



HAL
open science

Effect of dependency in systems for multivariate surveillance

Eva Andersson

► **To cite this version:**

Eva Andersson. Effect of dependency in systems for multivariate surveillance. Communications in Statistics - Simulation and Computation, 2009, 38 (03), pp.454-472. 10.1080/03610910802531307 . hal-00514340

HAL Id: hal-00514340

<https://hal.science/hal-00514340>

Submitted on 2 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Effect of dependency in systems for multivariate surveillance

Journal:	<i>Communications in Statistics - Simulation and Computation</i>
Manuscript ID:	LSSP-2008-0149.R1
Manuscript Type:	Original Paper
Date Submitted by the Author:	13-Sep-2008
Complete List of Authors:	Andersson, Eva; Statistical research unit, Goteborg University and Department of Occupational and Environmental medicine, Goteborg University
Keywords:	surveillance, multivariate, likelihood ratio, optimal, dependency
Abstract:	Systems for multivariate on-line surveillance (e.g. outbreak detection), are investigated. Optimal systems for statistical surveillance are based on likelihood ratios. Three systems are compared; based on each marginal density, based on the joint density and based on the Hotelling's T ² . The effect of dependency between the monitored processes is investigated, and the effect of correlation between the change times. When the first change occurs immediately, the three methods give similar delay of an alarm, in the situation with independency. For late changes, T ² has the longest delay, both for independent processes and for processes with a positive covariance.



Effect of dependency in systems for multivariate surveillance

E Andersson

Statistical Research Unit, Göteborg University, Sweden

Dep of Occupational and Environmental medicine, Sahlgrenska University Hospital,
Göteborg, Sweden

Corresponding author: E Andersson

Short title: Dependency in multivariate on-line surveillance

Abstract

Systems for multivariate on-line surveillance (e.g. outbreak detection), are investigated. Optimal systems for statistical surveillance are based on likelihood ratios. Three systems are compared; based on each marginal density, based on the joint density and based on the Hotelling's T^2 . The effect of dependency between the monitored processes is investigated, and the effect of correlation between the change times. When the first change occurs immediately, the three methods give similar delay of an alarm, in the situation with independency. For late changes, T^2 has the longest delay, both for independent processes and for processes with a positive covariance.

Keywords: Surveillance; Multivariate; Likelihood ratio; Optimal; Dependency.

1. Introduction

In many situations it is important to monitor a process in order to detect an important change as soon as possible. Examples are turning point detection in business cycles

1
2
3 (Neftci (1982), Hamilton (1989), Andersson, et al. (2004)), in hormone cycles (Royston
4
5 (1991)), in influenza cycles (Baron (2002)) and in financial cycles (Frisén (2007), Bock
6
7 (2008)). Other areas are detection of growth retardation of foetuses (Petzold, et al.
8
9 (2004)) or monitoring of prematurely born children (Löfqvist, et al. (2006)). Yet another
10
11 is detection of an increased level, emerging from a source and spreading spatially (Järpe
12
13 (1999)). On-line monitoring is based on repeated decisions: at each time point, a new
14
15 observation becomes available and a new decision is made as to whether the process has
16
17 changed or not. The methodology of statistical surveillance is appropriate.
18
19
20
21

22 Statistical surveillance is a methodology for discriminating between two events, "the
23
24 change has occurred" and "the change has not occurred". The time of change is unknown.
25
26 The time scale can differ between applications (days, weeks, months), but common to all
27
28 surveillance are the repeated decisions, made at each time point. The decision is made
29
30 using an alarm statistic and an alarm limit. In industrial quality control charts (e.g. xbar-
31
32 charts), an alarm can be given as soon as an observation crosses the alarm limit
33
34 (Shewhart (1931)). There is always a risk for a false alarm, but the parameters of the
35
36 surveillance method are chosen so that we know the false alarm property. For motivated
37
38 alarms, i.e. when the change actually happens, we want a quick detection. Since we have
39
40 repeated decisions, size and power are not appropriate measures. Instead we have a trade-
41
42 off between false alarms and the delay of a motivated alarm. The false alarms can be
43
44 controlled by a fixed average run length, ARL^0 . Besides the Shewhart method, the
45
46 EWMA method (Roberts (1959)) and the CUSUM method (Page (1954)) can be
47
48 mentioned.
49
50
51
52
53
54
55
56
57
58
59
60

In many situations we monitor several processes, which may change simultaneously or at different times. One example is business cycle monitoring, where changes in the import and export series probably can be seen before changes in unemployment, but where all these series measure the underlying business cycle. There are different approaches in multivariate monitoring: the data can be reduced to a scalar at each time or we can use separate alarm systems for each process in combination with an inference rule like union intersection. Also, multivariate versions of univariate methods have been suggested, for example MEWMA and MCUSUM. Important aspects are the correlation between the change times and the dependency between the processes themselves.

The aim of this paper is to compare different multivariate surveillance systems. In Section 3, optimality in surveillance is discussed and in Section 4 some of the different approaches to multivariate surveillance are reviewed. In Section 5 the results of a simulation study are presented and Section 6 contains a discussion.

2. Model

In this paper we study the situation when two processes may change at different time points (τ_X and τ_Y) and where the τ :s can be dependent.

The aim is to detect an increase in the expected value. With two series, we want to detect the first change (the first increase). The observations (X, Y) , up to decision time s , are modeled according to

$$\begin{pmatrix} m(1) \\ m(2) \\ \dots \\ m(s) \end{pmatrix} = \begin{pmatrix} x(1) & y(1) \\ x(2) & y(2) \\ \dots & \dots \\ x(s) & y(s) \end{pmatrix} = \begin{pmatrix} \mu_X(1) & \mu_Y(1) \\ \mu_X(2) & \mu_Y(2) \\ \dots & \dots \\ \mu_X(s) & \mu_Y(s) \end{pmatrix} + \begin{pmatrix} \varepsilon_X(1) & \varepsilon_Y(1) \\ \varepsilon_X(2) & \varepsilon_Y(2) \\ \dots & \dots \\ \varepsilon_X(s) & \varepsilon_Y(s) \end{pmatrix} \quad (1)$$

The stochastic term at time t , $(\varepsilon_X(t) \ \varepsilon_Y(t))$, follows a bivariate normal distribution, with expected value zero and covariance matrix

$$\sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad (2)$$

where ρ =correlation between $X(t)$ and $Y(t)$.

The variables $X(t)$ and $Y(t)$ are (possibly) dependent but not $X(t)$ and $X(t-j)$, $Y(t)$ and $Y(t-j)$ or $X(t)$ and $Y(t-j)$. **Surveillance of processes with autocorrelation is treated in Frisé and Sonesson (2005), Petzold, et al. (2004), Knoth and Schmid (2004) and Okhrin and Schmid (2007).** The aim is to detect the first change in either of the μ vectors (μ_X and μ_Y). One application might be detection of the start of an influenza epidemic (an increase from a constant base line). The Swedish institute for infectious disease control receives weekly information about confirmed cases (laboratory diagnosed) and suspected cases. The development is roughly captured through the following parametric model for μ , which will be used for both X and Y in a simulation study below.

$$\mu(t): \begin{cases} \mu^D(t): \beta_{01}, & t < \tau \\ \mu^{Cr}(t): \beta_{02} + \beta_1 \cdot (t+1-\tau) + \beta_2 \cdot (t+1-\tau)^3, & t \geq \tau \end{cases} \quad (3)$$

where $\beta_1 < 0$, $\beta_2 > 0$. Thus, for $\tau_X = \tau_Y$, X and Y have the same distribution and they are independent (or dependent) conditional on τ . The model above might be too crude in a real situation, but is used here to demonstrate the inferential aspects of multivariate surveillance.

A bivariate Geometric distribution is used for (τ_X, τ_Y) , with parameters (v_{01}, v_{10}, v_{11}) . The marginal distributions are $\tau_X \sim \text{Geo}(v_{1.} = v_{10} + v_{11})$ and $\tau_Y \sim \text{Geo}(v_{.1} = v_{01} + v_{11})$. The two marginal intensities are assumed to be equal ($v_{1.} = v_{.1} = 0.10$). The value of v_{11} will be

varied, thereby yielding different correlations, here $\{0, 0.80\}$. The correlation between τ_X and τ_Y is hence forward denoted ψ .

3. Optimality and evaluation of multivariate surveillance

Statistical surveillance is a methodology for discriminating between the two events, namely C="the change has occurred" and D="the change has not occurred". By an alarm system we decide whether D or C is the most likely.

Statistical surveillance is used for on-line detection of an important change in the underlying process, for instance a change in the expected value. For e.g. daily data, a new decision is made each day, based on the available data. When there is enough evidence of a change, an alarm is called. The alarm limit in the surveillance system is set so that the alarm system yields a specified false alarm risk.

In **univariate surveillance** we observe the process X at each decision time $s=1, 2, \dots$. At time s the vector of available observations is denoted $x_s=\{x(1),x(2),\dots,x(s)\}$. At an unknown time, τ , there is a change in the distribution of X , so that

$$X(s) \square \begin{cases} f(x(s)|D), & s < \tau \\ f(x(s)|C), & s \geq \tau \end{cases}$$

The events D and C are specified according to the change of interest (a simple case is a change in the mean, from μ_0 to μ_1). In a multivariate situation the aim is often to detect the first change in any of the p monitored processes, $\tau_{(1)}=\min[\tau_1,\dots,\tau_p]$.

The time of alarm, t_A , is defined as the first time that the alarm statistic exceeds the alarm limit. The alarm limit is set so that the false alarm property has a specified value. In quality control, it is common to use the average run length to the first false alarm, $ARL^0=E[t_A | \tau=\infty]$. Sometimes the median run length is used, MRL^0 , instead of the

1
2
3 **arithmetic mean.** In multivariate surveillance, the false discovery rate can be used (see
4
5 e.g. Wong, et al. (2003)). In many theoretical work (e.g. Shiryaev (1963)), the false alarm
6
7 probability is used, which summarize the false alarm distribution using the distribution of
8
9 τ , such that

$$10$$

$$11$$

$$12$$

$$13 \text{PFA} = \sum_{i=1}^{\infty} P(t_A < i) \cdot P(\tau = i).$$

$$14$$

$$15$$

$$16$$

17 For an on-line system, the ability to detect a change quickly is important, i.e. a short
18
19 delay for motivated alarms. For most surveillance methods, the delay of an alarm
20
21 depends on when the change does occur, in relation to the start of the surveillance. The
22
23 delay is often longest when the change occurs at the start ($\tau=1$). The conditional expected
24
25 delay of an alarm (see Frisén and Wessman (1999)) is defined as
26
27

$$28$$

$$29 \text{CED}(t) = E[t_A - \tau | t_A \geq \tau, \tau = t].$$

$$30$$

$$31$$

32 Many evaluations are made using $\tau=1$, e.g. $\text{CED}(1)$ which is equivalent to ARL^1-1 .
33
34 However it is important to consider other change point times also. **In multivariate**
35
36 **surveillance regarding the first change, $\tau_{(1)}$, the following delay measure has been**
37
38 **suggested (Wessman (1999))**

$$39$$

$$40 \text{CED}(t_1, t_2) = E[t_A - \tau_{(1)} | t_A \geq \tau_{(1)}, \tau_X = t_1, \tau_Y = t_2].$$

$$41$$

$$42$$

43
44 Optimal alarm systems are based on the full likelihood ratio between C and D (Shiryaev
45
46 (1963), Frisén and de Maré (1991)),

$$47$$

$$48 \frac{f(x_s | C)}{f(x_s | D)},$$

$$49$$

$$50$$

$$51$$

$$52$$

53 where $x_s = \{x(1), x(2), \dots, x(s)\}$. **The full likelihood ratio minimizes the expected delay for**
54
55 **a fixed false alarm probability (Frisén (2003)). Optimality is further discussed in Section**
56
57
58
59
60

4.1.3. If we want to detect a change at the current time s , then $C=\{\tau=s\}$, $D=\{\tau>s\}$. When the observations are independent and normally distributed, then the optimal full likelihood ratio reduces to the last observation, $x(s)$.

$$\frac{f(x_s|\tau=s)}{f(x_s|\tau>s)} = \frac{f(x(1)|\tau=s) \cdot f(x(2)|\tau=s) \cdot \dots \cdot f(x(s)|\tau=s)}{f(x(1)|\tau>s) \cdot f(x(2)|\tau>s) \cdot \dots \cdot f(x(s)|\tau>s)} = \frac{f(x(s)|\tau=s)}{f(x(s)|\tau>s)}.$$

This is the Shewhart method, used in e.g. an xbar-chart. Contrary to using only the latest observation, the CUSUM and EWMA methods cumulate the observations. CUSUM is the maximum of the likelihood ratios at the decision time s . The EWMA method uses an exponentially weighted moving average and for certain values of the smoothing constant, EWMA is approximately the same as LR (Frisén and Sonesson (2006)).

When it is important to detect if there has been a change since the start of the surveillance, we specify $C = \{\tau \leq s\} = \{\{\tau=1\}, \{\tau=2\}, \dots, \{\tau=s\}\}$ and $D = \{\tau > s\}$. Then the full likelihood ratio consists of s components

$$w_1 \frac{f(x_s|\tau=1)}{f(x_s|\tau>s)} + w_2 \frac{f(x_s|\tau=2)}{f(x_s|\tau>s)} + \dots + w_s \frac{f(x_s|\tau=s)}{f(x_s|\tau>s)} > k_s,$$

where $w_i = P(\tau=i)/P(\tau \leq s)$ and $k_s = k \cdot P(\tau \geq s)/P(\tau < s)$. Contrary to $C=\{\tau=s\}$ we now consider a composite event $C=\{\tau \leq s\}$ and the full likelihood ratio will include all observations $\{x(1), \dots, x(s)\}$. Here it is optimal to use all observations in the sense that the expected delay is minimized.

When $P(C) = 1 - P(D)$, the likelihood ratio is equivalent to using the posterior probability (alarm when $P(C|x_s) > k$). This is often used in hidden Markov model approaches (HMM), see e.g. Koskinen and Öller (2004).

4. Methods

4.1 Different approaches to multivariate surveillance

In a multivariate setting there are p processes $\{X_1, X_2, \dots, X_p\}$ for which

$$X_j(s) \square \begin{cases} f(x_j(s)|D), s < \tau_j \\ f(x_j(s)|C), s \geq \tau_j \end{cases} .$$

An overview of multivariate surveillance is given in Sonesson and Frisén (2005), who categorize different approaches to multivariate surveillance; reduction of dimensionality, reduction to one scalar statistic, parallel surveillance, vector accumulation and simultaneous solution. In this paper we compare reduction to scalar, parallel surveillance and simultaneous solution.

