Universidade Federal de Lavras

Departamento de Ciências Exatas

Prof. Daniel Furtado Ferreira

6ª Lista de Exercícios Práticos   Distribuição de Probabilidades Discretas

1) Considere ninhadas de $n = 3$ filhotes de coelhos. Construir o espaço amostral considerando os nascimentos de fêmeas e machos, utilizando um diagrama de árvore e considerar os eventos *nascer macho* e *nascer fêmea* como equiprováveis.

   a) Sendo X a ocorrência de fêmeas, construa a distribuição de probabilidade de X;

   b) Calcule as probabilidades dos seguintes eventos por meio da distribuição de probabilidade construída:

      i) nascimento de exatamente duas fêmeas.

      ii) nascimento de pelo menos um macho.

      iii) nascimento de pelo menos duas fêmeas.

      iv) nascimento de no máximo uma fêmea.

   c) Suponha que você faça uma amostragem de 500 ninhadas de 3 filhotes. Em quantos, em média, você espera encontrar com exatamente 1 fêmea?

2) Considere nascimentos de $n = 4$ filhotes de coelhos de um determinada raça. Nesta raça há um distúrbio genético e a probabilidade de nascer fêmea é 5/8. Sendo X a ocorrência de fêmeas e utilizando a distribuição binomial obter:

   a) a distribuição de probabilidade de X, ou seja, os valores e as probabilidades associadas aos respectivos valores x;

   b) a média e variância da variável aleatória X, com distribuição binomial;

   c) o número esperado (médio) de ninhadas em uma amostra de 1.000 ninhadas de tamanho $n = 4$ para cada valor da variável aleatória X.

3) Numa lâmina verificou-se que existiam em média 4 bactérias/cm^2. A lâmina foi subdividida em 600 quadrados de 1 cm^2. Qual é o modelo probabilístico adequado para modelar a ocorrência de bactérias por cm^2, supondo que a distribuição espacial segue um padrão aleatório? Em quantos dos 600 quadrados, em média, você espera encontrar no máximo 1 bactéria? Qual é a probabilidade de se encontrar mais de 2 bactérias por centímetro quadrado? Qual é a probabilidade de não encontrar bactérias em um quadrado tomado aleatoriamente destes 600 quadrados?

4) Um pesquisador da área de zootecnia conseguiu uma série de dados dos últimos 120 anos com o registro do número de uma doença rara em equinos da localidade em que trabalhava. Os dados obtidos foram:

Número de doenças (x)	0	1	2	3	4	5
Número de anos (F_i)	55	40	17	5	2	1

   a) Estime o número médio de doenças /ano;

   b) Calcule para cada valor da variável aleatória X, as probabilidades associadas. Suponha que X possua distribuição de Poisson e que a média amostral é o estimador do parâmetro λ da distribuição Poisson;

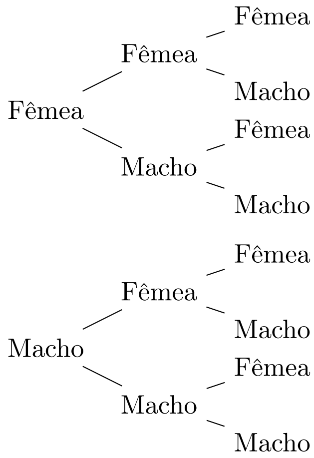
   c) Calcule a frequência esperada (em anos) para cada valor da variável aleatória X;

   d) Compare os resultados esperados com os observados. Com base nesta comparação, você pode afirmar que a distribuição de Poisson é adequada para explicar a ocorrência desta doença na região de estudo? Justifique.

5) Uma plantação de tomate possui em média 2 galhas de *M. incógnita* por planta. Qual é a probabilidade de que uma planta amostrada desta população não possua galha? Suponha que o modelo Poisson é apropriado para modelar a ocorrência de galhas de nematóide. Qual é a probabilidade de que em uma amostra de tamanho $n = 5$ plantas, as 5 não apresentem galhas?

