Using Products and Powers of Products to Test Uniformity

Maria de Fátima Brilhante
Universidade dos Açores (DM), CEAUL
Rua da Mãe de Deus, Apartado 1422, 9501-801 Ponta Delgada, Portugal
E-mail: fbrilhante@uac.pt

Sandra Mendonça
Universidade da Madeira (CCEE), CEAUL
Campus Universitário da Penteada, 9000-390 Funchal, Portugal
E-mail: smfm@uma.pt

Dinis Pestana, Fernando Sequeira
Universidade de Lisboa, Faculdade de Ciências (DEIO), CEAUL
Bloco C6, Piso 4, Campo Grande, 1749-016 Lisboa, Portugal
E-mails: dinis.pestana@fc.ul.pt, fsequeira@fc.ul.pt

Abstract. *The family of probability density functions* $\{f_{X_m}(x) = \left(mx - \frac{m-2}{2}\right) \mathrm{I}_{(0,1)}(x),\ m \in [-2, 0]\}$ *of mixtures of* $Beta(1,2)$ *and of a standard uniform provides an interesting theoretical framework to assess uniformity, corresponding to* $m = 0$*, versus more proneness to generate values near 0. Computational sample augmentation techniques using this family have been investigated in the context of uniformity goodness-of-fit tests, when combining p-values for meta-analysing independent tests on some hypothesis, with unexpected results: power decreases with computational sample augmentation.*
We compute the probability density function of products and of random powers of products of variables from that family, and investigate their usefulness in testing uniformity. We show, using simulation, that those transformations also have a negative impact on power of uniformity goodness-of-fit tests. The Kakutani inner product of the measures under the null hypothesis and under the alternative hypothesis confirms increased colinearity and poorer fit. In fact, the techniques investigated pull the entropy towards 0, which is the entropy of the standard uniform random variable. This enlightens why testing uniformity can be rather misleading.

Keywords. Uniformity, mixtures, products and powers, Kakutani product, entropy.

1. Introduction

The uniform random model fits exactly the ideal of equiprobability, interesting from a philosophical viewpoint, but as such an asset of rather limited interest as a randomness model in the scope of decision procedures. But on the other hand, in view of what Rohatgi and Saleh (2001, p. 208) call the probability integral transform theorem — the fact that, under very broad conditions, $F_X(X) \frown Uniform(0,1)$, and that even under broader conditions $Y = F_X^{\leftarrow}(U) \frown F_X$, where U is the standard uniform random variable and $F_X^{\leftarrow}$ denotes the generalized inverse of F_X —, the uniform model became central in all areas using Computational Statistics. An important consequence of the probability integral transform theorem is that performing similar independent tests, under H_0 the set of p-values is a sample from the standard uniform population. Combining p-values is a tool in meta-analysis, but testing uniformity using small samples can be misleading. Gomes *et al.* (2009) presented several 'recipes' of computational sample augmentation, but with rather unexpected effects on the power of the goodness-of-fit test. Brilhante *et al.* (2010) developed further results along similar lines.

Gomes *et al.* (2009) used, namely, the family of random variables X_m with probability density function (pdf) $f_{X_m}(x) =$

$\left(mx - \frac{m-2}{2}\right) \mathrm{I}_{(0,1)}(x)$ — where $\mathrm{I}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$ is the indicator function. In other words, X_m has a mixture pdf, of the standard uniform, with weight $1+\frac{m}{2}$, and of a $Beta(1,2)$ random variable, with weight $-\frac{m}{2}$. Those mixtures, and namely the $Beta(1,2)$ model itself, are important in investigating the truthfulness of a null hypothesis, in meta-analysis, combining the p-values of independent tests since for $m \in [-2,0)$ they are more prone to generate small values, namely below the usual significance level, than the uniform model, corresponding to $m = 0$. Using the characterization that shows that $\min\left(\frac{X_m}{X_p}, \frac{1-X_m}{1-X_p}\right)$ is uniform and independent of X_p if and only if either X_m or X_p is standard uniform in case both variables have support (0,1), computational sample augmentation is straightforward. However, if the goal is to test uniformity, sample augmentation has a negative impact, since it actualy decreases power.

