Lista 1. Distribuição Condicional Caso Discreto. (sexta 11/09/2020)

Exercício 1. A distribuição conjunta de duas variáveis discretas X, Y é dada pela seguinte esquema: valores de X são $0, 1, 2$ e valores de Y são 0 e 1. Sabemos que X dado que $Y = 0$ tem a distribuição binomial $B(2, p)$ (escrevemos isso assim: $X \mid Y = 0 \sim B(2, p)$) e $X \mid Y = 1 \sim B(2, 1-p)$, em que p é algum número em intervalo $(0, 1)$.

1. Supondo que a distribuição marginal do Y é uniforme: $\mathbb{P}(Y = 0) = \mathbb{P}(Y = 1) = 0.5$, preencher a tabela da distribuição conjunta

$Y \setminus X$	0	1	2
0			
1			

2. Achar a distribuição de $Z = 2X + Y$.
3. Achar as distribuições marginais de X e Y. As variáveis X e Y são independentes?
4. Achar a distribuição de $\mathbb{E}(Y \mid X)$.
5. Achar a distribuição de $\mathbb{V}ar(Y \mid X)$.

Solução:

1. Supondo que a distribuição marginal do Y é uniforme: $\mathbb{P}(Y = 0) = \mathbb{P}(Y = 1) = 0.5$, preencher a tabela da distribuição conjunta.

Usando a fórmula de condicional $\mathbb{P}(X = i, Y = j) = \mathbb{P}(X = i \mid Y = j)\mathbb{P}(Y = j)$ e utilizando anunciado $X \mid Y = 0 \sim B(2, p)$) e $X \mid Y = 1 \sim B(2, 1-p)$ obtemos:

$$\mathbb{P}(X = 0, Y = 0) = \mathbb{P}(X = 0 \mid Y = 0)\mathbb{P}(Y = 0) = \binom{2}{0} p^0 (1-p)^2 \frac{1}{2} = (1-p)^2 \frac{1}{2}$$

$$\mathbb{P}(X = 1, Y = 0) = \mathbb{P}(X = 1 \mid Y = 0)\mathbb{P}(Y = 0) = \binom{2}{1} p^1 (1-p)^1 \frac{1}{2} = 2p(1-p) \frac{1}{2}$$

$$\mathbb{P}(X = 2, Y = 0) = \mathbb{P}(X = 2 \mid Y = 0)\mathbb{P}(Y = 0) = \binom{2}{2} p^2 (1-p)^0 \frac{1}{2} = p^2 \frac{1}{2}$$

$$\mathbb{P}(X = 0, Y = 1) = \mathbb{P}(X = 0 \mid Y = 1)\mathbb{P}(Y = 1) = \binom{2}{0} (1-p)^0 p^2 \frac{1}{2} = p^2 \frac{1}{2}$$

$$\mathbb{P}(X = 1, Y = 1) = \mathbb{P}(X = 1 \mid Y = 1)\mathbb{P}(Y = 1) = \binom{2}{1} (1-p)^1 p^1 \frac{1}{2} = 2(1-p)p \frac{1}{2}$$

$$\mathbb{P}(X = 2, Y = 1) = \mathbb{P}(X = 2 \mid Y = 1)\mathbb{P}(Y = 1) = \binom{2}{2} (1-p)^2 p^0 \frac{1}{2} = (1-p)^2 \frac{1}{2}$$

Logo,

$Y \setminus X$	0	1	2
0	$(1-p)^2 \frac{1}{2}$	$p(1-p)$	$p^2 \frac{1}{2}$
1	$p^2 \frac{1}{2}$	$(1-p)p$	$(1-p)^2 \frac{1}{2}$

2. Achar a distribuição de $Z = 2X + Y$.

Adicionando os valores de Z em selas da tabela (em vermelho)

$Y \setminus X$	0	1	2
0	$\frac{(1-p)^2}{2} \setminus 0$	$p(1-p) \setminus 2$	$\frac{p^2}{2} \setminus 4$
1	$\frac{p^2}{2} \setminus 1$	$(1-p)p \setminus 3$	$\frac{(1-p)^2}{2} \setminus 5$

construimos a tabela da distribuição de Z:

