Desenvolvimentos assintóticos e polinómios ortogonais

Madalena Malva
Instituto Politécnico de Viseu - Escola Superior de Tecnologia de Viseu, Departamento de Matemática e CEAUL - Centro de Estatística e Aplicações da Universidade de Lisboa
Sandra Mendonça
Universidade da Madeira - Departamento de Matemática e Engenharias e CEAUL - Centro de Estatística e Aplicações da Universidade de Lisboa

Resumo: Neste trabalho são discutidos melhoramentos às expansões de Edgeworth usando pontos sela, e são fornecidos resultados no contexto da família exponencial natural (NEF) de Morris com função variância quadrática (QVF). Há apenas seis membros na família de Morris, e são discutidas as famílias de polinómios ortogonais associados, todos pertencentes à classe mais geral dos polinómios de Meixner. Expansões com base em pontos sela podem ser reinterpretados em termos de distribuições conjugadas, seguindo o trabalho de Cramér (1928) [*Skand. Aktuarietidskr.* 11: 13-74, 141-180], o que proporciona as bases gerais para a discussão de aproximações assintóticas no domínio do teorema limite central.

Palavras–chave: famílias exponenciais naturais, função variância quadrática, polinómios ortogonais, expansões de Edgeworth, expansões com base no método do ponto sela.

Abstract: We discuss the improvement of Edgeworth expansions using saddlepoints, and provide results in the context of Morris natural exponencial family (NEF) of quadratic variance function (QVF). There are only six members in Morris family, and we discuss the associated families of orthogonal polynomials, all of then belonging to a more general class, the Meixner polynomials. Saddlepoint expansions may be reinterpreted in terms of conjugate distributions, following the work of Cramér (1928) [*Skand. Aktuarietidskr.* 11: 13-74, 141-180], and this provides a general framework for the discussion of asymptotic approximations in the realm of the central limit theorem.

Keywords: natural exponencial families, quadratic variance function, orthogonal polynomials, Edgeworth expansions, saddlepoint approximation.

Classificação MSC2000: 60F05.

1 Introdução

O fascínio inicial pelo Teorema Limite Central obliterou, até certo ponto, a constatação do seu maior defeito: a velocidade de convergência pode ser, em muitas circunstâncias, excessivamente lenta. Os estudos feitos no início do século XX por Gram-Charlier e Edgeworth, com correcção das demonstrações em Cramér (1928), veio culminar, na teoria clássica, com os resultados de Berry (1941) e Esséen (1942 e 1945). No entanto, a velocidade de convergência conseguida era apenas $O(n^{-1/2})$.

Daniels (1954) usando pontos sela conseguiu velocidades de convergência da ordem $O(n^{-1})$. Mais relevante ainda, conseguiu relacionar os aspectos puramente analíticos com interpretações probabilísticas usando distribuições conjugadas, que parecem dever-se a Esscher (1932), e tinham já sido usadas com sucesso por Cramér nos seus estudos pioneiros sobre convergências.

O desenvolvimento da teoria das famílias exponenciais (Fisher (1934), Darmois (1935), Koopman (1936), Pitman (1936)) forneceu um contexto particularmente adequado para a discussão desta questão. Neste contexto, os cumulantes (inventados por Thiele em 1899, cf. Thiele (1903), e "reinventados" por Fisher (1929) trinta anos depois), embora não tenham uma interpretação óbvia como os momentos, permitem um tratamento analítico muito simples. Anote-se, ainda, que nas leis limite mais usadas — a gaussiana e a Poisson — têm caracterizações extremamente simples em termos dos cumulantes.

Morris (1982) estudou um conjunto de variáveis da família exponencial, contendo a gaussiana$(\mu\,,\sigma^2)$ e a Poisson(μ), caracterizada por

$$V(\mu) = v_0 + v_1\mu + v_2\mu^2\ , \quad \text{onde} \quad \mathrm{E}(X) = \mu\,,\ \mathrm{V}(\mu) = \mathrm{E}(X^2) - \mathrm{E}(X)^2\ , \quad (1)$$

i.e., a variância expressa como função do valor médio, ser um polinómio de grau máximo dois. Para além da gaussiana onde a solução trivial é obtida fazendo $v_0 = \sigma^2\,, v_1 = v_2 = 0$, demonstrou que as restantes soluções eram: a Poisson, gama, binomial, binomial negativa e a secante hiperbólica.