Herein we exploit an extension of the fact that if $U_1 \stackrel{\mathrm{d}}{=} U_2 \stackrel{\mathrm{d}}{=} U_3 \frown Uniform(0,1)$ are independent random variables, then $(U_1U_2)^{U_3} \frown Uniform(0,1)$, in the aim of investigating whether power divergence (an idea inspired by Cressie and Read, 1984) enhances power. In section 2 we prove that $(U_1U_2)^{X_\ell}$ is uniform iff $\ell = 0$, i.e. $X_\ell \frown Uniform(0,1)$, and that $(U_1U_2U_3)^{X_\ell}$ is uniform iff $\ell = -2$, i.e. $X_\ell \frown Beta(1,2)$. This inspired us to assess whether using products and powers of products is useful in testing uniformity.

Using simulation, we shall show that the answer is negative, and the Kakutani inner product of the measures of the test statistic under the null and under the alternative hypothesis confirms increased colinearity using the transformed variables. The explanation lies in the fact that the transformations used pull the entropy towards 0, which is the entropy of the standard uniform law.

In the sequel we shall need some functions related to the exponential integral, we transcribe their definition from Abramowitz and Stegun (1922, chap. 5):

$$\mathrm{Ei}(x) = -\fint_{-x}^{\infty} \frac{\mathrm{e}^{-t}}{t}\,\mathrm{d}t = \fint_{-\infty}^{x} \frac{\mathrm{e}^{t}}{t}\,\mathrm{d}t, \ x > 0.$$

$$\mathrm{li}(x) = \fint_{0}^{x} \frac{\mathrm{d}t}{\ln(t)} = \mathrm{Ei}(\ln(x)),$$

and for $n = 0, 1, \ldots; \mathrm{Re}(z) > 0$,

$$\mathrm{E}_n(z) = \int_1^{\infty} \frac{\mathrm{e}^{-zt}}{t^n}\,\mathrm{d}t,$$

$$\alpha_n(z) = \int_1^{\infty} t^n\,\mathrm{e}^{-zt}\,\mathrm{d}t =$$

$$= \frac{n!}{z^{n+1}}\,\mathrm{e}^{-z}\left(1 + z + \frac{z^2}{2} + \cdots + \frac{z^n}{n!}\right).$$

2. On products of random variables

In what follows, $U_1, U_2, \ldots$ denote independent and identically distributed (iid) standard uniform random variables.

Define also $U_n^* \frown Uniform(0, U_{n-1}^*)$, for $n = 1, 2, \ldots$, where $U_0^* = 1$. The pdf of U_n^* is

$$f_{U_n^*}(x) = \frac{(-\ln(x))^{n-1}}{(n-1)!}\,\mathrm{I}_{(0,1)}(x),$$

and therefore $U_n^* \stackrel{\mathrm{d}}{=} U_1U_2\cdots U_n$.

Most of our results stem out from a very humble and straightforward result:

Theorem

If T_α and $V_{\alpha+1}$ are independent random variables, $T_\alpha \frown Beta(\alpha,1)$ and $V_{\alpha+1} \frown Gamma(\alpha+1,1)$, then $T_\alpha V_{\alpha+1} \frown Gamma(\alpha,1)$.

In fact, $f_{T_\alpha V_{\alpha+1}}(x) =$

$$= \int_x^{\infty} \alpha \left(\frac{x}{y}\right)^{\alpha-1} \frac{y^\alpha\,\mathrm{e}^{-y}}{\Gamma(\alpha+1)}\,\frac{\mathrm{d}y}{y}\,\mathrm{I}_{(0,\infty)}(x)$$

$$= \frac{x^{\alpha-1}\,\mathrm{e}^{-x}}{\Gamma(\alpha)}\,\mathrm{I}_{(0,\infty)}(x).$$

Observations

1. From $f_{\frac{1}{X}}(x) = \frac{f_X\left(\frac{1}{x}\right)}{x^2}$, it follows that $T_\alpha \frown Beta(\alpha,1) \Longleftrightarrow \frac{1}{T_\alpha} \frown Pareto(\alpha)$.

