Recurso – Questão 46

Uma comissão da Câmara dos Deputados está analisando a distribuição de salários de determinada categoria funcional no Brasil. Durante um primeiro estudo promovido pelos deputados, um pesquisador calculou as médias, os desvios padrões e padronizou os valores dos salários brasileiros, para uma amostra de profissionais, com o intuito de comparar o desempenho salarial entre diferentes países. Esse pesquisador observou que, no Brasil, o salário médio dessa categoria é de R$ 6.000 com desvio padrão de R$ 800, enquanto, em outro país — P —, usado na comparação, o valor tabelado, em dólar, do salário médio dessa mesma categoria é $ 3.500 e o desvio padrão é $ 250.

Com base nessa situação hipotética, julgue os itens seguintes.

46. Se o referido pesquisador tivesse escolhido uma nova amostra de trabalhadores 9 vezes maior que a do primeiro estudo, o erro padrão seria reduzido a um terço, caso fossem considerados os mesmos valores para o desvio padrão e a média do estudo inicial.

RECURSO

Para resolver a questão acima é necessário conhecer a fórmula do **ERRO PADRÃO**:

$$erro\ padrão = \frac{desvio\ padrão}{\sqrt{n}}$$

A presente questão deve ser anulada, uma vez que aborda conteúdo não previsto no edital do certame.

Conforme o conteúdo programático estabelecido no edital, para a disciplina de *RACIOCÍNIO LÓGICO E NOÇÕES DE ESTATÍSTICA*, estão previstos os seguintes tópicos:

"RACIOCÍNIO LÓGICO E NOÇÕES DE ESTATÍSTICA: 1 Estruturas lógicas. 2 Lógica de argumentação: analogias, inferências, deduções e conclusões. 3 Lógica sentencial (ou proposicional). 3.1 Proposições simples e compostas. 3.2 Tabelas verdade. 3.3 Equivalências. 3.4 Leis de Morgan. 3.5 Diagramas lógicos. 4 Lógica de primeira ordem. 5 Princípios de contagem e probabilidade. 6 Operações com conjuntos. 7 Raciocínio lógico envolvendo problemas aritméticos, geométricos e matriciais. 8 Estatística e probabilidade básicas. 8.1 Medidas de tendência central. 8.2 Medidas de dispersão. 8.3 Formulação e validação de hipóteses. 8.4 Amostragem, viés e erro. 8.5 Normalização e padronização. 8.6 Outliers."

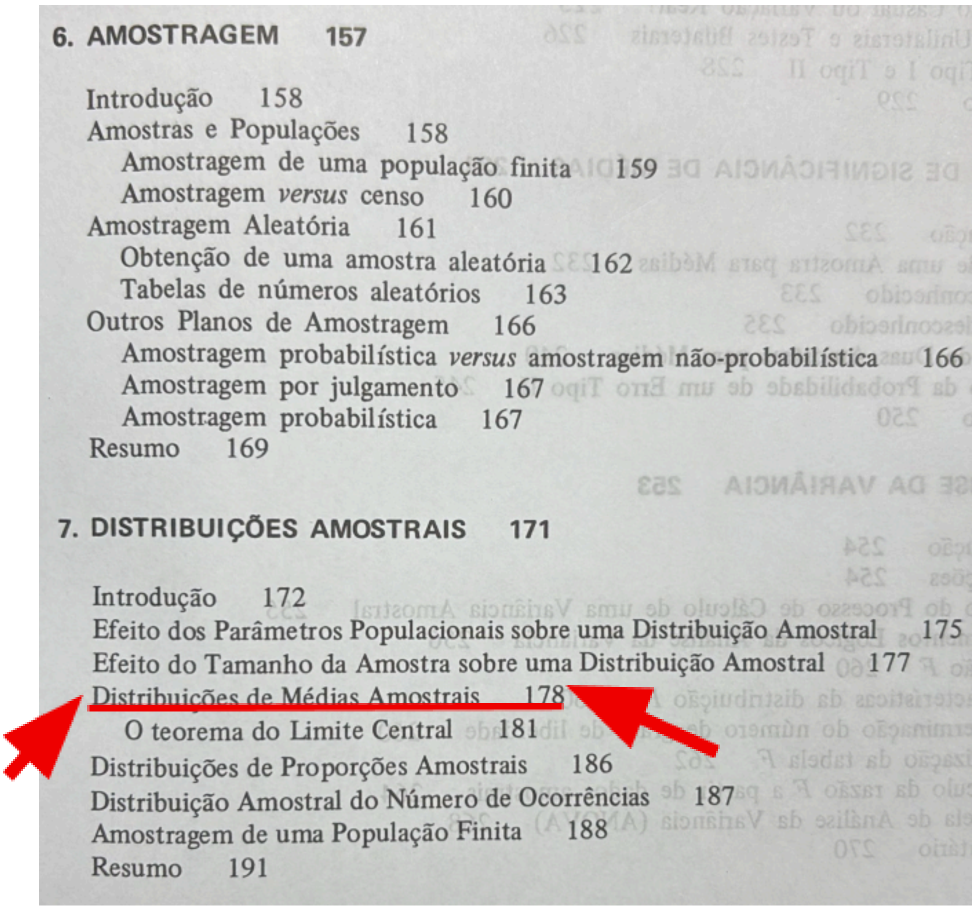
A questão em análise exige conhecimento acerca do assunto ERRO PADRÃO, o qual não se encontra entre os conteúdos expressamente previstos no edital, tampouco pode ser considerado como desdobramento direto dos tópicos listados.

