UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS EXATAS
DEPARTAMENTO DE ESTATÍSTICA

APOSTILA DA DISCIPLINA INFERÊNCIA ESTATÍSTICA I

CURSO DE ESTATÍSTICA

Prof. Paulo Ricardo Bittencourt Guimarães

1º SEMETRE DE 2002

PLANO DE ENSINO

Ficha nº 2

Disciplina: Inferência Estatística I **Código: CE209**

Turma : A **Curso : Estatística - 3º período**

Professor responsável: Paulo Ricardo Bittencourt Guimarães **Pré-requisito: CE203+CE205+CM008**

Programa:

I - Conceitos Básicos

1.1. Definição de Estatística, População alvo, Amostra aleatória, estatística e momentos amostrais. 1.2. Amostragem da distribuição normal: distribuições Qui-quadrado, t de Student e F de Snedecor, principais resultados.

II - Suficiência

2.1. Definição de estatística suficiente. 2.2. Teorema da Fatoração. 2.3. Família exponencial uniparamétrica.

III - Propriedades dos Estimadores Pontuais

3.1. Estimação pontual: definição de estimador e estimativa. 3.2. Definição de Consistência, desigualdade de Tchebychev. 3.3. Definição de Erro Médio Quadrático e estimador não-viciado.

IV - Métodos de Estimação

4.1. Método da Máxima Verossimilhança. 4.2. Método dos Momentos. 4.3. Método dos mínimos Quadráticos.

V - Propriedades Ótimas dos Estimadores

5.1. Definição de estatística completa e estatística ótima. 5.2. Teorema de Lehmann-Scheffé: estimador UMVU. 5.3. Invariância na locação e na escala.

Referências Bibliográficas:

1. Mood, A.M., Graybill, F.A., Boes, D.C. (1974). Introduction to the theory of Statistics. 3 ed. New York: McGraw Hill

2. Hogg & Craig – Introduction to Mathematical Statistical.

3. Bickel, P. J., Doksum, K.A. – Mathematical Statistics, Holden Day Inc.

4. Kreyszig, E. (1970). Introductory Mathematical Statistics. John Wiley & Sons.

5. Rohatgi, V.K. (1976). An Introduction to Probability Theory and Mathematical Statistics. John Wiley & Sons

ÍNDICE

I CONCEITOS BÁSICOS

1.1 INTRODUÇÃO

Latim: *INFERENTIA*
Ato de inferir (tirar por conclusão).
Admite-se uma proposição como verdadeira, que não seja conhecida diretamente, através da relação dela com outras proposições já sabidamente verdadeiras.

Casos especiais: Inferência Dedutiva (certa).
Inferência Indutiva (incerta, há um nível de probabilidade envolvido).

Ex.:
Dedutiva: Premissa principal - um dos ângulos de um triângulo retângulo tem sempre 90º.
Premissa secundária - T é um triângulo retângulo.
Conclusão: T tem um ângulo de 90º (particular → geral).

Indutiva: 10^7 sementes são plantadas, deseja-se saber quantas darão flores brancas e quantas vermelhas. Há um nível de probabilidade envolvido para cada flor (cor) (geral → particular).

A ESTATÍSTICA

OBJETIVOS PRINCIPAIS

a) Planejar o processo amostral e experimental;
b) Obter inferências sobre a população;
c) Estabelecer níveis de incerteza envolvidos nessas inferências.

1.2 NOÇÕES DE POPULAÇÃO E AMOSTRA

DEF.1 : POPULAÇÃO ALVO é a totalidade de elementos que estão sob discussão e das quais se deseja informação.

EXEMPLO 1:- Suponha que se tenha armazenado num depósito 10 milhões de sementes de flores das quais sabe-se que produzem flores brancas e vermelhas. Os 10 milhões de sementes é a população e deseja-se a informação: "quantas destes 10 milhões de sementes produzem flores brancas ?".