Hence the above result may alternatively be stated as:

Let $V_\alpha \frown Gamma(\alpha,1)$ and $W_\alpha \frown Pareto(\alpha)$ be independent random variables. Then $\frac{V_{\alpha+1}}{W_\alpha} \overset{\mathrm{d}}{=} V_\alpha \frown Gamma(\alpha,1)$.

2. From $Z \frown Beta(p,q) \iff 1-Z \frown Beta(q,p)$ and the integral probability transform, $\exp(-V_1) \frown Uniform(0,1)$; therefore, from well known additive properties of the Gamma family, $\exp(-V_n) \overset{\mathrm{d}}{=} U_1U_2\cdots U_n \overset{\mathrm{d}}{=} U_n^*$, with the U_k's iid standard uniform random variables, has pdf

$$f_{U_n^*}(x) = f_{\exp(-V_n)}(x) = f_{U_1U_2\cdots U_n}(x) =$$

$$= \frac{(-\ln(x))^{n-1}}{(n-1)!}\mathrm{I}_{(0,1)}(x)\,.$$

Consequently,

$$\left(U_1U_2\cdots U_{n+1}\right)^{T_n} \overset{\mathrm{d}}{=} U_1U_2\cdots U_n,$$

and in particular,

$$(U_1U_2)^{U_3} = U \frown Uniform(0,1),$$

and hence

$$(U_1U_2)^{U_3}(U_4U_5)^{U_6}\cdots(U_{3n-2}U_{3n-1})^{U_{3n}} \overset{\mathrm{d}}{=}$$

$$\overset{\mathrm{d}}{=}\left(U_1U_2\cdots U_{n+1}\right)^{T_n}.$$

Let us now investigate the distribution of $W_{n,\ell} = (U_n^*)^{X_\ell}$. The joint pdf of $(U_n^*, W_{n,\ell})$ is $f_{U_n^*,W_{n,\ell}}(u,w) =$

$$= f_{U_n^*}(u) f_{X_\ell}\left(\frac{\ln(w)}{\ln(u)}\right)\frac{1}{|w\,\ln(u)|}\mathrm{I}_{\{0\le u\le w\le 1\}}(u,v),$$

and therefore, for $n\ge 3$, $w\in(0,1)$, $f_{W_{n,\ell}}(w) =$

$$\int_0^w \frac{(-\ln(u))^{n-1}}{(n-1)!}\left(\ell\,\frac{\ln(w)}{\ln(u)} - \frac{\ell-2}{2}\right)\frac{\mathrm{d}u}{w|ln(u)|} =$$

$$= -\frac{\ell\ln(w)}{w}\,\frac{(-\ln(w))^{n-2}}{(n-1)!}\,\alpha_{n-3}(-\ln(w)) -$$

$$-\frac{\ell-2}{2w}\,\frac{(-\ln(w))^{n-1}}{(n-1)!}\,\alpha_{n-2}(-\ln(w)) =$$

$$= \frac{(-\ln(w))^{n-1}}{(n-1)!\,w}\times$$

$$\times\left[\ell\,\alpha_{n-3}(-\ln(w)) - \frac{\ell-2}{2}\,\alpha_{n-2}(-\ln(w))\right].$$

