Universidade do Minho 2020/2021

Probabilidades e Aplicações – CC e MAT (3º ano) 1º semestre

1 Transformadas

1.1 Transformada de Laplace de uma v.a.

Dada uma variável aleatória (*v.a.*) X, consideremos o valor médio $E(e^{-tX})$, para $t \in \mathbb{R}$. Se este valor médio existir numa vizinhança de $t = 0$, chama-se *transformada de Laplace* da *v.a.* X (ou da sua distribuição) à função L_X definida nessa vizinhança por

$$L_X(t) = E(e^{-tX}).$$

Sendo L_X a transformada de Laplace de X, tem-se $L_X(0) = 1$ e a transformada de Laplace de $Y = a + bX$ $(a, b \in \mathbb{R})$ é dada por

$$L_Y(t) = E(e^{-t(a+bX)}) = e^{-at}E(e^{-b\,tX}) = e^{-at}L_X(bt).$$

A transformada de Laplace identifica a distribuição, *i.e.*, a cada função de distribuição (*f.d.*) para a qual exista esta transformada, corresponde uma única transformada de Laplace, e reciprocamente, a cada transformada de Laplace corresponde uma única *f.d.* (este e outros resultados aqui referidos encontram-se demonstrados em Widder, *The Laplace Transform*). Outro resultado importante estabelece que a convergência em distribuição[1] é equivalente à convergência das transformadas de Laplace para uma função contínua na origem.

A transformada de Laplace L_X e os momentos[2] de X estão relacionados. De facto, se tal transformada existir numa vizinhança da origem, $\{t : |t| < t_0\}$, esta terá derivadas (de qualquer ordem) na origem, sendo válida a expansão (única) em série de Mclaurin,

$$L(t) = L(0) + \frac{1}{1!}L'(0)\,t + \frac{1}{2!}L''(0)\,t^2 + \dots$$

Nesse caso existirão todos os momentos de X (de qualquer ordem), dados pelo coeficiente de $\frac{1}{n!}(-t)^n$ naquela expansão, ou seja, dados pela fórmula (que em muitos casos simplifica o cálculo dos momentos):

$$E(X^n) = (-1)^n L_X^{(n)}(0).$$

[1] i.e., a convergência das *f.d.* nos pontos de continuidade da função limite, também chamada *convergência fraca* ou *convergência em lei.*

[2] recorde-se que o momento de ordem n, $n \in \mathbb{N}$, de uma *v.a.* X é o valor médio $E(X^n)$, caso este exista, i.e., caso $E(|X^n|) < +\infty$.

Como a transformada de Laplace identifica a distribuição, temos que a sequência dos momentos, no caso de existir transformada de Laplace, também identifica a distribuição.

No entanto, podem existir os momentos de todas as ordens de uma *v.a.* X e não existir transformada de Laplace, tal como acontece no caso da distribuição lognormal[3] ou no caso da *f.d.p.* dada por $f(x) = ce^{-|x|^\alpha}$, sendo $0 < \alpha < 1$. Nestes casos, é possível que existam diferentes distribuições que tenham rigorosamente a mesma sequência de momentos (de todas as ordens). Tal é o caso da distribuição lognormal: há exemplos de distribuições contínuas e discretas que têm exatamente os mesmos momentos que a lognormal. Portanto, a sequência dos momentos de uma *v.a.* não identifica a distribuição (a não ser que exista a transformada de Laplace).

Como a existência da transformada de Laplace implica a existência dos momentos de qualquer ordem, conclui-se que basta que não exista algum momento, como acontece por exemplo com a distribuição de Cauchy, para que não exista transformada de Laplace. Finalmente, no caso de não existirem momentos de uma *v.a.*, a partir de certa ordem, não é possvel identificar a distribuição a partir dos momentos que existem (vd. Rohatgi et al., *An Introduction to Probability and Statistics*).

