Aplicações das funções geradoras de probabilidade a variáveis aleatórias reais

Manuel L. Esquível
Departamento de Matemática, FCT/UNL e CMAF/UL

Resumo: Após referir trabalho prévio sobre as funções geradoras de probabilidade para variáveis aleatórias reais estendemos a estas variáveis aleatórias a lei uniforme dos grandes números e um teorema de limite central funcional para as funções geradoras de probabilidade empíricas. Apresentamos uma aplicação ao estudo de leis contínuas, nomeadamente a estimação de parâmetros da lei gaussiana e da lei gama por via de um estimador de mínimo de contraste construído a partir da função geradora de probabilidade empírica da amostra. Testamos o procedimento introduzido por simulação e demonstramos a consistência do estimador estudado.

Palavras–chave: funções geradoras de probabilidade, leis empíricas, estimação de parâmetros de leis contínuas.

Abstract: After recalling previous work on probability generating functions for real valued random variables we extend to these random variables uniform laws of large numbers and functional limit theorem for the empirical probability generating function. We present an application to the study of continuous laws, namely, estimation of parameters of Gaussian and gamma laws by means of a minimum contrast estimator that uses the empirical probability generating function of the sample. We test the procedure by simulation and we prove the consistency of the estimator.

Keywords: probability generating function, empirical laws, estimation of parameters of continuous laws.

Classificação MSC2000: 60–08, 60E10, 30B50.

1 Introdução

As técnicas de cálculo associadas às funções geradoras de probabilidade são usualmente utilizadas para estudar variáveis aleatórias discretas tomado valores inteiros. Recentemente, vários trabalhos desenvolveram aplicações interessantes destas técnicas a um leque alargado de objectivos como sejam a análise gráfica preliminar de dados, estimação, testes, etc (vejam-se as referências para uma escolha destes trabalhos).

Um estudo detalhado das condições de existência de funções geradoras de probabilidade para variáveis aleatórias discretas torna possível estender às variáveis aleatórias tomando valores reais os métodos disponíveis para as variáveis aleatórias tomando valores inteiros. Neste trabalho mostramos que leis contínuas admitindo funções geradoras de momentos não triviais (tais como as lei gaussiana e a lei gama) podem ser estudadas com o auxílio das funções geradoras de probabilidade.

2 A fgp para variáveis aleatórias discretas reais

No que vai seguir-se X é uma variável aleatória discreta tomando como valores a sucessão $(\alpha_k)_{k\in\mathbb{Z}}$ de valores reais convencionando-se, sem perda de generalidade, que $\alpha_k < 0$ para $k < 0$, $\alpha_0 = 0$ e $\alpha_k > 0$ para $k > 0$.

Para uma sucessão de números não negativos $(p_k)_{k\in\mathbb{Z}}$ tais que $\sum_{k=-\infty}^{+\infty} p_k = 1$, as probabilidades, temos que $\mathbb{P}[X = \alpha_k] = p_k$. A função geradora de probabilidade (fgp) de X é

$$\psi_X(t) = \mathbb{E}[t^X] = \sum_{k=-\infty}^{+\infty} p_k t^{\alpha_k}$$

para $t > 0$. O domínio natural de definição da fgp, $\mathbb{D}_X = \{t > 0 : \psi_X(t) < +\infty\}$ pode ser facilmente descrito como no resultado seguinte (veja-se [3]).

Teorema 2.1 *Seja X uma variável aleatória e seja ψ_X a sua fgp. Temos então que:*

1. *Se X toma um número finito de valores reais*
$$\mathbb{D}_X =]0, +\infty[\ .$$
2. *Se X toma um número infinito de valores reais sem pontos de acumulação*
$$\exists u_0, v_0 \in]-\infty, 0]\ \ ,\ \ ]e^{u_0}, e^{-v_0}[\subset \mathbb{D}_X \subset [e^{u_0}, e^{-v_0}]\ . \tag{1}$$
3. *Se X é uma variável aleatória discreta com caudas com decaímento exponencial, isto é, se para alguns $k, c > 0$ temos que $\mathbb{P}[\mid X \mid > x] \leq ke^{-cx}$ então, temos também a condição expressa pela fórmula (1).*

A fgp caracteriza a distribuição da variável aleatória. Duas variáveis aleatórias X and Y têm a mesma distribuição se e só se as respectivas fgp coincidem numa vizinhança de 1. Tem-se também um teorema de tipo teorema de Lévy para as funções características. Se uma sucessão de variáveis aleatórias $(X_n)_{n\in\mathbb{N}}$ tem a correspondente sucessão de fgp $(\psi_n)_{n\in\mathbb{N}}$ convergindo para ψ_X numa vizinhança de 1 então a sucessão converge em distribuição para X.