For the two 'extreme' cases $\ell=-2$ and $\ell=0$, for $w\in(0,1)$, we get respectively

$$f_{W_{n,-2}}(w) = \frac{2(-\ln(w))^{n-1}}{(n-1)!\,w}\times$$

$$\times\left[\alpha_{n-2}(-\ln(w)) - \alpha_{n-3}(-\ln(w))\right]$$

and

$$f_{W_{n,0}}(w) = \frac{(-\ln(w))^{n-1}}{(n-1)!\,w}\,\alpha_{n-2}(-\ln(w)).$$

For $n=1,2,3$, $\ell\in[-2,0]$, $w\in(0,1)$, we get

$$f_{W_{1,\ell}}(w) = \frac{\ell}{w}\,\mathrm{E}_2(-\ln(w)) + \frac{\ell-2}{2w}\,\mathrm{li}(w)$$

$$f_{W_{2,\ell}}(w) = \frac{\ell\,\ln(w)\,\mathrm{li}(w)}{w} - \frac{\ell-2}{2}$$

$$f_{W_{3,\ell}}(w) = -\frac{\ell\,\ln(w)}{2} - \frac{\ell-2}{4}(1-\ln(w))$$

and thus in the 'extreme' cases the pdf, for $w\in(0,1)$, has the expression $f_{W_{n,\ell}}$ given in the table below:

n	$\ell=-2$	$\ell=0$
1	$-2\,\frac{\mathrm{E}_2(-\ln(w))+\mathrm{li}(w)}{w}$	$-\frac{\mathrm{li}(w)}{w}$
2	$2-\frac{2\,\ln(w)\,\mathrm{li}(w)}{w}$	1
3	1	$\frac{1-\ln(w)}{2}$

More generally, observe that the pdf of $W_{n,0} = (U_n^*)^{U_{n+1}}$ is, for $n=2,3,\ldots$, $w\in(0,1)$

$$f_{W_{n,0}}(w) = \int_0^w \frac{(-\ln(t))^{n-1}}{(n-1)!}\,\frac{\mathrm{d}t}{w\,|\ln(t)|} =$$

$$= \frac{(-\ln(w))^{n-1}}{(n-1)!\,w}\alpha_{n-2}(-\ln(w)) =$$

$$\frac{1}{n-1}\sum_{k=1}^{n-1}\frac{(-\ln(w))^{k-1}}{(k-1)!},$$

i.e., $f_{W_{n,0}}(w)$ is the balanced mixture, with equal weights $\varrho_k = \frac{1}{n-1}$, $k=1,\ldots,n-1$, of the pdf's $f_k(w) = \frac{(-\ln(w))^{k-1}}{(k-1)!}\mathrm{I}_{(0,1)}(w)$ of U_k^*.

Observe also that the distribution function (df) of $U_n^* \overset{\mathrm{d}}{=} U_1 U_2 \cdots U_n$ is

$$F_{U_n^*}(x) = \begin{cases} 0 & x < 0 \\ \displaystyle\sum_{k=1}^{n} \frac{(-\ln(x))^{k-1}}{(k-1)!}\, x & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

and the df of $(U_n^*)^{U_{n+1}} \overset{\mathrm{d}}{=} (U_1 U_2 \cdots U_n)^{U_{n+1}}$ is $F_{(U_n^*)^{U_{n+1}}}(x) =$

$$= \begin{cases} 0 & x < 0 \\ \displaystyle\sum_{k=1}^{n-1} \frac{n-k}{n-1}\, \frac{(-\ln(x))^{k-1}}{(k-1)!}\, x & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

We now investigate, for the tractable case of three independent random variables which are mixtures of the standard uniform with a $Beta(1,2)$, the distribution of $(X_m X_p)^{X_\ell}$.

If X_m and X_p are independent, the pdf of $X_m X_p$ is $f_{X_m X_p}(x) =$

$$= \Big(mpx(-\ln(x) + \frac{(m-2)\,(p-2)}{4}\,(-\ln(x)) -$$

$$- \frac{(m-2)p + m(p-2)}{2}\,(1-x) \Big)\, \mathrm{I}_{(0,1)}(x).$$

Then the pdf $f_{m,p.\ell}$ of $Z_{m,p,\ell} = (X_m X_p)^{X_\ell}$ is, for $x \in (0,1)$, $f_{m,p.\ell}(x) =$

$$= -\frac{(m-2)(p-2)(\ell-2)}{8} - \frac{mp(\ell-2)}{4}\, x +$$

$$+ \frac{[(m-2)p + m(p-2)](\ell-2)}{4}\, \frac{\mathrm{E}_1(-\ln(x)) - \mathrm{E}_1(-\ln(x^2))}{x} +$$

