TPC 6: Parametrização dos Modelos - Turma 1

André Filipe Gomes Silvestre Nº104532 CDB1

Exercício (Turma 1)

1.

Suponha a experiência aleatória que consiste em selecionar aleatoriamente uma foto de um banco de fotos genéricas e verificar se inclui ($x = 1$) ou não ($x = 0$) gatinhos. Ou seja, considere a seguinte variável aleatória X − foto escolhida ao acaso tem gatinhos (1:sim, 0: não), assumida como tendo distribuição de *Bernoulli* de parâmetro p.

Pretende considerar amostras de dimensão 10 dessa população, $(X_1, X_2, \ldots, X_{10})$.

a. Construa a função de probabilidade conjunta de $(X_1, X_2, \ldots, X_{10})$.

b. Se $p = 0.1$, qual a probabilidade de observar a amostra $(1, 0, 1, 0, 0, 0, 0, 0, 0, 0)$? E se $p = 0.2$?

c. Como pode interpretar (no contexto apresentado) a estatística $T_1 = \sum_{i=1}^{10} X_i$? Qual o seu valor esperado?

d. Como pode interpretar (no contexto apresentado) a estatística $T_2 = \frac{\sum_{i=1}^{10} X_i}{10}$? Qual o seu valor esperado?

---

1. Resolução

a)

> **X -** foto escolhida ao acaso tem gatinhos (1 :sim, 0 :não), assumida como tendo distribuição de *Bernoulli* de parâmetro p

Dado que a função da distribuição de *Bernoulli* é dada por $X \sim \text{Bern}(p)$, então a função de probabilidade de X

$$f(x) = p^x(1-p)^{1-x}, \quad x = 0, 1$$

Como assumimos que $(X_1, X_2, ..., X_10)$ é amostra aleatória desta população, então a função conjunta será dada pelo produto das marginais (porque os X_i são independentes e identicamente distribuídos, com distribuição igual à população).

Logo, a *função de probabilidade conjunta* de $(X_1, X_2, \ldots, X_{10})$ é dada por:

$$f\left(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\right) = \prod_{i=1}^{10} p^{x_i}(1-p)^{(1-x_i)} = p\sum_{i=1}^{10} x_i \times (1-p)^{10-\sum_{i=1}^{10} x_i}$$

$$= p^{x_1+x_2+x_3+x_4+x_5+x_6+x_7+x_8+x_9+x_{10}}(1-p)^{10-(x_1+x_2+x_3+x_4+x_5+x_6+x_7+x_8+x_9+x_{10})}$$

b)

Seja a amostra a obter $(1, 0, 1, 0, 0, 0, 0, 0, 0, 0)$

Pretende-se

$$f(1, 0, 1, 0, 0, 0, 0, 0, 0, 0 \mid p = 0.1) = 0.1^2(1-0.1)^{10-2} = 0.0043$$

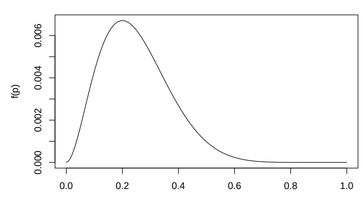
e

$$f(1, 0, 1, 0, 0, 0, 0, 0, 0, 0 \mid p = 0.2) = 0.2^2(1-0.2)^{10-2} = 0.0067$$

Graficamente representada por

```
n_amostra<-10
amostra<-c(1,0,1,0,0,0,0,0,0,0)
f_p<-function(p){p^sum(amostra)*(1-p)^(n_amostra-sum(amostra))*(0<=p & p<=1)}

curve(f_p,xlab="p",ylab="f(p)")
```



Para $p = 0.1$, $f(x_1, \ldots, x_{10} \mid p = 0.1) = f(x_1 \mid p = 0.1) \times f(x_2 \mid p = 0.1) \times \cdots \times f(x_{10} \mid p = 0.1) =$

```
f_p(0.1)
```

```
## [1] 0.004304672
```

E para $p = 0.2$, $f(x_1, x_2, \ldots, x_{10} \mid p = 0.2) =$

```
f_p(0.2)
```

```
## [1] 0.006710886
```

c)

Uma vez que a experiência aleatória consiste em selecionar aleatoriamente uma foto de um banco de fotos genéricas e verificar se inclui ($x = 1$) ou não ($x = 0$) gatinhos.

A estatística $T_1 = \sum_{i=1}^{10} X_i$ é a soma de todos os sucessos da experiência aleatória

> Ou seja, o nº de vezes que, ao selecionar aleatoriamente 10 fotos de um banco de fotos genéricas, verificar-se que inclui gatinhos

(**TC -** T_1 representa o número de sucessos nas 10 provas consideradas, ou seja, o número de fotos de gatinhos em 10 fotos visionadas.)

