor variables x_i. The three plots flag the outliers clearly without any problems. Moreover PRSRR probability plots in figure 5 show the same thing in 3-D. Using probability, the PRSRR

has been hung from PRSRR-x_i plane, and those outliers are located far from the shell of the other observations.

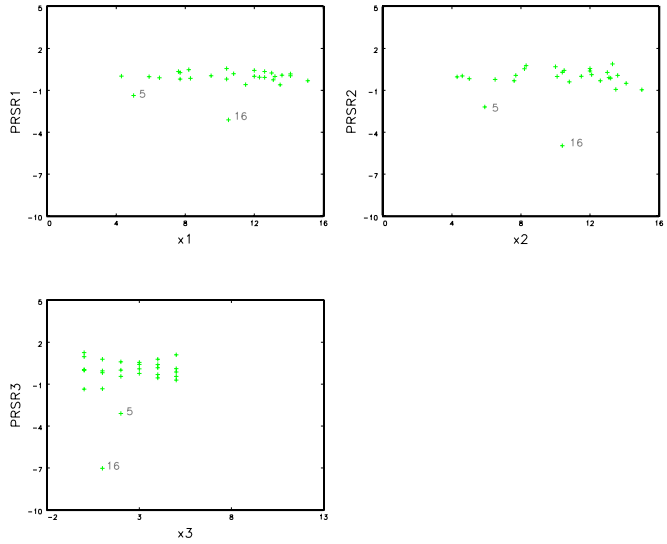

Figure 3: *PRSRR plots for Salinity data.*

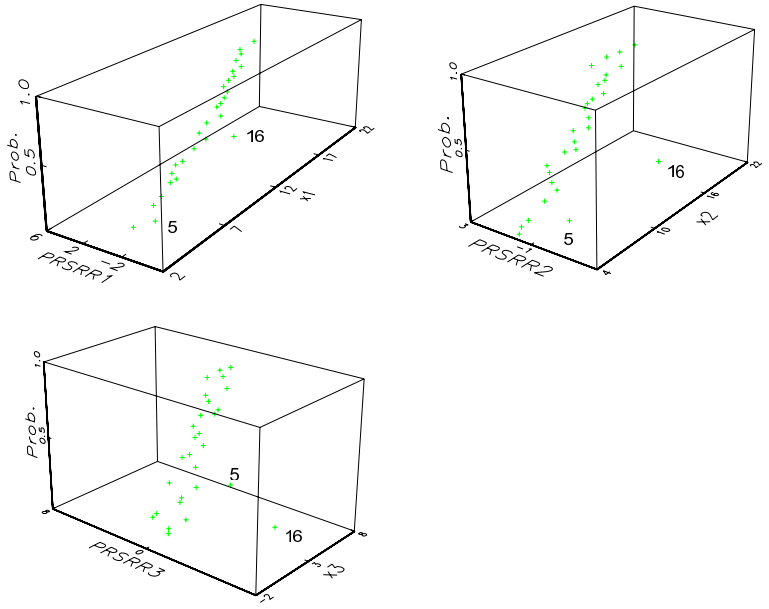

Figure 4: *PRSRR probability plots for esoteric data.*

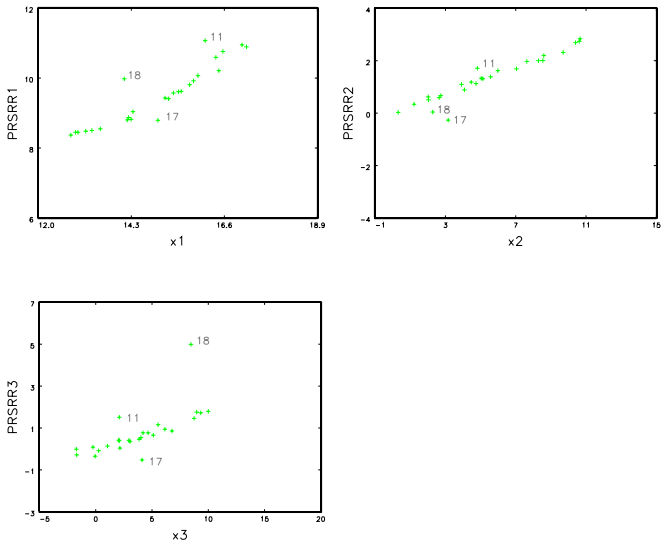

Figure 5*: *PRSRR plots for esoteric data.

5 Conclusions

Robust residuals might retain harmful information on regressor variables. But still those residuals play an important role in robust outlier diagnostics, specially when we are using RSRR or PRSRR through graphical or numerical display. The PRSRR and RSRR, PRSRR probability plots play an important role in outlier diagnostics, and their superiority in detecting multiple outliers is not affected by the enshroudment of outliers by each other, by other points, or by retaining harmful information on regressor variables that were caused by robust residuals in the residuals plots.

6 References

Atkinson, A.C. (1985). *Plots, Transformation, and Regression.* An introduction to graphical methods of diagnostic regression analysis. New York: Oxford University Press.

Barnett, V. and Lewis, T. (1994). *Outliers in Statistical Data.* 3rd edition. New York: John Wiley.

Belsley, D.A., Kuh, E. and Welsch, R.E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York: John Wiley.

Coakley, C. and Hettmansperger, T.P. (1993). A bounded influence, high break-down efficient regression estimator. *Journal American Statistical Association*, 88, 872-880.

Carroll, R.J. and Ruppert, D. (1985). Transformation in regression: a robust analysis. *Technometrics*, 27, 1-12.

Cook, R.D., and Weisberg, S. (1982). *Residuals and Influence in Regression.* New York: Chapman and Hall.

Cook, R.D., Hawkins, D.M. and Weisberg, S. (1992). Comparison of model misspecification diagnostics using residuals from least median of squares and least median of squares fits. *Journal American Statistical Association*, 87, 419-424.

Chatterjee, S. and Hadi, A.S. (1988). *Sensitivity Analysis in Linear Regression.* New York: John Wiley.

du Toit, S.H.C., Steyn, A.G.W. and Stumpf, R.H. (1986). *Graphical Exploratory Data Analysis.* New York: Springer-Verlag.

Hocking, R.R (1996). *Method and Application of Linear Models.* New York: John Wiley & Sons. Inc.

Hocking, R.R. and Pendleton, O.J. (1983). The regression dilemma. *Communication in Statistics, part A-Theory and Methods*, 67, 388-394.

Huber, P. J.(1973). Robust Regression: asymptotic, conjectures, and Monte Carlo. *Annals of Statistics*, 1, 799-821.

Hussain, S.S. and Sprent, P. (1983). Non-Parametric Regression. *Journal of the Royal Statistical Society Series A. Part 2*, 146, 182-191.

Jajo, N. K. (1999). *Robust diagnostics of outliers in linear regression*, Ph.D. Thesis. Nanaki University, Tianjin, China.

Jajo, N.K. and Wu, X.Z. (1998). Robust diagnostics of outliers in multiple linear regression. *Proceedings of joint Statistical conference. Commemorating the 100th anniversary of Peking University.* Organized by the Institute of Mathematical Statistics and the Department of Probability and Statistics, Peking University, Peking, P.R.China.

Kianifard, F. and Swallow, W.H. (1989). Using recursive residuals, calculated on adaptive ordered observations, to identify outliers in linear regression. *Biometrics*, 45, 571-585.

Kianifard, F. and Swallow, W.H. (1996). A review of the development and application of recursive residuals in linear models. *Journal American Statistical Association*, 88, 1254-1263.

Mckean, J.W., Sheather, S.J. and Hettmansperger, T.P. (1993). The use and interpretation of residuals based on robust estimation. *Journal American Statistical Association*, 88, 1254-1263.

Mckean, J.W., Sheather, S.J. and Hettmansperger, T.P. (1994). Robust and high-breakdown fits of polynomial models. *Technometrics*, 36, 409-415.

Mickey, M.R., Dunn, O.J. and Clark, V. (1967). Note on use of stepwise regression in detecting outliers. *Computer & Biomed. Res.*, 1, 105-111.

Rousseeuw, P.J. and Leroy, A.M. (1987). *Robust Regression and Outlier Detection.* New York: Wiley.

Ruppert, D. and Carroll, R.J. (1980). Trimmed least squares estimation in the linear regression model. *Journal American Statistical Association*, 75, 828-838.

Simpson, D.G., Ruppert, D. and Carroll, R. (1992). On one-step GM estimates and stability of influences in linear regression. *Journal American Statistical Association*, 87, 439-450.

Swallow, H.W. and Kianifard, F. (1996). Using a robust scale estimates in detecting multiple outliers in linear regression. *Biometrics*, 52, , 545-556.

Velilla, S. (1998). A note on the behavior of residual plots in regression. *Statistics & Probability Letters*, 37, 269-278.

SORT 29 (1) January-June 2005, 11-26

Positivity theorem for a general manifold

Rémi Léandre*

Université de Bourgogne

Abstract

We give a generalization in the non-compact case to various positivity theorems obtained by Malliavin Calculus in the compact case.

MSC: 60H07

Keywords: Malliavin calculus. Density

1 Introduction

For background about differential geometry, the reader can see the appendix.

Let us consider $X_0, X_1, \ldots, X_m$ $m+1$ smooth vector fields on R^d. Let us consider B_t^i m independent Brownian motions. We consider the equation in Stratonovitch sense on R^d:

$$dx_t(x) = X_0(x_t(x))dt + \sum_{i>0} X_i(x_t(x))dB_t^i \tag{1}$$

issued of $x \in R^d$. If we perform a change of coordinates through a diffeomorphism of R^d, the vector fields are transformed according this change of coordinates, and since in Itô-Stratonovitch Calculus, the Itô formula is the traditional one, equation (1) has a meaning independent of the system of coordinates chosen. This means that we can look at (1) on a manifold. On R^d, we can consider the quadratic form

* *Address for correspondence:* Institut de Mathématiques. Faculté des Sciences. Université de Bourgogne. 21000. Dijon. FRANCE. *e-mail:* Remi.leandre@u-bourgogne.fr.

Received: September 2003

Accepted: December 2004

$g(x)^{-1} : \xi \to \sum_{i>0} < X_i(x), \xi >^2$. This quadratic form depends smoothly on x. We say that we are in an elliptic situation if this quadratic form is non-degenerated. We can introduce the measure on R^d $d\mu(x) = detg(x)^{1/2}dx$ where dx is the Lebesgue measure on R^d. $d\mu$ is transformed intrinsically under a change of coordinates. We say that R^d endowed with the family of quadratic forms $g(x)$ is a Riemannian manifold endowed with the Riemannian measure $d\mu$. These considerations lead to the definition of a general (curved) Riemannian manifold. If X is a vector field on R^d, we can define its divergence with respect of the measure $d\mu$ by the following integration by parts formula:

$$\int_{R^d} X f d\mu = \int_{R^d} f div X d\mu \tag{2}$$

If f is a function on R^d, the differential $df(x)$ can be assimilated via the non-degenerated quadratic form $g(x)$ to a vector field $grad f$ (We suppose that we are in an elliptic situation). The Laplace-Beltrami operator is therefore $\Delta = div\, grad$. Moreover, (1) generates a Markov process whose generator is $L = X_0 + 1/2 \sum_{i>1} X_i^2$. We can find a drift $\tilde{X}_0$ such that

$$-L = \tilde{X}_0 + \Delta \tag{3}$$

All these considerations are invariant under a change of coordinates on R^d, and apply to a general manifold.

Let us consider a general Riemannian manifold M, endowed with its Riemannian measure dy. On the space $C_K^\infty(M)$ of smooth functions f with compact support on M, there is a canonical second order operator Δ called the Laplace-Beltrami operator. Δ is symmetric positive:

$$\int_M g(y)\Delta f(y)dy = \int_M f(y)\Delta g(y)dy \tag{4}$$

By analytical methods (see Gilkey(1995) for instance), we can solve the parabolic differential equation:

$$\frac{\partial}{\partial t}u_t = -\Delta u\,;\; u_0 = f \tag{5}$$

where f is smooth with compact support. We call $u_t = \exp[-t\Delta]f$. By classical analytical technics (see Gilkey, 1995), the semi-group has smooth heat-kernels:

$$u_t(x) = \int_M p_t(x,y) f(y) dy \tag{6}$$

where $(t,x,y) \to p_t(x,y)$ is smooth from $R^{+*} \times M \times M$ into R^{+*}. The fact that $p_t(x,y) > 0$ can be proved by analytical methods based upon the maximum principle.

Let us suppose that M is compact. By using a finite cover by small balls and some suitable partition of unity, we can write Δ under Hoermander's form:

$$-\Delta = X_0 + 1/2 \sum X_i^2 \tag{7}$$

for some suitable smooth vector fields on M (this decomposition is still true over each relatively compact open subset of M in the non compact case). Reciprocally, we can consider an Elliptic Hoermander's type operator L on M: the vector fields $X_i\, i \neq 0$ span in all points the tangent space of M if $L = X_0 + 1/2 \sum_{i>0} X_i^2$. In such a case, we can introduce a metric on M and a drift $\tilde{X}_0$ such that

$$-L = \tilde{X}_0 + \Delta \tag{8}$$

where Δ is the Laplace-Beltrami operator associated to the Riemannian metric. We can solve the linear parabolic equation:

$$\frac{\partial}{\partial t} u_t = L u_t \ ; \ u_0 = f \tag{9}$$

We get a semi-group $\exp[tL]$ and we can show that there exists a smooth in (x, y) strictly positive heat-kernel such that:

$$\exp[tL] f(x) = \int_M p_t(x, y) f(y) dy \tag{10}$$

The proofs of these results are based upon elliptic theory and the maximum principle (Gilkey, 1995). The natural geometrical object associated with these operators is the Riemannian metric.

There are other distances, called Carnot-Caratheodory distances, which are associated with non-integrable systems of subspaces of the tangent space of the manifold. It is a natural object in Sub-Riemannian geometry. The big difference is the following: if the Riemannian distance is Lipschitz, a Sub-Riemannian distance is in general only Hoelder. Hoermander type operators are associated to Sub-Riemannian geometry. They are of the shape $L = 1/2 \sum X_i^2$. If in all x, the Lie algebra spanned by the X_i is equal to the tangent space of M (Strong Hoermander's hypothesis), the semi-group $\exp[tL]$ has a smooth **strictly positive** heat-kernel $p_t(x, y)$ (Hoermander, 1967). We are motivated by an extension of this theorem to Hoermander's type operator with drift $L = X_0 + 1/2 \sum X_i^2$.

Let us consider a general manifold M and some smooth vector fields $X_i, i = 0, \ldots, m$. In x, let us introduce the Lie ideal spanned by the vector fields X_i, X_0 excluded in the Lie algebra spanned by all the vector fields X_i, X_0 included. It is constructed as follows: we consider the space $F_0(x)$ spanned by the vector fields $X_i, i \neq 0$ in x. We define inductively $F_n(x)$ the linear space spanned by the Lie Brackets between an element of $F_{n-1}(x)$ and a vector $X_i\ i = 0, \ldots, m$. We suppose that in x, $\cup F_n(x) = T_x(M)$. This Hypothesis is called weak Hoermander's hypothesis in x. Let us consider a Hoermander's type operator $L = X_0 + 1/2 \sum_{i>0} X_i^2$. Hoermander's theorem (Hoermander, 1967) states that the semi-group generated by L has a smooth density $p_t(x, y)$ if the weak Hoermander hypothesis is checked in all x. We want to know when $p_1(x, y) > 0$. For that, we consider

the control equation:

$$dx_t(h) = X_0(x_t(h))dt + \sum_{i>0} X_i(x_t(h))h_t^i dt \tag{11}$$

starting from x, where h_t^i belongs to $L^2([0,1])$. In R^d, when Hoermander's condition is satisfied in x and when the vector fields X_i are bounded with bounded derivatives of all orders, Ben Arous and Léandre (1991) have given the following criterion: $p_1(x,y) > 0$ if and only if there exists a h such that $x_1(h) = y$ and such that $h' \to x_1(h')$ is a submersion in h.

This last condition is called Bismut condition (Bismut, 1984).

The boundedness assumption in this theorem can be seen as a compactness assumption. Let us namely compactify R^n by adding a point at the infinity. We get the sphere S^n. The vector fields X_i bounded with bounded derivatives can be extended into smooth vector fields over S^n equal to 0 at the infinity.

Tools used by Ben Arous and Léandre were Malliavin Calculus. This theorem was generalized by Léandre (1990) for jump processes. An abstract version for diffusions was given by Aida, Kusuoka and Stroock (1993). Bally and Pardoux (1998) have given a version of this theorem for the case of a stochastic heat equation. A. Millet and M. Sanz Solé(1997) have given a positivity theorem for the case of a stochastic wave equation. Fournier (2001) has generalized the theorem of Léandre (1990) for the case of a non-linear jump process associated to the Boltzmann equation. Léandre (2003a) has studied the case of a delay equation on a manifold.

By using the mollifier in Malliavin sense introduced by Jones and Léandre (1997) and Léandre (1994), our goal is to remove the boundedness assumption in the theorem of Ben Arous and Léandre, and to generalize it to a general manifold M not necessarily complete. Our theorem is the following:

Main Theorem. *Let us suppose that in all points x of the manifold M, the weak Hoermander's hypothesis is checked. Then $p_1(x,y) > 0$ if and only if there exists an h such that $x_1(h) = y$ and such that $h' \to x_1(h')$ is a submersion in h.*

We refer for more details on Malliavin Calculus to the review of Meyer (1984), to the surveys of Léandre (1988), Léandre (1990), Kusuoka (1992) and Watanabe (1992) for the application of Malliavin Calculus to heat kernels in the compact or the bounded case. In the first part, we give a proof of the main theorem. In the second part, we give some extensions to other processes than diffusions.

2 Proof of the main theorem

Let us show that the condition is sufficient.

Let us introduce the solution of the stochastic differential equation in Stratonovitch sense, where B_t^i are some independent Brownian motions:

$$dx_t(x) = X_0(x_t(x))dt + \sum_{i>0} X_i(x_t(x))dB_t^i \tag{12}$$

starting from x. Let us introduce the exit time τ of the manifold. If f is a smooth function on M, we have classically (see Ikeda-Watanabe (1981), Nualart (1995)):

$$\int p_1(x,y)f(y)dy = E[f(x_1(x))1_{\tau>1}] \tag{13}$$

where $p_t(x,y)$ is the heat-kernel associated to the heat semi-group associated to the Hoermander's type operator $L = X_0 + 1/2\sum_{i>0} X_i^2$. In general, we cannot apply Malliavin Calculus to the diffusion $x_t(x)$. In order to be able to apply Malliavin Calculus, we introduce the mollifiers of Jones-Léandre (1997) and Léandre (1994). We consider a smooth function d from M into R^+, equal to 0 only in x and which tends to ∞ when y tends to infinity, the one compactification point of M. We consider a smooth function over $]-k,k[$ ($k \in R^+$), equal to 1 over $[-k/2, k/2]$ and which behaves as $\frac{1}{(k-y)^r}$ when $y \to k_-$. Outside $]-k,k[$, this function, called $g_k(y)$ is equal to ∞. We suppose that $g_k \geq 1$.

We choose a big integer r. We choose a smooth function from $[1,\infty[$ into $[0,1]$, with compact support, equal to 1 in 1 and which decreases.

The mollifier functional of Jones-Léandre (1997) is

$$F_k = h(\int_0^1 g_k(d(x_s(x)))ds) \tag{14}$$

Lemma. *F_k belongs to all the Sobolev spaces in the sense of Malliavin Calculus if r is big enough, and is equal to 1 if $\sup_s d(x_s(x)) \leq k/2$, is smaller than 1 if $\sup_s d(x_s(x)) > k/2$ and is equal to 0 almost surely if $\sup_s d(x_s(x)) \geq k$. Moreover, $F_k \geq 0$.*

Proof of the Lemma. The support property of F_k comes from the fact that the paths of the diffusion $s \to x_s(x)$ are in fact almost surely Hoelder with a Hoelder exponent strictly smaller than $1/2$, instead of being only continuous.

Let us show that F_k belongs to all the Sobolev spaces.

Let us introduce some smooth vector fields X_i^k which are equal to X_i for $d \leq k$ and which are equal to 0 if $d \geq k+1$. We consider the stochastic differential equation in Stratonovitch sense starting from x:

$$dx_t^k(x) = X_0^k(x_t^k(x))dt + \sum_{i>0} X_i^k(x_t^k(x))dB_t^i \tag{15}$$

Since we consider a Stratonovitch equation, its solution is the limit in all the L^p of the solution of the random ordinary differential equation got when we replace the Stratonovich differential dB_t^i by the random ordinary differential of the polygonal approximation of the leading Brownian motion. It is called Wong-Zakai approximation (Ikeda-Watanabe (1981)). This explains, as we will see later, that the rules of computations with this equation are formally the same as for the solution of an ordinary differential equation, unlike an Itô equation. We put

$$\tilde{F}_k = h(\int_0^1 g_k(d(x_s^k(x)))ds) \tag{16}$$

We get clearly $\tilde{F}_k = F_k$. The interest to use the diffusion $x_t^k(x)$ instead of the initial diffusion is that we can apply Malliavin Calculus to it. Let us recall quickly how we proceed (see Meyer (1984) for a detailed exposition). Since the vector fields X_i^k have compact support, we can exhibit a smooth version of $x \to x_t(x)$ (See Ikeda-Watanabe(1981) and Meyer (1981)). We put

$$\phi_t^k(x) = \frac{\partial}{\partial x} x_t^k(x) \tag{17}$$

which is the solution of the linear equation in Stratonovitch sense:

$$d\phi_t^k(x) = \frac{\partial}{\partial x} X_0^k(x_t^k(x))\phi_t^k(x)dt + \sum_{i>0} \frac{\partial}{\partial x} X_i^k(x_t^k(x))\phi_t^k(x)dB_t^i \tag{18}$$

If we perturb dB_t^i into $dB_t^i + \lambda h_t^i dt$, we get by Ikeda-Watanabe (1981) and Meyer (1981) a smooth version of the solution $x_t^k(\lambda, x)$. Moreover , $\frac{\partial}{\partial \lambda} x_t^k(0, x)$ is solution of the stochastic differential equation with second member which is deduced from the first one by taking formally the derivative of the equation of $x_t^k(\lambda, x)$. These formal considerations are justified because the vector fields have compact supports (see Ikeda-Watanabe (1981) and Meyer (1981)). We get, in Stratonovitch sense:

$$\begin{aligned} d\frac{\partial}{\partial \lambda} x_t^k(0, x) &= \frac{\partial}{\partial x} X_0^k(x_t^k(x)) \frac{\partial}{\partial \lambda} x_t^k(0, x)dt \\ &+ \sum_{i>0} \frac{\partial}{\partial x} X_i^k(x_t^k(x)) \frac{\partial}{\partial \lambda} x_t^k(0, x)dB_t^i + \sum_{i>0} X_i^k(x_t^k(x))h_t^i dt \end{aligned} \tag{19}$$

Since we consider Stratonovitch differential, we can solve (19) by the method of variation of constant. We get:

$$\frac{\partial}{\partial \lambda} x_t^k(0, x) = D_h x_t^k(x) = \phi_t^k(x) \int_0^t (\phi_s^k(x))^{-1} X_i^k(x_t^k(x))h_t^i dt \tag{20}$$

Therefore the random kernel of $Dx_t^k(x)$ is given by

$$Dx_t^k(x)(s) = \phi_t^k(x)(\phi_s^k(x))^{-1} X_i^k(x_s^k(x))$$

for $s \leq t$. Since the vector fields have compact supports, $\phi_t^k(x)$ as well as its inverse are bounded in all L^p for finite p. So the kernel of $Dx_t^k(x)$ are bounded in all the L^p (see Meyer (1984)).

Moreover, the path $t \to x_t^k(x)$ is Hoelder with Hoelder exponent strictly smaller than $1/2$. By Kolmogorov lemma (see Meyer (1981)), the Hoelder norm of the diffusion $t \to x_t^k(x), t \leq 1$ belongs to all the L^p. We deduce that for r big enough (see Jones-Léandre (1997) (2.14))

$$P\{\sup_t \frac{1}{\left(k - d(x_t^k(x)\right)^+} > \frac{1}{\epsilon}; \int_0^1 \frac{dt}{(k - d(x_t^k(x)))^{+r}} < C\} < C(p)\epsilon^p \tag{21}$$

for all p.

The kernel of the first derivative of $\tilde{F}_k$ is not 0 only when $\sup d(x_t^k(x)) \leq k$. It is given by

$$h'\left(\int_0^1 g_k(d(x_t^k(x))dt\right) \int_0^1 g_k'(d(x_t^k(x)))d'(x_t^k(x))Dx_t^k(x)(s)dt \tag{22}$$

It remains to use the inequality

$$\begin{aligned} &|\int_0^1 g_k'(d(x_t^k(x)))d'(x_t^k(x))Dx_t^k(x)(s)dt| \\ &\leq \left(\int_0^1 (g_k'(d(x_t^k(x))\right)^2 dt)^{1/2} \left(\int_0^1 (d'(x_t^k(x))Dx_t^k(s)(s))^2 dt\right)^{1/2} \end{aligned} \tag{23}$$

and to use (21) in order to deduce that $D\tilde{F}_k(s)$ is bounded in all the L^p. The same holds for the derivatives of higher order of $\tilde{F}_k$. □

We introduce the auxiliary measure μ_k:

$$\mu_k : f \to E[F_k f(x_1(x))] \tag{24}$$

To the measure μ_k, we can apply Malliavin Calculus. Namely, $\mu_k[f] = E[\tilde{F}_k f(x_1^k(x))]$. In particular μ_k has a density q_k smaller than $p_1(x, y)$. In particular, if there exists a h such that $x_1(h) = y$ and $h' \to x_1(h')$ is a submersion in h, we can find k large enough such that $q_k(y) > 0$, by the positivity theorem of Ben Arous and Léndre (1991) in the compact case with the extra-condition that $\tilde{F}_k$ has to be strictly positive. This shows that the condition is sufficient.

In order to show that the condition is necessary, we remark that if $p_1(x,y) > 0$ in y, $q_k(y)$ is still strictly positive for k large enough , because for k enough large, for ϵ small

$$|E[(1_{\tau>1} - F_k) f(x_1(x))]| \leq \epsilon \|f\|_\infty \tag{25}$$

where $\|f\|_\infty$ denotes the uniform norm of f.

Therefore, it is enough to apply the Ben Arous-Léandre result in the other sense.

Remark: Let us suppose that Hoermander's condition is satisfied only in x. We can suppose that h is decreasing and that g_k decreases to 1, such that F_k increases to $1_{\tau>1}$. By Malliavin Calculus, μ_k has a density q_k, which increases. Let us consider the function $f = 1_A$ for a set A of measure 0 for the Lebesgue measure over M. We have:

$$\mu_k[f] = 0 \tag{26}$$

But

$$\mu_k[f] = E[F_k f(x_1(x))] = 0 \tag{27}$$

anf $F_k f(x_1(x))$ increases and tends to $1_{\tau>1} f(x_1(x))$, which is in L^1. We deduce that

$$E[1_{\tau>1} f(x_1(x))] = 0 \tag{28}$$

This means that the the law of $x_1(x)$ has a density without to suppose that Hoermander's hypothesis is satisfied in all points.

Remark: The localization procedure given in this work is a localization procedure of all the paths between 0 and $x_1(x)$, when we cannot apply Malliavin Calculus to all the diffusions $x_t(x)$. It is different of various localization procedures, developped in Léandre (1988) for instance, in order to get some estimates of hypoelliptic heat-kernels in small time, which were used when we can apply the machinery of the Malliavin Calculus to all the diffusion $x_t(x)$: namely, in Léandre (1988), we consider vector fields X_i on R^n with bounded derivatives of all orders in order to apply Malliavin Calculus. This allows to get a rough estimate of the heat kernel. Nash inequality (Carlen-Kusuoka-Stroock (1987)) allows to get rough estimates of the heat kernel: in Léandre (2002) we mix the localization procedures developped in this part and the Nash inequality, in order to localize the estimates which were got previously by Malliavin Calculus (see Kusuoka (1992), Léandre (1988), Léandre (1990), Watanabe (1992)) under the restrictions of Malliavin Calculus, and to avoid the classical boundedness assumption of Malliavin Calculus.

Remark: Since the Laplace-Beltrami operator is an elliptic Hoermander's type operator on each locally compact open subset of the manifold, we can apply the previous localization method to show that the heat-kernel associated to the Laplace-Beltrami operator on a Riemannian manifold is strictly positive.

Remark: If the drift X_0 is identically equal to 0, this theorem recovers the fact that the heat kernel associated to the operator $1/2 \sum X_i^2$ under the strong Hoermander's hypothesis is strictly positive, by using the technics of Léandre (1988) Theorem II.1.

3 Extensions

The main novelty of the Malliavin Calculus with respect to its preliminary forms (See works of Hida, Elworthy, Albeverio, Fomin, Berezanskii..) is the following: it can be applied to diffusions, and can differentiate some functionals which are only almost surely defined. There are other examples of Wiener functionals, almost surely defined, where we can apply the Malliavin Calculus and where we can get some positivity theorems. We sketch the proof only.

Nualart-Sanz (1985) consider some smooth vector fields $X_i, i = 0, \ldots, d$ on R^n with derivatives at each order bounded. They consider d independent Brownian sheets $B^i(s,t)$ $s \geq 0, t \geq 0$. Let us recall that it is a Gaussian process indexed by $R^+ \times R^+$ defined by:

$$E[B(s,t)B(s',t')] = (s \wedge s')(t \wedge t') \tag{29}$$

and

$$E[B(s,t)] = 0 \tag{3.2}$$

They consider the Cairoli equation (δ denotes the Itô integral):

$$x_{(s,t)} = x + \sum_{i>0} \int_{[0,s]\times[0,t]} X_i(x_{(u,v)}) \delta B^i(u,v) + \int_{[0,s]\times[0,t]} X_0(x_{(u,v)}) du dv \tag{30}$$

By Malliavin Calculus, Nualart-Sanz (1985) can show if $st > 0$, that $x_{(s,t)}$ has a law having a smooth density with respect of the Lebesgue measure on R^n if in all x, the vector fields X_i $i \neq 0$ span R^n (Nualart-Sanz (1985) study in fact a more degenerated situation). Millet-Sanz (1997) have shown that this density is strictly positive under this non degenerate assumption (They establish in fact under a more general assumption a necessary and sufficient condition for this density to be strictly positive).

By using the fact that the path $(s,t) \to x_{(s,t)}$ is Hoelder, we can remove the hypothesis that the derivative at each order of the vector fields are bounded. If we remove these hypothesis, the two-parameter diffusion can blow up. We introduce O the measurable set where $x_{(u,v)}$ does not blow up on $[0,s] \times [0,t]$. By using the technics of Part III, we can prove the following theorem:

Theorem III.1. *Let us suppose that the vector fields are smooth, and that in all x the vector fields X_i $i > 0$ span R^n. Let us consider the measure: $f \to E[1_O f(x_{(s,t)})]$. This*

measure is bounded below by a measure having a strictly positive smooth density with respect of the Lebesgue measure.

We can restrict the Brownian sheet $B(t,x)$ to the set $R^+ \times [0,1]$ and study the Walsh equation:

$$\frac{\partial}{\partial t}x(t,x) = \frac{\partial^2}{\partial x^2}x(t,x) + \psi(x(t,x)) + \phi(x(t,x))\frac{\partial^2}{\partial t \partial x}B(t,x) \tag{31}$$

where ψ and ϕ are bounded smooth functions with bounded derivatives at each order and $\frac{\partial^2}{\partial t \partial x}B(t,x)$ is the formal white noise associated to $B(t,x)$. We refer to Walsh (1986) for a complete study of this stochastic heat equation. We consider the initial smooth condition $x(0,x) = x_0(x)$ and the Neumann boundary conditions:

$$\frac{\partial}{\partial x}x(t,0) = \frac{\partial}{\partial x}x(t,1) = 0 \tag{32}$$

We suppose that $\phi > 0$ in order to simplify the exposition.

Pardoux-Zhang (1993) have shown that under these conditions, we can apply the Malliavin Calculus to the solution $x_t(x)$ of (31). The final result of Bally-Pardoux (1998) is the following: let us consider $0 \leq x_1 < x_2 < \cdots < x_d \leq 1$. Under these assumptions the law of $x(t,x_1), \ldots, x(t,x_d)$ has a strictly positive smooth density.

But we remark that $(t,x) \to x(t,x)$ is almost surely Hoelder, if $x \to x_0(x)$ is smooth (see Walsh (1986)). Let us introduce the measurable set O where the solution $x(s,x)$ does not blow-up on $[0,t] \times [0,1]$. We can get by using the technics of the third part:

Theorem III.2. *Let us suppose that ϕ and ψ are smooth and that $\phi > 0$. Let us consider $0 \leq x_1 < x_2 \cdots < x_d \leq 1$. Let us consider the measure over R^d:*

$$f \to E[1_O f(x(t,x_1), \ldots, x(t,x_d))] \tag{33}$$

This measure is bounded below by a measure having a strictly positive density on R^d.

The last studied extension is the case of a delay equation on a manifold M. Let us consider a compact Riemannian manifold. If $t \to x_t$ is a semi-martingale on M, we can define the parallel transport from t to t' on the tangent bundle of M endowed with the Levi-Civita connection (See appendix) $\tau_{t',t}$ for $t < t'$. Let us consider some smooth vector fields X_i on M and d independent Brownian motions B^i.

Léandre-Mohammed (2001) have introduced and studied the following delay equation on a manifold in Statonovitch sense:

$$dx_t = \tau_{t,t-\delta} \sum X_i(x_{t-\delta}) dB^i_t \tag{34}$$

with initial condition on $[-\delta, 0]$ equal to the finite energy path $s \to \gamma_s$ defined on $[-\delta, 0]$. The parallel transport considered is the stochastic parallel transport associated to the solution.

Let us suppose that the vector fields X_i span in all points the tangent space. Under these assumptions, Léandre (2003a) has shown that if $t > 0$ the law of x_t has a strictly smooth density with respect of the Riemannian measure, by using the Malliavin Calculus. We remark that $t \to x_t$ is Hoelder. This will allow us to remove the compactness hypothesis on M. If M is not compact, $s \to x_s$ can blow up on $[0, t]$. Let us introduce the measurable set O where $s \to x_s$ does not blow-up on $[0, t]$. By using the technics of the part III, we get the following theorem:

Theorem III.3. *Let us suppose that the smooth vector fields X_i on the non-compact manifold M span in all points the tangent space of M. Let us introduce the measure $f \to E[1_O f(x_t)]$. This measure is bounded below by a measure having a smooth strictly positive density on M.*

4 Appendix: a brief review about stochastic differential geometry

We refer to Elworthy (1982), Emery (1989) and Ikeda-Watanabe (1981) for an extensive study of the material of this part.

Let us recall that a smooth manifold of finite dimension M is locally homeomorphic to an open subset of R^n and that the transition function between different local charts are smooth. We can define the algebra $C^\infty(M)$ of smooth functionals over it. Let (Ω, F_s, P) be a filtered probability space. A continuous semi martingale x_s with values in M is a process such that, by definition, $f(x_s)$ is a semi-martingale with values in R for any smooth functions f.

A vector field X is an operator on $C^\infty(M)$ such that

$$X(fg) = gX(f) + fX(g) \tag{35}$$

We check that $XY - YX$ is still a vector field called the Lie bracket $[X, Y]$ of the two vector fields X and Y. Let $X_i, i = 0, \ldots, d$ some smooth vector fields with compact support: $X_i f$ is equal to zero if the support of f does not intersect the support of X_i. Let B_t^i some independent Brownian motions over R. We introduce the solution of the Stratonovitch differential equation:

$$dx_t(x) = X_0(x_t(x))dt + \sum_{i>0} X_i(x_t(x)) dB_t^i \; ; \; x_0(x) = x \tag{36}$$

This means that for all smooth functions f, the process $x_t(x)$ has to satisfy:

$$f(x_t(x)) = f(x) + \int_0^t X_0 f(x_s(x))ds + \sum_{i>0} \int_0^t X_i f(x_s(x))dB_s^i \tag{37}$$

This differential equation has a unique solution which is a semi-martingale. We can extend this notion to the case where the vector have no-compact supports, if we take care that the solution of (37) can have an blowing-up time $\tau(x)$.

Let L be the operator $X_0 + 1/2 \sum_{i>0} X_i^2$. We can consider the semi-group $\exp[tL]$. It has the following stochastic representation:

$$\exp[tL] f(x) = E[1_{\tau(x)>t} f(x_t(x))] \tag{38}$$

We can consider a vector field as a section of a linear bundle $T(M)$ called the tangent bundle, by looking at a trivialization of M and patching together these trivializations modulo linear maps in the fiber related to the differential of the transition diffeomorphism between the local charts of M. A Riemannian metric is a strictly positive quadratic form over $T_x(M)$, which is intrinsic (it depends consistently upon the different change of trivialization of $T(M)$), and which depends smoothly on x. Let us write in local coordinates the metric $\sum g_{i,j}(x)dx \otimes dx^j$. We can see that the measure $det(g_{i,j}(x))^{-1/2} \prod dx_i$ is intrisically defined. This allows to define the Riemannian measure dx on M.

The application $X(x) \to Xf(x)$ defines a continuous form on the tangent space. By duality, we can write $Xf(x) =< X, grad f >_{T_x(M)}$. Moreover, we have some integration by parts formulas:

$$int_M Xf(x)dx = \int_M f(x)\, divX(x)dx \tag{39}$$

The Laplace-Beltrami operator is defined intrinsically by:

$$\Delta f = div\, grad\, f \tag{40}$$

We can write Δ in local coordinates. For that, let us recall that there is a unique differential operator ∇_Y for a vector field Y acting on the vector fields, which satisfies to the following requirements:

$$\nabla_Y(fX) = f\nabla_Y X + (Yf)X \tag{41}$$

$$\nabla_{\lambda Y + \lambda' Y'} X = \lambda \nabla_Y X + \lambda' \nabla_{Y'} X \tag{42}$$

$$X < Y, Z >=< \nabla_X Y, Z > + < Y, \nabla_X Z > \tag{43}$$

$$\nabla_X Y - \nabla_Y X = [X, Y] \tag{44}$$

(43) says that the connection ∇ is metric. (44) says that the connection is without torsion. In local coordinates, if $X = \sum \lambda_i \frac{\partial}{\partial x_i}$ and $Y = \sum \mu_j \frac{\partial}{\partial x_j}$, we have:

$$\nabla_X Y = \sum \lambda_i \frac{\partial}{\partial x_i} Y + \sum \Gamma_{j,k} \lambda_j \mu_k \frac{\partial}{\partial x_i} \tag{45}$$

The set of $\Gamma^i_{j,k}$, called the Christoffel-Symbols of the Levi-Civita connection, defines in local coordinates a 1-form A with values in the endomorphism of $T_x(M)$.

Let $(g^{i,j}) = (g_{i,j})^{-1}$. In local coordinates,

$$\Delta f = g^{i,j} \left(\frac{\partial^2}{\partial x_i \partial x_j} f(x) - \Gamma^k_{i,j} \frac{\partial}{\partial x_k} f(x) \right) \tag{46}$$

(We use Einstein summation convention).

The Laplace-Beltrami operator is locally an Hoermander's type operator. It generates a semi-group called the heat semi-group on the manifold.

If x_s is a semi-martingale with values in M, we can solve in local coordinates the linear equation:

$$d\tau_t = -A_{dx_t} \tau_t \tag{47}$$

Since, in Stratonovitch Calculus, the Itô formula is the traditional one:

$$f(x_t) = f(x) + \int_0^t < df(x_s), dx_s > \tag{48}$$

the local linear differential equations (47) patch together, and we get a global process which is an isometry from $T_{x_0}(M)$ to $T_{x_t}(M)$ called the stochastic parallel transport along the semi-martingale x_t.

This allows to get the construction of Eells-Elworthy-Malliavin of the Brownian motion starting from x on the Riemannian manifold M:

$$dx_s(x) = \tau_s dB_s \tag{49}$$

where $s \to \tau_s$ is the stochastic parallel transport for the Levi-Civita connection along the solution $x_s(x)$ and B_s a linear Brownian motion in $T_x(M)$. (49) has a unique solution up to an exit stopping time $\tau(x)$. We get:

$$\exp[-t/2\Delta] f(x) = E[1_{\tau(x)>t} f(x_t(x))] \tag{50}$$

Remark. The equation (49) can be extended in the degenerated case by using Langerock' connection, in order to get geometrical degenerated operators (Léandre (2004)).

5 References

Aida, S. Kusuoka, S., Stroock, D.W. (1993). On the support of Wiener functionals. In *Asymptotics Problems in Probability Theory: Wiener Functionals and Asymptotics*. K.D. Elworthy S. Kusuoka N. Ikeda (eds). Pitman Res. Math. Series. 284, 3-34.

Bally, V., Pardoux, E. (1998). Malliavin Calculus for white noise driven parabolic S.P.D.E.S. *Potential analysis*. 9.1., 27-64.

Ben Arous G., Léandre R. (1991). Décroissance exponentielle du noyau de la chaleur sur la diagonale (II). *Probab. Theory Rel. Fields* 90.3, 377-402.

Bismut, J.M. (1984). *Large deviations and the Malliavin Calculus*. Progress in Math. 45. Basel:Birkhauser (1984).

Bochnak, J., Coste, M., Roy, M.F. hspace-0.1cm(1998). *Real Algebraic Geometry*. Heidelberg: Springer.

Carlen, E., Kusuoka, S., Stroock, D.W. (1987). Upper bounds for symmetric transition functions. *Ann. I.H.P*. 23, 145-187.

Carverhill, A.P., Elworthy, K.D. (1983). Flows of stochastic dynamical systems: the functional analytical approach. *Z.W.* 65 245-267.

Elworthy, K.D. (1982). *Stochastic Differential Equations on Manifolds*. L.M.S. Lectures Notes Series 20. Cambridge: Cambridge University Press.

Emery, M. (1989). *Stochastic Calculus in Manifolds*. Heidelberg: Springer (1989).

Fournier, N. (2001). Strict positivity of the solution to a 2-dimensional spatially homogeneous equation without cutoff. *Annales de l'Institut Henri Poincaré-Probabilites et Statistiques*, 31, 481-502.

Gilkey, P. (1995). *Invariance theory, the heat equation and the Atiyah-Singer theorem*. New York: C.R.C. Press.

Hoermander, L. (1967). Hypoelliptic second order equations. *Act. Math.* 119, 147-171.

Ikeda, N., Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. Amsterdam: North-Holland.

Jones, J.D.S., Léandre, R. (1997). A stochastic approach to the Dirac operator over the free loop space. *Proc. Steklov. Inst.* 217, 253-282.

Kusuoka, S. (1992). More recent theory of Malliavin Calculus. *Sugaku* 5, 155-173.