$$+ \frac{[(m-2)p + m(p-2)]\ell}{2}\, \frac{\mathrm{E}_2(-\ln(x^2)) - \mathrm{E}_2(-\ln(x))}{x} +$$

$$+ \frac{(m-2)(p-2)\ell}{4}\, \frac{(-\ln(x))\,\mathrm{E}_1(-\ln(x))}{x} +$$

$$+ mp\ell\, \frac{(-\ln(x))\,\mathrm{E}_1(-\ln(x^2))}{x},$$

with obvious simplifications when $m = 0$ and/or $p = 0$ and/or $\ell = 0$ (i.e., the corresponding random variable standard uniform), or two or three of the mixture weights m, p and ℓ are equal.

3. Testing uniformity

Fig. 1 below, corresponding to the power functions of the Kolmogorov–Smirnov test $H_0 : m = 0$ *vs.* $H_A : m \in [-2, 0)$ obtained by simulation (5 000 runs) for the special case $n = 12$, when using samples from X_m and from $Z_{0,0,m}$, respectively, is exemplar of what always happens: transformed data, using random powers of products in the family X_m decreases the Kolmogorov–Smirnov test power to reject a false null uniformity hypothesis.

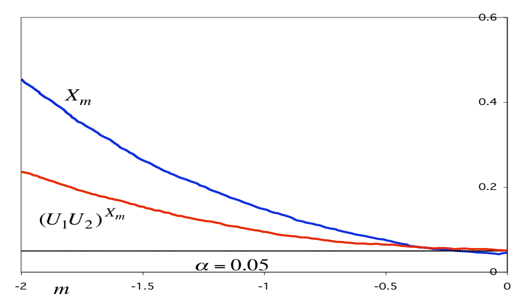

Figure 1: Power function, Kolmogorov–Smirnov test $H_0 : m = 0$ *vs.* $H_A : m \in [-2, 0)$, using samples of size $n = 12$ from X_m and from $Z_{0,0,m}$.

Let f_0 and f_A denote the pdf of some test statistic under H_0 and under H_A, respectively. The Kakutani inner product of the corresponding probability measures $\mathcal{P}_0$ and $\mathcal{P}_A$

$$H_q(\mathcal{P}_0, \mathcal{P}_A) = \int f_0^q\, f_A^{1-q}(x)\mathrm{d}x,\ \ q \in [0,1]$$

has been used by Chernoff (1952) to construct bounds for the probability of error, both of the first and of the second kind, cf. also Kraft (1955) and Hellman and Raviv (1970). Stuck (1976), comments that '*the Kakutani inner product can be thought of intuitively as the amount of 'colinearity' or 'overlap' of two probability measures, with the larger the Kakutani inner product the larger the 'overlap'*"

Observe that when testing $H_0 :\ m = 0$ *vs.* $H_A :\ m \in [-2, 0)$

$$H_q(\mathcal{P}_0, \mathcal{P}_A) = \frac{(2-m)^{2-q} - (2+m)^{2-q}}{2^{2-q}(q-2)m}\,.$$

whose mnimum is 0.942 On the other hand, the minimum of the Kakutani inner product when using the Kolmogorov–Smirnov for testing uniformity goodness of fit when the alternative is $Z_{m,p.\ell}$ is 0.972, showing increased overlap.

4. Conclusions

Although testing $H_0 : m = 0$ (uniformity) *vs.* $H_A : m \in [-2, 0)$ in the X_m family seems an interesting tool in assessing the goodness of combining p-values in meta-analysis, our simulation results indicate that testing uniform goodness-of-fit has some unexpected features. Computational sample augmentation doesn't increase power, contrarily to what is the natural naïve hope for dealing with small samples, and the same happens with random powers of products divergence statistics. On the other hand goodness-of-fit using the original sample cannot be very good, since the Kakutani inner product reveals that there is a considerable dose of overlap of the measures under H_0 and under H_A, cf. Fig. 2.