É interessante anotar que às medidas de probabilidade associadas a cada uma daquelas soluções estão associados polinómios ortogonais da classe de Meixner. Estes polinómios são de grande importância nomeadamente no contexto das expansões assintóticas, pois em muitos casos os polinómios p_j de grau j , em x , que aparecem quer nas expansões de Edgeworth quer nas de Gram-Charlier, são ortogonais em relação à distribuição associada a $f(x)$, que é neste contexto considerada a aproximação de primeira ordem, como por exemplo a função densidade ou de distribuição da gaussiana reduzida.

Na secção 2 descrevemos sucintamente os polinómios ortogonais referidos, estabelecendo um teorema geral sobre o seu termo director, decorrente da expressão recursiva para o seu cálculo.

Discutimos seguidamente as expansões de Edgeworth e as expansões usando pontos de sela.

Finalmente, apresentamos um exemplo que mostra a que ponto esta teoria melhora aproximações à convolução de densidades, subjacente a estudos sobre a média.

Na discussão final, chamamos a atenção para o facto do seu poder regularizador do integral de convolução que leva a que a densidade da soma (e portanto da média) seja "quase-simétrica" em torno do ponto sela, aligeirando assim o efeito do terceiro cumulante (o terceiro cumulante padronizado é o clássico coeficiente de assimetria de Pearson) permitindo assim passar de aproximações $O(n^{-1/2})$ para aproximações $O(n^{-1})$. Um comentário sobre a aplicação desses resultados à variáveis aleatórias (v.a.) de Lévy (estável com $\alpha = \frac{1}{2}$) mostra que as aproximações podem ser excelentes, mesmo quando falham as condições de regularidade exigidas ao abordar a questão com a função geradora de cumulantes.

São assim objectivos deste artigo

1. Realçar a simplicidade de obtenção de aproximações para densidades de somas (médias) melhores do que o Teorema Limite Central, nomeadamente na região de grandes desvios.

2. Mostrar a simplicidade acrescida no caso de parente de famílias exponencial, nomeadamente recorrendo a funções geradoras de cumulantes.

3. Divulgar a importante família de Morris, e estabelecer as conexões entre a problemática assintótica e a utilização de polinómios ortogonais.

4. Exemplificar detalhadamente, com especial relevo para o caso de soma de uniformes, cujo tratamento analítico clássico é complicado.

5. Chamar a atenção para o campo de investigação que se abre com um exemplo inesperado; a aproximação no caso de parente Lévy é afinal exacta, o que mostra que a validade do método do ponto sela se verifica em condições muito diversas das assumidas em geral — fica assim em aberto estudar nesta perspectiva os domínios de atracção de estáveis não gaussianas.

2 As famílias exponenciais naturais com função de variância quadrática e polinómios ortogonais

Uma família de distribuições paramétricas com conjunto de parâmetros naturais $\Theta \subset \mathbb{R}$ é uma família exponencial univariada se as v.a. Y governadas por estas distribuições satisfizerem

$$P_\theta(Y \in A) = \int_A \exp\{\theta T(y) - \psi(\theta)\}\xi(dy) \tag{2}$$

para alguma medida ξ não dependente de $\theta \in \Theta$, $A \subset \mathbb{R}$ um conjunto mensurável e T uma função de valores reais mensurável. A *observação natural* em $A \equiv \mathbb{R}$ é $X = T(Y)$. A sua distribuição pertence à família exponencial natural (NEF),

$$P_\theta(Y \in A) = \int_A \exp\{\theta x - \psi(\theta)\}dF(x) \tag{3}$$

com F uma medida de Stieltjes em $\mathbb{R}$. Se $0 \in \Theta$ e $\psi(0) = 0$ então F é uma função de distribuição. Caso contrário, para qualquer $\theta_0 \in \Theta$ seja $dF_0(x) = \exp\{\theta_0 x - \psi(\theta_0)\}dF$. Então F_0 é uma função distribuição e gera a mesma família exponencial que F. Assim, sem perda de generalidade, assumimos que F em (3) é uma função distribuição.