Léandre, R. (1983). Un exemple en théorie des flots stochastiques.In *Séminaire de Probabilités XVII*. Azéma J. Yor M. (eds). Lecture Notes. Maths. 986. Heidelberg: Springer, 158-161.

Léandre, R. (1988). Appliquations quantitatives et qualitatives du Calcul de Malliavin. In *French-Japanese seminar* M. Métivier S. Watanabe (eds). Lecture Notes. Maths. 1322. Heidelberg: Springer, 109-133. English translation: in *Geometry of Random Motion*. R. Durrett M. Pinsky (eds). Contemporary Maths 75. Providence: A.M.S., 173-197.

Léandre, R. (1990). Strange behaviour of the heat kernel on the diagonal. In *Stochastic Processes, Physics and Geometry*. S. Albeverio and al. (eds). Singapore: World Scientific, 516-527.

Léandre, R. (1994). Brownian motion over a Kaehler manifold and elliptic genera of level N. In *Stochastic Analysis and Applications in Physics* A. Cardoso and al. (eds). Nato ASI Series. Dordrecht: Kluwer, 193-219.

Léandre, R. (2002). Stochastic mollifer and Nash inequality. To appear in *Proceedings of the Satellite Conference of Stochastic Analysis of I.C.M.* Ma Z.M. Roeckner M. (eds.)

Léandre, R. (2003*a*). Positivity theorem for a stochastic delay equation on a manifold. *Proceedings of the 8th Vilnius conference*. V. Paulauskas, H. Pragarauskas, R. Rudzkis(edt). *Acta. Appli. Math.*78, 273-284.

Léandre, R. (2003*b*). Malliavin Calculus for a general manifold. In *Equations aux derivées partielles 2002/2003*. J.M. Bony J.Y. Chemin, J. Sjoestrand (eds). Publi. Ecole Polytechnique. Exposé XXIV.

Léandre, R. (2004). A geometrical hypoelliptic diffusion. Preprint Inst. Mathétiques de Bourgogne (374).

Léandre, R., Mohammed, S. (2001). Stochastic functional differential equations on manifolds. *Probability Theory and Related Fields*, 121, 117-135.

Malliavin, P. (1978). Stochastic calculus of variation and hypoelliptic operators. In *Stochastic Analysis* K. Itô (ed.) Tokyo: Kinokuyina, 155-263.

Meyer, P.A. (1981). Flot d'une équation différentielle stochastique. In *Séminaire de Probabilités XV.* J. Azéma M. Yor (eds). Lecture Notes. Maths. 850. Heidelberg: Springer, 100-117.

Meyer, P.A. (1984). Le Calcul de Malliavin et un peu de pédagogie. In *R.C.P. 25.* Volume 34. Public. Univ. Strasbourg, 17-43.

Millet, A., Sanz-Solé, M. (1997). Points of positive density for the solution to a hyperbolic SPDE. *Potential Analysis* 7, 623-659.

Norris, J. (1986). Simplified Malliavin Calculus. in *Séminaire de Probabilités XX*. J. Azéma, M. Yor edit. Lecture Notes. Math. 1204. Heidelberg: Springer, 101-130.

Nualart, D. (1995). *The Malliavin Calculus and related topics.* Heidelberg: Springer.

Nualart, D., Sanz-Solé, M. (1985). Malliavin Calculus for Two-Parameter Wiener Functionals. *Z.W.* 70, 573-590.

Pardoux, E., Zhang, T. (1993). Absolute continuity for the law of the solution of a parabolic SPDE. *J.F.A.* 112, 447-458.

Walsh, J.B. (1986). An introduction to stochastic partial differential equations. In *Ecole d'Eté de Probabilités de Saint-Flour XV* Hennequin P. (eds). Lect. Notes. Math. 1180. Heidelberg: Springer, 266-437.

Watanabe, S. (1992). Stochastic analysis and its applications. *Sugaku* 5, 21-71.

SORT 29 (1) January-June 2005, 27-42

Correspondence analysis and two-way clustering*

Antonio Ciampi[1], Ana González Marcos[2] and Manuel Castejón Limas[2]

[1]*McGill University, Montreal, Canada,* [2]*University of León*

Abstract

Correspondence analysis followed by clustering of both rows and columns of a data matrix is proposed as an approach to two-way clustering. The novelty of this contribution consists of: *i)* proposing a simple method for the selecting of the number of axes; *ii)* visualizing the data matrix as is done in micro-array analysis; *iii)* enhancing this representation by emphasizing those variables and those individuals which are 'well represented' in the subspace of the chosen axes. The approach is applied to a 'traditional' clustering problem: the classification of a group of psychiatric patients.

MSC: 62H25

Keywords: block clustering, selecting number of axes, data visualization.

1 Introduction

Cluster analysis is often introduced as the family of techniques aiming to describe and represent the structure of the pairwise dissimilarities amongst objects. Usually objects are observational units or variables. The dissimilarity between a pair of units is defined as a function of the values taken by *all* the variables on the two units. In a dual way, the dissimilarity between two variables is defined as a function of the values taken by the two variables on the set of *all* units. However, some modern clustering problems, such as those arising in micro-array analysis and text mining, pose a new challenge: not only to describe dissimilarities relationships among individuals and variables, but

* *Address for correspondence:* Antonio Ciampi, Department of Epidemiology and Statistics. antonio.ciampi@mcgill.ca, dieago@unileon.es and diemcl@unileon.es.

Received: November 2003
Accepted: January 2005

also to discover groups of variables and of individuals such that the variables are useful in describing the dissimilarities amongst the individuals and vice versa. To this end, techniques known as two-way clustering and crossed classification clustering have been developed, with the aim of producing homogeneous blocks in the data matrix (Tibshirani *et al.* 1999).

Correspondence analysis (CA), as other biplot techniques, offers the remarkable feature of jointly representing individuals and variables. As a result of such analyses, not only does one gain insight in the relationship amongst individuals and amongst variables, but one can also find an indication of which variables are important in the description of which individuals (Gordon, 1999). It is therefore natural to develop clustering algorithms based on the coordinates of a CA, and indeed practitioners of "analyse des données" commonly advocated this well before the advent of micro-array and text mining.

More recently, in an early attempt to develop clustering methods for micro-array data Tibshirani *et al.* 1999 used the coordinates associated to the first vectors of the singular value decomposition of the data matrix to simultaneously rearrange its rows and columns. They eventually abandoned this approach to concentrate on block clustering. In this work we explore their early idea further. Instead of using only the first axis, we select a few important axes of a CA and apply clustering algorithms to the corresponding coordinates of both rows and columns. Then, instead of ordering rows and columns by the value of the respective first coordinates, we use one of the orderings of rows and columns induced by the classification thus obtained.

The motivation of this work does not come from micro-array or textual analysis, but from a more traditional type of data analysis problem arising from the area of criminal psychology. The problem and the data are briefly described in Section 2; the general approach is described in Section3; Section 4 is devoted to the analysis of our data and finally we end with a brief discussion.

2 The data analysis problem: identifying 'psychopaths' in a data set

In criminal psychology, one is interested in identifying a type of individual that can be considered as especially dangerous, roughly corresponding to what is commonly termed 'psychopath'. One of us (AC) was asked to help identify such a group in a data set consisting of the values taken by 20 variables on 404 patients living in an institution for offenders with recognized psychiatric problems. The source of the data cannot be disclosed. The 20 variables are the items of a standard questionnaire used to identify dangerous individuals. It is current practice to add the values of the 20 variables for each individual and classify an individual as dangerous if this sum exceeds a specific threshold. The expert psychologist was not satisfied with the standard approach, mainly because his intuition was that a 'psychopath' should be defined by only a few personality

variables, and in particular ‘glibness and superficial charm’ and a ‘grandiose sense of self-worth’.

The list of the variables with their meaning is given in Table 1. All variables are ordered categorical, ranging from 0 to 2, with 0 denoting absence, 1 moderate degree and 2 high degree of a certain characteristic. The variables are classified as *behavioural* or *personality related*; this is indicated in the table by the letters B and P respectively.

Table 1:

LABEL	NAME	B/P	MEAN
PCL1	Glibness, superficial charm	P	0.1782
PCL2	Grandiose sense of self worth	P	0.3441
PCL3	Need for stimulation, proness to boredom	P	0.6361
PCL4	Pathological lying	P	0.3936
PCL5	Conning, manipulative	P	0.6386
PCL6	Lack of guilt or remorse	P	1.0594
PCL7	Shallow affect	P	0.5544
PCL8	Callous, lack of empathy	P	0.6015
PCL9	Parasitic life-style	B	0.8490
PCL10	Poor behavioural controls	B	0.8837
PCL11	Promiscuous sexual behaviour	B	0.6188
PCL12	Early behaviour problems	B	0.7005
PCL13	Lack of realistic long term plans	P	0.8218
PCL14	Impulsivity	P	0.7995
PCL15	Irresponsibility	B	0.9084
PCL16	Failure to accept respons for actions	P	0.9604
PCL17	Many short-term marital relationships	B	0.3663
PCL18	Juvenile delinquency	B	1.1188
PCL19	Revocation of conditional release	B	1.0544
PCL20	Criminal versatility	B	1.2747

For instance, PCL1 is classified as personality related (P); it is a number between 0 and 2 according to the level of ‘glibness, superficial charm’ shown by the subject in a video taped interview.

From the analyst’s point of view, some specific aspects of the problem can be identified: *a)* data are recorded on a 3-point ordered scale; *b)* some kind of variable selection should be useful in separating signal from noise; *c)* perhaps some individuals may provide noise and removing them might result in a crisper classification.

The method developed aims at dealing with these aspects. We are looking for clusters, in particular for one ‘stable’ cluster or ‘taxon’ that can be easily identified. There is an additional difficulty: this cluster is probably a rather small one, as, fortunately, psychopaths are rare, even in a population detained for crimes.

3 Clustering individuals by the coordinates of a correspondence analysis

We will consider data in the form of a two-way contingency table and the problem of clustering the rows and the columns of this table. Extensions to data matrices where rows are observational units and columns continuous or categorical ordered variables are easy and will be discussed later. The chi-squared distance between rows and columns is perhaps the most natural dissimilarity that can be defined between pairs of rows and columns. It is well known that if we apply (simple) correspondence analysis (CA) to our table, we obtain a new set of coordinates for both rows and columns, in which the chi-squared distance becomes the classical Euclidean distance, or Pythagorean distance. Therefore, the chi-square distance between two objects calculated directly from the contingency table is equivalent to the Pythagorean distance between the two objects as represented in the factor space of the CA. But the factor space has the interesting property of decomposing the total inertia of the data so that the first axis is associated to the greatest proportion of the total inertia, the second axis to the second largest proportion and so on. In other words, CA defines a sequence of subspaces containing an increasing proportion of the total inertia. It is at the root of the practice of CA that taking only a few axes often clarifies the relationship amongst objects, separating, in some sense, interesting 'signal' from the uninteresting 'noise'.

All this is well known and applied in current data analysis practice (Greenacre, 1984, 1993). Here we propose a few tools for taking maximum advantage of this approach.

3.1 Selecting the number of axes of the CA

One fundamental problem in CA is the identification of the 'important' dimensions. In practice, the selection is done informally, by studying the axes corresponding to the first few eigenvalues of the CA and retaining those which are interpretable. However, when CA is the first step of an algorithm as in our case, it is important to have a slightly formalized selection procedure. We propose one based on the following reasoning. Suppose we had a good idea of the distribution of the eigenvalues of a family of contingency tables similar to ours but generated under the independence assumption. Then if the eigenvalues of our data appear to follow this distribution, we may conclude that none of the axes contains interesting information. On the other hand, if the first k axes of our data matrix contain information and the others do not, then we would expect that the first k eigenvalues markedly differ from the behaviour of the first k eigenvalues of matrices generated under the independence assumption. More formally, this leads to the following rule:

1. Perform a CA of the data matrix and draw the scree plot of the ordered eigenvalues of the CA.

2. Generate an artificial data matrix by randomly permuting the responses (rows) of each variable (column).
3. Perform the CA of the generated data and superimpose the scree plot of its eigenvalues to the graph obtained in 1.
4. Repeat 2 to 3 a number of times.
5. Look at the graph containing all the scree plots. If the graph of the real data is indistinguishable from the band formed by the simulated scree plots, or if the former is consistently below the latter, conclude that there is no structure in the data, *i.e* the two variables defining the contingency table are independent. Otherwise, identify the point of intersection of the real data's scree plot with the band of simulated scree plots and conclude that the number of interesting axes is the largest one at the left of the abscissa of the intersection.

We remark that this rule can also be seen as a formalization of the elbow rule, which is very popular in factorial analyses.

3.2 Visualization of the data matrix

Consider a $R \times C$ contingency table with elements n_{ij}, $i = 1 \ldots R$, $j = 1 \ldots C$. A visual representation of such data is obtained by associating different colors to segments of the range of the n_{ij}'s and drawing a picture in the form of $R \times C$ grid with colors replacing the n_{ij}'s. This is exactly what is done in micro-array analysis, but it would be applicable to any range of data, even negative data. Normally this picture is rarely useful, unless the rows and the columns have been previously permuted in an appropriate way, aiming to extract information. What we propose here is to cluster the points representing rows and columns in the (reduced) factor space with Euclidean distance by a hierarchical clustering algorithm (e.g. Ward). Then the rows and columns of the picture can be rearranged using (one of) the ordering(s) induced by the clustering. The hierarchical trees can also be drawn in the picture, imitating again what is current practice in micro-array analysis.

Obviously this representation is applicable to any data matrix with non-negative entries, and is especially useful for a cases $\times$ variables rectangular matrix with variables which are measured in the same units or which have been preliminarily scaled and centred.

3.3 Taking out poorly represented variables and/or cases

As is well known, CA provides useful aids to interpretation, among which the quality of the representation of each object on each object factorial axis: this is defined as the square of the cosines of the angle that the object forms with the object axis. By summing

these quantities for the first k axes, one obtains the quality of the representation of the object in the object factorial subspace spanned by the first k axes.

In our approach, we propose to cluster both variables and individuals using only their coordinates on a few chosen axes. This was motivated by the aim of decreasing noise. Now, it may be useful to further reduce noise by representing graphically only the individuals and the variables that are well represented on the subspace spanned by the chosen axes. In this paper we distinguish well-represented objects from poorly represented objects by defining as well-represented an object which is better represented in the factorial subspace than outside of it (quality of the representation on the factorial plane > 50%). As we will demonstrate in the next section, it is useful to remove the individuals and/or the poorly represented variables, and to repeat the analysis with the sub-matrix of the original data matrix consisting of only the well-represented variables and/or individuals.

3.4 Applying the approach to other types of data

CA, initially developed for contingency tables, is formally applicable to any data matrix with non-negative entries. In particular it is applicable to a data matrix of the form cases $\times$ variables as long as the variables can only assume non-negative values. Because of the *distributional invariance* property of the chi-square distance, on which CA is based, the application is particularly well justified if the object of the analysis is to study profiles. An example where CA is useful and relevant is the case of nutrition data, when one wishes to find dietary patterns (clusters) based on the proportion of each nutrient that an individual absorbs rather than on absolute values. Thus two individuals are considered similar if they have a very close profile, i.e. if they absorb similar proportions of each nutrient regardless of the total amount, which can be quite different.

Another well-justified application of CA to situations other than the two-way contingency table is to cases $\times$ variables data matrices when the variables are ordinal, with levels represented by non-negative integers. Here, however, one needs to use the artifice of doubling the number of variables, as explained for example in (Greenacre, 1993). Thus for each ordinal variable X_i taking values between, say, 0 and p_i, one creates its non-negative 'double', $p_i - X_i$, and performs CA on the data matrix consisting of all variables and their 'doubles'.

4 Application to our data

The first step of our analysis was to double the variables, i.e. create for each PCLi its double 2-PCLi. The resulting 404 $\times$ 40 matrix was then treated by CA. The graph of the eigenvalues is given in Figure 1. The graph also contains a few graphs based on simulated matrices with random permutations of the responses of each variable.

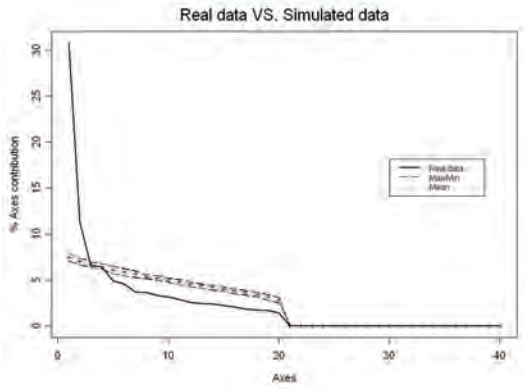

Figure 1: *Real data VS. Simulated data.*

The graph of the real matrix and the band of graphs of the simulated matrices, intersect at a point corresponding to three factorial axes. Before this point the graph of the real matrix is above the band, and lies below it after the intersection. We chose therefore the three-dimensional subspace spanned by the first three factors to represent both individuals and variables. We applied Ward's clustering algorithm with the Euclidean distance to the coordinates of both subject-points and variable-points in this subspace. To verify that the choice of three axes is a sensible one, we plotted our original data matrix (doubled variables are omitted) with both rows and columns ordered according to the clustering for varying number of axes. This is shown in Figure 2.

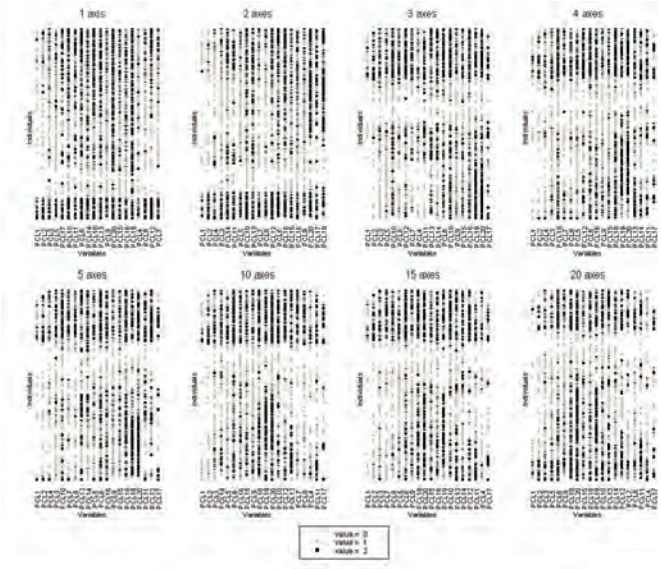

Figure 2: *Clustering with the correspondence analysis coordinates: Euclidean distance.*

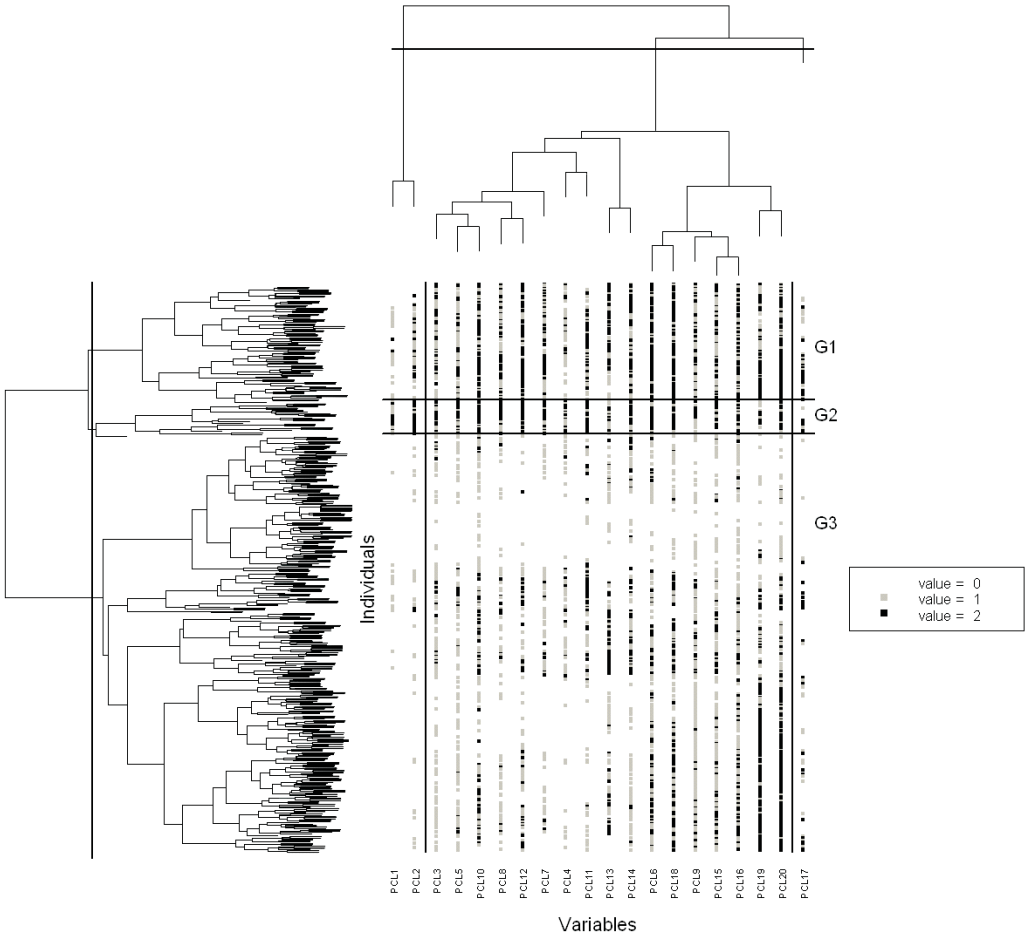

Figure 3: *Clustering with the correspondence analysis coordinates: 3 groups.*

It seems that taking a higher number of axes does not improve the general pattern obtained with three axes. On the other hand, three axes give a much neater picture than one axis only, which would have been the choice of [Tibshirani *et al.* 1999]. Figure 3 is a larger picture of the data matrix with rows and columns ordered according to clustering obtained from the three axes choice; in it we also show the hierarchical classification trees of both columns and rows, with tentative 'cuts' yielding three clusters of variables and three clusters of individuals. Note that two branches come together at the distance between the two clusters being merged.

Interestingly, the clustering of variables places two personality characteristics (PCL1 and PCL2) in one cluster and all the remaining variables in the other with the exception of one behavioural characteristic (PCL17), which remains isolated. A description of this first clustering of subjects is given in Figure 4.

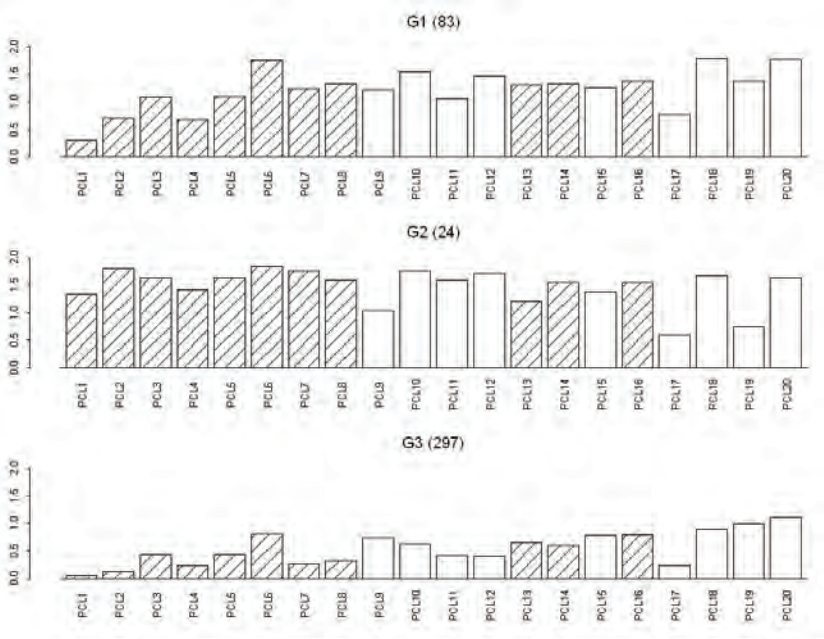

Figure 4: *Cluster description.*

One might interpret group 2 (24 patients) as including the 'psychopaths', since the individuals in this group have high values of PCL1 and PCL2; they also have high value of other personality and behavioural variables. Group 3 includes the majority (297 patients), characterized by generally low levels of all variables. Group 1 (83 patients) is similar to group 2 but some variables, and in particular PCL1 and PCL2, are not as high on the average. Notice that if we had cut the tree so as to have two clusters, we would not have seen the difference between group 2 and group 1. On the other hand, we have also looked at finer cuts up to the eight-cluster partition (data not shown) and found the following features. Group 1 remains a distinct entity even in the eight-cluster solution. Group 2 splits only once, at the 4 cluster cut, into two sub-clusters of 23 and 1 patients respectively, with the isolated patient being characterized by having low values for the behavioural variables. On the other hand, group 3 is the one that appears less stable, splitting into up to 5 sub-clusters.

Next, as explained in Section 2.3, we proceeded to an elimination of variables and individuals and applied our algorithm to what is left of the data matrix. We decided to keep only the rows and columns with quality of the representation on the 3-dimensional factorial subspace greater than 50%. We are left with a 171×8 matrix. The 8 remaining variables are 5 of the 11 personality variables (PCL1, PCL2, PCL7, PCL8, and PCL13) and 3 of the 9 behavioural variables (PCL18, PCL19, PCL20). This is an indication that personality variables in general, and PCL1 and PCL2 in particular, are more relevant to the goal of identifying clusters in the whole data set.

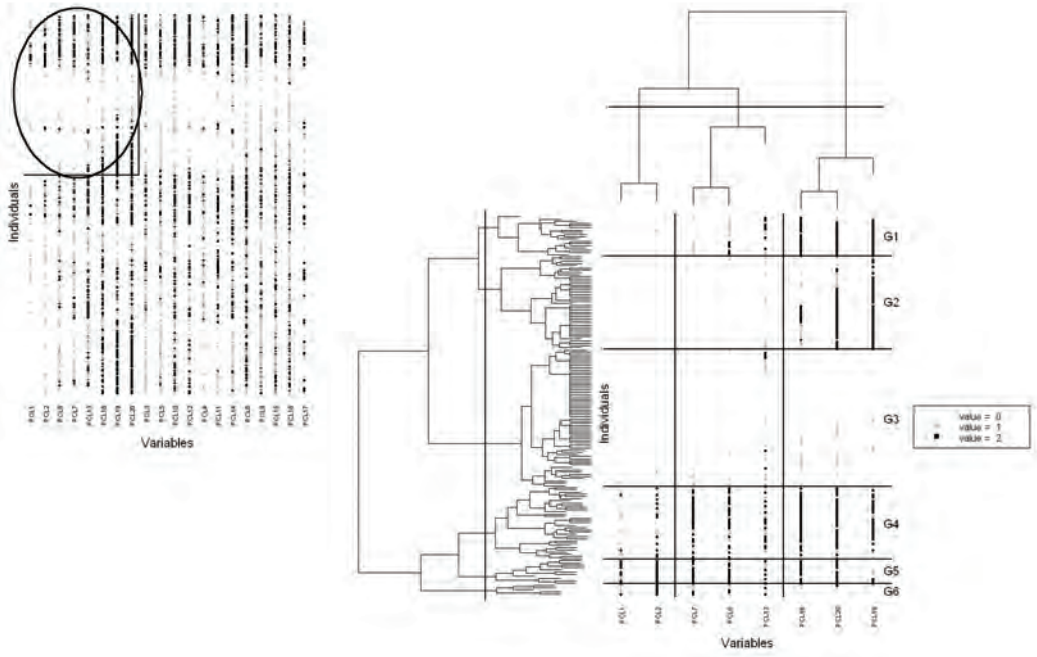

Figure 5: *Clustering with the correspondence analysis coordinates of the representative variables and individuals.*

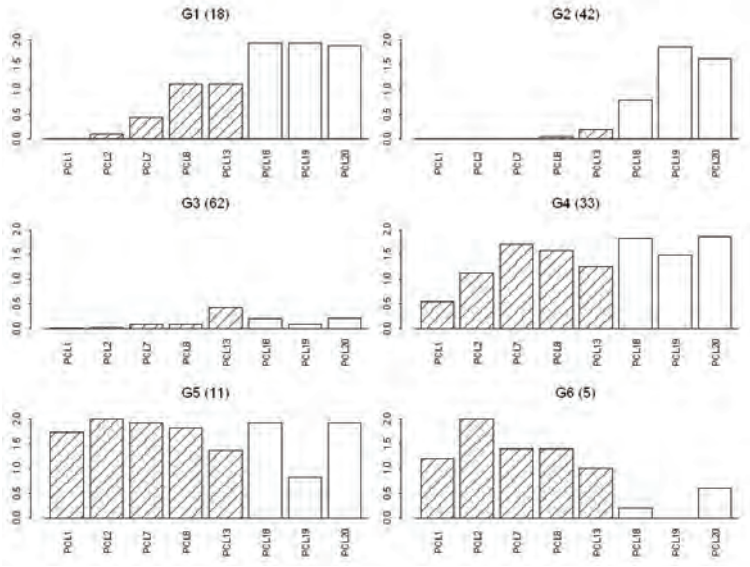

Figure 6: *Cluster description.*

Applying our algorithm to the resulting 171 × 8 data matrix, we obtained the following results. Our modified scree plot approach (not shown here) suggested choosing two factorial axes to represent both individuals and variables. The left side of Figure 5 shows the original data matrix, but now the rows and columns that are poorly represented in the three-dimensional factorial subspace, are shown at the margin. The right portion of the figure shows the row- and column-clustering for the data matrix with the poorly represented rows and columns taken out. A tentative cut of the two trees suggests six clusters of subjects and three clusters of variables.

These clusters are described in Figure 6.

Now it is group 5 which can be seen as consisting of 'psychopaths'. This group has high values of PCL1 and PCL2, but also of many other personality and behavioural variables. Interestingly, the smaller cluster of five individuals consists of individuals with high levels of PCL1 and PCL2 and low levels of the behavioural variables.

Next, given the emphasis of the expert on finding clusters of *subjects*, we re-introduced the poorly representative individuals as supplementary and repeated the clustering for the rows. Figure 7 shows the results of the clustering with the well-represented columns and the entire set of rows. Again, we cut the dendrogram of the rows to obtain six clusters of subjects.

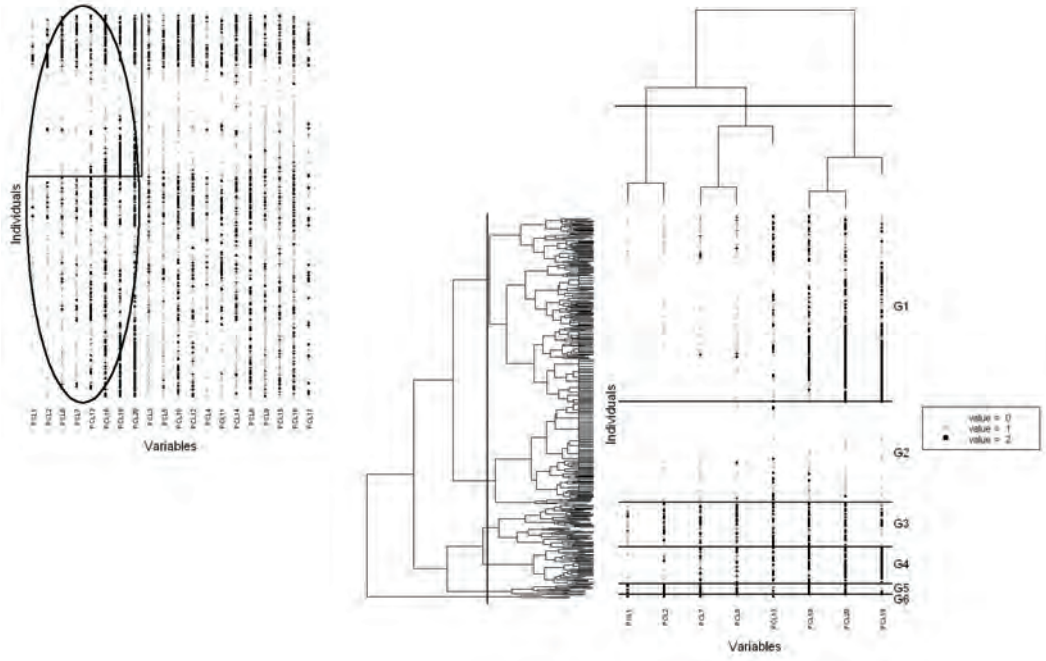

Figure 7: *Clustering with the correspondence analysis coordinates of the representative variables and individuals: poorly represented individuals added as supplementary.*

Comparing these clusters, which are described in Figure 8, with those shown in Figure 6, we observe nearly the same profiles.

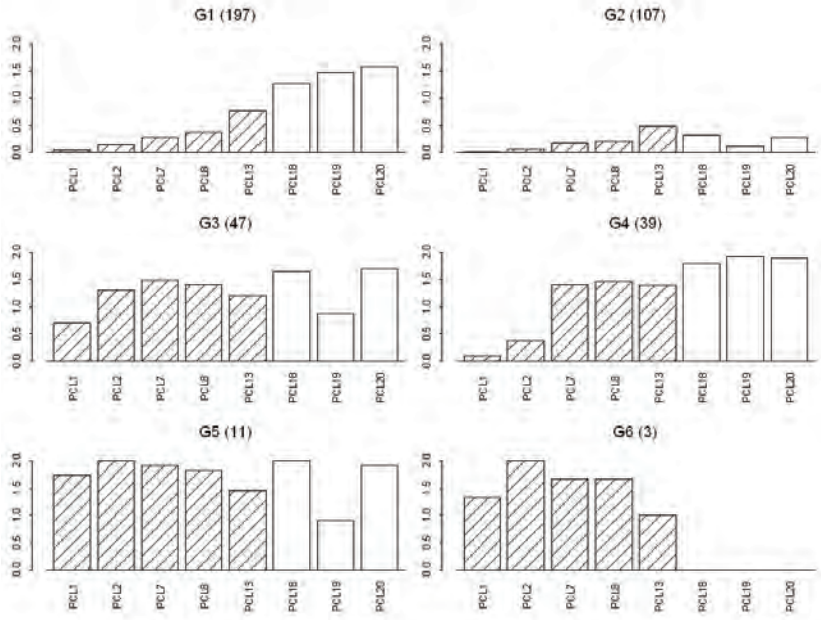

Figure 8: *Cluster description.*

The differences are indeed minor. We have verified that the group of ‘psychopaths’ consists of the same 11 patients in both cases (group 5). The small group of ‘psychopaths’ with a normal behaviour (group 6 in both clusterings), consists of 3 of the 5 individuals of the earlier clustering (the other 2 patients join group 3 of the new clustering). We have also considered coarser solutions and found that the group of the ‘psychopath’ with normal behaviour is clearly distinguishable from the rest even for the two-cluster cut, while the cluster of the ‘psychopaths’ appears at the four-cluster cut.

For these data it appears that removing variables and/or individuals that are poorly represented in the reduced subspace of the ‘important’ first factors, results in a sharper classification. Moreover, the results of the analysis correspond to the expert’s original intuition: psychopaths are an identifiable ‘taxon’, but they are better identified by the personality variables than by the behavioural variables, and, in particular, by PCL1 and PCL2. Our method succeeds in correctly identifying the important variables and obtaining a rather clear hierarchical classification of our patient group. An interesting and somewhat surprising result is the identification of the very small subgroups consisting of individuals with high values of the personality variables PCL1 and PCL2 and low values of the behavioural variables (group 6). Indeed this led our expert to comment that these individuals are ‘psychopaths’ that appear quite normal in their daily

social behaviour and specialize in crimes that are cleverly disguised. He added that the reason why there are so few of them in our psychiatric prison sample, might be that they rarely get caught! Thus group 6 can be considered as a variant of the typical 'psychopath'.

5 Discussion

In this work we have shown by example how CA can be used as a powerful tool in clustering, particularly in two-way clustering, where clustering of both rows and columns (observational units and variables) is of interest. The basic idea consists of first obtaining a representation of both units and variables as points in a subspace of the factor space identified by the CA. Next, a standard hierarchical clustering algorithm is applied to the points of this subspace.

This basic idea, as recognized in the introduction, is far from new. Indeed, data analysts commonly use it, in spite of lack of a strong theoretical foundation. However, in view of recent theoretical work, the idea acquires a new strength: roughly speaking, it appears that if there are clusters, then CA is the best representation to discover them (Caussinus, H. & Ruiz-Gazen, A. (2003), Caussinus, H. & Ruiz-Gazen, A. (1995)). Furthermore, recent work in unsupervised machine learning (Bengio *et al.* (2003), Ng *et al.* (2002)) seems to indicate that 'spectral' data reduction algorithms applied to the matrix of the pairwise distances between points, provide impressive results in retrieving 'unusual' cluster shapes. This is not the same as the basic idea developed in this work, since in our case the spectral decomposition is applied to the data matrix and not the distance matrix (see, however, Greenacre (2000)). Nevertheless, the connection is intriguing. In any case, further theoretical work along the two lines of research mentioned above seems to be highly promising for providing a deeper theoretical justification to the common practice of applying clustering algorithms to reduced data.

The novel contributions of this work are: *i)* proposing a simple method for selecting the number of axes previous to clustering; *ii)* proposing a visualization of the data matrix which generalizes the one current in micro-array analysis; *iii)* enhancing this visualization by emphasizing those variables and those observational units which are 'well represented' in the subspace of the chosen axes. Each of these contributions is grounded more on intuition than on theoretical results. Also, in this paper we have simply presented the ideas and their motivation, illustrating them by the analysis of a non-trivial problem. Each of these contributions should be considered as themes for further research.

The problem of selecting the dimension of the subspace on which to represent the data is all pervasive in data reduction and model building. Many approaches have been proposed and ours is just one within the family of computational intensive proposals. The same approach can be applied to the problem of selecting the number of clusters

or, equivalently for hierarchical clustering algorithms, the level at which to cut the dendrogram. We have, in our case, preferred to not propose a single cut, in keeping with the exploratory aim of our analysis.

The visualization of the data matrix with the aid of CA and clustering may be improved at many levels. Our priority, however, is to extend the approach to the representation of multiple categorical variables, starting from some version of multiple correspondence analysis. As for the enhancement of the visualization by emphasizing the well represented objects, we have already outlined some possibilities that, we feel, deserve to be explored. For instance, one could start by using only the first axis (as in Tibshirani et al., 1999), obtain a visualization of the data, and then pull out the objects that are not well represented. The next step would be to repeat the same approach starting from the second factorial axis, and so on: one obtains as many classification schemes as there are 'important' axes, and each such scheme applies to a subset of individuals and variables, with possibly overlapping subsets.

6 Acknowledgments

This work has been partially funded with a research scholarship granted by the State Secretary of Education and Universities of the Spanish Ministry of Education, Culture and Sport. The authors want to recognize the hospitality of the Department of Epidemiology and Statistics members during the visits of A. González and M. Castejón to McGill University between 2002 and 2004.

7 References

Bengio, Y., Vincent, P., Paiement, J-F., Delalleau, O., Ouimet, M., and Le Roux, N. (2003). 'Spectral clustering and Kernel PCA are learning eigenfunctions'. Technical Report 1239, Département d'Informatique et Recherche Opérationnelle, Université de Montréal. http://www.iro.umontreal.ca/~lisa/publications.html.

Caussinus, H. & Ruiz-Gazen, A. (1995). Metrics for finding typical structures by means of principal component analysis. In *Data Science and its Applications*, Y. Escoufier & C. Hayashi (eds)., Tokyo: Academic Press, 177-192.

Caussinus, H. & Ruiz-Gazen, A. (2003). Which structures do generalized principal component analysis display? The case of multiple correspondence analysis. To appear in *Multiple Correspondence Analysis and Related Methods* (eds. Michael Greenacre and Jörg Blasius), London: Chapman & Hall, 2006.

Gordon, A.D. (1999). Classification, 2nd Edition, London: Chapman & Hall.

Greenacre, M.J. (1984). *Theory and Application of Correspondence Analysis*, London: Academic Press.

Greenacre, M. (1993). *Correspondence Analysis in Practice*. London: Academic Press.

Greenacre, M. (2000). Correspondence analysis of a square symmetric matrix. *Applied Statistics*, 49, 297-310.

Ng, A. Y., Jordan, M. I., and Y. Weiss, Y. (2002). On spectral clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems (NIPS)*, T. Dietterich, S. Becker and Z. Ghahramani (eds.), volume 14. Cambridge MA: MIT Press.

Tibshirani, R., Hastie T., Eisen, M., Ross, D., Botstein, D., and Brown, P. (1999). Clustering methods for the analysis of DNA microarray data. *Technical report*, Department of Statistics, Stanford University. http://www-stat.stanford.edu/~tibs/lab/publications.html.

SORT 29 (1) January-June 2005, 43-56

Information matrices for some elliptically symmetric distributions

Saralees Nadarajah[1], Samuel Kotz[2]

[1] *University of Nebraska,* [2] *The George Washington University*

Abstract

The Fisher information matrices are derived for three of the most popular elliptically symmetric distributions: the Pearson type II, Pearson type VII and the Kotz type distributions. We hope the results could be important to the many researchers working in this area.

MSC: 33C90, 62E99

Keywords: Elliptically symmetric Kotz type distribution, Elliptically symmetric Pearson type II distribution, Elliptically symmetric Pearson type VII distribution, Fisher information matrices

1 Introduction

The elliptically symmetric Pearson type II, Pearson type VII and the Kotz type distributions are given by the joint pdfs

$$f(x,y) = \frac{N+1}{\pi\sqrt{1-\rho^2}}\left(1-\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^N \quad (1)$$

(for $N > -1$ and $-1 < \rho < 1$),

* *Address for correspondence:* Saralees Nadarajah, Department of Statistics, University of Nebraska, Lincoln, NE 68583.
Samuel Kotz, Department of Engineering Management and Systems Engineering, The George Washington University, Washington, D.C. 20052.

Received: September 2004
Accepted: December 2004

$$f(x,y) = \frac{N-1}{\pi m\sqrt{1-\rho^2}}\left(1+\frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right)^{-N} \tag{2}$$

(for $N > 1$, $m > 0$ and $-1 < \rho < 1$), and

$$f(x,y) = \frac{sr^{N/s}\left(x^2+y^2-2\rho xy\right)^{N-1}}{\pi\Gamma(N/s)\left(1-\rho^2\right)^{N-1/2}}\exp\left\{-r\left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^{s}\right\} \tag{3}$$

(for $N > 0$, $r > 0$, $s > 0$ and $-1 < \rho < 1$), respectively. The bivariate t-distribution and the bivariate Cauchy distribution are special cases of (2) for $N = (m+2)/2$ and $m = 1$, $N = 3/2$, respectively. When $s = 1$, (3) is the original Kotz distribution introduced in Kotz (1975). When $N = 1$, $s = 1$ and $r = 1/2$, (3) reduces to a bivariate normal density. The parameter ρ is the correlation coefficient between the x and y components. For details on properties and applications of these distributions see Johnson (1987), Fang *et al.* (1990), Nadarajah (2003) and Kotz and Nadarajah (2004).