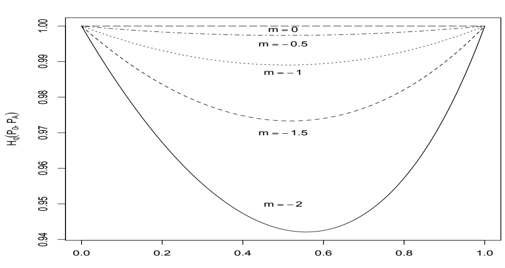

Figure 2: Kakutani inner product.

The entropy of X_m, $m \in [-2, 0]$ is

$$H(X_m) = -\int_0^1 \left(mx + \frac{m-2}{2}\right) \ln\left(mx + \frac{m-2}{2}\right) \mathrm{d}x$$

$$= 0.5 + \ln(2) + \frac{\ln\left(\left(\frac{2-m}{2+m}\right)^m\right)}{8} - \frac{\ln(4-m^2)}{2} + \frac{\ln\left(\frac{2-m}{2+m}\right)}{2m},$$

cf. Fig. 3 (for a detailed study of entropy, cf. Johnson, 2004).

Hence, as $\min\left(\frac{X_m}{X_p}, \frac{1-X_m}{1-X_p}\right) \overset{\mathrm{d}}{=} X_{\frac{mp}{6}}$, its entropy is always nearer the entropy of the standard uniform than the entropy of X_m

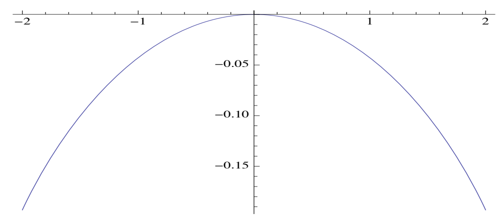

Figure 3: Entropy of X_m, $m \in [-2, 2]$.

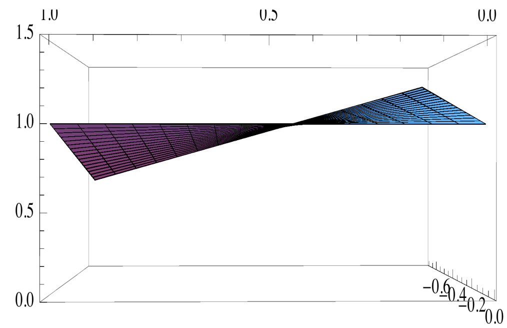

Figure 4: Probability density functions of $\min\left(\frac{X_m}{X_p}, \frac{1-X_m}{1-X_p}\right)$, $m, p \in [-2, 0]$.

or X_p. Fig. 4 shows that in the appropriate range $\frac{mp}{6} \in [-\frac{2}{3}, 0]$ the densities of $\min\left(\frac{X_m}{X_p}, \frac{1-X_m}{1-X_p}\right)$ are indeed quite near the uniform probability density function.

The same holds true for the probability density functions of $Z_{m,m,m}$, $m \in [-2, 0]$. In Fig. 5 we show the corresponding probability density functions:

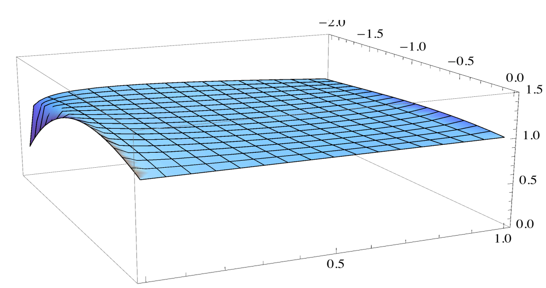

Figure 5: Probability density functions of $Z_{m,m,m}$, $m \in [-2, 0]$.

and in In Fig. 6 we show that in the range $m \in [-1, 0]$ those densities are almost indistinguishable of the standard uniform density:

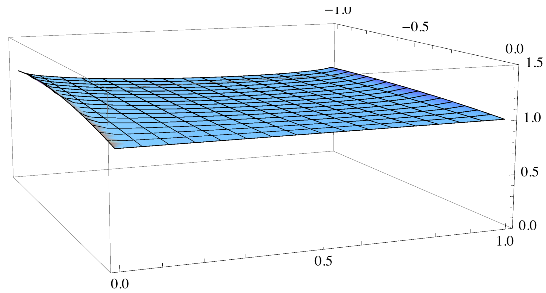

Figure 6: Probability density functions of $Z_{m,m,m}$, $m \in [-1,0]$.