A fgp de uma variável aleatória discreta tomando valores inteiros é muito útil para calcular as leis de probabilidade de somas de variáveis aleatórias independentes deste tipo sobretudo, porque para variáveis aleatórias tomando valores inteiros tem-se que $p_k = \psi_X^{(n)}(0)/n!$. Para variáveis aleatórias gerais (não necessariamente a valores inteiros), os programas que permitem o cálculo simbólico tornam possível o mesmo tipo de cálculos de um modo prático justificando assim o estudo das fgp para estas variáveis aleatórias.

Observação 2.2 *Para uma variável aleatória contínua com lei μ_X pode acontecer que o conjunto $\mathbb{D}_X := \{t > 0 : \int_{\mathbb{R}} t^x d\mu_X(x) < +\infty\}$ tenha um interior não vazio. Tal acontece por exemplo para as variáveis aleatórias com lei gaussiana ou com lei gama. Neste caso, usamos ainda $\psi_X(t) = \mathbb{E}[t^X]$ para $t \in \mathbb{D}_X$.*

3 Sobre o estimador empírico da fgp

Nesta secção mostramos como usar uma amostra de uma variável aleatória para estimar a fgp desta variável aleatória. Os resultados apresentados são extensões ao caso de variáveis aleatórias tomando valores reais de resultados já conhecidos para variáveis aleatórias discretas tomando valores inteiros (veja-se [12]).

Seja $(X_n)_{n\in\mathbb{N}}$ uma amostra de uma variável aleatória X tendo como lei μ_X e fgp $\psi_X(t) = \mathbb{E}[t^X]$ definida para $t \in \mathbb{D}_X$. Defina-se também a fgp empírica (fgpe) por:

$$\forall t > 0 \;\; \overline{\psi}_{X,n}(t) = \frac{1}{n}\sum_{i=1}^{n} t^{X_i} \; . \tag{2}$$

Como se tem que para todo o $t \in \mathbb{D}_X$ que $\mathbb{E}[\overline{\psi}_{X,n}(t)] = \psi_X(t)$ então tem-se que $(\overline{\psi}_{X,n}(t))_{n\in\mathbb{N}}$ é uma sucessão de estimadores não enviesados de $\psi_X(t)$. Uma aplicação trivial da lei forte dos grandes números mostra-nos que temos para todo o $t \in \mathbb{D}_X$ a consistência forte do estimador, isto é, quase certamente $\lim_{n\to+\infty} \overline{\psi}_{X,n}(t) = \psi_X(t)$.

Como consequência, uma aplicação fácil do teorema do limite central mostra que para todo o $t \in \mathbb{D}_X$, a sucessão $(n^{1/2}(\overline{\psi}_{X,n}(t) - \psi_X(t)))_{n\geq 1}$ converge em distribuição para $\mathcal{N}(0, \sqrt{\psi_X(t^2) - \psi_X(t)^2})$. Podem-se obter, também, uma lei uniforme dos grandes números e um teorema do limite central funcional, como veremos adiante.

Na demonstração do resultado seguinte, a lei uniforme dos grandes números para as fgpe, seguimos a ideia geral de [9] mas aplicando o lema de Fatou inverso em vez do teorema da convergência dominada o qual, pelo menos no nosso caso, não nos permite efectuar a demonstração. Pode-se encontrar um resultado deste tipo para as funções geradoras de momentos em [4].

Teorema 3.1 *Seja $[a, b] \subset \mathbb{D}_X \neq \emptyset$. Então, quase certamente, temos*

$$\lim_{n\to+\infty} \sup_{t\in[a,b]} \left|\overline{\psi}_{X,n}(t) - \psi_X(t)\right| = 0 \; .$$

Demonstração. Pela lei forte dos grandes números temos quase certamente

$$\forall k \in \mathbb{Z} \;\; p_k = \lim_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{X_i=\alpha_k\}} \; . \tag{3}$$

Considere-se agora ω_0 neste conjunto de probabilidade plena. Decompondo a soma e observando, como é hábito que $\alpha_k > 0$ para $k \geq 1$, $\alpha_k < 0$ para $k < -1$

e $\alpha_0 = 0$, tem-se

$$\sup_{t\in[a,b]} \left|\overline{\psi}_{X,n}(t,\omega_0) - \psi_X(t)\right| \leq \sup_{t\in[a,b]} \sum_{k=-\infty}^{+\infty} t^{\alpha_k} \left|p_k - \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{X_i=\alpha_k\}}(\omega_0)\right| =$$
$$= \sum_{k=1}^{+\infty} b^{\alpha_k} \left|p_k - \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{X_i=\alpha_k\}}(\omega_0)\right| + \left|p_0 - \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{X_i=0\}}(\omega_0)\right|$$
$$+ \sum_{k=-1}^{-\infty} a^{\alpha_k} \left|p_k - \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{X_i=\alpha_k\}}(\omega_0)\right| .$$

Iremos mostrar de seguida que o limite dos somatórios em k à direita é zero quando a dimensão da amostra n tende para o infinito. Trataremos apenas do primeiro somatório dado que para o segundo termo o resultado é trivial pela fórmula 3 e para o segundo somatório a demonstração é semelhante à que iremos apresentar para o primeiro somatório. Para este fim utilizaremos o lema de Fatou inverso.