The aim of this note is to calculate the Fisher information matrices corresponding to each of the pdfs given by (1), (2) and (3). This requires calculation of product moments of the form $E(X^mY^n)$. A transformation which aides this task is:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \sqrt{1+\rho}+\sqrt{1-\rho} & \sqrt{1+\rho}-\sqrt{1-\rho} \\ \sqrt{1+\rho}-\sqrt{1-\rho} & \sqrt{1+\rho}+\sqrt{1-\rho} \end{pmatrix}\begin{pmatrix} U \\ V \end{pmatrix}. \tag{4}$$

Under this transformation, one can easily see that (1), (2) and (3) reduce to

$$f(u,v) = \frac{N+1}{\pi}\left(1-u^2-v^2\right)^N, \tag{5}$$

$$f(x,y) = \frac{N-1}{\pi m}\left(1+\frac{u^2+v^2}{m}\right)^{-N}, \tag{6}$$

and

$$f(u,v) = \frac{sr^{N/s}\left(x^2+y^2\right)^{N-1}}{\pi\Gamma(N/s)}\exp\left\{-r\left(x^2+y^2\right)^{s}\right\}, \tag{7}$$

respectively. Furthermore, the product moments for (5), (6) and (3) are given by

$$E\left(U^pV^q\right) = \frac{N+1}{\pi}B\left(N+1,\frac{p+q}{2}+1\right)B\left(\frac{p+1}{2},\frac{q+1}{2}\right), \tag{8}$$

$$E\left(U^pV^q\right) = \frac{m^{(p+q)/2}(N-1)}{\pi}B\left(\frac{p+q}{2}+1, N-\frac{p+q}{2}-1\right)B\left(\frac{p+1}{2}, \frac{q+1}{2}\right), \quad (9)$$

and

$$E\left(U^pV^q\right) = \frac{\Gamma\left(\frac{2N+p+q}{2s}\right)}{\pi r^{(p+q)/(2s)}\Gamma\left(N/s\right)}B\left(\frac{p+1}{2}, \frac{q+1}{2}\right), \quad (10)$$

respectively, if both $p \geq 0$ and $q \geq 0$ are even integers (see, for example, Fang *et al.* (1990)). If either p or q is odd then the moment equals to zero. Hence, the product moment $E(X^mY^n)$ for any m and n can be calculated by combining (4), (8), (9) and (10). Certain relationships which express $E(X^mY^n)$ in terms of product moments of (U, V) are given in the Appendix (Section 5).

The calculations for the Kotz type distribution require additional moments of the form

$$J(m,n) = E\left[\left(\frac{X^2+Y^2-2\rho XY}{1-\rho^2}\right)^m\left\{\log\left(\frac{X^2+Y^2-2\rho XY}{1-\rho^2}\right)\right\}^n\right] \quad (11)$$

and

$$K(m,n) = E\left[XY\left(\frac{X^2+Y^2-2\rho XY}{1-\rho^2}\right)^m\left\{\log\left(\frac{X^2+Y^2-2\rho XY}{1-\rho^2}\right)\right\}^n\right]. \quad (12)$$

For this, apply the transformation

$$(X,Z) = \left(X, \frac{X^2+Y^2-2\rho XY}{1-\rho^2}\right).$$

Note that the jacobian of this transformation is:

$$|J| = \frac{\sqrt{1-\rho^2}}{2\sqrt{Z-X^2}}.$$

Thus, (11) and (12) can be reduced to

$$J(m,n) = \frac{2sr^{N/s}P(m,n)}{\pi\Gamma\left(N/s\right)\left(1-\rho^2\right)^{N-1}} \quad (13)$$

and

$$K(m,n) = \frac{2sr^{N/s}\rho Q(m,n)}{\pi\Gamma\left(N/s\right)\left(1-\rho^2\right)^{N-1}}, \quad (14)$$

respectively, where $P(m,n)$ and $Q(m,n)$ denote the integrals

$$P(m,n) = \int_0^\infty \int_0^{\sqrt{z}} \frac{z^{N+m-1}(\log z)^n \exp(-rz^s)}{\sqrt{z-x^2}} dxdz$$

and

$$Q(m,n) = \int_0^\infty \int_0^{\sqrt{z}} \frac{x^2 z^{N+m-1}(\log z)^n \exp(-rz^s)}{\sqrt{z-x^2}} dxdz.$$

Integrating with respect to the x variable yields

$$P(m,n) = \frac{\pi}{2} \int_0^\infty z^{N+m-1}(\log z)^n \exp(-rz^s)\, dz$$

and

$$Q(m,n) = \frac{\pi}{4} \int_0^\infty z^{N+m}(\log z)^n \exp(-rz^s)\, dz.$$

These integrals can be calculated by using equation (2.6.21.1) in Prudnikov *et al.* (1986, volume 1) to yield

$$P(m,n) = \frac{\pi}{2s} \left(\frac{\partial}{\partial \alpha}\right)^n \left[r^{-\alpha/s}\Gamma\left(\frac{\alpha}{s}\right)\right]\Bigg|_{\alpha=N+m} \tag{15}$$

and

$$Q(m,n) = \frac{\pi}{4s} \left(\frac{\partial}{\partial \alpha}\right)^n \left[r^{-\alpha/s}\Gamma\left(\frac{\alpha}{s}\right)\right]\Bigg|_{\alpha=N+m+1}. \tag{16}$$

Hence, (11) and (12) can be calculated by combining (13)–(14) and (15)–(16).

The exact forms of the information matrices are given in Sections 2, 3 and 4. The calculations use the digamma function defined by $\Psi(x) = d\log\Gamma(x)/dx$.

2 Information matrix for Pearson II

If (x,y) is a single observation from (1) then the log-likelihood function can be written as

$$\log L(N,\rho) = \log(N+1) - \log\pi - \frac{1}{2}\log\left(1-\rho^2\right) + N\log\left(1 - \frac{x^2+y^2-2\rho xy}{1-\rho^2}\right).$$

The first-order derivatives are:

$$\frac{\partial \log L}{\partial N} = \frac{1}{N+1} + \log\left(1 - \frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)$$

and

$$\frac{\partial \log L}{\partial \rho} = \frac{\rho}{1-\rho^2} - \frac{2N\left\{\rho x^2 + \rho y^2 - \left(1+\rho^2\right)xy\right\}}{\left(1-\rho^2\right)^2}\left(1 - \frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^{-1}.$$

The second-order derivatives are:

$$\frac{\partial^2 \log L}{\partial N^2} = -\frac{1}{(N+1)^2},$$

$$\frac{\partial^2 \log L}{\partial N \partial \rho} = -\frac{2\left\{\rho x^2 + \rho y^2 - \left(1+\rho^2\right)xy\right\}}{\left(1-\rho^2\right)^2}\left(1 - \frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^{-1},$$

and

$$\frac{\partial^2 \log L}{\partial \rho^2} = \frac{1+\rho^2}{\left(1-\rho^2\right)^2} - \frac{4N\left\{\rho x^2 + \rho y^2 - \left(1+\rho^2\right)xy\right\}^2}{\left(1-\rho^2\right)^4}\left(1 - \frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^{-2}$$

$$- \frac{2N\left[\left(1-\rho^2\right)\left(x^2+y^2-2\rho xy\right) + 4\rho\left\{\rho x^2 + \rho y^2 - \left(1+\rho^2\right)xy\right\}\right]}{\left(1-\rho^2\right)^3}.$$

Now, we can compute the elements of the Fisher information matrix. It is clear that

$$E\left(-\frac{\partial^2 \log L}{\partial N^2}\right) = \frac{1}{(N+1)^2}.$$

By applying (8) and (17)–(19), one gets

$$E\left(-\frac{\partial^2 \log L}{\partial N \partial \rho}\right) = \frac{(N+1)B(2,N)\rho}{1-\rho^2}.$$

Finally, application of (8) and (17)–(24) yields

$$E\left(-\frac{\partial^2 \log L}{\partial \rho^2}\right) = \frac{1}{2\left(1-\rho^2\right)^2}\Big\{2N(N+1)B(3,N-1)\rho^2 + 4N(N+1)B(2,N)\rho^2 - 2\rho^2 + N(N+1)B(3,N-1) + 4N(N+1)B(2,N) - 2\Big\}.$$

3 Information matrix for Pearson VII

If (x,y) is a single observation from (2) then the log-likelihood function can be written as

$$\log L(N,m,\rho) =$$

$$\log(N-1) - \log\pi - \log m - \frac{1}{2}\log\left(1-\rho^2\right) - N\log\left(1 + \frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right).$$

The first-order derivatives are:

$$\frac{\partial \log L}{\partial N} = \frac{1}{N-1} - \log\left(1 + \frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right),$$

$$\frac{\partial \log L}{\partial m} = \frac{N\left(x^2+y^2-2\rho xy\right)}{m^2\left(1-\rho^2\right)}\left(1 + \frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right)^{-1} - \frac{1}{m},$$

and

$$\frac{\partial \log L}{\partial \rho} = \frac{\rho}{1-\rho^2} - \frac{2N\left\{\rho x^2 + \rho y^2 - \left(1+\rho^2\right)xy\right\}}{m\left(1-\rho^2\right)^2}\left(1 + \frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right)^{-1}.$$

The second-order derivatives are:

$$\frac{\partial^2 \log L}{\partial N^2} = -\frac{1}{(N-1)^2},$$

$$\frac{\partial^2 \log L}{\partial N \partial m} = \frac{x^2+y^2-2\rho xy}{m^2\left(1-\rho^2\right)}\left(1 + \frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right)^{-1},$$

$$\frac{\partial^2 \log L}{\partial N \partial \rho} = -\frac{2\left\{\rho x^2 + \rho y^2 - \left(1+\rho^2\right)xy\right\}}{\left(1-\rho^2\right)^2}\left(1+\frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right)^{-1},$$

$$\frac{\partial^2 \log L}{\partial m^2} = \frac{1}{m^2} - \frac{2N\left(x^2+y^2-2\rho xy\right)}{m^3\left(1-\rho^2\right)}\left(1+\frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right)^{-1} + \frac{N\left(x^2+y^2-2\rho xy\right)^2}{m^4\left(1-\rho^2\right)^2}\left(1+\frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right)^{-2},$$

$$\frac{\partial^2 \log L}{\partial m \partial \rho} = \frac{2N\left\{\rho x^2 + \rho y^2 - \left(1+\rho^2\right)xy\right\}}{m^2\left(1-\rho^2\right)^2}\left(1+\frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right)^{-1} - \frac{2N\left(x^2+y^2-2\rho xy\right)\left\{\rho x^2 + \rho y^2 - \left(1+\rho^2\right)xy\right\}}{m^3\left(1-\rho^2\right)^3}\left(1+\frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right)^{-2}$$

and

$$\frac{\partial^2 \log L}{\partial \rho^2} = \frac{1+\rho^2}{\left(1-\rho^2\right)^2} + \frac{2N\left[\left(1-\rho^2\right)\left(x^2+y^2-2\rho xy\right) + 4\rho\left\{\rho x^2 + \rho y^2 - \left(1+\rho^2\right)xy\right\}\right]}{m^2\left(1-\rho^2\right)^3} \times \left(1+\frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right)^{-1} \times \left[4\left\{\rho x^2 + \rho y^2 - \left(1+\rho^2\right)xy\right\}^2\left(1+\frac{x^2+y^2-2\rho xy}{m\left(1-\rho^2\right)}\right)^{-1} - m\right].$$

Now, we can compute the elements of the Fisher information matrix. It is clear that

$$E\left(-\frac{\partial^2 \log L}{\partial N^2}\right) = \frac{1}{(N-1)^2}.$$

By applying (9) and (17)–(19), one gets

$$E\left(-\frac{\partial^2 \log L}{\partial N \partial m}\right) = -\frac{(N-1)B(2,N-1)}{m}$$

and

$$E\left(-\frac{\partial^2 \log L}{\partial N \partial \rho}\right) = \frac{(N-1)B(2,N-1)m\rho}{1-\rho^2}.$$

By applying (9) and (17)–(24), one gets

$$E\left(-\frac{\partial^2 \log L}{\partial m^2}\right) = \frac{2N(N-1)B(2,N-1)+N(1-N)B(3,N-1)-1}{m^2}$$

and

$$E\left(-\frac{\partial^2 \log L}{\partial m \partial \rho}\right) = \frac{N(N-1)\{B(3,N-1)-B(2,N-1)\}\rho}{m\left(1-\rho^2\right)}.$$

Finally, application of (9), (17)–(19) and (25)–(31) yields

$$\begin{aligned} E\left(-\frac{\partial^2 \log L}{\partial \rho^2}\right) = \frac{1}{\left(1-\rho^2\right)^2}\Big\{ & 2N(1-N)B(4,N-2)m\rho^8 + 3N(1-N)B(4,N-2)m\rho^6 \\ & +11N(N-1)B(4,N-2)m\rho^4 + 5N(1-N)B(4,N-2)m\rho^2 \\ & +2N(N-1)B(2,N-1)\rho^2 - \rho^2 + 2N(N-1)B(2,N-1) \\ & +N(1-N)B(4,N-2)m - 1\Big\}. \end{aligned}$$

4 Information matrix for Kotz type

If (x,y) is a single observation from (3) then the log-likelihood function can be written as

$$\begin{aligned} \log L(N,r,s,\rho) = & \log s + \frac{N\log r}{s} - \log\pi - \log\Gamma\left(\frac{N}{s}\right) + \left(\frac{1}{2}-N\right)\log\left(1-\rho^2\right) \\ & +(N-1)\log\left(x^2+y^2-2\rho xy\right) - r\left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^s. \end{aligned}$$

The first-order derivatives are:

$$\frac{\partial \log L}{\partial N} = \frac{\log r}{s} - \frac{1}{s}\Psi\left(\frac{N}{s}\right) - \log\left(1-\rho^2\right) + \log\left(x^2+y^2-2\rho xy\right),$$

$$\frac{\partial \log L}{\partial r} = \frac{N}{rs} - \left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^s,$$

$$\frac{\partial \log L}{\partial s} = \frac{1}{s} - \frac{N\log r}{s^2} + \frac{N}{s^2}\Psi\left(\frac{N}{s}\right) - r\left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^s \log\left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)$$

and

$$\frac{\partial \log L}{\partial \rho} = \frac{(2N-1)\rho}{1-\rho^2} - \frac{2(N-1)xy}{x^2+y^2-2\rho xy} - \frac{2rs\left\{\rho x^2+\rho y^2-\left(1+\rho^2\right)xy\right\}}{\left(1-\rho^2\right)^2}\left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^{s-1}.$$

The second-order derivatives are:

$$\frac{\partial^2 \log L}{\partial N^2} = -\frac{1}{s^2}\Psi'\left(\frac{N}{s}\right),$$

$$\frac{\partial^2 \log L}{\partial N \partial r} = \frac{1}{rs},$$

$$\frac{\partial^2 \log L}{\partial N \partial s} = -\frac{\log r}{s^2} + \frac{1}{s^2}\Psi\left(\frac{N}{s}\right) + \frac{N}{s^3}\Psi'\left(\frac{N}{s}\right),$$

$$\frac{\partial^2 \log L}{\partial N \partial \rho} = \frac{2\rho}{1-\rho^2} - \frac{2xy}{x^2+y^2-2\rho xy},$$

$$\frac{\partial^2 \log L}{\partial r^2} = -\frac{N}{r^2 s},$$

$$\frac{\partial^2 \log L}{\partial r \partial s} = -\frac{N}{rs^2} - \left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^s \log\left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right),$$

$$\frac{\partial^2 \log L}{\partial r \partial \rho} = -\frac{2s\left\{\rho x^2+\rho y^2-\left(1+\rho^2\right)xy\right\}}{\left(1-\rho^2\right)^2}\left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^{s-1},$$

$$\frac{\partial^2 \log L}{\partial s^2} = -\frac{1}{s^2} + \frac{2N \log r}{s^3} - \frac{2N}{s^3}\Psi\left(\frac{N}{s}\right) - \frac{N^2}{s^4}\Psi'\left(\frac{N}{s}\right) - r\left\{\log\left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)\right\}^2 \left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^s,$$

$$\frac{\partial^2 \log L}{\partial s \partial \rho} = -\frac{2r\left\{\rho x^2 + \rho y^2 - \left(1+\rho^2\right)xy\right\}}{\left(1-\rho^2\right)^2}\left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^{s-1} \times\left\{1 + s\log\left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)\right\},$$

and

$$\frac{\partial^2 \log L}{\partial \rho^2} = \frac{(2N-1)\left(1+\rho^2\right)}{\left(1-\rho^2\right)^2} - \frac{4(N-1)x^2y^2}{\left(x^2+y^2-2\rho xy\right)^2} - \frac{4rs(s-1)\left\{\rho x^2+\rho y^2-\left(1+\rho^2\right)xy\right\}^2}{\left(1-\rho^2\right)^4} \times\left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^{s-2} - \frac{2rs\left[\left(1-\rho^2\right)\left(x^2+y^2-2\rho xy\right) + 4\rho\left\{\rho x^2+\rho y^2-\left(1+\rho^2\right)xy\right\}\right]}{\left(1-\rho^2\right)^3} \times\left(\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right)^{s-1}.$$

Now, we can compute the elements of the Fisher information matrix. It is clear that

$$E\left(-\frac{\partial^2 \log L}{\partial N^2}\right) = \frac{1}{s^2}\Psi'\left(\frac{N}{s}\right).$$

$$E\left(-\frac{\partial^2 \log L}{\partial N \partial r}\right) = -\frac{1}{rs},$$

$$E\left(-\frac{\partial^2 \log L}{\partial N \partial s}\right) = \frac{\log r}{s^2} - \frac{1}{s^2}\Psi\left(\frac{N}{s}\right) - \frac{N}{s^3}\Psi'\left(\frac{N}{s}\right),$$

and

$$E\left(-\frac{\partial^2 \log L}{\partial r^2}\right) = \frac{N}{r^2 s}.$$

By applying (10) and (19), one gets

$$E\left(-\frac{\partial^2 \log L}{\partial N \partial \rho}\right) = -\frac{\rho}{1-\rho^2}.$$

By applying (10) and (17)–(19), one gets

$$E\left(-\frac{\partial^2 \log L}{\partial r \partial \rho}\right) = \frac{N\rho}{r\left(1-\rho^2\right)}.$$

By applying (10) and (17)–(24), one gets

$$E\left(-\frac{\partial^2 \log L}{\partial \rho^2}\right) = \frac{1+Ns\left(1+2\rho^2\right)}{2\left(1-\rho^2\right)^2}.$$

By applying (13), one gets

$$E\left(-\frac{\partial^2 \log L}{\partial r \partial s}\right) = \frac{N}{rs^2} + \frac{N\left\{\Psi\left(1+\frac{N}{s}\right)-\log r\right\}}{rs^2\left(1-\rho^2\right)^{N-1}}$$

and

$$E\left(-\frac{\partial^2 \log L}{\partial s^2}\right) = \frac{1}{s^2} - \frac{2N\log r}{s^3} + \frac{2N}{s^3}\Psi\left(\frac{N}{s}\right) + \frac{N^2}{s^4}\Psi'\left(\frac{N}{s}\right) + \frac{N\left\{\Psi'\left(1+\frac{N}{s}\right)+\Psi^2\left(1+\frac{N}{s}\right)-2\log r\Psi\left(1+\frac{N}{s}\right)+(\log r)^2\right\}}{s^3\left(1-\rho^2\right)^{N-1}}.$$

Finally, application of (17)–(19), (13) and (14) yields

$$E\left(-\frac{\partial^2 \log L}{\partial s \partial \rho}\right) = \frac{N\rho}{s\left(1-\rho^2\right)} + \frac{2N\rho\left\{\Psi\left(1+\frac{N}{s}\right)-\log r\right\}}{s\left(1-\rho^2\right)^N} - \frac{N\rho\left\{\Psi\left(1+\frac{N}{s}\right)-\log r\right\}}{s\left(1-\rho^2\right)^{N+1}}.$$

5 Appendix

The following relationships are needed for the calculation of the elements of the Fisher information matrices. These relations follow directly from the transformation (4).

$$E\left(X^2\right) = \frac{1+\sqrt{1-\rho^2}}{2}E\left(U^2\right)+\frac{1-\sqrt{1-\rho^2}}{2}E\left(V^2\right), \quad (17)$$

$$E\left(Y^2\right) = \frac{1-\sqrt{1-\rho^2}}{2}E\left(U^2\right)+\frac{1+\sqrt{1-\rho^2}}{2}E\left(V^2\right), \quad (18)$$

$$E\left(XY\right) = \frac{\rho}{2}\left\{E\left(U^2\right)+E\left(V^2\right)\right\}, \quad (19)$$

$$E\left(X^4\right) = \frac{2\sqrt{1-\rho^2}+2-\rho^2}{4}E\left(U^4\right)-\frac{2\sqrt{1-\rho^2}-2+\rho^2}{4}E\left(V^4\right)+\frac{3\rho^2}{2}E\left(U^2V^2\right), \quad (20)$$

$$E\left(Y^4\right) = \frac{2\sqrt{1-\rho^2}+2-\rho^2}{4}E\left(V^4\right)-\frac{2\sqrt{1-\rho^2}-2+\rho^2}{4}E\left(U^4\right)+\frac{3\rho^2}{2}E\left(U^2V^2\right), \quad (21)$$

$$E\left(XY^3\right) = \frac{\rho\left\{1-\sqrt{1-\rho^2}\right\}}{4}E\left(U^4\right)+\frac{\rho\left\{1+\sqrt{1-\rho^2}\right\}}{4}E\left(V^4\right)+\frac{3\rho}{2}E\left(U^2V^2\right), \quad (22)$$

$$E\left(X^3Y\right) = \frac{\rho\left\{1-\sqrt{1-\rho^2}\right\}}{4}E\left(V^4\right)+\frac{\rho\left\{1+\sqrt{1-\rho^2}\right\}}{4}E\left(U^4\right)+\frac{3\rho}{2}E\left(U^2V^2\right), \quad (23)$$

$$E\left(X^2Y^2\right) = \frac{\rho^2}{4}\left\{E\left(U^4\right)+E\left(V^4\right)\right\}+\left(1+\frac{\rho^2}{2}\right)E\left(U^2V^2\right), \quad (24)$$

$$E\left(X^6\right) = \frac{4-3\rho^2+\left(4-\rho^2\right)\sqrt{1-\rho^2}}{8}E\left(U^6\right)+\frac{4-3\rho^2-\left(4-\rho^2\right)\sqrt{1-\rho^2}}{8}E\left(V^6\right)$$
$$+\frac{15\rho^2\left\{1-\sqrt{1-\rho^2}\right\}}{8}E\left(U^2V^4\right)+\frac{15\rho^2\left\{1+\sqrt{1-\rho^2}\right\}}{8}E\left(U^4V^2\right), \quad (25)$$

$$E\left(Y^6\right) = \frac{4-3\rho^2+\left(4-\rho^2\right)\sqrt{1-\rho^2}}{8}E\left(V^6\right) + \frac{4-3\rho^2-\left(4-\rho^2\right)\sqrt{1-\rho^2}}{8}E\left(U^6\right)$$
$$+\frac{15\rho^2\left\{1-\sqrt{1-\rho^2}\right\}}{8}E\left(U^4V^2\right) + \frac{15\rho^2\left\{1+\sqrt{1-\rho^2}\right\}}{8}E\left(U^2V^4\right), \quad (26)$$

$$E\left(XY^5\right) = \frac{\rho\left\{2-\rho^2-2\sqrt{1-\rho^2}\right\}}{8}E\left(U^6\right) + \frac{\rho\left\{2-\rho^2+2\sqrt{1-\rho^2}\right\}}{8}E\left(V^6\right)$$
$$+\frac{5\rho\left\{2+\rho^2+2\sqrt{1-\rho^2}\right\}}{8}E\left(U^2V^4\right)$$
$$+\frac{5\rho\left\{2+\rho^2-2\sqrt{1-\rho^2}\right\}}{8}E\left(U^4V^2\right), \quad (27)$$

$$E\left(X^5Y\right) = \frac{\rho\left\{2-\rho^2-2\sqrt{1-\rho^2}\right\}}{8}E\left(V^6\right) + \frac{\rho\left\{2-\rho^2+2\sqrt{1-\rho^2}\right\}}{8}E\left(U^6\right)$$
$$+\frac{5\rho\left\{2+\rho^2+2\sqrt{1-\rho^2}\right\}}{8}E\left(U^4V^2\right)$$
$$+\frac{5\rho\left\{2+\rho^2-2\sqrt{1-\rho^2}\right\}}{8}E\left(U^2V^4\right), \quad (28)$$

$$E\left(X^2Y^4\right) = \frac{\rho^2\left\{1-\sqrt{1-\rho^2}\right\}}{8}E\left(U^6\right) + \frac{\rho^2\left\{1+\sqrt{1-\rho^2}\right\}}{8}E\left(V^6\right)$$
$$+\frac{4+11\rho^2+\left(4+\rho^2\right)\sqrt{1-\rho^2}}{8}E\left(U^2V^4\right)$$
$$+\frac{4+11\rho^2-\left(4+\rho^2\right)\sqrt{1-\rho^2}}{8}E\left(U^4V^2\right), \quad (29)$$

$$E\left(X^4Y^2\right) = \frac{\rho^2\left\{1-\sqrt{1-\rho^2}\right\}}{8}E\left(V^6\right)+\frac{\rho^2\left\{1+\sqrt{1-\rho^2}\right\}}{8}E\left(U^6\right)$$
$$+\frac{4+11\rho^2+\left(4+\rho^2\right)\sqrt{1-\rho^2}}{8}E\left(U^4V^2\right)$$
$$+\frac{4+11\rho^2-\left(4+\rho^2\right)\sqrt{1-\rho^2}}{8}E\left(U^2V^4\right), \quad (30)$$

and

$$E\left(X^3Y^3\right) = \frac{\rho^3}{8}\left\{E\left(U^6\right)+E\left(V^6\right)\right\}+\frac{3\rho\left(4+\rho^2\right)}{8}\left\{E\left(U^2V^4\right)+E\left(U^4V^2\right)\right\}. \quad (31)$$

7 References

Fang, K. -T., Kotz, S. and Ng, K. W. (1990). *Symmetric Multivariate and Related Distributions*. London: Chapman and Hall.

Johnson, M. E. (1987). *Multivariate Statistical Simulation*. New York: John Wiley and Sons.

Kotz, S. (1975). Multivariate distributions at a cross-road. In *Statistical Distributions in Scientific Work*, **1**, editors G. P. Patil, S. Kotz and J. K. Ord. Dordrecht (eds)., The Netherlands: D. Reidel Publishing Company.

Kotz, S. and Nadarajah, S. (2004). *Multivariate t Distributions and Their Applications*. New York: Cambridge University Press.

Nadarajah, S. (2003). The Kotz type distribution with applications. *Statistics*, 37, 341–358.

Prudnikov, A. P., Brychkov, Y. A. and Marichev, O. I. (1986). *Integrals and Series* (volumes 1, 2 and 3). Amsterdam: Gordon and Breach Science Publishers.

SORT 29 (1) January-June 2005, 57-100

Automatic error localisation for categorical, continuous and integer data

Ton de Waal*

Netherlands

Abstract

Data collected by statistical offices generally contain errors, which have to be corrected before reliable data can be published. This correction process is referred to as statistical data editing. At statistical offices, certain rules, so-called edits, are often used during the editing process to determine whether a record is consistent or not. Inconsistent records are considered to contain errors, while consistent records are considered error-free. In this article we focus on automatic error localisation based on the Fellegi-Holt paradigm, which says that the data should be made to satisfy all edits by changing the fewest possible number of fields. Adoption of this paradigm leads to a mathematical optimisation problem. We propose an algorithm for solving this optimisation problem for a mix of categorical, continuous and integer-valued data. We also propose a heuristic procedure based on the exact algorithm. For five realistic data sets involving only integer-valued variables we evaluate the performance of this heuristic procedure.

MSC: 03B05, 03B35, 62-02, 68T15, 90C10, 90C11

Keywords: branch-and-bound, categorical data, continuous data, error localisation, Fourier-Motzkin elimination, integer-valued data, statistical data editing.

1 Introduction

Data collected by statistical offices generally contain errors. In order to be able to publish reliable statistical information these errors have to be corrected. This correction process is referred to as statistical data editing. At statistical offices, certain rules, so-called edits, are often used to determine whether a record, *i.e*. the data of an individual respondent, is consistent or not. An example of such an edit is that the sum of the profit and the

* *Address for correspondence:* Statistics Netherlands. PO Box 4000. 2270 JM Voorburg. Netherlands, tel. +31 70 337 4930, fax +31 70 337 5964, e-mail: twal@cbs.nl.

Received: August 2004

Accepted: January 2005

costs of an enterprise should equal its turnover. Inconsistent records are considered to contain errors, while consistent records are considered error-free. If a record contains errors, the erroneous fields in this record need to be identified. In former days and often still nowadays, detected errors or inconsistencies were reported and explained on paper computer output or on a computer screen. Subsequently, subject-matter specialists corrected the errors by consulting the questionnaire, or by re-contacting the supplier of the information. This traditional form of statistical data editing, called manual (or interactive) editing, leads to statistical data of good quality, but is very costly in terms of resources and timeliness.

Several studies (*cf.* Granquist, 1995, 1997; Granquist and Kovar, 1997) have demonstrated that in order to obtain reliable publication figures only the most influential errors have to be edited manually. This observation, which has been confirmed by practical experience at several statistical offices, allows one to improve the efficiency of the statistical data editing process. For instance, at Statistics Netherlands most structural business surveys are nowadays treated by a combination of selective (or significance) editing (*cf.* Lawrence and McKenzie, 2000; Hoogland, 2002; Hedlin, 2003), automatic editing, and macro-editing (*cf.* Granquist, 1990). After data entry, simple checks, such as range checks, and simple automatic corrections, for instance in cases where a respondent filled in a financial figure in Euros instead of the requested thousands of Euros, are applied. Next, selective editing is applied to split the data into a critical stream and a non-critical stream. The former stream consists of those records that are the most likely ones to contain influential errors; the latter stream consists of the remaining records. The records in the critical stream are edited in a traditional, manual manner. The records in the non-critical stream are edited automatically. The final editing step we apply is macro-editing. Macro-editing consists of verifying whether the figures to be published seem plausible. Macro-editing is applied after outliers have been detected, raising weights have been determined, and the publication figures have been computed. It can lead to the detection of errors that would go unnoticed with selective editing or automatic editing. Only after the macro-editing step has been successfully completed can the publication figures be published. For more information on the statistical data editing strategy of Statistics Netherlands for structural business surveys we refer to De Jong (2002) and Hoogland (2002).

In this article we focus on automatic editing. Generally speaking, statistical data editing can be subdivided into the error localisation step and the imputation step. In the error localisation step the errors in the data are detected, in the imputation step the erroneous data are replaced by more accurate data and the missing values are filled in. To automate the statistical data editing process, both error localisation and imputation need to be automated.

In this article we restrict ourselves to discussing how to automate the error localisation step. To this end, one generally uses the (generalised) paradigm of Fellegi and Holt (1976) as a guiding principle to identify the errors. This (generalised) paradigm

says that the data of a record should be made to satisfy all edits by changing the fewest possible (weighted) number of fields. Here each variable is given a weight, the so-called reliability weight, which is a measure for the level of confidence in the values of this variable. The higher the reliability weight of a variable, the more reliable its observed values are considered to be. In the original form of the Fellegi-Holt paradigm each variable was given a reliability weight of 1. Using the (generalised) Fellegi-Holt paradigm, as we do in the present article, the error localisation problem can be formulated as a mathematical optimisation problem.

Solving this mathematical optimisation problem is, however, a non-trivial matter. Overviews of various algorithms for solving the error localisation problem based on the Fellegi-Holt paradigm have been given by Liepins, Garfinkel and Kunnathur (1982), and De Waal and Coutinho (2005). Most algorithms and software packages for solving this problem described in the literature are either designed for categorical (discrete) data or for continuous data. Algorithms for categorical data have been proposed by Fellegi and Holt (1976), Garfinkel, Kunnathur and Liepins (1986), Winkler (1998), Bruni, Reale and Torelli (2001), Bruni and Sassano (2001), and Boskovitz, Goré and Hegland (2003). Software packages for categorical data include SCIA (*cf.* Barcaroli *et al.*, 1995) by ISTAT, and DISCRETE (*cf.* Winkler and Petkunas, 1997) by the US Bureau of the Census. Algorithms for solving the problem for continuous data have been proposed by Fellegi and Holt (1976), Sande (1978), McKeown (1984), Garfinkel, Kunnathur and Liepins (1988), Ragsdale and McKeown (1996), and Riera-Ledesma and Salazar-González (2003). Software packages for continuous data include GEIS (*cf.* Kovar and Whitridge, 1990) by Statistics Canada, SPEER (*cf.* Winkler and Draper, 1997) by the US Bureau of the Census, AGGIES (*cf.* Todaro, 1999) by NASS, a SAS program developed by the Central Statistical Office of Ireland (*cf.* Central Statistical Office, 2000), and CherryPi (*cf.* De Waal, 1996) by Statistics Netherlands. The latter program is nowadays a module of version 1.0 of SLICE, a general software framework for automatic editing and imputation developed by Statistics Netherlands (*cf.* De Waal, 2001). Algorithms for solving the error localisation problem in a mix of categorical and continuous data are proposed by Sande (1978), Schaffer (1987), De Waal (2003a and 2003b), and De Waal and Quere (2003). In the present article we extend the latter algorithm to a mix of categorical, continuous, and *integer-valued* data. With the exception of De Waal (2003a), part of which formed the basis of the present article, such an algorithm has not yet been described in the literature before.

The remainder of this article is organised as follows. Section 2 sketches the error localisation problem by means of an example. Section 3 describes the edits we consider in this article. Section 4 formulates the error localisation problem as a mathematical optimisation problem. Section 5 sketches the algorithm for solving the error localisation problem in a mix of categorical and continuous data proposed by De Waal and Quere (2003). Section 7 extends this algorithm to a mix of categorical, continuous and integer data. Essential in this extended algorithm is Fourier-Motzkin elimination for

integer data, which we describe in Section 6. This elimination method is due to Pugh (*cf*. Pugh, 1992; Pugh and Wonnacott, 1994), who applied this technique to develop so-called array data dependence testing algorithms. Section 8 discusses a heuristic approach based on the exact algorithm described in Section 7. This heuristic procedure is easier to implement and maintain than the exact algorithm. Computational results for this heuristic procedure are given in Section 9. We conclude the article with a brief discussion in Section 10. In the present article we do not discuss the statistical quality of automatically edited data. For evaluation studies of this aspect of automatic editing we refer to Hoogland and Van der Pijll (2003), and Pannekoek and De Waal (2005).

2 An illustration of the error localisation problem

In this section we illustrate the error localisation problem for a mix of continuous and integer data by means of an example. We also sketch the idea of our solution method for such data, which basically consists of testing whether all integer-valued variables involved in a solution to the corresponding continuous error localisation problem, *i.e*. the error localisation problem where all numerical variables are assumed to be continuous, can indeed attain integer values.

Suppose a set of edits is given by

$$T = P + C, \tag{1}$$

$$0.5T \leq C, \tag{2}$$

$$C \leq 1.1T, \tag{3}$$

$$T \leq 550N, \tag{4}$$

$$320N \leq C \tag{5}$$

$$T \geq 0, \tag{6}$$

$$C \geq 0, \tag{7}$$

$$N \geq 0, \tag{8}$$

where T denotes the turnover of an enterprise, P its profit, C its costs, and N the number

of employees. The turnover, profit and costs are continuous variables, the number of employees an integer one. The original edits, (1) to (8) in this case, are called the *explicit* edits.

Let us consider a specific record with values $T = 5{,}060$, $P = 2{,}020$, $C = 3{,}040$ and $N = 5$. This record fails edit (4). We apply the Fellegi-Holt paradigm, and try to make the record satisfy all edits by changing as few variables as possible. As T and N occur in the failed edit, it might be possible to satisfy all edits by changing the value of one of these variables only. However, if we were to change the value of T, we would also need to change the value of P or C in order not to violate (1). We therefore start by considering the option of changing N. We first treat N as a continuous variable. To test then whether N can be changed so that all edits (1) to (8) become satisfied, we eliminate N by means of Fourier-Motzkin elimination (*cf.* Duffin, 1974; Chvátal, 1983; Schrijver, 1986; see also Section 5 of the present article). We combine all upper bounds on N (in this case only (5)) with all lower bounds on N (in this case (4) and (8)), to eliminate N from these edits. We obtain a new constraint, given by

$$320T \leq 550C \quad \text{(combination of (4) and (5))} \tag{9}$$

An edit such as (9) that is implied by the original set of edits is called an *implicit*, or *implied*, edit.

The constraints not involving N, *i.e.* (1), (2), (3), (6), (7) and (9) are all satisfied by the original values of T, P and C. A fundamental property of Fourier-Motzkin elimination is that a set of (in)equalities can be satisfied if and only if the set of (in)equalities after the elimination of a variable can be satisfied. This implies that the edits (1) to (8) can be satisfied by changing the value of N only. That is, if N were continuous, the (only) optimal solution to the above error localisation problem would be: change the value of N. However, N is an integer-valued variable. So, we need to test whether a feasible *integer* value for N exists. By filling in the values for T, P, and C in (4) and (5) we find $9.2 \leq N \leq 9.8$. In other words, a feasible integer value for N does not exist. Changing the value of N is hence not a solution to this error localisation problem.

The next best solution to the continuous error localisation problem is given by: change the values of T, P and C (see Section 5 for an algorithm to obtain this solution). This is obviously also a feasible solution to the error localisation problem for continuous and integer data under consideration, as in this solution variable N retains its original value, *i.e.* 5, which is integer. It is the (only) optimal solution to our problem as this is the best solution to the corresponding continuous error localisation problem for which all integer-valued variables can indeed attain integer values.

In this example it is quite easy to check whether a solution to the continuous error localisation problem is also a solution to the error localisation problem for continuous and integer data. In general, this is not the case, however. In Sections 6 and 7 we describe in detail how to test whether integer variables involved in a solution to the continuous error localisation problem can indeed attain feasible integer values.

3 The edits

3.1 Formal definition of edits

We denote the categorical variables by v_i (i=1,...,m) and the numerical variables by x_j (j=1,...,n). For categorical data we denote the domain, *i.e.* the set of the possible values, of variable i by D_i. I denotes the index set of the integer variables, *i.e.* x_j (j=1,...,n) is an integer-valued variable if and only if $j \in I$ and a continuous variable otherwise. We assume that each edit k (k=1, ..., K) is given by

$$\text{IF } v_i \in F_i^k \text{ (for all } i=1,\ldots,m) \text{ THEN } (x_1,\ldots,x_n) \in \{\mathbf{x} | a_{1k}x_1 + \cdots + a_{nk}x_n + b_k \geq 0\}, \quad (10a)$$

or by

$$\text{IF } v_i \in F_i^k \text{ (for all } i=1,\ldots,m) \text{ THEN } (x_1,\ldots,x_n) \in \{\mathbf{x} | a_{1k}x_1 + \cdots + a_{nk}x_n + b_k = 0\}. \quad (10b)$$

Edit k is satisfied by a record (v_1,..., v_m, x_1,..., x_n) if (10a), respectively (10b) holds true. The a_{jk} (j=1,...,n; k=1,...,K) and the b_j (j=1,...,n) are assumed to be rational numbers. $F_i^k \subseteq D_i$ for all $i$and k. We often refer to edits of type (10a) as *inequality* edits and to edits of type (10b) as *balance* edits.

The condition after the IF-statement, *i.e.* "$v_i \in F_i^k$ (for all i=1,...,m)", is called the IF-condition of the edit. The condition after the THEN-statement is called the THEN-condition. A categorical variable v_i is said to *enter* an edit k given by (10) if $F_i^k \subset D_i$ and $F_i^k \neq D_i$, *i.e.* if F_i^k is strictly contained in the domain of variable i. That edit is then said to be *involved with* this categorical variable. A numerical variable x_j is said to *enter* the THEN-condition of edit k given by (10) if $a_{jk} \neq 0$. That THEN-condition is then said to be *involved with* this numerical variable. By multiplying the a_{jk} (j=1,...,n; k=1,...,K) and the b_j (j=1,...,n) involved in the THEN-condition of the k-th (k=1,...,K) edit of type (10) by an appropriately chosen integer we can ensure that in each THEN-condition these coefficients become integral and that their greatest common divisor in this THEN-condition equals 1. Such an edit is then said to be *normalised*.

If the set in the THEN-condition of (10) is the entire n-dimensional real vector space, then the edit is always satisfied and may be discarded. If the set in the THEN-condition of (10) is empty, then the edit is failed by any record for which the IF-condition holds true, *i.e.* for any record for which $v_i \in F_i^k$ for all i=1,...,m. If a set $F_i^k = \emptyset$ for some i=1,...m, then the edit is by definition satisfied and may be discarded. If the IF-condition of an edit does not hold true for a particular record, the edit is satisfied, irrespective of the values of the numerical variables (provided they are not missing).

All edits given by (10) and all integrality constraints have to be satisfied simultaneously. We assume that the edits and integrality constraints can indeed be satisfied simultaneously. We also assume that for each variable entering the edits a value

has to be filled in. Any field for which the value is missing is hence considered to be erroneous. An edit involved with a missing value is considered failed. Any non-integral value for an integer-valued variable is also considered erroneous.

3.2 Examples of edits

Below we illustrate what kinds of edits can be expressed in the form (10) by means of a number of examples, which are taken from De Waal (2003b).

(i) $Turnover - Profit \geq 0$.

This is an example of a numerical edit. The edit can be formulated in our standard form as:

IF $v_i \in D_i$ (for all i=1,…,m) THEN ($Profit$, $Turnover$) $\in$ {($Profit$, $Turnover$)| $Turnover - Profit \geq 0$}.

In the remaining examples we will be less formal with our notation, as we will omit the terms "$v_i \in D_i$" from the edits.

(ii). IF (*Gender* = *"Male"*) THEN (*Pregnant* = *"No"*).

This is an example of a categorical edit. It can be formulated in our standard form as:

IF ((*Gender* = *"Male"*) AND (*Pregnant* = *"Yes"*)) THEN $\emptyset$.

(iii) IF (*Occupation* = *"Statistician"*) THEN (*Income* $\geq$ 1,000 Euro). (11)

This is an example of a mixed edit. Conditional on certain categorical values, a certain numerical constraint has to be satisfied.

(iv) IF ((*Occupation* = *"Statistician"*) OR (*Education* = *"University"*)) THEN (*Income* $\geq$ 1,000 Euro).

This edit can be split into two edits given by (11) and

IF (*Education* = *"University"*) THEN (*Income* $\geq$ 1,000 Euro).

