RT-MAE 2005-30

On Runs in Exchangeable Bernoulli Processes

Esteves, L. G.,
Peixoto, C.M.
and
Simonis, A.

Palavras-Chave: Number of runs of fixed length, longest run, Bernoulli trials, exchangeable processes, De Finett´s Representation Theorem.
Classificação AMS: 60E05, 60G09.

- Novembro de 2005 -

On Runs in Exchangeable Bernoulli Processes*

Luís Gustavo Esteves†

Cláudia Peixoto‡

Adilson Simonis§

University of São Paulo

Abstract

The exact probability distribution for the number of runs of fixed length in n successive Bernoulli trials of a $\{0,1\}^{\infty}$-valued exchangeable process is derived. The distribution of the longest run is also obtained and a few problems for further investigation are posed.

Key words: Number of runs of fixed length, longest run, Bernoulli trials, exchangeable processes, De Finetti's Representation Theorem.

1 Introduction

The Distribution Theory of Runs is an old field of investigation in probability (according to Mood (1940), it seems it dates back to the end of the 19th century with the work of Karl Pearson (1897)) and still draws a lot of attention in academy.

Much of the interest in the Theory of Runs may be attributed to its potential applications in a broad spectrum of activities, not to mention its purely mathematical issues. Early, Wald

*Correspondence to: Adilson Simonis, Departamento de Estatística, IME- Universidade de São Paulo, Rua do Matão, 1010, Cidade Universitária, CEP 05508-090, São Paulo - SP, Brazil. E-mail: asimonis@ime.usp.br

†Supported by FAPESP through Grant 2003/10105-2,

‡Supported by FAPESP through Grant 2003/09930-9,

§Supported by FAPESP through Grant 2004/07276-2 and PRONEX/CNPq/FAPERJ *Probabilidade e Processos Estocásticos.*

and Wolfowitz (1940) derived not only the distribution of the total number of runs in random sequences of elements of two kinds but also its asymptotic behavior in connection with the problem of (hypothesis) testing the equality of two continuous distribution functions. Mood (1940) developed the formulae for the number of runs of fixed length from binomial and multinomial populations. Later on, Burr and Cane (1961) studied the longest run of consecutive observations having a specified attribute. Feller (1968) stated the "Poissonity" of the limiting distribution of the number of success runs in a sequence of Bernoulli variables via generating function. In the eighties, Philippou and Makri (1986) and Hirano (1986) independently obtained the exact distribution for the number of success runs of given length and for the longest success run in n independent Bernoulli trials. Godbole (1990(a), 1990(b)) derived an alternative expression to the one established by Philippou and Makri (1986) and Hirano (1986) and determined its asymptotic form in a different fashion from Feller's (1968). In the last decade, the Theory of Runs continued to make great advances: in Godbole (1991) runs were examined in connection with the development of models for the occurrence of genetic patterns in ADN sequences; Fu and Koutras (1994) and Lou (1996) focused the Distribution Theory of Runs based on a finite Markov Chain imbedding technique. More recently, Aki and Hirano (1999) investigated waiting time problems for runs in Markov bivariate trials and Vaggelatou (2003) derived the distribution of the length of the longest run in a multi-state Markov chain. Erylmaz (2005) obtained the distribution of the length of the longest hydrologic risk period applying runs to hidrology.

Certainly, the authors do not intend the aforementioned references to cover the least of the bibliography of runs; in fact, just to emphasize its importance. For a detailed account of the historical development of the Theory of Runs, see Balakrishnan and Koutras (2002).

In this work, we develop expressions for the number of runs of fixed length and for the longest run in successive Bernoulli trials of a $\{0,1\}^{\infty}$-valued exchangeable process. To the best of the author's knowledge, these results have not been presented in the literature.

For the reader's guidance, section 2 is devoted to the derivation of the main result, namely the distribution of the number of runs of fixed length, illustrated then with a few examples. We end that section up with the distribution of the longest run. In section 3, we make our final comments and point some directions for future inquires.

