

MAE0314 - Análise Estatística

Fernando Henrique Ferraz Pereira da Rosa

24 de março de 2004

Lista 1¹

2.2.1 Considere a distribuição de Pareto bivariada definida por:

$$f(x, y) = c(x + y - 1)^{-p-2}, \quad x, y > 1$$

onde $p > 0$. Mostre que c deve ser igual a $p(p+1)$. Calcule a distribuição acumulada conjunta e marginal de x e y . Se $p > 1$ mostre que x e y tem esperança $\frac{p}{p-1}$, e se $p > 2$ mostre que a matriz de covariância é dada por

$$\frac{1}{(p-1)^2(p-2)} \begin{bmatrix} p & 1 \\ 1 & p \end{bmatrix}.$$

O que acontece com as esperanças se $0 < p \leq 1$? Se $p > 2$ mostre que $\text{corr}(x, y) = 1/p$, e que a variância generalizada e a variância total são

$$\frac{(p+1)}{(p-1)^3(p-2)^2} \quad \text{e} \quad \frac{2p}{(p-1)^2(p-2)}.$$

Calcule a curva de regressão e a curva cedástica de x em y .

Para que $f(x, y)$ seja uma função de densidade temos que a condição abaixo deve ser satisfeita:

$$\int_1^\infty \int_1^\infty c(x + y - 1)^{-p-2} dx dy = 1 \Rightarrow \int_1^\infty \underbrace{\int_1^\infty (x + y - 1)^{-p-2} dx}_{I} dy = \frac{1}{c} \quad (1)$$

Resolvamos (I) da equação 1. Tomemos a transformação $u = x + y - 1$, $du = dx$:

$$(I) = \int_y^\infty u^{-p-2} du = \frac{u^{-p-1}}{-p-1} \Big|_y^\infty = \frac{y^{-p-1}}{p+1} \quad (2)$$

¹Powered by L^AT_EX 2_ε and R 1.8.1

onde a última passagem é válida somente para $p > 1$. Voltando então (I) na primeira expressão obtemos:

$$\begin{aligned} \frac{1}{c} &= \int_1^{\infty} \frac{y^{-p-1}}{p+1} dy \Rightarrow \frac{p+1}{c} = \int_1^{\infty} y^{-p-1} dy = \frac{y^{-p}}{-p} \Big|_1^{\infty} = \frac{1}{p} \\ &\Rightarrow \frac{p+1}{c} = 1p \Rightarrow c = p(p+1) \end{aligned}$$

Mostrando a identidade pedida. Assim a densidade conjunta da Pareto bivariada é dada por:

$$f(x, y) = p(p+1)(x+y-1)^{-p-2} \quad x, y > 1. \quad (3)$$

Para calcularmos a função de distribuição acumulada conjunta, basta usarmos sua definição:

$$F(a, b) = P(X \in]1, a), Y \in]1, b) = \int_1^a \int_1^b \underbrace{p(p+1)(x+y-1)^{-p-2}}_I dy dx \quad (4)$$

Calculemos (I) em 4, tomando a transformação $u = x + y - 1$, $du = dy$:

$$\begin{aligned} \frac{(I)}{p(p+1)} &= \int_x^{x+b-1} u^{-p-2} dy = \frac{u^{-p-1}}{-p-1} \Big|_x^{x+b-1} = \frac{(x+b-1)^{-p-1}}{-p-1} + \frac{x^{-p-1}}{p+1} \\ &= \frac{-(x+b-1)^{-p-1} + x^{-p-1}}{p+1} \\ &\Rightarrow (I) = p(-(x+b-1)^{-p-1} + x^{-p-1}) \end{aligned}$$

Voltando a expressão encontrada para (I) em 4, obtemos:

$$F(a, b) = p \left(\underbrace{\int_1^a -(x+b-1)^{-p-1} dx}_{II} + \underbrace{\int_1^a x^{-p-1} dx}_{III} \right) \quad (5)$$