(v) IF (*Tax on Wages* > 0) THEN (*Number of Employees* ≥ 1). (12)

Edit (12) is not in our standard form (10), because the IF-condition involves a numerical variable. To handle such an edit, we introduce an auxiliary categorical variable *TaxCond* with domain {"*false*", "*true*"} during a pre-processing step. Initially, *TaxCond* is set to "*true*" if *Tax on Wages* > 0 in the unedited record, and to "*false*" otherwise. Its reliability weight is set to zero. We now replace edit (12) by the following three edits of type (10):

IF (*TaxCond* = "*false*") THEN (*Tax on Wages* ≤ 0),
IF (*TaxCond* = "*true*") THEN (*Tax on Wages* $\geq \varepsilon$),
IF (*TaxCond* = "*true*") THEN (*Number of Employees* ≥ 1),

where ε is a sufficiently small positive number. The initial value of *TaxCond* may be considered erroneous by the algorithm proposed in this article.

4 A mathematical formulation of the error localisation problem

In this section we give a mathematical formulation of the error localisation problem for a mix of categorical, continuous and integer data. For each record $(v_1^0, \ldots, v_m^0, x_1^0, \ldots, x_n^0)$ in the data set that is to be edited automatically we have to determine, or more precisely: have to ensure the existence of, a synthetic record $(v_1, \ldots, v_m, x_1, \ldots, x_n)$ such that

$$\sum_{i=1}^{m} w_i^c \delta(v_i^0, v_i) + \sum_{j=1}^{n} w_j^r \delta(x_j^0, x_j) \tag{13}$$

is minimised subject to the conditions that all edits $k = 1,\ldots,K$ of type (10) become satisfied, x_j is integer for $j \in I$, and the remaining x_j are continuous. Here w_i^c is the reliability weight of categorical variable i (i=1,…,m), w_j^r the reliability weight of numerical variable j (j=1,…,n), $\delta(y^0, y) = 1$ if $y^0 \neq y$ or y^0 is missing, and $\delta(y^0, y) = 0$ if $y^0 = y$. The objective function (13) is the weighted number of fields that have to be changed. The variables for which the value in the synthetic record differs from the original value plus the variables for which the original value was missing together form an optimal solution to the error localisation problem. Whenever we refer to the error localisation problem in the remainder of this article, we will mean the above mathematical optimisation problem.

Note that if $w_i^c = 1$ for all i=1,…,n and w_j^r for all j=1,…,m, then minimising the objective function (13) reduces to the original paradigm of Fellegi and Holt (1976). The objective function (13) is the most natural generalisation of the paradigm of Fellegi and Holt.

Our aim is to find several, preferably all, optimal solutions to the error localisation problem. The reason for pursuing this goal instead of finding only one optimal solution is that the actual statistical problem of automatic statistical data editing is more comprehensive than the above optimisation problem. This statistical problem is the problem of obtaining high quality data from a data set with errors in an efficient manner. Statistical aspects, such as the probability distribution of the corrected data, need to be taken into account besides the (weighted) number of fields that need to be modified. By generating several optimal solutions to our optimisation problem, we gain the option to later use a secondary, statistical criterion to select one optimal solution

that is best from a statistical point of view (*cf*. Stoop, 2003). The variables involved in the selected optimal solution are set to missing. They are subsequently imputed during the imputation step, for instance by means of regression imputation or donor imputation (see Kalton and Kasprzyk, 1986, and Kovar and Whitridge, 1995, for an overview of imputation methods). In the present article we will not examine the process of selecting one optimal solution from several optimal solutions, neither will we examine the imputation process.

5 A branch-and-bound algorithm for categorical and continuous data

In this section we sketch the branch-and-bound algorithm for solving the error localisation problem for a mix of categorical and continuous data proposed by De Waal and Quere (2003). We assume for the moment that no values are missing. The algorithm is based on constructing a binary tree. An example of such a tree is given in Figure 1.

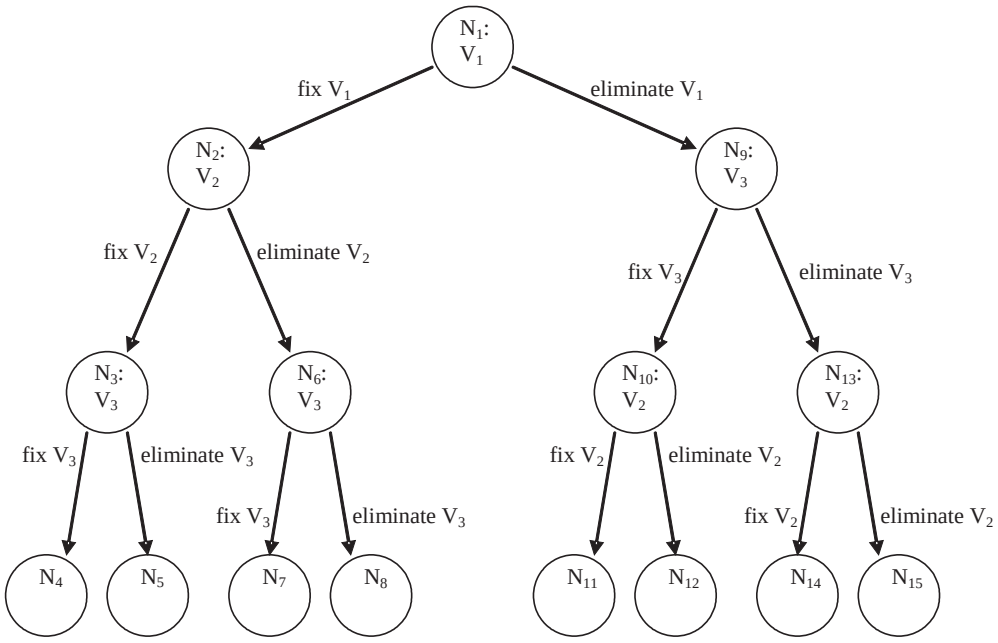

Figure 1: *A binary tree involving three variables.*

In each node of this tree we select a variable that has not yet been selected in any predecessor node. If all variables have been selected, we have reached a terminal node of the tree. After selection of a variable two branches are constructed: in one branch the selected variable is fixed to its original value, in the other branch the selected variable is eliminated from the set of current edits. For instance, in node N_1 of Figure 2 variable V_1

is selected. In the left-hand branch variable V_1 is fixed to its original value, in the right-hand branch V_1 is eliminated. A variable that has either been fixed or eliminated is said to have been *treated* (for the corresponding branch of the tree). Fixing a variable to its original value corresponds to assuming that this value is correct, eliminating a variable from the set of current edits corresponds to assuming that the original value of this variable is incorrect and has to be modified. In the algorithm, all continuous variables are selected before any categorical variable is. When a variable is fixed or eliminated, the set of current edits is updated. The set of edits corresponding to the root node of our tree is the original set of edits.

We now discuss how to update the set of current edits. We distinguish between fixing and eliminating a variable, and also between categorical and continuous variables. To fix a variable, either continuous or categorical, to its original value we substitute this value in all current edits. Note that, given that we fix this variable to its original value, the new set of current edits is a set of (implicit) edits for the remaining variables in the tree, *i.e*. the remaining variables have to satisfy the new set of edits. As a result of fixing the selected variable to its original value some edits may become satisfied, for instance when a categorical variable is fixed to a value such that the IF-condition of an edit can never become true anymore. These edits may be discarded from the new set of edits. Conversely, some edits may become violated. In such a case this branch of the binary tree cannot result in a solution to the error localisation problem.

Eliminating a variable amounts to generating a set of implicit edits that do not involve this variable. That set of implicit edits becomes the set of current edits corresponding to the new branch of the tree. If a continuous variable is to be eliminated, we basically apply Fourier-Motzkin elimination (*cf*. Duffin, 1974; Schrijver, 1986) to eliminate that variable from the set of edits. Care has to be taken in order to ensure that the IF-conditions of the resulting edits are correctly defined. In particular, if we want to eliminate a continuous variable x_r from the set of current edits, we start by copying all edits not involving x_r from the set of current edits to the new set of edits.

Next, we examine all edits involving x_r pair-wise. Suppose we consider the pair consisting of edits s and t. We start by checking whether the intersection of the IF-conditions is non-empty, *i.e*. whether the intersections $F_i^s \cap F_i^t$ are non-empty for all i=1,...,m. If any of these intersections is empty, we do not consider this particular combination of edits anymore. If all intersections are non-empty, we try to construct an implicit edit. If edit s is a balance edit (10*b*), we use the equality

$$x_r = -\frac{1}{a_{rs}} \left(b_s + \sum_{j \neq r} a_{js} x_j \right)$$

to eliminate x_r from the THEN-condition of edit t. Similarly, if edit s is an inequality edit (10*a*) and edit t is a balance equality, the equality in the THEN-condition of edit t is used to eliminate x_r from the THEN-condition of edit s.

If both edits s and t are inequality edits, we check whether the coefficients of x_r in those inequalities have opposite signs, *i.e.* whether $a_{rs} \times a_{rt} < 0$. If the coefficients of x_r in the two inequalities do not have opposite signs, we do not consider this particular combination of edits anymore. If the coefficients of x_r in the two inequalities do have opposite signs, one of the inequalities can be written as a lower bound on x_r and the other as an upper bound on x_r. Combining these two bounds leads to an inequality not involving x_r. To be precise, we generate the THEN-condition:

$$(x_1, \ldots, x_n) \in \{\mathbf{x} | \tilde{a}_1 x_1 + \cdots + \tilde{a}_n x_n + \tilde{b} \geq 0\}, \tag{14}$$

where

$$\tilde{a}_j = |a_{rs}| \times a_{jt} + |a_{rt}| \times a_{js} \quad \text{for all} \quad j = 1, \ldots, n$$

and

$$\tilde{b} = |a_{rs}| \times b_t + |a_{rt}| \times b_s.$$

The above THEN-condition forms the THEN-condition of a new implied edit. Note that x_r indeed does not enter this THEN-condition. The IF-condition of this implicit edit is given by the intersections $F_i^s \cap F_i^t$ for all i=1,...,m. Intuitively it is clear that the IF-condition of the new implicit edit is given by the *intersections* $F_i^s \cap F_i^t$ (i=1,...,m), because two (numerical) THEN-conditions can only be combined into the (numerical) THEN-condition of an implicit edit for the overlapping part of their corresponding categorical IF-conditions. Note that if we eliminate a continuous variable in any of the ways described above, the resulting set of edits is a set of implicit edits for the remaining variables in the tree. That is, this resulting set of edits has to be satisfied by the remaining variables in the tree. Repeatedly applying the above elimination process until all continuous variables have been eliminated results in edits with two kinds of THEN-conditions, namely edits with a THEN-condition that is trivially true, *e.g.* "$1 \geq 0$", and edits with a THEN-condition that is trivially false, *e.g.* "$0 \geq 1$". The edits with a THEN-condition of the former kind are deleted.

As we already mentioned before, categorical variables are treated, *i.e.* fixed or eliminated, after all continuous variables have been treated. So, when the categorical variables may be selected all edits in the set of current edits have the following form:

$$\text{IF} v_i \in F_i^k \quad (\text{for} \quad i = 1, \ldots, m) \ \text{THEN}(x_1, \ldots, x_n) \in \emptyset. \tag{15}$$

To eliminate categorical variable v_r from the set of edits given by (15), we start by copying all edits not involving v_r to the new set of edits. Next, we basically apply the method of Fellegi and Holt to the IF-conditions to generate the IF-conditions of the new edits (*cf.* Fellegi and Holt, 1976). In the terminology of Fellegi and Holt, field v_r is

selected as the generated field. We start by determining all index sets S such that

$$\bigcup_{k \in S} F_r^k = D_r \tag{16}$$

and

$$\bigcap_{k \in S} F_i^k \neq \emptyset \quad \text{for all} \quad i = 1, \ldots, r-1, r+1, \ldots, m. \tag{17}$$

From these index sets we select the *minimal* ones, *i.e.* the index sets S that obey (16) and (17), but none of their subsets obey (16). Given such a minimal index set S we construct the edit given by

$$\text{IF} v_r \in D_r, v_i \in \bigcap_{k \in S} F_i^k \quad (\text{for } i = 1, \ldots, r-1, r+1, \ldots, m) \text{ THEN} (x_1, \ldots, x_n) \in \emptyset.$$

Note that if we eliminate a categorical variable in the way described above, the resulting set of edits is a set of implicit edits for the remaining variables in the tree. That is, this resulting set of edits has to be satisfied by the remaining variables in the tree.

If values are missing in the original record, the corresponding variables only have to be eliminated (and not fixed) from the set of current edits. These variables are considered erroneous, and have to be imputed.

We have now explained how the set of current edits changes if we fix or eliminate a variable. After all categorical variables have been treated we are left with a set of relations involving no unknowns. This set of relations may be the empty set, in which case it obviously does not contain any self-contradicting relations. A self-contradicting relation is given by

$$\text{IF } v_i \in D_i \quad (\text{for } i = 1, \ldots, m) \text{ THEN } (x_1, \ldots, x_n) \in \emptyset.$$

We have the following theorem.

Theorem 1 *A set of relations obtained after all categorical variables have been treated contains no self-contradicting relations if and only if the variables that have been eliminated in order to reach the corresponding terminal node of the tree can be imputed consistently, i.e. modified such that all original edits can be satisfied.*

Theorem 1 follows from a repeated application of the following theorem, which is proved in De Waal and Quere (2003), and the fact that eliminating a variable amounts to generating a set of implied edits for the remaining variables

Theorem 2 *Suppose that the index set of the variables in a certain node is given by T_0, and the set of current edits corresponding to that node by Ω_0. Suppose furthermore that*

to obtain a next node variable r is either fixed or eliminated. Denote the index set of the resulting variables by T_1 ($T_1 = T_0 - \{r\}$) *and the set of edits corresponding to this next node by* Ω_1. *Now, if there exist values* u_i *for* $i \in T_1$ *that satisfy the edits in* Ω_1, *then there exists a value* u_r *for variable r such that the values* u_i *for* $i \in T_0$ *satisfy the edits in* Ω_0.

In the algorithm we check for each terminal node of the tree whether the variables that have been eliminated in order to reach this node can be imputed consistently. Of all sets of variables that can be imputed consistently we select the ones with the lowest sums of reliability weights. In this way we find all optimal solutions to the error localisation problem.

The algorithm may seem rather slow because an extremely large binary tree has to be generated to find all optimal solutions, even for moderately-sized problems. Fortunately, the situation is not nearly as bad as it may seem. First, balance edits can often be handled more efficiently than we described here (*cf.* De Waal and Quere, 2003; De Waal, 2003a). Second, if the minimum number of fields that have to be changed in order to make a record pass all edits is large, we feel that the record should not be edited automatically. In our opinion, the quality of such a record is too low to correct it automatically. We suggest that such a record should either be edited manually, or be discarded completely. This is similar to selective editing (*cf.* Lawrence and McKenzie, 2000; Hedlin, 2003) where (only) the very influential and very contaminated records are selected for manual correction. Such very influential and very contaminated records can be corrected by using subject-matter knowledge or, in the worst case, by re-contacting the supplier of the data. By specifying an upper bound on the number of fields that may be changed, the size of the tree can drastically be reduced. Third, the size of the tree can also be reduced during the execution of the algorithm, because it may already become clear in an intermediate node of the tree that the corresponding terminal nodes cannot generate an optimal solution to the problem. For instance, by fixing the wrong variables we may make the set of current edits infeasible, which may be noticed in an intermediate node. Fourth, the value of the objective function can be used as an incumbent in order to reduce the size of the tree. This value cannot decrease while going down the tree. So, if the objective value exceeds the value of an already found (possibly suboptimal) solution, we can again conclude that the corresponding terminal nodes cannot generate an optimal solution to the problem. In other words, the objective value of the best already found solution is used as the bound in our branch-and-bound scheme. During the execution of the algorithm the bound is updated.

In their article Fellegi and Holt (1976) describe a method for solving the error localisation problem that is also based on generating implicit edits. They propose to generate a, generally very large, set of implicit edits, which they refer to as the *complete set of edits*, for all records simultaneously before the error localisation problem is actually solved for each record. Given this "complete" set of edits, the error localisation problem can, for each record, be formulated as a set-covering problem (see, *e.g.*, Nemhauser and Wolsey, 1988, for more on the set-covering problem in general). In

contrast, in our algorithm we generate implicit edits while solving the error localisation problem for each record separately. The implicit edits are generated conditional on which variables have been selected for elimination and on the observed data in the record under consideration. As a result, our sets of implicit edits are generally much smaller than the "complete" set of edits generated by the approach of Fellegi and Holt. Over the years, generating this "complete" set of edits has often proven to be infeasible for large or even moderately-sized problems (*cf.* Winkler, 1996). This has been confirmed by some earlier experiments at Statistics Netherlands. Generating the smaller sets of implicit edits in our algorithm has turned out to be feasible for most records arising in practice. We refer to De Waal and Quere (2003) and Section 9 of the present article for confirmation of this assertion.

6 Fourier-Motzkin elimination in integer data

An important technique used in the algorithm described in Section 5 is Fourier-Motzkin elimination for eliminating a continuous variable from a set of linear (in)equalities. Fourier-Motzkin elimination can be extended to integer data in several ways. For example, Dantzig and Eaves (1973) and Williams (1976 and 1983) describe extensions of Fourier-Motzkin elimination to integer programming problems. Unfortunately, these methods seem too time-consuming in many practical cases. Pugh (1992) proposes an alternative extension that he refers to as the Omega test. Pugh (1992) and Pugh and Wonnacott (1994) claim a good performance of this test for many practical cases. Below we briefly explain the Omega test. For more details we refer to Pugh (1992), and Pugh and Wonnacott (1994).

The Omega test has been designed to determine whether an integer-valued solution to a set of linear (in)equalities exists. Suppose linear (in)equality k (k=1,...,K) is given by

$$a_{1k}x_1 + \cdots + a_{nk}x_n + b_k \geq 0,$$

or by

$$a_{1k}x_1 + \cdots + a_{nk}x_n + b_k = 0.$$

To simplify our notation we define $x_0 = 1$ and $a_{0k} = b_k$ (k=1,...,K), and re-write the above linear (in)equality as

$$a_{0k}x_0 + a_{1k}x_1 + \cdots + a_{nk}x_n \geq 0, \tag{18a}$$

respectively as

$$a_{0k}x_0 + a_{1k}x_1 + \cdots + a_{nk}x_n = 0. \tag{18b}$$

Without loss of generality we assume that redundant equalities have been removed, and that all (in)equalities are normalised, *i.e.* that all a_{jk} (j=0,...,n; k=1,...,K) are integer and the greatest common divisor of the a_{jk} in each constraint k equals 1. All variables x_j (j=0,...,n) are integer-valued in this section.

We start by "eliminating" all equalities until we arrive at a new problem involving only inequalities. In this context, we say that all equalities have been eliminated once we have transformed the original system of (in)equalities (18) into an equivalent system of (in)equalities of the following type:

$$x'_k = \sum_{j>k} a'_{jk} x'_j \quad \text{for } k = 0, \ldots, s-1, \tag{19a}$$

$$\sum_{j \geq s} a'_{jk} x'_j \geq 0 \quad \text{for } k = s, \ldots, K', \tag{19b}$$

where s is the number of equalities in the system (19), and the a'_{jk} are integer. The x'_j are a permutation of the x_j, possibly supplemented by some additional, auxiliary variables (see Subsection 6.1). We call a set of (in)equalities (18) equivalent to a set of (in)equalities (19) if a solution to the system (18) can be extended to a corresponding solution to the system (19), and conversely a solution to the system (19) is also a solution to the system (18) if we disregarded the additional variables. In (19), the first s x'_j, which are only involved in equalities, are expressed in terms of the remaining variables, which may also be involved in inequalities. Owing to the possible introduction of additional variables, the system (19) may have more equalities than the original system (18), so $K' \geq K$. The original system (18) has an integer-valued solution if and only if the system (19b) has an integer-valued solution. Namely, an integer solution for the x'_j ($j \geq s$) to the system (19b) yields an integer solution to the system (19), *i.e.* the system consisting (19a) plus (19b), by applying back-substitution to the x'_j ($j < s$). In other words, to check whether a system (18) has an integer-valued solution, we only need to check whether the inequalities (19b) of the equivalent system (19) have an integer-valued solution. In this sense the equalities of (18) have been eliminated once we have transformed a system given by (18) into an equivalent system given by (19).

6.1 Eliminating equalities

We now discuss how to eliminate an equality. As usual we denote the number of numerical, in this subsection: integer-valued, variables by n. We define the operation $c \overline{\bmod} d$ involving two integers c and d by

$$c \overline{\bmod} d = c - d \lfloor c/d + 1/2 \rfloor, \tag{20}$$

where $\lfloor u \rfloor$ denotes the largest integer less than or equal to u. If d is odd, the value of $c \overline{\bmod} d$lies in $[-(d-1)/2, (d-1)/2]$. If d is even, the value of $c \overline{\bmod} d$ lies in

$[-d/2, d/2-1]$. If $\lfloor c/d \rfloor < 1/2$, then $c \overline{\bmod} d = c \bmod d$. If $\lfloor c/d \rfloor \geq 1/2$, then $c \overline{\bmod} d = -c \bmod d$. Here, the mod d operator assumes values in [0, d-1].

To eliminate an equality s given by

$$\sum_{j=0}^{n} a_{js} x_j = 0, \tag{21}$$

we select an r such that $a_{rs} \neq 0$ and $|a_{rs}|$ has the smallest value among the a_{js} (j=0,...,n). If $|a_{rs}| = 1$, we eliminate the equality by using this equality to express x_r in terms of the other variables, and substitute this expression for x_r into the other (in)equalities. Otherwise, we define $\gamma = |a_{rs}| + 1$. Now we introduce a new variable σ defined by

$$\gamma\sigma = \sum_{j=0}^{n} (a_{js} \overline{\bmod} \gamma)\, x_j. \tag{22}$$

This variable σ is integer-valued. This can be shown as follows.

$$\sum_{j=0}^{n} (a_{js} \overline{\bmod} \gamma)\, x_j = \sum_{j=0}^{n} \left(a_{js} - \gamma \left\lfloor a_{js} \middle/ \gamma + 1/2 \right\rfloor \right) x_j = -\sum_{j=0}^{n} \gamma \left\lfloor a_{js} \middle/ \gamma + 1/2 \right\rfloor x_j, \tag{23}$$

where we have used (21). So, σ equals $-\sum_{j=0}^{n} \left\lfloor a_{js} \middle/ \gamma + 1/2 \right\rfloor x_j$, which is integer because the x_j (j=0,...,n) and their coefficients in (23) are integer.

It is easy to see that $a_{rs} \overline{\bmod} \gamma = -\text{sign}(a_{rs})$. Now, we use (22) to express x_r in terms of the other variables.

$$x_r = -\text{sign}(a_{rs}) q\sigma + \sum_{j=0, j\neq r}^{n} \text{sign}(a_{rs})(a_{js} \overline{\bmod} q)\, x_j \tag{24}$$

Substituting (24) into the original equality (21) gives

$$-|a_{rs}|\, \gamma\sigma + \sum_{j=0, j\neq r}^{n} (a_{js} + |a_{rs}|\, (a_{js} \overline{\bmod} \gamma))\, x_j = 0. \tag{25}$$

Because $|a_{rs}| = \gamma - 1$, (25) can be written as

$$-|a_{rs}|\, \gamma\sigma + \sum_{j=0, j\neq r}^{n} (a_{js} - (a_{js} \overline{\bmod} \gamma) + \gamma(a_{js} \overline{\bmod} \gamma))\, x_j = 0 \tag{26}$$

Using (20) on (26), and dividing by γ gives

$$-|a_{rs}|\,\sigma + \sum_{j=0, j\neq r}^{n} \left(\left\lfloor a_{js} \big/ \gamma + 1/2 \right\rfloor + (a_{js} \overline{\bmod}\ \gamma) \right) x_j = 0. \tag{27}$$

In (27) all coefficients are integer-valued.

It is clear that if the coefficient of variable x_j (j=0,…,n) equals zero in (21), the corresponding coefficient in (27) also equals zero. It is also clear that the absolute value of the coefficient of σ in (27) is equal to the absolute value of the coefficient of x_r in (21). However, for all other variables with a non-zero coefficient in (21) the absolute value of the corresponding coefficient in (27) is smaller than the absolute value of the coefficient in (21). To prove this statement we first re-write the coefficient of x_j ($j \neq r$) in (27) in the following way:

$$\left\lfloor a_{js} \big/ \gamma + 1/2 \right\rfloor + (a_{js} \overline{\bmod}\ \gamma) = \left\lfloor a_{js} \big/ \gamma + 1/2 \right\rfloor + a_{js} - \gamma \left\lfloor a_{js} \big/ \gamma + 1/2 \right\rfloor =$$

$$-|a_{rs}| \left\lfloor \frac{a_{js}}{|a_{rs}|+1} + \frac{1}{2} \right\rfloor + a_{js} \equiv \hat{a}_{js},$$

where we have used again that $\gamma = |a_{rs}| + 1$. We now consider the cases where a_{js} is positive and negative separately. If $a_{js} > 0$, then $a_{js} \geq |a_{rs}|$ by our choice of r. Suppose $a_{js} = \lambda\ |a_{rs}|$, where $\lambda \geq 1$. We then have

$$\hat{a}_{js} = a_{js} \left(-\frac{1}{\lambda} \left\lfloor \frac{\lambda\,|a_{rs}|}{|a_{rs}|+1} + \frac{1}{2} \right\rfloor + 1 \right).$$

Using

$$1 \leq \left\lfloor \frac{\lambda\,|a_{rs}|}{|a_{rs}|+1} + \frac{1}{2} \right\rfloor \leq \lambda$$

for all possible values of $|a_{rs}|$, we obtain $0 \leq \hat{a}_{js} \leq (1 - \frac{1}{\lambda}) a_{js}$. Hence, we can conclude that $\left|\hat{a}_{js}\right| < \left|a_{js}\right|$. In a similar way, one can show that if $a_{js} < 0$, then too $\left|\hat{a}_{js}\right| < \left|a_{js}\right|$. This is left for the reader to verify.

After a repeated application of the above substitution rule, where each time a new variable is introduced and an old variable is eliminated, to the original equality (21) and its derived form(s) (27), the equality is transformed into an equality in which (at least) one of the coefficients has absolute value 1. The corresponding variable can then be expressed in terms of the other variables. We substitute this expression into the other (in)equalities. The equality has then been eliminated.

This process continues until we have eliminated all equalities and we have obtained a system of the form (19). In the next subsection we explain how integer variables can

be eliminated from a set of linear inequalities (19b), but first we give an example of how equalities are eliminated.

Example 1 *We repeat part of an example given by Pugh (1992). In this example, four constraints have been specified:*

$$7x + 12y + 31z = 17, \tag{28}$$

$$3x + 5y + 14z = 7, \tag{29}$$

$$1 \leq x \leq 40, \tag{30}$$

Note that (30) stands for two inequalities. We wish to eliminate equality (28). Note that $\gamma = 8$*, and using (22) we introduce a variable* σ *defined by*

$$8\sigma = -x + 4y + z + 1. \tag{31}$$

We eliminate x from (28) to (30). Applying rule (27) to constraint (28) yields

$$-7\sigma - 2y + 3z = 3, \tag{32}$$

and applying rule (24) to constraints (29) and (30) yields

$$-24\sigma - 7y + 11z = 10, \tag{33}$$

$$1 \leq -8\sigma - 4y - z - 1 \leq 40. \tag{34}$$

The absolute values of the coefficients of y and zin (32) are smaller than the absolute values of the corresponding coefficients in (28). The system (31) to (34) is equivalent to the system (28) to (30).

6.2 Eliminating an integer variable from a set of inequalities

When an integer variable is eliminated from a set of inequalities involving only integer-valued variables, two different regions are determined. The first region is referred to as the *real shadow*. This is simply the region described by the set of inequalities that results if we apply the standard form of Fourier-Motzkin elimination. That is, the real shadow

results if we treat the integer variable that is being eliminated as continuous. The second region is referred to as the *dark shadow*. This dark shadow is constructed in such a way that if it contains a feasible (integer) solution, then the existence of a feasible (integer) solution to the original inequalities is guaranteed.

We describe the construction of the dark shadow. Suppose that two inequalities

$$ax \leq \alpha \tag{35}$$

and

$$bx \geq \beta \tag{36}$$

are combined to eliminate the integer variable x. Here a and b are positive integer constants, and α and β are linear expressions that may involve all variables except x. Each variable involved in α or β is assumed to have an integer coefficient. The real shadow obtained by eliminating x from the pair of inequalities (35) and (36) is defined by

$$a\beta \leq b\alpha. \tag{37}$$

We define the real shadow obtained by eliminating a variable $x$from a set of inequalities S to be the region described by the inequalities in S not involving x, and the inequalities (37) generated by all pairs of upper bounds (35) on x and lower bounds (36) on $x$in S.

Now, consider the case in which there is an integer value larger than or equal to $a\beta$ and smaller than or equal to $b\alpha$, but there is no integer solution for x to $a\beta \leq abx \leq b\alpha$. Let $q = \lfloor \beta/b \rfloor$, then by our assumptions we have

$$abq < a\beta \leq b\alpha < ab(q+1).$$

We clearly have $a(q+1) - \alpha > 0$. Since the values of a, b, α and β are integer, we have $a(q+1) - \alpha \geq 1$, and hence

$$ab(q+1) - b\alpha \geq b. \tag{38}$$

Similarly, we obtain

$$a\beta - abq \geq a. \tag{39}$$

Combining (38) and (39), we arrive at

$$b\alpha - a\beta \leq ab - a - b.$$

In other words, if

$$b\alpha - a\beta \geq ab - a - b + 1 = (a-1)(b-1), \tag{40}$$

then an integer solution for x necessarily exists.

To be able to satisfy (35) and (36) by choosing an appropriate integer value for x it is sufficient that (40) holds true. We therefore define the dark shadow obtained by eliminating variable x from the pair of inequalities (35) and (36) by the region described by (40). Note that if (40) holds true, there is an integer value larger than or equal to $a\beta$ and smaller than or equal to $b\alpha$. We define the dark shadow obtained by eliminating a variable $x$from a set of inequalities S to be the region described by the inequalities in S not involving x, and the inequalities (40) generated by all pairs of upper bounds (35) on x and lower bounds (36) on $x$in S.

We now consider a set of inequalities S with only integer-valued coefficients and variables. If the real shadow and the dark shadow resulting from the elimination of x from S are identical, we say that the elimination, or projection, is *exact*. In that case, an integer solution exists if and only if an integer solution to the real/dark shadow exists. If the real shadow and the dark shadow are not identical, we have the following possibilities:

- If the dark shadow has an integer solution, the set of inequalities S has an integer solution.
- If the real shadow does not contain a feasible (integer) solution, there is no integer solution to the set of inequalities S.
- In all other cases, it is not yet clear whether an integer solution to the set of inequalities S exists.

In the latter case we know that if an integer solution to the set of inequalities S were to exist, a pair of constraints $ax \leq \alpha$ and $\beta \leq bx$ would exist such that $ab-a-b \geq b\alpha - a\beta$ and $b\alpha \geq abx \geq a\beta$. From this we can conclude that in such a case an integer solution to the set of inequalities S would satisfy $ab - a - b + a\beta \geq abx \geq a\beta$. We can check whether an integer solution to the set of inequalities S exists by examining all possibilities. Namely, we determine the largest coefficient a_{max} of x for all upper bounds (35) on x. For each lower bound $\beta \leq bx$ we then test whether an integer solution exists to the original constraints S combined with $bx = \beta + p$ for each integer p satisfying $(a_{\max}b - a_{\max} - b)/a_{\max} \geq p \geq 0$. That is, in the latter case we examine $\lfloor (a_{\max}b - a_{\max} - b)/a_{\max} \rfloor + 1$ subproblems of the original problem. These subproblems are referred to as *splinters*.

The theory discussed so far shows that if the dark shadow or one of the splinters has an integer solution, then the original set of inequalities S has an integer solution. Conversely, because we examine all possibilities, it also holds true that if the original set of inequalities S has an integer solution then the dark shadow or one of the splinters has an integer solution. So, we have demonstrated the following theorem.

Theorem 3 *If and only if an integer solution to the dark shadow or one of the splinters exists, then an integer solution to the original set of inequalities S exists.*

Note that if the original set of inequalities S involves n integer variables, the dark shadow and the splinters involve only n-1 integer variables (for the splinters the added equality $bx = \beta + p$ first has to be eliminated in order to arrive at a system of inequalities involving n-1 variables). We have now explained how we can check whether a feasible integer value exists for an integer variable involved in a set of linear inequalities by eliminating this variable. In the next subsection we examine how we can test whether an integer solution exists for several variables simultaneously by eliminating these variables.

6.3 Eliminating several integer variables from a set of inequalities

Suppose we want to determine whether an integer solution exists for a set of linear inequalities involving n variables. We solve this problem by eliminating these n variables. During the elimination process the original problem may split into several subproblems owing to the splinters that may arise. We apply the procedure sketched below. We focus on the idea underlying the procedure; the computational efficiency of the procedure is ignored here.

We construct a list of subproblems. At the start of the procedure the only (sub)problem is the original problem involving all n variables. We treat each subproblem that may arise separately. We now consider one of those subproblems. We eliminate all variables involved in this subproblem by means of standard Fourier-Motzkin elimination, *i.e.* we repeatedly determine the real shadow until all variables have been eliminated. If the final real shadow without any unknowns is inconsistent, the subproblem does not have a continuous solution, let alone an integer solution. In such a case this subproblem can be discarded.

If the final real shadow of a subproblem is consistent, and a continuous solution hence exists, we examine the subproblem again and test whether there is an integer solution to this subproblem. For this subproblem we iteratively select a variable from the set of variables that have not yet been eliminated. The selected variable will be eliminated, using the method of Subsection 6.2. In order to keep the number of computations limited we choose the variable so that the elimination will be exact if possible. As a secondary aim we may then also minimise the number of constraints resulting from the combination of upper and lower bounds. If an exact elimination is

not possible, we select a variable with coefficients as close as possible to zero. For such a variable the number of splinters will be relatively small. Testing all splinters for integer solutions can be quite time-consuming, so creating splinters and testing them for integer solutions should be avoided as much as possible. For the subproblem under consideration, we determine the dark shadow and the splinters (if any) by eliminating the selected variable, using the method of Subsection 6.2. The dark shadow and the splinters define new subproblems, and are added to the list of subproblems. After this, we have dealt with the subproblem under consideration, and it is deleted from the list of subproblems. We continue this process until all variables have been eliminated from all subproblems on the list of subproblems. The final "subproblems", or better: final sets of relations, involve only numbers and no unknowns. As in the continuous case (see Section 5), such a relation can be self-contradicting, *e.g.* "$0 \geq 1$". We have the following theorem.

Theorem 4 *If any of the final sets of relations does not contain a self-contradicting relation, the original set of inequalities has an integer solution. Conversely, if all final sets of relations contain a self-contradicting relation, the original set of inequalities does not have an integer solution.*

Proof. This follows from a repeated application of Theorem 3. □

7 Error localisation in categorical, continuous and integer data

In this section we integrate the Omega test described in Section 6 with the branch-and-bound approach for solving the error localisation problem for categorical and continuous data proposed by Quere and De Waal (2003) (see Section 5). The result of this integration is an algorithm for solving the error localisation problem for categorical, continuous, and integer-valued data. The idea of this algorithm is to test whether the integer-valued variables involved in a solution to the continuous error localisation problem, *i.e.* the error localisation problem where all numerical variables are assumed to be continuous, can attain integer values. This is illustrated in Figure 2 below.

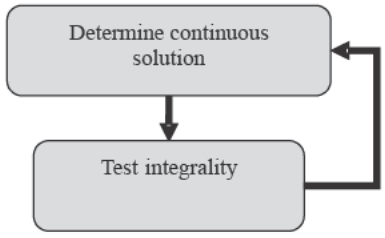

***Figure 2**: The basic idea of the error localisation algorithm.*

For a given combination of categorical values, our integrality test reduces to the Omega test. In other words, we basically apply the Omega test on each possible combination of categorical values. What complicates the issue is that we do not explicitly enumerate and test all possible combinations of categorical values. Before we describe the algorithm, we first explain in Subsection 7.1 how balance edits involving integer variables can be "eliminated" and in Subsection 7.2 how integer variables can be eliminated from inequality edits. Finally, Subsection 7.3 describes our algorithm for solving the error localisation problem for categorical, continuous and integer data.

As usual, the edits are given by (10). For notational convenience, we define $x_0 = 1$ and $a_{0k} = b_k$ for k=1,...,K, where K is the number of edits, like we also did in Section 6.

7.1 Error localisation: eliminating balance edits involving integer variables

In our integrality test (see Subsection 7.3) integer variables are treated after all continuous variables have been treated and before any categorical variable is treated. That is, once the integer variables are treated all edits involve only categorical and integer variables. If integer variables are involved in balance edits, we first "eliminate" these edits. We select a balance edit, and basically apply the technique explained in Subsection 6.1 to arrive at an equality in which the absolute value of the coefficient of an integer variable equals 1. During this process the IF-condition of the edit under consideration does not alter. To be more precise, if the selected edit s is given by

$$\text{IF} v_i \in F_i^s \quad (\text{for } i = 1, \ldots, m) \quad \text{THEN } (x_1, \ldots, x_n) \in \{\mathbf{x} | \sum_{j=0}^{n} a_{js} x_j = 0\}, \qquad (41)$$

with the a_{js} (j=0,...,n) integer coefficients and the x_j (j=0,...,n) integer variables, we transform this edit into

$$\text{IF } v_i \in F_i^s \quad (\text{for } i = 1, \ldots, m) \text{ THEN } (\tilde{x}_1, \ldots, \tilde{x}_{\tilde{n}}) \in \{\tilde{\mathbf{x}} | \sum_{j=0}^{\tilde{n}} \tilde{a}_{js} \tilde{x}_j = 0\}, \qquad (42)$$

where the $\tilde{a}_{js}$ (j=0,...,$\tilde{n}$) are integer coefficients, and the $\tilde{x}_j$ (j=0,..., $\tilde{n}$) are the transformed integer variables, possibly supplemented by some auxiliary integer variables owing to the elimination of the equality. The total number of variables $\tilde{x}_j$ is denoted by $\tilde{n}$ ($\tilde{n} \geq n$). In (42), at least one integer variable, say $\tilde{x}_r$, has a coefficient $\tilde{a}_{rs}$ with $|\tilde{a}_{rs}| = 1$. Below we describe the procedure to transform (41) into (42). For notational convenience, we write $\tilde{n}$ again as n. Likewise, we write the transformed coefficients $\tilde{a}_{js}$ (j=0,..., $\tilde{n}$) and transformed variables $\tilde{x}_j$ (j=0,..., $\tilde{n}$) again as a_{js} and x_j. It is important to keep in mind, though, that these coefficients and variables may differ from the original coefficients and variables.

Because auxiliary variables may need to be introduced during the elimination process of a balance edit, we may in fact need to introduce some auxiliary balance edits of which the THEN-conditions are given by equations of type (22) (or equivalently: of type (24)), and the IF-conditions by the IF-condition of the selected edit s. In each of these auxiliary equations, the new auxiliary variable is expressed in terms of the other integer variables x_j (j=1,...,n), *i.e.* the original integer variables and the already generated auxiliary variables. The other edits are written in terms of the new auxiliary variable by applying the substitution (24) to the numerical THEN-conditions as far as this is permitted by the IF-conditions. The IF-conditions of these other edits are changed by the substitution process. In particular, an edit t given by

$$\text{IF } v_i \in F_i^t \quad (\text{for } i = 1, \ldots, m) \text{ THEN } (x_1, \ldots, x_n) \in \{\mathbf{x} | \sum_{j=0}^{n} a_{jt} x_j \geq 0\}, \tag{43}$$

involving x_r in its THEN-condition gives rise to (at most) two edits given by

$$\begin{aligned} &\text{IF } v_i \in F_i^t \cap F_i^s \quad (\text{for } i = 1, \ldots, m) \\ &\text{THEN } (x_1, \ldots, x_n) \in \{\mathbf{x} | -\text{sign}(a_{rs}) a_{rt} \gamma\sigma + \sum_{j=0, j\neq r}^{n} (a_{jt} + sign(a_{rs}) a_{rt} (a_{js} \overline{\text{mod}}\ \gamma)) x_j \geq 0\}, \end{aligned} \tag{44}$$

and

$$\begin{aligned} &\text{IF } v_i \in F_i^t - F_i^s \quad (\text{for } i = 1, \ldots, m) \\ &\text{THEN } (x_1, \ldots, x_n) \in \{\mathbf{x} | \sum_{j=0}^{n} a_{jt} x_j \geq 0\}. \end{aligned} \tag{45}$$

In (43) to (45) the inequality sign may be replaced by an equality sign. Edits of type (44) for which $F_i^t \cap F_i^s = \emptyset$ (for some i=1,...,m), and edits of type (45) for which $F_i^t - F_i^s = \emptyset$ (for some i=1,...,m) may be discarded. Edits given by (43) not involving x_r are not modified.

Once we have obtained an edit of type (42) with a coefficient a_{rs} such that $|a_{rs}| = 1$, we use the THEN-condition of this edit to express the variable x_r in terms of the other variables. That is, we use

$$x_r = -sign(a_{rs}) \sum_{j=0, j\neq r}^{n} a_{js} x_j. \tag{46}$$

This expression for x_r is then substituted into the THEN-conditions of the other edits as far as this is permitted by the IF-conditions. The IF-conditions of these other edits are changed by the substitution process. In particular, owing to this substitution process an edit given by (43) involving x_r in its THEN-condition gives rise to (at most) two edits

given by (45) and

$$\text{IF } v_i \in F_i^t \cap F_i^s \text{ (for } i = 1, \ldots, m)$$
$$\text{THEN } (x_1, \ldots, x_n) \in \{\mathbf{x} | \sum_{j=0, j\neq r}^{n} (a_{jt} - sign(a_{rs}) a_{rt} a_{js}) x_j \geq 0\}, \quad (47)$$

In (43), (45), and (47) the inequality sign may be replaced by an equality sign. Edits of type (47) for which $F_i^t \cap F_i^s = \emptyset$ (for some i=1,...,m), and edits of type (45) for which $F_i^t - F_i^s = \emptyset$ (for some i=1,...,m) may be discarded. Edits given by (43) not involving x_r are not modified.

The new system of edits is equivalent to the original system of edits, in the sense that a solution to the original system of edits corresponds to a solution to the new system, and vice versa. Namely, for the categorical values for which we can use equation (24) or (46) to eliminate variable x_r, we do this (see (44) and (47)). For the categorical values for which we cannot use equation (24) or (46) to eliminate x_r, we simply leave x_r untouched (see (45)). Note that the IF-conditions of an edit of type (44) or (47) where x_r has been eliminated and an edit still involving x_r have an empty overlap. An edit of type (44) or (47) and an edit still involving x_r will hence never be combined when eliminating integer variables from inequality edits (see Subsection 7.2 for the elimination of integer variables from inequality edits).