Resolução

1) O dendrograma (diagrama de árvore) é:



O espaço amostral é obtido considerando todos os ramos desse diagrama de árvore e resulta em:

$$\Omega = \{(FFF), (FFM), (FMF), (FMM), (MFF), (MFM), (MMF), (MMM)\}.$$

Cada ponto deste é equiprovável, pois as probabilidades de nascimento de fêmea e macho são iguais e cada elemento de Ω tem probabilidade $(1/2)^3 = 1/8$.

a) A distribuição de probabilidade da variável aleatória X, definida como sendo o número de fêmeas, é:

x	0	1	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Por exemplo, a probabilidade $P(X = 1)$ é obtida pelo número total de pontos de Ω em que $X = 1$, que no caso é igual a 3, em relação ao número total de pontos, que é igual a 8.

b) Os resultados para os eventos solicitados são:

i) $P(X = 2) = \dfrac{3}{8} = 0{,}3750 = 37{,}50\%$;

ii) Se Y representa o número de machos, o evento $Y \geq 1$ equivale a $X \leq 2$, pois se houver 1 macho na ninhada, isto implica em duas fêmeas; se houver 2 machos, implica em 1 fêmea; e se houver 3 machos, implica em 0 fêmea. Assim, a probabilidade do evento é:
$P(Y \geq 1) = P(X \leq 2) = P(X = 2) + P(X = 1) + P(X = 0) = \dfrac{3}{8} + \dfrac{3}{8} + \dfrac{1}{8} = \dfrac{7}{8} = 0{,}8750 = 87{,}50\%$;

iii) $P(X \geq 2) = P(X = 2) + P(X = 3) = \dfrac{3}{8} + \dfrac{1}{8} = \dfrac{1}{2} = 0{,}50 = 50\%$;

iv) $P(X \leq 1) = P(X = 1) + P(X = 0) = \dfrac{3}{8} + \dfrac{1}{8} = \dfrac{1}{2} = 0{,}50 = 50\%$.

c) O número esperado NE de ninhadas de 3 filhotes com exatamente 1 fêmea é dado pelo produto da probabilidade do evento $P(X = 1)$ pelo número total de ninhadas, ou seja, é $NE = (3/8) \times 500 \approx 188$ ninhadas. Assim, das 500 esperamos que 188 tenham exatamente 1 fêmea.

2) Neste caso assumimos que X, definida como o número de fêmeas, possui distribuição binomial com parâmetros $n = 4$ e $p = 5/8$. A probabilidade de sucesso não é $1/2$, de acordo com as leis de Mendel, pois há um distúrbio genético na raça. Usualmente usamos a seguinte notação para dizer a mesma coisa que acabamos de explicar: $X \sim$ Binomial$(n = 4, p = 5/8)$.

a) Para obtermos a distribuição de probabilidade de X, podemos utilizar o modelo binomial dado por:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

para $x = 0, 1, \ldots, 4$.

Por exemplo, para a probabilidade $P(X=0)$ temos:

$$P(X=x) = \binom{4}{0}\left(\frac{5}{8}\right)^0\left(1-\frac{5}{8}\right)^{4-0} = 1\times 1\times\left(\frac{3}{8}\right)^4 = 0{,}0198.$$

Assim, temos a seguinte distribuição de probabilidades:

x	0	1	2	3	4
$P(X=x)$	0,0198	0,1318	0,3296	0,3662	0,1526

b) A média é: $\mu_X = np = 4\times 5/8 = 2{,}5$; e a variância é: $\sigma^2_X = np(1-p) = 4\times(5/8)\times(3/8) = 0{,}9375$.

c) O número esperado NE de ninhadas é dado pelo produto da probabilidade do evento de interesse $P(X=x)$ pelo número total de ninhadas. Por exemplo, para $X=0$, temos $P(X=0)\times 1.000 = 0{,}0198\times 1.000 = 19{,}8 \approx 20$ ninhadas com nenhuma fêmea. Os demais valores (arredondados) estão apresentados na tabela a seguir:

x	0	1	2	3	4
$P(X=x)$	0,0198	0,1318	0,3296	0,3662	0,1526
NE	20	132	330	366	153

3) Resposta para cada uma das questões:

a) O modelo probabilístico adequado é o modelo Poisson, principalmente se pudermos supor que a distribuição das bactérias pela lâmina seja aleatória. Assim, $X \sim \text{Poisson}(\lambda = 4)$.

b) NE= $P(X\leq 1)\times 600$. Para obtermos a probabilidade do evento de interesse devemos usar o modelo Poisson dado por:

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!},$$

para $x = 0, 1, 2, \ldots$ e $\lambda = 4$.
Assim, $P(X\leq 1) = P(X=1) + P(X=0)$, sendo

$$P(X=0) = \frac{e^{-4}\times 4^0}{0!} = 0{,}0183 \qquad \text{e} \qquad P(X=1) = \frac{e^{-4}\times 4^1}{1!} = 0{,}0733.$$