4.1.1 Reduction to one scalar statistic at each time

The p X -processes can be reduced to a scalar at each time, for example by a (weighted) mean. Wessman (1998) showed that when the processes have identical change times ($\tau_1 = \tau_2 = \dots = \tau_p = \tau$), there exists a sufficient univariate reduction of the variables $\{X_1, \dots, X_p\}$. Thus, without loss of information, the multivariate data can be reduced to a scalar statistic and then univariate surveillance can be applied. The sufficiency also holds when the changes times are not identical but the lag times between them are known.

When the changes occur at different time points, a problem with a reduction is to determine which variable that caused the alarm. Javaheri and Houshmand (2001) suggest a follow-up with discriminant analysis. Jolayemi (2000) constructs multiple control regions for two assignable causes. Mason, et al. (1995) decompose T_2 into independent components, each reflecting an individual variable X_j .

1
2
3 Wikström, et al. (1998) use principal component analysis to derive the most important
4 principal components and then univariate CUSUM and univariate EWMA is applied.
5
6

7
8 Abu-Shawiesh and Abdullah (2001) study surveillance of two correlated processes
9 where both scale and location change. The T2 statistic is based on robust estimates of
10 location and scale. This approach is compared to the ordinary T2. Mason, et al. (2003)
11 find that AR processes result in a U-shaped T2 curve.
12
13
14

15
16 Another reduction is to use the minimum and maximum values at each time, as is done
17 in Sepúlveda and Nachlas (1997). Kang and Albin (2000) model the variable Y as a
18 function of the variable X and monitor the slope and intercept by T2 statistic as well as
19 the residuals (deviations from reference line).
20
21
22
23
24
25

26
27 Aparisi, et al. (2001) reduce the covariance matrix through the determinant or the
28 trace. This scalar is monitored, using univariate Shewhart, EWMA and CUSUM.
29 Guerrero-Cusumano (1995) uses an entropy measure instead of the determinant.
30
31
32

33
34 Stoumbos and Jones (2000) reduce the multivariate data to a probability measure
35 which shows how “central” the observation is.
36
37

38
39 Lu, et al. (1998) study correlated variables and the proportion of non-conforming
40 units. The multivariate data are reduced to a scalar by weighting together the non-
41 conforming units and then monitored by univariate Shewhart.
42
43
44

45
46 Cheng and Liu (2000) use a rank measure for how outlying an observation is and then
47 the univariate rank variable is monitored using a Shewhart approach.
48
49

50
51 Koskinen and Öller (2004) suggest the use of a weighted index, which is monitored
52 using univariate methods. A similar approach is used in Talluri and Sarkis (2002).
53
54
55
56
57
58
59
60

1
2
3 Qiu and Hawkins (2001) use the anti-ranks of the p observations at each time and
4 summarize the anti ranks in a statistic which is monitored using univariate CUSUM.
5
6

7
8 There are several suggested methods which are a multivariate variation of a univariate
9 method, such as MEWMA and MCUSUM. Crosier (1988) and Pignatiello and Runger
10 (1990) used an MCUSUM on the same form as as the univariate CUSUM, only with
11 matrixes instead of scalars.
12
13

14
15
16
17 Bodden and Rigdon (1999) use a multivariate EWMA and smooth all p processes
18 using the same constant, λ , and then the vector of the smoothed values is reduced by the
19 T_2 statistic. This approach is also used in Love and Linderman (2003) and Molnau, et al.
20 (2001) and by Stoumbos and Sullivan (2002), who show that if the smoothing constant
21 equals 1 (no smoothing) and the process is not normally distributed, then the ARL^0 is
22 over estimated if the alarm limits are determined under normality assumptions. But for a
23 small smoothing constant, the ARL^0 is not so biased, even if the normality assumption is
24 violated. A small smoothing constant gives approximately equal weight to all
25 observations, which is approximately normal according to the central limit theorem.
26
27
28
29
30
31
32
33
34
35
36
37

38
39 Lowry, et al. (1992) use EWMA smoothing with separate λ values, and then reduction
40 by the T_2 statistic. The λ values are determined so as to minimize the ARL_1 . Yumin
41 (1996) suggests that if the X processes are correlated, they should be transformed into
42 principal components, which are then smoothed separately. Runger, et al. (1999) reduce
43 the dimension by a transform similar to principal component analysis, then the
44 transformations are smoothed using the same λ and then the T_2 statistic is calculated
45 from the smoothed series. Gan (1997) constructs a control chart with the smoothed
46 variance on one axis and the smoothed mean on the other, called a combined EWMA
47
48
49
50
51
52
53
54
55
56
57
58
59
60

chart. The advantage, to a T2 reduction, is that the chart shows whether it is the mean or the variance that is out-of control.

Hawkins (1991) applied a univariate CUSUM to a linear combination of the variables $\{X_1, X_2, \dots, X_p\}$. Wessman (1998) showed that if the p processes have identical change times, then data can, without loss of information, be reduced to a univariate index.

4.1.2 Parallel surveillance

The marginal density of each process can be monitored. One drawback is that no information about the dependency structure is used. The surveillance system for X_1 is only concerned with τ_1 , and correspondingly for X_2, X_3 etc. For process X_j at time s , we have alarm statistic $p(x_{js})$ and alarm limit k_s^j . An alarm is called when $p(x_{js}) > k_s^j$, where $x_{js} = \{x_j(1), x_j(2), \dots, x_j(s)\}$.

The time of alarm of X_j is

$$t_{Aj} = \min t: p(x_{jt}) > k_t^j.$$

The time of alarm for the whole system is defined as

$$t_A = \min \{t_{A1}, t_{A2}, \dots, t_{Ap}\}. \quad (4)$$

Does, et al. (1999) uses separate surveillance for of each process (each principal component) in a case study. In Woodall and Ncube (1985) a univariate CUSUM is used for each of the p processes.

4.1.3 Simultaneous solution

In a simultaneous solution, the joint distribution of the p processes can be used. Shiryaev (1963) showed that the likelihood ratio between C and D is optimal in the sense that is maximized the expected utility, $E[u]$, where u equals

$$u(t_A, \tau) = \begin{cases} h(t_A - \tau) & , t_A < \tau \\ a_1 \cdot (t_A - \tau) + a_2, & t_A \geq \tau \end{cases}$$

The optimality was proven for the situation when τ follows a Geometric distribution (with parameter ν). If the function $h(t_A - \tau)$ is a constant, b , the utility is maximized when the delay, $t_A - \tau$, is minimized (Frisén (2003)). The minimal expected delay was shown to hold also for a situation where τ is not Geometrically distributed (Andersson (2004)).

Here we study the situation when we want to detect the first change, $\tau_{(1)} = \min[\tau_1, \tau_2, \dots, \tau_p]$ and then the utility can be written as

$$u(t_A, \tau_{(1)}) = \begin{cases} h(t_A - \tau_{(1)}) & , t_A < \tau_{(1)} \\ a_1 \cdot (t_A - \tau_{(1)}) + a_2, & t_A \geq \tau_{(1)} \end{cases}$$

where t_A is the time of alarm for the whole system (e.g. (4)). The distribution of $\tau_{(1)}$ depends on the p -variate distribution for (τ_1, \dots, τ_p) . For $p=2$, we use the bivariate Geometric distribution, see e.g. Marshall and Olkin (1997). Then $\tau_{(1)}$ follows a Geometric distribution (Sun and Basu (1995)) and the optimality holds: the likelihood ratio maximizes the expected utility.

Wessman (1999) investigated the situation where $C = \{\tau_{(1)} = s\}$. In this paper we investigate $C = \{\tau_{(1)} \leq s\} = \{\{\tau_{(1)} = 1\}, \dots, \{\tau_{(1)} = s\}\}$. The D event is specified as $\{\tau_{(1)} > s\}$. Below the optimal surveillance system is derived from the likelihood ratio. Since $P(C) = 1 - P(D)$, the likelihood ratio is equivalent to the posterior probability

$$P(C | m_s) = P(m_s \cap C) / (P(m_s \cap C) + P(m_s \cap D)) > k,$$

where $m_s = \{m(1), \dots, m(s)\} = \{\{x_1(1), \dots, x_p(1)\}, \dots, \{x_1(s), \dots, x_p(s)\}\}$ and k is a constant limit.

Hence forward we express the likelihood function by $f(*)$ instead of $P(*)$, so that

$$\frac{f(m_s \cap C)}{f(m_s \cap D)} > \frac{k}{1-k}$$

For $C = \{\tau_{(1)} \leq s\}$ and $D = \{\tau_{(1)} > s\}$, the likelihood ratio equals

$$\sum_{t=1}^s w_s^t \cdot L(s, t) = w_s^1 \cdot \frac{f(m_s | \tau_{(1)} = 1)}{f(m_s | \tau_{(1)} > s)} + w_s^2 \cdot \frac{f(m_s | \tau_{(1)} = 2)}{f(m_s | \tau_{(1)} > s)} + \dots + w_s^s \cdot \frac{f(m_s | \tau_{(1)} = s)}{f(m_s | \tau_{(1)} > s)} > k_s.$$

The weights, $w_s^t = P(\tau_{(1)} = t) / P(\tau_{(1)} \leq s)$, and the limit, $k_s = k \cdot P(\tau_{(1)} > s) / P(\tau_{(1)} \leq s)$, depends on the distribution of $\tau_{(1)}$, which in turn depends on the joint distribution of (τ_1, \dots, τ_p) . The time of alarm is defined as

$$t_A = \min t: LR(m_s) > k_s.$$

4.2 Three suggested methods for multivariate surveillance

In this paper we compare results from three methods for multivariate surveillance; the T2 method (a reduction), a method for parallel surveillance and a simultaneous solution, derived from the optimal, full likelihood ratio.

4.2.1 The T2 method

An early multivariate method is the T2 method of Hotelling (1947). The statistic is assumed to have a known covariance matrix. The T2 is an example of a reduction. We assume that X and Y have the same distribution, conditional on τ_X and τ_Y . Thus an alarm is given when

$$T2(s) = \frac{(x(s) - \mu^D(s))^2}{(1 - \rho^2)\sigma^2} + \frac{(y(s) - \mu^D(s))^2}{(1 - \rho^2)\sigma^2} - \frac{2\rho(x(s) - \mu^D(s))(y(s) - \mu^D(s))}{(1 - \rho^2)\sigma^2} > k.$$

4.2.2 Parallel likelihood ratio systems

The parallel surveillance system in this paper, denoted LRpar, is based on separate likelihood ratio systems for X and Y. For X the likelihood ratio alarm statistic is

$$LR_X(s) = \sum_{j=1}^s \frac{P(\tau = j)}{P(\tau \leq s)} \exp \left(\frac{\sum_{t=1}^s (x(t) - \mu^D(t))^2}{2\sigma^2} - \frac{\sum_{t=1}^s (x(t) - \mu^{Cj}(t))^2}{2\sigma^2} \right) > k \frac{P(\tau_X > s)}{P(\tau_X \leq s)}$$

where μ^D and μ^C are defined in (3) and τ is assumed to follow a geometric distribution with intensity ν . The time of alarm for X is defined as

$$t_{AX} = \min[LR_X(s) > k^X_s]$$

where $k^X_s = k \cdot (P\tau_X > s) / P(\tau_X \leq s)$. The $LR_Y(s)$ statistic and t_{AY} are defined correspondingly.

The time of alarm for the LRpar system is the first time for which either alarm system gives an alarm, i.e. $t_A = \min\{t_{AX}, t_{AY}\}$. If X and Y are independent, the distribution of t_A is a direct function of the distributions of t_{AX} and t_{AY} .

The two alarm limits, k^X_s and k^Y_s , are adjusted to yield a specified false alarm property for the whole system. If there is no particular information regarding the cost for false alarms, then it is natural to have the same false alarm property for the two methods (e.g. equal ARL^0). If inspection and restoration costs are not the same for the processes, Serel, et al. (2000) suggest different type I errors.

4.2.3 Simultaneous solution using the joint likelihood ratio

Recall from Section 4.1.3 that the optimal surveillance method for $C = \{\tau_{(1)} \leq s\}$ consists of s weighted partial likelihood ratios

$$\sum_{t=1}^s w_s^t \cdot L(s, t),$$

$$\text{where } L(s, t) = \frac{f(m_s | \tau_{(1)} = t)}{f(m_s | \tau_{(1)} > s)}.$$

The event $\{\tau_{(1)} = t\}$ consists of three sub-events:

$\{\tau_X = t \cap \tau_Y > t\}$, event CD, “a change in X at time t”

$\{\tau_X > t \cap \tau_Y = t\}$, event DC, “a change in Y at time t”

$\{\tau_X = t \cap \tau_Y = t\}$, event CC, “change in both X and Y at time t”.

Thus $w_s^t \cdot L(s, t)$ is separated into $w_s^{t0} \cdot L(s, t, CD) + w_s^{0t} \cdot L(s, t, DC) + w_s^{tt} \cdot L(s, t, CC) =$

$$w_s^{t0} \cdot \frac{f(m_s | \tau_X = t, \tau_Y > t)}{f(m_s | \tau_X > t, \tau_Y > t)} + w_s^{0t} \cdot \frac{f(m_s | \tau_X > t, \tau_Y = t)}{f(m_s | \tau_X > t, \tau_Y > t)} + w_s^{tt} \cdot \frac{f(m_s | \tau_X = t, \tau_Y = t)}{f(m_s | \tau_X > t, \tau_Y > t)}.$$

Since we want to detect changes in the μ vectors, the expression above is written as

$$w_s^{t0} \frac{f(x_s, y_s | \mu^{Ct}, \mu^D)}{f(x_s, y_s | \mu^D, \mu^D)} + w_s^{0t} \frac{f(x_s, y_s | \mu^D, \mu^{Ct})}{f(x_s, y_s | \mu^D, \mu^D)} + w_s^{tt} \frac{f(x_s, y_s | \mu^{Ct}, \mu^{Ct})}{f(x_s, y_s | \mu^D, \mu^D)},$$

where $f(x_s, y_s)$ is the bivariate normal distribution in (1) and (2). When X and Y are independent X and Y, the expression above simplifies to

$$w_s^{t0} \frac{f(x_s | \tau_X = t)}{f(x_s | \tau_X > s)} + w_s^{0t} \frac{f(y_s | \tau_Y = t)}{f(y_s | \tau_Y > s)} + w_s^{tt} \frac{f(x_s | \tau_X = t)}{f(x_s | \tau_X > s)} \cdot \frac{f(y_s | \tau_Y = t)}{f(y_s | \tau_Y > s)},$$

where the first two components are those used in LRpar. The weights are

$$w_s^{t0} = \frac{P(\tau_X = t \cap \tau_Y > t)}{P(\tau_{(1)} \leq s)}, \quad w_s^{0t} = \frac{P(\tau_X > t \cap \tau_Y = t)}{P(\tau_{(1)} \leq s)}, \quad w_s^{tt} = \frac{P(\tau_1 = t \cap \tau_2 = t)}{P(\tau_{(1)} \leq s)},$$

where the bivariate distribution of (τ_X, τ_Y) is governed by the intensity parameters $v_{1.}$, $v_{.1}$

and v_{11} . The information regarding (τ_X, τ_Y) is also included in the alarm limit,

$k_s = k \cdot P(\tau_{(1)} > s) / P(\tau_{(1)} \leq s)$, see Section 4.1.3.