Z	0	1	2	3	4	5
$\mathbb{P}$	$\frac{(1-p)^2}{2}$	$\frac{p^2}{2}$	$p(1-p)$	$(1-p)p$	$\frac{p^2}{2}$	$\frac{(1-p)^2}{2}$

3. Achar as distribuições marginais de X e Y. As variáveis X e Y são independentes?

A distribuição de Y já é dada em anunciado: $Y \sim B(\frac{1}{2})$ (Bernoulli com probabilidade de sucesso $1/2$). Pela tabela acharemos a distribuição de X:

$Y \setminus X$	0	1	2
0	$\frac{(1-p)^2}{2}$	$p(1-p)$	$\frac{p^2}{2}$
1	$\frac{p^2}{2}$	$(1-p)p$	$\frac{(1-p)^2}{2}$
	$p^2-p+\frac{1}{2}$	$2p(1-p)$	$p^2-p+\frac{1}{2}$

logo

X	0	1	2
$\mathbb{P}$	$p^2-p+\frac{1}{2}$	$2p(1-p)$	$p^2-p+\frac{1}{2}$

As variáveis não são independentes: isso vimos pelo anunciado – a distribuição condicional de X muda dependendo do estado de Y. O que também pode ser verificado direto, por exemplo:

$$\frac{(1-p)^2}{2} = \mathbb{P}(X=0, Y=0) \neq \mathbb{P}(X=0)\mathbb{P}(Y=0) = \left(p^2 - p + \frac{1}{2}\right)\frac{1}{2}.$$

4. Achar a distribuição de $\mathbb{E}(Y \mid X)$.

Sabemos que os valores de Y são 0 e 1. Então a distribuição condicional de Y dado $X = a$ é Bernoulli, com uma probabilidade de sucesso $\mathbb{P}(Y = 1 \mid X = a)$. Assim, a esperança condicional é igual á $\mathbb{P}(Y = 1 \mid X = a)$. Logo, os valores de $\mathbb{E}(Y \mid X)$ são:

$$\mathbb{P}(Y=1 \mid X=0) = \frac{\frac{p^2}{2}}{p^2-p+\frac{1}{2}} \text{ com probabilidade } \mathbb{P}(X=0) = p^2 - p + \frac{1}{2}$$
$$\mathbb{P}(Y=1 \mid X=1) = \frac{p(1-p)}{2p(1-p)} = \frac{1}{2} \text{ com probabilidade } \mathbb{P}(X=1) = 2p(1-p)$$
$$\mathbb{P}(Y=1 \mid X=2) = \frac{\frac{(1-p)^2}{2}}{p^2-p+\frac{1}{2}} \text{ com probabilidade } \mathbb{P}(X=2) = p^2 - p + \frac{1}{2}$$

logo

$\mathbb{E}(Y \mid X)$	$\frac{\frac{p^2}{2}}{p^2-p+\frac{1}{2}}$	$\frac{1}{2}$	$\frac{\frac{(1-p)^2}{2}}{p^2-p+\frac{1}{2}}$
$\mathbb{P}$	$p^2-p+\frac{1}{2}$	$2p(1-p)$	$p^2-p+\frac{1}{2}$

Observe a formula iterativa da esperança, $\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y \mid X))$:

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

5. Achar a distribuição de $\mathbb{V}ar(Y \mid X)$.

Com o mesmo raciocínio como do item anterior, logo escrevemos os valores com respectivas probabilidades:

$$\mathbb{V}ar(Y \mid X=0) = \mathbb{P}(Y=1 \mid X=0)\mathbb{P}(Y=0 \mid X=0) = \frac{\frac{p^2}{2}}{p^2-p+\frac{1}{2}} \cdot \frac{\frac{(1-p)^2}{2}}{p^2-p+\frac{1}{2}} \text{ com probabilidade } \mathbb{P}(X=0) = p^2-p+\frac{1}{2}$$
$$\mathbb{V}ar(Y \mid X=1) = \mathbb{P}(Y=1 \mid X=1)\mathbb{P}(Y=0 \mid X=1) = p(1-p)\cdot(1-p)p = p^2(1-p)^2 \text{ com probabilidade } \mathbb{P}(X=1) = 2p(1-p)$$
$$\mathbb{V}ar(Y \mid X=2) = \mathbb{P}(Y=1 \mid X=2)\mathbb{P}(Y=0 \mid X=2) = \frac{\frac{(1-p)^2}{2}}{p^2-p+\frac{1}{2}} \cdot \frac{\frac{p^2}{2}}{p^2-p+\frac{1}{2}} \text{ com probabilidade } \mathbb{P}(X=2) = p^2-p+\frac{1}{2}$$