Para resolver (II) em 5 tomamos $u = x + b - 1$, $du = dx$:

$$(II) = \int_b^{a+b-1} -u^{-p-1} dx = \frac{-u^{-p}}{-p} \Big|_b^{a+b-1} = \frac{-(a+b-1)^{-p} + b^{-p}}{-p}$$

Resolvemos (III) diretamente:

$$(III) = \frac{x^{-p}}{-p} \Big|_1^a = \frac{a^{-p} - 1}{-p}$$

Voltando (II) e (III) em 5 obtemos:

$$\begin{aligned} F(a, b) &= (a + b - 1)^{-p} - b^{-p} - a^{-p} + 1 \\ &= 1 + (a + b - 1)^{-p} - (a^{-p} + b^{-p}) \end{aligned}$$

Para encontrar as funções de distribuições de x e y agora, vamos primeiro encontrar as funções de densidade de x e y , a partir das marginais da densidade conjunta. Começemos calculando a densidade marginal de x . Para isso basta integrarmos (3) em y :

$$f_x(x) = \int_1^\infty p(p+1)(x+y-1)^{-p-2} dy = p(p+1) \int_1^\infty (x+y-1)^{-p-2} dy$$

Notemos entretanto que resolver essa integral é resolver a mesma integral (I) que resolvemos na equação (2), trocando y por x . Temos então que:

$$f_x(x) = p(p+1) \frac{x^{-p-1}}{p+1} = px^{-p-1} \quad (6)$$

Analogamente obtemos:

$$f_y(y) = py^{-p-1} \quad (7)$$

Agora basta usar a definição para calcular as funções de probabilidade:

$$F_X(a) = P(X \in]1, a)) = \int_1^a px^{-p-1} dx = p \frac{x^{-p}}{-p} \Big|_1^a = 1 - a^{-p}$$

Analogamente obtemos:

$$F_Y(b) = 1 - b^{-p}$$

Podemos concluir a partir daí que x e y não são independentes pois $F_X F_Y \neq F_{XY}$.

Calculemos agora a esperança de x e y :

$$E(X) = \int_1^\infty xpx^{-p-1} dx = p \int_1^\infty x^{-p} dx = p \frac{x^{-p+1}}{-p+1} \Big|_1^\infty = \frac{p}{p-1}.$$

Donde a última igualdade só é válida de $p > 1$, caso contrário a integral é divergente. Analogamente temos que $E(Y) = \frac{p}{p-1}$.

Para calcularmos a matriz de covariâncias precisamos em primeiro lugar de $Var(X)$, $Var(Y)$ e $Cov(X, Y)$. Notemos que:

$$Var(X) = E(X^2) - E^2(X)$$

Onde apenas a primeira parcela dessa diferença é desconhecida. Achemos essa esperança:

$$E(X^2) = \int_1^\infty x^2 p x^{-p-1} dx = p \int_1^\infty x^{-p+1} dx = p \left. \frac{x^{-p+2}}{-p+2} \right|_1^\infty = \frac{p}{p-2}.$$

Onde a última igualdade só é válida para $p > 2$, caso contrário a integral é divergente. A partir daí temos que:

$$Var(X) = \frac{p}{p-2} - \left(\frac{p}{p-1} \right)^2 = \frac{p}{(p-1)^2(p-2)}$$

Analogamente obtemos que $Var(Y) = Var(X)$. Para calcular $Cov(X, Y)$ notemos que:

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

onde a única quantidade desconhecida até o momento é $E(XY)$. Obtenhamos essa quantidade:

$$\begin{aligned} E(XY) &= \int_1^\infty \int_1^\infty xy f(x, y) dy dx = \int_1^\infty \int_1^\infty xyp(p+1)(x+y-1)^{-p-2} dy dx \\ &= \int_1^\infty p(p+1)x \underbrace{\int_1^\infty y(x+y-1)^{-p-2} dy}_{I} dx \end{aligned} \quad (8)$$