O conteúdo cobrado na questão, **ERRO PADRÃO**, é associado aos seguintes temas da disciplina Estatística: **DISTRIBUIÇÕES AMOSTRAIS ou ESTIMAÇÃO DE PARÂMETROS ou INTERVALO DE CONFIANÇA**. Nenhum desses temas que normalmente aparecem em conteúdos programáticos de ESTATÍSTICA em concursos da banca CEBRASPE está no conteúdo programático do concurso elaborado pela CEBRASPE para CÂMARA DOS DEPUTADOS cargo POLICIAL LEGISLATIVO FEDERAL.

Apresentando duas bibliografias para comprovar que o tema (**ERRO PADRÂO**) **NÃO FAZ PARTE** de conteúdo presente na ementa do concurso:

LIVRO 1
ESTATÍSTICA APLICADA À ADMINISTRAÇÃO
AUTOR: WILLIAN J. STEVENSON
EDITORA HARBRA

Foto do sumário (tema 7 – página 178 até 181)



Página 181 faz parte do tema DISTRIBUIÇÕES DE MÉDIAS AMOSTRAIS

A foto desta página exibe a fórmula do erro padrão, também chamado de desvio padrão da distribuição amostral, necessária para a resolução da questão.

Quando a população é muito grande ou infinita, o desvio padrão da distribuição amostral de média é

$$\boxed{\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}}$$

onde

$\sigma_{\bar{x}}$ = desvio padrão da distribuição amostral
σ_x = desvio padrão da população
n = tamanho da amostra

A fórmula do desvio padrão nos diz, com efeito, que a quantidade de dispersão na distribuição amostral depende de duas coisas:

1. A dispersão na população.
2. A raiz quadrada do tamanho da amostra.

Por exemplo, em qualquer população, o aumento do tamanho das amostras extraídas resultará em menor variabilidade entre as possíveis médias amostrais. E se o mesmo tamanho de amostra é usado com diferentes populações, as populações com maior quantidade de dispersão (σ_x) tenderão a gerar a maior quantidade de variabilidade entre as médias de amostras extraídas delas.

O Teorema do Limite Central

A capacidade de usar amostras para fazer inferências sobre parâmetros populacionais depende do conhecimento da distribuição amostral. Acabamos de ver como se determinam a média e o desvio padrão, mas precisamos ainda de outra informação: a forma da distribuição amostral. Já dissemos antes que há uma tendência para as distribuições de médias e de proporções se apresentarem aproximadamente normais. No caso das médias amostrais, pode-se demonstrar matematicamente que, se uma população tem distribuição normal, a distribuição das médias amostrais extraídas da população também tem distribuição normal, para qualquer tamanho de amostra.

LIVRO 2

ESTATÍSTICA APLICADA

AUTOR: JOHN E. FREUND & GARY A. SIMON

EDITORA BOOKMAN

Item 10.6 – DISTRIBUIÇÕES AMOSTRAIS – página 189

10.6 DISTRIBUIÇÕES AMOSTRAIS

Introduziremos agora o conceito de **distribuição amostral** de uma estatística, provavelmente o conceito mais fundamental da inferência estatística. Como veremos, este conceito está estreitamente relacionado com a idéia de variação aleatória, ou flutuações aleatórias, já mencionada anteriormente para enfatizar a necessidade de medir a variabilidade dos dados. Neste capítulo, vamos concentrar-nos principalmente na média amostral e sua distribuição amostral, mas, em alguns exercícios da página 197 e em capítulos posteriores, vamos considerar também distribuições amostrais de outras estatísticas.

