Aula 19: Distribuições condicionais

Disciplina: PGE950 - Probabilidade

Programa de Pós-graduação em Estatística da UFPE

Prof. Pablo M. Rodriguez

http://www.pablo-rodriguez.org

Universidade Federal de Pernambuco

CCEN

Centro de Ciências Exatas e da Natureza

Conteúdo desta aula

- Distribuições condicionais.
- Definições e exemplos.

Lembrete: Probabilidade condicional

Lembrete!

No que segue se considera um espaço de probabilidade $(\Omega, \mathcal{F}, P)$.

Seja $A \in \mathcal{F}$ e $C \in \mathcal{F}$ tal que $P(C) > 0$.

A probabilidade condicional de A dado C é definida por

$$P(A|C) = \frac{P(A \cap C)}{P(C)}.$$

Observação!

Seja $C \in \mathcal{F}$, $P(C) > 0$. *Se* $P_C : \mathcal{F} \longrightarrow [0,1]$ *é definida por* $P_C(A) := P(A|C)$, *para todo* $A \in \mathcal{F}$ *então* P_C *é uma medida de probabilidade em* $\mathcal{F}$.

Lembrete!

Podemos verificar que P_C é uma probabilidade em $\mathcal{F}$ notando que:

- Vale **A1**: $P_C(A) = P(A|C) \geq 0$, para todo $A \in \mathcal{F}$.
- Vale **A2**: $P_C(\Omega) = P(\Omega|C) = \dfrac{P(\Omega \cap C)}{P(C)} = \dfrac{P(C)}{P(C)} = 1.$
- Vale **A3**: se $A_n \in \mathcal{F}$, para $n \in \mathbb{N}$, e $A_i \cap A_j = \emptyset$ se $i \neq j$, então

$$P_C\left(\bigcup_{n=1}^{\infty} A_n\right) = \frac{P\left(\left\{\bigcup_{n=1}^{\infty} A_n\right\} \bigcap C\right)}{P(C)} = \frac{P\left(\bigcup_{n=1}^{\infty} \{A_n \cap C\}\right)}{P(C)}.$$

Logo

$$P_C\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \frac{P(A_n \cap C)}{P(C)} = \sum_{n=1}^{\infty} P_C(A_n).$$

Distribuição condicional de X dado A

Se X é uma variável aleatória em $(\Omega, \mathcal{F}, P)$ e $A \in \mathcal{F}$, $P(A) > 0$, então:

$$P(X \in B|A) = \frac{P(\{X \in B\} \cap A)}{P(A)},$$

com $B \in \mathcal{B}(\mathbb{R})$, é a distribuição condicional de X dado A.

Observação!

A função de distribuição condicional de X dado A define-se como:

$$F_X(x|A) := P(X \leq x|A) = \frac{P(\{X \leq x\} \cap A)}{P(A)},\ x \in \mathbb{R}.$$

Distribuição condicional: variáveis discretas

Sejam X e Y v.a. discretas com função de probabilidade conjunta $p(i,j)$. A função de probabilidades de X dado que $Y = j$, ou função de probabilidade condicional, define-se como

$$p_{X|Y}(i|j) := \frac{p(i,j)}{p_Y(j)},$$

para todo j tal que $p_Y(j) > 0$.

Observação!

De fato, note que $p_{X|Y}(i|j) := P(X = i|Y = j) = \dfrac{P(X = i, Y = j)}{P(Y = j)}$.

A função de distribuição condicional de X dado que $Y = j$ é dada por

$$F_{X|Y}(i|j) := P(X \leq i|Y = j) = \sum_{k \leq i} p_{X|Y}(k|j).$$

Exemplo 19.1

Considere o lançamento de dois tetraedros regulares, com lados etiquetados com os números 1 a 4, se:

$$X = \textit{número do } 1^\circ \textit{ tetraedro que está em contato com a superfície}$$

e

$$Y = \textit{maior das faces que estão em contato com a superfície,}$$

então a função de probabilidade conjunta de X e Y é dada por

$\mathbf{x}$	$(1,1)$	$(1,2)$	$(1,3)$	$(1,4)$	$(2,2)$	$(2,3)$	$(2,4)$	$(3,3)$	$(3,4)$	$(4,4)$
$p(\mathbf{x})$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{4}{16}$

Determine a função de probabilidade condicional de X dado $Y = 3$.

