Gabarito Lista 3

1) Amostra aleatória é o processo em que selecionamos n indivíuos de uma população onde cada indivíduo tem a mesma probabilidade de ser escolhido. Ela é simples pois cada um dos indivíduos dessa população pode ser sorteado mais de uma vez, ou seja, é um sorteio com reposição.

Exemplo: Uma fábrica produz 1000 peças por dia de um certo eletrodoméstico. Um funcionário deseja saber a proporção de peças defeituosas que são fabricadas durante esse período. Para isso ele irá selecionar uma peça da caixa da produção diária de forma aleatória verificando se ela está defeituosa ou não. Após isso o funcionário devolve a peça à caixa e repete o procedimento n vezes.

2) Um estimador não-eviesado é aquele que possui valor esperado igual ao valor do parâmetro populacional. Por exemplo: $E[\overline{x}] = \mu$.

Um estimador será consistente se a sua variância vai para zero quando o tamanho amostral n aumenta (mais formalmente é quando o n tende ao infinito).

3) O μ possui apenas um valor que nós não conhecemos. Caso construirmos inúmeros intervalos de confiança, o μ jamais mudará de valor e sim o os valores obtidos do intervalo. Então se o μ é um valor fixo, não há aleatoriedade desse valor e portanto não podemos dizer que a probabilidade de μ pertencer ao intervalo de confiança é 95 %. Diferentemente do μ, os valores dos intervalos de confiança serão diferentes para cada amostra, logo é correto dizer que há uma confiança de 95 % do intervalo conter o parâmetro populacional.
A diferença é bem sutil e muitas vezes nada intuitívo, e uma outra forma de expressar isso é pensar no seguinte: *a probabilidade do intervalo conter μ mas não a probabilidade de μ estar no intervalo.*

4)

a) O erro tipo I (α) é o valor da probabilidade de rejeitarmos H_0 dado que a hipótese nula é verdadeira.

b) O erro tipo II (β) é o valor da probabilidade de não rejeitarmos H_0 dado que ela é falsa.

c) O poder do teste é valor que corresponde a 1 - β, em palavras é a probabilidade de rejeitarmos H_0 dado que ela falsa.

d) O valor P é a probabilidade de obtermos um valor da estatística do teste mais extremo que o observado na amostra utilizada assumindo que a hipótese nula é verdadeira.

5) Será utilizado um exemplo pafecido com o da última aula de exercícios.

Um grupo de pesquisadores deseja realizar uma pesquisa com o objetivo de estimar a idade média da população do estado de São Paulo. Em uma primeira etapa, escolheram de forma aleatória 1000 pessoas do estado de São Paulo. Cada uma delas colocou em um papel a sua idade. Em termos estatísticos temos:

X_i = idade do i-ésimo indivíduo na amostra.

Amostra: X_1, X_2, X_3, ..., X_{998}, X_{999}, X_{1000}

a) População de estudo: População de São Paulo.

Unidades amostrais: cada indivíduo da população de São Paulo.

b) X_i = idade do i-ésimo indivíduo na amostra

c) Foi utilizada a técnica de amostragem aleatória simples

d) O modelo probabilístico para a variável idade é de uma distribuição normal com média μ e variância σ^2.

e) Um estimador para μ será a média amostral $\hat{\mu}$ da amostra obtida.

f) Independentemente dos valores do intervalo sabemos que de cada 100 intervalos que construirmos, com um nível de confiança de 95 % ele conterá o parâmetro populacional μ.

g) Um possível teste de hipótese para este caso seria:

Hipótese nula H_0: $\hat{\mu} = \mu$
Hipótese alternativa H_1: $\hat{\mu} \neq \mu$