Estimação de Parâmetros Populacionais:

Podemos fazer estimativas **por ponto** ou **por intervalo**.

Estimação por ponto:

Seja X uma característica que na população possui distribuição Normal com média μ e variância $\boldsymbol{\sigma^2}$ (desvio padrão $\boldsymbol{\sigma}$).

Seja X_1, X_2, X_3, ..., X_n uma amostra aleatória de tamanho ***n*** extraída desta população. Os parâmetros μ e $\boldsymbol{\sigma^2}$ podem ser estimados com base na amostra.

Temos uma estimação por ponto quando, a partir de uma amostra, procura-se obter um único valor de um certo parâmetro populacional.

Parâmetro		Estimador
Média Aritmética Populacional	μ	$\overline{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$
Variância Populacional	$\boldsymbol{\sigma^2}$	$S^2_{(n)} = \dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n}$ **ou** $S^2_{(n-1)} = \dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1}$

Estimador:

Parâmetro é o verdadeiro valor de uma característica de interesse, medida na população e, de forma geral, é raramente conhecido. O estimador é a medida estatística que descreve o parâmetro, em termos amostrais.

Consideremos uma amostra (X_1, X_2, X_3, ..., X_n) de uma variável aleatória que descreve uma característica de interesse de uma população e seja θ um parâmetro que se deseja estimar. Um estimador do parâmetro θ é qualquer função das observações de X_1, X_2, X_3, ..., X_n.

Estimativa:

É o valor numérico obtido pelo estimador (ou estatística) em uma amostra.

Vício (tendência ou viés):

É dado por: **viés[T]** = $E[T - \theta] = E[T] - \theta$, ou seja, a diferença entre a média do estimador e o parâmetro que se quer estimar.

Assim: $\boxed{\text{viés}[T] = E[T] - \theta}$.

Erro Quadrático Médio (EQM):

É dado por: **EQM[T]** = $E[T - \theta]^2 = Var[T] + \text{viés}^2[T]$.

Assim: $\boxed{EQM[T] = Var[T] + \text{viés}^2[T]}$.

Ou seja, o EQM nada mais é do que a variância do estimador acrescida do quadrado da sua tendenciosidade (viés).

Propriedades dos Estimadores:

1) Não viciado (não-viesado ou não-tendencioso):

Definição:

Seja T = g (X_1, X_2, X_3, ..., X_n) um estimador de θ.

O estimador T é dito estimador não viciado de θ se E[T] = θ, para todo θ.

Ora, se viés[T] = E[T] – θ então, sendo E[T] = θ para todo θ, **viés[T] = 0**, ou seja, não haverá viés.

2) Eficiência de um estimador:

Seja θ um parâmetro a ser estimado e $\hat{\theta}$ um estimador.

$\hat{\theta}$ será um estimador eficiente de θ se forem satisfeitas as seguintes condições:

a) $\hat{\theta}$ for não viciado (viés = 0);

b) $\text{Var}(\hat{\theta}) \leq \text{Var}(\tilde{\theta})$, onde $\tilde{\theta}$ é qualquer outro estimador não viciado de θ.

3) Consistência de um estimador:

Seja ainda θ um parâmetro a ser estimado e $\hat{\theta}$ um estimador.

Se o limite de probabilidade de $\hat{\theta}$ for igual a θ, ou seja, $p\,lim\,\hat{\theta} = \theta$ e se $\lim_{n\to\infty} \text{EQM}(\hat{\theta}) = 0$, então $\hat{\theta}$ será um estimador consistente de θ.

Estimador para a Média Aritmética:

$\overline{X}$ é um estimador não viciado de μ, ou seja, E[$\overline{X}$] = μ.

Consideremos a Média Aritmética: $\overline{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$;

Temos que: $E[\overline{X}] = \dfrac{1}{n} \cdot \sum_{i=1}^{n} E(X_i) = \mu$.

Logo, $\overline{X}$ é um estimador não viciado de μ.

Estimador para a Variância:

A variância, definida como $S^2_{(n)}$ é um estimador viciado.

Entretanto, se for definido como $S^2_{(n-1)}$ é não viciado, pois $E\left[S^2_{(n)}\right] = \dfrac{n}{n-1} \cdot \sigma^2$ e $E\left[S^2_{(n-1)}\right] = \sigma^2$.

Considerando $S^2_{(n)} = \dfrac{1}{n} \cdot \sum_{i=1}^{n} (X_i - \overline{X})^2$, tem-se que $E\left[S^2\right] = E\left[\dfrac{\sigma^2}{n}\left(\dfrac{nS^2}{\sigma^2}\right)\right] = \dfrac{\sigma^2}{n}(n-1)$, porque $\dfrac{nS^2}{\sigma^2}$ segue uma distribuição Qui-Quadrado com (n – 1) graus de liberdade $\left(\chi^2(n-1)\right)$.

Assim, o estimador não viciado de $\boldsymbol{\sigma^2}$ é: $\dfrac{n \cdot S^2}{n-1}$.

QUESTÕES DE CONCURSOS:

1) [NCE/UFRJ- Tecnologista Junior - IBGE-2002] A amostra a seguir foi obtida:

32 27 35 36 20

A variância amostral, obtida a partir do estimador não viesado usual da variância populacional, é igual a:

(A) 29,3
(B) 33,2
(C) 38,9
(D) 43,5
(E) 52,1

Resolução comentada:

A média amostral será: $\overline{X} = \dfrac{32+27+35+36+20}{5} = \dfrac{150}{5}$ = 30.

