

CE-009: Introdução a Estatística - Turma A

Avaliações Semanais (1º semestre 2017)

Avaliação 01

- Um determinado exame tem probabilidade de 0,85 de detectar uma doença enquanto que um segundo tipo de exame tem probabilidade de 0,70. A doença é considerada presente se algum dos exames é positivo. Considere que um material com a doença vai ser testado por ambos exames.
 - Descreva os eventos relevantes com uma notação apropriada.
 - Forneça o espaço amostral na forma de um conjunto e aponte suas características (finito ou infinito, enumerável ou não enumerável, equiprovável ou não).
 - Defina em notação o evento “a doença é detectada” e forneça o conjunto que define este evento.
 - Qual a probabilidade da doença ser detectada?
 - Qual a suposição feita no cálculo da probabilidade do anterior?
 - Se três materiais com a doença forem testados com ambos exames, qual a probabilidade de que todos deem (falso) “negativo”.

Solução:

- (a) **Notação:**

A : doença detectada no primeiro exame
 B : doença detectada no segundo exame
 DD : a doença é detectada
 $P[A] = 0,85$
 $P[B] = 0,70$

- (b) $\Omega = \{(AB), (A\bar{B}), (\bar{A}B), (\bar{A}\bar{B})\}$.
O espaço amostral é *finito*, *enumerável* e *não equiprovável*.
- (c) $DD = \{(AB), (A\bar{B}), (\bar{A}B)\}$.
- (d) $P[DD] = P[A \cup B] = P[A] + P[B] - P[A \cap B] = P[A] + P[B] - P[A] \cdot P[B] = 0,85 + 0,70 - 0,85 \cdot 0,70 = 0.955$
- (e) Independência entre os resultados dos dois exames.
- (f) Supondo independência entre materiais e exames:
 $P[DD]^3 = 0.871$

-
- Um material genético de feijão em desenvolvimento foi testado quanto à resistência a duas doenças que afetam comumente a cultura. Os resultados de 100 exames são resumidos na tabela a seguir.

resistência à doença B	resistência à doença A	
	alta	baixa
alta	80	9
baixa	6	5

Denote por A o evento *o material tem alta resistência à doença A* e por B o evento *o material alta resistência à doença B*.

- Obtenha: $P[A]$, $P[A \cap B]$, $P[A^c]$, $P[A^c \cap B^c]$, $P[A^c \cup B]$.
- Obtenha: $P[A|B]$, $P[B|A]$, $P[A|B^c]$, $P[B^c|A]$, $P[B|A^c]$.
- Se um material é selecionado ao acaso qual a probabilidade de ter:
 - alta resistência a A e baixa a B?
 - alta resistência a B e baixa a A?
- os eventos alta resistência a ambas doenças são mutuamente exclusivos? (justifique)
- os eventos alta resistência a ambas doenças são independentes? (justifique)

Solução:

```
> m <- matrix(c(80,6,9,5),ncol=2,dimnames=list(c("A", "A^c"),c("B", "B^c")))
> m
      B B^c
A    80  9
A^c  6  5
> mp <- prop.table(m)
> mp
      B B^c
A    0.80 0.09
A^c 0.06 0.05
(a) •  $P[A] = 0.86$ 
    •  $P[A \cap B] = 0.8$ 
    •  $P[A^c] = 0.14$ 
    •  $P[A^c \cap B^c] = 0.05$ 
    •  $P[A^c \cup B] = P[A^c] + P[B] - P[A^c \cap B] = 0.94$ 
(b) •  $P[A|B] = \frac{P[A \cap B]}{P[B]} = 0.9$ 
    •  $P[B|A] = \frac{P[A \cap B]}{P[A]} = 0.93$ 
    •  $P[A|B^c] = \frac{P[A \cap B^c]}{P[B^c]} = 0.55$ 
    •  $P[B^c|A] = \frac{P[B^c \cap A]}{P[A]} = 0.43$ 
    •  $P[B|A^c] = \frac{P[A^c \cap B]}{P[A^c]} = 0.64$ 
(c) •  $P[B \cap A^c] = 0.06$ 
    •  $P[A \cap B^c] = 0.09$ 
(d) Não, pois  $P[A \cap B] \neq 0$ , isto é, os eventos ter alta resistência a ambas doenças possuem intersecção, por isso não são mutuamente exclusivos. No contexto do exemplo, isto significa, por exemplo, que é possível ter resistência a ambas doenças ao mesmo tempo.
(e)  $P[A \cap B] \neq P[A] \cdot P[B]$ , isto é, o produto das marginais difere dos valores observados, por isso sabemos que os eventos não são independentes. No contexto do exemplo, as chances de ter resistência a uma doença são diferentes quando se tem ou não resistência à outra. Dizendo de outra forma, a probabilidade marginal (em todo universo) de se ter uma das doenças é diferente da condicional (no subgrupo que tem a outra doença), i.e.  $P[A] \neq P[A|B]$ .
> addmargins(mp)
      B B^c Sum
A    0.80 0.09 0.89
A^c 0.06 0.05 0.11
Sum 0.86 0.14 1.00
> outer(rowSums(mp), colSums(mp), "*")
      B B^c
A    0.7654 0.1246
A^c 0.0946 0.0154
> ## note tb a ordem entre os termos!
> #outer(apply(m/sum(m),2,sum),apply(m/sum(m),1,sum),"*")
```