Portanto,

$$P(X\leq 1) = P(X=0) + P(X=1) = 0{,}0183 + 0{,}0733 = 0{,}0916 = 9{,}16\%$$

e o número esperado de quadrados com no máximo 1 bactéria é NE= $P(X\leq 1)\times 600 = 0{,}0916\times 600 \approx 55$.

c) $P(X>2) = P(X=3) + P(X=4) + \ldots$. Como esta soma possui um número infinito de termos e sabendo que a soma de todos as probabilidades é igual a 1, então

$$\begin{aligned} P(X>2) =& 1 - [P(X=0) + P(X=1) + P(X=2)] = 1 - \left(0{,}0183 + 0{,}0733 + \frac{e^{-4}\times 4^2}{2!}\right) \\ =& 1 - (0{,}0183 + 0{,}0733 + 0{,}1465) = 1 - 0{,}2381 = 0{,}7619 = 76{,}19\%. \end{aligned}$$

d) $P(X=0) = 1{,}83\%$.

4) A ocorrência de doenças por ano pode ser modelada, se for aleatória, pelo modelo Poisson. Assim, $X \sim \text{Poisson}(\hat{\lambda = \bar{X}}$.

a) A média é

$$\bar{X} = \frac{0\times 55 + 1\times 44 + \cdots + 5\times 1}{120} = \frac{102}{120} = 0{,}85 \text{ doenças/ano.}$$

b) O parâmetro λ pode ser estimado por $\hat{\lambda} = \bar{X} = 0{,}85$. Assim, podemos estimar a distribuição de probabilidade utilizando o modelo Poisson por:

$$P(X = x) = \frac{e^{-0,85} \times 0,85^x}{x!}.$$

Para $X = 0$ e $X = 1$, para fins de ilustração, temos:

$$P(X = 0) = \frac{e^{-0,85} \times 0,85^0}{0!} = 0,4274 \qquad P(X = 1) = \frac{e^{-0,85} \times 0,85^1}{1!} = 0,3633.$$

Estas e as demais probabilidades são apresentadas na tabela seguinte:

x	0	1	2	3	4	5	6 ou mais
$P(X = x)$	0,4274	0,3633	0,1544	0,0437	0,0093	0,0016	0,0003

c) A frequência esperada é dada pelo produtos das probabilidades por 120. Logo,

x	0	1	2	3	4	5	6 ou mais
$P(X = x)$	0,4274	0,3633	0,1544	0,0437	0,0093	0,0016	0,0003
FE	51,29	43,60	18,53	5,24	1,12	0,19	0,04

d) Como as frequências observadas e esperadas estão relativamente próximas, podemos considerar que o modelo Poisson é adequado para modelar a ocorrência da doença na região estudada.

5) Na primeira parte do exercício, devemos considerar uma variável $X \sim$ Poisson($\lambda = 2$) referente a distribuição do número de galhas de nematóides por planta. A probabilidade desejada é $P(X = 0)$, cuja resposta é:

$$P(X = 0) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-2} \times 2^0}{0!} = 0,1353 = 13,53\%.$$

Assim, a probabilidade de uma planta sorteada ao acaso não possuir galhas é de 13,53%.

A segunda parte da questão refere-se a probabilidade de que em uma amostra de tamanho $n = 5$ desta população, todas as 5 não apresentem galhas. Se definirmos a variável aleatória (V.A.) Y como sendo o número de plantas sem galhas de nematóides, veremos que esta variável possui distribuição binomial com probabilidade de sucesso, a planta não apresentar galhas, dada por $p = P(X = 0)$. A probabilidade de sucesso é dada pela probabilidade obtida na primeira parte do exercício. Assim, $Y \sim$ binomial($n = 5, p = 0,1353$) e a probabilidade desejada é calculada por:

$$\begin{aligned} P(Y = 5) =& \binom{n}{y} p^y (1-p)^{n-y} = \binom{5}{5} \times 0,1353^5 \times (1 - 0,1353)^{5-5} = 0,1353^5 = 0,000045 \\ =& 0,0045\%. \end{aligned}$$

Logo, a probabilidade de amostrarmos 5 plantas e todas elas não apresentarem galhas é de apenas 0,0045%.