4.2.4 Alarm regions for the three methods

The methods T2, LRpar and LRjoint are made comparable by adjusting their respective alarm limits so that the probability of false alarm, PFA, equals 0.1, where

$$\text{PFA} = \sum_{i=1}^{\infty} P(t_A < i) \cdot P(\tau_{(1)} = i).$$

The alarm regions are illustrated in Figure 1, for different values of ψ (correlation between the change times) and different values of ρ (dependency between X and Y).

<FIGURE 1>

For LRpar, an alarm is called if at least one of X or Y deviates largely from the D-state. For LRjoint (when X and Y are independent), an alarm is called when either X or Y or (X+Y) deviates from the D-state. The same holds for T2. For both LRjoint and T2, the shape of the alarm region changes as ρ and ψ changes, but not for LRpar.

5. Results

The methods (T2, LRpar and LRjoint) are made comparable by all having PFA=0.1, where PFA is summarized using the bivariate Geometric distribution that results in a specified ψ (correlation between τ_X and τ_Y).

LRpar and LRjoint have similar run length distributions (thus similar **median run length**, MRL^0), whereas T2 has a higher MRL^0 , see Figure 2. For all methods, the run length distribution is similar for $\rho=0$ and $\rho=0.5$ (dependency between X and Y), but very different for $\psi=0$ versus $\psi=0.8$ (correlation between τ_X and τ_Y). The density of $\tau_{(1)}$ is used to weight together the alarm times in PFA. As ψ tends to 1, the alarms tend to be more uniformly distributed and MRL^0 is longer. For both T2 and LRpar, the alarm limit

needs to be changed in order to maintain PFA=0.10 when a different density for $\tau_{(1)}$ is used. For LRjoint, the alarm statistic itself is changed when $\psi=0.8$, since the components CD, DC and CC are weighted differently.

<FIGURE 2>

5.1 Motivated alarms: the first change occurs immediately

CED(t_1, t_2) is presented for $\tau_{(1)}=1$, i.e. when the first change occurs immediately.

<FIGURE 3>

5.1.1 The effect of correlation between the change times

The delay is shortest when both changes occur immediately ($\tau_X=\tau_Y=1$), independent of ψ .

When independent change times are assumed ($\psi=0$), T2 has the shortest delay, but the three methods give similar delay. For the two likelihood ratio methods, the similarity was indicated by the similar alarm regions in Figure 1. At ($\tau_X = \tau_Y = 1$), LRjoint has slightly shorter delay, compared to LRpar. LRjoint consists of three components (CD, DC, CC, see section 4.2.2), whereas LRpar is based on two components (CD and DC).

When a positive correlation is assumed between the change times ($\psi=0.8$), LRjoint is superior at ($\tau_X = \tau_Y = 1$). LRjoint assigns a large weight to the CC component and when the changes actually occur simultaneously, LRjoint has a short delay.

5.1.2 The effect of the X-processes under surveillance being dependent

For immediate changes in both processes ($\tau_X = \tau_Y = 1$), the T2 has the shortest delay and LRpar the longest, when X and Y are independent ($\rho=0$). However, for $\rho=0.5$, LRpar has shortest delay and LRjoint longest. This is further discussed in 4.2.2.

5.2 Motivated alarms: the effect of the change times

The delay $CED(t_1, t_2)$ is presented for $t_1 = \{1, 5, 10\}$ and $t_2 = 1, 2, \dots, 14$.

<FIGURE 4a>

<FIGURE 4b>

<FIGURE 4c>

5.2.1 The effect of correlation between the change times

For all methods, CED is longer when $\psi > 0$. For LRpar, the delay is generally longer as a result of the higher alarm limit. The same holds for T2. Also for LRjoint, the delay is generally longer, but there is also a larger difference between $CED(t, t)$ and $CED(t, \infty)$, exemplified by $CED(5, 5) = 3.5$ and $CED(5, 14) = 4.2$. For $\psi > 0$ the LRjoint works well when the changes do actually occur simultaneously.

Both for $\psi = 0$ and $\psi = 0.8$, simultaneous changes have shortest delay for all method, i.e. $CED(t, t) < CED(t, j)$, $j \neq t$. For $CED(t, t)$, both X and Y are in C-state at each time point.

For the two LR methods, later changes have shortest delay (e.g. $CED(1, 1) > CED(5, 5)$), whereas T2 has the same delay for e.g. $(\tau_X = \tau_Y = 1)$ as for $(\tau_X = \tau_Y = 5)$.

As one change time tends to infinity, the CED-curve depends only on the other change time ($CED(\infty, t)$ depends only on t).

5.2.2 The effect of the processes under surveillance being dependent

For LRpar the CED-curves are practically the same for $\rho = 0$ and $\rho = 0.5$.

For LRjoint, the simultaneous change situation, $\tau_X = \tau_Y$, is quicker detected if the processes are independent. The same holds for T2.

For T2, the CED-curve only depends on the distance $(t_1 - t_2)$, see Andersson (2008).

Both for $\rho=0$ and for $\rho=0.5$, T2 has longer CED than both the LR methods, in most situations.

For $\rho=0$, LRpar and LRjoint have similar CED-curves. For $\rho=0.5$, LRpar has a shorter delay for simultaneous changes ($\tau_X = \tau_Y$), compared to LRjoint, Figure 4a-4c. One reason for the long delay of LRjoint is the assumption that $\psi=0$. Another was pointed out in Wessman (1999) who investigated multivariate surveillance for the situation $C=\{\tau_{(1)}=s\}$ and showed that if changes occur in all processes, the probability of detecting it is lower if the observations are highly correlated.

For $\rho=0$, the $CED(t,t)$ is much smaller than $CED(t,\infty)$ for all three methods. But for $\rho=0.5$, when we use LRjoint or T2, the $CED(t,t)$ is approximately the same as $CED(t,\infty)$, exemplified for LRjoint by $CED(5,5)=3.3$ and $CED(5,14)=3.3$.

When $\rho=0.5$, both LRjoint and T2 yield long delay for “almost simultaneous changes”, so that $CED(t, |t-1|) > CED(t,\infty)$, see Figure 4b and 4c (exemplified for LRjoint by $CED(5,6)=3.5$ and $CED(5,14)=3.3$). Consider T2, based on the Mahalanobis distance $M_{\tau_X, \tau_Y}(t) = (X'(t))^2 + (Y'(t))^2 - 2\rho(X'(t))(Y'(t))$. In a simple shift situation (0 to θ) with simultaneous changes, we have

$$E[M_{\tau_X, \tau_Y}(t)] = \theta^2(2 - 2\rho), \text{ for all values of } t.$$

When $\tau_X \neq \tau_Y$, we have

$$E[M_{\tau_X, \tau_Y}(t)] = \theta^2, \text{ for } \tau_{(1)} \leq t < \tau_{(2)}$$

$$E[M_{\tau_X, \tau_Y}(t)] = \theta^2(2 - 2\rho), \text{ for } t \geq \tau_{(2)}$$

1
2
3 This makes $E[M_{i,i}(t)]$ smaller than (or equal to) $E[M_{i,j}(t)]$, when ρ is close to 1,
4
5 indicating longer delay for simultaneous changes. In this simulation study, however, the
6
7 change is from a constant level to an increasing function, making the expressions more
8
9 complex. For ρ close to 1, we again find smaller values for $E[M_{i,i}(t)]$ than for $E[M_{i,j}(t)]$,
10
11 indicating a long delay for simultaneous changes (and vice versa for ρ close to 0). But for
12
13 $\rho=0.5$ the values of $E[M_{i,j}(t)]$ for different t are not necessarily monotone as a function
14
15 of $\tau_X - \tau_Y$ and this causes the non-monotonic CED-function (for T2 we have $CED(10,5) =$
16
17 4.1 , $CED(10,9)=4.3$, $CED(10,10)=4.0$).
18
19
20
21
22
23

24 6. Discussion

25
26 Warning systems are used in many areas: public health, bio terrorism, radiation,
27
28 pregnancy, intensive care patients. Very often, several processes can be used to detect an
29
30 underlying change. Then we need a warning (or surveillance) system for multivariate
31
32 data. When monitoring more than one process, we must consider the dependency
33
34 between the processes and the correlation between the change times. **Surveillance of**
35
36 **autocorrelated processes is an important area, though not treated in this paper.**
37
38
39

40
41 Three methods of multivariate surveillance are compared. In the Hotellings T2 the
42
43 multivariate data is reduced to a scalar at each time point. In the LRpar method, the
44
45 likelihood ratio method is applied to each marginal process. The LRjoint method is
46
47 optimal according to the Shiryaev criterion and is derived from the joint likelihood ratio.
48
49

50
51 The alarm limits of each of the three systems are adjusted to yield the same false alarm
52
53 probability.
54

55
56 We investigate the effect of a positive correlation between the change times (the
57
58 correlation between τ_X and τ_Y is denoted ψ). We also investigate the effect of a positive
59
60

1
2
3 covariance between the processes (X and Y), conditional on τ_X and τ_Y (the covariance is
4 denoted ρ). The LRjoint is the only method that uses the information of ψ in the alarm
5
6
7
8
9 statistic. Both T2 and LRjoint incorporate the information of ρ .

10
11 The evaluation is made using the delay of an alarm, in relation to the first change time.

12
13 The methods are first compared for independent change times ($\psi=0$). For *immediate*
14
15 *changes in both processes* ($\tau_X=\tau_Y=1$), T2 has the shortest delay, followed by LRjoint and
16
17 then LRpar. T2 allocates the alarms early, whereas the two LR methods have few early
18
19 alarms. LRjoint has a slightly shorter delay than LRpar, since LRjoint includes three
20
21 components, corresponding to a change in either X or Y or both, whereas LRpar includes
22
23 only two components (change in either X or Y) which is not optimal for simultaneous
24
25 changes. However, the difference is small since $\psi=0$ (we do not expect many
26
27 simultaneous changes). For *later simultaneous changes* ($\tau_X=\tau_Y$), LRjoint is slightly better
28
29 than LRpar, whereas T2 here yields a long delay. T2 allocates the alarms early and also
30
31 T2 uses only observations from the current time point, thus it is not based on more data at
32
33 later time points (as opposed to the LR methods). The T2 use of only the observations at
34
35 the current time point corresponds to the Shewhart method in univariate surveillance. T2
36
37 is a reduction, which is sufficient at simultaneous change times, but the method is not
38
39 always optimal for $\tau_X=\tau_Y$. For *different change times* ($\tau_X\neq\tau_Y$), LRpar has the same delay
40
41 as LRjoint and T2 has the longest delay, because of the alarm allocation.

42
43
44
45
46
47
48
49
50 Next we investigate the effect of a positive correlation between the change times
51
52 ($\psi=0.8$). For $\tau_X=\tau_Y=1$, LRjoint has the shortest delay, followed by T2 and then LRpar.
53
54 LRjoint uses the information that the change times are likely to occur simultaneously and
55
56 T2 benefits from allocating the alarm early. For $\tau_X=\tau_Y=t$, $t>1$, LRjoint has the shortest
57
58
59
60

1
2
3 delay, followed by LRpar and then T2. As for $\psi=0$, T2 only works well for early
4 simultaneous changes. For $\tau_X \neq \tau_Y$, LRpar gives slightly shorter delay than LRjoint, and
5
6 T2 gives the longest delay. LRjoint uses the information that $\psi>0$, which is a
7
8 disadvantage when the changes actually do occur at different times.
9
10

11
12
13 We also compare the methods in the situation when X and Y have a positive
14 covariance ($\rho=0.5$). For $\tau_X=\tau_Y=1$, LRpar has the shortest delay, followed by T2 and then
15 LRjoint. For $\tau_X=\tau_Y=t, t>1$, LRpar gives shorter delay than LRjoint (although LRjoint now
16 works better than T2). For LRjoint, one reason for the long delay for $\tau_X = \tau_Y$ is the
17 assumption of $\psi=0$, which results in a low weight for the “simultaneous change
18 component”. Both T2 and LRjoint do include the components of the Mahalanobis
19 distance, which tend to be smaller for simultaneous changes when ρ is large, see
20
21 Wessman (1999). If X and Y are independent and change simultaneously, this is a
22 stronger indication than if they change simultaneously when they are positively
23 dependent and thus the delay is shorter in the independent situation. The values of the
24 Mahalanobis distance depends on the type of change (here we have a change from a
25 constant level to an increasing function) as well as ρ . In a simple situation with a shift,
26 the Mahalanobis distance (M) is a monotone function of the distance between the change
27 times: for ρ close to 0, the delay will be shortest for simultaneous changes and for ρ close
28 to 1 the delay will be longest for simultaneous changes. This also holds for a more
29 complex change, but for $\rho=0.5$ (which was investigated here) the association is not so
30 clear-cut, which results in a non-monotonic delay curve.
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

54 In this paper we only deal with positive dependency. A negative correlation between
55 the change times, $\psi<0$, implies that the change times do not coincide. Then it would be
56
57
58
59
60

1
2
3 best to have a very small weight for the component for simultaneous changes. If the
4
5 processes themselves have a negative covariance ($\rho < 0$), the implication is that $X = -Y$.
6
7 Then the alarm region would constitute of large values of $(X+Y)$ and the alarm region for
8
9 $\rho < 0$ would be similar to that of a positive correlation between the change times ($\psi > 0$).
10
11
12
13