... continuação do Exemplo 19.1. Lembre que

$$p_{X\,|\,Y}(i|3) = \frac{p(i,3)}{P_Y(3)}$$

e, dado que $Y = 3$, X toma os valores de $i \in \{1, 2, 3\}$. Então, como

$$p_Y(3) = \sum_{i=1}^{4} p(i,3) = \frac{1}{16} + \frac{1}{16} + \frac{3}{16} = \frac{5}{16}$$

resulta ser

$$p_{X\,|\,Y}(1|3) = p_{X\,|\,Y}(2|3) = \frac{1/16}{5/16} = \frac{1}{5}, \quad p_{X\,|\,Y}(3|3) = \frac{3/16}{5/16} = \frac{3}{5}.$$

Exemplo 19.2

Se $X \sim Poisson(\lambda_1)$, $Y \sim Poisson(\lambda_2)$, *e são independentes, então:*

$$X|X+Y=n \sim B\left(n, \frac{\lambda_1}{\lambda_1+\lambda_2}\right).$$

Isto é, a distribuição condicional de X *dado que* $X+Y=n$, *é binomial de parâmetros* n *e* $\lambda_1/(\lambda_1+\lambda_2)$.

Para verificar, note que

- *Dado que* $X+Y=n$, X *toma valores em* $\{0,1,2,\ldots,n\}$.
- *Para cada* $i \in \{0,1,\ldots,n\}$ *temos que*

$$P(X=i|X+Y=n) = \frac{P(X=i)P(Y=n-i)}{P(X+Y=n)}.$$

- $X+Y \sim Poisson(\lambda_1+\lambda_2)$.

Exemplo 19.3

Suponha que n lançamentos independentes de uma moeda que tem probabilidade p de dar cara a cada lançamento são realizados. Se

$$X_i = \begin{cases} 1, & \textit{se o } i\textit{-ésimo lançamento resulta em cara,} \\ 0, & \textit{caso contrário.} \end{cases}$$

para $i \in \{1, 2, \ldots, n\}$ *e*

$$Y = \sum_{i=1}^{n} X_i,$$

então, dado que $Y = k$*, o vetor aleatório discreto* $(X_1, X_2, \ldots, X_n)$ *tem distribuição uniforme no conjunto*

$$\mathcal{S} = \left\{ (x_1, x_2, \ldots, x_n) : x_i \in \{0, 1\}, i \in \{1, 2, \ldots, n\}, \ e \ \sum_{i=1}^{n} x_i = k \right\}.$$

Note que $|\mathcal{S}| = \binom{n}{k}$.

Universidade Federal de Pernambuco

CCEN Centro de Ciências Exatas e da Natureza

... continuação do Exemplo 19.3. Seja $\mathbf{s} \in \mathcal{S}$ e note que

$$P(\mathbf{X} = \mathbf{s} | Y = k) = \frac{P(\mathbf{X} = \mathbf{s}, Y = k)}{P(Y = k)}$$

e como $\{\mathbf{X} = \mathbf{s}\} \subset \{Y = k\}$ então

$$P(\mathbf{X} = \mathbf{s} | Y = k) = \frac{P(\mathbf{X} = \mathbf{s})}{P(Y = k)} = \frac{p^k(1-p)^{n-k}}{\binom{n}{k} p^k (1-p)^{n-k}} = \frac{1}{\binom{n}{k}}.$$

Observação!

Dado que resultam k caras, temos a mesma probabilidade de obter qualquer uma das $\binom{n}{k}$ ordenações possíveis de k caras e $n-k$ coroas.

Distribuição condicional: variáveis continuas

Sejam X e Y duas v.a. continuas com função densidade conjunta $f(x,y)$. A função densidade condicional de X dado que $Y=y$, define-se como

$$f_{X|Y}(x|y) := \frac{f(x,y)}{f_Y(y)},$$

para todo y tal que $f_Y(y) > 0$.

Observação!