2 Main Result

In the sequel, usual probability distributions in the Theory of Runs are obtained for sequences of exchangeable Bernoulli variables. Such derivation is based on the celebrated De Finetti's Representation Theorem and on some combinatory analysis.

Let $(X_n)_{n\geq 1}$ be a $\{0,1\}^\infty$-valued exchangeable process and $\mathbb{P}$ its probability measure. As usual, the event $\{X_n = 1\}$, $n \geq 1$, will be named a success (more precisely, a success in the n-th stage of the process) and its complement a failure. Also, for each $n \in \mathbb{N}$, let $S_n = \sum_{i=1}^{n}(1 - X_i)$ be the number of failures in the first n steps of the process. By a run of length k, $k \geq 1$, we mean a sequence of k consecutive successes followed by a failure. In this way, we are interested in the probability distribution of the quantity

$$N_n^{(k)} = \sum_{i=1}^{n-k} X_i X_{i+1} \ldots X_{i+k-1}(1 - X_{i+k}), \quad n > k \tag{2.1}$$

that is, the number of runs of length k in the first n stages of the process $(X_n)_{n\geq 1}$. We should mention that the definition of a run considered here is a little different from that proposed by Feller (1968).

Now we state our main result.

Theorem 2.1 *Let $(X_n)_{n\geq 1}$ be an exchangeable process taking values on $\{0,1\}^\infty$. Let $N_n^{(k)}$ be as defined in (2.1). Then, if $\mathbb{P}$ denotes the probability measure of the process,*

$$\mathbb{P}(N_n^{(k)} = t) = \int_0^1 \sum_{j=0}^{n} (1-\theta)^j \theta^{n-j} \binom{j}{t} \sum_{i=0}^{j-t} (-1)^i \binom{j-t}{i} \binom{n-(t+i)k}{j} d\mu(\theta),$$

$t = 0, \ldots, [\frac{n}{k+1}]$, where μ is De Finetti's measure associated with the process and $[x]$ is the (usual) notation for the largest integer less than or equal to x.

Proof: We adopt the usual convention $\binom{a}{b} = 0$, whenever $a < b$.

For $n \in \mathbb{N}$ fixed, let $A_{n,t} = \{(x_1, \ldots, x_n) \in \{0,1\}^n : \sum_{i=1}^{n-k} x_i \ldots x_{i+k-1}(1 - x_{i+k}) = t\}$.

We then have,

$$\mathbb{P}(N_n^{(k)} = t) = \sum_{(x_1,\ldots,x_n)\in A_{n,t}} \mathbb{P}(\{(x_1, \ldots, x_n)\}) = \sum_{A_{n,t}} \mathbb{P}(\{(x_1, \ldots, x_n)\}), \tag{2.2}$$

As $(X_n)_{n\geq 1}$ is exchangeable, De Finetti's Representation Theorem holds (De Finetti (1937), Heath and Sudderth (1976)) and (2.2) may be rewritten as

$$\mathbb{P}(N_n^{(k)}=t)=\sum_{A_{n,t}}\int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}\,d\mu(\theta)=\int_0^1\sum_{A_{n,t}}\theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}\,d\mu(\theta),$$

where the last equality follows from Tonelli's Theorem. Setting

$$A_{n,t,j}=\{(x_1,\dots,x_n)\in A_{n,t}:\sum_{i=1}^n(1-x_i)=j\},\ j=0,1,\dots,n,$$

it follows that

$$\mathbb{P}(N_n^{(k)}=t)=\int_0^1\sum_{j=0}^n|A_{n,t,j}|\ \theta^{n-j}(1-\theta)^j d\mu(\theta), \tag{2.3}$$

where $|A|$ represents the cardinality of the set A.

