DEST

Inferência Estatística

Lista 4

Lista de Exercícios 5 - Resolução

Nome: Miqueias T.
Data: 28 de agosto de 2025

Questão 1

Um experimento genético envolve uma população de moscas de frutas que consiste em 1 macho (Mike) e 3 fêmeas, chamadas Ana, Bárbara e Cristina. Suponha que duas moscas de frutas sejam selecionadas aleatoriamente com reposição.

a) *Depois de listar as 16 diferentes amostras possíveis, ache a proporção de fêmeas em cada amostra e, então, use uma tabela para descrever a distribuição amostral da proporção de fêmeas.*

b) *Ache a média da distribuição amostral.*

c) *A média da distribuição amostral (item b) é igual à proporção populacional de fêmeas?*

Solução:

a) A população total é de 4 moscas (M, A, B, C). A amostragem é com reposição, então há $4 \times 4 = 16$ amostras possíveis. A tabela abaixo lista todas as amostras e a proporção de fêmeas ($\hat{p}$) em cada uma.

1ª Mosca	**2ª Mosca**			
(Amostra)	**Macho (M)**	**Ana (A)**	**Bárbara (B)**	**Cristina (C)**
Macho (M)	(M,M) $\to \hat{p} = 0$	(M,A) $\to \hat{p} = 0.5$	(M,B) $\to \hat{p} = 0.5$	(M,C) $\to \hat{p} = 0.5$
Ana (A)	(A,M) $\to \hat{p} = 0.5$	(A,A) $\to \hat{p} = 1$	(A,B) $\to \hat{p} = 1$	(A,C) $\to \hat{p} = 1$
Bárbara (B)	(B,M) $\to \hat{p} = 0.5$	(B,A) $\to \hat{p} = 1$	(B,B) $\to \hat{p} = 1$	(B,C) $\to \hat{p} = 1$
Cristina (C)	(C,M) $\to \hat{p} = 0.5$	(C,A) $\to \hat{p} = 1$	(C,B) $\to \hat{p} = 1$	(C,C) $\to \hat{p} = 1$

Contando as ocorrências de cada proporção:

- $\hat{p} = 0$: 1 vez
- $\hat{p} = 0.5$: 6 vezes
- $\hat{p} = 1$: 9 vezes

A distribuição amostral da proporção de fêmeas é:

Proporção Amostral ($\hat{p}$)	**Probabilidade P($\hat{p}$)**
0	1/16
0.5	$6/16 = 3/8$
1.0	9/16

b) A média da distribuição amostral ($\mu_{\hat{p}}$) é o valor esperado de $\hat{p}$:

$$\mu_{\hat{p}} = E[\hat{p}] = \sum \hat{p} \cdot P(\hat{p})$$

$$\mu_{\hat{p}} = \left(0 \cdot \frac{1}{16}\right) + \left(0.5 \cdot \frac{6}{16}\right) + \left(1.0 \cdot \frac{9}{16}\right)$$

$$\mu_{\hat{p}} = 0 + \frac{3}{16} + \frac{9}{16} = \frac{12}{16} = \frac{3}{4} = 0.75$$

c) A proporção populacional de fêmeas (p) é de 3 fêmeas para 4 moscas no total, logo $p = 3/4 = 0.75$.

Sim, a média da distribuição amostral ($\mu_{\hat{p}} = 0.75$) é exatamente igual à proporção populacional ($p = 0.75$). Isso ocorre porque a proporção amostral $\hat{p}$ é um estimador não viciado da proporção populacional p.

Questão 2

As idades (anos) dos quatro presidentes dos Estados Unidos quando foram assassinados no exercício do cargo são 56 (Lincoln), 49 (Garfield), 58 (McKinley) e 46 (Kennedy).

a) *Supondo que duas das idades sejam selecionadas com reposição, liste as 16 diferentes amostras possíveis.*

b) *Ache a média de cada uma das 16 amostras e, então, resuma a distribuição amostral das médias no formato de uma tabela que represente uma distribuição de probabilidade.*

Solução:

a) As 16 amostras possíveis de tamanho 2, com reposição, são:

(56, 56)	(56, 49)	(56, 58)	(56, 46)
(49, 56)	(49, 49)	(49, 58)	(49, 46)
(58, 56)	(58, 49)	(58, 58)	(58, 46)
(46, 56)	(46, 49)	(46, 58)	(46, 46)

b) Primeiro, calculamos a média ($\bar{x}$) para cada uma das 16 amostras:

56.0	52.5	57.0	51.0
52.5	49.0	53.5	47.5
57.0	53.5	58.0	52.0
51.0	47.5	52.0	46.0

Agora, resumimos a distribuição amostral das médias, agrupando os valores idênticos e calculando suas probabilidades (frequência/16).