Toda a função de distribuição F_0 que possua função geradora de momentos (f.g.m.) numa vizinhança de zero gera uma NEF do seguinte modo. Defina-se a função geradora de cumulantes (f.g.c.) $\psi(\theta)$ em Θ, o maior intervalo para o qual a f.g.c existe, por

$$\psi(\theta) = \log\left(\int \exp(\theta x) dF_0(x)\right), \ \theta \in \Theta. \tag{4}$$

Então as funções de distribuição F_θ, $\theta \in \Theta$, definidas por

$$dF_\theta(x) = \exp(\theta x - \psi(\theta)) dF_0(x), \tag{5}$$

formam uma NEF, com F_θ função de distribuição. A NEF assim gerada é chamada *família conjugada*.

Poucas NEFs têm função de variância quadrática (QVF) dada por

$$V(\mu) = v_0 + v_1\mu + v_2\mu^2. \tag{6}$$

Se X tem uma distribuição pertencente a uma NEF e $V(\mu)$ é quadrática, dizemos que X tem uma distribuição NEF-QVF.

Seja $f(x,\theta)$ uma densidade de uma NEF-QVF proporcional a $\exp(\theta x - \psi(\theta))$ relativamente a alguma medida (como em (3)) e defina-se

$$P_m(x,\mu) = \frac{V^m(\mu)\frac{d^m f(x,\theta)}{d\mu^m}}{f(x,\theta)} \quad \text{para} \quad m = 0, 1, 2, \ldots\,. \tag{7}$$

Mostra-se que $P_m(x,\mu)$ é um polinómio de grau m em x e μ e com termo principal x^m, e que P_m é uma família de polinómios ortogonais. De (7) vem que

$$P_{m+1} = (P_1 - mV')P_m + VP'_m \quad , \quad m \geq 1 \tag{8}$$

com $P'_m = \frac{\partial P_m(x,\mu)}{\partial \mu}$ e $P_m^{(r)} = \frac{\partial^r P_m(x,\mu)}{\partial \mu^r}$.

Teorema *O conjunto $\{P_m(x,\mu) : m = 0, 1, \ldots\}$ é, para as famílias NEF-QVF, um sistema ortogonal de polinómios com respeito a $f(x,\theta) = \exp(x\theta - \psi(\theta))$. $P_m(x,\theta)$ tem grau exacto igual a m, tanto em μ como em x, com termo director x^m, e é gerado por*

$$P_{m+1} = (P_1 - mV')P_m - m[1 + (m-1)v_2]VP_{m-1} \tag{9}$$

para $m \geq 1$ *com* $P_0 = 1$ *e* $P_1 = x - \mu$. *Defina-se* $a_0 = 1$ *e para* $m \geq 1$

$$a_m = m! \prod_{i=0}^{m-1} (1 + iv_2). \tag{10}$$

Então para $m \geq 1$, $r = 0, 1, ..., m$, *as derivadas em ordem a* μ *são*

$$P_m^{(r)} = (-1)^r \frac{a_m}{a_{m-r}} P_{m-r}. \tag{11}$$

Finalmente, $E_\theta(P_m) = 0$ *para* $m \geq 1$ *e*

$$E_\theta(P_m P_n) = \delta_{mn} a_m V^m \ , \ m, n \geq 0. \tag{12}$$

Demonstração Defina-se $b_m = (m+1)(1 + mv_2)$. Temos então que

$$\frac{a_m}{a_{m-1}} = b_{m-1}. \tag{13}$$

Comecemos por mostrar que a igualdade

$$P_m' = -b_{m-1} P_{m-1} \ ,$$

que para $m = 1$ é verdadeira pois $P_1' = -b_0 \, P_0$ e $P_0 = b_0 = 1$, é também válida para qualquer outro natural positivo. Assuma-se que a igualdade é verdadeira para $m \geq 1$. De (8) tem-se que

$$P_{m+1}' = -(1 + 2\, m\, v_2 + b_{m-1})\, P_m = -b_m \, P_m \ .$$

Indução sobre m prova que $P_m^{(1)} = -b_{m-1}\, P_{m-1}$. Iterando a expressão anterior $(r-1)$ vezes obtém-se

$$P_m^{(r)} = (-1)^r \,(b_{m-1} \cdots b_{m-r})\, P_{m-r} = (-1)^r \,\frac{a_m}{a_{m-r}}\, P_{m-r} \ ,$$

o que prova (11).