Defina-se $f_n(k)$ e $g_n(k)$ por

$$f_n(k) := b^{\alpha_k} \left|p_k - \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{X_i=\alpha_k\}}(\omega_0)\right| \leq b^{\alpha_k} \left(p_k + \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{X_i=\alpha_k\}}(\omega_0)\right) =:$$
$$=: g_n(k) .$$

Observe-se que para todo o $n \in \mathbb{N}$,

$$\sup_{m\geq n} \frac{1}{m}\sum_{i=1}^{m} \mathbb{I}_{\{X_i=\alpha_k\}}(\omega_0) \leq \sup_{m\geq 1} \frac{1}{m}\sum_{i=1}^{m} \mathbb{I}_{\{X_i=\alpha_k\}}(\omega_0) \leq$$
$$\leq \begin{cases} p_k \text{ ou} \\ (1/m_1)\sum_{i=1}^{m_1} \mathbb{I}_{\{X_i=\alpha_k\}}(\omega_0) , \end{cases}$$

onde m_1 é tal que

$$\sup_{m\geq 1} (1/m) \sum_{i=1}^{m} \mathbb{I}_{\{X_i=\alpha_k\}}(\omega_0) = (1/m_1) \sum_{i=1}^{m_1} \mathbb{I}_{\{X_i=\alpha_k\}}(\omega_0) .$$

Sendo μ_c a medida de contagem sobre $\mathbb{Z}$, temos para todo o $n \in \mathbb{N} - \{0\}$:

$$\sum_{k=1}^{+\infty} \sup_{m \geq n} f_m(k) = \int_{\mathbb{N}-\{0\}} \sup_{m \geq n} f_m(k) d\mu_c(k) \leq \int_{\mathbb{N}-\{0\}} \sup_{m \geq n} g_m(k) d\mu_c(k) \leq$$

$$\leq \sum_{k=1}^{+\infty} b^{\alpha_k} p_k + \max\left(\sum_{k=1}^{+\infty} b^{\alpha_k} p_k, \sum_{k=1}^{+\infty} (b^{\alpha_k} \frac{1}{m_1} \sum_{i=1}^{m_1} \mathbb{I}_{\{X_i = \alpha_k\}}(\omega_0)) \right) =$$

$$= \psi_X(b) + \max\left(\psi_X(b), \frac{1}{m_1} \sum_{i=1}^{m_1} b^{\alpha_{k(i,\omega_0)}} \right) < +\infty \ ,$$

observando que para ω_0 dado e $i = 1, \ldots, m_1$ existe apenas um $k = k(i, \omega_0)$ tal que se verifica $\mathbb{I}_{\{X_i = \alpha_{k(i,\omega_0)}\}}(\omega_0) \neq 0$. Podemos agora aplicar o lema de Fatou inverso com toda a segurança para concluir que

$$0 \leq \limsup_{n \to +\infty} \sum_{k=1}^{+\infty} b^{\alpha_k} \left| p_k - \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\{X_i = \alpha_k\}}(\omega_0) \right| \leq$$

$$\leq \sum_{k=1}^{+\infty} \limsup_{n \to +\infty} \left| p_k - \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\{X_i = \alpha_k\}}(\omega_0) \right| = 0 \ ,$$

tal como se pretendia. □

Observação 3.2 *O teorema permanece válido sob a hipótese mais fraca de se ter $(X_n)_{n \in \mathbb{N}}$ sucessão estacionária e ergódica dado que nesse caso se tem ainda a fórmula 3.*

Também dispomos de um princípio de invariância para a fgpe.

Teorema 3.3 *A sucessão de processos $(n^{1/2}(\overline{\psi}_{X,n}(t) - \psi_X(t))_{n \geq 1}$ no espaço das funções contínuas com a norma uniforme converge fracamente para um processo Gaussiano tendo média zero e covariância $\psi_X(st) - \psi_X(s)\psi_X(t)$ num qualquer intervalo fechado contido em $\mathbb{D}_X/2$.*

Demonstração. Este resultado pode ser deduzido directamente do teorema 2.3 em [4] onde o resultado é enunciado e demonstrado para a função geradora de momentos $\tilde{\mu}_X(t) = \mathbb{E}[e^{tX}]$, observando que $\psi_X(t) = \tilde{\mu}_X(\ln(t))$. □

4 Aplicações da fgp

Os resultados na secção precedente permitem-nos contemplar o estudo de variáveis aleatórias discretas e reais da mesma forma em que é usualmente feito na literatura, sobre este assunto, para as variáveis aleatórias discretas com valores inteiros.