ou em forma de tabela

$\mathbb{V}ar(Y \mid X)$	$\frac{\frac{(1-p)^2p^2}{4}}{\left(p^2-p+\frac{1}{2}\right)^2}$	$p^2(1-p)^2$
$\mathbb{P}$	$2\left(p^2-p+\frac{1}{2}\right)$	$2p(1-p)$

□

Exercício 2. X_1, X_2 têm distribuição geométrica com parâmetro p e $1-p$ respectivamente, $X_1 \sim Geom(p), X_2 \sim Geom(1-p)$. X_1, X_2 são variáveis aleatórias independentes. Achar distribuição $\mathbb{P}(X_1 = k \mid X_1 + X_2 = n)$ e esperança $\mathbb{E}(X_1 = k \mid X_1 + X_2 = n)$.

Solução: Escolhemos, como a distribuição geométrica seguinte distribuição $\mathbb{P}(Y = n) = (1-p)^{n-1}p$ para $n = 1, 2, 3, \ldots$ (número de tentativas até o primeiro sucesso). Logo

$$\begin{aligned}
\mathbb{P}(X_1 = k \mid X_1 + X_2 = n) &= \frac{\mathbb{P}(X_1 = k)\mathbb{P}(X_2 = n-k)}{\mathbb{P}(X_1 + X_2 = n)} = \frac{(1-p)^{k-1}p \cdot p^{n-k-1}(1-p)}{\sum_{k=1}^{n-1}(1-p)^{k-1}p \cdot p^{n-k-1}(1-p)} \\
&= \frac{(1-p)^k p^{n-k}}{\sum_{r=1}^{n-1}(1-p)^r p^{n-r}} = \frac{(1-p)^k p^{n-k}}{p^n \sum_{r=1}^{n-1}(\frac{1-p}{p})^r} = \frac{(1-p)^k p^{n-k}}{\frac{p(1-p)}{2p-1}\left(p^{n-1} - (1-p)^{n-1}\right)} = \frac{(2p-1)(1-p)^{k-1}p^{n-k-1}}{p^{n-1} - (1-p)^{n-1}}
\end{aligned}$$

onde $k = 1, 2, \ldots, n-1$ e $n \geq 2$ e usamos a fórmula

$$\sum_{k=1}^{n-1} x^k = x\frac{1 - x^{n-1}}{1-x}.$$

Logo, pela definição da esperança, temos

$$\begin{aligned}
\mathbb{E}(X_1 \mid X_1 + X_2 = n) &= \sum_{k=1}^{n-1} k\frac{(2p-1)(1-p)^{k-1}p^{n-k-1}}{p^{n-1} - (1-p)^{n-1}} = \frac{(2p-1)}{p^{n-1} - (1-p)^{n-1}} \sum_{k=1}^{n-1} k(1-p)^{k-1}p^{n-k-1} \\
&= \frac{\frac{2p-1}{p}}{1 - \left(\frac{1-p}{p}\right)^{n-1}} \sum_{k=1}^{n-1} k\left(\frac{1-p}{p}\right)^{k-1} = \frac{1 - \frac{1-p}{p}}{1 - \left(\frac{1-p}{p}\right)^{n-1}} \frac{d}{dx}\left(\sum_{k=1}^{n-1} x^k\right)\Bigg|_{x=\frac{1-p}{p}}
\end{aligned}$$

agora

$$\frac{d}{dx}\left(\sum_{k=1}^{n-1} x^k\right) = \frac{d}{dx}\left(\frac{x - x^n}{1-x}\right) = \frac{1 - nx^{n-1} + (n-1)x^n}{(1-x)^2},$$

obtemos

$$\mathbb{E}(X_1 \mid X_1 + X_2 = n) = \frac{1-x}{1-x^{n-1}} \cdot \frac{1 - nx^{n-1} + (n-1)x^n}{(1-x)^2}\Bigg|_{x=\frac{1-p}{p}} = \frac{1 - nx^{n-1} + (n-1)x^n}{(1-x)(1-x^{n-1})}\Bigg|_{x=\frac{1-p}{p}}$$