Média Amostral ($\bar{x}$)	**Frequência**	**Probabilidade P($\bar{x}$)**
46.0	1	$1/16$
47.5	2	$2/16 = 1/8$
49.0	1	$1/16$
51.0	2	$2/16 = 1/8$
52.0	2	$2/16 = 1/8$
52.5	2	$2/16 = 1/8$
53.5	2	$2/16 = 1/8$
56.0	1	$1/16$
57.0	2	$2/16 = 1/8$
58.0	1	$1/16$
Total	**16**	**1.0**

Questão 3

Repita o Exe 2 usando a mediana em lugar de médias.

Solução: O procedimento é análogo ao do Exercício 2, mas em vez da média, calculamos a mediana para cada uma das 16 amostras. As amostras possíveis são as mesmas.

É importante notar que para uma amostra de tamanho n=2, com valores $\{x_1, x_2\}$, a mediana é definida como a média dos dois valores centrais, ou seja, $(x_1 + x_2)/2$. Este cálculo é **idêntico ao da média** para uma amostra de tamanho 2. Portanto, os valores numéricos das medianas amostrais serão os mesmos das médias amostrais do exercício anterior.

Primeiro, calculamos a mediana para cada uma das 16 amostras:

56.0	52.5	57.0	51.0
52.5	49.0	53.5	47.5
57.0	53.5	58.0	52.0
51.0	47.5	52.0	46.0

Em seguida, resumimos a distribuição amostral das medianas, agrupando os valores idênticos e calculando suas probabilidades. Como os valores são os mesmos do exercício anterior, a distribuição de probabilidade também será idêntica.

Mediana Amostral	**Frequência**	**Probabilidade**
46.0	1	1/16
47.5	2	2/16 = 1/8
49.0	1	1/16
51.0	2	2/16 = 1/8
52.0	2	2/16 = 1/8
52.5	2	2/16 = 1/8
53.5	2	2/16 = 1/8
56.0	1	1/16
57.0	2	2/16 = 1/8
58.0	1	1/16
Total	**16**	**1.0**

Questão 4

Considere o seguinte problema (adaptado de Magalhães & Lima, 2006): Um fabricante afirma que sua vacina contra gripe imuniza em 80% dos casos. Uma amostra de 25 indivíduos entre os que tomaram a vacina foi sorteada e testes foram feitos para verificar a imunização ou não desses indivíduos.

a) *No contexto do problema identifique: a população, o parâmetro de interesse, o estimador, a estimativa, a distribuição amostral.*

b) *Se o fabricante estiver correto, qual é a probabilidade da proporção de imunizados na amostra ser inferior a 0,75? E superior a 0,85?*

Solução:

a) Identificação dos elementos estatísticos:

- **População:** O conjunto de todos os indivíduos que tomaram ou poderiam tomar a vacina.
- **Parâmetro de interesse:** A verdadeira proporção populacional (p) de indivíduos que são imunizados pela vacina. O valor afirmado pelo fabricante é $p = 0.80$.
- **Estimador:** É a estatística amostral usada para estimar o parâmetro p. Neste caso, é a proporção amostral, representada por $\hat{p} = \frac{X}{n}$, onde X é o número de indivíduos imunizados na amostra e n é o tamanho da amostra.
- **Estimativa:** É o valor numérico do estimador obtido a partir de uma amostra específica. O problema não fornece o resultado da amostra, mas se, por exemplo, 21 dos 25 indivíduos fossem imunizados, a estimativa seria $\hat{p} = \frac{21}{25} = 0.84$.
- **Distribuição Amostral:** É a distribuição de probabilidade de todas as possíveis proporções amostrais ($\hat{p}$) que poderiam ser obtidas de amostras de tamanho $n = 25$. Como a variável de interesse (ser ou não imunizado) é binária, o número de sucessos X em n ensaios segue uma distribuição Binomial, $X \sim B(n, p)$. A distribuição de $\hat{p}$ é, portanto, diretamente derivada da Binomial.

b) Para resolver este item, assumimos que a afirmação do fabricante é verdadeira, ou seja, $p = 0.80$ e $n = 25$. A variável X (número de imunizados na amostra) segue $X \sim B(25, 0.80)$.