4.1 Aplicações às variáveis aleatórias discretas

Como consequência dos resultados na secção 3 podemos aplicar às variáveis aleatórias discretas tomando valores reais quaisquer, os processos de estimação e teste desenvolvidos para variáveis aleatórias discretas tomando valores inteiros. Veja-se [10] para uma panorâmica completa destas técnicas e por exemplo os trabalhos [5], [7], [4], [17], [11], [12], [2], [16] e finalmente [15], para estudos detalhados de alguns métodos estatísticos particulares. A análise detalhada destes métodos para exemplos relevantes de variáveis aleatórias discretas reais está ainda por fazer.

4.2 Aplicações a variáveis gerais

Nesta subsecção mostramos como a fgp pode ser utilizada para estudar a lei de uma variável não necessariamente discreta. A ideia que justifica esta aproximação (veja-se [14, p.131]) é que estamos fisicamente limitados a observar apenas um número finito (e por isso discreto) de valores que uma variável aleatória toma. Por isso, justifica-se afirmar que só podemos *apreender* realmente variáveis aleatórias discretas.

A medida empírica definida na forma usual a partir de uma amostra de uma dada variável aleatória, veja-se [6], aproxima a lei dessa variável aleatória sob todos os aspectos desejáveis (lei forte uniforme, teorema do limite central functional, etc) num espectro variado de situações. Esta medida empírica é a ferramenta apropriada para variáveis aleatórias com valores vectoriais.

O resultado simples seguinte mostra que qualquer lei em $\mathbb{R}$ pode ser aproximada por uma família de leis de variáveis aleatórias discretas, construída como um histograma.

É bem conhecido que qualquer medida finita num espaço localmente compacto pode ser aproximada por uma sucessão de combinações lineares de medidas de Dirac (veja-se [8, p. 99]).

Numa linha de pensamento paralela, um resultado clássico mostra que o histograma, construído sobre uma amostra de uma dada variável aleatória é uma função em escada (aleatória) convergindo em probabilidade para a densidade, em cada ponto de continuidade desta densidade (veja-se [13, p. 367]). Tendo em mente estes resultados é natural pensar-se que a distribuição de uma variável aleatória pode ser aproximada por uma sucessão de de distribuições aleatórias construídas sobre a amostra.

Teorema 4.1 *Seja para cada $n \in \mathbb{N}$, $(I_k^n)_{k\in\mathbb{Z}}$ uma partição da recta real tal que se verifique, $\lim_{n\to+\infty} max_{k\in\mathbb{N}} |I_k^n| = 0$ e seja para cada $n \in \mathbb{N}$, $(\alpha_k^n)_{k\in\mathbb{Z}}$ a sucessão das extremidades à esquerda dos intervalos da partição de ordem n. Seja $(X_n)_{n\in\mathbb{N}}$ uma amostra de uma variável aleatória X tendo como lei μ_X. Defina-se*

$$\overline{\mu}_{N,n} := \sum_{k\in\mathbb{Z}} \frac{\#\{i : X_i \in I_k^n\}}{N} \delta_{\alpha_k^n} \ . \tag{4}$$

Então, $(\overline{\mu}_{N,n})_{N,n\in\mathbb{N}}$ é uma família de medidas de probabilidade aleatórias convergindo estreitamente em probabilidade para μ_X, isto é, para cada f contínua e limitada

$$\lim_{N,n\to+\infty} \overline{\mu}_{N,n}(f) = \mu_X(f) \ .$$

Demonstração. Um cálculo simples mas entediante mostra que para cada f contínua e limitada se tem

$$\mathbb{E}[\overline{\mu}_{N,n}(f)] = \sum_{k\in\mathbb{Z}} \mu_X(I_k^n) f(\alpha_k^n)$$

$$\mathbb{V}[\overline{\mu}_{N,n}(f)] = \frac{1}{N}\left[\sum_{k\in\mathbb{Z}} \mu_X(I_k^n) f^2(\alpha_k^n) - \left(\sum_{k\in\mathbb{Z}} \mu_X(I_k^n) f(\alpha_k^n)\right)^2\right] \ .$$

Como consequência, pela definição usual do integral de Stieltjes temos que $\lim_{n\to+\infty} \mathbb{E}[\overline{\mu}_{N,n}(f)] = \mu_X(f)$ e $\lim_{n\to+\infty} N \times \mathbb{V}[\overline{\mu}_{N,n}(f)] = N \times (\mu_X(f^2) - \mu_X(f)^2)$ pelo que o resultado anunciado fica demonstrado. □

4.3 Estimação de parâmetros com a fgp

Introduzimos seguidamente um método de estimação para parâmetros de variáveis aleatórias contínuas baseado na fgpe. Trata-se de uma técnica de um tipo habitualmente tido como útil apenas para variáveis aleatórias discretas. Descrevemos primeiramente o método, depois apresentamos um protocolo para testar o método e finalmente, demonstramos a consistência do estimador em dois casos distintos.