Sejam y e $\Delta y \approx 0$ tais que $P(y \leq Y \leq y + \Delta y) > 0$. Note que

$$\lim_{\Delta y \to 0} P(X \leq x | y \leq Y \leq y + \Delta y) = \int_{-\infty}^{x} \frac{f(x,y)}{f_Y(y)} dx \text{ e que } \frac{f(x,y)}{f_Y(y)} \text{ é densidade.}$$

Lembrete: Distribuição normal bi-variada

X e Y têm distribuição normal bi-variada se sua densidade conjunta é dada por

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\left\{\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2-\frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right\}}$$

para $-\infty < x, y < \infty$, onde $\mu_X, \mu_Y, \sigma_X, \sigma_Y$ e ρ são constantes tais que $\sigma_X > 0, \sigma_Y > 0$ e $-1 < \rho < 1$.

Então:

- $X|Y = y \sim \mathcal{N}\left(\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y), \sigma_X^2(1-\rho^2)\right)$.
- $Y|X = x \sim \mathcal{N}\left(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_x), \sigma_Y^2(1-\rho^2)\right)$.
- X e Y são independentes se, e somente se, $\rho = 0$. Verificar!

De fato, dado que $Y = y$, então

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = C\, e^{-\left\{\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]\right\}},$$

onde $C := C(\sigma_X^2, \sigma_Y^2, \rho, \mu_X, y)$ é uma constante. Logo,

$$\begin{aligned} f_{X|Y}(x|y) &= C_1\, e^{-\left\{\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right) - \rho\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]^2\right\}} \\ &= C_1\, e^{-\left\{\frac{1}{2\sigma_X^2(1-\rho^2)}\left[x-\mu_X - \frac{\rho\sigma_X}{\sigma_Y}(y-\mu_Y)\right]^2\right\}}, \end{aligned}$$

com $C_1 := C_1(\sigma_X^2, \sigma_Y^2, \rho, \mu_X, y)$ uma constante, que encontra-se de

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y)\, dx = 1.$$

Finalmente, $X|Y = y \sim \mathcal{N}\left(\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y), \sigma_X^2(1-\rho^2)\right)$.

Exemplo 19.4

Um pai e seu filho (adulto) de certa população são selecionados ao acaso. Considere:

$$X = \text{ altura do pai em metros, e } Y = \text{ altura do filho em metros.}$$

Suponha que X *e* Y *têm uma distribuição normal bi-variada com*

$$\mu_X = \mu_Y = 1,75; \;\; \sigma_X = \sigma_Y = 0,05; \;\; \rho = 0,5.$$

Então, dado que $X = 1,88$, Y *tem distribuição*

$$\mathcal{N}\left(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_x), \sigma_Y^2(1-\rho^2)\right),$$

isto é, $Y|X = 1,88 \sim \mathcal{N}(1,815; 0,00019)$.

Distribuição condicional: continua dado discreta

Seja X uma variável aleatória continua com densidade $f(x)$ e seja Y discreta. A densidade condicional de X dado que $Y = i$ define-se como:

$$f_{X|Y}(x|i) = \frac{P(Y = i|X = x)}{P(Y = i)} f(x).$$

$$\begin{aligned}
f_{X|Y}(x|i) &= \lim_{\Delta x \to 0} \frac{P(x \leq X \leq x + \Delta x | Y = i)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \left\{ \frac{P(Y = i|x \leq X \leq x + \Delta x)}{P(Y = i)} \frac{P(x \leq X \leq x + \Delta x)}{\Delta x} \right\} \\
&= \frac{P(Y = i|X = x)}{P(Y = i)} f(x).
\end{aligned}$$

Lembrete: distribuição Beta

$X \sim Beta(\alpha, \beta)$, $\alpha, \beta > 0$, se sua deinsidade é dada por

$$f(x) = \begin{cases} \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, & 0 < x < 1 \\ 0, & \text{caso contrário.} \end{cases}$$

onde

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx.$$

Observação!