14 **References**

- 15 Abu-Shawiesh M;Abdullah M. A new robust bivariate control chart for location.
16
17 Communications in Statistics-Simulation and Computation, **2001**, 30, (3), 513-530.
18
19 Andersson E. The impact of intensity in surveillance of cyclical processes.
20
21 Communications in Statistics-Simulation and Computation, **2004**, 33, (4), 889-913.
22
23 Andersson E. Hotelling's T2 method in multivariate on-line surveillance. On the delay of
24
25 an alarm. Research report: Statistical Research Unit, Göteborg University, Sweden,
26
27 **2008**.
- 28 Andersson E;Bock D;Frisén M. Detection of turning points in business cycles. Journal of
29
30 Business Cycle Measurement and Analysis, **2004**, 1, (1), 93-108.
- 31 Aparisi F;Jabaloyes J;Carrión A. Generalized variance chart design with adaptive sample
32
33 sizes. The bivariate case. Communications in Statistics-Simulation and Computation,
34
35 **2001**, 30, (4), 931-949.
- 36
37 Baron M. Bayes and asymptotically pointwise optimal stopping rules for the detection of
38
39 influenza epidemics. *Case Studies in Bayesian Statistics*, vol 6; Gatsonis C, Kass RE,
40
41 Carriquiry A, et al., Eds.; Springer-Verlag: New York, **2002**; 153-163.
- 42 Bock D. Aspects on the control of false alarms in statistical surveillance and the impact
43
44 on the return of financial decision systems. Journal of Applied Statistics, **2008**, 35,
45
46 (2), 213 - 227.
- 47 Bodden KM;Rigdon SE. A program for approximating the in-control ARL for the
48
49 MEWMA chart. Journal of Quality Technology, **1999**, 31, (1), 120-123.
- 50 Cheng AY;Liu RY. Monitoring multivariate aviation safety data by data depth: Control
51
52 charts and threshold systems. IIE Transactions, **2000**, 32, (9), 861-872.
- 53
54 Crosier RB. Multivariate Generalizations of Cumulative Sum Quality-Control Schemes.
55
56 Technometrics, **1988**, 30, 291-303.
57
58
59
60

- 1
2
3 Does RJMM;Roes KCB;Trip A. Handling multivariate problems with univariate control
4 charts. *Journal of Chemometrics*, **1999**, *13*, (3-4), 353-369.
- 5
6
7 Frisén M. Statistical surveillance. Optimality and methods. *International Statistical*
8
9 *Review*, **2003**, *71*, (2), 403-434.
- 10
11 Frisén M. Financial surveillance. Wiley, **2007** 272.
- 12
13 Frisén M;de Maré J. Optimal Surveillance. *Biometrika*, **1991**, *78*, 271-80.
- 14
15 Frisén M;Sonesson C. Optimal surveillance. *Spatial surveillance for public health*
16
17 Lawson A, Kleinman K, Eds.; Wiley, **2005**; 31-57.
- 18
19 Frisén M;Sonesson C. Optimal surveillance based on exponentially weighted moving
20
21 averages. *Sequential analysis*, **2006**, *25*, (4), 379-403.
- 22
23 Frisén M;Wessman P. Evaluations of likelihood ratio methods for surveillance.
24
25 Differences and robustness. *Communications in Statistics-Simulations and*
26
27 *Computations*, **1999**, *28*, (3), 597-622.
- 28
29 Gan FF. Joint monitoring of process mean and variance. *Nonlinear Analysis*, **1997**, *30*,
30
31 (7), 4017-4024.
- 32
33 Guerrero-Cusumano J-L. Testing variability in multivariate quality control: A conditional
34
35 entropy measure approach. *Information Sciences*, **1995**, *86*, (1-3), 179-202.
- 36
37 Hamilton JD. A new approach to the economic analysis of nonstationary time series and
38
39 the business cycle. *Econometrica*, **1989**, *57*, 357-384.
- 40
41 Hawkins DM. Multivariate Quality Control Based on Regression-Adjusted Variables.
42
43 *Technometrics*, **1991**, *33*, 61-75.
- 44
45 Hotelling H. Multivariate Quality Control. *Techniques of statistical analysis* Eisenhart
46
47 C, Hastay MW, Wallis WA, Eds.; McGraw-Hill: New York, **1947**.
- 48
49 Javaheri A;Houshmand AA. Average run length comparison of multivariate control
50
51 charts. *Journal of Statistical Computation and Simulation*, **2001**, *69*, (2), 125-140.
- 52
53 Jolayemi JK. A model for the statistical design of multivariate control charts with
54
55 multiple control regions. *Applied Mathematics and Computation*, **2000**, *109*, (1), 73-
56
57 91.
- 58
59 Järpe E. Surveillance of spatial patterns. *Communications in Statistics. Theory and*
60
61 *Methods*, **1999**, *28*, 3009-3025.

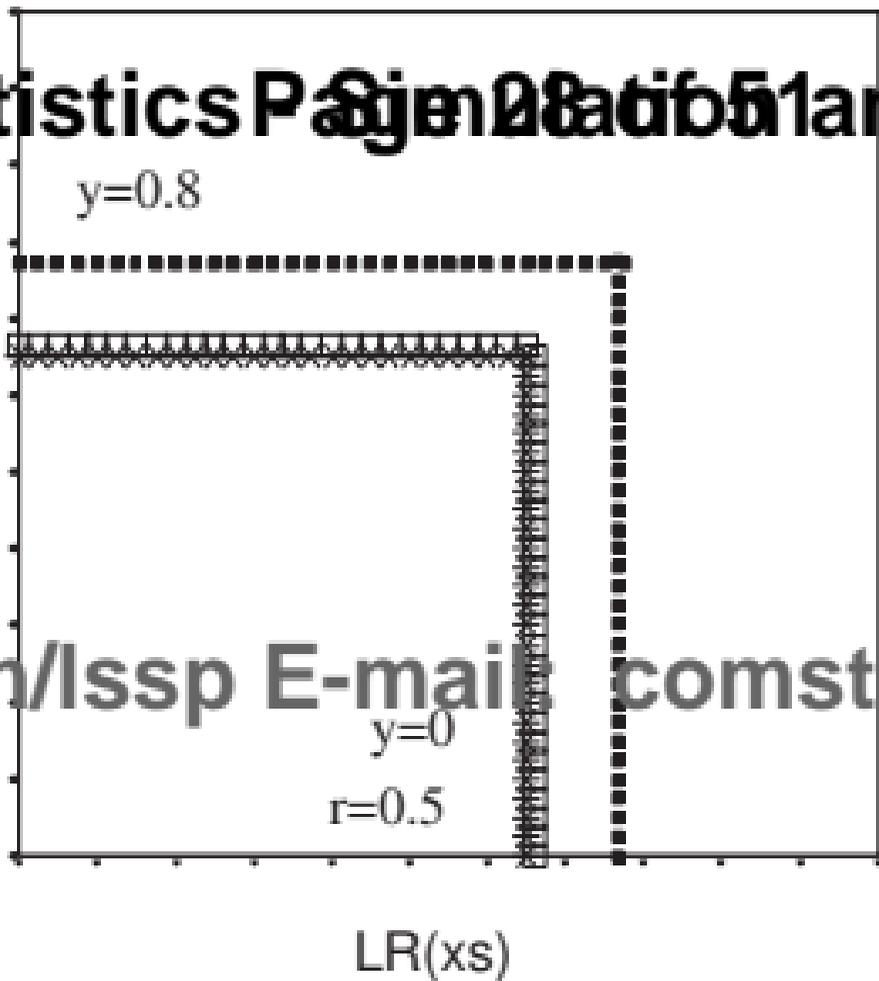
- 1
2
3 Kang L;Albin SL. On-line monitoring when the process yields a linear profile. Journal of
4 Quality Technology, **2000**, 32, (4), 418-426.
5
6
7 Knoth S;Schmid W. Control charts for time series: a review. *Frontiers in statistical*
8 *quality control*, vol 7; Lenz H-J, Wilrich P-T, Eds.; Physica: Heidelberg, **2004**; 210-
9 236.
10
11
12 Koskinen L;Öller L-E. A Classifying Procedure for Signalling Turning Points. Journal of
13 Forecasting, **2004**, 23, (3), 197-214.
14
15
16 Love TE;Linderman K. A Weibull process failure mechanism for the economic design of
17 MEWMA control charts. Journal of Statistical Computation and Simulation, **2003**,
18 73, (3), 195-202.
19
20
21 Lowry CA;Woodall WH;Champ CW;Rigdon SE. A multivariate exponentially weighted
22 moving average control chart. *Technometrics*, **1992**, 34, 46-53.
23
24
25 Lu XS;Xie M;Goh TN;Lai CD. Control chart for multivariate attribute processes.
26 *International Journal of Production Research*, **1998**, 36, (12), 3477-3489.
27
28
29 Löfqvist C;Andersson E;Sigurdsson J, et al. Longitudinal Postnatal Weight and Insulin-
30 like Growth Factor I Measurements in the Prediction of Retinopathy of Prematurity.
31 *Archives of Ophthalmology*, **2006**, 124, 1711-1718.
32
33
34 Marshall AW;Olkin I. A new method for adding a parameter to a family of distributions
35 with application to the exponential and Weibull families. *Biometrika*, **1997**, 84, (3),
36 641-652.
37
38
39 Mason RL;Chou YM;Sullivan JH;Stoumbos ZG;Young JC. Systematic patterns in T-2
40 charts. Journal of Quality Technology, **2003**, 35, (1), 47-58.
41
42
43 Mason RL;Tracy ND;Young JC. Decomposition of T2 for multivariate control chart
44 interpretation. Journal of Quality Technology, **1995**, 27, (2), 99-108.
45
46
47 Molnau WE;Runger GC;Montgomery DC;Skinner KR;Loredo EN;Prabhu SS. A
48 program for ARL calculation for multivariate EWMA charts. Journal of Quality
49 Technology, **2001**, 33, (4), 515-521.
50
51
52 Neftci S. Optimal prediction of cyclical downturns. Journal of Economic Dynamics and
53 Control, **1982**, 4, 225-41.
54
55
56 Okhrin Y;Schmid W. Surveillance of univariate and multivariate linear time series.
57 *Financial surveillance* Frisén M, Ed. Wiley: Chichester, **2007**.
58
59
60

- 1
2
3 Page ES. Continuous inspection schemes. *Biometrika*, **1954**, *41*, 100-114.
4
5 Petzold M;Sonesson C;Bergman E;Kieler H. Surveillance in longitudinal models.
6
7 Detection of intra-uterine growth restriction. *Biometrics*, **2004**, *60*, (4), 1025-1033.
8
9 Pignatiello JJ;Runger GC. Comparisons of Multivariate CUSUM Charts. *Journal of*
10
11 *Quality Technology*, **1990**, *22*, 173-186.
12
13 Qiu PH;Hawkins D. A rank-based multivariate CUSUM procedure. *Technometrics*,
14
15 **2001**, *43*, (2), 120-132.
16
17 Roberts SW. Control Chart Tests Based on Geometric Moving Averages. *Technometrics*,
18
19 **1959**, *1*, 239-250.
20
21 Royston P. Identifying the fertile phase of the human menstrual cycle. *Statistics in*
22
23 *Medicine*, **1991**, *10*, 221-240.
24
25 Runger GC;Keats JB;Montgomery DC;Scranton RD. Improving the performance of the
26
27 multivariate exponentially weighted moving average control chart. *Quality and*
28
29 *Reliability Engineering Internato*, **1999**, *15*, (3), 161-166.
30
31 Sepúlveda A;Nachlas JA. A simulation approach to multivariate quality control.
32
33 *Computers & Industrial Engineering*, **1997**, *33*, (1-2), 113-116.
34
35 Serel DA;Moskowitz H;Tang J. Univariate (X)over-bar control charts for individual
36
37 characteristics in a multinormal model. *Iie Transactions*, **2000**, *32*, (12), 1115-1125.
38
39 Shewhart WA. *Economic Control of Quality of Manufactured Product*. MacMillan and
40
41 Co., **1931**.
42
43 Shiryaev AN. On optimum methods in quickest detection problems. *Theory of*
44
45 *Probability and its Applications.*, **1963**, *8*, 22-46.
46
47 Sonesson C;Frisén M. Multivariate surveillance. *Spatial surveillance for public health*
48
49 *Lawson A, Kleinman K, Eds.; Wiley*, **2005**.
50
51 Stoumbos ZG;Jones LA. On the properties and design of individuals control charts based
52
53 on simplicial depth. *Nonlinear Studies*, **2000**, *7*, (2), 147-178.
54
55 Stoumbos ZG;Sullivan JH. Robustness to non-normality of the multivariate EWMA
56
57 control chart. *Journal of Quality Technology*, **2002**, *34*, (3), 260-276.
58
59 Sun K;Basu AP. A characterization of a bivariate geometric distribution. *Statistics &*
60
Probability Letters, **1995**, *23*, 307-311.

- 1
2
3 Talluri S;Sarkis J. A methodology for monitoring system performance. International
4 Journal of Production Research, **2002**, *40*, (7), 1567-1582.
5
6
7 Wessman P. Some Principles for surveillance adopted for multivariate processes with a
8 common change point. Communications in Statistics. Theory and Methods, **1998**, *27*,
9 1143-1161.
10
11
12 Wessman P. The surveillance of several processes with different change points. Research
13 report # 1999:2: Department of Statistics, Göteborg University, Sweden, **1999**.
14
15
16 Wikström C;Albano C;Eriksson L, et al. Multivariate process and quality monitoring
17 applied to an electrolysis process: Part I. Process supervision with multivariate
18 control charts. Chemometrics and Intelligent Laboratory Systems, **1998**, *42*, (1-2),
19 221-231.
20
21
22
23 Wong WK;Moore A;Cooper G;Wagner M. WSARE: What's Strange about Recent
24 Events? Journal of Urban Health, **2003**, *80*, i66-i75.
25
26
27 Woodall WH;Ncube MM. Multivariate Cusum Quality Control Procedures.
28 Technometrics, **1985**, *27*, 285-292.
29
30 Yumin L. An improvement for mewma in multivariate process control*1. Computers &
31 Industrial Engineering, **1996**, *31*, (3-4), 779-781.
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ns in Statistics Page 10 of 11