We continue "eliminating" balance edits until for each possible combination of categorical values the associated set of numerical THEN-conditions is either a system of type (19) or the empty set. The latter possibility occurs if a combination of categorical values is not allowed by the edits. Note that the balance edits will be eliminated after finitely many steps. Namely, for each possible combination of categorical values we in fact implicitly apply the elimination process of Subsection 6.1, which terminates after a finite number of steps.

After the termination of the above elimination process, we delete all balance edits. We are then left with a set of edits with linear inequalities involving only integer variables as THEN-conditions. Because auxiliary variables may have been introduced to eliminate the balance edits, the total number of integer variables in this system of edits may be larger than the original number of integer variables. How we deal with a set of inequality edits involving only integer variables is explained in the next subsection.

7.2 Error localisation: eliminating integer variables from inequality edits

In this subsection we assume that each THEN-condition is either a linear inequality involving only integer variables or the empty set. When an integer variable is eliminated from a set of inequality edits, a dark shadow and possibly several splinters are generated. Below we describe how this dark shadow and these splinters are defined. We start by selecting an integer variable that we want to eliminate, say x_r. The current edits involving x_r are combined into implicit edits not involving x_r. We consider all edits

involving x_r pair-wise. Such a pair of edits is given by

$$\text{IF } v_i \in F_i^s \quad (\text{for } i = 1, \ldots, m) \text{ THEN } (x_1, \ldots, x_n) \in \{\mathbf{x} \mid \sum_{j=0}^{n} a_{js}\, x_j \geq 0\} \tag{48}$$

and

$$\text{IF } v_i \in F_i^t \text{ (for } i = 1, \ldots, m) \text{ THEN } (x_1, \ldots, x_n) \in \{\mathbf{x} | \sum_{j=0}^{n} a_{jt} x_j \geq 0\}, \tag{49}$$

where all involved numerical variables are integer-valued. We assume that the a_{js}, respectively the a_{jt}, (j=0,...,n) are normalised.

The real shadow obtained by eliminating x_r from the pair of edits (48) and (49) is defined only if $a_{rs} \times a_{rt} < 0$. Its THEN-condition is then given by (14), and its IF-condition by $v_i \in F_i^t \cap F_i^s$ (for i=1,...,m) . The dark shadow is also only defined if $a_{rs} \times a_{rt} < 0$. In that case one coefficient is larger than zero, say $a_{rs} > 0$, and the other coefficient is less than zero, $a_{rt} < 0$. The dark shadow obtained by eliminating x_r from the pair of edits (48) and (49) is then defined by

$$\begin{aligned} &\text{IF } v_i \in F_i^s \cap F_i^t \text{ (for } i = 1, \ldots, m) \\ &\text{THEN} \mathbf{x} \in \{\mathbf{x} | \sum_{j=0}^{n} (a_{rs} a_{jt} - a_{rt} a_{js}) x_j \geq (a_{rs} - 1)(-a_{rt} - 1)\}. \end{aligned} \tag{50}$$

If $F_i^t \cap F_i^s$ is empty for some i=1, ..., m, edit (50) is deleted. As for the real shadow, the IF-condition of the dark shadow (50) is given by the intersections $F_i^s \cap F_i^t$ (i=1,...,m), because two numerical THEN-conditions can only be combined into an implicit numerical THEN-condition for the overlapping parts of their corresponding categorical IF-conditions. Note that for this overlapping part the THEN-condition of the dark shadow is given by (40). The dark shadow obtained by eliminating x_r from a set of inequality edits is by definition given by the edits not involving x_r plus the dark shadows (50), assuming they exist, for all pairs of edits (48) and (49).

Defining the splinters obtained by eliminating x_r from a set of inequality edits is more complicated than in Section 6. The reason is that here we want to define splinters for different combinations of categorical values simultaneously, whereas Section 6 considers the case without any categorical variables. We describe one possibility to define splinters; for an alternative possibility we refer to De Waal (2003a). We write the inequality edits involving x_r as

$$\text{IF } v_i \in F_i^k \quad (\text{for } i = 1, \ldots, m) \text{ THEN } a_{rk} x_r \geq - \sum_{j=0, j \neq r}^{n} a_{jk} x_j. \tag{51}$$

For negative coefficients a_{rk}, the THEN-condition of (51) provides an upper bound on x_r. For positive coefficients a_{rk}, the THEN-condition of (51) provides a lower bound on x_r. We start by determining the smallest negative coefficient a_{rq} of x_r for all edits (51), *i.e.* a_{rq} is the coefficient of x_r in all upper bounds on x_r with the largest absolute value. For each lower bound on x_r, we then test whether an integer solution exists to the original edits combined with

$$\text{IF } v_i \in F_i^k \quad (\text{for } i = 1, \ldots, m) \text{ THEN } a_{rk} x_r = -\sum_{j \neq r} a_{jk} x_j + p \tag{52}$$

for each integer p satisfying $(-a_{rq}a_{rk} + a_{rq} - a_{rk}) \Big/ (-a_{rq}) \geq p \geq 0$. For each possible combination of categorical values, all splinters required according to the Omega test described in Subsection 6.2 are taken into consideration. For some combinations of categorical values, more splinters than necessary are taken into consideration. These superfluous splinters increase the computing time, but do no harm otherwise. We have the following theorem.

Theorem 5 *The original set of edits with linear inequalities involving only integer variables as THEN-conditions has a solution if and only if the dark shadow or a splinter resulting from the elimination of variable x_r has a solution.*

Theorem 5 follows immediately by noting that for arbitrary, fixed categorical values, it reduces to Theorem 3.

We now eliminate all integer-valued variables from the original set of inequality edits. During this process we may have to consider several different sets of edits and corresponding variables owing to the splinters that may arise. We consider each such set of edits (and corresponding variables) separately. If a set of edits contains a balance edit, which happens if this set of edits is a splinter, we apply the technique of Subsection 7.1 to eliminate that equality from this set. For a set of edits involving only inequality edits, we select a variable that has not yet been eliminated and proceed to eliminate this variable using the technique of this subsection. We continue until all integer variables in all sets of edits have been eliminated, and we are left with one or more sets of edits involving only categorical variables. The theory of Subsection 7.1 and a repeated application of Theorem 5 yields the following theorem.

Theorem 6 *A set of edits with THEN-conditions involving only integer variables has a solution if and only if any of the sets of edits involving only categorical variables arising after the elimination of all integer-valued variables has a solution.*

To check the existence of a solution to a set of edits involving only categorical variables one can use the methodology sketched in Section 5.

7.3 Error localisation: algorithm for categorical, continuous and integer data

After the preparations in the previous subsections, we are now able to state our algorithm for solving the error localisation problem for a mix of categorical, continuous and integer data. We denote the error localisation problem in categorical, continuous and integer data under consideration by P_I. To solve P_I we first apply the branch-and-bound algorithm presented in Section 5 without taking into account that some of the variables are integer-valued, *i.e.* we first treat the integer variables as being continuous. We denote the problem where integer variables are treated as continuous ones by P_C. P_C is the continuous error localisation problem.

Let c_{obj} denote the value of the objective function (13) for the best currently found solution to P_I, and S the set of currently best solutions to P_I. We initialise c_{obj} to ∞, and S to $\emptyset$. A solution to P_C not involving any integer variables is automatically also a solution to P_I. So, whenever we find a solution to P_C not involving any integer variables for which (13) is less than c_{obj}, we update c_{obj} with that value of (13) and set S equal to the current solution to P_C. Also, whenever we find a solution to P_C not involving any integer variables for which (13) is equal to c_{obj}, we add the current solution to P_C to S.

Whenever we find a solution to P_C involving integer variables for which (13) is at most equal to c_{obj}, we consider P_I. We test whether the variables involved in the current solution to P_C also constitute a solution to P_I. The basic idea of this test is illustrated in Figure 3 below.

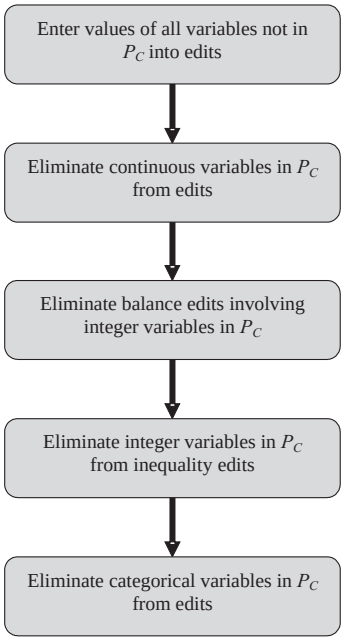

Figure 3: *The basic idea of the integrality test.*

For the integrality test we first fill in the values of the variables not involved in the current solution to P_C into the edits. Subsequently, we eliminate the continuous variables involved in the solution to P_C. This yields a system of edits (10) in which only the integer-valued and categorical variables involved in the solution to P_C occur.

Next, we eliminate all balance edits with integer-valued variables involved in the current solution to P_C in the manner described in Subsection 7.1. Subsequently, we eliminate all integer-valued variables involved in the current solution to P_C from all inequality edits in the manner described in Subsection 7.2. During this latter elimination process the original problem may be split into several subproblems owing to the splinters that may arise. Finally, we eliminate all categorical variables from each of these subproblems. For each subproblem we end up with a set of relations not involving any unknowns. Such a set of relations may be empty. If a set of relations we obtain in this way does not contain a self-contradicting relation, which is for instance (by definition) the case if the set of relations is empty, we have found a solution to P_I. In that case, if the value of (13) for the current solution to P_I is less than c_{obj} we update c_{obj} accordingly and set S equal to the current solution of P_I, else we add the current solution to P_I to S. If all sets of relations involving no unknowns contain a self-contradicting relation, none of the subproblems leads to a solution to P_I and the solution to P_C under consideration is not a solution to P_I. In that case c_{obj} is not updated, and we continue with finding solutions to P_C.

Note that in the above approach, the relatively time-consuming integrality test is only invoked once a solution to P_C with an objective value of c_{obj} or less involving integer-valued variables has been found, so generally only rather infrequently. We have the following theorem.

Theorem 7 *The above procedure finds all optimal solutions to P_I.*

Proof. We start by noting that Theorem 1, Subsection 7.1 and Theorem 6 show that if and only if any of the final sets of relations involving no unknowns obtained by eliminating all variables involved in a solution to P_C does not contain a self-contradicting relation, the original set of edits can be satisfied by modifying the values of the variables involved in this solution. Now, the branch-and-bound algorithm for categorical and continuous data can be used to find all solutions to P_C with an objective value (13) of c_{obj} or less, for any given value of c_{obj}. For each solution to P_C with a value for (13) equal to or less than c_{obj}, we test whether it is also a solution to P_I. The result of this test is conclusive. We update c_{obj} whenever we have found a better solution to P_I than the best one found so far. In other words, all potentially optimal solutions to P_I are considered by the procedure, and all optimal solutions to P_I are indeed identified as such. □

We illustrate the algorithm by means of a simple example involving only two integer-valued variables.

Example 2 *We consider a case with only two variables x_1 and x_2, and three edits given by*

$$-2x_2 + 5 \geq 0,$$

$$5x_1 - x_2 \geq 0, \tag{53}$$

$$-3x_1 + 2x_2 \geq 0.$$

Both variables are integer-valued, and their reliability weights equal one. The original, incorrect record is given by $x_1 = 1$, and $x_2 = 1$. We initialise c_{obj} to ∞, and S to $\emptyset$. We start by solving P_C. We select a variable, say x_1, and construct two branches: in the first branch we eliminate x_1 from the set of current edits, in the second branch we fix x_1 to its original value. If we eliminate x_1 from the set of current edits, we obtain (53) and $x_2 \geq 0$ as our new set of current edits. This new set of current edits is satisfied by the original value of x_2. Hence, we have found a solution to P_C, namely: change x_1. We test whether this is also a solution to P_I. To this end, we start by filling in the original value of x_2 into the original set of edits. We obtain the following set of edits involving only x_1.

$$5x_1 - 1 \geq 0, \tag{54}$$

$$-3x_1 + 2 \geq 0. \tag{55}$$

The dark shadow obtained by eliminating x_2 from (54) and (55) (see (50)) is given by

$$7 \geq 8,$$

which is clearly a self-contradicting relation. We therefore have to consider the splinters. In this simple case there are three splinters. For the first splinter we have to add the constraint

$$5x_1 = 1$$

to (54) and (55) (see (52)), for the second one the constraint

$$5x_1 = 2,$$

and for the third one the constraint

$$5x_1 = 3.$$

It is clear that none of these three splinters has an integer solution for x_1. This would also follow if we were to apply the proposed algorithm. We conclude that although changing x_1 is a solution to P_C, it is not a solution to P_I.

After this intermezzo during which we tested whether changing the value of only x_1 is a solution to $P_I$we continue with finding solutions to P_C. We now consider the branch where x_1 is fixed to its original value. The corresponding set of current edits is given by (53),

$$-x_2 + 5 \geq 0, \tag{56}$$

$$2x_2 - 3 \geq 0. \tag{57}$$

By eliminating x_2 we see that changing the value of only x_2 is a solution to P_C. We check whether this is also a solution to P_I. We fill in the original value of x_1 into the original set of edits. We obtain the system (53), (56) and (57). The dark shadow of (53) and (57) obtained by eliminating x_2 (see (50)) is given by

$$4 \geq 1,$$

and the dark shadow of (56) and (57) obtained by eliminating x_2 by

$$7 \geq 0.$$

The above relations are not self-contradicting, so we can conclude that changing the value of x_2 is a solution to P_I. We can even conclude that this is the only optimal solution to P_I. A feasible value for x_2 is 2.

The method described in this section may appear to be very slow in many cases. Indeed, it is not difficult to design a set of edits for which the method is extremely slow. However, we argue that in practice the situation is not so bad. First, like we already mentioned, the time-consuming algorithm to check potential solutions to P_I is only invoked once a new solution to P_C with an objective value less than or equal to the current value of c_{obj} involving integer variables has been found. In practice, the number of times that such a solution to P_C is found is in most cases rather limited.

Second, whenever we find a solution to P_C with an objective value less than or equal to the current value of c_{obj} we only have to test whether the variables involved in this particular solution also form a solution to P_I. Moreover, often one is only interested in solutions to the error localisation problem with a few variables, say 10 or less. We already argued in Section 5 that if a record requires more than, say, 10 values to be changed, it should not be edited automatically in our opinion as the statistical quality of the automatically edited record would be too low. This implies that the relatively time-consuming test described in this section involves only a few variables.

Third, the integrality test becomes only really time-consuming when many splinters

have to be considered. However, in most edits, either explicit or implicit ones, encountered in practice the coefficients of the integer variables equal –1 or +1. This is especially true for balance edits. For an integer variable with coefficient –1 or +1 the elimination from inequality edits will be exact, *i.e.* the dark shadow and the real shadow coincide and no splinters have to be generated. For balance edits involving integer variables with coefficients -1 or +1 no auxiliary variables have to be introduced in order to eliminate these edits. For such a balance edit the elimination can be performed very fast.

Finally, we can also resort to a heuristic approach based on the exact algorithm. In the next section such a heuristic procedure is described.

8 A heuristic procedure

At Statistics Netherlands we originally aimed to develop a software package for a mix of categorical and continuous data only. In order to achieve this aim a number of algorithms were considered. For an overview of the algorithms considered we refer to De Waal (2003a). Most of those algorithms have been implemented in prototype software, and have subsequently been evaluated. For an assessment of several algorithms on continuous data we refer to De Waal and Coutinho (2005). As a consequence of our work the algorithm described in Section 5 has been implemented in SLICE, our general software framework for automatic editing and imputation (*cf.* De Waal, 2001). Later the wish to extend the implemented algorithm to include integer-valued data arose. The algorithm described in Section 7 was developed to fulfil that wish. However, once this algorithm was developed we considered it to be too complex to implement and maintain in production software. We therefore decided not to implement the exact algorithm of Section 7, but instead to develop a simpler heuristic procedure based on the exact algorithm. That heuristic procedure, which is described below, has been implemented in version 1.5 of SLICE.

Only the integrality test for the integer-valued variables involved in a solution to P_C, *i.e.* a potential solution to P_I, differ for the exact algorithm and the heuristic procedure. In our heuristic procedure, we do not examine splinters; neither do we introduce auxiliary variables in order to eliminate balance edits.

Whenever we have to eliminate an integer-valued variable x_r from a pair of edits s and t in the heuristic checking procedure, we distinguish between two cases. If either edit s or edit t (or both) is a balance edit involving x_r, we examine whether the coefficient of x_r in the corresponding normalised THEN-condition (if both edits are balance edits, we examine both normalised THEN-conditions) equals +1 or -1. If this is not the case, we make the conservative assumption that no feasible integer value for x_r exists, and we reject the potential solution to the P_I. If both edits s and t are inequality edits, we eliminate x_r from these edits by determining the dark shadow (see (50)).

If several integer variables are involved in the solution to P_C under consideration, we repeatedly apply the above procedure until all these variables have been eliminated. If the resulting set of edits involving only categorical variables has a solution, the solution to P_C is also a solution to P_I (see Theorem 6). On the other hand, if the resulting set of edits does not have a solution we make the conservative assumption that the current solution to P_C is not a solution to P_I. This assumption is conservative as we do not check the splinters.

The above heuristic procedure is considerably easier to implement and maintain than the exact algorithm of Section 7. The price we have to pay for using the heuristic procedure instead of the exact algorithm of Section 7 is that we sometimes conclude that an integer solution does not exist, whereas in fact it does.

9 Computational results

In this section we provide computational results for the heuristic procedure described in Section 8. The exact algorithm described in Section 7 has not been implemented in a computer program, and has therefore not been evaluated. The experiments have been performed on a 1500 MHz PC with 256 MB of RAM. This PC was connected to a local area network.

The heuristic procedure has been tested on five realistic data sets. We have used realistic data sets rather than randomly generated synthetic data for our evaluation study, because we feel that the properties of realistic data are completely different than those of randomly generated data. Considering that a production version of SLICE for categorical and continuous data already existed, we decided to implement the heuristic procedure described in Section 8 directly in SLICE (version 1.5), without implementing it in prototype software first.

Our experiments have therefore been carried out by means of SLICE 1.5. This software package has been designed for use in the day-to-day routine at our statistical office. It has been optimised for robustness against mis-use and for ease of maintainability. It uses well-tested components that facilitate debugging. Moreover, the software stores different kinds of metadata, such as which fields are identified as being erroneous. SLICE 1.5 has not been optimised for speed. The speed of the software can definitely be improved upon. Compared to the prototype software for categorical and continuous data on which this production software is based, the production software is about 16 times or more slower (see De Waal and Quere, 2003, where similar data and edits were used as in the present article). The prototype software, however, could handle only a mix of categorical and continuous data, not integer-valued data.

SLICE 1.5 allows the user to specify several parameters, such as a maximum for the number of errors in a record, a maximum for the number of missing values in a record, the maximum computing time per record, the maximum number of (explicit

and implicit) edits in a node of the binary search tree, and a maximum for the number of determined solutions. In our evaluation experiments we did not set a limit for the number of missing values in a record. We have set the maximum number of (explicit and implicit) edits in a node to 3,000, and the maximum computing time per record to 60 seconds. In our experiments we have varied the maximum number of errors and the maximum for the number of determined solutions.

If a record cannot be made to satisfy all edits by changing at most the specified maximum number of errors, it is discarded by SLICE 1.5. A record is also discarded by SLICE 1.5 if it contains more missing values than the specified maximum. Whenever SLICE 1.5 has found N_{sol} solutions with the lowest value c_{low} for the objective function (13) found so far, where N_{sol} is the specified maximum number of determined solutions, it from then on searches only for solutions to the error localisation problem for which the value of the objective function (13) is strictly less than c_{low}. After SLICE 1.5 has solved the error localisation problem for a record, it returns at most N_{sol} solutions with the lowest value for the objective function (13). Owing to the use of the heuristic procedure of Section 8, these determined solutions may be suboptimal. In case the maximum number of edits in a node exceeds 3,000 or the maximum computing time per record exceeds 60 seconds, SLICE 1.5 returns the best solutions (if any) it has determined so far. So, even if the maximum number of edits in a node or the maximum computing time per record is exceeded, the heuristic procedure implemented in SLICE 1.5 may return a solution. For some records the heuristic procedure of SLICE 1.5 could not find a solution at all.

For Statistics Netherlands improving the efficiency of the data editing process for economic, and hence mainly numerical, data is much more important than for social, and hence mainly categorical, data. Therefore, the heuristic procedure has only been evaluated for purely numerical test data. In fact, all variables in the five data sets were integer-valued ones. The five evaluation data sets come from a wide range of business surveys. In Table 1 we give a brief description of each data set.

Table 1: *Description of the five evaluation data sets*

Name	Description
Data set A	structural business survey on enterprises in the photographic sector
Data set B	structural business survey on enterprises in the building and construction industry
Data set C	structural business survey on the retail sector
Data set D	survey on environmental expenditures
Data set E	Annual Business Inquiry

Data set E, the so-called Annual Business Inquiry data set, is one of the evaluation data sets from the EUREDIT project. The EUREDIT project (see http://www.cs.york.ac.uk/euredit) was a large international research and development project on statistical data editing and imputation involving 12 institutes, including

Statistics Netherlands, from seven different countries. The project lasted from March 2000 till March 2003. Important aims were the evaluation of current "in-use" methods for data editing and imputation, and the development and evaluation of a selected range of new or recent techniques for data editing and imputation. For more information on the methods examined in the EUREDIT project we refer to Chambers (2004). Owing to confidentiality reasons the branch of industry to which the businesses in data set E belong has not been made public.

To the best of our knowledge the five evaluation data sets are representative for many other data sets from business surveys. A good performance on the five evaluation data sets hence suggests that the performance on many other business survey data sets arising in practice will also be acceptable.

In Table 2 below we give a summary of the characteristics of the five evaluation data sets. In this table the number of integer-valued variables, the number of non-negativity constraints (*i.e.* constraints expressing that a variable should have a non-negative value), the number of inequality edits (excluding the non-negativity constraints), the number of balance edits, the total number of records, the number of inconsistent records (*i.e.* records failing edits or containing missing values), the total number of missing values, and the average number of errors per inconsistent record are listed.

Table 2: *Characteristics of the five evaluation data sets*

	Data set A	Data set B	Data set C	Data set D	Data set E
Number of integer variables	76	53	51	54	26
Number of non-negativity constraints	70	36	49	54	22
Number of inequality edits[a]	2	16	7	0	15
Number of balance edits	18	20	8	21	3
Total number of records	274	1,478	4,217	1,039	1,425
Number of inconsistent records	157	1,402	2,152	378	1,141
Total number of missing values	0	0	0	2,230	195
Average number of errors per inconsistent record	2.7	2.7	1.6	6.2	2.6

[a] Excluding non-negativity constraints

In all balance edits corresponding to the five evaluation data sets, the coefficients of the involved variables equal -1 or +1. Also, in all inequality edits corresponding to data sets A and C, the coefficients of the involved variables equal -1 or +1. In the equality edits corresponding to data sets B and E, however, many coefficients of the involved variables are not equal to -1 or +1.

We have compared the solutions determined by the heuristic procedure implemented in SLICE 1.5 to the optimal solutions. For purely numerical data, the edits (10) reduce to linear constraints, and the error localisation problem can easily be formulated as an integer programming problem (see, *e.g.*, Schaffer, 1987, and Riera-Ledesma and Salazar-González, 20033). We have therefore used a solver for integer programming

problems to determine the optimal solutions. For our evaluation study we have used CPLEX (*cf*. ILOG CPLEX 7.5 Reference Manual, 2001). Note that although the error localisation problem for numerical (either continuous or integer-valued) data can quite easily be solved by a solver for integer programming problems, the error localisation problem for a mix of numerical (either continuous or integer-valued) and categorical data quickly becomes very hard to solve for such a solver.

In Table 3 we give the number of records for which the heuristic procedure of SLICE 1.5, with the maximum number of errors set to 10, found an optimal solution, the number of records for which it could not find a (possibly suboptimal) solution at all, and the number of records for which it did find solution but exceeded the maximum computing time per record. In our evaluation study the maximum number of edits in a node was never exceeded. Note that records for which the heuristic procedure exceeded the maximum computing time may still be solved to optimality by this procedure.

Table 3: *Number of records that were optimally solved, could not be solved, and for which the maximum computing time per record was exceeded*

	Data set A	Data set B	Data set C	Data set D	Data set E
Number of optimally solved records	120	1,347	2,150	378	1039
Number of unsolved records	4	30	2	0	2
Number of records for which the maximum computing time (60 seconds) was exceeded	3	14	11	0	0

As described in Sections 7 and 8, the heuristic procedure of SLICE 1.5 consists of two parts: a branch-and-bound algorithm where all numerical variables are treated as being continuous ones and an integrality test. In order to assess the slow-down of the algorithm owing to the integrality test, we compare the computing times of the heuristic procedure to the computing times if all variables were continuous ones rather than integer-valued ones. The computing times if all variables were continuous ones are, for various maximum numbers of errors and maximum numbers of determined solutions, given in Table 4 below.

Table 4: *Computing times of the algorithm if the variables were continuous ones (in seconds)*

Parameters	Data set A	Data set B	Data set C	Data set D	Data set E
6 errors, 10 solutions	807	1,725	5,158	228	702
8 errors, 10 solutions	1,023	3,163	5,173	631	704
10 errors, 10 solutions	1,131	6,074	5,187	1,384	706
10 errors, 1 solution	1,088	5,626	5,065	1,348	651

The computational results of the heuristic procedure for various maximum numbers of errors and maximum numbers of determined solutions are given in Table 5 below. In this table we also give the increase in computing time (in per cents of the computing time for the corresponding case for continuous data) owing to the integrality test.

Table 5: *Computing times of the heuristic procedure in seconds (between brackets the increase in computing time owing to the integrality test in per cents of the computing time for continuous data)*

Parameters	Data set A	Data set B	Data set C	Data set D	Data set E
6 errors, 10 solutions	949 (18%)	1,877 (9%)	5,210 (1%)	238 (4%)	783 (12%)
8 errors, 10 solutions	1,056 (3%)	3,505 (11%)	5,223 (1%)	644 (2%)	962 (37%)
10 errors, 10 solutions	1,152 (2%)	6,959 (15%)	5,307 (2%)	1,397 (1%)	1,129 (60%)
10 errors, 1 solution	1,089 (0%)	6,652 (18%)	5,184 (2%)	1,348 (0%)	1,073 (65%)

For data sets A to D, the increase in computing time owing to the integrality test is rather small, namely between 0% and 18%. For data set E, however, the increase in computing time owing to the integrality test is quite large (up to 65%).

The effect of increasing the maximum number of errors on the relative computing time of the integrality test depends on the data set under consideration. For data sets A and D, the relative increase in computing time owing to the integrality test becomes less with increasing maximum numbers of errors. For data set B the relative increase in computing time owing to the integrality test gradually becomes more with increasing maximum numbers of errors. For data set C, this relative increase in computing time is more or less stable for different maximum numbers of errors. Finally, for data set E the relative increase in computing time grows rapidly with increasing maximum numbers of errors.

Determining several solutions instead of one leads to a limited increase in computing time. In Tables 4 and 5 we have given the computing times of SLICE 1.5, with the maximum number of errors set to 10, for the maximum number of determined solutions set to 10 and for the maximum number of determined solutions set to 1. In Table 5 the largest relative increase in computing time when determining at most 10 solutions instead of only one is for data set A. The computing time increases from 1,089 seconds to 1,152 seconds, an increase of approximately 6%. The largest relative increase in computing time owing to determining at most 10 solutions instead of only one is slightly higher in Table 4, namely approximately 8% for data sets B and D.

In Table 6 below we give the total number of erroneous fields according to the heuristic procedure of SLICE 1.5 and the exact algorithm implemented by means of CPLEX for the records that could be solved, possibly in a suboptimal manner, by means of the heuristic procedure.

The number of fields that are unnecessarily identified as being erroneous by the heuristic procedure was very small in our evaluation study. In other words, for the data

Table 6: *Total number of erroneous fields in solved records according to the heuristic procedure and the exact algorithm*

	Data set A	Data set B	Data set C	Data set D	Data set E
Exact algorithm for integer data (CPLEX)	378	3,424	3,526	2,362	2,919
Heuristic procedure (SLICE 1.5)	381	3,482	3,526	2,362	2,919

sets used in our evaluation study, the quality of the solutions determined by the heuristic procedure in terms of the total number of fields identified as erroneous is very good. In the worst case, data set B, the surplus of fields identified as being erroneous by the heuristic procedure in comparison to the number of fields identified as being erroneous by the exact algorithm implemented by means of CPLEX is less than 2% of the latter number of fields.

Finally, we examine the quality of the heuristic procedure in terms of the number of optimal solutions determined. We set both the maximum number of errors and the maximum number of solutions per record to 10. The reason for selecting the latter number is that for records with more than 10 optimal solutions to the error localisation problem, it is very hard to later select the correct solution, *i.e.* correctly identify the erroneous fields, anyway. For the records for which the heuristic procedure succeeded in determining an optimal solution, we compare the number of optimal solutions determined by the heuristic procedure to the number of optimal solutions determined by the exact algorithm implemented by means of CPLEX. The results are given in Table 7.

Table 7: *Number of optimal solutions of the heuristic procedure and the exact algorithm (between brackets the number of optimal solutions determined by the heuristic procedure in per cents of the number of optimal solutions determined by the exact algorithm)*

	Data set A	Data set B	Data set C	Data set D	Data set E
Exact algorithm for integer data (CPLEX)	701	6,609	11,404	474	6,207
Heuristic procedure (SLICE 1.5)	701(100%)	6,477(98%)	11,404 (100%)	474(100%)	4,828(78%)

For data sets A, C and D the heuristic procedure determined the same number of optimal solutions as the exact algorithm. Data sets B and E, the data sets for which the coefficients of the variables involved in the corresponding inequality edits often are unequal to -1 or +1, the number of optimal solutions determined by the heuristic procedure is less than the number of optimal solutions determined by the exact algorithm implemented by means of CPLEX. In particular, this is the case for data set E, where the number of optimal solutions determined by the heuristic procedure is only 78% of the number of optimal solutions determined by the exact algorithm. Data set E is

the only data set for which the number of inequality edits with coefficients unequal to -1 or +1 for the involved variables clearly outnumbers the number of balance edits, which probably explains our result. Note that despite the fact that the number of optimal solutions determined by the heuristic procedure for data set E is clearly less than the actual number of optimal solutions, the heuristic procedure does succeed in solving all records to optimality, except for two records for which it could not find a solution at all.

10 Discussion

In this article we have developed an exact algorithm for solving the error localisation problem for a mix of categorical, continuous and integer data. This algorithm is quite complex to implement and maintain in a software system, especially in a software system that is meant to be used routinely in practice. Based on this exact algorithm we have therefore also developed a much simpler heuristic procedure. This heuristic procedure has been implemented in our production software, SLICE 1.5. In this article we have also examined the performance of the heuristic procedure.

The exact algorithm and the heuristic procedure described in this article have a number of theoretical drawbacks. Both the exact algorithm and the heuristic procedure are extensions to an exact algorithm for continuous and categorical data (see Section 5). The computing time of this latter exact algorithm can, theoretically, be exponential in its input parameters, such as the number of variables, the number of edits and the maximum number of errors. For some data sets in our evaluation study, namely data sets B and D, this exponential increase in the computing time owing to an increase of the maximum number of errors is, unfortunately, also observed in practice. For some practical instances of the error localisation problem, this exponential increase in the computing time may be a problem. For such instances, one has to resort to other heuristic approaches, such as setting fields that are likely to be erroneous to “missing” in a pre-processing step (the exact algorithm for continuous and categorical data is generally faster for records with many missing values than for records with many erroneous values), or to an alternative algorithm altogether (see Section 1 and De Waal and Coutinho, 2005, for references to some papers on alternative approaches).

The computing time of the integrality test of the exact algorithm can, theoretically, also be exponential in the number of variables, the number of edits, and the maximum number of errors. In our evaluation study on the heuristic procedure, the increase in computing time owing to the integrality test is limited for most evaluation data sets. However, for data E the increase in computing time owing to the integrality test grows rapidly when increasing the maximum number of errors. Again, for some practical instances of the error localisation problem, this rapid increase in the computing time may be a problem, and one may have to resort to other approaches.

In principle, the number of erroneous fields identified by the heuristic procedure may

be (much) higher than the number of erroneous fields identified by an exact algorithm. In our evaluation study this has, however, not occurred. The number of fields identified as being erroneous by the heuristic procedure is for all evaluation data sets almost equal, and often even precisely equal, to the number of fields identified as being erroneous by an exact algorithm implemented by means of CPLEX.

Another potential drawback of the heuristic procedure is that the number of optimal solutions determined by this procedure can be (much) less than for an exact algorithm. In our evaluation study this has also not occurred. For most evaluation data sets, the number of optimal solutions determined by the heuristic procedure is equal or almost equal to the number of optimal solutions determined by an exact algorithm implemented by means of CPLEX. The only exception is data set E, where the number of optimal solutions determined by the heuristic procedure drops to about 78% of the number of optimal solutions determined by an exact algorithm implemented by means of CPLEX. Whereas the actual average number of optimal solutions is 5.4 (=6,207/1,039, see Tables 3 and 7) per optimally solved record if the maximum number of optimal solutions determined is set to 10, the heuristic procedure determines only 4.2 optimal solutions on the average. Fortunately, for our purposes at Statistics Netherlands this is an acceptable result.

As mentioned before, at Statistics Netherlands we aimed to implement an algorithm for a mix of categorical, continuous and integer data. Given the fact we had already implemented the algorithm for continuous and categorical data described in Section 5 in our production software, our main choice to be made was whether we would implement the exact algorithm described in Section 7 or the heuristic procedure of Section 8 in that production software. Considering the complexity of implementing and maintaining the exact algorithm in production software, we decided to implement the heuristic procedure instead the exact algorithm. Our, admittedly limited, experience with the heuristic procedure so far suggests that we have made a good choice here. For Statistics Netherlands, the benefits of using the heuristic procedure, in particular a considerable simplification in developing and maintaining the software in comparison to the exact algorithm of Section 7, outweigh the disadvantages, possibly worse and less solutions, of using the heuristic procedure instead of the exact algorithm. Despite the earlier mentioned theoretical drawbacks of the heuristic procedure, its computing speed and the quality of its solutions thus far appear to be fully acceptable for application in practice at Statistics Netherlands.

Acknowledgement

The author wishes to thank three anonymous referees for their useful comments on an earlier version of this article.

References

Barcaroli, G., C. Ceccarelli, O. Luzi, A. Manzari, E. Riccini and F. Silvestri (1995). The methodology of editing and imputation of qualitative variables implemented in SCIA. Internal Report, ISTAT, Rome.

Boskovitz, A., R. Goré and M. Hegland (2003). A logical formalisation of the Fellegi-Holt method of data cleaning. Report, Research School of Information Sciences and Engineering, Australian National University, Canberra.

Bruni, R., A. Reale, and R. Torelli (2001). Optimization techniques for edit validation and data imputation. *Proceedings of Statistics Canada Symposium 2001 "Achieving Data Quality in a Statistical Agency: a Methodological Perspective" XVIII-th International Symposium on Methodological Issues.*

Bruni, R. and A. Sassano (2001). Logic and optimization techniques for an error free data collecting. Report, University of Rome "La Sapienza".

Chambers, R. (2004). Methods Investigated in the EUREDIT Project. In: *Methods and Experimental Results from the EUREDIT Project*, J.R.H. Charlton (ed.) http://www.cs.york.ac.uk/euredit/.

Central Statistical Office (2000). Editing and calibration in survey processing. Report SMD-37, Ireland.

Chvátal, V. (1983). *Linear Programming*. W.H. Freeman and Company: New York.

Dantzig, G.B. and B. Curtis Eaves (1973). Fourier-Motzkin elimination and its dual. *Journal of Combinatorial Theory (A) 14*, 288-297.

De Jong, A. (2002). Uni-Edit: standardized processing of structural business statistics in the Netherlands. UN/ECE Work Session on Statistical Data Editing, Helsinki.

De Waal, T. (1996). CherryPi: a computer program for automatic edit and imputation. UN/ECE Work Session on Statistical Data Editing, Voorburg.

De Waal, T. (2001). SLICE: generalised software for statistical data editing. In *Proceedings in Computational Statistics*, J.G. Bethlehem and P.G.M. Van der Heijden (eds.), New York: Physica-Verlag, 277-282.

De Waal, T. (2003*a*). *Processing of Erroneous and Unsafe Data*. Ph.D. Thesis, Erasmus University, Rotterdam

De Waal, T. (2003*b*). Solving the error localization problem by means of vertex generation. *Survey Methodology*, 29, 71-79.

De Waal, T. and W. Coutinho (2005). Automatic editing for business surveys: an assessment for selected algorithms. *International Statistical Review*, forthcoming.

De Waal, T. and R. Quere (2003). A fast and simple algorithm for automatic editing of mixed data. *Journal of Official Statistics*, 19, 383-402.

Duffin, R.J. (1974). On Fourier's analysis of linear inequality systems. *Mathematical Programming Studies*, 1, 71-95.

Fellegi, I.P. and D. Holt (1976). A systematic approach to automatic edit and imputation. *Journal of the American Statistical Association*, 71, 17-35.

Garfinkel, R.S., A.S. Kunnathur and G.E. Liepins (1986). Optimal imputation of erroneous data: categorical data, general edits. *Operations Research*, 34, 744-751.

Garfinkel, R.S., A.S. Kunnathur and G.E. Liepins (1988). Error localization for erroneous data: continuous data, linear constraints. *SIAM Journal on Scientific and Statistical Computing*, 9, 922-931.

Granquist, L. (1990). A review of some macro-editing methods for rationalizing the editing process. *Proceedings of the Statistics Canada Symposium*, 225-234.

Granquist, L. (1995). Improving the traditional editing process. In *Business Survey Methods*, Cox, Binder, Chinnappa, Christianson & Kott (eds.), John Wiley & Sons, Inc., 385-401.

Granquist, L. (1997). The new view on editing. *International Statistical Review*, 65, 381-387.

Granquist, L. and J. Kovar (1997). Editing of survey data: how much is enough?. In *Survey Measurement*

and Process Quality, Lyberg, Biemer, Collins, De Leeuw, Dippo, Schwartz & Trewin (eds.), John Wiley & Sons, Inc., 415-435.

Hedlin, D. (2003). Score functions to reduce business survey editing at the U.K. office for national statistics. *Journal of Official Statistics*, 19, 177-199.

Hoogland, J. (2002). Selective editing by means of plausibility indicators. UN/ECE Work Session on Statistical Data Editing, Helsinki.

Hoogland, J. and E. Van der Pijll (2003). Evaluation of automatic versus manual editing of production statistics 2000 trade & transport. UN/ECE Work Session on Statistical Data Editing, Madrid.

ILOG CPLEX 7.5 Reference Manual (2001). ILOG, France.

Kalton, G. and D. Kasprzyk (1986). The treatment of missing survey data. *Survey Methodology*, 12, 1-16.

Kovar, J. and P. Whitridge (1990). Generalized edit and imputation system: overview and applications. *Revista Brasileira de Estadistica*, 51, 85-100.

Kovar, J. and P. Whitridge (1995). Imputation of business survey data. In *Business Survey Methods*, Cox, Binder, Chinnappa, Christianson & Kott (eds.), New York: John Wiley & Sons, Inc., 403-423.

Liepins, G.E., R.S. Garfinkel and A.S. Kunnathur (1982). Error localization for erroneous data: A survey. *TIMS/Studies in the Management Sciences*, 19, 205-219.

McKeown, P.G. (1984). A mathematical programming approach to editing of continuous survey data. *SIAM Journal on Scientific and Statistical Computing*, 5, 784-797.

Nemhauser, G.L. and L.A. Wolsey (1988). *Integer and Combinatorial Optimisation.* John Wiley & Sons, New York.

Pannekoek, J. and T. De Waal (2005). Automatic editing and imputation for business surveys: the dutch contribution to the EUREDIT project. *Journal of Official Statistics*, forthcoming.

Pugh, W. (1992). The Omega test: a fast and practical integer programming algorithm for data dependence analysis. *Communications of the ACM 35*, 102-114.

Pugh, W. and D. Wonnacott (1994). Experiences with constraint-based array dependence analysis. In *Principles and Practice of Constraint Programming, Second International Workshop*. *Lecture Notes in Computer Science 768*, Berlin: Springer-Verlag.

Ragsdale, C.T. and P.G. McKeown (1996). On solving the continuous data editing problem. *Computers & Operations Research*, 23, 263-273.

Riera-Ledesma, J. and J.J. Salazar-González (2003). New algorithms for the editing and imputation problem. UN/ECE Work Session on Statistical Data Editing, Madrid.

Sande, G. (1978). An algorithm for the fields to impute problems of numerical and coded data. Technical report, Statistics Canada.

Schaffer, J. (1987). Procedure for solving the data-editing problem with both continuous and discrete data types. *Naval Research Logistics*, 34, 879-890.

Schrijver, A. (1986). *Theory of Linear and Integer Programming*. New York: John Wiley & Sons.

Stoop, J.R. (2003). The best piece of CherryPie (in Dutch). Internal report (BPA number: 2098-03-TMO), Voorburg: Statistics Netherlands.

Todaro, T.A. (1999). Overview and evaluation of the AGGIES automated edit and imputation system. UN/ECE Work Session on Statistical Data Editing, Rome.

Williams, H.P. (1976). Fourier-Motzkin elimination extension to integer programming. *Journal of Combinatorial Theory (A)*, 21, 118-123.

Williams, H.P. (1986). Fourier's method of linear programming and its dual. *American Mathematical Monthly*, 93, 681-695.

Winkler, W.E. (1996). State of statistical data editing and current research problems. UN/ECE Work Session on Statistical Data Editing, Rome.

Winkler, W.E. (1998). Set-covering and editing discrete data. Statistical Research Division Report 98/01, US Bureau of the Census, Washington, D.C.

Winkler, W.E. and L.A. Draper (1997). The SPEER edit system. *Statistical Data Editing (Volume 2); Methods and Techniques*, United Nations, Geneva.

Winkler, W.E. and T.F. Petkunas (1997). The DISCRETE edit system. *Statistical Data Editing (Volume 2)*; *Methods and Techniques*. United Nations, Geneva.

SORT 29 (1) January-June 2005, 101-118

Estimation of the spectral density of a homogeneous random stable discrete time field

N.N. Demesh*, S.L. Chekhmenok

Belarus State University

Abstract

In earlier papers, 2π-periodic spectral data windows have been used in spectral estimation of discrete-time random fields having finite second-order moments. In this paper, we show that 2π-periodic spectral windows can also be used to construct estimates of the spectral density of a homoge-neous symmetric α-stable discrete-time random field. These fields do not have second-order moments if $0 < \alpha < 2$. We construct an estimate of the spectrum, calculate the asymptotic mean and variance, and prove weak consistency of our estimate.