Avaliação 02

- Em um lote estão misturadas 20 sementes de uma determinada cultivar (A) de uva com 5 de outra (B). Não é possível distinguir facilmente as sementes, o que só pode ser feito em um exame detalhado. Considere as diferentes possibilidades abaixo e calcule as probabilidades solicitadas.
 - Se forem retiradas ao acaso de uma só vez quatro sementes para inspeção, qual a probabilidade de obter ao menos uma de B ?
 - Se forem retiradas ao acaso e inspecionadas as sementes uma a uma, retornando a semente retirada ao lote, qual a probabilidade de que sejam retiradas três ou mais de A antes de se retirar a primeira de B ?
 - Finalmente, considere que as sementes serão retiradas ao acaso e inspecionadas uma a uma, retornando a semente retirada ao lote antes da próxima retirada e serão feitas exatamente quatro retiradas, Qual a probabilidade de obter ao menos uma de B ?

Solução:

(a)

 X : Número sementes do tipo B entre quatro retiradas

$$x \in \{0, 1, 2, 3, 4\}$$

$$X \sim \text{HG}(N = 25, K = 5, n = 4)$$

$$P[X = x] = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}} = \frac{\binom{5}{x} \binom{25-5}{4-x}}{\binom{25}{n}}$$

$$P[X \geq 1] = P[X = 1] + P[X = 2] + P[X = 3] + P[X = 4] = 1 - P[X = 0] = 1 - \frac{\binom{5}{0} \binom{20}{4}}{\binom{25}{4}} = 0.617$$

(b)

 X : Número de sementes retiradas de A até retirar a primeira de B

$$X \sim G(p = 5/25 = 0, 2)$$

$$x \in \{0, 1, 2, 3, 4, 5, \dots\}$$

$$P[X = x] = (1 - p)^x \cdot p = (0, 8)^x 0, 2$$

$$P[X \geq 3] = P[X = 3] + P[X = 4] + P[X = 5] + \dots = 1 - P[X < 3] = 1 - \{P[X = 0] + P[X = 1] + P[X = 2]\} = 1 - \{(0, 8)^0 0, 2 + (0, 8)^1 0, 2 + (0, 8)^2 0, 2\} = 0.422$$

(c)

 X : Número sementes do tipo B entre quatro retiradas (com reposição)

$$x \in \{0, 1, 2, 3, 4\}$$

$$X \sim B(n = 4, p = 0, 20)$$

$$P[X = x] = \binom{n}{x} p^x (1 - p)^{n-x} = \binom{4}{x} 0, 2^x 0, 8^{4-x}$$

$$P[X \geq 1] = P[X = 1] + P[X = 2] + P[X = 3] + P[X = 4] = 1 - P[X = 0] = 1 - \binom{4}{0} 0, 2^0 0, 8^4 = 0.59$$

Soluções computacionais com o programa R:

```
> (pa <- phyper(0, m=5, n=20, k=4, lower=F))
```

```
[1] 0.617
```

```
> (pb <- pgeom(2, p=0.25, lower=F))
```

```
[1] 0.4219
```

```
> (pc <- pbinom(0, size=4, prob=0.2, lower=F))
```

```
[1] 0.5904
```

2. Registros mostram que em determinado bairro são registradas, em média, 1,7 violações de residências por dia. Assuma uma distribuição de probabilidades possivelmente adequada e calcule as probabilidades de que hajam

- três ou mais ocorrências em um dia;
- nenhuma ocorrência em um dia;
- entre uma e quatro ocorrências em um dia.

Solução: X : Número diário de violações

$$x \in \{0, 1, 2, 3, 4, \dots\}$$

$$X \sim P(\lambda = 1, 7)$$

$$P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!} = \frac{e^{-1.7} 1.7^x}{x!}$$

$$(a) P[X \geq 3] = P[X = 3] + P[X = 4] + P[X = 5] + \dots = 1 - \{P[X = 0] + P[X = 1] + P[X = 2]\} = 1 - \left\{ \frac{e^{-1.7} 1.7^0}{0!} + \frac{e^{-1.7} 1.7^1}{1!} + \frac{e^{-1.7} 1.7^2}{2!} \right\} = 1 - e^{-1.7} \left\{ 1 + 1.7 + \frac{1.7^2}{2} \right\} = 0.243$$

$$(b) P[X = 0] = \frac{e^{-1.7} 1.7^0}{0!} = 0.183$$

$$(c) P[1 \leq X \leq 4] = P[X = 1] + P[X = 2] + P[X = 3] + P[X = 4] \dots = \left\{ \frac{e^{-1.7} 1.7^1}{1!} + \frac{e^{-1.7} 1.7^2}{2} + \frac{e^{-1.7} 1.7^3}{3!} + \frac{e^{-1.7} 1.7^4}{4!} \right\} = 1 - e^{-1.7} \left\{ 1.7^1 + \frac{1.7^2}{2} + \frac{1.7^3}{6} + \frac{1.7^4}{24} \right\} = 0.788$$

Soluções computacionais com o programa R:

```
> (pa <- ppois(2, lam=1.7, lower=F))
```

```
[1] 0.2428
```

```
> (pb <- dpois(0, lam=1.7))
```

```
[1] 0.1827
```

```
> (pc <- diff(ppois(c(0,4), lam=1.7)))
```

```
[1] 0.7877
```