Thus, in order to obtain the probability distribution of $N_n^{(k)}$ in (2.3), we only need to evaluate $|A_{n,t,j}|, j=1,\dots,n$. Clearly, $|A_{n,t,0}|=0$. By definition, a n−uple in $A_{n,t,j}$ has exactly j coordinates equal to 0 and $n-j$ equal to 1. We consider then, for each $x\in\{0,1\}^n$ with j coordinates equal to 0, $j=1,\dots,n$, the quantities $Y_1^{\underset{\sim}{x}},\dots,Y_{j+1}^{\underset{\sim}{x}}$ as follows: $Y_1^{\underset{\sim}{x}}$ is the number of 1's preceeding the first 0 in the n-uple (that is, $Y_1^{(x_1,\dots,x_n)}=\min\{i\in\{1,\dots,n\}:x_i=0\}-1$), $Y_i^{\underset{\sim}{x}}$ is the number of 1's between the $(i-1)$-th and the i−th 0's in the n-uple, $2\leq i\leq j$ (or $Y_i^{(x_1,\dots,x_n)}=\min\{k\in\{1,\dots,n\}:\sum_{l=1}^k(1-x_l)=i\}-\min\{k\in\{1,\dots,n\}:\sum_{l=1}^k(1-x_l)=i-1\}-1$) and $Y_{j+1}^{\underset{\sim}{x}}$ is the number of 1's after the (last) j−th 0 in the n-uple ($Y_{j+1}^{\underset{\sim}{x}}=n-\min\{k\in\{1,\dots,n\}:\sum_{l=1}^k(1-x_l)=j\}$). Clearly, $Y_1^{\underset{\sim}{x}}+\dots+Y_{j+1}^{\underset{\sim}{x}}=n-j$, if $\underset{\sim}{x}$ possesses j 0's. Thus, a point $\underset{\sim}{x}=(x_1,\dots,x_n)\in\{0,1\}^n$ belongs to $A_{n,t,j}$ if, and only if, there are exactly t of the j first quantities $Y_1^{\underset{\sim}{x}},\dots,Y_j^{\underset{\sim}{x}}$ greater than $k-1$ and $j-t$ of them smaller than k. In this way, the problem of determining $|A_{n,t,j}|$ is equivalent to the combinatory problem of calculating the number of nonnegative integer solutions of $Y_1+\dots+Y_{j+1}=n-j$ with t of the variables $Y_1,\dots,Y_j$ being at least k and $n-t$ of them being strictly smaller than k.

Defining $B_i=\{Y_i\geq k\}, i=1,\dots,j+1$, notice that

$$A_{n,t,j}=\bigcup_{\substack{I\subset\{1,\dots,j\}:\\|I|=t}}\left[(\cap_{i\in I}B_i)\cap(\cap_{i\notin I}B_i^c)\right].$$

It follows that

$$|A_{n,t,j}|=\sum_{\substack{I\subset\{1,\dots,j\}:\\|I|=t}}\left|(\cap_{i\in I}B_i)\cap(\cap_{i\notin I}B_i^c)\right|=\binom{j}{t}\left|(\cap_{i=1}^t B_i)\cap(\cap_{i=t+1}^j B_i^c)\right| \tag{2.4}$$

As

$$(\cap_{i=1}^{t} B_i) \cap (\cap_{i=t+1}^{j} B_i^c) = (\cap_{i=1}^{t} B_i) - (\cap_{i=1}^{t} B_i) \cap (\cup_{i=t+1}^{j} B_i),$$

we obtain that

$$\left|(\cap_{i=1}^{t} B_i) \cap (\cap_{i=t+1}^{j} B_i^c)\right| = \left|(\cap_{i=1}^{t} B_i)\right| - \left|(\cap_{i=1}^{t} B_i) \cap (\cup_{i=t+1}^{j} B_i)\right|.$$