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Lembrete: distribuição Beta

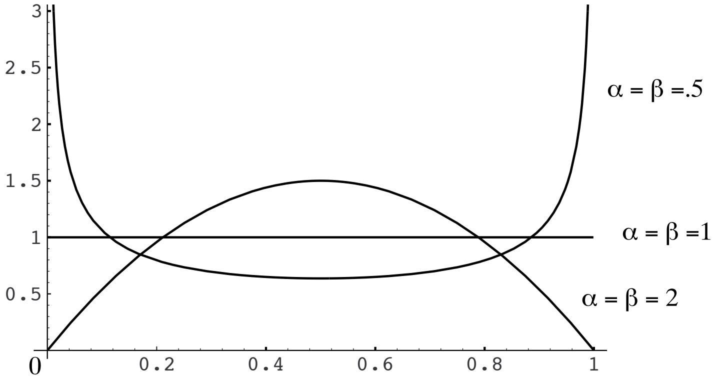

Figura: Densidade Beta para $\alpha = \beta \in \{0,5; 1; 2\}$.

Exemplo 19.5

Suponha que n lançamentos sucessivos da mesma moeda são realizados, mas que não sabemos qual é a probabilidade da moeda usada dar cara. Suponha que tal probabilidade é uma v.a. $U(0,1)$. Então, dado que dos n lançamentos i resultam em cara, a probabilidade da moeda dar cara tem distribuição $Beta(i+1, n-i+1)$.

De fato, se X a probabilidade da moeda escolhida dar cara e Y é o número de caras nos n lançamentos, então:

$$X \sim U(0,1) \quad e \quad Y|X = p \sim B(n,p).$$

Logo, $f_{X\,|\,Y}(p|i) = \dfrac{P(Y=i|X=p)f_X(p)}{P(Y=i)} = \dfrac{\binom{n}{i}p^i(1-p)^{n-i}}{P(Y=i)}$. *Isto é,*

$$f_{X\,|\,Y}(p|i) = Cp^i(1-p)^{n-i}$$

onde C é uma constante que não depende de p.

Exemplo 19.6

Uma droga é efetiva a um tratamento com probabilidade p, independentemente do resultado de outros tratamentos. Suponha que não conhecemos o valor de p mas que experiências anteriores nos permitem considerar p como uma v.a. com distribuição Beta(α, β). Suponha que damos a droga para n indivíduos, e que ela é efetiva em i dos n casos. Vamos encontrar a densidade condicional de X, dado que $Y = i$, onde Y é o número de casos nos quais a droga é efetiva.

Neste caso, notemos que

$$Y|X = p \sim B(n, p)$$

e que

$$f_{X\,|\,Y}(p|i) = \frac{P(Y = i|X = p)f_X(p)}{P(Y = i)}$$

Logo,

$$P(Y = i) = \int_0^1 P(Y = i|X = p)\, f_X(p)\, dp.$$

... continuação do Exemplo 19.6. Portanto,

$$P(Y = i) = \binom{n}{i} \frac{B(\alpha + i, \beta + n - i)}{B(\alpha, \beta)},$$

e

$$f_{X\,|\,Y}(p|i) = \frac{1}{B(\alpha + i, \beta + n - i)} p^{\alpha+i-1}(1-p)^{\beta+n-i-1},$$

isto é

$$X\,|\,Y = i \sim Beta(\alpha + i, \beta + n - i).$$

Observação!

Se $X \sim U(0,1)$ *então* $X \sim Beta(1,1)$.

Em geral!

Se X e Y são variáveis aleatórias definidas em $(\Omega, \mathcal{F}, P)$. A função

$$P(X \in B | Y = y), \;\; B \text{ boreliano e } y \in \mathbb{R},$$

chama-se distribuição condicional para X dada Y se:

i. $\forall y \in \mathbb{R}$ fixo, $P(X \in B | Y = y)$ é uma probabilidade em $\mathcal{B}$,

ii. $\forall B \in \mathcal{B}$ fixo, $P(X \in B | Y = y)$ é função mensurável de y e:

$$\int_{-\infty}^{y} P(X \leq x | Y = t) dF_Y(t) = P(X \leq x, Y \leq y), \;\; \forall (x, y) \in \mathbb{R}^2.$$

Definição!

Para $B \in \mathcal{B}$ e $y \in \mathbb{R}$ a probabilidade condicional de que $X \in B$ dado que $Y = y$, define-se por $P(X \in B | Y = y) = \lim_{\Delta y \to 0} P(X \in B | y \leq Y \leq y + \Delta y)$.

Bom estudo!