Calulemos (I) na equação (8), tomando a substituição $u = x + y - 1$, $du = dy$, $y = u - x + 1$:

$$\begin{aligned} (I) &= \int_x^\infty (u-x+1)u^{-p-2} du = \int_x^\infty u^{-p-1} - xu^{-p-2} + u^{-p-2} du \\ &= \left[\frac{u^{-p}}{-p} - \frac{xu^{-p-1}}{-p-1} + \frac{u^{-p-1}}{-p-1} \right]_x^\infty = \left[\frac{-u^{-p}}{p} + \frac{xu^{-p-1}}{p+1} - \frac{u^{-p-1}}{p+1} \right]_x^\infty \\ &= \frac{x^{-p}}{p} - \frac{x^{-p}}{p+1} + \frac{x^{-p-1}}{p+1} \end{aligned}$$

Voltando (I) em (8) temos:

$$\begin{aligned}
E(XY) &= \int_1^{\infty} (p+1)x^{-p+1} - px^{-p+1} + px^{-p} dx \\
&= \left[\frac{(p+1)x^{-p+2}}{-p+2} - \frac{px^{-p+2}}{-p+2} + \frac{px^{-p+1}}{-p+1} \right]_1^{\infty} \\
&= \left[-\frac{(p+1)x^{-p+2}}{p-2} + \frac{px^{-p+2}}{p-2} - \frac{px^{-p+1}}{p-1} \right]_1^{\infty} \\
&= \frac{p+1}{p-2} + \frac{p}{p-1} - \frac{p}{p-2} = \frac{p}{p-1} + \frac{1}{p-2} \\
&= \frac{p^2 - p - 1}{(p-1)(p-2)}
\end{aligned}$$

Temos assim que:

$$\begin{aligned}
Cov(X, Y) &= \frac{p^2 - p - 1}{(p-1)(p-2)} - \left(\frac{p}{p-1} \right)^2 = \frac{(p-1)(p^2 - p - 1) - p^2(p-2)}{(p-1)^2(p-2)} \\
&= \frac{p^3 - p^2 - p - 2 + p + 1 - p^3 + 2p^2}{(p-1)^2(p-2)} = \frac{1}{(p-1)^2(p-2)}.
\end{aligned}$$

E portanto a matriz de covariâncias será dada por:

$$S = \begin{bmatrix} \frac{p}{(p-1)^2(p-2)} & \frac{1}{(p-1)^2(p-2)} \\ \frac{1}{(p-1)^2(p-2)} & \frac{p}{(p-1)^2(p-2)} \end{bmatrix} = \frac{1}{(p-1)^2(p-2)} \begin{bmatrix} p & 1 \\ 1 & p \end{bmatrix}$$

Provando a identidade pedida. Se $0 < p \leq 1$ a integral utilizada para calcular a esperança é divergente, logo a esperança seria “infinita”. Notemos agora que:

$$corr(x, y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}} = \frac{\frac{1}{(p-1)^2(p-2)}}{\frac{p}{(p-1)^2(p-2)}} = \frac{1}{p}.$$

Provando a outra identidade pedida. Calculemos agora a variância generalizada:

$$|S| = \frac{p^2 - 1}{(p-1)^2(p-2)} = \frac{(p+1)(p-1)}{(p-1)^2(p-2)} = \frac{(p+1)}{(p-1)^3(p-2)}$$

E a variância total:

$$tr(S) = \frac{p}{(p-1)^2(p-2)} + \frac{p}{(p-1)^2(p-2)} = \frac{2p}{(p-1)^2(p-2)}$$

O que prova as outras duas identidades pedidas. Por fim, calculemos a curva da regressão de x em y :

$$\begin{aligned} E(X|Y) &= \int_1^\infty x f(x|y) dx = \int_1^\infty x \frac{f(x,y)}{f(y)} dx \\ &= \int_1^\infty \frac{x p(p+1)(x+y-1)^{-p-2}}{p y^{-p-1}} dx \\ &= (p+1)y^{p+1} \int_1^\infty x(x+y-1)^{-p-2} dx \end{aligned}$$