From the above equality, we have

$$\begin{aligned}\left|(\cap_{i=1}^{t} B_i) \cap (\cap_{i=t+1}^{j} B_i^c)\right| &= \binom{n-kt}{j} - \left|\cup_{i=t+1}^{j} (B_i \cap (\cap_{l=1}^{t} B_l))\right| \\ &= \binom{n-kt}{j} - \sum_{i=1}^{j-t} (-1)^{i+1} \binom{j-t}{i} \binom{n-(t+i)k}{j}.\end{aligned}$$

Therefore,

$$\left|(\cap_{i=1}^{t} B_i) \cap (\cap_{i=t+1}^{j} B_i^c)\right| = \sum_{i=0}^{j-t} (-1)^i \binom{j-t}{i} \binom{n-(t+i)k}{j}. \tag{2.5}$$

From (2.3), (2.4) and (2.5), we obtain the result.

■

Theorem 2.1 gives the exact distribution of the number of runs of length k for exchangeable Bernoulli variables, extending in a sense Godbole´s result in [9] for the i.i.d. case. The extension arises ultimately from the conditional uniformity of any finite exchangeable sequence of Bernoulli trials (not only of i.i.d. ones) given the sum of its components (or the number of successes/failures in the sequence), a minimal sufficient statistic for this whole family of distributions (as a matter of fact, this is equivalent to finite exchangeability in the 0-1 case). The combinatorial argument is therefore quite similar to that of Godbole (1990a), except for the setting based on nonnegative integer solutions of a equation (in place of Godbole´s presentation using urn schemes). Also, the argument holds for the usual definition of a run (Feller (1968)).

The distribution of $N_n^{(k)}$ in Theorem 2.1 depends explicitly on De Finetti´s measure of the process and, for this reason, may be not appealing at first glance. Nevertheless, it holds for a large class of stochastic processes, without mentioning the exact distributions in the Theory of Runs do not seem sympathetic in general.

In this work, we have tackled neither the weak convergence of $N_n^{(k)}$ nor the derivation of exact probability distributions for runs in broader classes of exchangeable processes (the latter problem hinted by the existence of De Finetti´s type theorems for more general models). Next, we illustrate Theorem 2.1 with three examples.

Example 2.1.1 *Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. Bernoulli random variables with parameter θ_0, $0<\theta_0<1$. In this case, $\mu(\{\theta_0\})=1$, that is, De Finetti´s measure of the process is degenerate at θ_0, and the integral in Theorem 2.1 reduces to an expression similar to that obtained by Godbole (1990a),*

$$\mathbb{P}(N_n^{(k)}=t)=\sum_{j=0}^{n}(1-\theta_0)^j\theta_0^{n-j}\binom{j}{t}\sum_{i=0}^{j-t}(-1)^i\binom{j-t}{i}\binom{n-(t+i)k}{j}.$$

Example 2.1.2 *(Pólya´s urn scheme) Let $(X_n)_{n\geq 1}$ be a stochastic process with probability measure $\mathbb{P}$ defined by, $\forall\ n\in\mathbb{N}$, $\forall\ (x_1,\ldots,x_n)\in\{0,1\}^n$,*

$$\mathbb{P}(X_1=x_1,\ldots,X_n=x_n)=\frac{\Gamma(\frac{a}{c}+\sum_{i=1}^{n}x_i)\Gamma(\frac{b}{c}+n-\sum_{i=1}^{n}x_i)\Gamma(\frac{a+b}{c})}{\Gamma(\frac{a}{c})\Gamma(\frac{b}{c})\Gamma(\frac{a+b}{c}+n)}=$$
$$=\textit{Pólya}((x_1,\ldots,x_n)/((a,b),c)),$$

where $\Gamma(\cdot)$ is the gamma function and a,b,c natural numbers. $(X_n)_{n\geq 1}$ describes the evolution of a Pólya-Eggenberger urn model (Pólya and Eggenberger (1923)) initially with a white balls and b black balls in the urn, or, with initial configuration (a,b) and with c balls being added to the urn at each stage (X_n denotes then the indicator of a white ball being drawn in the $n-$th step).