1. Considere-se uma variável aleatória X tendo como lei $\mu_{X(\theta)}$ onde θ é um parâmetro desconhecido da lei da variável aleatória onde, para maior simplicidade, o parâmetro está definido num certo conjunto compacto $\Theta \subset \mathbb{R}$. Suponha-se que para cada $\alpha \in \Theta$ a fgp $\psi_{X(\alpha)}$ está bem definida num conjunto $\mathbb{D}_{X(\alpha)}$ com interior não vazio.

2. Observe-se uma amostra de X e considere-se $\overline{\psi}_{X,n}$ a fgpe baseada na amostra.

3. Considere-se um conjunto de pontos $t_1, t_2, \ldots, t_M$ em $\bigcap_{\alpha\in\Theta} \mathrm{Int}\mathbb{D}_{X(\alpha)}$ que supomos não vazio e defina-se o contraste

   $$O_n(\alpha) := \sum_{i=1}^{M} \left(\overline{\psi}_{X,n}(t_i) - \psi_{X(\alpha)}(t_i)\right)^2 \ ,$$

   e o estimador do mínimo de contraste $\overline{\theta}_n$ do parâmetro desconhecido θ tal que

   $$O(\overline{\theta}_n) = \min\{O_n(\alpha) : \alpha \in \Theta\} \ .$$

Para testar este método por simulação propomos o protocolo seguinte.

- Passo 1: Escolha-se um valor para o parâmetro desconhecido θ. Faça-se $j = 1$. Escolha-se r o número de repetições da simulação.
- Passo 2: Simule-se uma amostra de $X(\theta)$.
- Passo 3: Determine-se pelo método descrito acima, $\overline{\theta}_{1,j}$ um valor estimado de θ mas também por um método standard e conhecido $\overline{\theta}_{2,j}$ um outro valor estimado de θ. Se $j < r$ incremente-se j e volte-se ao Passo 2. Se $j = r$ vá-se para o Passo 4.
- Passo 4: Calcule-se a média e o desvio padrão das famílias de estimativas $U = (\overline{\theta}_{1,j})_{j=1,\dots,r}$ e $V = (\overline{\theta}_{2,j})_{j=1,\dots,r}$ e comparem-se os métodos comparando estas médias e estes desvios padrões.

Apresentamos seguidamente uma aplicação deste protocolo para o teste do algoritmo em duas situações diferentes.

Para a lei gaussiana $\mathcal{N}(\theta, 1)$, isto é, com média igual a θ e desvio padrão igual a 1, estima-se θ pelo método habitual considerando para cada repetição j da simulação, $\theta_{2,j}$ dado pela média aritmética da amostra. Os pontos $t_1, \dots, t_M$ foram escolhidos ao acaso próximos de 1, mais precisamente, $t_1 = .8, t_2 = .85, t_3 = .9, t_4 = .95, t_5 = .98, t_6 = 1.05, t_7 = 1.08, t_8 = 1.09, t_9 = 1.1, t_{10} = 1.2$. Os resultados, para um valor objectivo $\theta = 2$, figuram na tabela 1.

Tabela 1: Resultados para a lei gaussiana.

r	Média U	Desv. Pad. U	Média V	Desv. Pad. V
10	1.99842	0.246679	1.99985	0.2467220
50	1.99401	0.149717	1.99484	0.1498410
100	2.00760	0.099236	2.00748	0.0996039
500	1.99908	0.052912	1.99890	0.0520873
1000	2.00133	0.032009	2.00124	0.0319557

Na situação seguinte consideramos a distribuição gama com parâmetros λ e α tendo uma densidade dada por $\mathcal{G}(\lambda, \frac{1}{\alpha}) = \frac{\alpha^\lambda}{\Gamma(\lambda)} e^{-\alpha x} x^{\lambda-1}$. É fácil ver que se $X = X(\lambda, \alpha) \in \mathcal{G}(\lambda, 1/\alpha)$ então para $t \in \mathbb{D}_{X(\lambda,\alpha)} =]0, e^\alpha[$ tem-se que $\psi_{X(\lambda,\alpha)}(t) = \mathbb{E}[t^X] = \frac{\alpha^\lambda}{(\alpha-\ln(t))^\lambda}$ Como $\mathbb{E}[X(\lambda, \alpha)] = \lambda/\alpha$ e $\mathbb{V}[X(\lambda, \alpha)] = \lambda/\alpha^2$ uma forma trivial para estimar α a partir de uma amostra consiste em calcular $\mathbb{E}[X(\lambda, \alpha)]/\mathbb{V}[X(\lambda, \alpha)]$. Usámos o protocolo definido acima para estimar $\alpha = 2$ com $\lambda = 2.1$ e com os pontos $t_1, \dots, t_M$ escolhidos tal como no exemplo anterior. Os resultados obtidos são apresentados na tabela 2.