Exemplo 1.1 *Se* $X \frown Exp(\lambda)$, $\lambda > 0$, *a transformada de Laplace existe e é dada por*

$$L(t) = E(e^{-tX}) = \int_0^{+\infty} e^{-tx}\lambda e^{-\lambda x}dx = \lambda \int_0^{+\infty} e^{-(t+\lambda)x}dx = \frac{-\lambda}{\lambda+t}\, e^{-(\lambda+t)x}\Big|_0^{+\infty} = \frac{\lambda}{\lambda+t},$$

se $t > -\lambda$ *(para* $t+\lambda < 0$ *o integral não converge pois* $\lim\limits_{x\to+\infty} e^{-(\lambda+t)x} = +\infty$*). Logo*

$$L^{(n)}(t) = (-1)^n n! \frac{\lambda}{(\lambda+t)^{n+1}},$$

donde $E(X^n) = (-1)^n L^{(n)}(0) = \dfrac{n!}{\lambda^n}$. *Em particular,* $E(X) = \dfrac{1}{\lambda}$ *e* $E(X^2) = \dfrac{2}{\lambda^2}$, *donde* $\mathrm{var}(X) = \dfrac{1}{\lambda^2}$.

As transformadas de Laplace têm larga aplicação na determinação da distribuição da soma de *v.a.* independentes. Dadas duas *v.a.*, X e Y, com *f.d.* F e G, respetivamente, chama-se *convolução* destas duas distribuições à *f.d.* (ou à *f.d.p.*/*f.m.p.* no caso contínuo/discreto) de $X+Y$, que representamos por $F*G$. Esta operação, que é comutativa e associativa, traduz-se, no caso de X e Y serem independentes e absolutamente contínuas, com *f.d.p.* f e g, respectivamente, na fórmula

$$F*G\ (s) = P(X+Y \le s) = \int_{-\infty}^{+\infty} P(X \le s-y \mid Y=y)\ g(y)dy = \int_{-\infty}^{+\infty} F(s-y)\ g(y)dy$$

[3] Diz-se que Y tem distribuição lognormal se $Y = e^X$, sendo $X \frown N(\mu,\sigma)$.

ou seja, reduz-se ao cálculo de $\int_{-\infty}^{+\infty} F(s-y)\ g(y)dy$ ou $\int_{-\infty}^{+\infty} G(s-y)\ f(y)dy$, que poderá ser complicado (repare que a fórmula em causa é uma versão generalizada do Teorema da Probabilidade Total; a fórmula para o caso discreto é semelhante[4]). No entanto, a transformada de Laplace de $X+Y$ vai ser simplesmente o produto das transformadas de Laplace das parcelas. De facto (recorrendo a que o valor médio do produto de *v.a.* independentes é igual ao produto dos valores médios, e a que funções de *v.a.* independentes são ainda *v.a.* independentes), temos

$$L_{X+Y}(t) = E\left(e^{-t(X+Y)}\right) = E\left(e^{-tX}e^{-tY}\right) = E\left(e^{-tX}\right)E\left(e^{-tY}\right) = L_X(t)L_Y(t).$$

O resultado generaliza-se imediatamente a n parcelas:

Dadas n *v.a.* independentes $X_1, X_2, \ldots, X_n$, com respetivas transformadas de Laplace $L_1(t), L_2(t), \ldots, L_n(t)$, então a transformada de Laplace da sua soma, $S_n = X_1 + X_2 + \ldots + X_n$, é o produto dessas mesmas transformadas, i.e., é dada por $L_{S_n}(t) = L_1(t)L_2(t)\ldots L_n(t)$.
Em particular, se as n *v.a.* forem *i.i.d.*, com transformada de Laplace L, então a transformada de Laplace de S_n é dada por $L_{S_n}(t) = L^n(t)$.

