MAE0119 INTRODUÇÃO À ESTATÍSTICA E PROBABILIDADE

FOLHA DE RESPOSTAS - LISTA 5

Nome: Beatriz Viana Costa **Número USP: 13673214**

1.

a)

Dados do exercício:

- *Erro máximo = 0,02;*
- *Intervalo de confiança = 0,92.*

Temos a variável aleatória X, tal que $\{X_1, ..., X_i\}$ são os eventos da amostra observada, de forma que formam uma sequência de variáveis aleatórias independentes com distribuição Bernoulli:

$$X_i = \begin{cases} \text{1, se a pessoa é ouvinte do programa de rádio} \\ \text{0, se a pessoa não é ouvinte do programa de rádio} \end{cases}$$

Dessa forma podemos realizar uma apriximação com a normal, onde $\mu = p$ e $\sigma^2 = \frac{p(1-p)}{n}$, assim temos:

$$X_i \sim N\ (p, \tfrac{p(1-p)}{n})$$

Temos então que:

$$P(|\bar{X} - p| < 0,02) = 0{,}92$$
$$P(-0{,}02 < \bar{X} - p < 0{,}02) = 0{,}92$$
$$P(-\tfrac{0{,}02}{\sqrt{VAR(\bar{X})}} < Z < \tfrac{0{,}02}{\sqrt{VAR(\bar{X})}})$$

Temos então que dividir nossa análise em dois casos, para quando $\hat{p} = 0,15$ e para quando $\hat{p} = 0,3$, e então considerar este intervalo.

I. Para $\hat{p} = 0,15$:

$$VAR(\bar{X}) = \tfrac{0{,}15*0{,}85}{n} = \tfrac{0{,}1275}{n}$$
$$\sqrt{VAR(\bar{X})} = \tfrac{0{,}357}{\sqrt{n}}$$
$$P(-\tfrac{\sqrt{n}*0{,}02}{0{,}357} < Z < \tfrac{\sqrt{n}*0{,}02}{0{,}357}) = 0{,}92$$
$$\tfrac{\pm\sqrt{n}*0{,}02}{0{,}357} = 1{,}75 \implies (\pm\sqrt{n})^2 = 31,238^2 = n = 976$$

II. Para $\hat{p} = 0,3$:

$$VAR(\bar{X}) = \tfrac{0{,}3*0{,}7}{n} = \tfrac{0{,}21}{n}$$
$$\sqrt{VAR(\bar{X})} = \tfrac{0{,}458}{n}$$
$$P(-\tfrac{\sqrt{n}*0{,}02}{0{,}458} < Z < \tfrac{\sqrt{n}*0{,}02}{0{,}458}) = 0{,}92$$
$$\tfrac{\pm\sqrt{n}*0{,}02}{0{,}458} = 1{,}75 \implies (\pm\sqrt{n})^2 = (40,075)^2 = n = 1606$$

Ou seja, o tamanho de uma amostra aleatória simples para estimar a verdadeira proporção com erro máximo de 0,02 e probabilidade de 92% deve estar dentro do intervalo de [976; 1606] pessoas.

b)

Dados:

- *Tamanho da amostra n = 625;*
- *150 são ouvintes da rádio;*
- *Grau de confiança γ = 0,94.*

Temos a proporção amostral como sendo:

$$\hat{p} = \bar{X} = \frac{150}{625} = 0,24$$

E, portanto, conseguimos definir a distribuição e encontrar a variância:

$$\bar{X} \sim N(p, \frac{p(1-p)}{n})$$
$$VAR(\bar{X}) = \frac{0{,}24*0{,}76}{625} = 0,00029184$$
$$\implies \sqrt{VAR(\bar{X})} \approx 0,017$$

Seja E o erro máximo, temos a seguinte relação:

$$P(|\bar{X} - p| < E) = 0{,}94$$
$$P(-E < \bar{X} - p < E) = 0{,}94$$
$$P(-\frac{E}{0{,}017} < Z < \frac{E}{0{,}017}) = 0{,}94$$
$$ZE_{\overline{2}} = 1{,}88$$
$$\frac{E}{0{,}017} = 1,89 \implies E \approx 0,03$$

2.

a)

Dados:

- *Desvio padrão σ = 5cm;*
- *Tamanho da amostra n = 36;*
- *Média $\bar{X}$ = 150cm;*
- *Grau de confiança γ = 0,95.*

Estamos procurando o intervalo de confiança.
Temos a seguinte relação:

$$P(150 - 1{,}96 * \frac{5}{\sqrt{36}} < \mu < 150 + 1{,}96 * \frac{5}{\sqrt{36}}) = 0{,}95$$
$$\implies P\ (148{,}37 < \mu < 151{,}63) = 0{,}95$$