It is easily seen that $(X_n)_{n\geq 1}$ is exchangeable with De Finetti´s measure given by a Beta density with parameters $\frac{a}{c}$ and $\frac{b}{c}$ (Johnson and Kotz (1977)). Thus, $(X_n)_{n\geq 1}$ satisfies the conditions of Theorem 2.1 and, in this case, $N_n^{(k)}$ has the following distribution

$$\mathbb{P}(N_n^{(k)}=t)=$$
$$=\int_0^1\Big\{\sum_{j=0}^{n}(1-\theta)^j\theta^{n-j}\binom{j}{t}\sum_{i=0}^{j-t}(-1)^i\binom{j-t}{i}\binom{n-(t+i)k}{j}\Big\}\frac{\Gamma(\frac{a+b}{c})}{\Gamma(\frac{a}{c})\Gamma(\frac{b}{c})}\theta^{\frac{a}{c}-1}(1-\theta)^{\frac{b}{c}-1}d\theta$$
$$=\sum_{j=0}^{n}\binom{j}{t}\sum_{i=0}^{j-t}(-1)^i\binom{j-t}{i}\binom{n-(t+i)k}{j}\frac{\Gamma(\frac{a+b}{c})}{\Gamma(\frac{a}{c})\Gamma(\frac{b}{c})}\Big\{\int_0^1\theta^{\frac{a}{c}+n-j-1}(1-\theta)^{\frac{b}{c}+j-1}d\theta\Big\}.$$

Therefore,

$$\mathbb{P}(N_n^{(k)}=t)=\frac{\Gamma(\frac{a+b}{c})}{\Gamma(\frac{a}{c})\Gamma(\frac{b}{c})}\sum_{j=0}^{n}\binom{j}{t}\frac{\Gamma(\frac{a}{c}+n-j)\Gamma(\frac{b}{c}+j)}{\Gamma(\frac{a+b}{c}+n)}\sum_{i=0}^{j-t}(-1)^i\binom{j-t}{i}\binom{n-(t+i)k}{j}.$$

Example 2.1.3 *(Pólya-Eggenberger urn model with unknown initial composition). Let $(X_n)_{n\geq 1}$ be a stochastic process the probability measure of which, $\mathbb{P}$, is given by, $\forall\, n \in \mathbb{N}$, $\forall\, (x_1, \ldots, x_n) \in \{0,1\}^n$*

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \sum_{l=1}^{\infty} p_l.Pólya((x_1, \ldots, x_n)\big|((a_l, b_l), 1)),$$

where $\{((a_l, b_l), p_l)\}_{l\in\mathbb{N}}$, with $(a_l, b_l) \in \mathbb{N}^2$, $p_l \geq 0, \forall\, l \in \mathbb{N}$, and $\sum_{l=1}^{\infty} p_l = 1$, is a probability measure on $(\mathbb{N}^2, \mathcal{P}(\mathbb{N}^2))$. In this example, $(X_n)_{n\geq 1}$ describes the evolution of a variant version of Pólya-Eggenberger urn model, that is to say, with $c = 1$ and random (unknown, under the subjectivistic view of probability) initial composition. The probability measure $\{((a_l, b_l), p_l)\}_{l\in\mathbf{N}}$ should be considered (still under the Bayesian standpoint) as a numerical transcription of personal uncertainty about the initial composition of the urn (Esteves, Wechsler, Iglesias and Pereira (2003)). $(X_n)_{n\geq 1}$ is of course exchangeable and possesses De Finetti´s measure given by a mixture of Beta densities (more precisely, the Beta distribution with parameters a_l and b_l is weighted by $p_l, l = 1, 2, \ldots$). The distribution of the number of runs of length k in n successive stages of a Pólya-Eggenberger urn model with unknown initial composition is then

$$\mathbb{P}(N_n^{(k)} = t) = \sum_{l=1}^{\infty} p_l \frac{\Gamma(a_l + b_l)}{\Gamma(a_l)\Gamma(b_l)} \sum_{j=0}^{n} \binom{j}{t} \frac{\Gamma(a_l + n - j)\Gamma(b_l + j)}{\Gamma(a_l + b_l + n)} \sum_{i=0}^{j-t} (-1)^i \binom{j-t}{i} \binom{n-(t+i)k}{j}.$$

We finish this section presenting the distribution of the longest run in n successive stages of a $\{0,1\}^{\infty}$-valued exchangeable process resultant from Theorem 2.1.