Sendo uma soma de *Bernoulli* independentes com o mesmo p, a sua distribuição é *Binomial*, pelo que

$$E\left[\sum_{i-1}^{10} X_i\right] = n \times p = 10p$$

Logo, o valor esperado é dado por

$$E[X] = np = 10p$$

onde

- n, é a dimensão da a.a. *(amostra aleatória)*, que é, neste caso, $n = 10$.

e

- p, o parâmetro da distribuição de *Bernoulli*

d)

Mantendo as mesmas considerações da alínea anterior,

a estatística $T_2 = \frac{\sum_{i=1}^{10} X_i}{10}$ é a média amostral de *sair uma foto escolhida ao acaso que tem gatinhos*

(**TC -** T_2 representa o número médio de fotos de gatinhos em 10 fotos visionadas.)

$$E\left[\frac{\sum_{i-1}^{10} X_i}{10}\right] = \frac{E\left[\sum_{i-1}^{10} X_i\right]}{10} = \frac{10p}{10} = p$$

Logo, o seu valor esperado, $E[.]$, é

$$E[\overline{X}] = p$$

$\therefore$ Daí esta estimativa ser um ***estimador*** *não enviesado para* p.

Exercício (Turma 2)

1.

Suponha a experiência aleatória que consiste em fazer scroll, sem qualquer regra, num banco de fotos genéricas e contar quantas fotos de gatinhos vê em 5 minutos.

Ou seja, considere a seguinte variável aleatória $X\breve{}$ número de fotos de gatinhos visionadas em 5 minutos.

Assuma que está em condições de considerar X como tendo distribuição de *Poisson* de parâmetro λ.

Pretende considerar amostras de dimensão 10 dessa população, $(X_1, X_2, \ldots, X_{10})$.

a. Construa a função de probabilidade conjunta de $(X_1, X_2, \ldots, X_{10})$.

b. Se $\lambda = 3.9$, qual a probabilidade de observar a amostra $(3, 0, 4, 5, 1, 2, 4, 7, 6, 1)$? E se $\lambda = 4.1$?

c. Como pode interpretar (no contexto apresentado) a estatística $T_1 = \sum_{i=1}^{10} X_i$? Qual o seu valor esperado?

d. Como pode interpretar (no contexto apresentado) a estatística $T_2 = \frac{\sum_{i=1}^{10} X_i}{10}$? Qual o seu valor esperado?

---

1. Resolução

a)

Como

> **X -** número de fotos de gatinhos visionadas em 5 minutos, assumida como tendo distribuição *Poisson* de parâmetro λ,

então a **função de probabilidade** de X é

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

Como assumimos que $(X_1, X_2, \ldots, X_{10})$ é **amostra aleatória** desta população, então a função conjunta será dada pelo produto das marginais (porque os X_i são *independentes e identicamente distribuídos*, com distribuição igual à população).

Logo temos

$$f(x_1, x_2, \ldots, x_{10}) = \prod_{i=1}^{10} \frac{e^{-\lambda}\lambda^{x_i}}{x_1!} = \frac{e^{-10\lambda}\lambda^{\sum_{t=1}^{10} x_i}}{\prod_{1=1}^{10} x_1!}$$

b)

Pretende-se

$$f(3,0,4,5,1,2,4,7,6,1 \mid \lambda = 3.9) = \frac{e^{-39} \times 3.9^{\Sigma x_i}}{\Pi x_i!}$$

e

$$f(3,0,4,5,1,2,4,7,6,1 \mid \lambda = 4.1) = \frac{e^{-41} \times 4.1^{\Sigma x_t}}{\Pi x_t!}$$

```r
x<-c(3,0,4,5,1,2,4,7,6,1)
sum_x<-sum(x)

fact_x <- factorial(x)

f1 <- function(m){
  val <- (exp(-10*m)/(prod(fact_x)))*(m^(sum_x))
  return(val)
  }

p1<-f1(3.9)
p2<-f1(4.1)
```

Ou seja,

$$f(3,0,4,5,1,2,4,7,6,1 \mid \lambda = 3.9) = 1.2277197 \times 10^{-10} \times 10^{-10}$$

e

$$f(3,0,4,5,1,2,4,7,6,1 \mid \lambda = 4.1) = 8.6545702 \times 10^{-11} \times 10^{-11}$$

c)

A estatística $T_1 = \sum_{i=1}^{10} X_1$ representa o número de fotos de gatinhos no total dos 10 períodos de 5 minutos, ou seja, o número de fotos de gatinhos visionadas em 50 minutos.

Sendo uma soma de *Poissons* independentes com o mesmo λ, a sua distribuição é Poisson com parâmetro $10 \times \lambda$, pelo que

$$E\left[\sum_{1=1}^{10} X_i\right] = 10 \times \lambda$$

d)

A estatística $T_2 = \frac{\sum_{i=1}^{10} X_i}{10}$ representa o número médio de fotos de gatinhos visionadas em cada período.

Pelas propriedades do valor esperado, $E[.]$, teremos

$$E\left[\frac{\sum_{1=1}^{10} X_1}{10}\right] = \frac{E\left[\sum_{1=1}^{10} X_1\right]}{10} = \frac{10\lambda}{10} = \lambda$$