Fig. 7 below

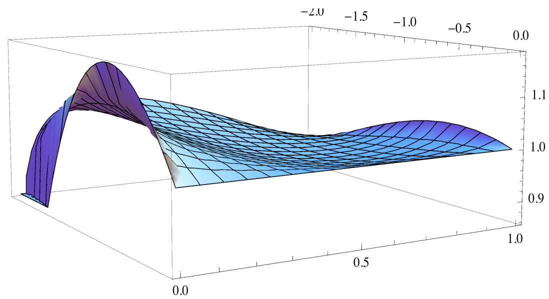

Figure 7: Probability density functions of $Z_{m,m,m}$, $m \in [-2,0]$.

is similar to Fig. 5, but the smaller plot range enhances the details, showing that for values of m near -2 there is indeed more chances of distinguishing $Z_{m,m,m}$ from $Z_{0,0,0} = U$. Note however that the plot range lies in [0.8,1.2].

In Johnson (2004) it is shown that the normal law has maximum entropy in the family of random variables with finite variance with support in the real line — and this is the basis of Linnik's proof of the central limit theorem — and that the exponential has maximum entropy in the family of random variables with finite expectation whose support is a positive half-line (and we point out that this is the reason why Rényi's rarefaction has exponential limit).

Our results show that the extremal properties of the standard uniform entropy, in the family of random variables with support in a positive segment, lead to a similar status of the standard uniform: it serves as an excellent approximation for a wide variety of models, as those that we have investigated

References

[1] Abramowitz M, Stegun I A. Handbook of Mathematical Functions. New York: Dover; 1992.

[2] Brilhante M F, Pestana D, Sequeira F. Combining p-values and random p-values, Lisboa, Notas Comunic. CEAUL 2010.

[3] Chernoff H A measure of the asymptotic efficiency for tests of a hypothesis based on a sum of observations. Annals of Mathematical Statistics 1952; 23: 493–507.

[4] Cressie N, Read T R C. Multinomial goodness-of-fit tests, Journal Royal Statistical Society Series B: 1984; 46: 440–464.

[5] Gomes M I, Pestana D, Sequeira F, Mendonça S, Velosa S. Uniformity of offsprings from uniform and non-uniform parents. In Luzar-Stiffler, V., Jarec, I. and Bekic, Z. (eds.), Proceedings of the ITI 2009, 31th International Conference on Information Technology Interfaces 2009; 243–248.

[6] Hellman M E, Raviv J. Probability of error, equivocation, and the Chernoff bound. IEEE Transactions Information Theory 1970; 16: 368–372.

[7] Johnson O. Information Theory and the Central Limit Theorem, London: Imperial College Press; 2004.

[8] Kraft CH. Some conditions for consisteency and uniform consistency of statistical procedures. Univ. California (Berkeley) Publ. Statistics 1955; 2: 125–142.

[9] Linnik Y. An information–theoretic proof of the central limit theorem. Theory of Probability Applications 4: 288–299.

[10] Rényi A. A characterization of the Poisson process, MTA Mat. Kut. Int. K?ozl. 1956: 1, 519Ð527.

[11] Rohatgi V K, Saleh A K Md E. An Introduction to Probability and Statistics. New York: Wiley; 2001.

[12] Stuck B W. Distinguishing stable probability measures, I: Discrete time. Bell System Tech.łJ. 1976; 55: 1125–1182.

Acknowledgements

Research partially suported by FCT/OE.