Theorem 2.2 *Let $(X_n)_{n\geq 1}$ be an exchangeable process taking values on $\{0,1\}^{\infty}$. Let R_n be the longest run in the first n steps of the process. Then,*

$$\mathbb{P}(R_n = r) = \int_0^1 \sum_{j=0}^{n} (1-\theta)^j \theta^{n-j} \sum_{i=0}^{j} (-1)^i \binom{j}{i} \left[\binom{n-(r+1)i}{j} - \binom{n-ri}{j}\right] d\mu(\theta),$$

$r = 0, \ldots, n-1$. ($R_n = 0$ represents no occurrence of runs).

Proof: It is easy to check that

$$R_n \leq r \;\text{ if and only if }\; N_n^{(r+1)} = 0.$$

Thus,

$$\begin{aligned}\mathbb{P}(R_n = r) &= \mathbb{P}(R_n \leq r) - \mathbb{P}(R_n \leq r-1)\\ &= \mathbb{P}(N_n^{(r+1)} = 0) - \mathbb{P}(N_n^{(r)} = 0).\end{aligned}$$

Theorem 2.1 and straightforward calculations yield

$$\mathbb{P}(R_n = r) = \int_0^1 \sum_{j=0}^{n} (1-\theta)^j \theta^{n-j} \sum_{i=0}^{j} (-1)^i \binom{j}{i} \left[\binom{n-(r+1)i}{j} - \binom{n-ri}{j} \right] d\mu(\theta).$$

■

3 Conclusions

The exact distribution of the number of runs of fixed length in successive Bernoulli trials of a $\{0,1\}^\infty$-valued exchangeable process is obtained from De Finetti's Representation Theorem. In addition, the formula for the longest run is also deduced. It should be emphasized that the asymptotic forms of such distributions have not been investigated (including other definitions of run), being possibly the subject of a future work. Other run statistics, as the time of the occurrence of the last run of fixed length, for instance, have not been dealt with either. It is also open to speculation whether De Finetti's type Theorems for other parametric models may lead to the probability distributions of run statistics in broader classes of exchangeable processes.

References

[1] Aki, S.; Hirano, K. (1999). Sooner and later waiting time problems for runs in Markov dependent bivariate trials. Ann. Inst. Statist. Math., **51**, 17-29.

[2] Burr, E. J.; Cane, G. (1961). Longest run of consecutive observations having a special attribute. Biometrika, **48**, 461-465.

[3] De Finetti, B. (1937). Lé prevision: ses lois logiques, ses sources subjectives. Ann. Inst. H. Poincaré. Prob. Statist., **7**,1-68.

[4] Erylmaz, S. (2005). The Longest Run Statistic Associated with Exchangeable Binary Variables. Turkish J. Eng. Env. Sci., **29**, 105-111.

[5] Balakrishnan, N; Koutras, M. V. (2002). Runs and Scans with Applications. Wiley Series in probability and Statistics.

[6] Esteves, L. G.; Wechsler, S.; Iglesias, P. L.; Pereira, A. L. (2003). A variant version of the Pólya-Eggenberger Urn Model. In Bayesian Statistics **7**, 511-518. J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith and M. West eds.. Oxford University Press.

[7] Feller, W. (1968). An Introduction to Probability Theory and Its Applications. Vol. I. Wiley, New York. 3^{rd} ed.