DEF. 2: AMOSTRA ALEATÓRIA: Sejam as variáveis [X_1, X_2, ... ,X_n] que têm a densidade conjunta $f_{x_1,x_2,...,x_n}(x_1,x_2,...,x_n)$ que fatora nas densidades marginais seguintes:

$$f_{\underline{x}}(\underline{x}) = f_{x_1,x_2,...,x_n}(x_1,x_2,...,x_n) = f_{x_1}(x_1).f_{x_2}(x_2)...f_{x_n}(x_n)$$

com f(.) sendo a densidade comum a todas as X_i. Então (X_1, X_2, ... , X_n) é definida como amostra aleatória (a.a.) de tamanho n da população com densidade f (.).

Nota: Um valor específico da a.a. X_1;...;X_n é denotado por x_1;x_2;...;x_n.

No exemplo das 10^7 sementes, pode-se ter:
flor - branca → valor 0
- vermelha → valor 1

X_i = 0 se branca
1 se vermelha

Assim, uma a.a. (n = 50) pode ser: $x_1 = 0$; $x_2 = 1$; ...; $x_{50} = 0$

EXEMPLO 2:- No exemplo tomado, pode-se associar 1 a flor branca e 0 a flor vermelha, então existe um número associado a cada elemento da população. Se a amostragem de sementes é feita de tal forma que as variáveis X_1, X_2, ... ,X_n são independentes e têm a mesma densidade a amostra é dita aleatória.

DEF. 3: POPULAÇÃO AMOSTRADA: Seja [X_1, X_2, ... ,X_n] a.a. da população com densidade f (.), então esta população é chamada população amostrada.

Resultados com base na a.a. só são válidas para a população amostrada, a menos que a população alvo seja a população amostrada.

EXEMPLO 3: Suponha que um sociólogo deseja entender os hábitos religiosos dos homens com 20 anos de idade em certo país. Ele extrai uma amostra de homens com 20 anos de uma grande cidade para estudar. Neste caso tem-se:

- População alvo: homens com 20 anos do país;
- População amostrada: homens com 20 anos da cidade grande amostrada.

Então ele pode fazer conclusões válidas apenas para os elementos da grande cidade (população amostrada), mas pode usar o seu julgamento pessoal para extrapolar os resultados obtidos para a população alvo, com muita cautela e certas reservas.

EXEMPLO 4:- Um pesquisador agrícola está estudando a produção de certa variedade de trigo em determinado estado. Ele tem a sua disposição 5 fazendas, espalhadas pelo estado, nas quais ele pode plantar trigo e observar a produção. A população amostrada, neste caso, consiste das produções de trigo nestas 5 fazendas, enquanto a população alvo consiste das produções de trigo em todas as fazendas do estado.

1.3 ESTATÍSTICA : MOMENTOS AMOSTRAIS

DEF. 4: ESTATÍSTICA:

Uma estatística é uma função das variáveis observáveis, e é por si própria uma variável observável que não contém qualquer parâmetro desconhecido.

EXEMPLO 5: Seja a a.a. $[X_1, X_2, \ldots, X_n]$. A média amostral $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ é uma estatística.

EXEMPLO 6: Seja a v.a $X \sim N(\mu, \sigma^2)$ com μ e σ^2 desconhecidos, então $x - \mu$ não é estatística porque não é observável, é função do parâmetro desconhecido μ.

EXEMPLO 6: Seja uma v.a. com distribuição $N(\mu;\sigma^2)$. Quais são Estatísticas?

a) $X^2-\mu$

b) $\frac{X}{\sigma^2}$

c) X^2-3

d) $X-4$

e) $X - \log X^3$

DEF.5: MOMENTOS AMOSTRAIS

Seja [$X_1, X_2, \dots, X_n$] uma a.a. da densidade $f_X(x)$. Então, o <u>k-ésimo momento amostral</u> é definido por:

$$M_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$$

OBS 1: Se k=1 $\quad M_1 = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$ (média amostral)

E, o k-ésimo momento amostral em torno de $\overline{X}$ é definido por

$$M'_k = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x})^k$$

OBS 2: Momentos amostrais são exemplos de estatísticas.