Calculando o somatório dos quadrados dos desvios em relação à média, teremos:

$$\sum_{i=1}^{5}\left(X_i - \overline{X}\right)^2 = (32-30)^2 + (27-30)^2 + (35-30)^2 + (36-30)^2 + (20-30)^2 \Rightarrow$$

$$\sum_{i=1}^{5}\left(X_i - \overline{X}\right)^2 = 2^2 + (-3)^2 + 5^2 + 6^2 + (-10)^2 \Rightarrow \sum_{i=1}^{5}\left(X_i - \overline{X}\right)^2 = 4+9+25+36+100 = 174.$$

Como queremos o estimador não viesado, basta dividir o resultado acima por n – 1. Logo:

$$S^2 = \frac{\sum_{i=1}^{5}\left(X_i - \overline{X}\right)^2}{5-1} = \frac{174}{4}$$ = **43,5** (Letra D).

2) [NCE/UFRJ- Tecnologista Junior - IBGE-2002] X_1, X_2, X_3 e X_4 é uma amostra aleatória simples de uma distribuição Poisson com parâmetro 1. Se **$Y = X_1 + X_2 + X_3 + X_4$**.

Então o desvio padrão de Y é igual a:

(A) 2
(B) 3
(C) 4
(D) 5
(E) 6

Resolução comentada:

Lembremos que, na distribuição de Poisson, a esperança e a variância são iguais ao parâmetro λ, no caso igual a 1. Logo, $V(X_1) = V(X_2) = V(X_3) = V(X_4) = 1$.

As variáveis X_1, X_2, X_3 e X_4 serão independentes e basta lembrar de uma das propriedades da variância, para variáveis independentes: "a variância da soma será a soma das variâncias". Portanto, a variância de Y:

$V(Y) = V(X_1) + V(X_2) + V(X_3) + V(X_4)$ = 1 + 1 + 1+ 1 $\Rightarrow$ V(Y) = 4.

Como queremos o desvio padrão, DP(Y) = $\sqrt{V(Y)}$ = $\sqrt{4}$ = **2** (Letra A).

3) [NCE/UFRJ- Tecnologista Junior - IBGE-2002] A média e a variância de um estimador T de um parâmetro θ valem, respectivamente, **E[T] =** θ **– 1** e **Var[T] =** θ**/100**.

O erro quadrático médio de T em relação a θ é então igual a:

(A) –1
(B) 0,99θ
(C) 1 + 0,01θ
(D) 1
(E) 1 – θ

Resolução comentada:

O EQM será dado por (ver página 1): $\boxed{\text{EQM[T]} = \text{Var[T]} + \text{viés}^2\text{[T]}}$.

O viés por: $\boxed{\text{viés[T]} = \text{E[T]} - \theta}$. Mas E[T] = θ – 1.

Logo: viés[T] = θ – 1 – θ $\Rightarrow$ **viés[T] = –1**.

Substituindo, na fórmula do EQM, a variância (dada no enunciado da questão) e o viés:

$\text{EQM[T]} = \dfrac{\theta}{100} + (-1)^2$ $\Rightarrow$ EQM[T] = **1 + 0,01**θ (Letra C).

4) [FCC - Analista em Estatística MPE/PE-2006] Com relação à teoria geral de amostragem, é INCORRETO afirmar:

(A) Quanto menor o erro padrão da estimativa, menor será a confiabilidade e a precisão da estimativa.
(B) Em uma amostra por conglomerados a população é dividida em sub-populações distintas.
(C) A realização de amostragem aleatória simples só é possível se o pesquisador possuir uma lista completa de cada unidade amostral.
(D) Um estimador é considerado não viciado quando sua esperança é igual ao valor populacional que está sendo pesquisado.
(E) Amostragem estratificada consiste na divisão de uma população em grupos segundo alguma característica conhecida. Os estratos da população devem ser mutuamente exclusivos.

Resolução comentada:

Alternativa incorreta na Letra A, pois quanto menor o erro, maior será a precisão. Demais itens corretos.

5) [FCC - Estatístico MPU-2007] Com relação à teoria geral de amostragem, é correto afirmar que:

(A) na amostragem aleatória simples, a seleção das unidades amostrais só pode ser realizada sem reposição.
(B) a amostragem por conglomerados em geral é mais eficiente e menos econômica quando comparada com o método de amostragem aleatória simples.
(C) na amostragem estratificada, os estratos da população não necessitam ser mutuamente exclusivos.
(D) o aumento do tamanho da amostra tem como conseqüência o aumento do erro padrão das estimativas.
(E) o viés ou vício de um estimador de um parâmetro é a diferença entre o seu valor esperado e o valor do parâmetro.

Resolução comentada:

(A) ERRADO. A AAS pode ser com ou sem reposição.
(B) ERRADO. A AAC é mais eficiente e mais econômica do que a AAS
(C) ERRADO. Os estratos deverão ser mutuamente exclusivos.
(D) ERRADO. Se aumentarmos o tamanho da amostra o erro da estimativa será menor.
(E) **CORRETO**. É exatamente a definição de viés ou vício de um estimador.