Como o cálculo exato com a Binomial pode ser trabalhoso, podemos usar a **aproximação da Binomial pela Normal**, pois as condições são satisfeitas:

$$np = 25 \cdot 0.80 = 20 \geq 5$$

$$n(1-p) = 25 \cdot 0.20 = 5 \geq 5$$

Calculamos a média (μ) e o desvio padrão (σ) da distribuição de X:

$$\mu = np = 20$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{25 \cdot 0.80 \cdot 0.20} = \sqrt{4} = 2$$

Agora, convertemos as proporções em número de indivíduos ($X = n \cdot \hat{p}$):

- **Probabilidade de $\hat{p} < 0.75$:** Corresponde a $X < 25 \cdot 0.75 = 18.75$. Como X é um número inteiro, isso é $P(X \leq 18)$. Usando a **correção de continuidade**, aproximamos por $P(X < 18.5)$. Padronizando (cálculo do Z-score):

$$Z = \frac{18.5 - \mu}{\sigma} = \frac{18.5 - 20}{2} = \frac{-1.5}{2} = -0.75$$

A probabilidade é $P(Z < -0.75)$. Pela tabela da Normal Padrão, isso é aproximadamente **0.2266**.

- **Probabilidade de $\hat{p} > 0.85$:** Corresponde a $X > 25 \cdot 0.85 = 21.25$. Como X é um número inteiro, isso é $P(X \geq 22)$. Usando a **correção de continuidade**, aproximamos por $P(X > 21.5)$. Padronizando:

$$Z = \frac{21.5 - \mu}{\sigma} = \frac{21.5 - 20}{2} = \frac{1.5}{2} = 0.75$$

A probabilidade é $P(Z > 0.75)$, que é $1 - P(Z < 0.75)$. Pela tabela da Normal Padrão, isso é $1 - 0.7734 = \mathbf{0.2266}$.

Portanto, a probabilidade da proporção amostral ser inferior a 0,75 é de aproximadamente 22,66%, e a probabilidade de ser superior a 0,85 também é de aproximadamente 22,66%.

Questão 5

Uma variável aleatória Y tem distribuição normal, com média 100 e desvio padrão 10.

a) *Qual a P(90 < Y < 110)?*

b) *Se $\bar{Y}$ for a média de uma amostra de 16 elementos retirados dessa população, calcule P(90 < $\bar{Y}$ < 110).*

Solução: A variável aleatória da população, Y, segue uma distribuição Normal com $\mu = 100$ e $\sigma = 10$. Escrevemos: $Y \sim N(100, 10^2)$.

a) Para encontrar a probabilidade de uma única observação Y, primeiro padronizamos os valores 90 e 110 para a escala Z (Normal Padrão), usando a fórmula $Z = (Y - \mu)/\sigma$.

Para $Y = 90$:

$$Z_1 = \frac{90 - 100}{10} = \frac{-10}{10} = -1$$

Para $Y = 110$:

$$Z_2 = \frac{110 - 100}{10} = \frac{10}{10} = 1$$

A probabilidade desejada é $P(-1 < Z < 1)$. Utilizando a tabela da distribuição Normal Padrão:

$$P(-1 < Z < 1) = P(Z < 1) - P(Z < -1)$$

$$P(-1 < Z < 1) = 0.8413 - 0.1587 = 0.6826$$

Portanto, $P(90 < Y < 110) = 0.6826$, ou 68,26%. Este é o resultado esperado pela Regra Empírica para um desvio padrão em torno da média.

b) Agora, estamos interessados na distribuição da **média amostral**, $\bar{Y}$, para amostras de tamanho $n = 16$.

Pelo Teorema Central do Limite (e como a população já é Normal), a distribuição de $\bar{Y}$ também é Normal. A média da distribuição amostral é a mesma da população ($\mu_{\bar{Y}} = \mu = 100$), mas o desvio padrão, chamado de **erro padrão**, é menor.