Logo, o intervalo de confiança de 95% para a média populacional é 148,37cm - 151,63cm, ou seja, 150 $\pm$ 1,63.

b)

Estamos procurando o tamanho da amostra aleatória simples de modo que o intervalo 150 $\pm$ 0,98 tenha 95% de confiança.
Tendo os dados do exercício anterior, temos a seguinte relação:

$$P(-0{,}98 < \bar{X} - \mu < 0{,}98) = 0{,}95$$
$$P(\frac{-\sqrt{n}*0{,}98}{5} < Z < \frac{\sqrt{n}*0{,}98}{5}) = 0{,}95$$
$$\implies \pm\frac{\sqrt{n}*0{,}98}{5} = 1,96$$
$$\pm\sqrt{n} = \frac{5*1{,}96}{0{,}98} = 10$$
$$\implies (\pm\sqrt{n})^2 = n = 100$$

Ou seja, o tamanho de amostra deve ser de 100 crianças para que haja 95% de confiança.

3.

a)

Dados:

- *Tamanho da amostra n = 1600;*
- *Quantidade de pessoas que possuem a característica desejada x = 300;*
- *Coeficiente α = 0,05 a probabilidade de rejeição da hipótese nula quando esta é verdadeira.*

Sabendo que $\{X_1, ..., X_n\}$ é definido como o conjunto das amostras e que X_i pode ser definida com a distribuição de Bernoulli, temos:

$$X_i \sim N\ (p, \frac{p(1-p)}{n})$$

$$X_i = \begin{cases} \text{1, se a pessoa assiste ao programa} \\ \text{0, caso contrário} \end{cases}$$

- *Hipótese nula H_0 = pelo menos 1/4 dos possuidores de televisão assistem o programa, ou seja, p $\geq$ 0,25;*
- *Hipótese alternativa H_1 = a quantidade de possuidores de televisão que assistem ao programa é menor que 1/4, ou seja, p < x_{c1}.*

Além disso podemos também encontrar $\hat{p}$ e a variância de $\bar{X}$:

$$\hat{p} = 0,25 * 0,75 = 0,1875$$
$$VAR(\bar{X}) = \sqrt{\frac{0,1875}{1600}} = 0,01$$

Iremos então procurar a região crítica (RC) utilizando o ponto crítico x_{c1}. Obtemos então a seguinte relação:

$$P(\bar{x} < x_{c1}|p = 0,25) = 0,05$$
$$P(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < \frac{x_{c1}-0,25}{0,01/\sqrt{1600}}) = 0,05$$
$$P(Z < z_{c1}) = 0,05 \quad z_{c1} = -1,65$$
$$x_{c1} = 0,25 + z_{c1}\frac{0,01}{40} = 0,2495875$$

Colhida a amostra, se $\bar{x}_{obs}$ é tal que $\bar{x}_{obs} < 0,2495875$, rejeitamos a hipótese nula, pois o conjunto de números reais menores que 0,2495875 é denominado região crítica.

$$z_{obs} = \frac{\hat{p}-\bar{x}}{VAR(\bar{X})} = \frac{0,1875-0,25}{0,01} = -6,25$$
$$\implies \bar{x}_{obs} = 0,2484375$$

Portando, como $\bar{x}_{obs} < \bar{x}$ encontramos evidências para rejeitar a hipóitese nula, assim os produtores devem tomar a decisão de mudar o programa de televisão.

b)

Vamos obter o valor-p.

$$P(\text{Valores mais extremos contra } H_0|H_0 \text{ é verdadeira})$$
$$P = P(\hat{p} \leq 0,2484375|H_0 \text{ verdadeira}) = P(\hat{p} \leq 0,2484375|\ p = 0,25)$$
$$P(Z \leq \frac{0,2484375-0,25}{0,01}) = P(Z \leq -0,15625)$$
$$= 1 - A(0,15625) = 1 - 0,5596 = 04404 \leq \alpha = 0,05; \text{ rejeita } H_0$$

4.

Dados:

- *900 crianças contrairam a doença;*
- *Das 450 que não receberam a vacina, 120 sofreram efeitos posteriores;*
- *Das 450 que receberam vacina, 150 sofreram efeitos posteriores.*

Estamos procurando o nível descritivo, ou seja, valor-p.
Tabela do exercício anexada na foto.

Iremos definir a hipótese nula H_0 *de que as chances de sofrer efeitos posteriores é a mesma para indivíduos que receberam e que não receberam a vacina.*
E também iremos definir a hipótese alternativa H_a *de que as chances de sofrer efeitos posteriores é diferente para indivíduos que receberam a vacina e para aqueles que não receberam.*

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

Valores esperados calculados com regra de três com base na tabela: $e_i = 135$, $e_2 = 135, e_3 = 315$ e $e_4 = 315$. *Substituindo os valores:*

$$\chi^2 \approx 4,890108$$

5.