Para ilustrar o conceito de distribuição amostral, vamos construir a da média de uma amostra aleatória de tamanho $n = 2$ extraída, sem reposição, de uma população finita de tamanho $N = 5$, cujos elementos são os números 3, 5, 7, 9 e 11. A média desta população é

$$\mu = \frac{3+5+7+9+11}{5} = 7$$

e seu desvio-padrão é

$$\sigma = \sqrt{\frac{(3-7)^2+(5-7)^2+(7-7)^2+(9-7)^2+(11-7)^2}{5}} = \sqrt{8}$$

Se tomamos, agora, uma amostra aleatória de tamanho $n = 2$ desta população, há $\binom{5}{2} = 10$ possibilidades

3 e 5	3 e 7	3 e 9	3 e 11	5 e 7
5 e 9	5 e 11	7 e 9	7 e 11	9 e 11

e suas médias são 4, 5, 6, 7, 6, 7, 8, 8, 9 e 10. Como cada amostra tem probabilidade $\frac{1}{10}$, obtemos a seguinte distribuição amostral da média de uma amostra aleatória de tamanho $n = 2$ da população dada:

$\bar{x}$	Probabilidade
4	$\frac{1}{10}$
5	$\frac{1}{10}$
6	$\frac{2}{10}$
7	$\frac{2}{10}$
8	$\frac{2}{10}$
9	$\frac{1}{10}$
10	$\frac{1}{10}$

A Figura 10.2 mostra um histograma desta distribuição.

Um exame desta distribuição amostral revela algumas informações apropriadas relativas ao problema de estimar a média da população dada com base em uma amostra aleatória de tamanho 2. Por exemplo, vemos que, para $\bar{x} = 6$, 7 ou 8, há uma probabilidade de $\frac{6}{10}$ de uma média amostral não diferir por mais de 1 da média populacional $\mu = 7$. Também para ...

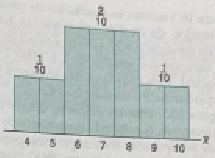


FIGURA 10.2
Distribuição amostral da média.

$\frac{8}{10}$ de uma média amostral não diferir por mais de 2 da média populacional $\mu = 7$. Assim, se não conhecêssemos a média da população dada e quiséssemos estimá-la com a média de uma amostra aleatória de tamanho $n = 2$, o processo acima nos daria uma idéia do tamanho possível do nosso erro.

Podem-se obter outras informações úteis sobre a distribuição amostral da média calculando-se sua média e seu desvio-padrão μ, onde os índices servem para distinguir entre esses parâmetros e os da população original. Com as definições de média e variância de uma distribuição de probabilidades das páginas 150 e 151, obtemos

$$\mu_{\bar{x}} = 4 \cdot \tfrac{1}{10} + 5 \cdot \tfrac{1}{10} + 6 \cdot \tfrac{2}{10} + 7 \cdot \tfrac{2}{10} + 8 \cdot \tfrac{2}{10} + 9 \cdot \tfrac{1}{10} + 10 \cdot \tfrac{1}{10} = 7$$

e

$$\sigma^2_{\bar{x}} = (4-7)^2 \cdot \tfrac{1}{10} + (5-7)^2 \cdot \tfrac{1}{10} + (6-7)^2 \cdot \tfrac{2}{10} + (7-7)^2 \cdot \tfrac{2}{10} + (8-7)^2 \cdot \tfrac{2}{10} + (9-7)^2 \cdot \tfrac{1}{10} + (10-7)^2 \cdot \tfrac{1}{10} = 3$$