MSC: 60G60

Keywords: homogeneous stable fields, spectral density estimate

1 Introduction

The use of methods of spectral analysis of stochastic processes and random fields in many areas of scientific research has been considerably extended. Particular attention has been focused on methods of spectral analysis of continuous-time and discrete-time stationary processes and homogeneous fields.

As a consequence of this research the topic of spectral analysis of processes and fields having finite second-order moments is better represented in the literature. The study of random processes and fields having no second-order moments is therefore of particular importance. In α-stable stochastic processes and random fields, $0 < \alpha \leq 2$,

* *Address for correspondence:* N.N. Demesh. Department of Information and Program-Mathematical Software of Automated Productions. Faculty of Applied Mathematics and Computer Sciences. Belarus State University. F. Skoriny ave.,4, Minsk, 220000, Belarus, *e-mail*: demesh@bsu.by.

Received: October 2003
Accepted: February 2005

only Gaussian processes and fields ($\alpha = 2$) have finite variances. In the case where $0 < \alpha < 2$, there exist finite moments of order $p \in (0, \alpha)$ only; see, e.g., section 2.1 in Zolotarev (1986). Therefore traditional methods cannot be applied for solving practical problems and the development of a special theory is required.

The research in multidimensional stable distributions required to develop a theory of random processes and fields (Paulauskas (1976), Press (1972)). A further development of this theory was influenced by the publication of results which are related to the representation of characteristic functions of multidimensional symmetric stable distributions and to the spectral representation of symmetric stable processes and fields. The notion of spectral density of a stable process introduced by Masry and Cambanis (1984) and the construction of estimates of the spectral density of a continuous-time stable random process are especially important for further research.

Different estimates of the spectral density of a homogeneous symmetric α-stable discrete-time random field have been introduced and studied in many articles: Hosoya (1978), Masry and Cambanis (1984), Sabre (1995), Sabre (2000), to mention a few. However, either one-dimensional processes (Masry and Cambanis (1984)) or two-dimensional fields (Sabre (2000)) have been studied in these articles. The difficulty of studying stable fields consists in the fact that these fields do not have finite second-order moment either and, if $0 < \alpha \leq 1$, they do not have finite first-order moments either. In the literature (see, e.g., Sabre (1995)), mainly spectral windows usually used for continuous-time fields are considered. The problem of development of the corresponding technique for the case of discrete-time is less investigated. In this paper, we construct and study an estimate of the spectral density of a homogeneous symmetric α-stable ($0 < \alpha < 2$) discrete-time random field by means of 2π-periodical spectral windows used for fields having finite second-order moments (Leonenko and Ivanov (1989)).

2 Assumptions and notations

Let us give our basic notations to be used in the article. Denote by $\mathbb{N} = \{1, 2, \ldots\}$ the set of natural numbers, $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ the set of integer numbers, $\Pi^n = [-\pi, \pi]^n$, $P^n = [-\tau_1, \tau_1] \times [-\tau_2, \tau_2] \times \cdots \times [-\tau_n, \tau_n]$ an integer lattice of n-dimensional parallelepiped where $\tau_j \in \mathbb{N}$, $j = \overline{1, n}$. $T = (T_1, \ldots, T_n)$ the n-dimensional vector having $T_j = 2\tau_j + 1$, $j = \overline{1, n}$; $N_T = T_1 \times \cdots \times T_n$; $< a, b >$ the inner product of the vectors $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$. The notation $T \longrightarrow \infty$ means that $T_j \longrightarrow \infty$, $j = \overline{1, n}$, we write $\|a\| = < a, a >^{\frac{1}{2}}$ and the notation $x_T \cong y_T$ means that $x_T - y_T \longrightarrow 0$ as $T \longrightarrow \infty$.

Denote by $D(p) = \int_{-\infty}^{\infty} |u|^{-1-p} (1 - cos(u))\, du$, $F(p, \alpha) = \int_{-\infty}^{\infty} |u|^{-1-p} \left(1 - e^{-|u|^{\alpha}}\right) du$, $c_\alpha = \frac{1}{\alpha\pi} \int_0^{\pi} |cos(u)|^{\alpha}\, du$, and

$$k(p, \alpha) = \frac{D(p)}{F(p, \alpha)\, (c_\alpha)^{\frac{p}{\alpha}}}. \tag{1}$$

Let $h_T(t)$, $t = \left(\frac{t_1}{\tau_1}, \frac{t_2}{\tau_2}, \ldots, \frac{t_n}{\tau_n}\right)$, be an n-dimensional data window,

$$H^{(T)}(\lambda) = \sum_{t \in P^n} h_T(t) \exp(-i < t, \lambda >), \tag{2}$$

$$A_T = \left[\int_{\Pi^n} \left| H^{(T)}(\lambda) \right|^{\alpha} d\lambda \right]^{-\frac{1}{\alpha}}. \tag{3}$$

Let us consider numerical sequences $M_{T_j} \in \mathbb{N}$ and $L_{T_j} \in \mathbb{N}$, $j = \overline{1,n}$, where $M_{T_j} \longrightarrow \infty$ as $T_j \to \infty$ but $\frac{M_{T_j}}{T_j} \longrightarrow 0$ as $T_j \to \infty$, $j = \overline{1,n}$; $L_{T_j} \longrightarrow \infty$, $\frac{M_{T_j}}{L_{T_j}} \longrightarrow 0$, $\frac{L_{T_j}}{T_j} \longrightarrow 0$ as $T_j \longrightarrow \infty$ for all $j = \overline{1,n}$. Denote by

$$M_T = \prod_{j=1}^{n} M_{T_j} \tag{4}$$

$$L_T = \prod_{j=1}^{n} L_{T_j} \tag{5}$$

Let $w_T(l) = w\left(\frac{l_1}{M_{T_1}}, \frac{l_2}{M_{T_2}}, \ldots, \frac{l_n}{M_{T_n}}\right)$ be a n-dimensional correlation window, $l = (l_1, l_2, \ldots, l_n)$, where $w(x)$, $x \in \mathbb{R}^n$, satisfies the following conditions:

$$\sup_{x \in \mathbb{R}^n} w(x) = w(0) = 1,$$

$$0 \le w(x) \le 1, x \in \mathbb{R}^n,$$

$$\int_{\mathbb{R}^n} w^2(x) dx < \infty, \tag{6}$$

and $W_T(\nu)$, $\nu \in \Pi^n$, is a nonnegative spectral data window of the form

$$W_T(\nu) = \frac{1}{(2\pi)^n} \sum_{l_1=-M_{T_1}}^{M_{T_1}} \sum_{l_2=-M_{T_2}}^{M_{T_2}} \cdots \sum_{l_n=-M_{T_n}}^{M_{T_n}} w_T(l) \exp(-i < \nu, l >), \nu \in \Pi^n. \tag{7}$$

Examples of spectral data windows can be found in Corollary 2, Appendix 2 (Example 3). See also Section 4.1 in Brillinger (1975) and Section 4.4 in Trush (1999).

3 Main results

The definition of a symmetric α-stable random field is due to Nolan (1988). If X is a α-stable random variable, $0 < \alpha \le 2$, then put

$$\|X\|_\alpha = \left[-\ln\left(E\exp\{iX\}\right)\right]^{\frac{1}{\alpha}}. \tag{8}$$

This is a norm in the space of symmetric α-stable random variables. Of course, $\|X\|_2^2 = \frac{\mathrm{var}(X)}{2}$ in the Gaussian case. Therefore one can think of $\|X\|_p^p$ to be a generalization of the variance. It is known (Nolan (1988)) that for any $0 < p < \alpha$ there exists a $C(p,\alpha)$ such that

$$E|X|^p = C(p,\alpha)\|X\|_\alpha$$

for every symmetric α-stable random variable X.

Let $0 < \alpha \le 2$ and $\mathbb{T} \subset \mathbb{Z}^n$. A complex-valued random field $X(t), t \in \mathbb{T}$, is called α-stable if for any $m \ge 1$, $a_j \in \mathbb{C}$, $t^{(j)} \in \mathbb{T}$, $j = \overline{1,m}$, every linear combination $\sum\limits_{j=1}^{m} a_j X\left(t^{(j)}\right)$ is a symmetric α-stable random variable. Then by (8)

$$E\exp\left\{i\sum_{j=1}^{m} a_j X\left(t^{(j)}\right)\right\} = \exp\left\{-\left\|\sum_{j=1}^{m} a_j X\left(t^{(j)}\right)\right\|\right\}.$$

Therefore $\left\|\sum\limits_{j=1}^{m} a_j X\left(t^{j}\right)\right\|_\alpha$ determines completely the distribution of $\left(X\left(t^{(1)}\right), X\left(t^{(2)}\right), \ldots, X\left(t^{(m)}\right)\right)$. Each α-stable random field $X(t)$, $t \in \mathbb{T}$, has a representation as a stochastic integral: there exists a measurable space (Ω, F, P) and a collection $\{f(t,\cdot) : t \in \mathbb{T}\} \subset L^\alpha(\Omega, F, P)$ such that $X(t) = \int f(t,u)\,\zeta(du)$. Then ζ is the α-stable random measure generated by P. A method for constructing ζ is described in Hardin (1982).

Later we shall consider a special class of symmetric α-stable fields which are called harmonizable random fields. Let $X(t), t \in \mathbb{Z}^n$, be a homogeneous symmetric α-stable $(0 < \alpha < 2)$ discrete-time random field having the spectral representation of the form

$$X(t) = \int_{\Pi^n} \exp(i < t, \lambda >) d\xi(\lambda)$$

where $t = (t_1, t_2, \ldots, t_n)$, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, and where $\xi(\lambda)$ is α-stable random field with independent increments such that $[E|d\xi(\lambda)|^p]^{\frac{\alpha}{p}} = C(p,\alpha)\varphi(\lambda)d\lambda$ for $p \in (0,\alpha)$. The constant $C(p,\alpha)$ depends on p and α only, the function $\varphi(\lambda)$, $\lambda \in \Pi^n$, is a nonnegative even 2π-periodic function in each argument on $\mathbb{R}^n$. This function is called spectral density of $X(t), t \in \mathbb{Z}^n$. If $\alpha = 2$, the field $X(t), t \in \mathbb{Z}^n$, is Gaussian, the function $\varphi(\lambda)$, $\lambda \in \Pi^n$, is the "usual" spectral density and a standard spectral analysis is then applied. If $0 < \alpha < 2$, the function $\varphi(\lambda)$, $\lambda \in \Pi^n$, does not represent spectral density in the usual sense. However, it was shown in Nolan (1988) that for linear prediction and filtering the role of this function is quite similar to that played by the spectral density function in second-order fields.

Let

$$\{X(t_1, t_2, \ldots, t_n), t = (t_1, t_2, \ldots, t_n) \in P^n\} \tag{9}$$

be N_T observations of the field $X(t), t \in \mathbb{Z}^n$ on P^n.

Definition 1 *The statistic*

$$d_T(\lambda) = A_T \sum_{t \in P^n} cos(< t, \lambda >) h_T(t) X(t), \lambda \in \Pi^n, \tag{10}$$

is called the finite modified Fourier transform of observations (9) of the field $X(t), t \in \mathbb{Z}^n$, where $h_T(t), t \in \mathbb{Z}^n$, is an n-dimensional data window and A_T is given by (3).

Definition 2 *The statistic*

$$I_T(\lambda) = k(p, \alpha)|d_T(\lambda)|^p, \lambda \in \Pi^n, \tag{11}$$

is called the modified periodogram of observations (9) of the field $X(t), t \in \mathbb{Z}^n$ where $d_T(\lambda), \lambda \in \Pi^n$, is given by (10) and $k(p, \alpha)$ is defined by (1) .

As an estimate of the function $[\varphi(\lambda)]^{\frac{p}{\alpha}}, \lambda \in \Pi^n$, we consider a smoothed periodogram of the form

$$\widehat{f_T}(\lambda) = \int_{\Pi^n} W_T(\nu) I_T(\lambda + \nu) d\nu, \lambda \in \Pi^n, \tag{12}$$

where $\nu = (\nu_1, \nu_2, \ldots, \nu_n), d\nu = d\nu_1 d\nu_2 \cdots d\nu_n$, and where $I_T(\lambda), \lambda \in \Pi^n$, is given by (11) and the spectral window $W_T(\nu), \nu \in \Pi^n$, is defined by (7). Let us consider the statistic

$$\widehat{\varphi_T}(\lambda) = \left[\widehat{f_T}(\lambda)\right]^{\frac{\alpha}{p}}, \lambda \in \Pi^n, \tag{13}$$

as an estimate of the spectral density $\varphi(\lambda), \lambda \in \Pi^n$.

Definition 3 *The function $\varphi(\lambda), \lambda \in \Pi^n$, is said to satisfy the Hölder inequality at a point $\lambda^{(0)} \in \Pi^n$ with index γ, $0 < \gamma \leq 1$, if for any λ close enough to $\lambda^{(0)}$, the following inequality holds:*

$$\left|\varphi(\lambda) - \varphi\left(\lambda^{(0)}\right)\right| \leq A\left(\lambda^{(0)}\right) \left\|\lambda - \lambda^{(0)}\right\|^{\gamma} \tag{14}$$

where $0 < A\left(\lambda^{(0)}\right) < +\infty$.

Definition 4 *A set of even 2π-periodic in each argument functins $G_T(\lambda), \lambda \in \Pi^n$, is called a kernel on Π^n if it satisfies the following conditions:*

$$\int_{\Pi^n} G_T(\lambda) d\lambda = 1 \text{ for any } T = (t_1, t_2, \ldots, t_n),$$

$$\int\limits_{\Pi^n\setminus\{\|\lambda\|\le\delta\}} G_T(\lambda)\,d\lambda \longrightarrow 0 \text{ for any fixed } \delta\in(0,\pi) \text{ as } T\to\infty.$$

We begin with considering the mathematical expectation and the variance of estimator (12) in general. Then we examine the use of specific data windows $h_T(t)$ and spectral data windows $w_T(l)$ and we prove convergence in probability of statistics (13). The proofs of all results in this section are given in Appendix 1.

Theorem 1 *Assume that the spectral density $\varphi(\lambda), \lambda\in\Pi^n$, of the field $X(t), t\in\mathbb{Z}^n$, is bounded on Π^n, positive and satisfies the Hölder inequality (14) at the point $\lambda^{(0)}\in\Pi^n$. Let $|H_T(\lambda)|^\alpha = \left|A_T H^{(T)}(\lambda)\right|^\alpha$, $\lambda\in\Pi^n$, be a kernel on Π^n where A_T and $H^{(T)}(\lambda)$, $\lambda\in\Pi^n$ are defined by (3) and (2), respectively. Assume also that*

$$\int_{\Pi^n\setminus\{\|\lambda\|\le\delta\}} |H_T(\lambda)|^\alpha\,d\lambda = o\left(\int_{\Pi^n}\|\lambda\|^\gamma |H_T(\lambda)|^\alpha\,d\lambda\right) \text{ for any fixed } \delta\in(0,\pi) \text{ as } T\to\infty, \tag{15}$$

$$\int_{\Pi^n}\|\lambda\|^\gamma |H_T(\lambda)|^\alpha\,d\lambda \longrightarrow 0 \text{ as } T\to\infty, \tag{16}$$

$$\int_{\Pi^n\setminus\{\|\nu\|\le\delta\}} W_T(\nu)\,d\nu = o\left(\int_{\Pi^n}\|\nu\|^\gamma W_T(\nu)\,d\nu\right) \text{ for any fixed } \delta\in(0,\pi) \text{ as } T\to\infty, \tag{17}$$

$$\int_{\Pi^n}\|\nu\|^\gamma W_T(\nu)\,d\nu \longrightarrow 0 \text{ as } T\to\infty. \tag{18}$$

Then

$$\left|E\widehat{f_T}\left(\lambda^{(0)}\right) - \left[\varphi\left(\lambda^{(0)}\right)\right]^{\frac{p}{\alpha}}\right| \le$$

$$\le C_1 A\left(\lambda^{(0)}\right)\left(\int_{\Pi^n}\|\nu\|^\gamma W_T(\nu)\,d\nu\,(1+o(1)) + \int_{\Pi^n}\|\mu\|^\gamma |H_T(\mu)|^\alpha\,d\mu\,(1+o(1))\right) \text{ as } T\to\infty \tag{19}$$

where

$$C_1 = \frac{p}{2\alpha}\max_{\nu\in\Pi^n}\left|\left[\psi_T\left(\lambda^{(0)}+\nu\right)\right]^{\frac{p}{\alpha}-1} + \left[\varphi\left(\lambda^{(0)}\right)\right]^{\frac{p}{\alpha}-1}\right|, \tag{20}$$

$$\psi_T(\lambda) = \int_{\Pi^n}|H_T(\nu)|^\alpha\,\varphi(\lambda+\nu)\,d\nu, \lambda\in\Pi^n. \tag{21}$$

Examples of the data windows satisfying conditions (15) and (16) and also spectral data windows satisfying conditions (17) and (18) can be found in Appendix 2 (Examples 1 and 3, respectively).

Corollary 1 *Assume that the correlation window $w_T(l)$ can be written as*

$$w_T(l) = \prod_{j=1}^{n} w_{T_j}\left(l_j\right), \tag{22}$$

where $w_{T_j}\left(l_j\right) = w_j\left(\frac{l_j}{M_{T_j}}\right)$, $j = \overline{1,n}$, *are even functions satisfying the following conditions:*

$$\sup_{x\in\mathbb{R}} w_j(x) = w_j(0) = 1,$$

$$0 \le w_j(x) \le 1, x \in \mathbb{R},$$

$$\int_{-\infty}^{+\infty} w_j^2(x)\,dx < \infty, j = \overline{1,n}.$$

Suppose that the function

$$H_T(\lambda) = \prod_{j=1}^{n} H_{T_j}\left(\lambda_j\right) \tag{23}$$

where $\left|H_{T_j}\left(\lambda_j\right)\right|^{\alpha} = \left|A_{T_j}H^{(T_j)}\left(\lambda_j\right)\right|^{\alpha}$ *are kernels on* Π, $A_{T_j} = \left[\int_{\Pi}\left|H^{(T_j)}\left(\lambda_j\right)\right|^{\alpha} d\lambda_j\right]^{-\frac{1}{\alpha}}$, $H^{(T_j)}\left(\lambda_j\right) = \sum_{t_j=-\tau_j}^{\tau_j}\exp\left(-it_j\lambda_j\right)h_j\left(\frac{t_j}{\tau_j}\right)$, $h_j\left(\frac{t_j}{\tau_j}\right)$ *are one-dimensional data windows* $j = \overline{1,n}$. *Then*

$$\left|E\widehat{f_T}\left(\lambda^{(0)}\right) - \left[\varphi\left(\lambda^{(0)}\right)\right]^{\frac{p}{\alpha}}\right| \le$$

$$\le C_2\left(\max_{j=\overline{1,n}}\int_{\Pi}\left|\nu_j\right|^{\gamma} W_{T_j}\left(\nu_j\right)d\nu_j(1+o(1)) + \max_{j=\overline{1,n}}\int_{\Pi}\left|\mu_j\right|^{\gamma}\left|H_{T_j}\left(\mu_j\right)\right|^{\alpha} d\mu_j(1+o(1))\right)$$

$$as\ T \to \infty$$

where

$$C_2 = 2^n n C_1 A\left(\lambda^{(0)}\right). \tag{24}$$

Corollary 2 *Let the coordinate* j *of the number of observations* N_T *be given in the form* $T_j = 2k_j\left(m_j - 1\right) + 1$, $k_j \in \left\{\frac{1}{2}\right\} \cup \mathbb{N}$, $m_j \in \mathbb{N}$, $j = \overline{1,n}$. *Assume that the assumptions in Corollary 1 hold and the functions* $w_j(x) = \widetilde{w}(x)$, $j = \overline{1,n}$, *are of the form*

$$\widetilde{w}(x) = \begin{cases} 1 - |x|, \text{ if } |x| \le 1, \\ 0, \text{ if } |x| > 1, \end{cases} \tag{25}$$

and the functions $h_j\left(\frac{t_j}{\tau_j}\right) = h\left(k_j, m_j\right)$ *are polynomial data windows. These windows can be represented as*

$$\left|H_{T_j}\left(\lambda_j\right)\right|^{\alpha} = \frac{1}{B_{T_j,\alpha}}\left|\frac{\sin\left(\frac{m_j\lambda_j}{2}\right)}{\sin\left(\frac{\lambda_j}{2}\right)}\right|^{2k_j\alpha}, \lambda_j \in \Pi, \tag{26}$$

(see Section 4.4 in Trush (1999) and Demesh (1988)) where

$$B_{T_j,\alpha} = \int_{-\pi}^{\pi}\left|\frac{\sin\left(\frac{m_j\lambda_j}{2}\right)}{\sin\left(\frac{\lambda_j}{2}\right)}\right|^{2k_j\alpha} d\lambda_j,\ j = \overline{1,n}.$$

Then

$$\left|E\widehat{f_T}\left(\lambda^{(0)}\right)-\left[\varphi\left(\lambda^{(0)}\right)\right]^{\frac{p}{\alpha}}\right|\leq C_2C_3\begin{cases}\max_{j=\overline{1,n}}\dfrac{1}{M_{T_j}^{\gamma}}, \textit{ if } 0<\gamma<1,\\ \\ \max_{j=\overline{1,n}}\dfrac{\ln\left(M_{T_j}\right)}{M_{T_j}}, \textit{ if } \gamma=1,\end{cases} \tag{27}$$

for $k_j>\frac{1+\gamma}{2\alpha}$, $j=\overline{1,n}$, *where* C_2 *is defined by (24) and where*

$$C_3=\begin{cases}\frac{2^{\gamma}\pi}{\left(1-\gamma^2\right)}, \textit{ if } 0<\gamma<1,\\ 2\pi, \textit{ if } \gamma=1.\end{cases}$$

If $k_j=\frac{1}{2}$, then the function

$$h\left(\frac{t_j}{\tau_j}\right)=\begin{cases}1, \text{ if } |t_j|\leq\tau_j,\\ 0, \text{ if } |t_j|>\tau_j,\end{cases}$$

is the unit data window, $j=\overline{1,n}$. If $k_j=1$, then the function

$$h\left(\frac{t_j}{\tau_j}\right)=\begin{cases}1-|t_j|, \text{ if } |t_j|\leq\tau_j,\\ 0, \text{ if } |t_j|>\tau_j,\end{cases}$$

is the triangle data window, $j=\overline{1,n}$.

Theorem 2 *Assume that the spectral density* $\varphi(\lambda),\lambda\in\Pi^n$, *of the field* $X(t),t\in\mathbb{Z}^n$ *is bounded on* Π^n, *positive and continuous at the point* $\lambda^{(0)}\in\Pi^n$. *Suppose that the expression* $|H_T(\lambda)|^{\alpha}=\left|A_TH^{(T)}(\lambda)\right|^{\alpha}$, $\lambda\in\Pi^n$, *is a kernel on* Π^n *where* A_T *and* $H^{(T)}(\lambda)$, $\lambda\in\Pi^n$ *are defined by (3) and (2), respectively, and let*

$$\max_{s,r=\overline{1,L_T},s\neq r}\int_{\Pi^n}\left|H_T\left(\nu^{(s)}-\nu\right)H_T\left(\nu^{(r)}-\nu\right)\right|^{\frac{\alpha}{2}}d\nu\longrightarrow 0 \textit{ as } T\to\infty. \tag{28}$$

Then

$$var\,\widehat{f_T}\left(\lambda^{(0)}\right)=O\left(\frac{M_T}{L_T}\right)+O\left(\int_{\Pi^n}\left|H_T\left(\nu^{(s)}-\nu\right)H_T\left(\nu^{(r)}-\nu\right)\right|^{\frac{\alpha}{2}}d\nu\right) \tag{29}$$

where L_T *and* M_T *are given in (5) and (4), respectively.*

Examples of data windows satisfying condition (28) can be found in Appendix 2 (Example 2).

Corollary 3 *Assume that* $H_T(\lambda)$, $\lambda\in\Pi^n$, *satisfies (23) and suppose that the assumptions in Theorem 2 hold. Then*

$$var\widehat{f_T}\left(\lambda^{(0)}\right) = O\left(\frac{M_T}{L_T}\right) + O\left(\max_{j=\overline{1,n},s,r=\overline{1,L_T},s\neq r}\left(\int_{\Pi}\left|H_{T_j}\left(\nu_j^{(s)}-\nu_j\right)H_{T_j}\left(\nu_j^{(r)}-\nu_j\right)\right|^{\frac{\alpha}{2}}d\nu_j\right)\right). \tag{30}$$

Corollary 4 *Assume that* $H_{T_j}\left(\lambda_j\right)$, $j = \overline{1,n}$, *satisfy (26) and suppose that the assumptions in Corollary 3 hold. Let also* $k_j = k$, $j = \overline{1,n}$. *Then*

$$var\widehat{f_T}\left(\lambda^{(0)}\right) = O\left(m^{-(\beta-q)n}\right) + O\left(m^{-\frac{2k^2\alpha^2(1-\beta)-1}{1+2k\alpha}}\right) \tag{31}$$

where $M_{T_j} = m^q$, $L_{T_j} = m^\beta$, $0 < q < \beta < 1$, $j = \overline{1,n}$.

For the data windows in Corollary 2 and the spectral data windows described in Corollary 4, we obtain the rate of convergence in probability.

Theorem 3 *Suppose that the assumptions in Theorem 1 and Theorem 2 hold. Then for any* $\epsilon > 0$

$$P\left(\left|\widehat{\varphi_T}\left(\lambda^{(0)}\right)-\varphi\left(\lambda^{(0)}\right)\right| > \epsilon\right) \longrightarrow 0 \text{ as } T \longrightarrow \infty$$

where $\widehat{\varphi_T}(\lambda)$, $\lambda \in \Pi^n$, *is given by (13).*

Corollary 5 *Suppose that the assumptions in Corollary 2 and Corollary 4 hold. Then*

$$a_m\left|\widehat{\varphi}\left(\lambda^{(0)}\right)-\varphi\left(\lambda^{(0)}\right)\right| \xrightarrow{P} 0$$

where

$$a_m = \begin{cases} m^{\frac{1}{2}\frac{\gamma n\left(2k^2\alpha^2-1\right)}{n\gamma(1+2k\alpha)+k^2\alpha^2(n+2\gamma)}}\frac{1}{\ln(m)} \text{ if } 0<\gamma<1, \\ m^{\frac{1}{2}\frac{n\left(2k^2\alpha^2-1\right)}{n(1+2k\alpha)+k^2\alpha^2(n+2)}}\frac{1}{\ln^2(m)} \text{ if } \gamma = 1, \end{cases}$$

when $2k^2\alpha^2 > 1$, $2k\alpha > 1+\gamma$.

4 Appendix 1

Proof of Theorem 1. As it follows from Trush and Orlova (1994), we have $EI_T(\lambda) = [\psi_T(\lambda)]^{\frac{p}{\alpha}}$, $\lambda \in \Pi^n$. Then

$$\left|E\widehat{f_T}\left(\lambda^{(0)}\right)-\left[\varphi\left(\lambda^{(0)}\right)\right]^{\frac{p}{\alpha}}\right| = \left|E\int_{\Pi^n}W_T(\nu)\,I_T\left(\lambda^{(0)}+\nu\right)d\nu-\left[\varphi\left(\lambda^{(0)}\right)\right]^{\frac{p}{\alpha}}\right| =$$

$$= \left|\int_{\Pi^n}W_T(\nu)\left[\psi_T\left(\lambda^{(0)}+\nu\right)\right]^{\frac{p}{\alpha}}d\nu-\left[\varphi\left(\lambda^{(0)}\right)\right]^{\frac{p}{\alpha}}\right| =$$

$$= \left| \int_{\Pi^n} W_T(\nu) \left[\left[\psi_T\left(\lambda^{(0)}+\nu\right) \right]^{\frac{p}{\alpha}} - \left[\varphi\left(\lambda^{(0)}\right) \right]^{\frac{p}{\alpha}} \right] d\nu \right| \leq$$

$$\leq \int_{\Pi^n} W_T(\nu) \left| \left[\psi_T\left(\lambda^{(0)}+\nu\right) \right]^{\frac{p}{\alpha}} - \left[\varphi\left(\lambda^{(0)}\right) \right]^{\frac{p}{\alpha}} \right| d\nu.$$

By the inequality

$$|a^q - b^q| \leq \frac{q}{2} |a-b| \left(a^{q-1} + b^{q-1} \right) \tag{32}$$

valid for all $a, b > 0$, $q \in (0,1]$, $q \geq 2$, and assuming $a = \psi_T\left(\lambda^{(0)}+\nu\right), b = \varphi\left(\lambda^{(0)}\right)$, $q = \frac{p}{\alpha}$, we obtain

$$\int_{\Pi^n} W_T(\nu) \left| \left[\psi_T\left(\lambda^{(0)}+\nu\right) \right]^{\frac{p}{\alpha}} - \left[\varphi\left(\lambda^{(0)}\right) \right]^{\frac{p}{\alpha}} \right| d\nu \leq$$

$$\leq \frac{p}{2\alpha} \max_{\nu \in \Pi^n} \left| \left[\psi_T\left(\lambda^{(0)}+\nu\right) \right]^{\frac{p}{\alpha}-1} + \left[\varphi\left(\lambda^{(0)}\right) \right]^{\frac{p}{\alpha}-1} \right| \times$$

$$\times \int_{\Pi^n} W_T(\nu) \left| \psi_T\left(\lambda^{(0)}+\nu\right) - \varphi\left(\lambda^{(0)}\right) \right| d\nu \leq$$

$$\leq C_1 \int_{\Pi^n} W_T(\nu) \int_{\Pi^n} |H_T(\mu)|^\alpha \left| \varphi\left(\lambda^{(0)}+\nu+\mu\right) - \varphi\left(\lambda^{(0)}\right) \right| d\mu d\nu.$$

By (14), (15), (17) and (20), we have for any fixed δ, $0 < \delta < \pi$,

$$C_1 \int_{\Pi^n} W_T(\nu) \int_{\Pi^n} |H_T(\mu)|^\alpha \left| \varphi\left(\lambda^{(0)}+\nu+\mu\right) - \varphi\left(\lambda^{(0)}\right) \right| d\mu d\nu =$$

$$= C_1 \int_{\|\nu\| \leq \delta} W_T(\nu) \int_{\|\mu\| \leq \delta} |H_T(\mu)|^\alpha \left| \varphi\left(\lambda^{(0)}+\nu+\mu\right) - \varphi\left(\lambda^{(0)}\right) \right| d\mu d\nu +$$

$$+ C_1 \int_{\Pi^n \setminus \{\|\nu\| \leq \delta\}} W_T(\nu) \int_{\|\mu\| \leq \delta} |H_T(\mu)|^\alpha \left| \varphi\left(\lambda^{(0)}+\nu+\mu\right) - \varphi\left(\lambda^{(0)}\right) \right| d\mu d\nu +$$

$$+ C_1 \int_{\Pi^n} W_T(\nu) \int_{\Pi^n \setminus \{\|\mu\| \leq \delta\}} |H_T(\mu)|^\alpha \left| \varphi\left(\lambda^{(0)}+\nu+\mu\right) - \varphi\left(\lambda^{(0)}\right) \right| d\mu d\nu \leq$$

$$\leq C_1 A\left(\lambda^{(0)}\right) \int_{\|\nu\| \leq \delta} W_T(\nu) \int_{\|\mu\| \leq \delta} |H_T(\mu)|^\alpha \|\nu+\mu\|^\gamma d\mu d\nu +$$

$$+ C_1 \max_{\nu,\mu \in \Pi^n} \left| \varphi\left(\lambda^{(0)}+\nu+\mu\right) - \varphi\left(\lambda^{(0)}\right) \right| \int_{\Pi^n \setminus \{\|\mu\| \leq \delta\}} |H_T(\mu)|^\alpha d\mu +$$

$$+ C_1 \max_{\nu,\mu \in \Pi^n} \left| \varphi\left(\lambda^{(0)}+\nu+\mu\right) - \varphi\left(\lambda^{(0)}\right) \right| \int_{\Pi^n \setminus \{\|\nu\| \leq \delta\}} W_T(\nu) d\nu \leq$$

$$\leq C_1 A\left(\lambda^{(0)}\right) \int_{\Pi^n} \|\nu\|^\gamma W_T(\nu) d\nu (1+o(1)) + C_1 A\left(\lambda^{(0)}\right) \int_{\Pi^n} \|\mu\|^\gamma |H_T(\mu)|^\alpha d\mu (1+o(1))$$

as $T \to \infty$.

Hence

$$\left|E\widehat{f_T}\left(\lambda^{(0)}\right)-\left[\varphi\left(\lambda^{(0)}\right)\right]^{\frac{p}{\alpha}}\right|\leq$$

$$\leq C_1A\left(\lambda^{(0)}\right)\left(\int_{\Pi^n}\|\nu\|^{\gamma}W_T(\nu)\,d\nu(1+o(1))+\int_{\Pi^n}\|\mu\|^{\gamma}|H_T(\mu)|^{\alpha}\,d\mu(1+o(1))\right)$$

$$\text{as } T\to\infty.$$

Theorem 1 is proved. □

Proof of Corollary 1. By (14) and (22), we obtain $W_T(\nu)=\prod_{j=1}^{n}W_{T_j}\left(\nu_j\right)$ where

$$W_{T_j}\left(\nu_j\right)=\sum_{l_j=-M_{T_j}}^{M_{T_j}}w_j\left(\frac{l_j}{M_{T_j}}\right)\exp\left(-i\nu_jl_j\right).$$

Then

$$\int_{\Pi^n}\|\nu\|^{\gamma}W_T(\nu)\,d\nu=\int_{\Pi^n}W_{T_1}(\nu_1)\times W_{T_2}(\nu_2)\times\cdots\times W_{T_n}(\nu_n)\left(\sum_{j=1}^{n}\nu_j^2\right)^{\frac{\gamma}{2}}d\nu_1d\nu_2\cdots d\nu_n\leq$$

$$\leq\int_{\Pi^n}W_{T_1}(\nu_1)\times W_{T_2}(\nu_2)\times\cdots\times W_{T_n}(\nu_n)\left(2^n\sum_{j=1}^{n}\left|\nu_j\right|^{\gamma}\right)d\nu_1d\nu_2\cdots d\nu_n=$$

$$=2^n\sum_{j=1}^{n}\int_{\Pi}\left|\nu_j\right|^{\gamma}W_{T_j}\left(\nu_j\right)d\nu_j\leq 2^nn\max_{j=\overline{1,n}}\left(\int_{\Pi}\left|\nu_j\right|^{\gamma}W_{T_j}\left(\nu_j\right)d\nu_j\right).$$

Therefore

$$\int_{\Pi^n}\|\nu\|^{\gamma}W_T(\nu)\,d\nu\leq 2^nn\max_{j=\overline{1,n}}\left(\int_{\Pi}\left|\nu_j\right|^{\gamma}W_{T_j}\left(\nu_j\right)d\nu_j\right).\tag{33}$$

By (23), we obtain

$$\int_{\Pi^n}\|\mu\|^{\gamma}|H_T(\mu)|^{\alpha}\,d\mu=$$

$$\int_{\Pi^n}\left|H_{T_1}(\mu_1)\right|^{\alpha}\times\left|H_{T_2}(\mu_2)\right|^{\alpha}\times\cdots\times\left|H_{T_n}(\mu_n)\right|^{\alpha}\left(\sum_{j=1}^{n}\mu_j^2\right)^{\frac{\gamma}{2}}d\mu_1d\mu_2\cdots d\mu_n\leq$$

$$\leq\int_{\Pi^n}\left|H_{T_1}(\mu_1)\right|^{\alpha}\times\left|H_{T_2}(\mu_2)\right|^{\alpha}\times\cdots\times\left|H_{T_n}(\mu_n)\right|^{\alpha}\left(2^n\sum_{j=1}^{n}\left|\mu_j\right|^{\gamma}\right)d\mu_1d\mu_2\cdots d\mu_n\leq$$

$$\leq 2^n\sum_{j=1}^{n}\int_{\Pi}\left|\mu_j\right|^{\gamma}\left|H_{T_j}\left(\mu_j\right)\right|^{\alpha}d\mu_j\leq 2^nn\max_{j=\overline{1,n}}\left(\int_{\Pi}\left|\mu_j\right|^{\gamma}\left|H_{T_j}\left(\mu_j\right)\right|^{\alpha}d\mu_j\right).$$

Therefore

$$\int_{\Pi^n} \|\mu\|^\gamma \left|H_T(\mu)\right|^\alpha d\mu \le 2^n n \max_{j=\overline{1,n}} \left(\int_\Pi \left|\mu_j\right|^\gamma \left|H_{T_j}\left(\mu_j\right)\right|^\alpha d\mu_j \right). \tag{34}$$

By substituting (33) and (34) in (19), we have

$$\left| \widehat{f_T}\left(\lambda^{(0)}\right) - \left[\varphi\left(\lambda^{(0)}\right)\right]^{\frac{p}{\alpha}} \right| \le$$

$$\le C_2 \left(\max_{j=\overline{1,n}} \int_\Pi \left|\nu_j\right|^\gamma W_{T_j}\left(\nu_j\right) d\nu_j (1+o(1)) + \max_{j=\overline{1,n}} \int_\Pi \left|\mu_j\right|^\gamma \left|H_{T_j}\left(\mu_j\right)\right|^\alpha d\mu_j (1+o(1)) \right)$$

$$\text{as } T \to \infty.$$

Corollary 1 is proved. □

Proof of Corollary 2. By (25), we obtain

$$W_{T_j}\left(\nu_j\right) = \frac{1}{2\pi M_{T_j}} \frac{\sin^2\left(\frac{M_{T_j}\nu_j}{2}\right)}{\sin^2\left(\frac{\nu_j}{2}\right)}, \nu_j \in \Pi. \tag{35}$$

Since

$$\int_\Pi \left|\nu_j\right|^\gamma W_{T_j}\left(\nu_j\right) d\nu_j \le \begin{cases} \frac{2^\gamma \pi}{\left(1-\gamma^2\right)} \frac{1}{M_{T_j}^\gamma}, \text{ if } 0 < \gamma < 1, \\ 2\pi \frac{\ln\left(M_{T_j}\right)}{M_{T_j}}, \text{ if } \gamma = 1, \end{cases} \tag{36}$$

$j = \overline{1,n}$ (see Trush (1999)). It follows from Demesh (1988) that $\int_\Pi \left|\mu_j\right|^\gamma \left|H_{T_j}\left(\mu_j\right)\right|^\alpha d\mu_j \le \frac{\pi^\gamma\left(2k_j\alpha-1\right)}{2k_j\alpha(\gamma+1)} \frac{1}{T_j^\gamma}$, $k_j > \frac{1+\gamma}{2\alpha}$. By Corollary 1, we can write

$$\left| E\,\widehat{f_T}\left(\lambda^{(0)}\right) - \left[\varphi\left(\lambda^{(0)}\right)\right]^{\frac{p}{\alpha}} \right| \le$$

$$\le C_2 C_3 \begin{cases} \max_{j=\overline{1,n}} \left(\frac{2^\gamma \pi}{\left(1-\gamma^2\right)} \frac{1}{M_{T_j}^\gamma}(1+o(1)) + \frac{\pi^\gamma\left(2k_j\alpha-1\right)}{2k_j\alpha(\gamma+1)} \frac{1}{T_j^\gamma}(1+o(1)) \right), \\ \qquad\qquad \text{if } 0 < \gamma < 1 \text{ as } T \to \infty \\ \max_{j=\overline{1,n}} \left(2\pi \frac{\ln\left(M_{T_j}\right)}{M_{T_j}}(1+o(1)) + \frac{\pi^\gamma\left(2k_j\alpha-1\right)}{2k_j\alpha(\gamma+1)} \frac{1}{T_j^\gamma}(1+o(1)) \right), \quad \text{if } \gamma = 1 \text{ as } T \to \infty . \end{cases}$$

Since $\frac{M_{T_j}}{T_j} \longrightarrow 0$ as $T_j \to \infty$, the convergence to zero of the right-hand side in the last expression will be determined by $\max_{j=\overline{1,n}} \frac{1}{M_{T_j}^\gamma}$, if $0 < \gamma < 1$, and by $\max_{j=\overline{1,n}} \frac{\ln\left(M_{T_j}\right)}{M_{T_j}}$, if $\gamma = 1$. Corollary 2 is proved. □

Proof of Theorem 2. We divide the coordinate j of the parallelepiped Π^n into L_{T_j} equal parts. These partitionings generate the partitioning of the parallelepiped Π^n into L_T parts. Let Q_s be the parallelepiped with number s and let $\nu^{(s)}$ is some point belonging to Q_s, $s = \overline{1, L_T}$.

We have

$$\operatorname{var}\widehat{f_T}\left(\lambda^{(0)}\right) = \operatorname{var}\left(\int_{\Pi^n} W_T(\nu)\, I_T\left(\lambda^{(0)}+\nu\right) d\nu\right) \cong \operatorname{var}\left(\sum_{s=1}^{L_T} W_T\left(\nu^{(s)}\right) I_T\left(\lambda^{(0)}+\nu^{(s)}\right)\frac{(2\pi)^n}{L_T}\right) =$$

$$= \sum_{s=1}^{L_T} \operatorname{var}\left(W_T\left(\nu^{(s)}\right) I_T\left(\lambda^{(0)}+\nu^{(s)}\right)\frac{(2\pi)^n}{L_T}\right) \text{ as } T\to\infty. \tag{37}$$

As it follows from Orlova (1993), we have

$$\operatorname{var} I_T(\lambda) = V(p,\alpha)\left[\psi_T(\lambda)\right]^{\frac{2p}{\alpha}}, \lambda\in\Pi^n, \tag{38}$$

where $V(p,\alpha) = \frac{(k(p,\alpha))^2}{k(2p,\alpha)} - 1$ and where $\psi_T(\lambda)$, $\lambda\in\Pi^n$, is given by (21). By substituting (38) into (37), we can write

$$\operatorname{var}\widehat{f_T}\left(\lambda^{(0)}\right) \cong \left(\frac{(2\pi)^n}{L_T}\right)^2 \sum_{s=1}^{L_T} W_T^2\left(\nu^{(s)}\right) V(p,\alpha)\left[\psi_T\left(\lambda^{(0)}+\nu^{(s)}\right)\right]^{\frac{2p}{\alpha}} +$$

$$+\left(\frac{(2\pi)^n}{L_T}\right)^2 \sum_{s=1}^{L_T}\sum_{r=1,s\neq r}^{L_T} W_T\left(\nu^{(s)}\right) W_T\left(\nu^{(r)}\right) cov\left\{I_T\left(\lambda^{(0)}+\nu^{(s)}\right), I_T\left(\lambda^{(0)}+\nu^{(r)}\right)\right\} \le$$

$$\le \frac{(2\pi)^n}{L_T}\int_{\Pi^n} W_T^2(\nu)\, V(p,\alpha)\left[\psi_T\left(\lambda^{(0)}+\nu\right)\right]^{\frac{2p}{\alpha}} d\nu +$$

$$+ \max_{s,r=\overline{1,L_T},s\neq r} cov\left\{I_T\left(\lambda^{(0)}+\nu^{(s)}\right), I_T\left(\lambda^{(0)}+\nu^{(r)}\right)\right\}\left(\frac{(2\pi)^n}{L_T}\right)^2 \sum_{s=1}^{L_T}\sum_{r=1,s\neq r}^{L_T} W_T\left(\nu^{(s)}\right) W_T\left(\nu^{(r)}\right).$$

Since

$$V(p,\alpha)\left[\psi_T\left(\lambda^{(0)}+\nu\right)\right]^{\frac{2p}{\alpha}} < \infty \text{ and } \frac{(2\pi)^n}{L_T}\int_{\Pi^n} W_T^2(\nu)d\nu \cong \frac{M_T}{L_T}\int_{\mathbb{R}^n} w_T^2(x)dx \text{ as } T\longrightarrow\infty,$$

we have

$$\frac{(2\pi)^n}{L_T}\int_{\Pi^n} W_T^2(\nu)\, V(p,\alpha)\left[\psi_T\left(\lambda^{(0)}+\nu\right)\right]^{\frac{2p}{\alpha}} d\nu = O\left(\frac{M_T}{L_T}\right) \text{ as } T\longrightarrow\infty. \tag{39}$$

by (6) and by the properties of the sequences L_{T_j} and M_{T_j}, $j = \overline{1,n}$.