O erro padrão é calculado como:

$$\sigma_{\bar{Y}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{16}} = \frac{10}{4} = 2.5$$

Então, a distribuição da média amostral é $\bar{Y} \sim N(100, 2.5^2)$.

Agora, padronizamos os valores 90 e 110 usando os parâmetros desta nova distribuição:

Para $\bar{Y} = 90$:

$$Z_1 = \frac{90 - 100}{2.5} = \frac{-10}{2.5} = -4$$

Para $\bar{Y} = 110$:

$$Z_2 = \frac{110 - 100}{2.5} = \frac{10}{2.5} = 4$$

A probabilidade desejada é $P(-4 < Z < 4)$. Este intervalo contém quase toda a área da curva Normal.

$$P(-4 < Z < 4) = P(Z < 4) - P(Z < -4)$$

$$P(-4 < Z < 4) \approx 0.999968 - 0.000032 = 0.999936$$

Portanto, $P(90 < \bar{Y} < 110) \approx 0.9999$, ou 99,99%.

Comparação: Note como a probabilidade aumentou drasticamente da parte (a) para a (b). Isso ocorre porque a distribuição das médias amostrais é muito mais "estreita"(tem menor variabilidade) do que a distribuição dos dados individuais. É muito mais provável que a média de 16 observações esteja perto da média populacional do que uma única observação aleatória.

Questão 6

Utilizando algum recurso computacional ou tabela, calcule as probabilidades a seguir, conforme a distribuição da v.a. Y:

Solução:

Distribuição t de Student ($Y \sim t_{20}$)

- $P(-2.85 \leq Y \leq 2.85) = P(Y \leq 2.85) - P(Y \leq -2.85) = 0.9950 - 0.0050 = \mathbf{0.990}$
- $P(Y < -2.85) = \mathbf{0.0050}$
- $P(Y > 2.85) = 1 - P(Y \leq 2.85) = 1 - 0.9950 = \mathbf{0.0050}$
- $P(Y > 2.12) = 1 - P(Y \leq 2.12) = 1 - 0.9774 = \mathbf{0.0226}$
- $P(Y < -3.01) = \mathbf{0.0034}$

Distribuição Qui-Quadrado ($Y \sim \chi^2_{16}$)

- $P(8.91 < Y < 32.85) = P(Y < 32.85) - P(Y < 8.91) = 0.9925 - 0.0916 = \mathbf{0.9009}$
- $P(Y > 8.91) = 1 - P(Y \leq 8.91) = 1 - 0.0916 = \mathbf{0.9084}$
- $P(Y > 32.85) = 1 - P(Y \leq 32.85) = 1 - 0.9925 = \mathbf{0.0075}$
- $P(Y > 22.80) = 1 - P(Y \leq 22.80) = 1 - 0.8756 = \mathbf{0.1244}$
- $P(Y < 10.12) = \mathbf{0.1558}$

Distribuição F de Snedecor ($Y \sim F_{(10,7)}$)

- $P(Y > 3.18) = 1 - P(Y \leq 3.18) = 1 - 0.9419 = \mathbf{0.0581}$
- $P(Y > 0.15) = 1 - P(Y \leq 0.15) = 1 - 0.0016 = \mathbf{0.9984}$
- $P(Y > 5.35) = 1 - P(Y \leq 5.35) = 1 - 0.9928 = \mathbf{0.0072}$
- $P(Y < 7.41) = \mathbf{0.9986}$
- $P(Y < 1) = \mathbf{0.4815}$

Questão 7

Para cada uma das 3 distribuições propostas no exercício 6, encontre o valor de y tal que:

a) $P(Y < y) = 0.90$

b) $P(Y < y) = 0.025$

c) $P(Y < y) = 0.01$

d) $P(Y > y) = 0.975$

Solução: Neste exercício, faremos o processo inverso (cálculo de quantis). Para o item (d), note que $P(Y > y) = 0.975$ é o mesmo que $P(Y < y) = 1 - 0.975 = 0.025$. Portanto, a resposta do item (d) é a mesma do item (b) para todas as distribuições.