A equação (9) é consequência das equações (8) e (11) com $r = 1$. Que $P_m(x, \mu)$ é um polinómio de grau exacto tanto em x como em μ com termo director igual a x^m , segue intuitivamente de (9). Para $n < m$, e usando (7) tem-se

$$\mathrm{E}(X^m \, P_m) = \mathrm{V}^m \int x^m \, f^{(m)}(x, \theta)\, dF(x) = \mathrm{V}^m \,\frac{d^m}{d\,\mu^m}\, \mathrm{E}_\theta(X^m) = 0 \ ,$$

pois $\mathrm{E}_\theta(X^m)$ é um polinómio de grau quando muito igual a n em $\mu = \psi'(\theta)$, então para $m > n$, $\frac{d^m}{d\,\mu^m}\,\mathrm{E}_\theta(X^m) = 0$. Então P_m é ortogonal a todos os polinómios de grau inferior a m , e

$$\mathrm{E}_\theta(P_n\,P_m) = \delta_{m\,n}\,a_m\,\mathrm{V}^m$$

é verdadeira para $m \neq n$. Para provar a igualdade anterior para $m = n > 1$, vamos usar a igualdade (9) multiplicando-a por P_{m-1} e de seguida tomar médias, repetimos o processo com P_m e obtém-se

$$\mathrm{E}(\mathrm{P}_m^2) = b_{m-1}\,\mathrm{V}\,\mathrm{E}(P_{m-1}^2)\ .$$

Iterando obtém-se

$$\mathrm{E}(P_m^2) = b_{m-1}\cdots b_0\,\mathrm{V}^m\,\mathrm{E}(P_0) = a_m\,\mathrm{V}^m.$$

$\square$

Os polinómios anteriores são conhecidos individualmente por polinómios de **Hermite** se a distribuição associada é a gaussiana, **Poisson-Charlier** se a distribuição associada é a Poisson, **Laguerre generalizados** se a distribuição associada é a gama, **Krawtchouk** se a distribuição associada é a binomial, **Meixner** se a distribuição associada é a binomial negativa e **Pollaczek** se a distribuição associada é a secante hiperbólica generalizada (GHS) (mas apenas para subfamílias simétricas).

3 Expansões de Edgeworth

As expansões de Edgeworth aqui obtidas para somas padronizadas de variáveis aleatórias independentes e identicamente distribuídas, são importantes na teoria assintótica quando o integral de convolução referente à soma de variáveis aleatórias não pode ser calculado de modo explícito.

Seja X uma variável aleatória com função densidade $f(x)$ e função geradora de cumulantes $K(t)$. Os cumulantes padronizados de X são $\rho_r = \frac{k_r}{k_r^{-r/2}}, r \geq 2$. Tem-se que $k_1 = \mathrm{E}(X)$ e $k_2 = \mathrm{Var}(X) = \sigma^2$. Seja $X_1, \cdots, X_n$ uma amostra aleatória de X e sejam $S_n = \sum_{i=1}^{n} X_i$ e $S_n^* = \frac{S_n - n\,\mu}{\sigma\,\sqrt{n}}$. Como as v.a. são independentes e identicamente distribuídas , as funções geradoras de cumulantes de S_n e S_n^* são dadas por $K_{S_n(t)} = n\,K(t)$ e

$$K_{S_n^*(t)} = -\frac{\sqrt{n}\,\mu\,t}{\sigma} + nK\left(\frac{t}{\sigma\,\sqrt{n}}\right)\ , \tag{14}$$

respectivamente. A expansão de $K(t)$ em série de Taylor é equivalente a uma soma de funções dos cumulantes de X padronizados que substituída em (14) dá origem a

$$K_{S_n^*(t)} = \frac{t^2}{2} + \rho_3 \frac{t^3}{6\sqrt{n}} + \rho_4 \frac{t^4}{24\,n} + O\left(n^{-3/2}\right) . \tag{15}$$

