AULA 5 – Distribuição Binomial

Lista 5 Classe BD2020 – G A B A R I T O

Exercício 1

Discuta a validade do modelo binomial nos seguintes casos.

(a) Dos alunos da USP, sorteamos 5 e contamos quantos se declaram usuários regulares do CEPEUSP;

Nesse exemplo o ensaio de Bernoulli corresponde a observar se o aluno selecionado frequenta regularmente (sucesso) ou não (fracasso) o CEPEUSP. Cada sorteio será considerado independente dos demais e suporemos também que a probabilidade de sucesso é a mesma em cada um dos 5 sorteios. Sendo assim, o modelo binomial é válido nesse caso.

Exercício 1 (cont.)

(b) Escolhemos 20 lâmpadas ao acaso na prateleira de um supermercado, sendo 10 de uma fábrica e 10 de outra. Contamos o número total de lâmpadas defeituosas;

Nesse exemplo o ensaio de Bernoulli corresponde a observar se a lâmpada selecionada é defeituosa (sucesso) ou não (fracasso). Cada sorteio será considerado independente dos demais, mas não podemos supor que a probabilidade de sucesso seja a mesma em cada um dos 20 sorteios, pois se trata de dois fabricantes de lâmpadas com proporção de defeituosas diferentes em sua fabricação.

Sendo assim, o modelo binomial não é válido nesse caso.

Exercício 1 (cont.)

(c) Quinze automóveis 0 km de um mesmo fabricante e mesmo modelo são submetidos a um teste anti-poluição e contamos quantos passaram no teste;

Nesse exemplo o ensaio de Bernoulli corresponde a observar se o carro selecionado passa no teste (sucesso) ou não (fracasso). Cada sorteio será considerado independente dos demais e suporemos também que a probabilidade de sucesso e fracasso é a mesma em cada um dos 5 sorteios. Sendo assim, o modelo binomial é válido nesse caso.

Exercício 1 (cont.)

(d) Um motorista é submetido a um teste em que deva estacionar seu veículo num pequeno espaço (isto é popularmente chamado de fazer baliza).
Em 10 tentativas, contamos o número de vezes em que o motorista estacionou corretamente.

Nesse exemplo o ensaio de Bernoulli corresponde a observar se o motorista selecionado estacionou o carro corretamente (sucesso) ou não (fracasso). Os diferentes sorteios não podem ser considerados independentes dos demais pois há acúmulo de experiência. Sendo assim, o modelo binomial não é válido nesse caso.

Exercício 2

Uma empresa oferece quatro modalidades de serviço A, B, C e D cobrando 100, 200, 300 e 400 (unidades monetárias), respectivamente. Sabe-se que um cliente contrata a modalidade A com probabilidade 0,2; a modalidade B com probabilidade 0,4; a C com probabilidade 0,3 e a D com probabilidade 0,1. Defina por X a variável que representa o ganho da empresa por cliente.

(a) Construa a distribuição de probabilidades de X.

x	100	200	300	400
$P(X=x)$	0,2	0,4	0,3	0,1

Exercício 2 (cont.)

(b) Calcule o ganho médio da empresa por cliente.

$E(X)$ = 100 x 0,2 + 200 x 0,4 + 300 x 0,3 + 400 x 0,1 = 230.

Logo, o ganho médio por cliente é de 230 unidades monetárias

(c) Calcule o desvio padrão do ganho da empresa por cliente.

$$\begin{aligned} \text{Var(X)} &= (0{,}2 \text{ x} 100^2 + 0{,}4 \text{ x } 200^2 + 0{,}3 \text{ x } 300^2 + 0{,}1 \text{ x } 400^2) - 230^2 \\ &= 61000 - 52900 = 8100. \end{aligned}$$

Logo, o desvio padrão é $DP(X)$= 90.