Podemos imediatamente extrair uma conclusão preliminar. Para as amostras consideradas, o estimador introduzido comporta-se de forma comparável ao do

Tabela 2: Resultados para a lei gama.

r	Média U	Desv. Pad. U	Média V	Desv. Pad. V
10	2.14098	0.5723900	2.95111	1.946780
50	2.00263	0.2076540	2.12457	0.561632
100	2.02520	0.1251470	2.17005	0.334930
500	1.99878	0.0606117	2.03207	0.190385
1000	1.99331	0.0470203	1.99880	0.123071

estimador usual da média no caso Gaussiano e, comporta-se melhor no caso da distribuição gama relativamente ao estimador de α dado pela razão da média sobre a variância da amostra.

Mostremos agora que o estimador do mínimo de contraste usado nestes dois exemplos é consistente. Primeiramente, para o caso Gaussiano.

Teorema 4.2 *Seja $X = X(\theta) \in \mathcal{N}(\theta,\sigma)$ onde σ é dado e onde $\theta \in \Theta \subset \mathbb{R}$ é um conjunto compacto. Seja $(X_n)_{n\in\mathbb{N}}$ uma amostra de X e $\overline{\psi}_{X,n}(t) = (1/n)\sum_{i=1}^{n} t^{X_i}$ definido para $t > 0$. Para os pontos dados $t_1,\ldots,t_M$ no conjunto $\cap_{\alpha\in\Theta}\mathbb{D}_{X(\alpha)} =]0,+\infty[$, considere-se o processo de contraste $O_n(\theta,\alpha) := \sum_{i=1}^{M}(\overline{\psi}_{X,n}(t_i) - \psi_{X(\alpha)}(t_i))^2$ e $(\overline{\theta}_n)_{n\in\mathbb{N}}$ a sucessão de estimadores do mínimo de contraste, isto é verificando para todo o $n \in \mathbb{N}$ $O_n(\theta,\overline{\theta}_n) := \min\{O_n(\theta,\alpha) : \alpha \in \Theta\}$. Então, $(\overline{\theta}_n)_{n\in\mathbb{N}}$ converge em probabilidade para θ.*

Demonstração. Aplicaremos um teorema standard (veja-se [1, p. 93]) para estimadores do mínimo de contraste. De acordo com este resultado temos que mostrar que a função de contraste $K(\theta,\alpha)$ definida, em consequência da lei forte dos grandes números por,

$$K(\theta,\alpha) = \lim_{n\to+\infty} O_n(\theta,\overline{\theta}_n) = \sum_{i=1}^{M}(\psi_{X(\theta)}(t_i) - \psi_{X(\alpha)}(t_i))^2 \ ,$$

é uma função contínua da variável α definida em Θ. Tal é de facto o caso, não só para $K(\theta,\alpha)$ mas também para $O_n(\theta,\alpha)$, uma vez que $\psi_{X(\alpha)}(t)$ é uma função contínua de α. Observe-se que $K(\theta,\alpha) \geq 0$ e que $K(\theta,\theta) = 0$. Podemos assim concluir que para todo o $n \in \mathbb{N}$ o estimador do mínimo de contraste $\overline{\theta}_n$ existe. Defina-se agora

$$\forall k \geq 1 \ \ w(n,k) := \sup\left\{|O_n(\theta,\alpha) - O_n(\theta,\beta)| : |\alpha-\beta| < \frac{1}{k}\right\} \ .$$

Para se ter consistência do estimador temos que verificar que existe uma sucessão decrescente $(\epsilon_k)_{k\geq 1}$ tal que $\lim_{k\to+\infty}\epsilon_k = 0$ e verificando ainda para todo o

$k \geq 1$ que $\lim_{n\to+\infty} \mathbb{P}[w(n,k) \geq \epsilon_k] = 0$. Para esse efeito observe-se que

$$O_n(\theta,\alpha) - O_n(\theta,\beta) =$$
$$= \sum_{i=1}^{M} [2\overline{\psi}_{X,n}(t_i) \left(\psi_{X(\beta)}(t_i) - \psi_{X(\alpha)}(t_i)\right) + \left(\psi_{X(\alpha)}(t_i)^2 - \psi_{X(\beta)}(t_i)^2\right)]$$

e que se definirmos

$$\forall k \geq 1 \;\; v(k) := \max_{1\leq i\leq M} \sup \left\{ \left|\psi_{X(\beta)}(t_i) - \psi_{X(\alpha)}(t_i)\right| : |\alpha - \beta| < \frac{1}{k} \right\}$$

e

$$\forall k \geq 1 \;\; u(k) := \max_{1\leq i\leq M} \sup \left\{ \left|\psi_{X(\beta)}(t_i)^2 - \psi_{X(\alpha)}(t_i)^2\right| : |\alpha - \beta| < \frac{1}{k} \right\}$$

as sucessões $(v_k)_{k\geq 1}$ e $(u_k)_{k\geq 1}$ são decrescentes tendo-se então que

$$w(n,k) \leq (u_k + v_k) \left[\sum_{i=1}^{M} (1 + 2\overline{\psi}_{X,n}(t_i)) \right] . \tag{5}$$