Distribuição t de Student ($Y \sim t_{20}$)

a) $P(Y < y) = 0.90 \implies y = \mathbf{1.325}$

b) $P(Y < y) = 0.025 \implies y = -\mathbf{2.086}$

c) $P(Y < y) = 0.01 \implies y = -\mathbf{2.528}$

d) $P(Y > y) = 0.975 \implies P(Y < y) = 0.025 \implies y = -\mathbf{2.086}$

Distribuição Qui-Quadrado ($Y \sim \chi^2_{16}$)

a) $P(Y < y) = 0.90 \implies y = \mathbf{23.542}$

b) $P(Y < y) = 0.025 \implies y = \mathbf{6.908}$

c) $P(Y < y) = 0.01 \implies y = \mathbf{5.812}$

d) $P(Y > y) = 0.975 \implies P(Y < y) = 0.025 \implies y = \mathbf{6.908}$

Distribuição F de Snedecor ($Y \sim F_{(10,7)}$)

a) $P(Y < y) = 0.90 \implies y = \mathbf{2.624}$

b) $P(Y < y) = 0.025 \implies y = \mathbf{0.306}$

c) $P(Y < y) = 0.01 \implies y = \mathbf{0.244}$

d) $P(Y > y) = 0.975 \implies P(Y < y) = 0.025 \implies y = \mathbf{0.306}$

Questão 8

A máquina de empacotar um determinado produto o faz segundo uma distribuição normal, com média μ e desvio padrão 10 g.

a) *Em quanto deve ser regulado o peso médio μ para que apenas 10% dos pacotes tenham menos do que 500 g?*

b) *Com a máquina assim regulada, qual a probabilidade de que o peso total de 4 pacotes escolhidos ao acaso seja inferior a 2 kg?*

Solução: Seja Y o peso de um pacote. A distribuição é Normal, com desvio padrão conhecido $\sigma = 10$ g. Escrevemos $Y \sim N(\mu, 10^2)$.

a) Queremos encontrar o valor da média μ tal que a probabilidade de um pacote ter menos de 500 g seja de 10%. Matematicamente:

$$P(Y < 500) = 0.10$$

Para resolver, primeiro padronizamos a variável, transformando-a em uma Normal Padrão (Z):

$$P\left(\frac{Y-\mu}{\sigma} < \frac{500-\mu}{10}\right) = 0.10$$
$$P\left(Z < \frac{500-\mu}{10}\right) = 0.10$$

Agora, precisamos encontrar o valor de Z na tabela da Normal Padrão que acumula 10% (ou 0.10) de probabilidade à sua esquerda. Este valor é o quantil $z_{0.10}$. Consultando a tabela (ou usando um recurso computacional), encontramos que $z_{0.10} \approx -1.2816$.

Com isso, podemos igualar as expressões e resolver para μ:

$$\frac{500-\mu}{10} = -1.2816$$
$$500-\mu = 10 \times (-1.2816)$$
$$500-\mu = -12.816$$
$$\mu = 500 + 12.816 = 512.816$$

Portanto, o peso médio μ deve ser regulado em aproximadamente **512,82 g**.

b) Usando a média regulada em $\mu = 512.82$ g, queremos a probabilidade de que o **peso total** de 4 pacotes ($n = 4$) seja inferior a 2 kg.

Primeiro, garantimos a consistência das unidades: 2 kg = 2000 g.

A pergunta sobre o peso total (S_4) pode ser convertida em uma pergunta sobre a média amostral ($\bar{Y}$):

$$P(S_4 < 2000) \implies P(4 \cdot \bar{Y} < 2000) \implies P(\bar{Y} < \frac{2000}{4}) \implies P(\bar{Y} < 500)$$

Precisamos da distribuição da média amostral $\bar{Y}$. Como a população é Normal, $\bar{Y}$ também segue uma distribuição Normal, com os seguintes parâmetros:

- Média: $\mu_{\bar{Y}} = \mu = 512.82$ g
- Erro padrão: $\sigma_{\bar{Y}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{4}} = \frac{10}{2} = 5$ g

Assim, $\bar{Y} \sim N(512.82, 5^2)$. Agora, padronizamos o valor 500 para encontrar a probabilidade:

$$Z = \frac{500-\mu_{\bar{Y}}}{\sigma_{\bar{Y}}} = \frac{500-512.82}{5} = \frac{-12.82}{5} = -2.564$$

A probabilidade desejada é $P(Z < -2.564)$. Consultando a tabela da Normal Padrão:

$$P(Z < -2.564) \approx 0.0052$$

A probabilidade de que o peso total de 4 pacotes seja inferior a 2 kg é de **0,52%**.