Exercício 3

Três em cada quatro alunos de uma universidade fizeram cursinho antes de prestar vestibular. Se 16 alunos são selecionados ao acaso, calcule (use 4 casas decimais):

(a) a probabilidade de que exatamente 6 não tenham feito cursinho;

Obs. Aqui, cada seleção é um ensaio de Bernoulli correspondendo a observar se um aluno selecionado fez cursinho (sucesso) ou não (fracasso). Cada selção será considerada independente das demais e podemos supor que a probabilidade de sucesso e fracasso é a mesma em cada sorteio. Sendo assim, o modelo binomial é válido.

Defina a $v.a.$ X = número de alunos selecionados que fizeram cursinho, dentre os 16 selecionados

Uma vez que o modelo binomial é válido, temos $X \sim b(16; ¾)$.

A probabilidade pedida é $P(X=10)$, pois exatamente 6 não terem feito cursinho é equivalente a 10 que fizeram.

Da tabela (gerada pelo $Rcmdr$) $\Rightarrow P(X=10) = 0{,}1101$.

	Probability
0	2.328306e-10
1	1.117587e-08
2	2.514571e-07
3	3.520399e-06
4	3.432389e-05
5	2.471320e-04
6	1.359226e-03
7	5.825255e-03
8	1.966024e-02
9	5.242730e-02
10	**1.100973e-01**
11	1.801593e-01
12	2.251991e-01
13	2.078761e-01
14	1.336346e-01
15	5.345384e-02
16	1.002260e-02

Exercício 3 (cont.)

(b) a probabilidade de que pelo menos 12 tenham feito cursinho;

Podemos calcular usando a tabela de distribuição da binomial no *Rcmdr*:

Distribuições → *Distribuições discretas* → *Distribuição binomial* → *Probabilidades da binomial* (com n=16 e p=0.75)

Assim,

$$P(X \geq 12) = P(X=12) + P(X=13) + P(X=14) + P(X=15) + P(X=16)$$
$$= 0{,}2252 + 0{,}2079 + 0{,}1336 + 0{,}0535 + 0{,}0100 = 0{,}6302.$$

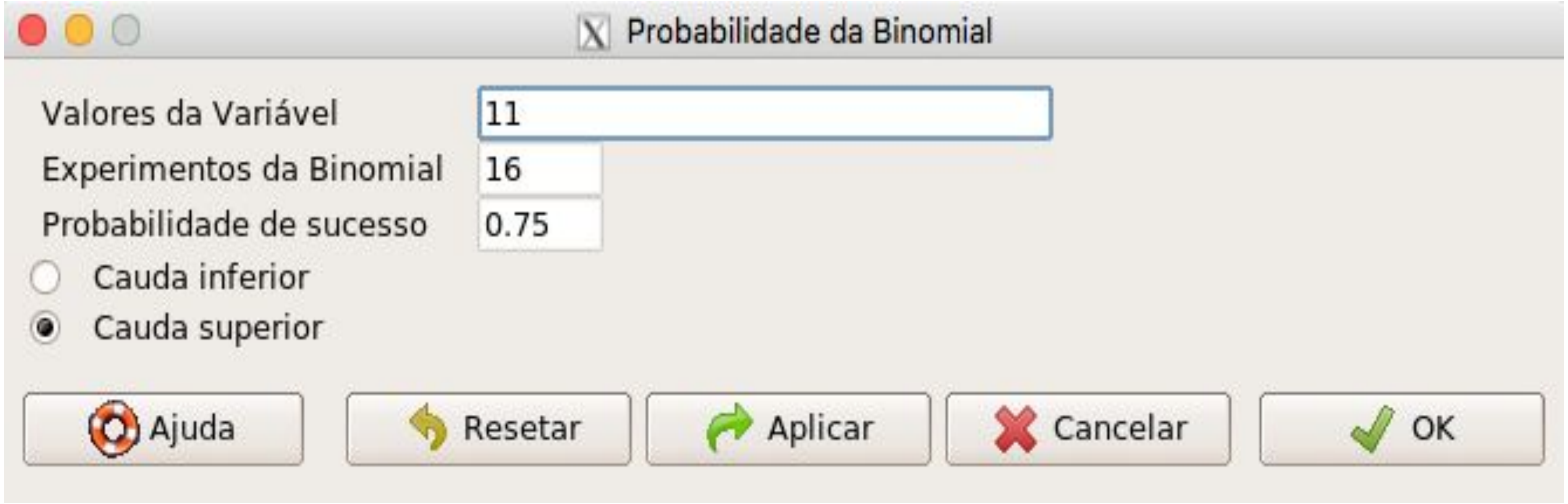
Saída no Rcmdr:

	Probability
0	2.328306e-10
1	1.117587e-08
2	2.514571e-07
3	3.520399e-06
4	3.432389e-05
5	2.471320e-04
6	1.359226e-03
7	5.825255e-03
8	1.966024e-02
9	5.242730e-02
10	1.100973e-01
11	1.801593e-01
12	**2.251991e-01**
13	**2.078761e-01**
14	**1.336346e-01**
15	**5.345384e-02**
16	**1.002260e-02**

Exercício 3 (b) (cont.)

Além disso, uma vez que $P(X \geq 12) = P(X > 11)$, podemos calcular a probabilidade caudal superior de X diretamente no *Rcmdr*:

Distribuições → Distribuições discretas → Distribuição binomial → Probabilidades da cauda da binomial



Saída no *Rcmdr*

```
> pbinom(c(11), size=16, prob=0.75, lower.tail=FALSE)
[1] 0.6301862
```

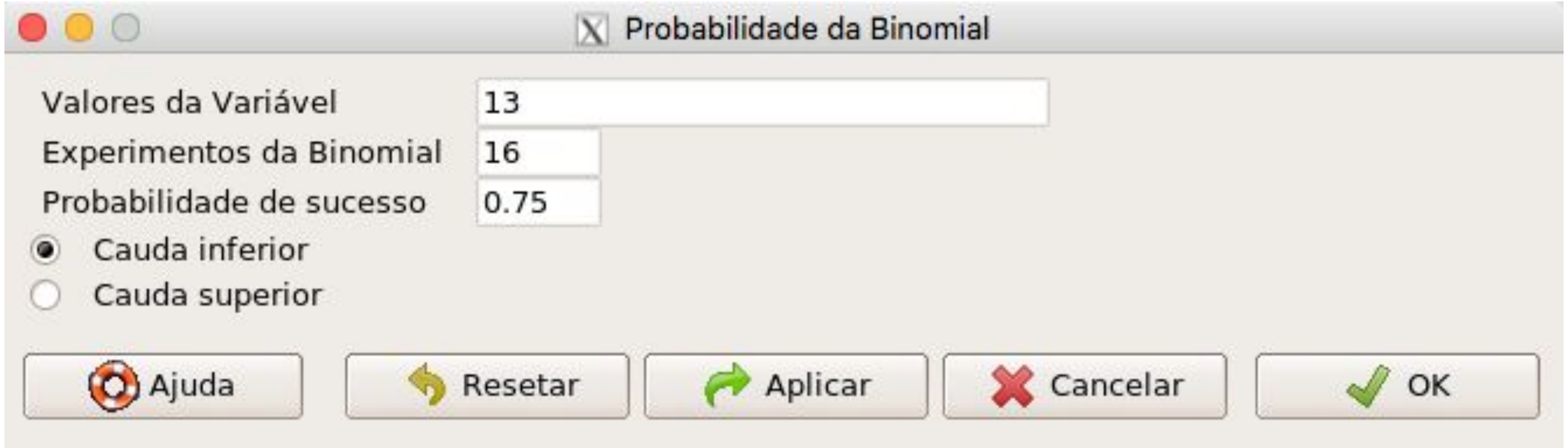
Logo, a probabilidade de que pelo menos 12 alunos tenham feito cursinho é 0,63

Exercício 3 (cont.)

(c) a probabilidade de que no máximo 13 tenham feito cursinho;

$P(X \leq 13)$ pode ser calculada diretamente no *Rcmdr* através da probabilidade de caudal sinferior de X: :

Distribuições → Distribuições discretas → Distribuição binomial → Probabilidades da cauda da binomial



Saída no *Rcmdr*:

> pbinom(c(13), size=16, prob=0.75, lower.tail=TRUE)

[1] 0.802889

Logo, a probabilidade de que no máximo 13 alunos tenham feito cursinho é 0,802.