Considerando-se agora $w_k := u_k + v_k$, o facto que a sucessão $(w_k)_{k\geq 1}$ é decrescente e os factos que $\mathbb{E}[M + 2\sum_{i=1}^{M} \overline{\psi}_{X,n}(t_i)] = M + 2\sum_{i=1}^{M} \psi_{X(\theta)}(t_i)$ e também $\mathbb{V}[M+2\sum_{i=1}^{M} \overline{\psi}_{X,n}(t_i)] = (4/n)\sum_{i,j=1}^{M} (\psi_{X(\theta)}(t_i t_j) - \psi_{X(\theta)}(t_i)\psi_{X(\theta)}(t_j))$, tem-se a seguinte cadeia de desigualdades para qualquer $c > 0$,

$$\mathbb{P}\left[w(n,k) \geq w_k \left((M + 2\sum_{i=1}^{M} \psi_{X(\theta)}(t_i) + c \right) \right] \leq_{(a)}$$
$$\leq_{(a)} \mathbb{P}\left[\sum_{i=1}^{M} \overline{\psi}_{X,n}(t_i) \geq \sum_{i=1}^{M} \psi_{X(\theta)}(t_i) + \frac{c}{2} \right] \leq$$
$$\leq \mathbb{P}\left[\left| (\sum_{i=1}^{M} \overline{\psi}_{X,n}(t_i)) - \sum_{i=1}^{M} \psi_{X(\theta)}(t_i) \right| \geq \frac{c}{2} \right] \leq \frac{4}{c^2} \mathbb{V}\left[M + 2\sum_{i=1}^{M} \overline{\psi}_{X,n}(t_i) \right] =$$
$$= \frac{16}{nc^2} \sum_{i,j=1}^{M} \left(\psi_{X(\theta)}(t_i t_j) - \psi_{X(\theta)}(t_i)\psi_{X(\theta)}(t_j) \right) \xrightarrow[n\to+\infty]{} 0 \,,$$

onde a desigualdade (a) resulta da fórmula 5. Se se definir $\epsilon_k := w_k(M + 2\sum_{i=1}^{M} \psi_{X(\theta)}(t_i) + c)$ a demonstração estará terminada logo que se mostrar que $\lim_{k\to+\infty} u_k = 0 = \lim_{k\to+\infty} v_k$. Com esse propósito em vista, observe-se que para $X(\theta) \in \mathcal{N}(\theta, \sigma)$ se tem que:

$$\psi_{X(\beta)}(t) - \psi_{X(\alpha)}(t) = \int_{\mathbb{R}} t^x (\alpha - \beta) \frac{\partial}{\partial\theta} \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2\sigma^2}} \right)_{\theta=(1-\lambda_x)\alpha+\lambda_x\beta} dx \,,$$

onde para cada $x \in \mathbb{R}$ se verifica $\lambda_x \in [0,1]$. Como consequência, para algum $\gamma \in \Theta$ independente de x, um dado $\epsilon > 0$ e alguma constante $c(\epsilon, t)$ dependendo só de ϵ e t,

$$\begin{aligned} \left|\psi_{X(\beta)}(t) - \psi_{X(\alpha)}(t)\right| &\leq \frac{|\alpha-\beta|}{\sigma\sqrt{2\pi}} \int_{\mathbb{R}} t^x \left(|\alpha| + |x| + |\alpha - \beta|\right) e^{-\frac{(x-\gamma)^2}{2\sigma^2}} dx \leq \\ &\leq |\alpha - \beta| \left((|\alpha| + |\alpha-\beta|)\psi_{X(\gamma)}(t) + c(\epsilon,t)(\psi_{X(\gamma+\epsilon)}(t) + \psi_{X(\gamma-\epsilon)}(t))\right) \ , \end{aligned} \tag{6}$$

mostrando-se assim que $\lim_{k\to+\infty} v_k = 0$. Pode-se também concluir que se tem $\lim_{k\to+\infty} u_k = 0$ usando o mesmo raciocínio uma vez que temos para algum $\gamma \in \Theta$

$$\left|\psi_{X(\beta)}(t)^2 - \psi_{X(\alpha)}(t)^2\right| \leq 2\left|\psi_{X(\gamma)}(t)\right| \left|\psi_{X(\beta)}(t) - \psi_{X(\alpha)}(t)\right| \ ,$$

permitindo-nos assim terminar. Observe-se que em qualquer caso, só temos um número finito de pontos $t_1, \ldots, t_M$ com que nos preocupar nas definições de v_k e u_k. □

Tratemos agora do caso da distribuição gama.