A expansão anterior revela que $K_{S_n^*}(t) \longrightarrow \frac{t^2}{2}$ quando $n \to \infty$, o que era esperado pois pelo teorema limite central S_n^* converge em distribuição para a distribuição gaussiana reduzida quando $n \to \infty$. Assim, a função geradora de momentos de $M_{S_n^*}(t)$ de S_n^* é obtida de (15) tomando exponenciais, e a função densidade de S_n^* obtém-se invertendo $M_{S_n^*}(t)$ termo a termo e usando a identidade

$$\int \exp(t\,x)\,\phi(x)\,H_r(x)\,dy = t^r \exp\left(\frac{t^2}{2}\right) .$$

A função densidade é assim dada por

$$f_{S_n^*}(x) = \phi(x)\left\{1 + \frac{\rho_3}{6\sqrt{n}}\,H_3(x) + \frac{\rho_4}{24\,n}\,H_4(x) + \frac{\rho_3^2}{72\,n}\,H_5(x)\right\} + O\left(n^{-3/2}\right) . \tag{16}$$

Integrando (16) obtém-se a expansão da função de distribuição de S_n^*

$$F_{S_n^*}(x) = \Phi(x) + \phi(x)\left\{1 + \frac{\rho_3 H_2(x)}{6\sqrt{n}} + \frac{\rho_4 H_3(x)}{24\,n} + \frac{\rho_3^2 H_4(x)}{72\,n}\right\} + O\left(n^{-3/2}\right) . \tag{17}$$

As fórmulas (16) e (17) são, respectivamente, as expansões de Edgeworth para as funções densidade e de distribuição de uma soma padrão. O termo de ordem $n^{-1/2}$ em (17) corrige, na aproximação "básica" feita à distribuição gaussiana, o principal efeito da assimetria enquanto que o termo de ordem n^{-1} corrige o efeito principal da curtose e o efeito secundário da assimetria.

A aproximação (16) pode não ser apropriada para aproximar valores na cauda da distribuição de S_n^* , pois os polinómios de Hermite não são limitados.

4 Expansões com base no ponto sela

Estas expansões são muito importantes na teoria assintótica pois aproximam com grande precisão as funções densidade e de distribuição.

Seja X uma v.a. contínua com função de distribuição $F(x)$. Assuma-se que a função densidade de probabilidade $f(x) = F'(x)$ existe e suponhamos que a função geradora de momentos

$$M(t) = \exp(K(t)) = \int_{-\infty}^{\infty} \exp(tx)f(x)\,dx$$

converge para t real, em algum intervalo não vazio contendo a origem ($K(t)$ é função geradora de cumulantes). Seja $-c_1 < t < c_2$ o maior de tais intervalos, onde $0 \leq c_1 \leq \infty$ e $0 \leq c_2 \leq \infty$ mas $c_1 + c_2 > 0$, ou seja, c_1 ou c_2 podem ser zero, mas não ambos (não é necessário que todos os momentos existam).

Seja $\overline{X}$ a média de n v.a. i.i.d. A função densidade de probabilidade $f(\overline{x}) = F'(\overline{x})$ é dada pela fórmula de inversão de Fourier

$$f(\overline{x}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-it\overline{x}) \varphi_{\overline{X}}(t)\, dt.$$

Como

$$\varphi_{\overline{X}}(t) = \left(M_X \left(i\frac{t}{n} \right) \right)^n$$

temos

$$f(\overline{x}) = \frac{n}{2\pi} \int_{\mathbb{R}} \exp(-iny\overline{x}) \left(M_X \left(i\, y \right) \right)^n \, dy\,. \tag{18}$$

Quando n é grande, uma aproximação para $f(\overline{x})$ pode ser encontrada escolhendo o caminho de integração que passa através do ponto sela da expressão a integrar e de modo que a expressão a integrar seja desprezável fora de uma vizinhança. Os pontos sela são tais que

$$(n(K(t) - t\overline{x}))' = 0 \Rightarrow nK'(t) - n\overline{x} = 0$$

i.e.,

$$K'(t) = \overline{x}\,. \tag{19}$$

Seja $t_0 \in (c_1, c_2)$ a única raiz real de (19) para cada $\overline{x}$ tal que $0 < F(\overline{x}) < 1$ e tal que $K''(t_0) > 0$. Escolha-se como caminho de integração a linha recta que passa por t_0 e é paralela ao eixo imaginário. Como $K(t) - t\overline{x}$ tem um mínimo em t_0 para t real, o módulo da expressão integranda tem no caminho escolhido um máximo em t_0. Mostra-se que para qualquer recta admissível paralela ao eixo imaginário a expressão integranda atinge o seu máximo em módulo apenas onde a recta intersecta o eixo real.