Exercício 3 (cont.)

(d) a probabilidade de que entre 9 e 14 (inclusive) tenham feito cursinho (use 4 decimais).

Logo, $P(9 \leq X \leq 14) = P(X=9) + P(X=10) + \ldots + P(X=14)$

$= 0{,}0524 + 0{,}1101 + 0{,}1801 + 0{,}2251 + 0{,}2079 + 0{,}1335 = 0{,}9092.$

Alternativamente,

$P(9 \leq X \leq 14) = P(X \leq 14) - P(X < 9) = P(X \leq 14) - P(X \leq 8)$, e podemos calcular a probabilidade através da diferença entre as probabilidades caudais:

Distribuições → Distribuições discretas → Distribuição binomial → Probabilidades das caudas da binomial

Saída no *Rcmdr* para k=14

```
> pbinom(c(14), size=16, prob=0.75, lower.tail=TRUE)
[1] 0.9365236
```

Saída no *Rcmdr* para k=8

```
> pbinom(c(8), size=16, prob=0.75, lower.tail=TRUE)
[1] 0.02712996
```

Logo, $P(9 \leq X \leq 14) = 0{,}9365 - 0{,}0271 = 0{,}9094.$

Portanto, a probabilidade de que entre 9 e 14 alunos tenham feito cursinho é 0,909.

Exercício 3 (cont.)

(e) Para um grupo de 80 alunos selecionados ao acaso, qual é o número esperado de alunos que fizeram cursinho? E o desvio padrão? (use 2 casas decimais)

Nesse caso, a $v.a.$ X, dada pelo número de alunos que fizeram cursinho, dentre 80 selecionados, tem distribuição $b(80; ¾)$.

Logo, $E(X) = n \times p$ = 80 x 0,75 = 60,

$Var(X) = n \times p \times (1-p)$ = 80 x 0,75 x 0,25 = 15 e, portanto, $DP(X)$=3,87.

Exercício 4

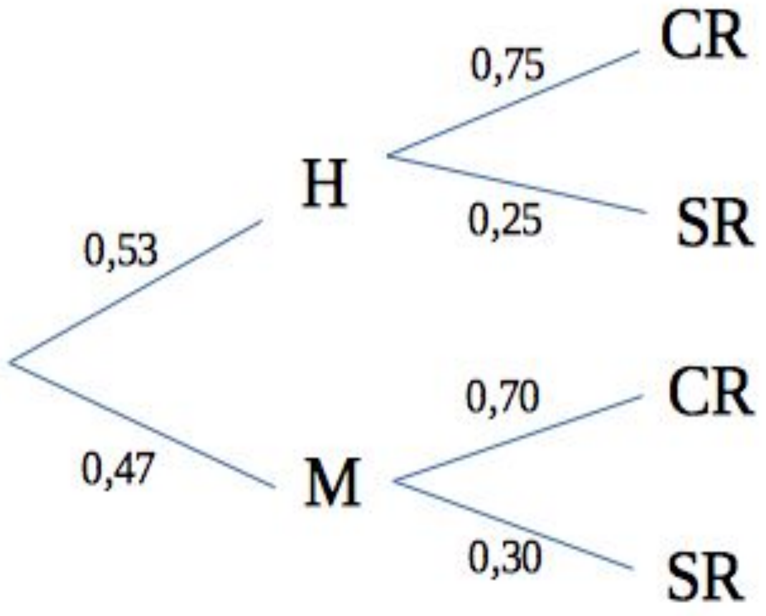
Suponhamos que 25% dos homens trabalhadores e 30% das mulheres trabalhadoras de uma população não tenham registro em carteira de trabalho. Suponhamos também que a população de trabalhadores seja constituída por 53% de homens e 47% de mulheres.

(a) Qual é a proporção de trabalhadores **sem** registro em carteira na população?

Considere os seguintes eventos:

$P(SR)$ **= ?**

H: o trabalhador é homem.
M: o trabalhador é mulher.
CR: o trabalhador tem registro na carteira de trabalho.
SR: o trabalhador não tem registro na carteira de trabalho.