RESULTADO 1:

Seja [$X_1, X_2, \dots, X_n$] uma a.a. da população com densidade $f_x(x)$. Então, a esperança do k-ésimo momento amostral é igual ao k-ésimo momento populacional, isto é:

$$E(M_k) = \mu_k = E(x^k)$$

e também,

$$V(M_k) = \frac{1}{n}[E(x^{2k}) - [E(x^k)^2] = \frac{1}{n}[\mu_{2k} - (\mu_k)^2]\text{, se } M_{2k} \text{ existe.}$$

EXERCÍCIO 1 : Prove o resultado anterior.

DEF.6.: VARIÂNCIA AMOSTRAL

Seja [$X_1, X_2, \dots, X_n$] uma a.a. da densidade f_x (x), então $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$ é definida como variância amostral.

OBS: Tanto s^2 quanto $M_2 = \hat{\sigma}^2$ medem a dispersão da amostra, contudo s^2 é melhor que $\hat{\sigma}^2$ porque $E(s^2) = m_2 = \sigma^2$, enquanto $E(M_2) = E(\hat{\sigma}^2) \neq m_2 = \sigma^2$

RESULTADO 2: Seja [$X_1, X_2, \dots, X_n$] uma a.a. da densidade $f_x(x)$, que tem média μ e variância finita σ^2, então :

$$E(\overline{x}) = \mu \qquad \text{e} \qquad V(\overline{x}) = \sigma^2_{\overline{x}} = \frac{\sigma^2}{n}$$

EXERCÍCIO 2: Prove o resultado 2.

RESULTADO 3: Seja [$X_1, X_2, \dots, X_n$] uma a.a. da densidade $f_x(x)$ e seja s^2 a variância amostral, então

$$E(s^2) = \sigma^2 \text{ e } V(s^2) = \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\sigma^4\right)$$

EXERCÍCIO 3: Prove a primeira parte do resultado 3.

1.4 AMOSTRAGEM DA DISTRIBUIÇÃO NORMAL

RESULTADO 4: Seja $\bar{x}$ a média amostral da a.a. $[X_1, X_2, \ldots, X_n]$ da distribuição N (μ,σ^2). Então $\bar{x} \sim$ N $(\mu,\sigma^2/n)$.

Uma v.a. X é normalmente distribuída se sua $f(\bullet)$ é dada por:

$$f_X(x) = f_X(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, \mu \text{ e } X \in R, \sigma > 0$$

REVISÃO: FUNÇÃO GERADORA DE MOMENTOS

$$\psi_X(t) = E(e^{tx})$$

Seja Y = ax + b → $\psi_Y(t) = e^{bt}.\psi_X(at)$

$$\psi_{X,Y}(t_1,t_2) = E(e^{t_1X+t_2Y})$$

Sejam as v.a's X e Y. Elas serão independentes se e somente se:

→ $\psi_{X,Y}(t_1,t_2) = \psi_X(t_1)\psi_Y(t_2)$

EXERCÍCIO 4: Prove o resultado 4.

DEF. 7: DISTRIBUIÇÃO QUI-QUADRADO (χ^2)

Se X é uma v.a com densidade $f_{ax}(x) = \dfrac{1}{\Gamma(\upsilon/2)}\left(\dfrac{1}{2}\right)^{\upsilon/2} x^{\frac{\upsilon}{2}-1} e^{-\frac{1}{2}x}$, x > 0, então X é dita ter distribuição qui-quadrado com ν graus de liberdade, onde ν é um inteiro positivo.

OBS.: Se X ~ χ^2_υ $\Rightarrow$ E(X) = ν e V(X) = 2ν

RESULTADO 5: Se as v.a's X_i, i = 1,2,..., n são normais e independentemente distribuídas com média μ, e variância σ_i^2, então a v.a $U = \sum_{i=1}^{n}\left(\dfrac{x_i - \mu_i}{\sigma_i}\right)^2 \sim x_n^2$.