Questão 9

Um estudo que investiga a relação entre idade e despesas médicas anuais amostra aleatoriamente 100 indivíduos em uma cidade da Califórnia. Espera-se que a amostra tenha uma média de idade semelhante à de toda a população.

a) *Se o desvio padrão das idades de todos os indivíduos em Davis for $\sigma = 15$, encontre a probabilidade de que a idade média dos indivíduos da amostra esteja dentro de dois anos da idade média de todos os indivíduos na cidade. (Dica: encontre a distribuição amostral da idade média da amostra e use o teorema do limite central. Você não precisa saber a média da população para responder, mas se isso facilitar, use um valor como $\mu = 30$.)*

b) *A probabilidade seria maior ou menor se $\sigma = 10$? Por quê?*

Solução: Seja $\bar{Y}$ a idade média da amostra e μ a idade média da população. Temos uma amostra de tamanho $n = 100$.

a) Queremos encontrar a probabilidade de que a idade média da amostra ($\bar{Y}$) esteja "dentro de dois anos"da média da população (μ). Matematicamente, isso significa que a diferença absoluta entre elas é menor ou igual a 2:

$$P(|\bar{Y} - \mu| \leq 2)$$

Isso é equivalente a $P(-2 \leq \bar{Y} - \mu \leq 2)$.

Como o tamanho da amostra $n = 100$ é grande, pelo **Teorema Central do Limite (TCL)**, a distribuição da média amostral $\bar{Y}$ é aproximadamente Normal. Os parâmetros dessa distribuição são:

- Média: $\mu_{\bar{Y}} = \mu$ (a própria média da população)
- Erro padrão: $\sigma_{\bar{Y}} = \frac{\sigma}{\sqrt{n}}$

Com $\sigma = 15$, calculamos o erro padrão:

$$\sigma_{\bar{Y}} = \frac{15}{\sqrt{100}} = \frac{15}{10} = 1.5 \text{ anos}$$

Para encontrar a probabilidade, padronizamos a inequação dividindo pelo erro padrão. Note que não precisamos saber o valor de μ:

$$P\left(\frac{-2}{1.5} \leq \frac{\bar{Y} - \mu}{1.5} \leq \frac{2}{1.5}\right)$$

O termo central, $\frac{\bar{Y}-\mu}{\sigma_{\bar{Y}}}$, é a nossa variável Z padronizada.

$$P(-1.33 \leq Z \leq 1.33)$$

Usando a tabela da Normal Padrão:

$$P(Z \leq 1.33) - P(Z \leq -1.33) = 0.9082 - 0.0918 = 0.8164$$

A probabilidade de que a idade média da amostra esteja dentro de 2 anos da média da população é de **81,64%**.

b) A probabilidade seria **maior**.

Cálculo: Se o desvio padrão da população fosse menor, $\sigma = 10$, o erro padrão também seria menor:

$$\sigma_{\bar{Y}} = \frac{10}{\sqrt{100}} = \frac{10}{10} = 1.0 \text{ ano}$$

A nova probabilidade seria:

$$P\left(\frac{-2}{1.0} \leq Z \leq \frac{2}{1.0}\right) = P(-2 \leq Z \leq 2)$$

$$P(Z \leq 2) - P(Z \leq -2) = 0.9772 - 0.0228 = 0.9544$$

A nova probabilidade seria de **95,44%**, que é maior que a anterior.

Por quê? Um desvio padrão populacional (σ) menor significa que os dados da população são menos dispersos e estão, em geral, mais próximos da média μ. Consequentemente, o erro padrão ($\sigma_{\bar{Y}}$) também diminui, o que torna a distribuição das médias amostrais mais "estreita"e concentrada em torno de μ. Com uma distribuição mais concentrada, a probabilidade de uma média amostral cair em um intervalo fixo (como ± 2 anos) ao redor do centro aumenta.