Usando o diagrama de árvore ao lado temos que

$P(SR) = P(H \cap SR) + P(M \cap SR)$

$= P(H)\, P(SR \mid H) + P(M)\, P(SR \mid M)$

$= 0{,}53 \times 0{,}25 + 0{,}47 \times 0{,}30 = 0{,}133 + 0{,}141 = 0{,}274.$

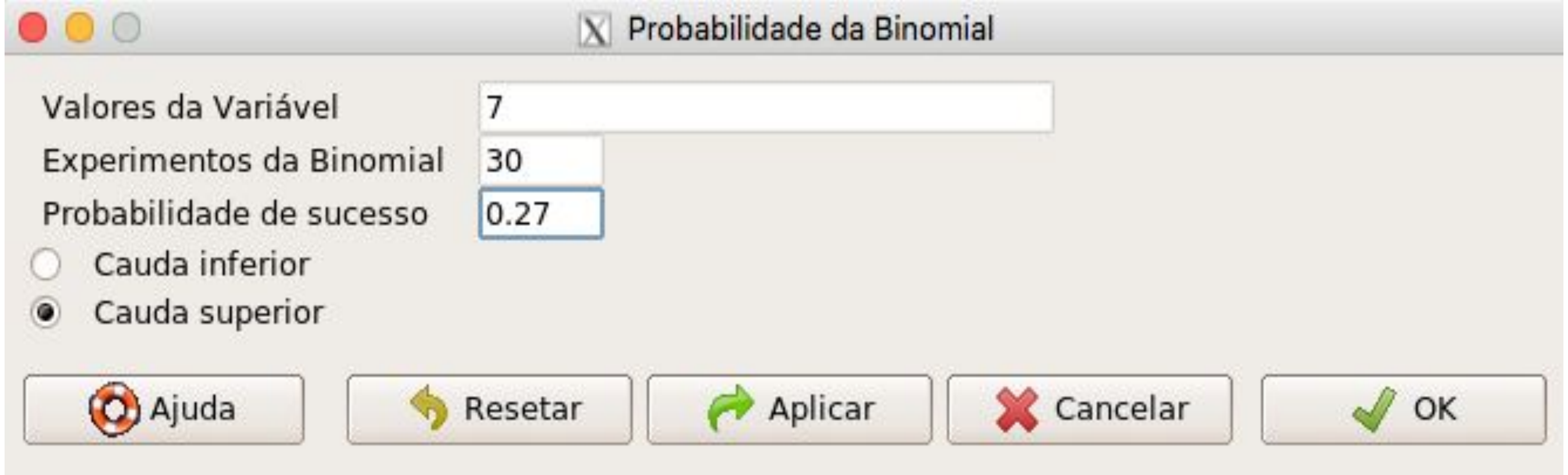
Assim, a proporção de trabalhadores **sem registro** em carteira de 27,4%.

Exercício 4 (cont.)

(b) Se 30 trabalhadores forem selecionados ao acaso dessa população, qual é a probabilidade de que pelo menos 8 não tenham registro em carteira?

Defina X a *v.a.* número de trabalhadores **sem** registro em carteira de trabalho, dentre os 30 selecionados $\Rightarrow X \sim b(30; 0{,}27)$

Como $P(X \geq 8) = P(X > 7)$, podemos usar o *Rcmdr* para calcular a probabilidade de cauda superior da binomial.



Saída no *Rcmdr:*

```
> pbinom(c(7), size=30, prob=0.27, lower.tail=FALSE)
[1] 0.5853472
```

Assim, a probabilidade de que pelo menos 8 trabalhadores não tenham registro em carteira é **0,585**.

Exercício 4 (cont.)

(c) Em média, quantos trabalhadores sem registro em carteira esperamos encontrar dentre os 30 selecionados? E qual é o desvio padrão do número de trabalhadores sem registro em carteira?

$E(X) = n \times p$ = 30 x 0,27 = 8,1.

$Var(X) = n \times p \times (1-p)$ = 30 × 0,27 × 0,73 = 5,91

$DP(X) \cong$ 2,43.