Em (8) entretanto já calculamos a integral acima mas para y . Para x as contas são as mesmas, obtendo:

$$\begin{aligned} E(X|Y) &= (p+1)y^{p+1} \left(\frac{y^{-p}}{p} - \frac{y^{-p}}{p+1} + \frac{y^{-p-1}}{p+1} \right) = \frac{(p+1)y}{p} - y + 1 \\ &= 1 + \frac{y}{p} \end{aligned}$$

Finalmente, calculemos a curva cedástica de x em y :

$$Var(X|Y) = E(X^2|Y) - E^2(X|Y) \quad (9)$$

Calculando a esperança desconhecida, temos:

$$\begin{aligned} E(X^2|Y) &= \int_1^\infty x^2 f(x|y) dx = \int_1^\infty x^2 \frac{f(x,y)}{f(y)} dx \\ &= \int_1^\infty x^2 \frac{p(p+1)(x+y-1)^{-p-2}}{p y^{-p-1}} dx \\ &= (p+1)y^{p+1} \underbrace{\int_1^\infty x^2(x+y-1)^{-p-2} dx}_I \end{aligned} \quad (10)$$

Vamos então resolver (I), pela regra do produto. Temos:

$$\begin{aligned} u &= x^2 & du &= 2x dx \\ v &= \frac{(x+y-1)^{-p-1}}{-p-1} & dv &= (x+y-1)^{-p-2} dx \end{aligned}$$

Como $\int u dv = uv - \int v du$ temos:

$$\begin{aligned}
(I) &= \frac{x^2(x+y-1)^{-p-1}}{-p-1} \Big|_1^\infty - \int_1^\infty \frac{(x+y-1)^{-p-1}}{-p-1} 2x dx \\
&= \frac{y^{-p-1}}{p+1} + \frac{2}{p+1} \underbrace{\int_1^\infty x(x+y-1)^{-p-1} dx}_{II} \quad (11)
\end{aligned}$$

Resolvamos agora (II), por substituição. Tomemos $u = x + y - 1$, $du = dx$, $x = u + 1 - y$:

$$\begin{aligned}
(II) &= \int_y^\infty (u - y + 1)u^{-p-1} du = \int_y^\infty u^{-p} + (1 - y)u^{-p-1} du \\
&= \left[\frac{u^{-p+1}}{-p+1} + (1 - y) \frac{u^{-p}}{-p} \right]_y^\infty = \left[-\frac{u^{-p+1}}{p-1} - (1 - y) \frac{u^{-p}}{p} \right]_y^\infty \\
&= \frac{y^{-p+1}}{p-1} + (1 - y) \frac{y^{-p}}{p} = \frac{y^{-p+1}}{p-1} + \frac{y^{-p}}{p} - \frac{y^{-p+1}}{p}
\end{aligned}$$

Voltemos agora (II) na equação 11:

$$(I) = \frac{y^{-p-1}}{p+1} + \frac{2}{p+1} \left(\frac{y^{-p+1}}{p-1} + \frac{y^{-p}}{p} - \frac{y^{-p+1}}{p} \right)$$

Voltemos agora (I) na equação 10:

$$\begin{aligned}
E(X^2|Y) &= (p+1)y^{p+1} \left(\frac{y^{-p-1}}{p+1} + \frac{2}{p+1} \left(\frac{y^{-p+1}}{p-1} + \frac{y^{-p}}{p} - \frac{y^{-p+1}}{p} \right) \right) \\
&= 1 + \frac{2y^2}{p-1} + \frac{2y}{p} - \frac{2y^2}{p}
\end{aligned}$$

Voltando então o resultado obtido em na equação 9:

$$\begin{aligned}
Var(X|Y) &= 1 + \frac{2y^2}{p-1} + \frac{2y}{p} - \frac{2y^2}{p} - \left(1 + \frac{y}{p} \right)^2 \\
&= 1 + \frac{2y^2}{p-1} + \frac{2y}{p} - \frac{2y^2}{p} - 1 - \frac{2y}{p} - \frac{y^2}{p^2} \\
&= \frac{2y^2}{p-1} - \frac{2y^2}{p} - \frac{y^2}{p^2} \\
&= \frac{2y^2p^2 - 2y^2(p-1)p - y^2(p-1)}{(p-1)p^2} \\
&= \frac{y^2(p+1)}{(p-1)p^2}
\end{aligned}$$

Que é a curva cedástica de x em y . Como ela não é constante, dizemos que a curva de regressão de x em y é heterocedástica.