Consider

$$\left(\frac{(2\pi)^n}{L_T}\right)^2 \sum_{s=1}^{L_T} \sum_{r=1, s\neq r}^{L_T} W_T\left(\nu^{(s)}\right) W_T\left(\nu^{(r)}\right) \leq \left(\frac{(2\pi)^n}{L_T}\right)^2 \sum_{s=1}^{L_T} \sum_{r=1}^{L_T} W_T\left(\nu^{(s)}\right) W_T\left(\nu^{(r)}\right) =$$

$$= \left(\frac{(2\pi)^n}{L_T} \sum_{s=1}^{L_T} W_T(\nu^{(s)})\right)^2 \cong \left(\int_{\Pi^n} W_T(\nu) d\nu\right)^2 = 1 \text{ as } T \longrightarrow \infty.$$

As it follows from Orlova (1993), we have

$$cov\left\{I_T\left(\lambda^{(0)} + \nu^{(s)}\right), I_T\left(\lambda^{(0)} + \nu^{(t)}\right)\right\} = O\left(\int_{\Pi^n} \left|H_T\left(\nu^{(s)} - \nu\right) H_T\left(\nu^{(r)} - \nu\right)\right|^{\frac{\alpha}{2}} d\nu\right).$$

By (39), we then obtain

$$\operatorname{var} \widehat{f_T}\left(\lambda^{(0)}\right) = O\left(\frac{M_T}{L_T}\right) + O\left(\int_{\Pi^n} \left|H_T\left(\nu^{(s)} - \nu\right) H_T\left(\nu^{(r)} - \nu\right)\right|^{\frac{\alpha}{2}} d\nu\right).$$

Theorem 2 is proved. □

Proof of Corollary 3. We have

$$\int_{\Pi^n} \left|H_T\left(\nu^{(s)} - \nu\right) H_T\left(\nu^{(r)} - \nu\right)\right|^{\frac{\alpha}{2}} d\nu = \int_{\Pi^n} \left|H_{T_1}\left(\nu_1^{(s)} - \nu_1\right) H_T\left(\nu_1^{(r)} - \nu_1\right)\right|^{\frac{\alpha}{2}} \times$$

$$\times \left|H_{T_2}\left(\nu_2^{(s)} - \nu_2\right) H_T\left(\nu_2^{(r)} - \nu_2\right)\right|^{\frac{\alpha}{2}} \times \cdots \times \left|H_{T_n}\left(\nu_n^{(s)} - \nu_n\right) H_T\left(\nu_n^{(r)} - \nu_n\right)\right|^{\frac{\alpha}{2}} d\nu_1 d\nu_2 \cdots d\nu_n.$$

The integral of the form $\int_{\Pi} \left|H_{T_j}\left(\nu_j^{(s)} - \nu_j\right) H_T\left(\nu_j^{(r)} - \nu_j\right)\right|^{\frac{\alpha}{2}} d\nu_j$, $j = \overline{1,n}$ is equal to 1 if $\nu_j^{(s)} = \nu_j^{(r)}$. Then

$$\int_{\Pi^n} \left|H_T\left(\nu^{(s)} - \nu\right) H_T\left(\nu^{(r)} - \nu\right)\right|^{\frac{\alpha}{2}} d\nu =$$

$$= O\left(\max_{j=\overline{1,n}, s, r=\overline{1,L_T}, s\neq r} \left(\int_{\Pi} \left|H_{T_j}\left(\nu_j^{(s)} - \nu_j\right) H_{T_j}\left(\nu_j^{(r)} - \nu_j\right)\right|^{\frac{\alpha}{2}} d\nu_j\right)\right).$$

By (29), we obtain the required result. □

Proof of Corollary 4. Since (5) and (4) hold, we have

$$L_T = O\left(m^{\beta n}\right), M_T = O\left(m^{qn}\right). \tag{40}$$

As it follows from Demesh (1988),

$$\int_{\Pi} \left| H_{T_j}\left(\nu_j^{(s)} - \nu_j\right) H_{T_j}\left(\nu_j^{(r)} - \nu_j\right) \right|^{\frac{\alpha}{2}} d\nu_j = O\left(m^{-\frac{2k^2\alpha^2(1-\beta)-1}{1+2k\alpha}}\right), \tag{41}$$

$\beta < 1 - \frac{1}{2k^2\alpha^2}$. Substitute (40) and (41) into (30) to obtain the required result. □

Proof of Theorem 3. We follow the lines of the proof in Masry and Cambanis (1984) for stable random processes.

Let $f\left(\lambda^{(0)}\right) = \left[\varphi\left(\lambda^{(0)}\right)\right]^{\frac{p}{\alpha}}$. By inequality (32) with $a = \widehat{f_T}\left(\lambda^{(0)}\right)$, $b = f\left(\lambda^{(0)}\right)$, $q = \frac{\alpha}{p}$, we have

$$\left|\widehat{\varphi_T}\left(\lambda^{(0)}\right) - \varphi\left(\lambda^{(0)}\right)\right| = \left|\left[\widehat{f_T}\left(\lambda^{(0)}\right)\right]^{\frac{\alpha}{p}} - \left[f\left(\lambda^{(0)}\right)\right]^{\frac{\alpha}{p}}\right| \leq$$

$$\leq \frac{\alpha}{2p}\left|\widehat{f_T}\left(\lambda^{(0)}\right) - f\left(\lambda^{(0)}\right)\right|\left(\left[\widehat{f_T}\left(\lambda^{(0)}\right)\right]^{\frac{\alpha}{p}-1} + \left[f\left(\lambda^{(0)}\right)\right]^{\frac{\alpha}{p}-1}\right).$$

By the equality

$$E\left|\widehat{f_T}\left(\lambda^{(0)}\right) - f\left(\lambda^{(0)}\right)\right|^2 = \operatorname{var}\widehat{f_T}\left(\lambda^{(0)}\right) + \left(E\widehat{f_T}\left(\lambda^{(0)}\right) - f\left(\lambda^{(0)}\right)\right)^2, \tag{42}$$

by (19) and (30), we have $E\left|\widehat{f_T}\left(\lambda^{(0)}\right) - f\left(\lambda^{(0)}\right)\right|^2 \longrightarrow 0$ as $T \longrightarrow \infty$. Therefore

$$\left[\widehat{f_T}\left(\lambda^{(0)}\right)\right]^{\frac{\alpha}{p}-1} + \left[f\left(\lambda^{(0)}\right)\right]^{\frac{\alpha}{p}-1} \xrightarrow{P} 2\left[f\left(\lambda^{(0)}\right)\right]^{\frac{\alpha}{p}-1} \text{ as } T \longrightarrow \infty.$$

Using the Chebyshev inequality for any $\epsilon > 0$ we obtain

$$P\left(\left|\widehat{\varphi_T}\left(\lambda^{(0)}\right) - \varphi\left(\lambda^{(0)}\right)\right| > \epsilon\right) \leq Const\frac{E\left|\widehat{f_T}\left(\lambda^{(0)}\right) - f\left(\lambda^{(0)}\right)\right|^2}{\epsilon^2} \longrightarrow 0$$

as $T \longrightarrow \infty$. Theorem 3 is proved. □

Proof of Corollary 5. If $\gamma \in (0, 1)$, then by (27), (31) and (42) we obtain

$$E\left|\widehat{f_T}\left(\lambda^{(0)}\right) - f\left(\lambda^{(0)}\right)\right|^2 = O\left(m^{-2q\gamma}\right) + O\left(m^{-(\beta-q)n}\right) + O\left(m^{-\frac{2k^2\alpha^2(1-\beta)-1}{1+2k\alpha}}\right).$$

By requiring that all terms on the right-hand side of the last expression tend to zero at the same rate, we obtain

$$\begin{cases} 2q\gamma = (\beta - q)\, n, \\ (\beta - q)\, n = \frac{2k\alpha^2(1-\beta)-1}{2+2k\alpha}. \end{cases}$$

Then

$$\beta = \frac{(n+2\gamma)\left(2k^2\alpha^2-1\right)}{(2\gamma n\,(1+2k\alpha)\,2k^2\alpha^2\,(n+2\gamma))} \text{ and } (\beta-q)\,n = \frac{\gamma n\left(2k^2\alpha^2-1\right)}{n\gamma\,(1+2k\alpha)+k^2\alpha^2\,(n+2\gamma)}.$$

Hence

$$E\left|\widehat{f_T}\left(\lambda^{(0)}\right)-f\left(\lambda^{(0)}\right)\right|^2 = O\left(m^{-\frac{\gamma n\left(2k^2\alpha^2-1\right)}{n\gamma(1+2k\alpha)+k^2\alpha^2(n+2\gamma)}}\right)$$

if $2k^2\alpha^2 > 1$. If $\gamma = 1$, then by (27), (31) and (42) we obtain

$$E\left|\widehat{f_T}\left(\lambda^{(0)}\right)-f\left(\lambda^{(0)}\right)\right|^2 = O\left(m^{-2q}\ln^2(m)\right)+O\left(m^{-(\beta-q)n}\ln^2(m)\right)+$$

$$+O\left(m^{-\frac{2k^2\alpha^2(1-\beta)-1}{1+2k\alpha}}\ln^2(m)\right).$$

By requiring that all items on the right-hand side of the last expression tend to zero at the same rate and making the same calculations as in the case $\gamma \in (0,1)$, we obtain

$$E\left|\widehat{f_T}\left(\lambda^{(0)}\right)-f\left(\lambda^{(0)}\right)\right|^2 = O\left(m^{-\frac{n\left(2k^2\alpha^2-1\right)}{n(1+2k\alpha)+k^2\alpha^2(n+2)}}\ln^2(m)\right)$$

if $2k^2\alpha^2 > 1$. By the Chebyshev inequality for a given $\epsilon > 0$, we have

$$P\left(a_n\left|\widehat{f_T}\left(\lambda^{(0)}\right)-f\left(\lambda^{(0)}\right)\right|>\epsilon\right) \le \frac{a_n^2}{\epsilon^2}E\left|\widehat{f_T}\left(\lambda^{(0)}\right)-f\left(\lambda^{(0)}\right)\right|^2 \le \frac{Const}{\epsilon^2}\frac{1}{\ln^2(m)} \longrightarrow 0$$

as $T \longrightarrow \infty$. Corollary 5 is proved. □

5 Appendix 2

Example 1 Let $h_T(t) = h(k_j, m_j)$ be a polynomial data window. In order to proof (15), we show that the rate of convergence to zero as $T \to 0$ of the integral

$$\int_{\Pi^n\setminus\{\|\lambda\|\le\delta\}} |H_T(\lambda)|^\alpha\, d\lambda \tag{43}$$

is greater than that of

$$\int_{\Pi^n} \|\lambda\|^\gamma\, |H_T(\lambda)|^\alpha\, d\lambda \tag{44}$$

Let us consider $H_T(\lambda) = \prod_{j=1}^{n} H_{T_j}\left(\lambda_j\right)$, $\lambda \in \Pi^n$. Then integral (43) is a product of integrals of the form $\int_{\Pi\setminus\{|\lambda_j|\le\delta\}}\left|H_{T_j}\left(\lambda_j\right)\right|^\alpha d\lambda_j$ for all $j = \overline{1,n}$. It follows from

Demesh (1988) that for any $0 < \delta < \pi$ fixed and for any j, $1 \le j \le n$, we have $\int_{\Pi\setminus\{|\lambda_j|\le\delta\}} \left|H_{T_j}\left(\lambda_j\right)\right|^{\alpha} d\lambda_j = O\left(\frac{1}{\tau_j^{2k_j\alpha-1}}\right)$ as $\tau_j \to \infty$ for polynomial data windows satisfying (26) The last expression tends to zero if $2k_j\alpha - 1 > 0$.

By (34), $\int_{\Pi^n} \|\mu\|^{\gamma} |H_T(\mu)|^{\alpha} d\mu \le 2^n n \max_{j=\overline{1,n}} \left(\int_{\Pi} |\mu_j|^{\gamma} \left|H_{T_j}\left(\mu_j\right)\right|^{\alpha} d\mu_j\right)$. For any j, $1 \le j \le n$, $\int_{\Pi} |\mu_j|^{\gamma} \left|H_{T_j}\left(\lambda_j\right)\right|^{\alpha} d\lambda_j = O\left(\frac{1}{\tau_j^{\gamma}}\right)$ as $\tau_j \to \infty$ (see Demesh (1988)). Therefore if $k_j > \frac{\gamma+1}{2\alpha}$, the integral $\int_{\Pi\setminus\{|\lambda_j|\le\delta\}} \left|H_{T_j}\left(\lambda_j\right)\right|^{\alpha} d\lambda_j$ tends to zero faster than the integral $\int_{\Pi} |\mu_j|^{\gamma} \left|H_{T_j}\left(\lambda_j\right)\right|^{\alpha} d\lambda_j$. Since $k_j \in \left\{\frac{1}{2}\right\} \cup \mathbb{N}$ and $\gamma \in (0,1]$, the condition $2k_j\alpha - 1 > 0$ hold. Hence integral (43) tends to zero faster than integral (44).

Example 2 Let $h_T(t) = h(k_j, m_j)$ be a polynomial data window. Consider $H_T(\lambda) = \prod_{j=1}^{n} H_{T_j}\left(\lambda_j\right)$. Then integral (28) is a product of one-dimensional integrals of the form

$$\max_{s,r=\overline{1,L_{T_j}},s\ne r} \int_{\Pi} \left|H_{T_j}\left(\nu_j^{(s)} - \nu_j\right) H_{T_j}\left(\nu_j^{(r)} - \nu_j\right)\right|^{\frac{\alpha}{2}} d\nu_j. \tag{45}$$

As it follows from Demesh (1988),the following condition holds for all $j, 1 \le j \le n$:

$$\max_{s,r=\overline{1,L_{T_j}},s\ne r} \int_{\Pi} \left|H_{T_j}\left(\nu_j^{(s)} - \nu_j\right) H_{T_j}\left(\nu_j^{(r)} - \nu_j\right)\right|^{\frac{\alpha}{2}} d\nu_j = O\left(\tau_j^{-\frac{2k_j^2\alpha^2-1}{2k_j\alpha+1}}\right) \text{ as } \tau_j \to \infty \tag{46}$$

if $2k_j^2\alpha^2 - 1 > 0$. Since each of the integrals of form (45) tends to zero, their product also tends to zero.

Example 3 Let $W_T(\nu) = \prod_{i=1}^{n} W_{T_i}(\nu_i)$ where $W_{T_i}(\nu_i)$, $\nu_i \in \Pi$, satisfy (35), $\gamma \in (0,1)$. Then the integral $\int_{\Pi^n\setminus\{\|\nu\|\le\delta\}} W_T(\nu)\, d\nu$ is a product of integrals of the form $\int_{\Pi\setminus\{|\nu_j|\le\delta\}} W_{T_j}\left(\nu_j\right) d\nu_j$. As it follows from Trush (1999), we have

$$\int_{\Pi\setminus\{|\nu_j|\le\delta\}} W_{T_j}\left(\nu_j\right) d\nu_j = O\left(\frac{1}{M_{T_j}}\right) \text{ as } T \to \infty. \tag{47}$$

By (33) and (36), we obtain

$$\int_{\Pi^n} \|\nu\|^{\gamma} W_T(\nu)\, d\nu = O\left(\frac{1}{M_{T_j}^{\gamma}}\right) \text{ as } T \to \infty \quad \text{if } \gamma \in (0,1) \tag{48}$$

Since $\gamma \in (0,1)$, integral (47) tends to zero faster than integral (48).

6 Acknowledgements

The authors would like to thank the Belarus State University and the organizing committee of BAS2003 for supporting the participation of the first author in the conference. We would also like to thank an anonymous referee for suggestions which have improved the presentation.

7 References

Brillinger, D.R. (1975). *Time Series: Data Analysis and Theory.* New York: Holt, Rinehart and Winston, Inc.

Demesh, N.N. (1988). Using polynomial kernels in evaluation of spectrum of discrete time random stable processes. *Preprint Academy of Science of Ukraine*, 1-31.

Hardin, C.D. (1982). On the spectral representation of symmetric stable processes. *Journal of Multivariate Analysis*, 12, 385-401.

Hosoya, Y. (1978). Discrete Time Stable Processes and their Certain Properties. *Annals of Probability*, 6, 107-133.

Leonenko, N.N., Ivanov, A.V. (1989). *Statistical Analysis of Random Fields*. Boston: Kluwer Academic Publishers.

Masry, E., Cambanis, S. (1984). Spectral Density Estimation for stationary stable processes. *Stochastic Process. Appl.*, 18, 1-31.

Nolan, J.P. (1988). Path properties of index-β stable fields. *The Annals of Probability*, 16, 1569-1607.

Orlova, E.N. (1993). Investigation of periodogram statistics of spectral density of a homogeneous stable random field. *Proceedings of the conference "Modern Problems of Computer Data Analysis and Modeling"*, 75-79.

Paulauskas, V.J. (1976). Some remarks on multivariate stable distributions. *Journal of Multivariate Analysis*, 6, 356-368.

Press, S.J. (1972). Multivariate stable distributions. *Journal of Multivariate Analysis*, 2, 444-462.

Sabre, R. (1995). Spectral Density Estimation for Stationary Random Fields. *Applied Mathematics (Warsaw)*, 23, 107-133.

Sabre, R. (2000). Discrete Estimation of Spectral Density for Symmetric Stable Process. *Statistica*, 60, 497-519.

Skorokhod, A.V. (1991). *Random Processes with Independent Increments* . Boston: Kluwer Academic Publishers.

Trush, N.N. (1999). *Asymptotical Properties of Time Series*. Minsk: BSU.

Trush, N.N., Orlova, E.N. (1994). Statistical properties of the periodogram of stable random field. *Herald of National Academy of Science of Belarus, ser.phys.-math. sciences*, 2, 33-41.

Zolotarev, V.M. (1986). *One-dimensional Stable Distributions*. v. 65, American Mathematical Society: Translations of Mathematical Monographs.

SORT 29 (1) January-June 2005, 119-138

The M/G/1 retrial queue: An information theoretic approach

J.R. Artalejo*, M.J. Lopez-Herrero

Complutense University of Madrid

Abstract

In this paper, we give a survey of the use of information theoretic techniques for the estimation of the main performance characteristics of the $M/G/1$ retrial queue. We focus on the limiting distribution of the system state, the length of a busy period and the waiting time. Numerical examples are given to illustrate the accuracy of the maximum entropy estimations when they are compared versus the classical solutions.

MSC: 60K25, 90B22

Keywords: Principle of maximum entropy, $M/G/1$ retrial queue, limiting distribution, busy period, waiting time

1 Introduction

In classical queueing theory it is usually assumed that any customer who cannot get service automatically upon arrival either joins a waiting line or leaves the system forever. However, the consideration of loss queueing models is just a first approximation to a more sophisticated situation. Usually the real behaviour of a blocked customer consists of leaving the service area temporarily but he returns to repeat his demand after some random time. This queueing behaviour is studied in the so-called retrial queues (see Falin and Templeton (1987) and Artalejo (1999a, 1999b) for a survey and bibliographical information).

* *Address for correspondence:* J.R. Artalejo, Department of Statistics and Operations Research, Faculty of Mathematics, Complutense University of Madrid, Madrid 28040, Spain.

Received: July 2004

Accepted: February 2005

This paper gives a survey on the application of information theoretic techniques for the estimation of the performance descriptors of the $M/G/1$ queue with retrials. More concretely, we use the principle of maximum entropy (PME) which provides an elegant methodology for computing a unique estimate for an unknown probability distribution based on information expressed in terms of some given mean value constraints. Although we focus on the $M/G/1$ queue, the methodology can be easily extended to other queueing models with retrials (see Artalejo and Martin (1994), Artalejo and Gomez-Corral (1995) and Aissani and Smail (2003)) and without retrials (see Kouvatsos (1994), Tadj and Hamdi (2001) and Wang *et al.* (2002)). For a review of other (non-queueing) applications of the information theoretic methods, we refer the reader to the book by Kapur (1989). As examples of more recent applications we mention the papers by Smolders and Urbach (2002) in the context of x-ray fluorescence spectroscopy and the paper by Oganian and Domingo-Ferrer (2003) published recently in this journal where the reciprocal of Shannon's entropy is used to measure disclosure risk in tabular data.

The paper will be organized as follows. In Section 2 we describe the mathematical model and summarize its main characteristics in terms of generating functions (discrete characteristics) or Laplace transforms (continuous characteristics). A brief overview of the maximum entropy formalism is given in Section 3. Then, the general theory is used to get maximum entropy estimations of the limiting distribution of the system state (Section 4), the length of a busy period (Section 5) and the waiting time (Section 6) (see Falin et al (1994), Lopez-Herrero (2004) and Artalejo et al. (2002)). The numerical experiments show the goodness of the maximum entropy solutions based on the first two moments and the value of the generating function (respectively the Laplace transform) at a given point.

2 The mathematical model

Primary customers arrive to a single server queueing system following a Poisson process of rate λ. Any customer finding the server busy is blocked and leaves temporarily the service area. Such customers join a group of unsatisfied customers called orbit. We assume that the access from the retrial group to the service facility is governed by the classical linear policy, i.e., the probability of a repeated attempt during the interval $(t, t+\Delta t)$, given that j customers were in orbit at time t, is $j\mu\Delta t + o\Delta(t)$. The service times follow a common distribution function $B(x)$ ($B(0) = 0$), with kth moment β_k and Laplace transform $\beta(s)$. The input flow of primary arrivals, intervals between repeated attempts and service times are mutually independent.

The system state at time t can be described by means of the process $X = \{(C(t), N(t), \xi(t));\ t \geq 0\}$, where $C(t)$ denotes the state of the server, 0 or 1 according to whether the server is free or busy, $N(t)$ is the number of customers in orbit at time

t and, if $C(t) = 1$, then $\xi(t)$ represents the elapsed time of the customer being served. In what follows, we neglect the component $\xi(t)$ and consider only the pair $(C(t), N(t))$ which takes values on the state space $S = \{0, 1\} \times \mathbb{N}$. We assume that $\rho = \lambda\beta_1 < 1$ so our queueing model is stable and the limiting probabilities $P_{ij} = \lim_{t\to\infty} P\{C(t) = i, N(t) = j\}$, $(i, j) \in S$, exist and are positive. Then, their corresponding partial generating functions $P_i(z) = \sum_{j=0}^{\infty} z^j P_{ij}$, $i \in \{0, 1\}$, are given by (see Falin and Templeton (1997)):

$$P_0(z) = (1-\rho)\exp\left\{-\frac{\lambda}{\mu}\int_z^1 \frac{1-\beta(\lambda-\lambda u)}{\beta(\lambda-\lambda u)-u}du\right\}, \tag{1}$$

$$P_1(z) = \frac{1-\beta(\lambda-\lambda z)}{\beta(\lambda-\lambda z)-z}P_0(z). \tag{2}$$

By routine differentiation of formulas (1)-(2) we get after some algebra the following expressions for the partial moments $m_i^k = \sum_{j=0}^{\infty} j^k P_{ij}$, for $i \in \{0, 1\}$ and $k = 0, 1, 2$:

$$m_0^0 = 1-\rho,\ m_1^0 = \rho \tag{3}$$

$$m_0^1 = \frac{\lambda\rho}{\mu},\ m_1^1 = \frac{\lambda^2\beta_2}{2(1-\rho)} + \frac{\lambda\rho^2}{\mu(1-\rho)}, \tag{4}$$

$$m_0^2 = m_0^1 + \frac{\lambda}{\mu}m_1^1, \tag{5}$$

$$m_1^2 = \frac{\lambda^3\beta_3}{3(1-\rho)} + \frac{\lambda^4\beta_2^2}{4(1-\rho)^2} + \frac{\lambda^2\beta_2}{2(1-\rho)} + \frac{\lambda^3\beta_2}{2\mu(1-\rho)^2} + \frac{\lambda\rho}{\mu(1-\rho)} + \frac{\lambda^4}{(1-\rho)^2}\left(\frac{\beta_1}{\mu}+\frac{\beta_2}{2}\right)^2 - m_0^2. \tag{6}$$

The busy period of the $M/G/1$ retrial queue, L, starts with the arrival of a primary customer who finds the system empty and ends at the first departure epoch in which the system becomes empty again. The analysis of L in terms of Laplace transforms leads to the following expression

$$L^*(s) = \frac{\int_0^{L_\infty^*(s)} \frac{\beta(s+\lambda-\lambda u)}{e(s,u)(\beta(s+\lambda-\lambda u)-u)}du}{\int_0^{L_\infty^*(s)} \frac{du}{e(s,u)(\beta(s+\lambda-\lambda u)-u)}},\ s > 0, \tag{7}$$

where $L_\infty^*(s)$ represents the Laplace transform for the busy period in the standard retrial queue without retrials given by $L_\infty^*(s) = \beta(s+\lambda-\lambda L_\infty^*(s))$ and $e(s,u)$ is

$$e(s,u) = \exp\left\{\frac{1}{\mu}\int_0^u \frac{s+\lambda-\lambda\beta(s+\lambda-\lambda v)}{\beta(s+\lambda-\lambda v)-v}dv\right\},\ 0 \le u < L_\infty^*(s).$$

The above expression provides a theoretical solution but it has serious limitations in practice. In particular, the moments of L cannot be obtained by direct differentiation. From the theory of regenerative processes, it is easy to get the following formula for the expectation:

$$E[L] = \frac{1}{\lambda}\left(\frac{1}{P_0(0)} - 1\right). \tag{8}$$

A direct method of calculation (see Artalejo and Lopez-Herrero (2000)) for the second moment yields

$$E[L^2] = \frac{1}{P_0(0)}\left(\frac{1}{(1-\rho)^2}\left(\beta_2 + \frac{2\rho\beta_1}{\mu}\right) - \int_0^1 \frac{2}{\lambda\mu(\beta(\lambda-\lambda t)-t)}\right. \tag{1}$$

$$\left.\times\left(1 - \frac{\lambda(1-t)\beta'(\lambda-\lambda t)}{\beta(\lambda-\lambda t)-t} - \frac{1}{1-\rho}\exp\left\{\frac{\lambda}{\mu}\int_t^1 \frac{1-\beta(\lambda-\lambda u)}{\beta(\lambda-\lambda u)-u}du\right\}\right)dt\right). \tag{9}$$

Unfortunately it does not seem possible to numerically invert the density function of L by applying well-known algorithms (see Press *et al.* (1992)) because the above solution (7) is derived when s is a real value, and such algorithmic methods require to evaluate the Laplace transform at any desired complex s. In this sense, the maximum entropy estimation developed in Section 5 provides an elegant alternative to solve this drawback.

Let us assume that a primary customer arrives to the system at time t. His virtual waiting time $W(t)$ is defined as the time that the customer spends in the orbit waiting for service. According to the definition, $W(t)$ excludes the service time. We consider the system at steady state so, in what follows, we simply denote $W(t)$ by W. The analysis of W is intricate because customers in the orbit operate under a random order discipline. Following the book by Falin and Templeton (1997) we observe that the Laplace transform of W is given by

$$W^*(s) = 1 - \rho + \frac{\lambda(1-\rho)}{s}\int_{L^*_{\infty}(s)}^1 \frac{(1-u)(\beta(\lambda-\lambda u)-\beta(s+\lambda-\lambda u))}{(\beta(\lambda-\lambda u)-u)(u-\beta(s+\lambda-\lambda u))} \tag{2}$$

$$\times\exp\left\{\int_u^1 \frac{s+\mu+\lambda-\lambda v}{\mu(\beta(s+\lambda-\lambda v)-v)}dv\right\}\exp\left\{\int_1^u \frac{\lambda-\lambda v}{\mu(\beta(\lambda-\lambda v)-v)}dv\right\}du. \tag{10}$$

An appeal to Little's formula gives

$$E[W] = \frac{\lambda\beta_2}{2(1-\rho)} + \frac{\rho}{\mu(1-\rho)}. \tag{11}$$

Recently, Artalejo *et al.* (2002) obtained the following explicit expression for the second moment:

$$E[W^2] = \frac{2\lambda\beta_3}{3(1-\rho)(2-\rho)} + \frac{\lambda^2\beta_2^2}{(1-\rho)^2(2-\rho)} + \frac{\lambda\beta_2}{\mu}\left(\frac{2}{(1-\rho)^2(2-\rho)} + \frac{\rho}{(1-\rho)^2}\right) + \frac{2\rho}{\mu^2(1-\rho)^2}. \quad (12)$$

Since s is real, we notice again that the most typical techniques for the numerical inversion of W do not apply. However, the above formulas (11)-(12) will be helpful in the sequel. In fact, they constitute the basis for the main value constraints needed to construct maximum entropy estimations.

3 The maximum entropy formalism

Some "classical" queueing techniques include the general framework of birth-and-death processes and methods of solution for non-Markovian stochastic processes (such as embedded Markov chains, supplementary variables, matrix-analytic techniques, etc.). One elegant alternative for this is given by information theoretic methods that use the principles of maximum entropy and minimum cross-entropy (if a prior distribution is available) to estimate probability distributions given information in the form of known mean values. We refer the reader to the survey paper by Kouvatsos (1994) and the references therein.

A novel reader having a first approach to the literature could feed the idea that maximum entropy solutions only provide a reasonable approximation to the true (but complex) queueing system modelled by "classical" techniques. Such interpretation of the information theoretic techniques is poor and trivial. The aim of the PME is to provide a self-contained method of inference for estimating uniquely an unknown probability distribution. The maximum entropy distribution gives the most random solution; i.e., it introduces the minimum additional information beyond what is implied in the original available mean constraints. It should be pointed out that information theoretic analysis neither pretends to replace the "classical" queueing solutions not to be an approximation to that "classical" results. The idea is just to apply the maximum entropy formalism in order to get the widest probability distribution subject to the known constraints. Hence, when in what follows we present "classical" queueing results versus maximum entropy solutions, we only pretend to display two alternative tools for analyzing an unique real underlying queueing phenomenon. It is so far of our intention to suggest a possible (philosophical or numerical) superiority of the "classical" methodology over the maximum entropy approach or vice versa.

We next summarize the maximum entropy formalism (see Shore and Johnson (1981) and Kouvatsos (1994)). The general theory is common for both the discrete and

continuous cases. Thus, we simply denote by $f(x)$ the corresponding mass probability function or density function associated with the queueing performance measure under study. We assume that $f(x)$ takes values in a state space χ, so we have the normalization condition

$$\int_{\chi} f(x)dx = 1. \tag{13}$$

The known information about $f(x)$ can be expressed in terms of mean value constraints of the form

$$\int_{\chi} F_k(x) f(x)dx = F_k, \qquad 1 \leq k \leq m, \tag{14}$$

for known functions $F_k(x)$ and known values F_k. We note that the structural form of the constraints (14) covers important special cases such as:

(a) $F_k(x) = x^k$ (central moments of order k).

(b) $F_k(x) = I_{(-\infty,x_k]}(x)$ (value of the distribution function at the point x_k).

(c) $F_k(x) = e^{-s_k x}$ (value of the Laplace transform or the moment generating function at the point s_k).

The PME states that, of all the distributions satisfying the mean value constraints (13) and (14), the minimal prejudiced is the one maximizing the Shannon's entropy functional

$$H(f) = -\int_{\chi} f(x) \ln f(x)dx. \tag{15}$$

Suppose that a prior distribution $g(x)$ is given as current estimate, then the principle of minimum cross-entropy generalizes the PME by stating that, of all the distributions satisfying the mean constraints, the minimum cross-entropy solution is chosen by minimizing the functional

$$H(f,g) = \int_{\chi} f(x) \ln \frac{f(x)}{g(x)}dx. \tag{16}$$

In fact, the PME corresponds to the particular case when the prior distribution $g(x)$ in (16) is uniformly distributed on the state space χ.

The maximization of $H(f)$ can be carried out with the help of the method of Lagrange's multipliers. If there exists a distribution that maximizes the entropy (15) and satisfies the constraints (13) and (14), then it has the following form

$$\hat{f}(x) = \exp\left\{-\hat{\alpha}_0 - \sum_{k=1}^{m} \hat{\alpha}_k F_k(x)\right\}, \qquad x \in \chi, \tag{17}$$

where $\hat{\alpha}_k$, for $0 \leq k \leq m$, are the Lagrangian multipliers. $\hat{\alpha}_0$ is determined from the normalization condition (13), so we obtain

$$\exp\{\hat{\alpha}_0\} = \int_{\chi} \exp\left\{-\sum_{k=1}^{m} \hat{\alpha}_k F_k(x)\right\} dx. \tag{18}$$

The rest of Lagrangian multipliers satisfy the relations

$$-\frac{\partial \hat{\alpha}_0}{\partial \hat{\alpha}_k} = F_k, \qquad 1 \leq k \leq m. \tag{19}$$

In general, it is impossible to solve (19) for $\hat{\alpha}_k$ explicitly. As an exception, we mention the special case where $m = 1$ and $F_1(x) = x$, which yields the explicit distribution

$$\hat{f}_1(x) = \frac{1}{F_1} e^{-x/F_1}, \qquad x \in \chi. \tag{20}$$

Suppose that we add the second moment as an additional constraint, then the pair $(\hat{\alpha}_1, \hat{\alpha}_2)$ must be computed numerically. By combining (14) and (17) we observe that a standard method for finding the optimal α_k is to solve the system

$$\int_{\chi} (F_i(x) - F_i) \exp\left\{-\sum_{k=1}^{m} \alpha_k (F_k(x) - F_k)\right\} dx = 0, \qquad 1 \leq i \leq m. \tag{21}$$

The above equations (21) for the Lagrangian multipliers are implicit and non-linear. It can be proved then that the problem of solving (21) is equivalent to minimizing the potential function

$$F(\alpha_1, ..., \alpha_m) = \ln \int_{\chi} \exp\left\{-\sum_{k=1}^{m} \alpha_k (F_k(x) - F_k)\right\} dx, \tag{22}$$

or, alternatively, the balanced function

$$G(\alpha_1, ..., \alpha_m) = \sum_{i=1}^{m} p_i \left(\int_{\chi} (F_i(x) - F_i) \exp\left\{-\sum_{k=1}^{m} \alpha_k (F_k(x) - F_k)\right\} dx \right)^2, \tag{23}$$

where $0 < p_i < 1$ and $\sum_{i=1}^{m} p_i = 1$.

The balanced function $G(\alpha_1, ...\alpha_m)$ in (23) takes the value 0 at the optimal solution $(\widehat{\alpha}_1, ..., \widehat{\alpha}_m)$ which provides a computational advantage over the potential function F in (22). For computing the minimum in (23) we will employ a method of direct search (see Nelder and Mead (1964)) which does not involve derivatives, avoiding problems arising when the Hessian of G is algorithmically almost singular. A complete discussion of this technical problem can be found in Agmon *et al.* (1979).

4 Maximum entropy estimation of the system state

After the preceding preliminaries we are ready to apply the maximum entropy methodology to the distribution of the system state in the $M/G/1$ retrial queue. Firstly, we assume that the available information consists of the marginal distribution of the server state and the partial expectations of the number of customers in orbit, so we know expressions (3) and (4).

Distinguishing the server state is important in order to provide a more detailed information. Then, the constraints (3) play the role of the normalization condition (13). According to (20), we expect to find a first order maximum entropy solution $\{\hat{P}^1_{ij}; i \in \{0,1\}, j \geq 0\}$ of geometric type. This is formalized in the following result.

Proposition 1 *If the available information is given by m_i^k, for $i \in \{0,1\}$ and $k \in \{0,1\}$, then according to the PME the estimation of the probability distribution of the system state is*

$$\hat{P}^1_{0j} = \frac{(m_0^0)^2}{m_0^0 + m_0^1}\left(\frac{m_0^1}{m_0^0 + m_0^1}\right)^j, \ j \geq 0, \tag{24}$$

$$\hat{P}^1_{1j} = \frac{(m_1^0)^2}{m_1^0 + m_1^1}\left(\frac{m_1^1}{m_1^0 + m_1^1}\right)^j, \ j \geq 0. \tag{25}$$

Proof. It is sufficient to consider the case $i = 0$. Applying the method of Lagrangian multipliers we get a solution $\hat{P}^1_{0j}$ of the form

$$\hat{P}^1_{0j} = uv^j, \ j \geq 0.$$

Since $\{\hat{P}^1_{0j}; j \geq 0\}$ satisfies the constraints m_0^0 and m_0^1, we find that

$$u = \frac{(m_0^0)^2}{m_0^0 + m_0^1}, \ v = \frac{m_0^1}{m_0^0 + m_0^1}.$$

This proves the desired expression (24). □

According to the geometric structural form, the first order estimations (24) and (25) are decreasing sequences. Nevertheless, the limiting probabilities $\{P_{ij}; j \geq 0\}$ may have a mode at any arbitrary level of the orbit, we say j_i^*, for $i \in \{0,1\}$. In particular, in the case of the $M/M/1$ retrial queue, the distribution is unimodal and the modes are given by

$$j_0^* = \begin{cases} 0, & \text{if } \lambda\rho < \mu, \\ \left[\dfrac{(\lambda-\mu)\rho}{\mu(1-\rho)}\right], & \text{if } \lambda\rho \geq \mu, \end{cases}$$

$$j_1^* = \left[\frac{\lambda\rho}{\mu(1-\rho)}\right],$$

where $[x]$ is the integer part of x. We observe that if $(\lambda-\mu)\rho/\mu(1-\rho)$ (respectively $\lambda\rho/\mu(1-\rho)$) is integer, then j_0^*-1 (respectively j_1^*-1) is also a mode.

In the light of the information about the modes, it is clear that the first order estimation will be accurate only when $j_0^* = j_1^* = 0$, which is equivalent to the inequality $\lambda\rho < \mu(1-\rho)$.

To illustrate the above comments, in Table 1 we consider an $M/M/1$ retrial queue with a small retrial rate $\mu = 0.05$, so that the distribution is sparse and $j_0^* = j_1^* = 6$. Hence, the maximum entropy solution $\hat{P}_{ij}^1$ gives a bad estimation of the probabilities P_{ij}. Hence, the necessity of deriving new estimations of the system state is clear.

Two initial reasons justify the use of two moment estimations. Firstly, in Falin and Templeton (1997) is mentioned that the number of customers in orbit is asymptotically Gaussian, as $\mu \to 0$. This fact agrees with the structural form (17)-(18) of the maximum entropy distribution. Furthermore, by treating j as a continuous variable, we easily see that the k-moment estimation has at most $k-1$ relative extremes.

By adapting the maximum entropy formalism to the case under consideration, we see that the Lagrangian multipliers can be obtained by minimizing the potential functions

$$F_i(\alpha_1^i, \alpha_2^i) = \ln \sum_{j=0}^{\infty} \exp\left\{-\sum_{k=1}^{2} \alpha_k^i \left(j^k - \frac{m_i^k}{m_i^0}\right)\right\}, \quad i \in \{0,1\}. \tag{26}$$

The computation of the infinite series on the right-hand side of (26) implies the consideration of a truncation threshold K which can be determined with the help of Tchebychev's inequality.

The second order estimations $\hat{P}_{ij}^2$ in Table 1 have modes at the seventh level of the orbit. The last row of the table gives the value of the Shannon entropy (15) for the classical distribution and the maximum entropy estimations. As expected, we observe that the entropy decreases when we increase the number of known moments.

It should be noted that the moments m_i^k are obtained by taking derivatives of the partial generating function $P_i(z)$, for $i \in \{0,1\}$, at the point $z = 1$. Hence, it should be interesting to improve the estimation by considering any other constraint providing information related to another different point z_0. To this end, we consider $P_i(z_0)$ which satisfies the structural form described in (14). Now the maximum entropy solution has

Table 1: $M/M/1$ *retrial queue with* $(\lambda, \mu) = (1.0, 0.05)$ *and* $\rho = 0.25$

j	P_{0j}	P_{1j}	$\widehat{P}^{1}_{0j}$	$\widehat{P}^{1}_{1j}$	$\widehat{P}^{2}_{0j}$	$\widehat{P}^{2}_{1j}$	$\widehat{P}^{2,1}_{0j}$	$\widehat{P}^{2,1}_{1j}$
0	0.00237	0.00059	0.09782	0.03125	0.01010	0.00285	0.00186	0.00043
1	0.01189	0.00312	0.08506	0.02734	0.01912	0.00545	0.01141	0.00294
2	0.03121	0.00858	0.07397	0.02392	0.03260	0.00941	0.03265	0.00899
3	0.05723	0.01645	0.06432	0.02093	0.05002	0.01471	0.05984	0.01734
4	0.08226	0.02468	0.05593	0.01831	0.06911	0.02077	0.08341	0.02520
5	0.09872	0.03085	0.04863	0.01602	0.08597	0.02652	0.09719	0.03048
6	0.10283	0.03342	0.04229	0.01402	0.09628	0.03061	0.09975	0.03240
7	0.09549	0.03222	0.03677	0.01227	0.09707	0.03194	0.09277	0.03119
8	0.08057	0.02819	0.03197	0.01073	0.08812	0.03013	0.07944	0.02763
9	0.06266	0.02271	0.02780	0.00939	0.07203	0.02570	0.06317	0.02274
10	0.04543	0.01703	0.02418	0.00822	0.05300	0.01982	0.04687	0.01747
11	0.03097	0.01200	0.02102	0.00719	0.03511	0.01382	0.03253	0.01257
12	0.02000	0.00800	0.01828	0.00629	0.02094	0.00871	0.02116	0.00847
13	0.01231	0.00507	0.01589	0.00550	0.01124	0.00496	0.01290	0.00536
14	0.00725	0.00308	0.01382	0.00481	0.00543	0.00255	0.00738	0.00318
15	0.00411	0.00179	0.01202	0.00421	0.00236	0.00119	0.00396	0.00177
SE	3.05175		3.51236		3.07099		3.05262	

the form

$$\hat{P}^{2,1}_{ij} = \exp\left\{-\left(\widehat{\alpha}^{i}_{0} + j\widehat{\alpha}^{i}_{1} + j^{2}\widehat{\alpha}^{i}_{2} + z_{0}^{j}\widehat{\alpha}^{i}_{2,1}\right)\right\},\ i \in \{0,1\},\ j \geq 0.$$

The Lagrangian coefficients $(\widehat{\alpha}^{i}_{0}, \widehat{\alpha}^{i}_{1}, \widehat{\alpha}^{i}_{2}, \widehat{\alpha}^{i}_{2,1})$ can be computed after a new appeal to the use of Nelder and Mead's algorithm.