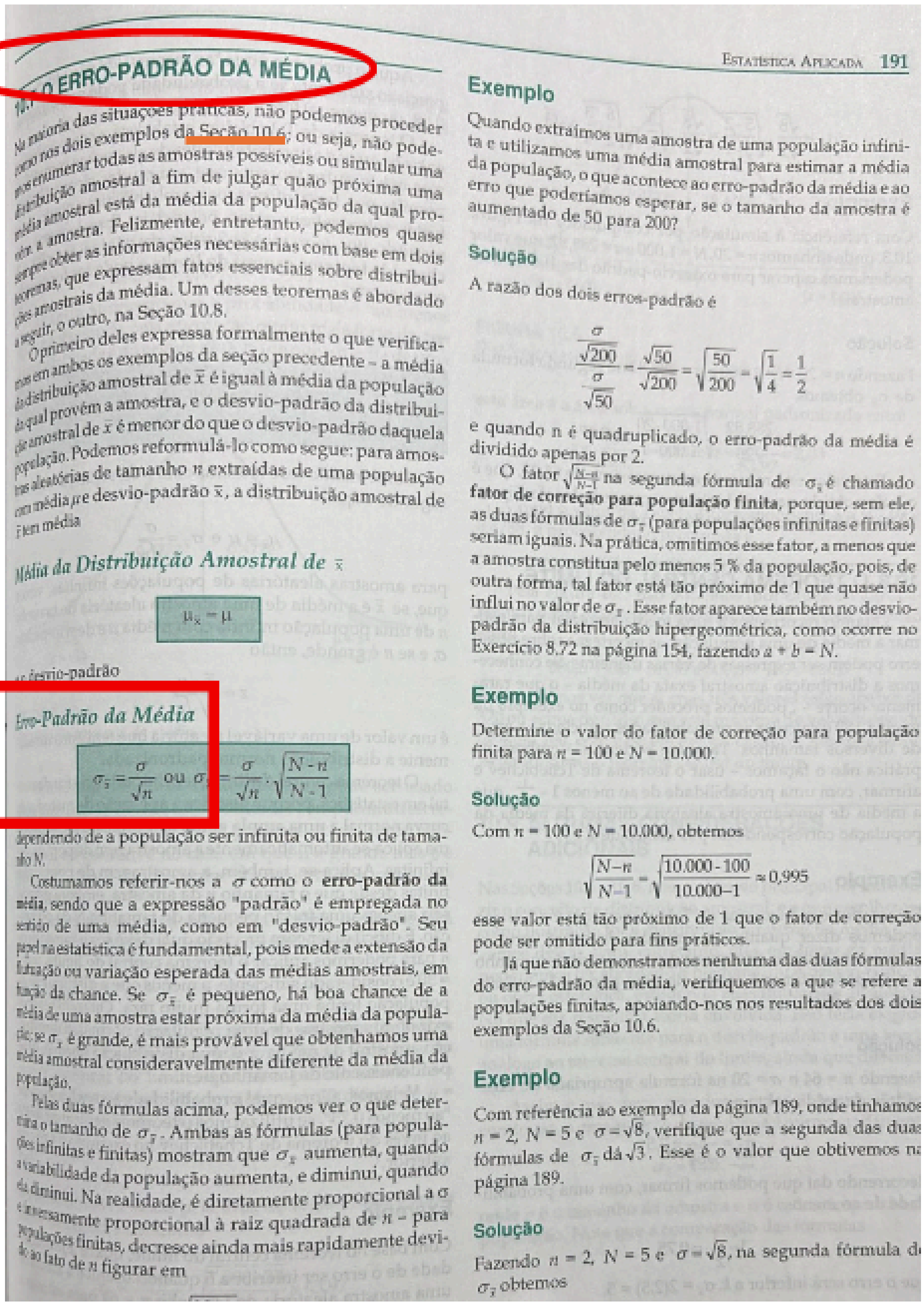
de forma que $\sigma_{\bar{x}} = \sqrt{3}$. Observe que, ao menos para este exemplo, $\mu_{\bar{x}}$ é igual a μ e é menor do que σ. Estas relações têm importância fundamental em estatísticas, e a elas voltaremos na Seção 10.7.

Para dar um exemplo de uma distribuição amostral, tomamos uma amostra muito pequena, de tamanho $n = 2$, de uma população muito pequena, de tamanho $N = 5$, mas seria difícil aplicar o mesmo método para construir a distribuição amostral da média de uma grande amostra de uma grande população. Teríamos de enumerar um número excessivamente grande de possibilidades. Por exemplo, para $n = 10$ e $N = 100$, teríamos de relacionar mais de 17 triliões de amostras.

Assim, para termos alguma idéia da distribuição amostral da média de uma amostra um tanto maior de uma grande população finita, lançaremos mão da **simulação por computador**. Isto é, deixaremos para o computador a tarefa de extrair repetidas amostras de uma dada população, determinar suas médias e descrever de várias maneiras a distribuição dessas médias. Isso nos dará alguma idéia sobre a forma global e sobre algumas das características-chave da real distribuição amostral da média para amostras aleatórias extraídas da população em estudo.

Item 10.7 – ERRO PADRÂO DA MÉDIA – página 191

Repare que o autor se refere ao item 10.6 para ensinar ERRO PADRÃO.



10.7 O ERRO-PADRÃO DA MÉDIA

Na maioria das situações práticas, não podemos proceder como nos dois exemplos da Seção 10.6; ou seja, não podemos enumerar todas as amostras possíveis ou simular uma distribuição amostral a fim de julgar quão próxima uma média amostral está da média da população da qual provém a amostra. Felizmente, entretanto, podemos quase sempre obter as informações necessárias com base em dois teoremas, que expressam fatos essenciais sobre distribuições amostrais da média. Um desses teoremas é abordado a seguir, o outro, na Seção 10.8.