- 2.2.3 Se $p \geq 2$ e $\Sigma(p \times p)$ é a matriz de equicorrelação, $\Sigma = (1-\alpha)I + \alpha 11'$, mostre que $\Sigma \geq 0$ se e somente se $-(p-1)^{-1} \leq \alpha \leq 1$. Se $-(p-1)^{-1} < \alpha < 1$, mostre que

$$\Sigma^{-1} = (1-\alpha)^{-1}[I - \alpha(1 + (p-1)\alpha)^{-1}11'].$$

Se $\mu_1' = (\delta, 0')$ e $\mu_2 = 0$ mostre que a distância de Mahalanobis entre eles é dada por

$$\Delta(\mu_1, \mu_2) = \delta \left[\frac{1 + (p-2)\alpha}{(1-\alpha)(1 + (p-1)\alpha)} \right]^{1/2}.$$

- 2.2.4 (Desigualdades de Fréchet). Sejam x e y variáveis aleatórias com f.d.p $H(x, y)$ e f.d.p's marginais $F(x)$ e $G(y)$. Mostre que

$$\max(F + G - 1, 0) \leq H \leq \min(F, G).$$

Quando $F = 0$ ou $G = 0$, ou $F = G = 0$, a relação é trivial. Resta mostrar que a relação é válida para F e G não nulas. Começemos mostrando a desigualdade da direita. Notemos que pelo Teorema de Bayes:

$$H(x, y) = P(X \leq x, Y \leq y) = P(X \leq X/Y \leq y)P(Y \leq y)$$

Portanto:

$$P(X \leq X/Y \leq y)P(Y \leq y) \leq \min(P(X \leq x), P(Y \leq y)) \Rightarrow H \leq \min(F, G)$$

Mostremos portanto que essa suposição é válida. Começemos mostrando que H é menor que F e então que também é menor que G , e logo será menor que o mínimo.

$$\underbrace{P(X \leq X/Y \leq y)}_{\leq 1} P(Y \leq y) \leq P(Y \leq y) \Rightarrow H \leq G$$

Analogamente, temos:

$$\underbrace{P(Y \leq y/X \leq x)}_{\leq 1} P(X \leq x) \leq P(X \leq x) \Rightarrow H \leq F$$

E assim:

$$H \leq F, H \leq G \Rightarrow H \leq \min(F, G)$$

O que prova o limite superior da desigualdade. Para a demonstração do limite inferior tomemos raciocínio análogo:

$$\begin{aligned} F + G - 1 \leq H &\Leftrightarrow P(X \leq x) + P(Y \leq y) - 1 \leq P(X \leq x/Y \leq y)P(Y \leq y) \\ &\Leftrightarrow \frac{P(X \leq x) - 1}{P(Y \leq y)} \leq \frac{P(X \leq x, Y \leq y)}{P(Y \leq y)} \Leftrightarrow P(X \leq x) - 1 \leq P(X \leq x, Y \leq y) \\ &\Leftrightarrow P(X \leq x) \leq P(X \leq x, Y \leq y) + 1 \Leftrightarrow F \leq H + 1 \end{aligned}$$

Donde a validade da última afirmação é imediata já que H e F são f.d.p's. Temos então portanto:

$$F + G - 1 \leq H, 0 \leq H \Rightarrow \max(F + G - 1, 0) \leq H$$

Assim fica provada a validade do limite inferior e a demonstração das desigualdades de Fréchet está completa.