No contorno próximo de t_0, considerando a expansão em série de $K(t)$ para $t = t_0$, obtém-se

$$K(t) - t\overline{x} = K(t_0) - t_0\,\overline{x} - \frac{K''(t_0)}{2} y^2 - \frac{K'''(t_0)}{6} iy^3 + \frac{K^{(iv)}(t_0)}{24} y^4 + \ldots\,.$$

Se considerarmos $y = \dfrac{v}{\sqrt{nK''(t_0)}}$, então

$$K(t) - t\overline{x} = K(t_0) - t_0\,\overline{x} - \frac{v^2}{2n} - \frac{1}{6}\lambda_3(t_0)\frac{iv^3}{n^{3/2}} + \frac{1}{24}\lambda_4(t_0)\frac{v^4}{n^2} + \ldots$$

onde $\lambda_j(t) = \dfrac{K^{(j)}(t)}{(K''(t))^{j/2}}$, para $j \geq 3$, obtendo-se

$$f(\overline{x}) \approx \left[\frac{n}{2\pi K''(t_0)}\right]^{1/2} \exp(n(K(t_0){-}t_0\overline{x})) \left[1 + \frac{1}{n}\left[\frac{1}{8}\lambda_4(t_0) - \frac{5}{24}\lambda_3^2(t_0)\right] + ...\right].$$

Seja $g(\overline{x}) = \left[\frac{n}{2\pi K''(t_0)}\right]^{1/2} \exp(n(K(t_0) - t_0\overline{x}))$, tem-se que

$$f(\overline{x}) \approx g(\overline{x}) \left[1 + \frac{1}{n}\left[\frac{1}{8}\lambda_4(t_0) - \frac{5}{24}\lambda_3^2(t_0)\right] + ...\right]. \tag{20}$$

A $g(\overline{x})$ **chamamos aproximação ponto sela de** $f(\overline{x})$.

Note-se a ausência em (20) do termo correspondente a $\dfrac{\lambda_3(t_0)}{n^{1/2}}$, relacionado com a assimetria de S_n/n. De facto, é o desaparecimento deste termo que permite a esta aproximação ter um erro da grandeza de n^{-1} ao invés da aproximação com erro da grandeza $n^{-1/2}$ na expansão de Edgeworth.

5 Exemplos

Suponhamos que $X \frown Uniforme(-1, 1)$. A função densidade de probabilidade da média de n v.a. independentes e uniformes em $(-1, 1)$, $\frac{S_n}{n} = \displaystyle\sum_{i=1}^{n} \frac{X_i}{n}$, é:

$$f(\overline{x}) = \frac{n^n}{2^n(n-1)!} \sum_{s=0}^{n} (-1)^s \mathrm{C}_s^n \left[\max\left(1 - \overline{x} - \frac{2s}{n}, 0\right)\right]^{n-1} \mathrm{I}_{[-1,1]}(x).$$

Vamos calcular $g(\overline{x})$. Temos

Função geradora de momentos: $M_X(t) = \dfrac{\sinh(t)}{t}$,

Função geradora de cumulantes: $K(t) = \log\left(\dfrac{\sinh(t)}{t}\right)$,

$K'(t) = \coth(t) - \dfrac{1}{t}$ e $K''(t) = \dfrac{1}{t^2} - \mathrm{cosech}^2(t)$.