Teorema 4.3 *Seja $X = X(\lambda, \theta) \in \mathcal{G}(\lambda, 1/\theta)$ onde λ se supõe ser dado e onde o parâmetro desconhecido é $\theta \in \Theta = [a,b] \subset \mathbb{R}$ um intervalo compacto. Seja $(X_n)_{n\in\mathbb{N}}$ uma amostra de X, $\overline{\psi}_{X,n}(t) = (1/n)\sum_{i=1}^n t^{X_i}$ definida para $t > 0$. Para $t_1, \ldots, t_M \in \cap_{\alpha\in\Theta}\mathbb{D}_{X(\lambda,\alpha)} =]0, e^a[$, considere-se o processo de contraste dado por $O_n(\theta, \alpha) := \sum_{i=1}^M (\overline{\psi}_{X,n}(t_i) - \psi_{X(\lambda,\alpha)}(t_i))^2$ e $(\overline{\theta}_n)_{n\in\mathbb{N}}$ a sucessão dos estimadores do mínimo de contraste, isto é, verificando para todo o $n \in \mathbb{N}$ a fórmula $O_n(\theta, \overline{\theta}_n) := \min\{O_n(\theta,\alpha) : \alpha \in \Theta\}$. Então $(\overline{\theta}_n)_{n\in\mathbb{N}}$ converge em probabilidade para θ.*

Demonstração. Substituindo a estimativa superior, dada na fórmula 6, pela seguinte

$$\left|\psi_{X(\beta)}(t) - \psi_{X(\alpha)}(t)\right| \leq |\alpha - \beta| \left(\frac{\lambda b^{\lambda-1}}{a^\lambda}\psi_{X(a,\lambda)}(t) + \frac{\lambda a^\lambda}{b^{\lambda+1}}\psi_{X(b,\lambda+1)}(t)\right) \ ,$$

a demonstração pode ser levada a cabo exactamente da mesma forma que a prova do caso Gaussiano. □

Agradecimentos

Este trabalho foi realizado com apoio parcial da FCT (Fundação para a Ciência e Tecnologia), programa POCTI (Portugal/FEDER-EU). Ao *referee* um obrigado pelas numerosas e justas correcções sugeridas.

Referências

[1] Dacunha-Castelle, D. e Duflo, M. (1983). *Probabilités et Statistiques, 2. Problèmes à Temps Mobiles.* Masson.

[2] Dowling, M. M. e Nakamura, M. (1997). Estimating parameters for discrete distributions via the empirical probability generating function. *Commun. Stat., Simulation Comput.*, Vol. 26(1), p. 301-313.

[3] Esquível, M. L. (2004). Probability generating functions for discrete real valued random variables (aceite para publicação).

[4] Feuerverger, A. (1989). On the empirical saddlepoint approximation. *Biometrika*, Vol. 76(3), p. 457-464.

[5] Feuerverger, A. e McDunnough, P. (1984). On statistical transform methods and their efficiency. *The Can. J. of Stat.*, Vol. 12(4), p. 303-317.

[6] Gaenssler, P. (1983). *Empirical Processes.* Lecture Notes-Monograph Series Vol. 3, Institute of Mathematical Statistics.

[7] Kocherlakota, S. e Kocherlakota, K. (1986). Goodness of fit tests for discrete distributions. *Commun. Statist.-Theor. Meth.*, Vol. 15(3), p. 815-829.

[8] Malliavin, P. (1995). *Integration and Probability.* Springer Verlag.

[9] Marques, M. S. e Pérez-Abreu, V. (1989). Law of large numbers and central limit theorem for the empirical probability generating function of stationary random sequences and processes. *Aportaciones Mat., Notas Invest.*, Vol. 2(4), p. 100-109.

[10] Nakamura, N. e Pérez-Abreu, V. (1993). Empirical probability generating function. An overview. *Insur. Math. Econ.*, Vol. 12(3), p. 349-366.

[11] Nakamura, N. e Pérez-Abreu, V. (1993). Exploratory data analysis for counts using the empirical probability generating function. *Commun. Stat., Theory Methods*, Vol. 22(3), p. 827-842.

[12] Nakamura, M. e Pérez-Abreu, V. (1993). Use of an empirical probability generating function for testing a Poisson model. *Can. J. Stat.*, Vol. 21(2), p. 149-156.

[13] Pestman, W. R. (1998). *Mathematical Statistics.* Walter de Gruyter.

[14] Prakasa Rao, B. L. S. (1999). *Statistical Inference for Diffusion Type Processes.* Arnold Hodder Headline Group.

[15] Rémillard, B. e Theodorescu, R. (2000). Inference based on the empirical probability generating function for mixtures of Poisson distributions. *Statistics and Decisions*, Vol. 18(4), p. 349-366.

[16] Rueda, R. e O'Reilly, F. (1999). Tests of fit for discrete distributions based on the probability generating function. *Commun. Stat., Simulation Comput.*, Vol. 28(1), p. 259-274.

[17] Rueda, R., Pérez-Abreu, V. e O'Reilly, F. (1991). Goodness of fit for the Poisson distribution based on the probability generating function. *Commun. Stat., Theory Methods*, Vol. 20(10), p. 3093-3110.