Questão 10

O teste de conhecimentos gerais chamado Graduate Record Examination (GRE) tem componentes que medem o raciocínio verbal e o raciocínio quantitativo. O exame verbal e o exame quantitativo têm cada um uma pontuação mínima de 200 e máxima de 800. Nos últimos anos, a pontuação total nos dois exames teve aproximadamente uma distribuição normal com uma média de cerca de 1050 e desvio padrão de cerca de 200.

a) *Qual a probabilidade de obter pontuação total (i) abaixo de 1200 e (ii) acima de 1200?*

b) *Dos participantes do teste GRE que pontuaram acima de 1.200, qual proporção deles teve pontuação acima de 1.400?*

c) *Um grupo de 25 alunos formou um grupo de estudos para se preparar para o GRE. Para eles, a média de suas 25 pontuações totais é 1200. Se eles fossem uma amostra aleatória dos alunos que estão fazendo o exame, explique por que isso teria sido um resultado muito incomum.*

Solução: Seja Y a pontuação total no GRE. A distribuição é Normal, com média $\mu = 1050$ e desvio padrão $\sigma = 200$. Logo, $Y \sim N(1050, 200^2)$.

a) Para encontrar as probabilidades, primeiro padronizamos o valor de 1200:

$$Z = \frac{Y - \mu}{\sigma} = \frac{1200 - 1050}{200} = \frac{150}{200} = 0.75$$

(i) A probabilidade de obter pontuação abaixo de 1200 é $P(Y < 1200) = P(Z < 0.75)$. Consultando a tabela da Normal Padrão, encontramos **0.7734**.

(ii) A probabilidade de obter pontuação acima de 1200 é $P(Y > 1200) = P(Z > 0.75)$. Isso é igual a $1 - P(Z \leq 0.75) = 1 - 0.7734 = \mathbf{0.2266}$.

b) Esta é uma questão de probabilidade condicional. Queremos encontrar $P(Y > 1400 \mid Y > 1200)$. A fórmula é:

$$P(Y > 1400 \mid Y > 1200) = \frac{P(Y > 1400 \text{ e } Y > 1200)}{P(Y > 1200)} = \frac{P(Y > 1400)}{P(Y > 1200)}$$

Já temos o denominador do item (a): $P(Y > 1200) \approx 0.2266$.

Agora, calculamos o numerador, $P(Y > 1400)$. Primeiro, padronizamos 1400:

$$Z = \frac{1400 - 1050}{200} = \frac{350}{200} = 1.75$$

Então, $P(Y > 1400) = P(Z > 1.75) = 1 - P(Z \leq 1.75) = 1 - 0.9599 = 0.0401$.

Finalmente, calculamos a proporção:

$$\frac{0.0401}{0.2266} \approx 0.1769$$

Portanto, aproximadamente **17,69%** dos participantes que pontuaram acima de 1200 também tiveram pontuação acima de 1400.

c) Para avaliar se uma média amostral de $\bar{Y} = 1200$ para $n = 25$ alunos é "incomum", calculamos a probabilidade de obter uma média amostral **tão alta ou maior** por puro acaso.

Primeiro, definimos a distribuição da média amostral $\bar{Y}$. Como a população é Normal, $\bar{Y}$ também é Normal.

- Média: $\mu_{\bar{Y}} = \mu = 1050$
- Erro padrão: $\sigma_{\bar{Y}} = \frac{\sigma}{\sqrt{n}} = \frac{200}{\sqrt{25}} = \frac{200}{5} = 40$

A distribuição é $\bar{Y} \sim N(1050, 40^2)$. Agora, calculamos $P(\bar{Y} \geq 1200)$. Padronizamos o valor 1200:

$$Z = \frac{\bar{Y} - \mu_{\bar{Y}}}{\sigma_{\bar{Y}}} = \frac{1200 - 1050}{40} = \frac{150}{40} = 3.75$$

A probabilidade é $P(Z \geq 3.75)$, que é $1 - P(Z < 3.75)$.

Um valor Z de 3.75 é extremamente alto. A probabilidade associada a ele é minúscula:

$$P(Z \geq 3.75) \approx 1 - 0.999912 = 0.000088$$

Explicação: A probabilidade de uma amostra aleatória de 25 alunos ter uma nota média de 1200 ou mais é de apenas 0,0088%. Como essa probabilidade é extremamente baixa, o resultado é considerado muito incomum. Isso sugere fortemente que o grupo de estudos não se comporta como uma amostra aleatória da população geral; o desempenho deles é significativamente superior, seja por serem alunos naturalmente mais aptos ou pelo efeito positivo do grupo de estudos.