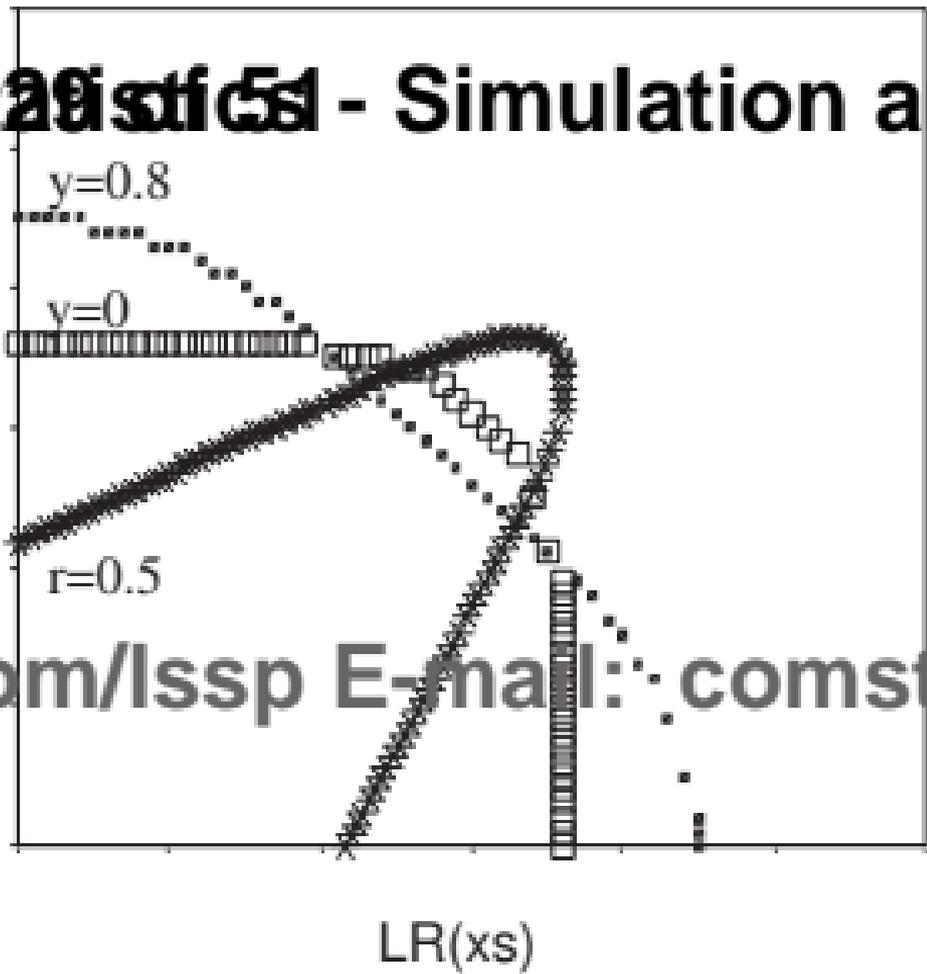
1
2
3
4
5
6



central.com/lssp E-mail: comsta

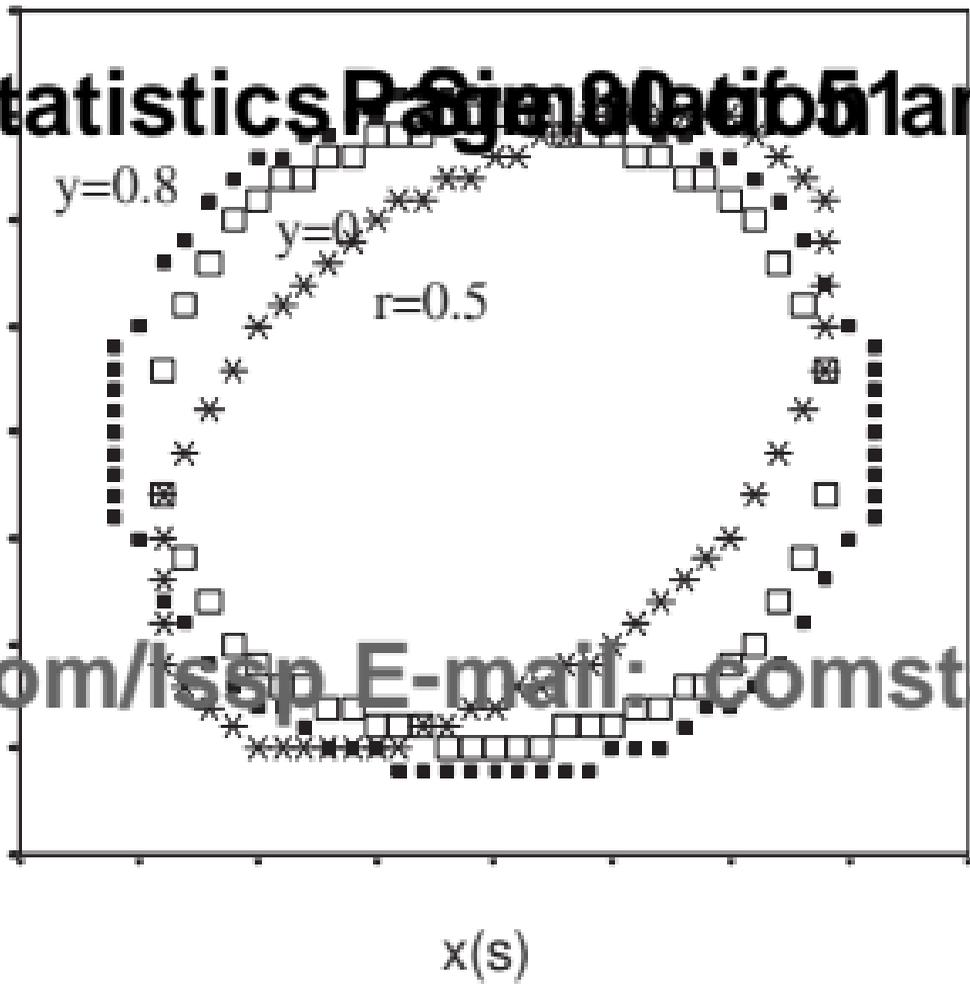
ns Ping Statistics - Simulation and

1
2 $N_{LR}(ys)$
3
4
5
6

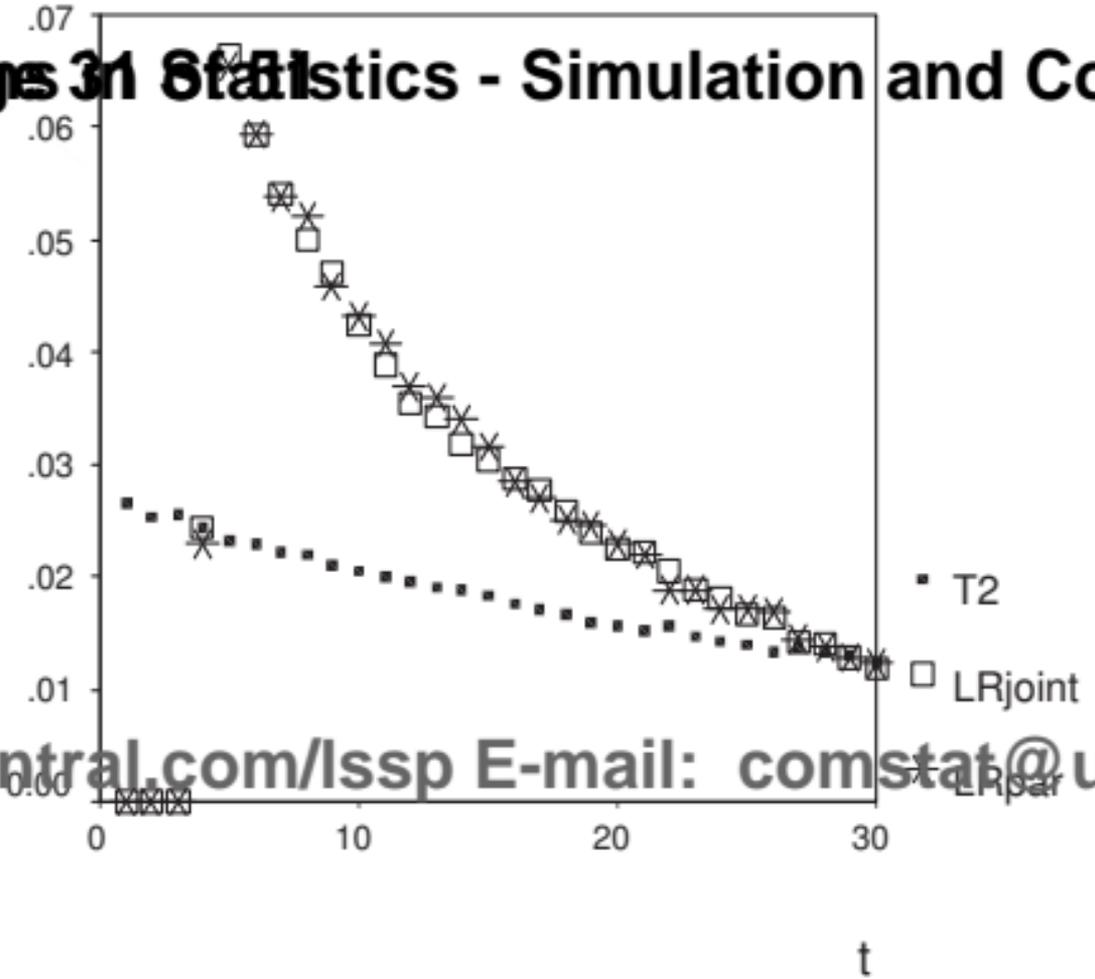


entral.com/lssp E-mail: .comsta

1
2
3
4
5
6



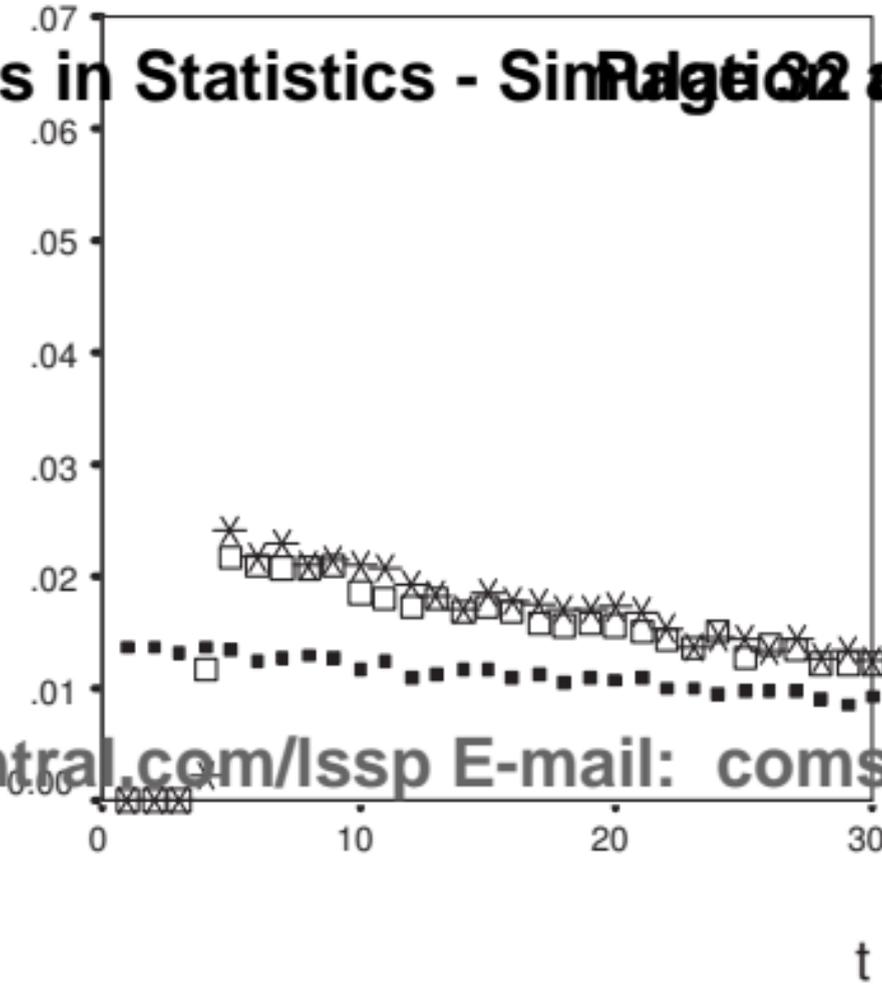
1
2
3
4
5
6
7
8
9
False alarm density



Applications in Statistics - Simulation of 5 Co

1
2
3
4
5
6
7
8
9

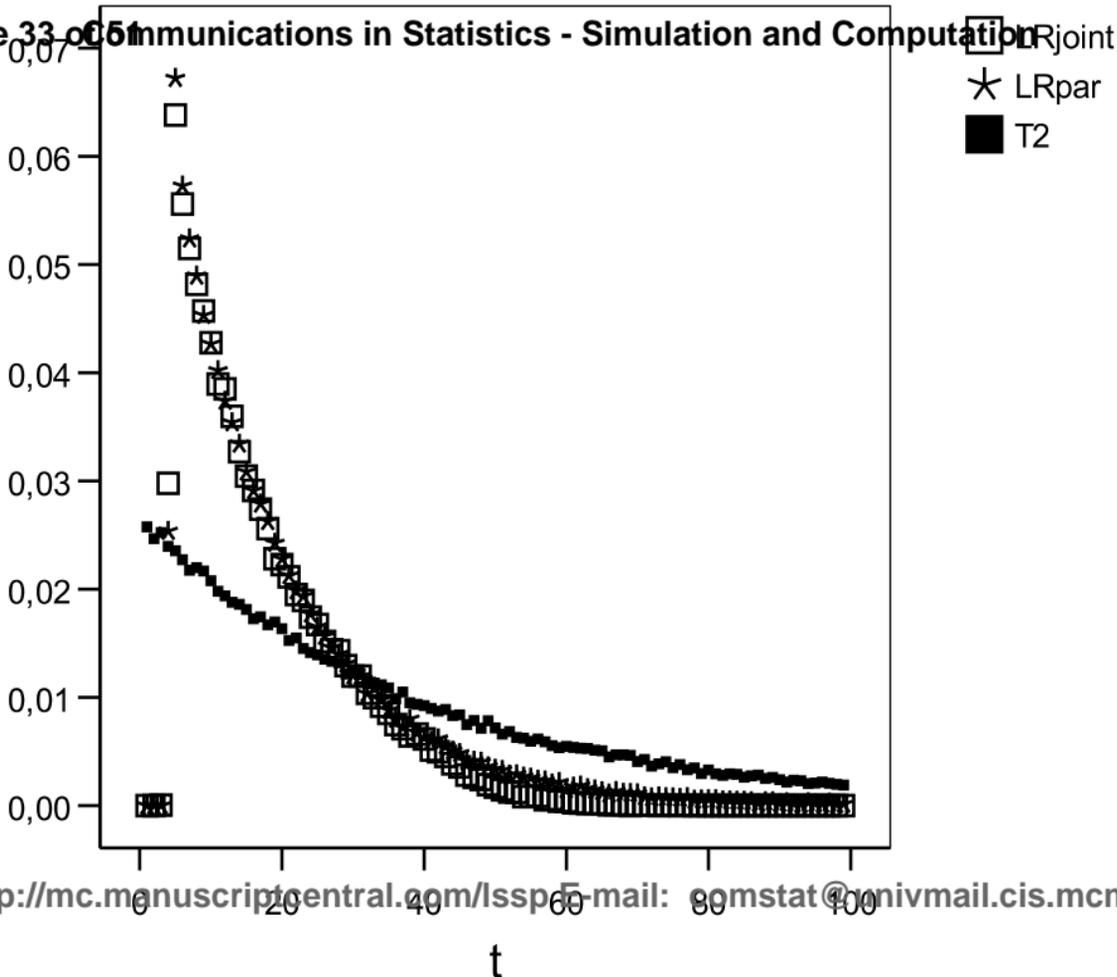
False alarm density



■ T2