As transformadas de Laplace das distribuições *binomial*, *Poisson*, *uniforme*, *normal* e *gama* calculam-se com facilidade. A partir destas, conclui-se imediatamente que a soma de n *v.a.* independentes X_i com distribuição $Poisson(\lambda_i)$, $bi(n_i, p)$, $gama(\alpha_i, \lambda)$, $N(\mu_i, \sigma_i)$, terá distribuição $Poisson(\sum \lambda_i)$, $bi(\sum n_i, p)$, $gama(\sum \alpha_i, \lambda)$, $N(\sum \mu_i, \sqrt{\sum \sigma_i^2})$, respetivamente.

No caso de *v.a.* não negativas, as transformadas de Laplace são caracterizadas por serem funções *completamente monótononas*, i.e., tais que $(-1)^n L^{(n)}(t) \geq 0$ (Bernstein).

Para um par aleatório (X, Y) define-se a sua transformada de Laplace por

$$L(t, u) = E(e^{-tX-uY})$$

desde que este valor médio exista numa vizinhança do ponto $(t, u) = (0, 0)$.

Tal como no caso univariado, esta transformada determina univocamente a distribuição conjunta do par. Se existir a transformada de Laplace do par (X, Y), existem os momentos conjuntos de qualquer ordem, que são dados por

$$E(X^m Y^n) = (-1)^{m+n} \left.\frac{\partial^{m+n} L(t, u)}{\partial t^m \partial u^n}\right|_{t=u=0}$$

[4]usa-se também a notação mais geral $\int \ldots dF(x)$, que equivale a $\int \ldots f(x)dx$ no caso contínuo e ao correspondente somatório no lugar do integral, no caso discreto.

Tem-se ainda que X e Y são independentes sse $L(t,u) = L(t,0)\ L(0,u)$. Aplicando este resultado, pode provar-se por exemplo que $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ e $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ são independentes, no caso de amostragem aleatória proveniente do modelo $N(\mu,\sigma)$.

1.2 Outras transformadas

Existem outras transformadas relevantes, entre as quais salientamos:

- a *função geradora de momentos* (*f.g.m.*), definida por $M(t) = E(e^{tX})$, caso este valor médio exista para $|t| < t_0$ (para algum $t_0 > 0$). Temos assim $M(t) = L(-t)$, para $|t| < t_0$. O nome desta transformada deriva do facto de o momento de ordem n ser obtido à sua custa pela fórmula $E(X^n) = M_X^{(n)}(0)$.

- a *função geradora de probabilidades* (*f.g.p.*), definida por $G(z) = E(z^X)$, que converge pelo menos para $|z| \leq 1$, utilizada sobretudo para *v.a.* discretas com suporte $\{0, 1, 2, \dots\}$. Neste caso, as probabilidades $p_n = P(X = n)$, $n = 0, 1, 2, \dots$, relacionam-se com a *f.g.p.* pela fórmula $p_n = \frac{1}{n!}G^{(n)}(0)$, que justifica o nome desta transformada. Esta função gera também os chamados *momentos factoriais* (caso estes existam), definidos por $E(X(X-1)\dots(X-r+1))$, para $r = 1, 2, \dots$, uma vez que $G^{(r)}(1) = E(X(X-1)\dots(X-r+1))$.

- a *função característica* (*f.c.*), $\phi : \mathbb{R} \longrightarrow \mathbb{C}$, definida por $\phi(t) = E(e^{itX})$, $t \in \mathbb{R}$, também conhecida por transformada de Fourier.

Note-se que a *f.c.* tem a vantagem (em relação à transformada de Laplace e à *f.g.m.*) de existir sempre (para qualquer t real e para qualquer *v.a.* X). De facto, como $|e^{itx}| \leq 1$, $\forall x \in \mathbb{R}$, resulta então $|\ E(e^{itX})\ | = |\int e^{itx} dF(x)\ | \leq \int |\ e^{itx}\ |\ dF(x) \leq \int dF(x) = 1$.