[8] Fu, J.; Koutras, M. (1994). Distribution Theory of Runs: Markov Chain Approach. Journal of the American Statistical Association, **89**, 1050-1058.

[9] Godbole, A. P. (1990a). Specific formulae for some success run distributions. Statistics and Probability Letters, **10**, 119-124.

[10] Godbole, A. P. (1990b). Degenerate and Poisson convergence criteria for success runs. Statistics and Probability Letters, **10**, 247-255.

[11] Godbole, A. P. (1991). Poisson approximations for runs and patterns of rare events. Adv. Appl. Prob.,**23**, 851-865.

[12] Heath, D.; Sudderth, W. (1976). De Finetti's Theorem on Exchangeable Variables. The American Statistician, Vol. 30, **4**, 188-189.

[13] Hirano, K. (1986). Some properties of the distributions of order k. In Fibonacci Numbers with Applications. Proceedings of the First International Conference on Fibonacci Numbers and Their Aplications. A. N. Philippou, A. F. Horadam and G. E. Bergum eds.. D. Reidel, Dordrecht.

[14] Lou, W. Y. W. (1996). On Runs and Longest Run Tests: A Method of Finite Markov Chain Imbedding. Journal of the American Statistical Association, **91**, 1595-1601.

[15] Mood, A. M. (1940). The Distribution Theory of Runs. The Annals of Mathematical Statistics, Vol. XI, **4**, 367-392.

[16] Pearson, K. (1897). The Chances of Death and Other Studies in Evolution. Vol. I. London.

[17] Philippou, A. N.; Makri, F. S. (1986). Successes, Runs and Longest Runs. Statistics and Probability Letters, **4**, 101-105.

[18] Vaggelatou, E. (2003). On the length of the longest run in a multi-state Markov Chain. Statistics and Probability Letters, **62**, 211-221.

[19] Wald, A.; Wolfowitz, J. (1940). On a Test Whether Two Samples are from the same Population. The Annals of Mathematical Statistics, Vol. XI, **2**, 147-162.

[20] Pólya, G.; Eggenberger, F. (1923). Über die Statistik verketteter Vorgange. Z. Angew. Math. Mech., **3**, 279-289.

[21] Johnson, N. L.; Kotz, S. (1977). Urn Models and Their Applications. New York, Wiley.

ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

2005-01 - DE SOUZA BORGES,W., GUSTAVO ESTEVES, L., WECHSLER, S. Process Parameters Estimation in The Taguchi On-Line Quality Monitoring Procedure For Attributes. 2005.19p. (RT-MAE-2005-01)

2005-02 - DOS ANJOS,U.,KOLEV, N. Copulas with Given Nonoverlapping Multivariate Marginals. 2005.09p. (RT-MAE-2005-02)

2005-03 - DOS ANJOS,U.,KOLEV, N. Representation of Bivariate Copulas via Local Measure of Dependence. 2005.15p. (RT-MAE-2005-03)

2005-04 - BUENO, V. C., CARMO, I. M. A constructive example for active redundancy allocation in a k-out-of-n:F system under dependence conditions. 2005. 10p. (RT-MAE-2005-04)

2005-05 - BUENO, V. C., MENEZES, J.E. Component importance in a modulated Markov system. 2005. 11p. (RT-MAE-2005-05)

2005-06 - BASAN, J. L., BRANCO, M.D´E., BOLFARINE, H. A skew item response model. 2005. 20p. (RT-MAE-2005-06)

2005-07 - KOLEV, N., MENDES, B. V. M., ANJOS, U. Copulas: a Review and recent developments. 2005. 46p. (RT-MAE-2005-07)

2005-08 - VENEZUELA, M. K., BOTTER, D. A., SANDOVAL, M. C. Diagnostic techniques in generalized estimating equations. 2005. 16p. (RT-MAE-2005-08)

2005-09 - BOLFARINE, H., LACHOS, V.H. Skew-probit measurement error models. 2005. 10p. (RT-MAE-2005-09)