The entries in Table 1 show that the estimation improves when we employ $\hat{P}^{2,1}_{ij}$, with $z_0 = 0.55$, instead of $\hat{P}^{2}_{ij}$. In particular, the estimations $\hat{P}^{2,1}_{ij}$ fit the modes of the classical probabilities P_{ij}.

Another different possibility is to employ as constraints the relationships $\lambda P_{1j} = (j+1)\mu P_{0,j+1}$, for $j \geq 0$, which express the conservation of flow across the level j of the orbit. The details and some numerical examples can be found in Falin *et al.* (1994).

In Table 2 we consider a second numerical example in which the system parameters are chosen to fix the traffic intensity $\rho = 0.9$. As a consequence of increasing the value of ρ, the limiting probabilities P_{ij} become sparse. Thus, a good estimation typically demands the use of higher truncation thresholds. For example, to calculate $\{\hat{P}^{2,1}_{1j};\ j \geq 0\}$ we take $K = 80$. The maximum entropy solution $\hat{P}^{2,1}_{ij}$ based on the values $P_i(0.55)$, for $i \in \{0,1\}$, seems to be an accurate estimation. In particular, it fits the modes $j^{*}_{0} = 0$ and $j^{*}_{1} = 1$ of the classical distribution P_{ij}.

The preceding numerical examples deal with a model with exponential service times. However, this assumption is not restrictive and similar conclusions can be obtained for other service time distributions. In fact, once the mean value constraints are fixed, the formalism is independent of the service time distribution.

Table 2: *$M/M/1$ retrial queue with $(\lambda, \mu) = (0.9, 5.0)$ and $\rho = 0.9$*

j	P_{0j}	P_{1j}	$\widehat{P}^1_{0j}$	$\widehat{P}^1_{1j}$	$\widehat{P}^2_{0j}$	$\widehat{P}^2_{1j}$	$\widehat{P}^{2,1}_{0j}$	$\widehat{P}^{2,1}_{1j}$
0	0.06606	0.05946	0.03816	0.07745	0.04492	0.07163	0.06453	0.05889
1	0.01070	0.06314	0.02360	0.07078	0.02383	0.06647	0.01386	0.06307
2	0.00568	0.06194	0.01459	0.06469	0.01294	0.06162	0.00556	0.06289
3	0.00371	0.05909	0.00902	0.05912	0.00719	0.05709	0.00315	0.06030
4	0.00265	0.05558	0.00557	0.05403	0.00408	0.05285	0.00215	0.05657
5	0.00200	0.05182	0.00344	0.04938	0.00237	0.04889	0.00163	0.05243
6	0.00155	0.04804	0.00213	0.04513	0.00141	0.04519	0.00131	0.04827
7	0.00123	0.04435	0.00131	0.04125	0.00086	0.04174	0.00109	0.04427
8	0.00099	0.04081	0.00081	0.03770	0.00053	0.03852	0.00092	0.04051
9	0.00081	0.03746	0.00050	0.03445	0.00034	0.03553	0.00078	0.03702
10	0.00067	0.03432	0.00031	0.03149	0.00022	0.03274	0.00067	0.03381
11	0.00056	0.03139	0.00019	0.02878	0.00014	0.03014	0.00058	0.03085
12	0.00047	0.02868	0.00011	0.02630	0.00010	0.02774	0.00050	0.02814
13	0.00039	0.02617	0.00007	0.02404	0.00007	0.02550	0.00043	0.02566
14	0.00033	0.02385	0.00004	0.02197	0.00005	0.02343	0.00037	0.02339
15	0.00028	0.02173	0.00002	0.02008	0.00003	0.02151	0.00032	0.02132
SE	3.52914		3.56020		3.55273		3.53329	

Table 3: *Sensitivity analysis of the function G on the Lagrangian multipliers*

	$(\lambda, \nu, \mu) = (1.0, 4.0, 0.05)$		$(\lambda, \nu, \mu) = (0.9, 1.0, 5.0)$	
	$i = 0$	$i = 1$	$i = 0$	$i = 1$
$\left(\widehat{\alpha}^i_1 + \varepsilon/2, \widehat{\alpha}^i_2 + \varepsilon/2\right)$	68.9365	94.6871	911.1997	1884403.5
$\left(\widehat{\alpha}^i_1 - \varepsilon/2, \widehat{\alpha}^i_2 + \varepsilon/2\right)$	52.7580	73.2812	857.8347	1745129.8
$\left(\widehat{\alpha}^i_1 - \varepsilon/2, \widehat{\alpha}^i_2 - \varepsilon/2\right)$	72.9542	100.4483	6224.688	14647723.4
$\left(\widehat{\alpha}^i_1 + \varepsilon/2, \widehat{\alpha}^i_2 - \varepsilon/2\right)$	55.6646	77.5092	5413.781	12915677.3
$\left(\widehat{\alpha}^i_1 + \varepsilon, \widehat{\alpha}^i_2, \widehat{\alpha}^i_{2,1}\right)$	0.7506	0.9492	0.5796	4596.297
$\left(\widehat{\alpha}^i_1 - \varepsilon, \widehat{\alpha}^i_2, \widehat{\alpha}^i_{2,1}\right)$	0.7545	0.9542	0.6063	4843.183
$\left(\widehat{\alpha}^i_1, \widehat{\alpha}^i_2 + \varepsilon, \widehat{\alpha}^i_{2,1}\right)$	177.7164	242.7601	184.5680	2869660.8
$\left(\widehat{\alpha}^i_1, \widehat{\alpha}^i_2 - \varepsilon, \widehat{\alpha}^i_{2,1}\right)$	207.7439	285.3387	1189.341	65279520.6
$\left(\widehat{\alpha}^i_1, \widehat{\alpha}^i_2, \widehat{\alpha}^i_{2,1} + \varepsilon\right)$	0.0002	0.0002	0.0010	0.2752
$\left(\widehat{\alpha}^i_1, \widehat{\alpha}^i_2, \widehat{\alpha}^i_{2,1} - \varepsilon\right)$	0.0002	0.0002	0.0010	0.2754

We conclude this section with some practical tips for the computation of Lagrangian multipliers. Some numerical results for sensitivity analysis are also presented in Table 3. Although the potential function F is a strictly convex function and therefore a Newton-Raphson method should converge for any initial guess $(\alpha_1, ..., \alpha_m)$, there are some practical problems. Due to the exponential structure of the maximum entropy solution (17), F becomes asymptotically linear along some directions. Thus, its Hessian eventually becomes algorithmically singular when the initial guess for the multiplier is

chosen close to the asymptotic region. It is also typical that F has a long valley in some direction, then the gradient of F is in the direction of the valley, but not necessarily in the direction of the optimal solution.

The discussion for the balanced function G is analogous but we recall that $G(\hat{\alpha}_1, ..., \hat{\alpha}_m) = 0$. Thus, in Table 3 we deal with G and discuss the sensitivity of the Lagrangian multipliers based on small changes in some determined directions. For the scenarios in Tables 1 and 2, we allow the optimal solutions $(\hat{\alpha}_1^i, \hat{\alpha}_2^i)$ and $(\hat{\alpha}_1^i, \hat{\alpha}_2^i, \hat{\alpha}_{2,1}^i)$ to change along ten directions described in the first column of the table. Then, the entries give the value of G for the choice $\varepsilon = 10^{-3}$. This sensitivity analysis shows the very strong incidence of the initial guess. The effect is stronger when we perturb the Lagrangian multiplier $\hat{\alpha}_2^i$ associated with the second order moment, and also when the distribution is very sparse (see the case $i = 1$ for the example with $\rho = 0.9$).

5 Maximum entropy estimation of the busy period

In this section we illustrate numerically the use of the PME to get an estimation for the density of L. Although the mathematical formalism and numerical techniques are common for both discrete and continuous distributions, the numerical effort to carry out the latter is considerably superior. More precisely, numerical implementation in a discrete case implies the estimation of a finite set of probabilities which can be done in a personal computer after a few minutes run. However, in a continuous situation is necessary to estimate a density function maybe defined over (0,+∞). It typically demands several hours of running time, and so often the program stops without converging to the optimal Lagrangian multipliers.

Initially we assume that the available information consists of the first and second moments of L, which are provided by formulas (8) and (9). After that, we add one more constraint by using the value of the Laplace transform $L^*(s)$ at a given real point $s = s_0$ (see equation (7)). The methodology described in Section 3 yields maximum entropy densities $\widehat{f_2}(x)$ and $\widehat{f_{2,1}}(x)$, respectively. Their corresponding functional forms look as follows

$$\widehat{f_2}(x) = \exp\left\{-\left(\widehat{\alpha}_0 + x\widehat{\alpha}_1 + x^2\widehat{\alpha}_2\right)\right\}, \ x \in (0, T),$$

$$\widehat{f_{2,1}}(x) = \exp\left\{-\left(\widehat{\alpha}_0 + x\widehat{\alpha}_1 + x^2\widehat{\alpha}_2 + e^{-s_0 x}\widehat{\alpha}_{2,1}\right)\right\}, \ x \in (0, T).$$

As a practical remark, we observe that the potential function F and the balanced function G, given in formulas (22) and (23) respectively, involve integrals defined over (0,+∞). Thus, solving the minimization problem implies firstly the consideration of a truncated interval $(0, T)$. The upper bound T may be chosen with the help of Tchebychev's inequality, such as $P\{L > T\} \leq 10^{-2}$.

Note that the maximum entropy densities satisfy the given constraints, in particular the first two moments of $\widehat{f_2}(x)$ and $\widehat{f_{2,1}}(x)$ coincide with the ones of $f_L(x)$, and so does

the Laplace transform associated with $\widehat{f}_{2,1}(x)$, at $s = s_0$, with $L^*(s)$. Thus, we propose to check the accuracy of the maximum entropy estimation by measuring the relative errors associated with their estimates for the Laplace transforms; i.e., we consider

$$E_2(s) = \left| \frac{L_2^*(s)}{L^*(s)} - 1 \right| \text{ and } E_{2,1}(s) = \left| \frac{L_{2,1}^*(s)}{L^*(s)} - 1 \right|,$$

where $L_2^*(s) = \int_0^T e^{-sx} \widehat{f}_2(x)dx$ and $L_{2,1}^*(s) = \int_0^T e^{-sx} \widehat{f}_{2,1}(x)dx$.

Table 4: *Comparing Laplace transforms in an $M/E_3/1$ retrial queue*

	$\rho = 0.25$		$\rho = 0.5$		$\rho = 0.75$	
s	$L^*(s)$	$L_2^*(s)$	$L^*(s)$	$L_2^*(s)$	$L^*(s)$	$L_2^*(s)$
0.05	0.89166	0.90632	0.84622	0.85323	0.76778	0.70380
0.1	0.83598	0.83324	0.76951	0.75077	0.67777	0.54867
0.25	0.69107	0.67587	0.61400	0.55568	0.53178	0.33193
0.5	0.53792	0.51711	0.46800	0.38979	0.40388	0.20069
1.0	0.35856	0.35355	0.30994	0.24492	0.26869	0.11219
1.5	0.25426	0.26907	0.22107	0.17874	0.19324	0.07787
3.0	0.11078	0.15700	0.09882	0.09880	0.08854	0.04061
4.5	0.05802	0.11090	0.05282	0.06828	0.04822	0.02747
6.0	0.03409	0.08574	0.03150	0.05217	0.02916	0.02075
10.0	0.01159	0.05343	0.01095	0.03202	0.01038	0.01256

We next analyze the length of a busy period in an $M/G/1$ retrial queue with Erlang service times; i.e., we have

$$B(x) = \int_0^x \frac{\nu^m}{(m-1)!} e^{-\nu x} x^{m-1} dx, \ x \geq 0,$$

where $m \in \{1, 2, ...\}$ and $\nu > 0$. In particular, we take $m = 3$ phases and $\beta_1 = m/\nu = 1$. Then, the arrival rate is chosen as $\lambda = 0.25$, 0.5 and 0.75. For a given λ, we assume that the retrial rate is $\mu = 2\lambda$. Table 4 presents a comparison between the classical Laplace transform $L^*(s)$ and the maximum entropy solution based on two moments $L_2^*(s)$. For most fixed $s > 0$, we observe that the classical and maximum entropy solutions are closer when the traffic intensity decreases.

We next improve the estimation by considering the maximum entropy solution $\widehat{f}_{2,1}(x)$. For practical purposes, the choice of s_0 can be done by taking into account that the behaviour of $f_L(x)$ and $L^*(s)$ near the boundaries of their domains is determined by the Tauberian relations

$$\lim_{s \to 0} sL^*(s) = \lim_{x \to +\infty} f_L(x) \text{ and } \lim_{s \to +\infty} sL^*(s) = \lim_{x \to 0} f_L(x). \tag{27}$$

Consequently, small values of s_0 provide a better description of the tail behaviour of $f_L(x)$, while a large value of s_0 describes better the behaviour near the origin.

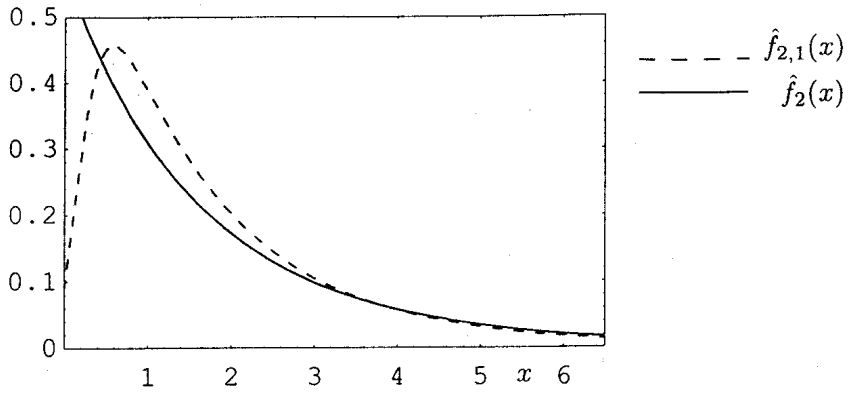

Figure 1: *ME estimations in an $M/E_3/1$ retrial queue with $\rho = 0.25$.*

In Table 5 and Figure 1 we consider and $M/E_3/1$ retrial queue with $\beta_1 = 1$ and $\rho = 0.25$, and we use as constraint the value of $L^*(s)$ at point $s_0 = 4.5$. We can observe in Table 5 that the relative errors are moderately small as s is close to zero. On the other hand, they are notably diminished when we compare large values of s and we employ $\widehat{f_{2,1}}(x)$ rather than $\widehat{f_2}(x)$. This fact and the boundary behaviour given in (27) indicate that, near the origin, the classical density function is much better described by $\widehat{f_{2,1}}(x)$ than by $\widehat{f_2}(x)$.

Table 5: *Relative errors in the $M/E_3/1$ retrial queue for $\rho = 0.25$*

s	$L^*(s)$	$L_2^*(s)$	$E_2(s)$	$L_{2,1}^*(s)$	$E_{2,1}(s)$
0.05	0.89166	0.90632	0.01644	0.90586	0.01593
0.1	0.83598	0.83324	0.00327	0.83101	0.00593
0.25	0.69107	0.67587	0.02200	0.66439	0.03861
0.5	0.53792	0.51711	0.03868	0.48998	0.08912
1.0	0.35856	0.35355	0.01395	0.30709	0.14352
1.5	0.25426	0.26907	0.05823	0.21404	0.15821
3.0	0.11078	0.15700	0.41719	0.09887	0.10750
4.5	0.05802	0.11090	0.91138	0.05802	0.00000
6.0	0.03409	0.08574	1.51504	0.03871	0.13564
10.0	0.01159	0.05343	3.60810	0.01841	0.58837

Figure 1 shows different shapes of the maximum entropy densities $\widehat{f_2}(x)$ and $\widehat{f_{2,1}}(x)$. Previous discussion permits to assert that the classical distribution near the origin should be better represented by $\widehat{f_{2,1}}(x)$; i.e., a bell-shaped function. Moreover, the figure

illustrates the importance of including information about the Laplace transform, because the mode of L is not reproduced unless a constraint on $L^*(s)$ is specified.

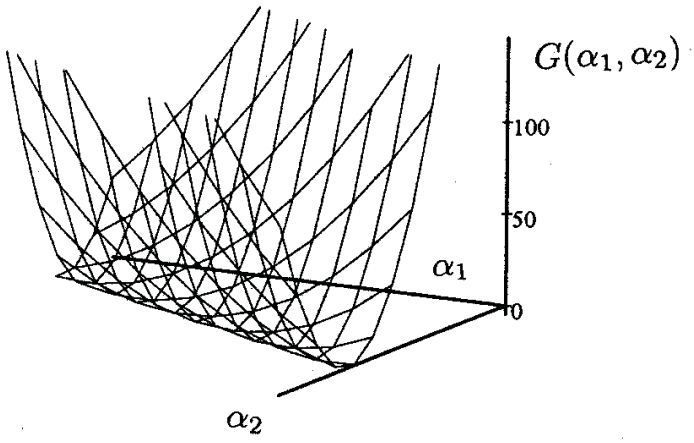

Figure 2: *The balanced function* $G(\alpha_1, \alpha_2)$.

In a second numerical example, see Table 6 and Figure 2, we consider an $M/H_2/1$ retrial queue so $B(x)$ is given by

$$B(x) = \sum_{i=1}^{2} p_i \left(1 - e^{-\nu_i x}\right), \ x \geq 0,$$

where $0 \leq p_1, p_2 \leq 1$, $p_1 + p_2 = 1$ and $\nu_1, \nu_2 > 0$. We consider that the mean service time is $\beta_1 = 1$ and the coefficient of variation $(\beta_2 - \beta_1^2)^{1/2}/\beta_1$ is 1.25. The parameters of the H_2 distribution cannot uniquely determined fitting the above values unless we add another additional condition. Thus, we assume that the distribution has balanced means; i.e., $p_1/\nu_1 = p_2/\nu_2$. The retrial rate is chosen as $\mu = \lambda/2$.

In Table 6 we compare the Laplace transforms associated with the maximum entropy densities based on two moments and two moments plus the value of $L^*(s)$ at the point $s_0 = 4.5$. The entries for the relative errors $E_{2,1}(s)$ are smaller than the errors $E_2(s)$, showing the superiority of the estimation $\widehat{f_{2,1}}(x)$.

In Figure 2, we plot the balanced function $G(\alpha_1, \alpha_2)$ in a neighbourhood of the Lagrangian multipliers $(\hat{\alpha}_1, \hat{\alpha}_2)$. In agreement with the comments expressed in Section 4, the surface shows a rapid growth of G along some directions. The existence of a valley is also observed.

For a numerical analysis of the number of customers served during a busy period, we refer to the paper by Lopez-Herrero (2002).

Table 6: *Relative errors in the $M/H_2/1$ retrial queue for $\rho = 0.25$*

s	$L^*(s)$	$L_2^*(s)$	$E_2(s)$	$L_{2,1}^*(s)$	$E_{2,1}(s)$
0.05	0.86545	0.82936	0.04170	0.83688	0.03301
0.1	0.80664	0.71491	0.11371	0.73749	0.08573
0.25	0.70620	0.50750	0.28135	0.57245	0.18939
0.5	0.60442	0.34316	0.43224	0.45084	0.25410
1.0	0.47820	0.20868	0.56361	0.35129	0.26539
1.5	0.39784	0.15000	0.62294	0.30388	0.23617
3.0	0.26601	0.08140	0.69398	0.23558	0.11438
4.5	0.20016	0.05586	0.72090	0.20016	0.00000
6.0	0.16049	0.04252	0.73503	0.17635	0.09887
10.0	0.10486	0.02598	0.75224	0.13675	0.30408

6 Maximum entropy estimation of the waiting time

We now present a maximum entropy analysis of the waiting time W based on the knowledge of $E[W]$, $E[W^2]$ and $W^*(s)$ at a given positive real point $s = s_0$ (see formulas (10)-(12)). Since the definition of W excludes the service time, we notice that the distribution function of W, $F_W(x)$, has a jump at $x = 0$ and is absolutely continuous in the interval $(0,+\infty)$. Thus, we have

$$\frac{dF_W(x)}{dx} = (1-\rho)u_0(x) + f_W(x),\ x \geq 0,$$

where $u_0(x)$ is the unit impulse at the origin defined by

$$u_0(x) = \begin{cases} +\infty, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0. \end{cases}$$

Hence, the estimation of the distribution of W reduces to the density $f_W(x)$ of the continuous contribution. We first assume that the available information consists of the first two moments $E[W]$ and $E[W^2]$. In a second step, we add the knowledge of $W^*(s)$ at the point $s = 0.25$. We denote both estimations by $\widehat{f_2}(x)$ and $\widehat{f_{2,1}}(x)$ respectively.

In Table 7 we consider the $M/M/1$ retrial queue with $\lambda = 0.5$, $\nu = 1.0$ and $\mu = 2.0$, so the traffic intensity is $\rho = 0.5$. We evaluate the accuracy of the maximum entropy solutions by comparing the classical Laplace transform $W^*(s)$ versus the maximum entropy versions $W_2^*(s)$ and $W_{2,1}^*(s)$, which are given by $W_2^*(s) = 1 - \rho + \int_0^T e^{-sx}\widehat{f_2}(x)dx$ and $W_{2,1}^*(s) = 1 - \rho + \int_0^T e^{-sx}\widehat{f_{2,1}}(x)dx$.

The entries $E_2(s)$ and $E_{2,1}(s)$ correspond to the relative errors which are defined analogously to those given in Section 5 for the busy period. We observe that the relative errors decrease when we employ $\widehat{f_{2,1}}(x)$ rather than $\widehat{f_2}(x)$. The upper bound T is 29.5.

Table 7: *Relative errors in the $M/M/1$ retrial queue for $\rho = 0.5$*

s	$W^*(s)$	$W_2^*(s)$	$E_2(s)$	$W_{2,1}^*(s)$	$E_{2,1}(s)$
0.01	0.98548	0.98548	1.27×10^{-6}	0.98548	7.58×10^{-8}
0.05	0.93549	0.93559	0.00010	0.93548	9.56×10^{-6}
0.1	0.88632	0.88682	0.00057	0.88627	5.98×10^{-5}
0.25	0.78842	0.79120	0.00325	0.78842	0.00000
0.3	0.76562	0.76928	0.00478	0.76579	0.00022
0.6	0.67751	0.68605	0.01261	0.67987	0.00407
0.9	0.63035	0.64237	0.05078	0.63532	0.00788
1.2	0.60109	0.61536	0.02374	0.60823	0.01187
1.5	0.58130	0.59699	0.02699	0.59007	0.01508
10.0	0.50633	0.51763	0.02232	0.51545	0.01801
20.0	0.50193	0.50898	0.01405	0.50781	0.01171

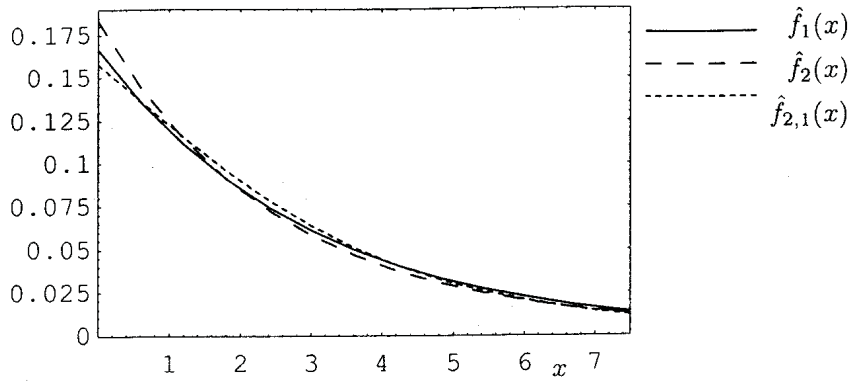

Figure 3: *ME estimations in an $M/M/1$ retrial queue with $\rho = 0.5$.*

For the same numerical example, in Figure 3 we display the maximum entropy densities $\widehat{f_1}(x)$, $\widehat{f_2}(x)$ and $\widehat{f_{2,1}}(x)$. In the light of the decreasing shape of the three densities, we conclude that all these solutions are enough close. However, at this point we remember the Tauberian relations (27), which give some light about the effect of the auxiliary point s_0. Accordingly, in Figure 4 we allow s_0 to take values 0.25, 0.5, 1.0 and 2.0. As far as s_0 increases we expect to get a better description of the behaviour of $f_W(x)$ near the origin $x = 0$. In fact, we observe that the densities associated with the values 0.25 and 0.5 are decreasing functions whereas the densities based on the values 1.0 and 2.0 exhibit a bell-shaped form.

Finally, in Figure 5 we plot the potential function $F(\alpha_1, \alpha_2)$. The resulting surface is complementary to Figure 2, where we consider the busy period and plot the balanced function G. Once more we observe the existence of a long valley and asymptotic linearity when we leave a neighbourhood of the Lagrangian multipliers.

The numerical results show that the use of the first two moments and the value of the Laplace transform in a given point is, in general, sufficient to obtain accurate estimations.

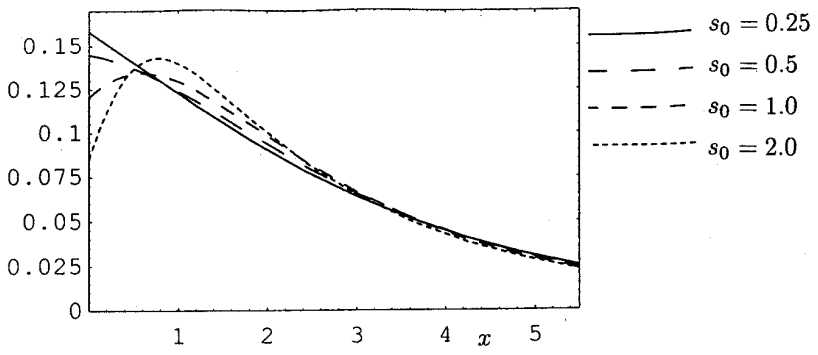

Figure 4: *The effect of the point* s_0.

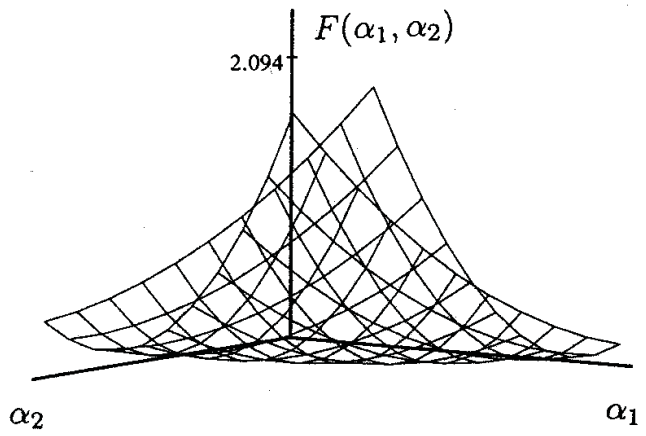

Figure 5: *The potential function* $F(\alpha_1, \alpha_2)$.

7 Acknowledgements

The authors thank the support received from the project BFM2002-02189.

8 References

Agmon, N., Alhassid, Y. and Levine, R.D. (1979). An algorithm for finding the distribution of maximal entropy. *Journal of Computational Physics*, 30, 250-258.

Aissani, A. and Smail, R. (2003). Information theoretic approximations for the $M/G/1$ retrial queue with unreliable server (preprint).

Artalejo, J.R. (1999*a*). Accessible bibliography on retrial queues. *Mathematical and Computer Modelling*, 30, 1-6.

Artalejo, J.R. (1999*b*). A classified bibliography of research on retrial queues: Progress in 1990-1999. *Top*, 7, 187-211.

Artalejo, J.R. and Gomez-Corral, A. (1995). Information theoretic analysis for queueing systems with quasi-random input. *Mathematical and Computer Modelling*, 22, 65-76.

Artalejo, J.R. and Lopez-Herrero, M.J. (2000). On the busy period of the $M/G/1$ retrial queue. *Naval Research Logistics*, 47, 115-127.

Artalejo, J.R., Falin, G.I. and Lopez-Herrero, M.J. (2002). A second order analysis of the waiting time in the $M/G/1$ retrial queue. *Asia-Pacific Journal of Operational Research*, 19, 131-148.

Artalejo, J.R. and Martin, M. (1994). A maximum entropy analysis of the $M/G/1$ queue with constant repeated attempts. In *Selected Topics on Stochastic Modelling*, R. Gutiérrrez and M.J. Valderrama (eds.), World Scientific, 181-190.

Falin, G.I., Martin, M. and Artalejo, J.R. (1994). Information theoretic approximations for the $M/G/1$ retrial queue. *Acta Informatica*, 31, 559-571.

Falin, G.I. and Templeton, J.G.C. (1997). *Retrial Queues.* London: Chapman and Hall.

Kapur, J.N. (1989). *Maximum Entropy Models in Sciences and Engineering.* New York: John Wiley.

Kouvatsos, D.D. (1994). Entropy maximization and queueing networks models. *Annals of Operations Research*, 48, 63-126.

Lopez-Herrero, M.J. (2002). On the number of customers served in the $M/G/1$ retrial queue: First moments and maximum entropy approach. *Computers and Operations Research*, 29, 1739-1757.

Lopez-Herrero, M.J. (2004). A maximum entropy approach for the busy period of the $M/G/1$ retrial queue. *Annals of Operations Research* (to appear).

Nelder, J.A. and Mead, R. (1964). A simplex method for function minimization. *Computer Journal*, 7, 308-313.

Oganian, A. and Domingo-Ferrer, J. (2003). A posteriori disclosure risk measure for tabular data based on conditional entropy. *SORT*, 27, 175-190.

Press, W.H., Teuklosky, S.A., Vetterling, W.T. and Flannery, B.P. (1992). *Numerical Recipes in Fortan. The Art of Scientific Computing.* Cambridge: Cambridge University Press.

Shore, J.E. and Johnson, R.W. (1981). Properties of cross-entropy minimization. *IEEE Transactions on Information Theory*, 27, 472-482.

Smolders, S.M.P. and Urbach, H.P. (2002). On the determination of dopant-concentration profiles by grazing emission X-ray fluorescence spectroscopy using the maximum-entropy method. *Journal of Engineering Mathematics*, 43, 115-134.

Tadj, L. and Hamdi, A. (2001). Maximum entropy solution to a quorum queueing system. *Mathematical and Computer Modelling*, 34, 19-27.

Wang, K.H., Chuang, S.L. and Pearn, W.L. (2002). Maximum entropy analysis to the N policy $M/G/1$ queueing systems with a removable server. *Applied Mathematical Modelling*, 26, 1151-1162.

Book reviews

PROBABILITY: A GRADUATE COURSE

Allan Gut

Springer, New York, 2005
603 pages

This is a graduate level text in probability theory. The book is clearly written and the emphasis is placed on theory, while there is mention of some applications. The reader needs to have a moderate background of measure theory.

Chapters 1 to 8 make up the body of a graduate probability course. It is a little surprising, but very useful, that there is a chapter, Chapter 3, dedicated exclusively to probabilistic inequalities. Also included is Chapter 9, which is of a more specialized nature, and Chapter 10 which deals with martingale theory.

The book begins with basics from the measure theory, such as σ-algebras, set theory, measurability and Lebesgue integration. Later it goes to the Borel-Cantelli lemmas, inequalities, characteristic functions and the classical limit theorems: the Law of Large Numbers, the Central Limit Theorem and the Law of the Iterated Logarithm. After an introduction to the generalizations and extensions of these three classical limit theorems, the book concludes with a chapter devoted to martingales, one of the most important tools in probability theory.

A list of notations and symbols precedes the main body of the text and an appendix with some mathematical tools and facts, a bibliography and an index conclude the book. At the end of each chapter there is a list of problems with some comments.

In the first chapter, *Introductory Measure Theory*, we find the basis concepts of measure theory necessary to read the book.

The second chapter is called *Random Variables*. Here, the author introduces random variables and presents the basis concepts (distributions, expectation, moments, independence, conditional distributions, Borel-Cantelli lemmas...) and concrete applications of probability models.

The third chapter is a collect of different inequalities. Inequalities play an important role in probability theory because very often one needs to estimate: certain probabilities by others, moments of sums by sums of moments, etc. . . In this chapter we find, among

other topics, tail probabilities estimated via moments (Markov's inequality, Chebyshev's inequality, Kolmogorov's inequality…), moment inequalities, Jensen's inequality and probability inequalities for maxima.

Chapter 4 is devoted to characteristic functions. In this chapter the author defines characteristic functions, proves some basic facts as uniqueness, inversion and the multiplication property and introduces the cumulant generating function, the probability generating function and the moment generating function.

The title of Chapter 5 is *Convergence*. This is an introductory chapter to the four following ones where the classical limit theorems are studied. In this chapter the author defines the various modes of convergence, proves the uniqueness of limits and the relations between them. The author also presents some results that are useful when one needs to investigate the convergence of a sequence of random variables.

In Chapter 6 the author presents the Weak and the Strong Laws of Large Numbers as well as some of their applications.

Chapter 7 is devoted to the Central Limit Theorem. In this chapter the Lindeberg-Lévy-Feller theorem, where the summands are independent but not identically distributed, and Lyapounov's version of the result are also proved. After this, the author presents the Berry-Esseen theorem, which is a convergence rate result for the central limit theorem. The remaining part of the chapter contains some rate results for tail probabilities and some comments on local limit theorems for discrete random variables and the concept of large deviations.

Continuing with the classical limit theorems, Chapter 8 is devoted to the Law of the Iterated Logarithm. In this chapter the main focus is on the Hartman-Wintner-Strassen theorem, which deals with the independent and identically distributed case.

The results given in Chapters 6, 7 and 8 are basically for sums of independent, identically distributed random variables. In Chapters 7 and 8 finite variance was an additional assumption. In Chapter 9 the author provides an introduction to some more general limit theorems.

Finally, Chapter 10 is devoted to martingales. Following some introductory material on conditional expectations and the definition of a martingale, the author presents us with some examples, convergence results, results for stopped martingales, regular martingales, uniformly integrable martingales, stopped random walks, reversed martingales and submartingales.

Xavier Bardina
Departament de Matemàtiques
Universitat Autònoma de Barcelona

TESTING STATISTICAL HYPOTHESES. THIRD EDITION. A CLASSIC LEHMANN AND ROMANO BOOK

Allan Gut

Springer-Verlag, New York, 2005
784 pages

It is at the same time easy and difficult to say something about a book that has become a classic, and this "Testing Statistical Hypotheses" is evidently the third edition of a book in this category. Almost all statisticians know and have at some time used the first edition published in 1959 or the second edition published in 1986.

In 1997 Lehman wrote in Statistical Science, vol. 12, a small but interesting article titled "The story of a book", where he recounts the lengthy genesis, philosophy, reception and publishing history of the first two editions. There he explained that with the money received for the book, he could not have built a house, not even a modest one, but it did enable him to buy a fancier car.

In the third edition, Lehman and Romano appear as authors, and the book is dedicated to the memory of Le Cam and Tukey. You see that all these names are top level and well-known among statisticians.

Jointly with the second edition of the companion volume "Theory of Point Estimation", written by Lehman and Casella, these constitute a set of obligatory references in all courses of advanced Statistics.

The present edition is divided into two parts. Part I treats small sample theory in chapters 1-10, while part II treats large sample theory in chapters 11-15.

As the authors note in the preface, the two principal additions in this edition have been the treatment of multiple comparisons and the asymptotic optimality, following the ideas of Le Cam and Tukey.

All the chapters have a remarkable collection of problems and finish with very interesting historical notes. The style makes the reading of the book very agreeable.

In summary, this is a recommendable book that is already a classic and has become an essential reference in this subject.

Review by Josep Lluís Solé
Departament de Matemàtiques
Universitat Autònoma de Barcelona

Information for authors and subscribers

Information for authors and subscribers

Submitting articles to SORT

Guidelines for submitting articles

SORT accepts for publication only original articles that have not been submitted to any other journal in the areas of Statistics, Operations Research, Official Statistics and Biometrics. Articles may be either theoretical or applied, and may include computational or educational elements. Publication will be exclusively in English, with an abstract in Catalan (abstract translation into Catalan can be done by the journal's staff).

All articles submitted to thematic sections will be forwarded for systematic peer review by independent specialists and/or members of the Editorial Board, except for those articles specifically invited by the journal or reprinted with permission. Reviewer comments will be sent to the primary author if changes are requested in form or content.

When an original article is received, the journal sends the author a dated receipt; this date will appear in the publication as "reception date". The date in which its final version is received will be the "acceptance date".

To submit an article, the author must send it in PDF format to the electronic address of the journal (sort@idescat.es) especifying his e-mail address and/or a phone number. Once the article has been approved, the author must send a final electronic version, following the instructions of the editor responsible for that article. It is recommended that this final version be submitted using a LaTeX 2_ε.

In any case, upon request the journal secretary will provide authors with LaTeX 2_ε templates and appropriate references to the MSC2000 classification of the American Mathematical Society. For more information about the criteria for using this classification system, authors may consult directly the Mathematics Subject Classification 2000 (MSC2000) at the website of the Institut d'Estadística de Catalunya (http://www.idescat.es/sort/Normes.stm).

Publishing rights and authors' opinions

Authored articles represent the opinions of the authors; the journal does not necessarily agree with the opinions expressed in authored articles.

Formal guidelines for articles

The first page of the article must contain the following items: title, name of the author(s), professional affiliation and complete address, and an abstract (100-200 words) followed by MSC2000 classification of the American Mathematical Society and the keywords. Before submitting an article, the author(s) would be well advised to ensure that the text uses correct English.

Bibliographic references have to be listed alphabetically at the end of the article, according to the following examples:

Citations
Mahalanobis (1936), Rao (1982b)

Journal articles
Pfeffermann, D. and Barnard, C. H. (1991). Some new estimators for small-area means with application to the assessment of farmland values. *Journal of Business & Economic Statistics*, 9, 73-84.

Books
Platek, R., Rao, J. N. K., Särndal, C. E. and Singh, M. P. (eds.) (1987). *Small Area Statistics: an International Symposium*. New York: John Wiley and Sons.

Parts of books
Lee, E.W., Wei, L.J. and Amato, D.A. (1992). Cox-type regression analysis for large numbers of small groups of correlated failure time observations. In *Survival Analysis: State of the Art*, J.P. Klein and P.K. Goel (eds.), Kluwer Academic Publishers, 237-247.

Web files or "pages"
Nielsen, S. F. (2001). *Proper and improper multiple imputation*
http://www.stat.ku.dk/~feodor/publications/ (10th May 2003).

Multiple publications by a single author are to be listed chronologically. Explanatory notes should be numbered sequentially and placed at the bottom of the corresponding page. Tables and figures should also be numbered sequentially and will be reproduced directly from the submitted originals if it is impossible to include them in the electronic text.

SORT *(Statistics and Operations Research Transactions)*
Institut d'Estadística de Catalunya (Idescat)
Via Laietana, 58
08003 Barcelona (Spain)
Tel.: +34-93 412 09 24 or +34-93 412 00 88
Fax: +34-93 412 31 45
E-mail: sort@idescat.es

How to cite articles published in SORT

Commenges, D. (2003). Likelihood for interval-censored observations from multi-state models. *SORT*, 27 (1), 1-12.

Subscription form
SORT *(Statistics and Operations Research Transactions)*

Name _____
_____
Organisation _____
_____
Street Address _____
_____
Zip/Postal code _____ City _____
State/Country _____ Tel. _____
Fax _____ NIF/VAT Registration Number _____
E-mail _____
Date _____

Signature

I wish to subscribe to **SORT *(Statistics and Operations Research Transactions)*** for the year 2005 (volume 29)

Annual subscription rates:
– Spain: €22 (VAT included)
– Other countries: €25 (VAT included)

Price for individual issues (current and back issues):
– Spain: €9/issue (VAT included)
– Other countries: €11/issue (VAT included)

Method of payment:

☐ Bank transfer to account number 2013-0100-53-0200698577

☐ Automatic bank withdrawal from the following account number
☐☐☐☐ ☐☐☐☐ ☐☐ ☐☐☐☐☐☐☐☐☐☐

☐ Check made payable to the Institut d´Estadística de Catalunya

Please send this subscription form (or a photocopy) to:

SORT *(Statistics and Operations Research Transactions)*
Institut d´Estadística de Catalunya (Idescat)
Via Laietana, 58
08003 Barcelona
SPAIN
Fax: +34-93-412 31 45

Bank copy

Authorisation for automatic bank withdrawal in payment for

SORT *(Statistics and Operations Research Transactions)*

The undersigned _____

authorises Bank/Financial institution _____

located at (Street Address) _____

Zip/postal code _____ City _____

Country _____

to draft the subscription to **SORT *(Statistics and Operations Research Transactions)*** from my account

number ☐☐☐☐ ☐☐☐☐ ☐☐ ☐☐☐☐☐☐☐☐☐☐

Date _____

Signature

SORT *(Statistics and Operations Research Transactions)*
Institut d´Estadística de Catalunya (Idescat)
Via Laietana, 58
08003 Barcelona
SPAIN
Fax: +34-93-412 31 45

Quatre modalitats de subscripció al DOGC

(Diari Oficial de la Generalitat de Catalunya)

Generalitat de Catalunya

Imprès, edició diària

Base de dades, actualització diària

DVD, edició semestral

A la carta, servei diari personalitzat

A més, per als subscriptors de l'edició impresa i del DVD, tramesa gratuïta d'un CD-ROM trimestral que conté les pàgines en format PDF (DOGC en imatges)

L'Administració més a prop

EADOP • Informació i subscripcions • Rocafort, 120 - Calàbria, 147 • 08015 Barcelona
Tel. 93.292.54.17 • Fax 93.292.54.18 • subsdogc@gencat.net • www.gencat.net/eadop