1.3 Distribuição gama

A função *gama* de Euler é definida por

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx$$

para $\alpha > 0$ (mais geralmente, esta função encontra-se definida para $\mathrm{Re}(\alpha) > 0$).

Prova-se facilmente (integrando por partes) que $\Gamma(\alpha+1) = \alpha\ \Gamma(\alpha)$ e como $\Gamma(1) = 1$, conclui-se que $\Gamma(n+1) = n!$, para $n = 1, 2, 3, \dots$. Repare-se que a função *gama* estende a noção de "factorial" de um número inteiro positivo a qualquer número real positivo.

Esta função encontra-se devidamente estudada e tabelada. Tem-se ainda $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Tanto no MATLAB como no R obtém-se o valor de $\Gamma(x)$ com o comando `gamma(x)`.

De uma *v.a.* X com suporte[5] $[0, +\infty[$ e *f.d.p.*

$$f(x) = \frac{\lambda^\alpha e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)},\ x > 0$$

diz-se que tem distribuição *gama* com parâmetros α ($\alpha > 0$) e λ ($\lambda > 0$), e escreve-se $X \frown Gama(\alpha, \lambda)$. Em particular, a distribuição $Gama(1, \lambda)$ coincide com a distribuição $Exp(\lambda)$. Na figura 1 encontram-se gráficos de densidades $Gama(\alpha, 1)$, para alguns valores de α. No R, as funções associadas a esta distribuição são `dgamma`, `pgamma`, `qgamma` e `rgamma`, e no MATLAB são `gammapdf`, `gammacdf`, `gammainv` e `gammarnd`.

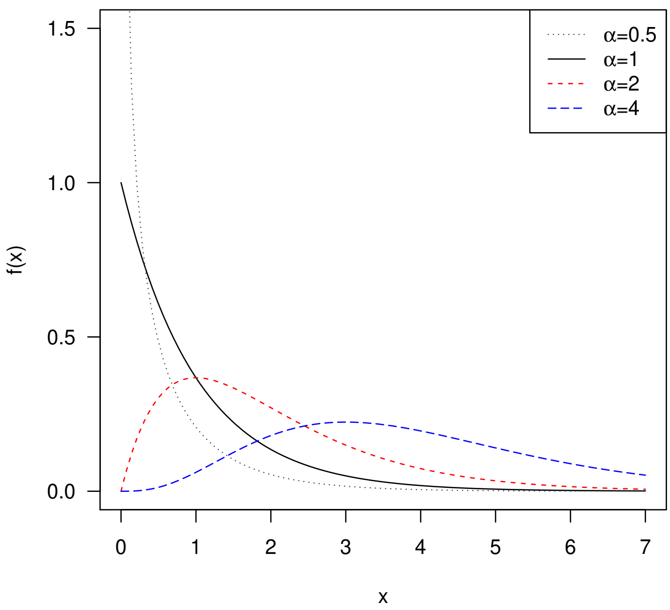

Figure 1: Densidades $Gama(\alpha, \lambda)$, para $\alpha = 0.5, 1, 2, 4$ e $\lambda = 1$

Note-se que $\delta = \frac{1}{\lambda}$ é um parâmetro de escala enquanto que α é um parâmetro de forma. Seguem-se alguns resultados sobre esta família de distribuições (as demonstrações ficam como exercício).

[5]rigorosamente, o suporte de uma distribuição de probabilidade (ou da correspondente *v.a.*) é o menor subconjunto fechado (de $\mathbb{R}$) que tem probabilidade 1.

1. A transformada de Laplace da distribuição $Gama(\alpha, \lambda)$ é $L(t) = \left(\frac{\lambda}{\lambda+t}\right)^{\alpha}$, $t > -\lambda$.