□

Exercício 3. X_1, X_2 têm distribuição uniforme. X_1 é uniforme em conjunto $\{0, 1, \ldots, n\}$, quando X_2 possua valores $\{0, 1, \ldots, m\}$. Supomos, por exemplo, que $n < m$. Usando a fórmula condicional ($\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X \mid Y))$) achar a probabilidade $\mathbb{P}(X_1 < X_2)$. (Dica: usar seguinte representação para a probabilidade $\mathbb{P}(X_1 < X_2) = \mathbb{E}(\mathbb{I}_{(X_1 < X_2)})$, em que $\mathbb{I}_{(X_1 < X_2)}$ indicador do evento $(X_1 < X_2)$: $\mathbb{I}_{(X_1 < X_2)} = 1$, se $X_1 < X_2$, e $\mathbb{I}_{(X_1 < X_2)} = 0$ caso contrário)

Solução:

$$\begin{aligned}
\mathbb{P}(X_1 < X_2) &= \sum_{k=0}^{m} \mathbb{P}(X_1 < X_2 \mid X_2 = k)\mathbb{P}(X_2 = k) \\
&= \sum_{k=1}^{n} \mathbb{P}(X_1 < X_2 \mid X_2 = k)\mathbb{P}(X_2 = k) + \sum_{k=n+1}^{m} \mathbb{P}(X_1 < X_2 \mid X_2 = k)\mathbb{P}(X_2 = k) \\
&= \frac{1}{m+1}\sum_{k=1}^{n} \mathbb{P}(X_1 < X_2 \mid X_2 = k) + \sum_{k=n+1}^{m} 1 \cdot \frac{1}{m+1} = \frac{1}{m+1}\left(\frac{1}{n+1} + \frac{2}{n+1} + \cdots + \frac{n-1}{n+1}\right) + \frac{m-n}{m+1} \\
&= \frac{1}{(m+1)(n+1)}\frac{(n-1)n}{2} + \frac{m-n}{m+1} = \frac{2m(n-1) - n(n-3)}{2(m+1)(n+1)}
\end{aligned}$$

onde usamos a fórmula

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

□

Exercício 4. A distribuição conjunta de $X, Y \in \{0, 1, \dots, n\}$ é dada pela fórmula

$$p(x,y) = \begin{cases} 0, & \text{se } x < y \\ a, & \text{se } x \geq y \end{cases}$$

1. achar a
2. X, Y são variáveis independentes?
3. qual é a distribuição de $\mathbb{E}(X \mid Y)$?

Solução: Aqui melhor escrever as probabilidades como a matrix $(p(x,y))_{x,y=0,1,\dots,n}$.

1. achar a

$$\sum_{x=0}^{n}\sum_{y=0}^{n} p(x,y) = \sum_{x=0}^{n}\sum_{y=0}^{n} a\mathbb{1}_{(x\geq y)} = \sum_{x=0}^{n} a(x+1) = a\frac{(n+1)(n+2)}{2} = 1 \Longrightarrow a = \frac{2}{(n+1)(n+2)}.$$

2. X, Y são variáveis independentes?

Não são independentes: por exemplo $0 = \mathbb{P}(X = 0, Y = 1) \neq \mathbb{P}(X = 0)\mathbb{P}(Y = 1) = a \cdot na \neq 0$

3. qual é a distribuição de $\mathbb{E}(X \mid Y)$?

Observe, que $X \mid Y = y$ tem distribuição uniforme em $\{y, y+1, \dots, n\}$, e o valor médio $\mathbb{E}(X \mid Y = y) = \frac{y+n}{2}$. Logo, a v.a. $\mathbb{E}(X \mid Y)$ tem valores

$$\mathbb{E}(X \mid Y = y) = \frac{y+n}{2} \text{ com probabilidade } \mathbb{P}(Y = y) = \frac{(n+1-y)a}{(n+1)} = \frac{n+1-y}{n+1},$$

onde $y = 0, 1, \dots, n$.

□

Referências

[1] S.M.Ross *Introduction to probability models.* Ninth Edition, Elsevier, 2007.