O primeiro deles expressa formalmente o que verificamos em ambos os exemplos da seção precedente – a média da distribuição amostral de $\bar{x}$ é igual à média da população da qual provém a amostra, e o desvio-padrão da distribuição amostral de $\bar{x}$ é menor do que o desvio-padrão daquela população. Podemos reformulá-lo como segue: para amostras aleatórias de tamanho n extraídas de uma população com média μ e desvio-padrão $\bar{x}$, a distribuição amostral de $\bar{x}$ tem média

Média da Distribuição Amostral de $\bar{x}$

$$\mu_{\bar{x}} = \mu$$

e desvio-padrão

Erro-Padrão da Média

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad \text{ou} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

dependendo de a população ser infinita ou finita de tamanho N.

Costumamos referir-nos a σ como o **erro-padrão da média**, sendo que a expressão "padrão" é empregada no sentido de uma média, como em "desvio-padrão". Seu papel na estatística é fundamental, pois mede a extensão da flutuação ou variação esperada das médias amostrais, em função da chance. Se $\sigma_{\bar{x}}$ é pequeno, há boa chance de a média de uma amostra estar próxima da média da população; se $\sigma_{\bar{x}}$ é grande, é mais provável que obtenhamos uma média amostral consideravelmente diferente da média da população.

Pelas duas fórmulas acima, podemos ver o que determina o tamanho de $\sigma_{\bar{x}}$. Ambas as fórmulas (para populações infinitas e finitas) mostram que $\sigma_{\bar{x}}$ aumenta, quando a variabilidade da população aumenta, e diminui, quando ela diminui. Na realidade, é diretamente proporcional a σ e inversamente proporcional à raiz quadrada de n – para populações finitas, decresce ainda mais rapidamente devido ao fato de n figurar em

Exemplo

Quando extraímos uma amostra de uma população infinita e utilizamos uma média amostral para estimar a média da população, o que acontece ao erro-padrão da média e ao erro que poderíamos esperar, se o tamanho da amostra é aumentado de 50 para 200?

Solução

A razão dos dois erros-padrão é

$$\frac{\frac{\sigma}{\sqrt{200}}}{\frac{\sigma}{\sqrt{50}}} = \frac{\sqrt{50}}{\sqrt{200}} = \sqrt{\frac{50}{200}} = \sqrt{\frac{1}{4}} = \frac{1}{2}$$

e quando n é quadruplicado, o erro-padrão da média é dividido apenas por 2.

O fator $\sqrt{\frac{N-n}{N-1}}$ na segunda fórmula de $\sigma_{\bar{x}}$ é chamado **fator de correção para população finita**, porque, sem ele, as duas fórmulas de $\sigma_{\bar{x}}$ (para populações infinitas e finitas) seriam iguais. Na prática, omitimos esse fator, a menos que a amostra constitua pelo menos 5 % da população, pois, de outra forma, tal fator está tão próximo de 1 que quase não influi no valor de $\sigma_{\bar{x}}$. Esse fator aparece também no desvio-padrão da distribuição hipergeométrica, como ocorre no Exercício 8.72 na página 154, fazendo $a + b = N$.

Exemplo

Determine o valor do fator de correção para população finita para $n = 100$ e $N = 10.000$.

Solução

Com $n = 100$ e $N = 10.000$, obtemos

$$\sqrt{\frac{N-n}{N-1}} = \sqrt{\frac{10.000-100}{10.000-1}} \approx 0{,}995$$

esse valor está tão próximo de 1 que o fator de correção pode ser omitido para fins práticos.

Já que não demonstramos nenhuma das duas fórmulas do erro-padrão da média, verifiquemos a que se refere a populações finitas, apoiando-nos nos resultados dos dois exemplos da Seção 10.6.

Exemplo

Com referência ao exemplo da página 189, onde tínhamos $n = 2$, $N = 5$ e $\sigma = \sqrt{8}$, verifique que a segunda das duas fórmulas de $\sigma_{\bar{x}}$ dá $\sqrt{3}$. Esse é o valor que obtivemos na página 189.

Solução

Fazendo $n = 2$, $N = 5$ e $\sigma = \sqrt{8}$, na segunda fórmula de $\sigma_{\bar{x}}$ obtemos

Ressalta-se ainda que, em concursos públicos, vigora o princípio da vinculação ao edital, de modo que a cobrança de conteúdo não previsto compromete a isonomia entre os candidatos.

Diante do exposto, solicita-se a anulação da questão.