□ LRjoint
× LRpar

ptcentral.com/lssp E-mail: comstat@u

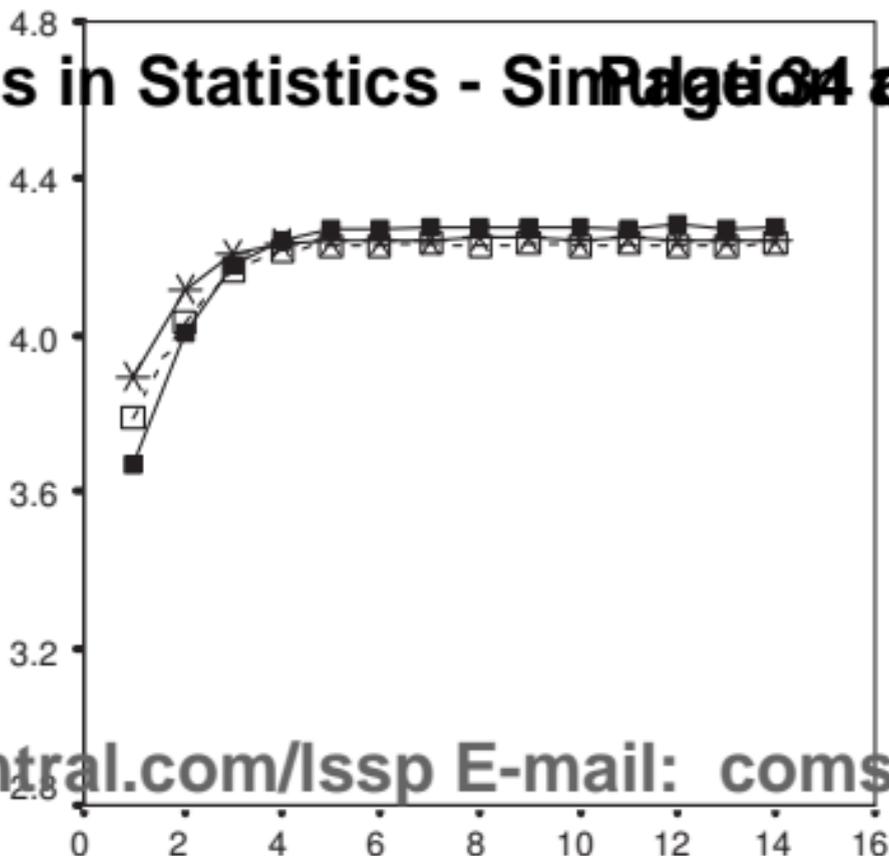
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22



Comparisons in Statistics - Simulation of 5 Co

1
2
3
4
5
6
7
8
9

CED($t_1=1, t_2$)

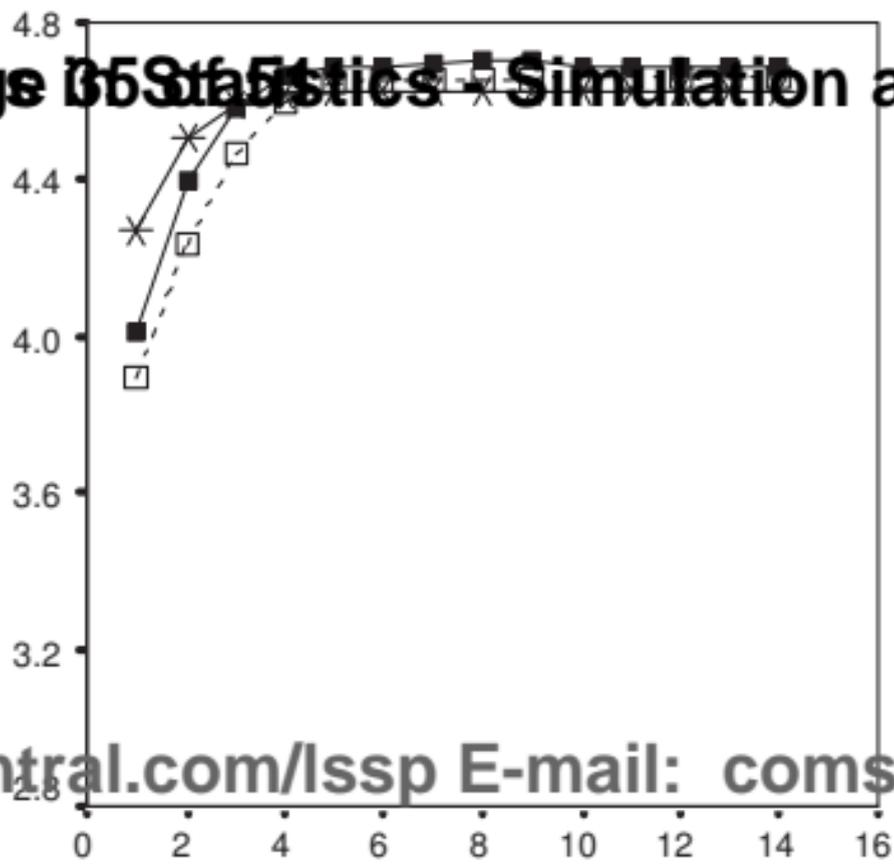


ptcentral.com/lssp E-mail: comstat@u

t_2

1
2
3
4
5
6
7
8
9

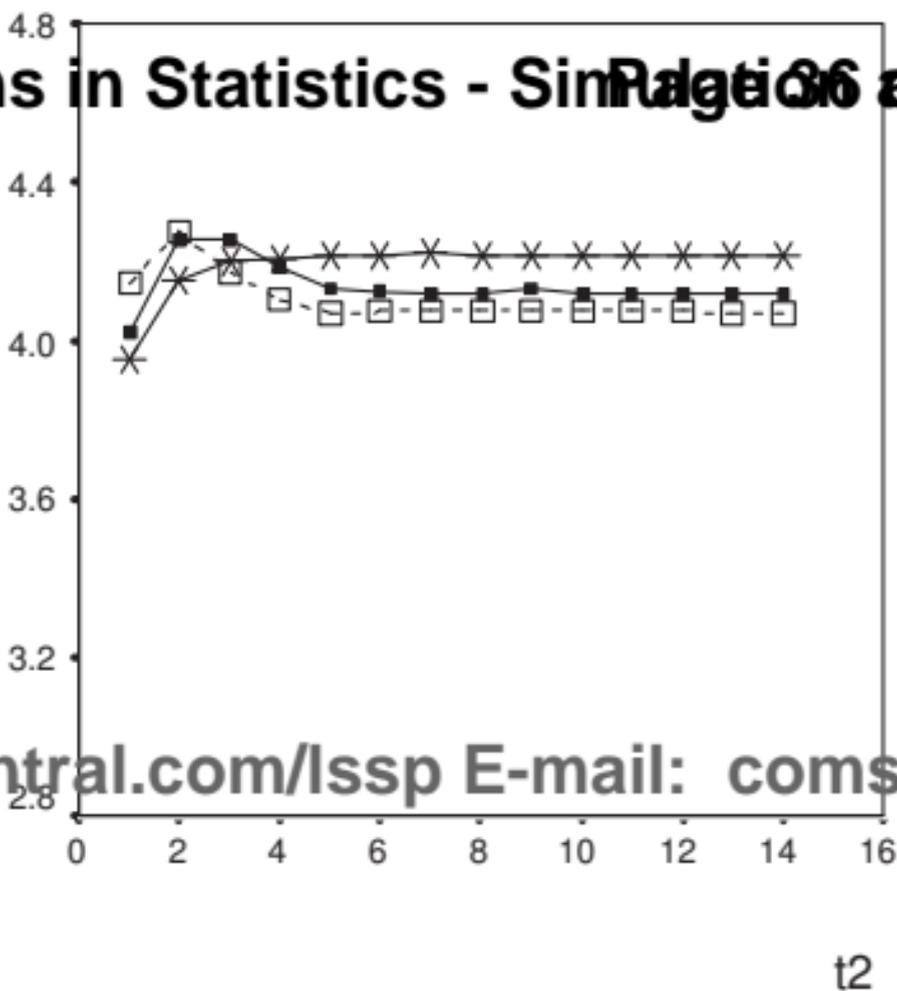
CED($t_1=1, t_2$)



t_2

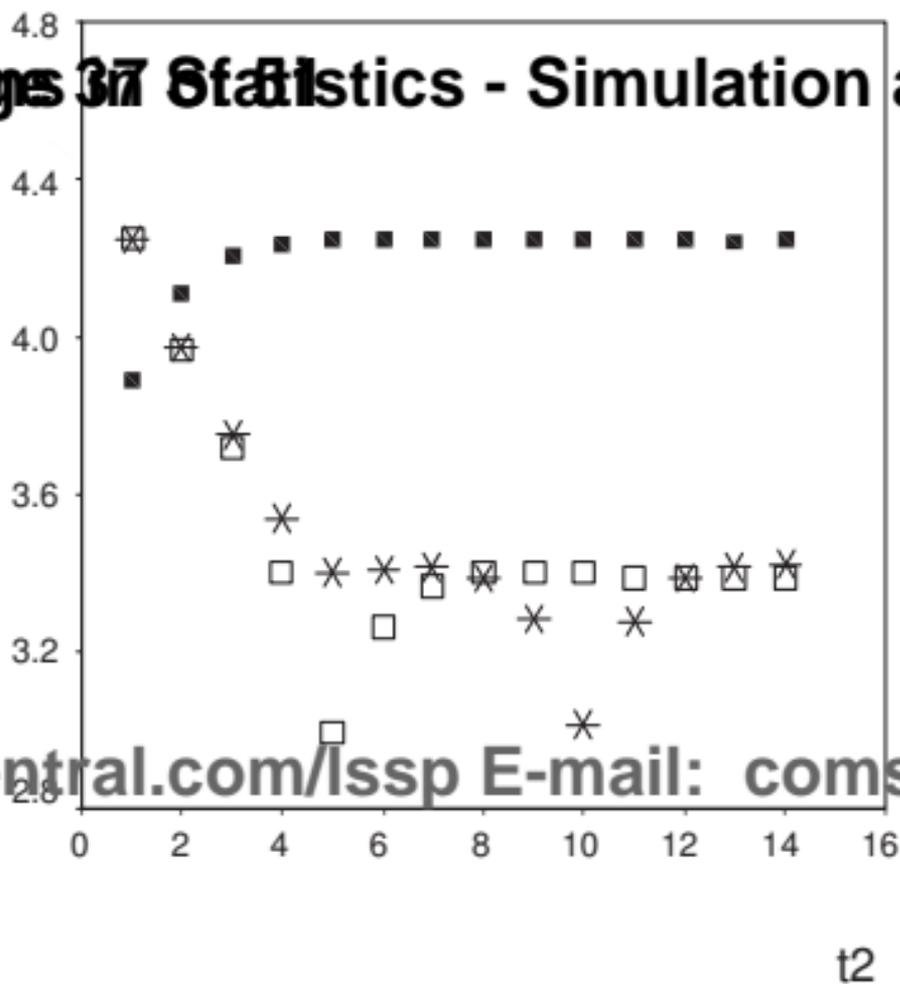
Comparisons in Statistics - Simulation of 5 Co