Além disso,

$$\exp(n(K(t_0) - t_0\overline{x})) = \left(\frac{\sinh(t_0)}{t_0}\right)^n \exp(-nt_0\overline{x}),$$

donde

$$g(\overline{x}) = \left(\frac{n}{2\pi}\right)^{1/2} \left\{\frac{1}{t_0^2} - \mathrm{cosech}^2(t_0)\right\}^{1/2} \left(\frac{\sinh(t_0)}{t_0}\right)^n \exp(-nt_0\overline{x}).$$

Quando t_0 é positivo e grande $\overline{x} \simeq 1 - \dfrac{1}{t_0}$, $K(t_0) \approx \log\left(\frac{\exp(t_0)}{2t_0}\right)$ e $K''(t_0) \approx \dfrac{1}{t_0^2}$. Então para $1-\overline{x}$ pequeno

$$g(\overline{x}) \approx \left(\frac{n}{2\pi}\right)^{1/2}\left(\frac{1}{2}\right)^{n}(1-\overline{x})^{n-1}\exp(n) \tag{21}$$

pois $K'(t_0) = \overline{x} \Rightarrow t_0 \simeq \dfrac{1}{1-\overline{x}}$, o que está de acordo com

$$f(\overline{x}) = \frac{n^n}{2^n(n-1)!}(1-\overline{x})^{n-1}\,,$$

quando $\overline{x} > 1 - \dfrac{2}{n}$ a menos de uma constante normalizadora. Um argumento análogo é válido para $\overline{x}$ próximo de -1. Para valores moderados de n sobre todo o domínio de x, prova-se que $\log(g(\overline{x}))$ toma valores muito próximos de $\log(f(\overline{x}))$.

Relativamente à expansão de Edgeworth para a função densidade de probabilidade da média de n variáveis, $X_1, ...X_n$, independentes e idênticas a uma variável $X \frown Uniforme(-1,1)$, S_n/n, vamos usar a fórmula (16).

Tendo em conta que

$$\rho_1 = 0, \quad \rho_2 = \frac{1}{3}, \quad \rho_3 = 0, \quad \text{e} \quad \rho_4 = -\frac{2}{15},$$

que

$$H_3\left(y\right) = y^3 - 3y,\ H_4\left(y\right) = y^4 - 6y^2 + 3 \text{ e } H_6\left(y\right) = y^6 - 15y^4 + 45y^2 - 15,$$

e que

$$f_{\frac{S_n}{n}}\left(y\right) = f_{S_n^*}\left(\frac{ny - E\left(S_n\right)}{\sqrt{var\left(S_n\right)}}\right)\frac{n}{\sqrt{var\left(S_n\right)}},$$

temos

$$f_{\frac{S_n}{n}}\left(y\right) = f_{S_n^*}\left(\frac{ny}{\sqrt{\frac{n}{3}}}\right)\frac{n}{\sqrt{\frac{n}{3}}} = f_{S_n^*}\left(y\sqrt{3n}\right)\sqrt{3n},$$

e, consequentemente,

$$f_{\frac{S_n}{n}}\left(y\right) = \frac{\exp\left(-\frac{y^2 3n}{2}\right)}{\sqrt{2\pi}}\left\{1 - \frac{\left(y\sqrt{3n}\right)^4 - 6\left(y\sqrt{3n}\right)^2 + 3}{180n}\right\}\sqrt{3n} + O\left(n^{-1}\right).$$

Curiosamente, a expansão dada pelo método do ponto de sela parece ser aplicável mesmo a situações de variáveis que não possuem cumulantes. Consideremos, por exemplo, a variável aleatória X com distribuição de Lévy (estável de parâmetros ($\alpha = \frac{1}{2}$ e $\beta = 1$)) com função densidade de probabilidade

$$f_X(y) = \frac{\exp\left(-\frac{1}{2y}\right)}{\sqrt{2\pi y^3}} \mathbf{I}_{(0,\infty)}(y).$$

Como é estável de parâmetro $\alpha = 1/2$ sabemos que

$$f_{\frac{S_n}{n^2}}(x) = \frac{\exp\left(-\frac{1}{2x}\right)}{\sqrt{2\pi x^3}} \mathbf{I}_{(0,\infty)}(x).$$

Consequentemente,

$$f_{S_n}(x) = f_{\frac{S_n}{n^2}}\left(\frac{x}{n^2}\right)\frac{1}{n^2} = n\frac{\exp\left(-\frac{n^2}{2x}\right)}{\sqrt{2\pi x^3}} \mathbf{I}_{(0,\infty)}(x)$$

e

$$f_{\frac{S_n}{n}}(x) = f_{S_n}(nx)\, n = \left(\frac{n}{2\pi x^3}\right)^{1/2} \exp\left(-\frac{n}{2x}\right) \mathbf{I}_{(0,\infty)}(x). \tag{22}$$