2005-10 - SALINAS, V.H., ROMEO, J.S., PEÑA, A. On Bayesian estimation of a survival curve: Comparative study and examples. 2005. 14p. (RT-MAE-2005-10)

2005-11 - LACHOS, V.H., BOLFARINE, H., VILCA, F., GALEA, M. Estimation and influence diagnostics for structural comparative calibration models under the skew-normal distributions. 2005. 24p. (RT-MAE-2005-11)

2005-12 - PEREIRA, B.B., PEREIRA, C.A.B. A likelihood to diagnostic tests in clinical medicine. 2005. 23p. (RT-MAE-2005-12)

2005-13 - BUENO, V.C. A note on component reliability importance through parallel improvement. 2005. 9p. (RT-MAE-2005-13)

2005-14 - SILVA, G.T., SEN, P.K., LIMA, A. C. P. Estimation of the mean quality adjusted survival using a multistate model for the sojourn times. 2005. 9p. (RT-MAE-2005-14)

2005-15 - PEREIRA, C.A., STERN, J.M., WECHSLER,S. Can a significance test be genuinely Bayesian? 2005. 27p. (RT-MAE-2005-15)

2005-16 - CYSNEIROS, A.H.M.A., FERARRI, S.L.P. An improved likelihood ratio test for varying dispersion in exponential family nonlinear models. 2005. 12p. (RT-MAE-2005-16)

2005-17 - FERRARI, S.L.P., SILVA, M.F., CRIBARI-NETO, F. Adjusted profile likelihoods for the Weibull shape parameter. 2005. 18p. (RT-MAE-2005-17)

2005-18 -BRANCO, M. D.,RODRIGUEZ,C.B. Bayesian Inference for the skewness parameter on skew normal distribution. 2005. 21p. (RT-MAE-2005-18)

2005-19 - AZEVEDO, C. L. N., NOBRE, J. S., O princípio da equivariância: conceitos e aplicações. 2005. 32p. (RT-MAE-2005-19)

2005-20 - BUENO, V.da C., A Series representation of a coherent system. 2005. 15p. (RT-MAE-2005-20)

2005-21 -BUENO, V.da C., Hot standby redundancy allocation in a k-out-of-n:f systems of dependent components. 2005. 14p. (RT-MAE-2005-21)

2005-22 -BUENO, V.da C., CARMO, I.M. do Active redundancy allocation for a k-out-of-n:F system under a point process transform. 2005. 14p. (RT-MAE-2005-22)

2005-23 - MONTENEGRO, L.C.; BOLFARINE, H., LACHOS, V.H. Inference and influence diagnostics in the grubbs model under the skew-normal distribution. 2005. 27p. (RT-MAE-2005-23)

2005-24 - REIGADA, S. M. B .; ELIAN, S.N. Diagnóstico em análise discriminante. 2005. 25p. (RT-MAE-2005-24)

2005-25 - BUENO, V. Da C.; CARMO, I.M. Active redundancy allocation for a k-out-of-n: F system using not identically spares. 2005. 14p. (RT-MAE-2005-25)

2005-26 - BUENO, V. da C. Hot standby redundancy allocation for a k-ou-of-n:F system under a point process transform. 2005. 12p.

2005-27 - PEREIRA, C. A.da B. Conceitos estatísticos: Reflexões. 2005. 22p.

2005-28 - SILVA, G. T. da; LIMA, A. C. P. de A bootstrap evaluation of Wald test for the variance of random effects in survival analysis.2005. 14p.

2005-29 - LACHOS, V. H., MONTENEGRO, L.C., BOLFARINE, H. Inference and Local Influence in the Null Intercept Measurement Error Model Under the Skew-Normal Distribution. 2005. 24p.

The complete list of "Relatórios do Departamento de Estatística", IME-USP, will be sent upon request.

Departamento de Estatística
IME-USP
Caixa Postal 66.281
05311-970 - São Paulo, Brasil