1
2
3
4
5
6
7
8
9



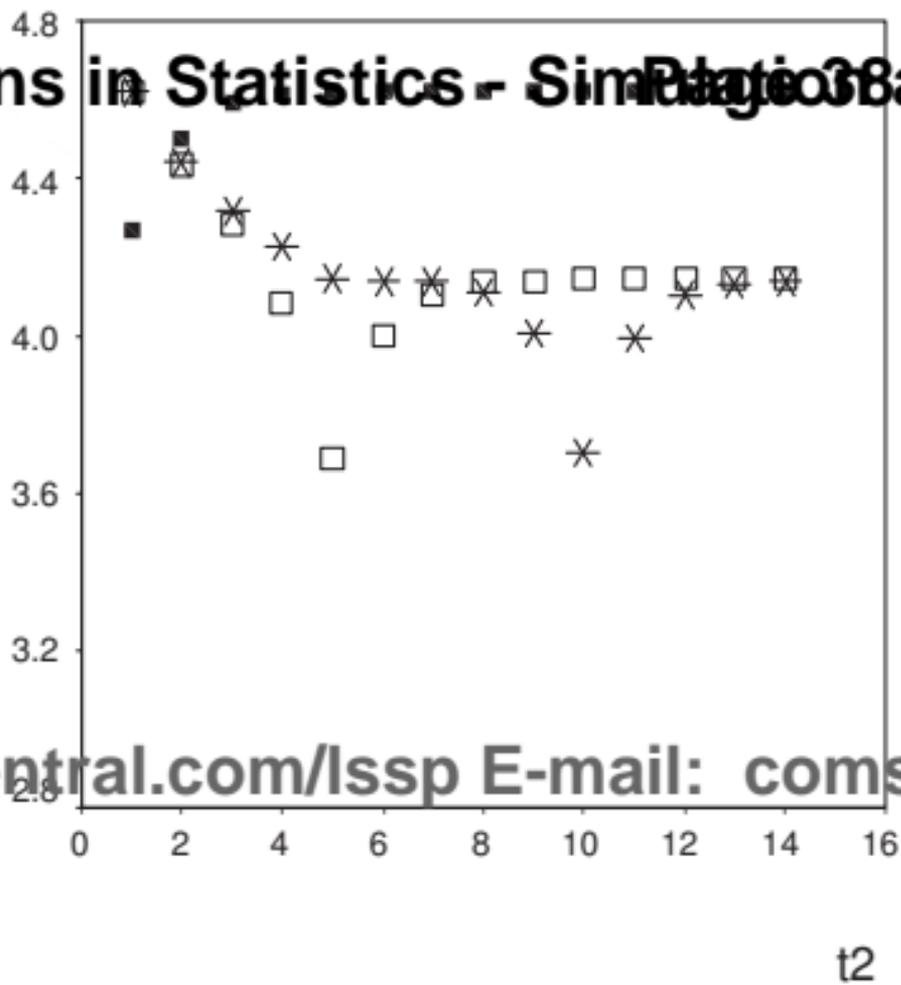
ptcentral.com/lssp E-mail: comstat@u

1
2
3
4
5
6
7
8
9



Applications in Statistics - Simulation of Barof5Con

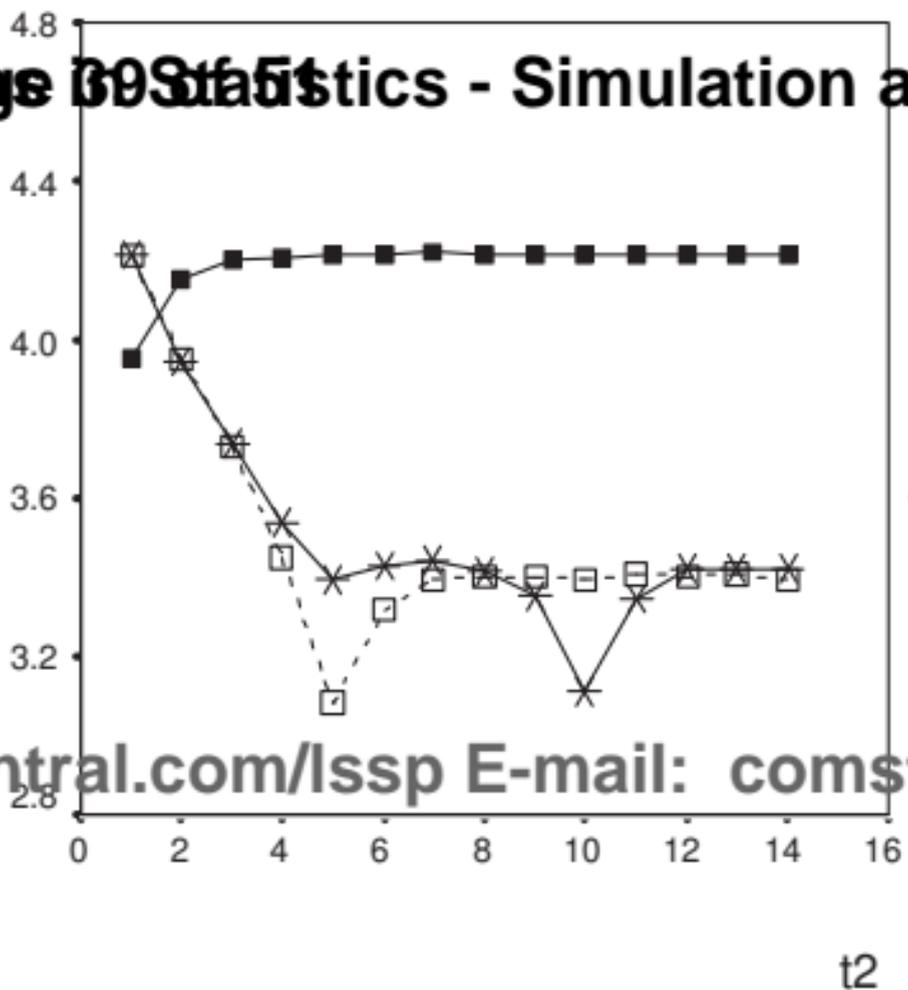
1
2
3
4
5
6
7
8
9



ipcentral.com/lssp E-mail: comstat@un

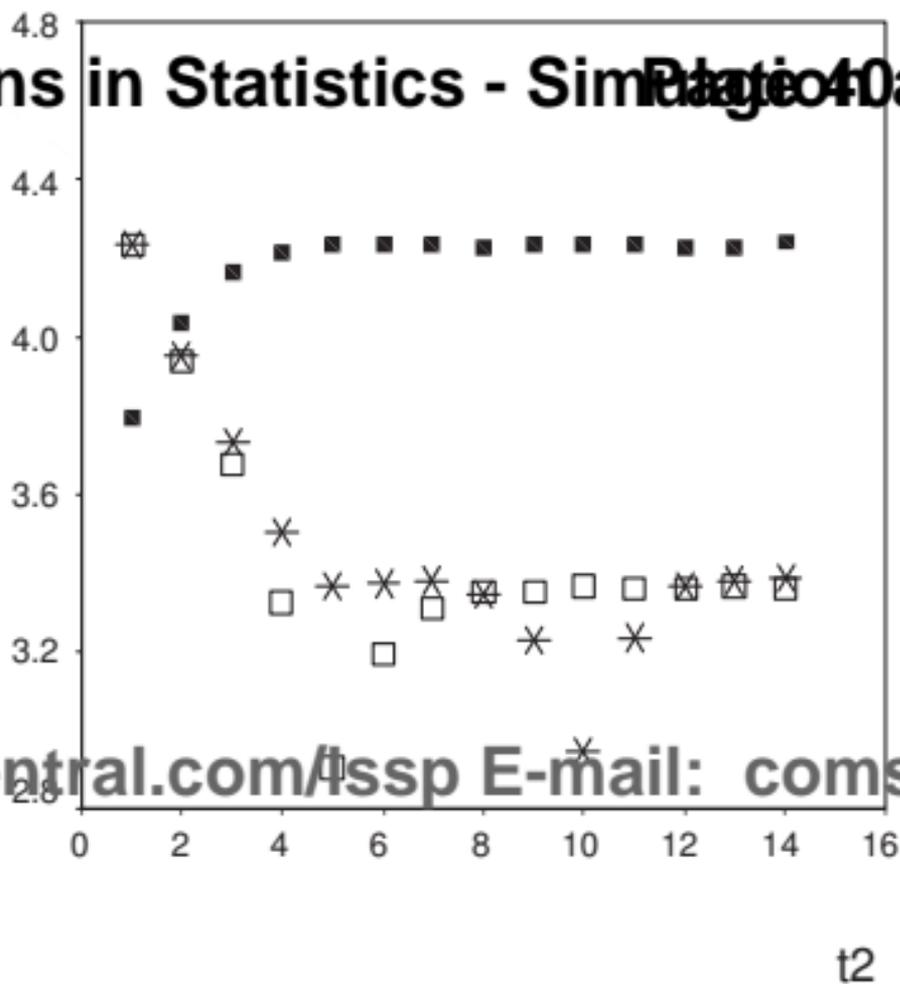
Page 9 Statistics - Simulation and Co

1
2
3
4
5
6
7
8
9



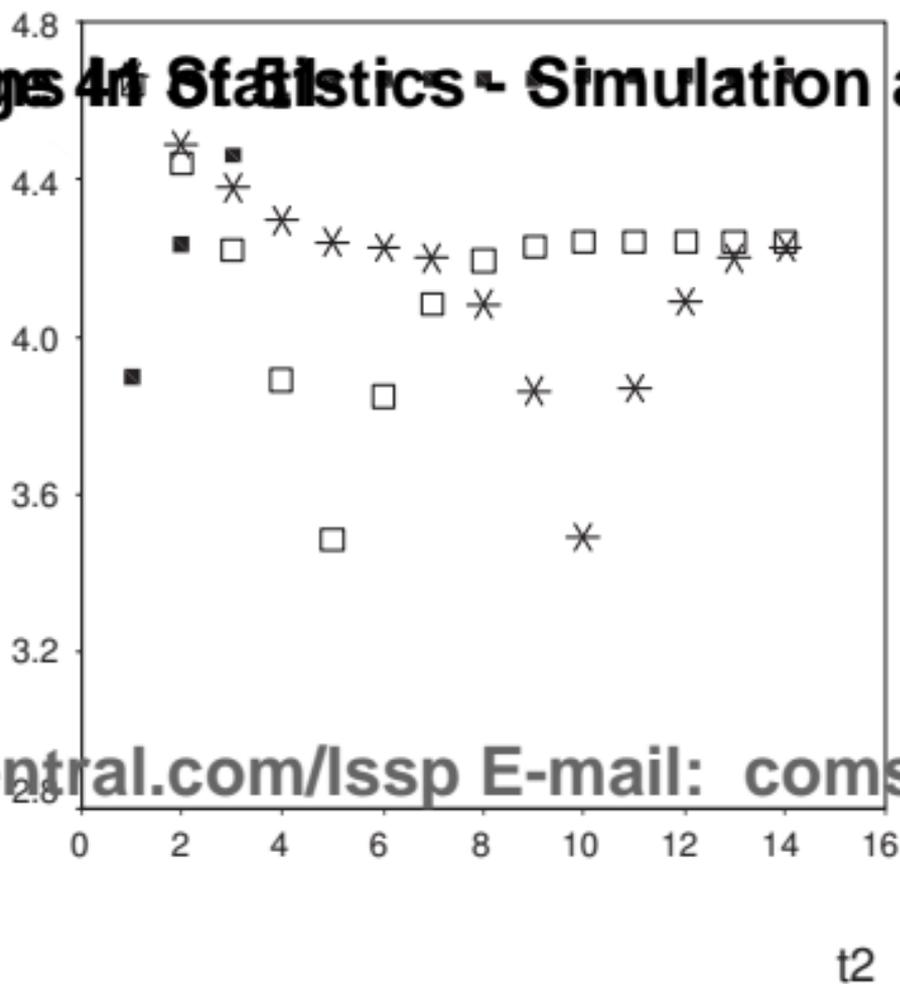
ptcentral.com/lssp E-mail: comstat@u

1
2
3
4
5
6
7
8
9



Chapter 4 Statistics - Simulation and Comparison

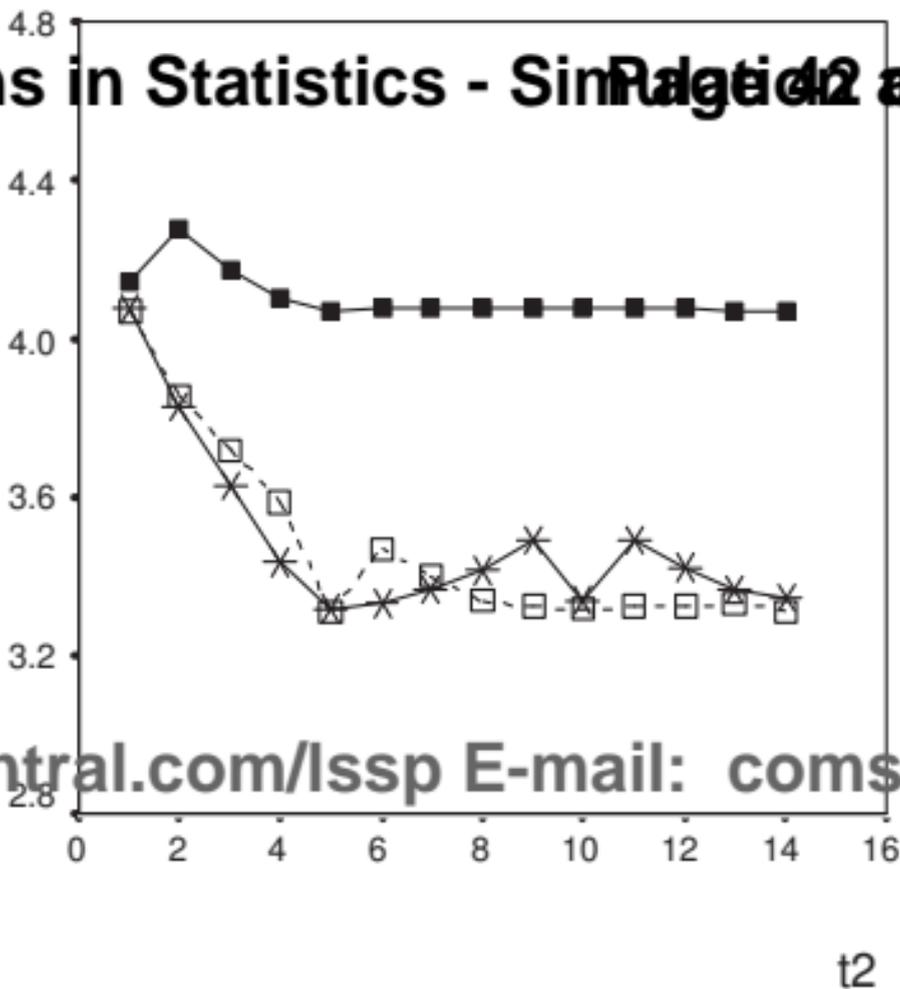
1
2
3
4
5
6
7
8
9



ipcentral.com/lssp E-mail: comstat@un

Applications in Statistics - SinRajeev and S Co

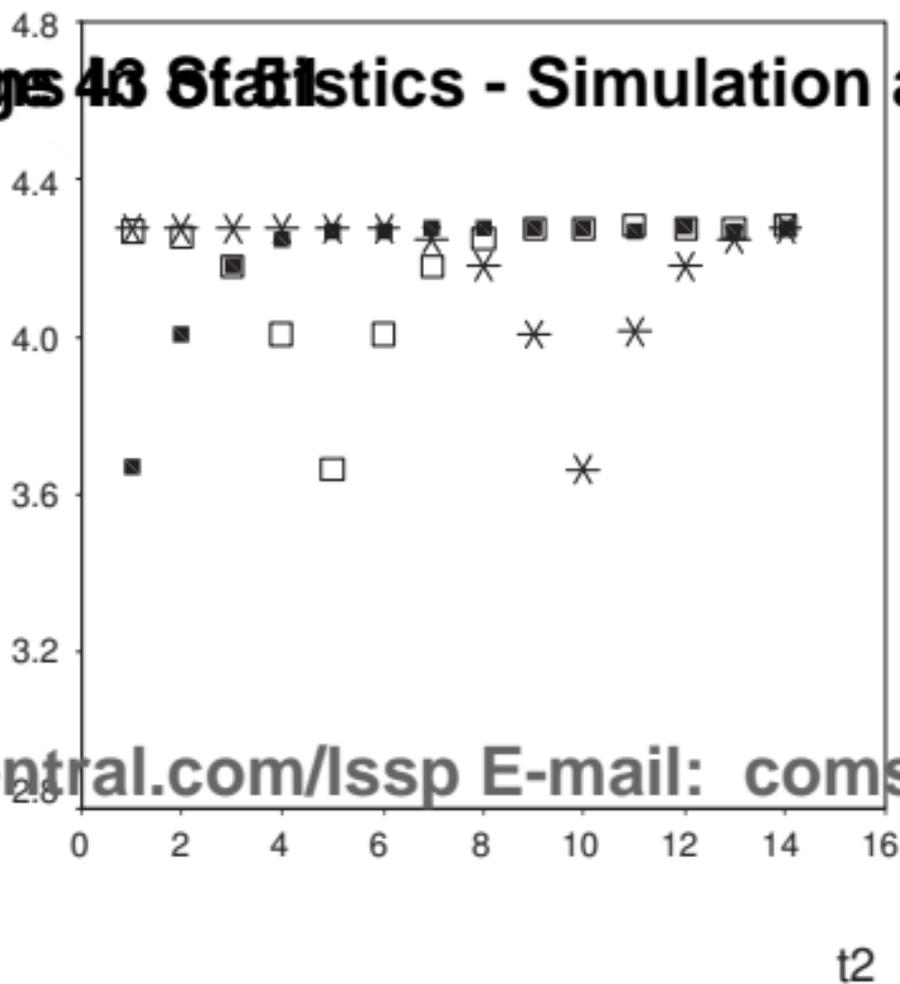
1
2
3
4
5
6
7
8
9



ptcentral.com/lssp E-mail: comstat@u

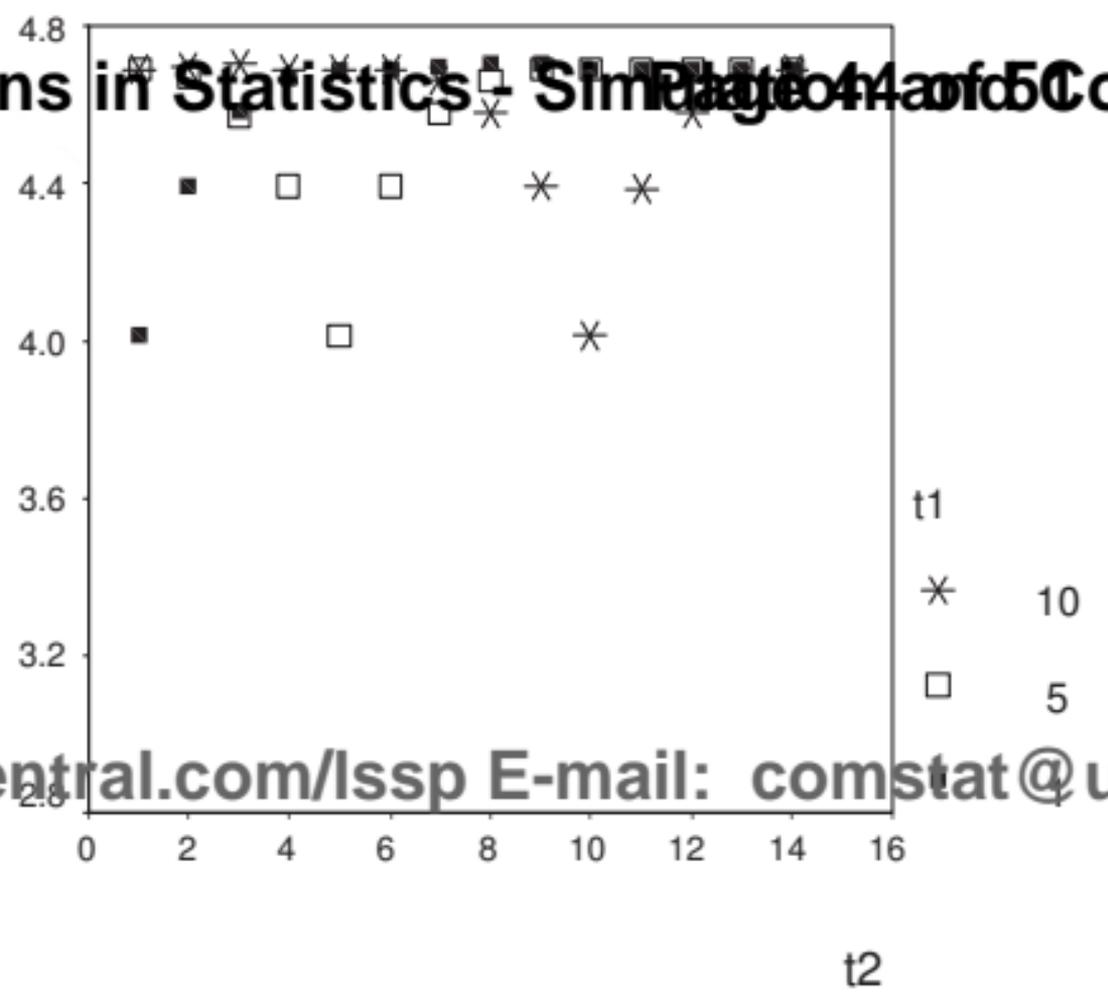
Chapter 43 Statistics - Simulation and Comparison

1
2
3
4
5
6
7
8
9



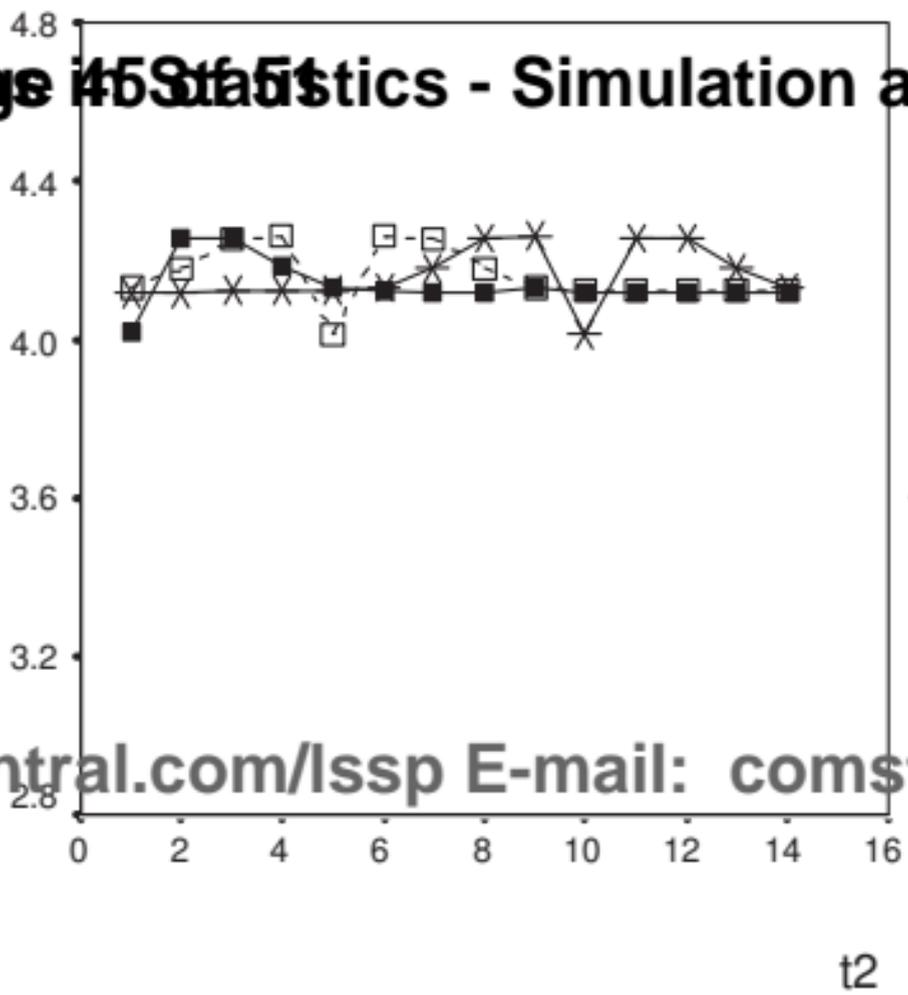
ripcentral.com/lssp E-mail: comstat@un

1
2
3
4
5
6
7
8
9



Page 45 Statistics - Simulation and Co

1
2
3
4
5
6
7
8
9



ptcentral.com/lssp E-mail: comstat@u

Figures and legends to figures