2. O valor médio e a variância da distribuição $Gama(\alpha, \lambda)$ são respetivamente $\mu = \frac{\alpha}{\lambda}$ e $\sigma^2 = \frac{\alpha}{\lambda^2}$. O momento de ordem n é dado pela fórmula $\mu_n = \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)}\lambda^{-n}$. Os coeficientes de assimetria e curtose são respetivamente $\beta_1 = \frac{2}{\sqrt{\alpha}}$ e $\beta_2 = 3 + \frac{6}{\alpha}$. Note-se que $\beta_1 > 0$ (assimetria positiva) e $\beta_2 > 3$ (superior à curtose da normal). Esta distribuição é amodal para $\alpha < 1$, e tem moda $\frac{\alpha-1}{\lambda}$ para $\alpha \geq 1$.

3. Recorrendo às transformadas de Laplace, prova-se que a soma de n *v.a.* independentes $Gama(\alpha_i, \lambda)$ tem distribuição $Gama(\alpha, \lambda)$, com $\alpha = \alpha_1 + \alpha_2 + \ldots + \alpha_n$.

4. Para n grande, a distribuição $Gama(n, \lambda)$ é aproximadamente normal. De facto, se $X \frown Gama(n, \lambda)$, então $X = \sum_{i=1}^{n} Y_i$, com $Y_1, Y_2, \ldots, Y_n$ *v.a. i.i.d.* $Gama(1, \lambda)$. Pelo TLC, resulta que $\frac{X-n/\lambda}{\sqrt{n}/\lambda}$ tem distribuição limite $N(0, 1)$. Mais geralmente, para α grande, a distribuição $Gama(\alpha, \lambda)$ é aproximadamente $N(\frac{\alpha}{\lambda}, \frac{\sqrt{\alpha}}{\lambda})$.

5. A distribuição $Gama(\frac{n}{2}, \frac{1}{2})$ é também conhecida por χ^2_n (*qui-quadrado* com n graus de liberdade). Dadas n *v.a.* $Z_1, Z_2, \ldots, Z_n$ *i.i.d.* com $Z \frown N(0, 1)$, prova-se que $Y = \sum_{i=1}^{n} Z_i^2 \frown \chi^2_n$, recorrendo às transformadas de Laplace. De facto, Z^2 tem *f.d.* $G(z) = P(Z^2 \leq z) = 2\Phi(\sqrt{x}) - 1$, $x > 0$, e *f.d.p.* $g(z) = e^{-x/2}\,(2\pi x)^{-\frac{1}{2}}$, $x > 0$, donde $Z^2 \frown Gama(\frac{1}{2}, \frac{1}{2})$. Mas como $Z_1, Z_2, \ldots, Z_n$ são mutuamente independentes, também o são $Z_1^2, Z_2^2, \ldots, Z_n^2$. Logo $Y \frown Gama\left(\frac{n}{2}, \frac{1}{2}\right)$, ou seja, $Y \frown \chi^2_n$.

1.4 Exercícios

1. Calcule as transformadas de Laplace das distribuições $Bi(n, p)$, $Poisson(\lambda)$, $U[a, b]$, $Exp(\lambda)$, $N(\mu, \sigma)$. e $Gama(\alpha, \lambda)$.

2. Calcule o valor médio, a variância e os coeficientes de assimetria e curtose das distribuições referidas no exercício anterior, recorrendo a transformadas de Laplace. Para as distribuições *normal* e *gama*, determine o momento de ordem r, $r > 0$.

3. Prove que a soma de *v.a.* $N(\mu_i, \sigma_i)$ $[Poisson(\lambda_i)$, $bi(n_i, p)]$ $Gama(\alpha_i, \lambda))$ independentes é ainda normal (*Poisson*, *binomial*, *gama*) e identifique os parâmetros.

4. Prove que a soma dos quadrados de n *v.a. i.i.d.* $N(0, 1)$ segue a lei $Gama(\frac{n}{2}, \frac{1}{2})$.

5. Se X for uma *v.a.* absolutamente contínua com *f.d.p.* par, e supondo que existe a sua transformada de Laplace, prove que esta também é par. Ilustre com exemplos.