2.2.6 Sejam x e y variáveis aleatórias tais que a densidade de y é

$$g(y) = \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2}, y > 0$$

e a densidade condicional de x dado y é

$$f(x|y) = (2\pi)^{-1/2} y^{1/2} e^{-yx^2/2}$$

Mostre que $E[x|y] = 0$ e que portanto $E_y[E[x|y]] = 0$, mas que entretanto a média incondicional de x não existe.

$$\begin{aligned} E[x|y] &= \int_{\mathbb{R}} x f(x|y) dx = \int_{\mathbb{R}} x (2\pi)^{-1/2} y^{1/2} e^{-yx^2/2} dx \\ &= (2\pi)^{-1/2} y^{1/2} \int_{\mathbb{R}} x e^{-yx^2/2} dx \end{aligned}$$

Tomemos $u = \frac{-yx^2}{2}$, $du = -yxdx$:

$$E[X|Y] = (2\pi)^{-1/2} y^{1/2} \int_{-\infty}^{\infty} x e^u \frac{du}{-yx} = 0$$

Assim $E_y[E[x|y]] = 0$. Notemos agora que:

$$f(x|y) = \frac{f(x, y)}{f(y)} \Rightarrow f(x, y) = f(x|y)f(y)$$

Assim:

$$f(x, y) = (2\pi)^{-1/2} y^{1/2} e^{-yx^2/2} \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2} = \frac{1}{2\pi} e^{-y \frac{(x^2+1)}{2}}$$

Integrando em y para obtermos a marginal de x temos:

$$\begin{aligned} f(x) &= \int_0^\infty \frac{1}{2\pi} e^{-y \frac{(x^2+1)}{2}} dy = \frac{1}{2\pi} \int_0^\infty e^{-y \frac{(x^2+1)}{2}} dy \\ &= \frac{1}{2\pi} \frac{2}{x^2+1} \underbrace{\int_0^\infty \frac{x^2+1}{2} e^{-y \frac{(x^2+1)}{2}} dy}_{f(u), \text{ onde } U \sim \exp\left(\frac{-2}{x^2+1}\right)} \\ &= \frac{1}{\pi(x^2+1)} \Rightarrow X \text{ segue distribuição Cauchy} \end{aligned}$$

Como x segue distribuição Cauchy, a esperança de x não existe².

2.6.2 *Uma Poisson multivariada.* Sejam u_0, u_1, \dots, u_p variáveis aleatórias Poisson com parâmetros $\lambda_0, \lambda_1 - \lambda_0, \dots, \lambda_p - \lambda_0$, respectivamente. Escreva a distribuição conjunta de $x_i = u_0 + u_i, i = 1, \dots, p$, e mostre que as distribuições marginais de x_1, \dots, x_p são todas Poisson. Para $p = 2$, mostre que a função de probabilidade é dada por

$$f(x, y) = \exp(-\lambda_1 - \lambda_2 + \lambda_0) \frac{a^x b^y}{x! y!} \sum_{r=0}^s \frac{x^{(r)} y^{(r)} \lambda_0^r}{a^r b^r r!},$$

onde $s = \min(x, y)$, $a = \lambda_1 - \lambda_0$, $b = \lambda_2 - \lambda_0$, $\lambda_1 > \lambda_0 > 0$, $\lambda_2 > \lambda_0 > 0$, e $x^{(r)} = x(x-1)\dots(x-r+1)$. Ainda,

$$E(y|x) = b + (\lambda_0/\lambda_1)x, \quad \text{Var}(y|x) = b + (a\lambda_0/\lambda_1^2)x.$$

Sobre

A versão eletrônica desse arquivo pode ser obtida em <http://www.feferraz.net>

Copyright (c) 1999-2005 Fernando Henrique Ferraz Pereira da Rosa.
É dada permissão para copiar, distribuir e/ou modificar este documento sob os termos da Licença de Documentação Livre GNU (GFDL), versão 1.2, publicada pela Free Software Foundation;
Uma cópia da licença em está inclusa na seção intitulada "Sobre / Licença de Uso".

²James R. Barry. (2002) *Probabilidade: um curso em nível intermediário*.