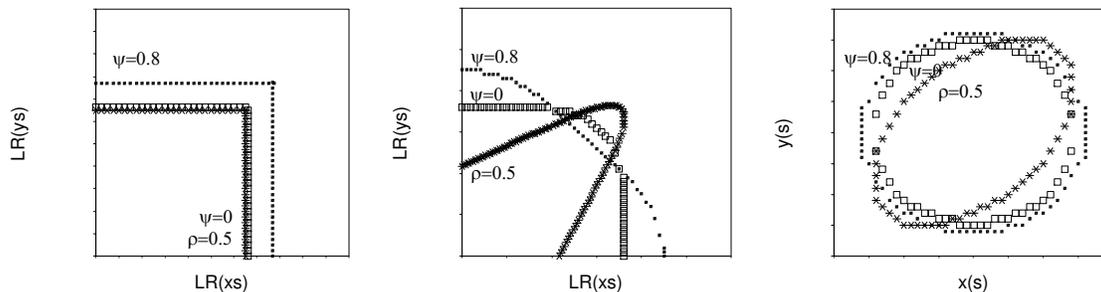


Figure 1: Alarm limit for LRpar (left), LRjoint (middle), T2 (right). Open square for $\{\psi=0, \rho=0\}$, filled square for $\{\psi=0.8, \rho=0\}$, star for $\{\rho=0.5, \psi=0\}$.

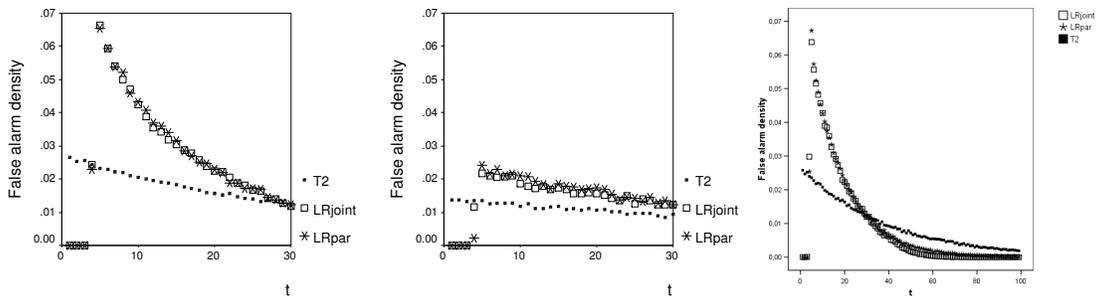


Figure 2: False alarm density, $P(t_A=t)$. Left: $\psi=0$, middle: $\psi=0.8$, right: $\rho=0.5$.

For Peer Review Only

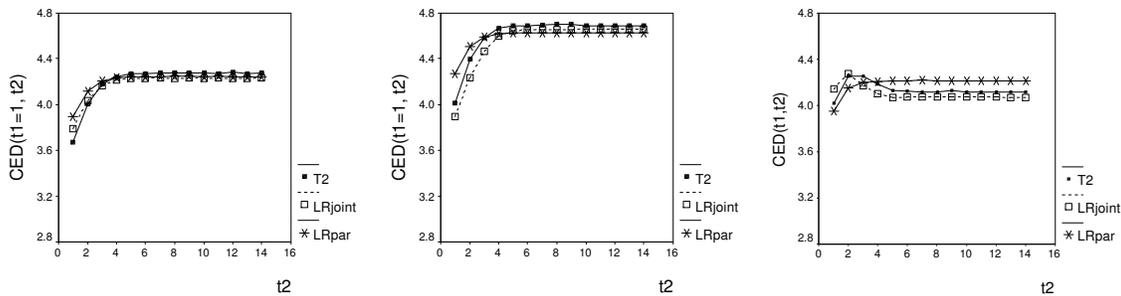


Figure 3: Conditional expected delay $CED(t_1, t_2)$. Left: $\psi=0$, middle: $\psi=0.8$, right $\rho=0.5$.

For Peer Review Only

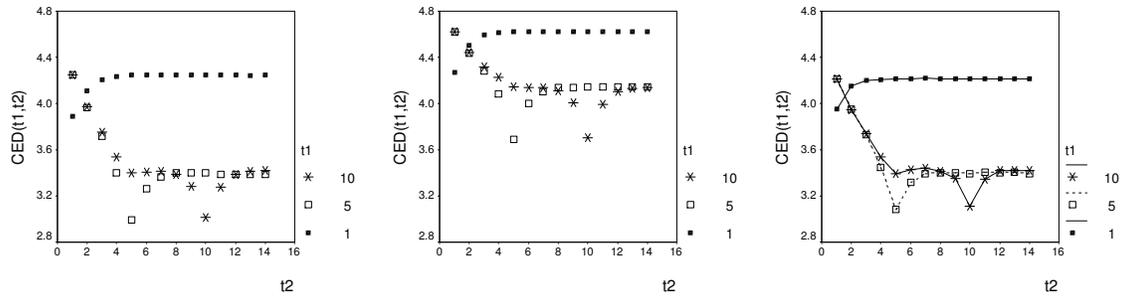


Figure 4a: Conditional expected delay, $CED(t_1, t_2)$ for LRpar, $t_1 = \{1, 5, 10\}$. Left: $\psi=0$, middle: $\psi=0.8$, right $\rho=0.5$.

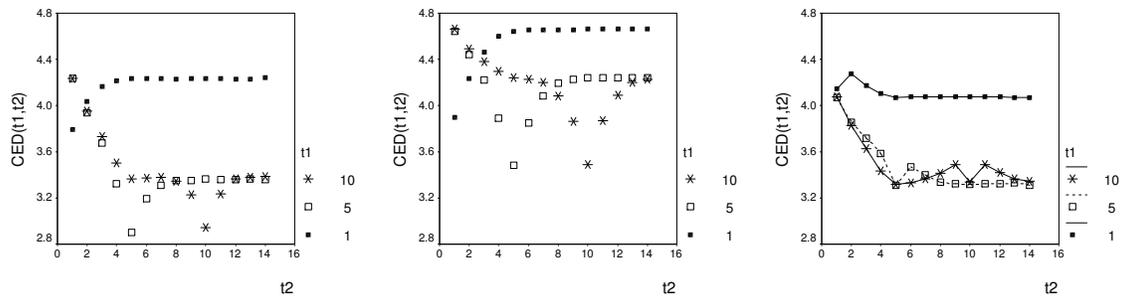


Figure 4b: Conditional expected delay, $CED(t_1, t_2)$ for LRjoint, $t_1 = \{1, 5, 10\}$. Left: $\psi=0$, middle: $\psi=0.8$, right $\rho=0.5$.

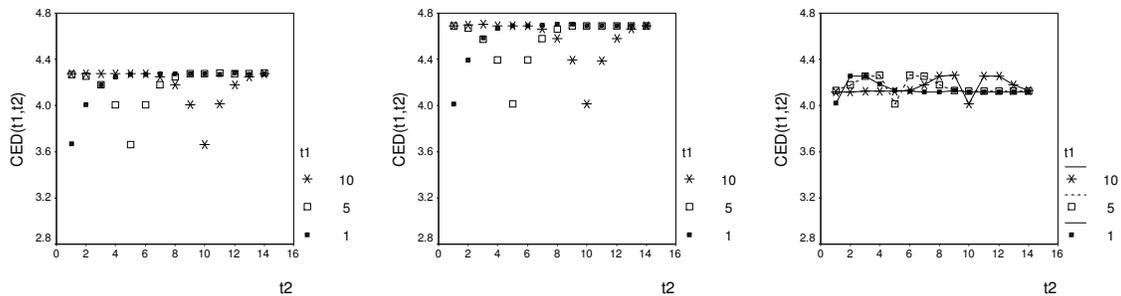


Figure 4c: Conditional expected delay, $CED(t_1, t_2)$ for T_2 , $t_1 = \{1, 5, 10\}$. Left: $\psi=0$, middle: $\psi=0.8$, right $\rho=0.5$.