**Introdução à Probabilidade e Estatística II
Exercícios para revisão e autoteste**“Estatística para Economistas”, Rodolfo Hoffmann
“Probabilidade e Estatística para engenharia e ciências”, Walpole e Myers
“Statics for Business & Economics”, Paul Newbold

**INFERÊNCIA ESTATÍSTICA**

**1.** Mostre que S² é um estimador não viciado do parâmetro σ².

**2.** Suponha que, para analisar a receita das famílias dos empregados de uma indústria, um pesquisador selecionou uma amostra aleatória de empregados cujos nomes constavam das folhas de pagamento das firmas dessa indústria. O procedimento é correto?

**3.** Afirma-se, comumente, que as famílias se tornaram menores, ou seja, que o número de filhos por casal é, atualmente, menor do que no passado. Suponha que, para verificar a afirmativa, um pesquisador selecionou uma amostra aleatória de famílias atuais e lhes perguntou quantos filhos tem ou tiveram, perguntou aos pais quantos irmãos tem ou tiveram e quantos irmãos seus próprios pais tiveram, seus avós e assim por diante, até onde for necessário. Critique o processo desse pesquisador.

**4.** Explique o que é um estimador não-tendencioso (não viesado).

**5.** Quantas amostras diferentes, de 2 elementos, podem ser obtidas de uma população de 5 elementos, sem reposição? E com reposição? Determine em cada caso, a probabilidade de se obter, em uma amostragem aleatória, as diferentes amostras. Indique os elementos da população por A,B,C,D e E.

**6.** Uma amostra de n=4 elementos de uma população infinita forneceu os seguintes valores: 4, 8, 4 e 4. Obtenha estimativas não viesadas da média, da variância e da variância da média de uma amostra com 4 elementos.

**7.** Uma amostra aleatória, obtida sem reposição, de uma população com 25 elementos forneceu os seguintes valores: 8,1,3,6,3,3,7,1 e 4. Estime a média da população e o desvio padrão da média da amostra.

**8.** Suponha que as medidas do comprimento do lado de um quadrado que vale µ geram, devido ao erro de medida, uma variável aleatória X de média µ e desvio padrão σ. Suponhamos ainda que, para estimar a área do quadrado, foram obtidos os valores X1, X2,..., Xn, relativos a n medidas independentes do comprimento do lado do quadrado. Sabemos que a área desse quadrado pode ser estimada ou calculando o quadrado da média das observações ou calculando a média aritmética dos quadrados dos valores observados. Compare o viés desses estimadores.

**9.** Se o desvio padrão do peso de crianças com 10 anos de idade é 2,5kg, qual é a probabilidade de o peso médio de 100 de tais crianças diferir por mais de 0,5kg da média da população de crianças de 10 anos?

**10**. Refazer a questão anterior admitindo que a população analisada tenha apenas 500 crianças.

**11.** Considere uma população com apenas 3 elementos, cujos valores são 4, 10 e 16. Determine a média, a variância (S²) e a amplitude dessa variável na população. A seguir, faça uma tabela com todas as amostras de 2 elementos que podem ser obtidas dessa população, fazendo amostragens sem reposição. Apresente, na tabela, a média da amostra, a amplitude da amostra, s² e s$\overbar{x}$². Verifique se a média da amostra, a amplitude da amostra e s² são estimadores não tendenciosos da média da população, da amplitude da população e de S², respectivamente.

**12.** Considere uma população com apenas 5 elementos, para os quais o valor de X é 1,2,3,5 e 9. Fazendo amostragem sem reposição, quais são as diferentes amostras de 3 observações que podem ser obtidas? Verifique, para essas amostras, que a média da amostra é um estimador não-tendencioso da média da população. Determine a tendenciosidade (ou viés) da mediana da amostra como estimador da mediana da população.

**13.** Quando um processo produtivo esta operando corretamente, a resistência em ohms dos componentes produzidos tem uma distribuição normal com desvio padrão 3,6. Uma amostra aleatória de 4 componentes foi retirada. Qual é a probabilidade da variância da amostra ser amor que 30? *(Statics for Business & Economics, 4ª ed, pg 247)*

**SOLUÇÕES**

1. $\sum\_{i=1}^{n}(Xi-\overline{X}$)² = $\sum\_{i=1}^{n}[\left(Xi-µ\right)-(\overline{X}-µ)]$²
 = $\sum\_{i=1}^{n}(Xi-µ$)² - n($\overline{X}$ - µ)²

Agora,
E[S²] = E[ $\frac{1}{n-1}\sum\_{i=1}^{n}(Xi-\overline{X}$)²]
 =$ \frac{1}{n-1}$ ($\sum\_{i=1}^{n} $σXi² -nσ$\overline{X}$²)

Entretanto,
σXi =σ p/ i=1,2,3,...,n,
e σ$\overline{X}$²= σ²/n.

Portanto,
E(S²)= $\frac{1}{n-1}$ (nσ²-n$ \frac{σ²}{n}$) = σ².

2. O procedimento não produz uma amostra aleatória das famílias das industrias, pois aquelas cujo número de membros empregados é maior tem maior probabilidade de serem selecionadas.

3. A amostragem é tendenciosa em favor das famílias maiores no passado. Assim, as famílias de gerações mais antigas com maior numero de filhos terão maior probabilidade de pertencer à amostra e as famílias de gerações anteriores que não tiveram filhos não poderão estar representadas na amostra.

4. Um estimador *a* qualquer de um parâmetro α é não tendencioso (não viesado, não viciado ou imparcial) se E(*a*)=α.

5. No caso de amostragem sem reposição existem 10 diferentes amostras, cada uma com probabilidade 0,1 de ser selecionada. No caso de amostragem com reposição existem 15 diferentes amostras (amostras que só diferem pela ordem dos elementos são consideradas iguais); a probabilidade de ser selecionada é 0,04 para as 5 amostras com elemento repetido e 0,08 para as demais.

6. $\overbar{X}=5,$ s²=4, s$\overbar{x}$²=1

7. $\overbar{X}=4$, s$\overbar{x}$²=2/3

8. E($\overbar{X}$²)-µ² = $\frac{σ²}{n}$; E($\frac{1}{n}\sum\_{}^{}Xi²$)-µ²=σ²

9. P(|Z|>2)= 0,0456

10. P(|Z|>2,236)= 0,0254

11. µ=10, S²=36 e amplitude α=12

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Amostra | $$\overbar{X}$$ | s² | s$\overbar{x}$² | Amplitude (α) |
| 4 e 10 | 7 | 18 | 3 | 6 |
| 4 e 16 | 10 | 72 | 12 | 12 |
| 10 e 16 | 13 | 18 | 3 | 6 |

Verifica-se que E($\overbar{X}$)=µ, E(s²)=S² e E(*a*)=8<α=12.

12. Na população, temos que µ = 4 e mediana sendo 3. Há 10 diferentes amostras igualmente prováveis. Verifica-se que E($\overbar{X}$) = µ. Indicando por *d* a mediana da amostra, verifica-se que E(*d*)=3,3. Então, o viés de *d* como estimador da mediana da população é 0,3.

13. 0.05 < P(sx²>30) < 0.10