Por outro lado $K(t) = -\sqrt{-2t}$ (definido para $t < 0$). Determinando a primeira e a segunda derivadas de K obtemos

$$K'(t) = (-2t)^{-1/2} \quad \text{e} \quad K''_X(t) = (-2t)^{-3/2}.$$

De $K'(t_0) = x \Rightarrow t_0 = -\dfrac{1}{2x^2}$ resulta $K''(t_0) = x^3$ e $K(t_0) = -\dfrac{1}{x}$. Substituindo na expressão de g, a aproximação ponto de sela de $f_{\frac{S_n}{n}}$, obtemos, para $x > 0$,

$$g(x) = \left(\frac{n}{2\pi x^3}\right)^{1/2} \exp\left(-\frac{n}{2x}\right).$$

Compare-se este último resultado com (22). A distribuição de Lévy não tem momentos inteiros e, no entanto, a aproximação dada por método do ponto de sela é, de facto, exacta. Este exemplo ilustra, portanto, que estas aproximações assintóticas podem ser válidas em situações menos regulares do que as que usámos ao estabelecer o corpo da teoria, abrindo largo campo para investigação sobre velocidades de convergência de somas de v.a. no domínio de atracção das estáveis não gaussianas.

Agradecimentos

Ao Professor Dinis Pestana pelo constante apoio prestado.

Os comentários detalhados do *referee* chamaram a nossa atenção para muitas questões pertinentes, o que permitiu uma revisão muito substancial deste trabalho.

Investigação subsidiada por FCT/POCTI/FEDER, Projecto VEXTRA e pelo programa PRODEP, Acção 5.3.

Referências

[1] Berry, A. C. (1941). The Accuracy of the Gaussian Approximation to the Sum of Independent Variates. *Trans. Amer. Math. Soc.*, Vol. 49, p. 122-136.

[2] Cramér, H. (1928). On the Composition of Elementary Errors. *Skand. Aktuarietidskr.*, Vol. 11, p. 13-74.

[3] Cramér, H. (1928). On the Composition of Elementary Errors. Second paper: statistical applications *Skand. Aktuarietidskr.*, Vol. 11, p. 141-180.

[4] Daniels, H. E. (1954). Saddlepoint Approximations in Statistics. *Ann. Math. Statist.*, Vol. 25, p. 631-650.

[5] Darmois, G. (1935). Sur les lois de probabilité à estimation exhaustive. *C. R. Ac. Sc. Paris*, Vol. 260, p. 1265-1266.

[6] Esscher, F. (1932). On the probability function in the collective theory of risk. *Skand. Aktuarietidskr.*, Vol. 15, p. 175-195.

[7] Esséen, C. G. (1942). On the Liapounoff Limit of Error in the Theory of Probability. *Ark. Mat. Astr. och Fys.*, Vol. 28(A) 9, p. 1-19.

[8] Esséen, C. G. (1945). Fourier Analysis of Distribution Functions. *Acta Math.*, Vol. 77, p. 1-125.

[9] Fisher, R. (1929). Moments and Product Moments of Sampling Distributions. *Proc. London Math. Soc. Series 2*, Vol. 30, p. 199-238.

[10] Fisher, R. (1934). Two new properties of mathematical likelihood. *Proc. Royal Soc. London, A*, Vol. 144, p. 295-307.

[11] Koopman, L. H. (1936). On distributions admiting a sufficient statistic. *Trans. Amer. Math. Soc.*, Vol. 39, p. 399-409.

[12] Morris, C. N. (1982). Natural Exponential Families With Quadratic Variance Functions. *Ann. Stat.*, Vol.10, p. 65-80.

[13] Pitman, E. J. G. (1936). Sufficient statistics and intrinsic accuracy. *Proc. Cambridge Phil. Soc.* Vol. 32, p. 567-579.

[14] Thiele, T. N. (1903). *Theory of Observations.*, Dayton, London. (Reimpresso em *Ann. Math. Stat.*, Vol. 2, p. 165-308, 1931).