EXERCÍCIO 5: a) Prove o resultado citado na observação anterior.
b) Prove o resultado 5.

RESULTADO 6 : Se $[z_1, z_2, \ldots, z_n]$ é uma a.a. de uma distribuição N(0,1), então :

(i) $\bar{z} \sim N\ (0, 1/n)$

(II) $\bar{z}$ e $\sum_{i=1}^{n}(z_i - \bar{z})^2$ são independentes

(III) $\sum_{i=1}^{n}(z_i - \bar{z})^2 \sim x_{n-1}^2$

COROLÁRIO : Se $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ é a variância amostral de uma $N(\mu,\sigma^2)$, então $U = \frac{(n-1)s^2}{\sigma^2}$ tem uma distribuição χ_{n-1}^2.

EXERCÍCIO 6: a) Prove o resultado 6.
b) Prove o corolário do resultado 6.

DEF. 8: DISTRIBUIÇÃO $\Im$
Se a v.a X tem a densidade abaixo, então X é definida como uma v.a com distribuição $\Im$ com ν_1 e ν_2 graus de liberdade.

$$f_X(x) = \frac{\Gamma\left(\frac{\upsilon_1+\upsilon_2}{2}\right)}{\Gamma\left(\frac{\upsilon_1}{2}\right)\Gamma\left(\frac{\upsilon_2}{2}\right)}\left(\frac{\upsilon_1}{\upsilon_2}\right)^{\upsilon_1/2}\frac{x^{(\upsilon_1-2)/2}}{\left[1+\left(\upsilon_1/\upsilon_2\right)x\right]^{\frac{\upsilon_1+\upsilon_2}{2}}} \qquad x > 0$$

RESULTADO 7: Seja U uma v.a com distribuição χ^2 com ν_1 graus de liberdade e seja V uma v.a χ^2 com ν_2 graus de liberdade, v.a's independentes. Então a v.a $X = \frac{U/\upsilon_1}{V/\upsilon_2}$ tem distribuição $\Im$ com ν_1 e ν_2 graus de liberdade.

COROLÁRIO: Se $[X_1, X_2, \ldots, X_{m+1}]$ é uma a.a. de tamanho m + 1 de uma população $N(\mu_x, \sigma_x^2)$ e se $[Y_1, Y_2, \ldots, Y_{n+1}]$ é uma a.a. de tamanho n +1 de uma população $N(\mu_y, \sigma_y^2)$ e se as duas amostras são independentes, então :

$$U = \frac{\sum_{i=1}^{m+1}(x_i - \bar{x})^2}{\sigma^2} \sim \chi^2_m, \quad V = \frac{\sum_{i=1}^{n+1}(y_i - \bar{y})^2}{\sigma^2} \sim \chi_n^2 \text{ , tal que } \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2/m}{\sum_{i=1}^{n}(y_i - \bar{y})^2/n} \sim \Im_{m,\,n}.$$

EXERCÍCIO 7: a) Prove o resultado 7
b) Prove o corolário do resultado 7.

RESULTADO 8: Seja X uma v.a $\Im_{\upsilon_1,\upsilon_2}$, então :

$$E(X) = \frac{\upsilon_2}{\upsilon_2 - 2} \quad \upsilon_2 > 2 \quad \text{e} \quad V(X) = \frac{2\upsilon_2^2(\upsilon_1+\upsilon_2-2)}{\upsilon_1(\upsilon_2-2)^2(\upsilon_2-4)} \quad \upsilon_2 > 4$$

EXERCÍCIO 8: Prove o resultado 8.

CARACTERÍSTICAS DA DISTRIBUIÇÃO F

i) Distribuição Assimétrica;
ii) Depende de m e n para sua caracterização;
iii) Há uma relação entre t_k e $F_{(1;k)}$: $t^2_k = F_{(1;k)}$
iv) Esperança e Variância
$E[x] = \frac{n}{n-2}$; (n > 2) e $V[x] = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$; (n > 4)
v) $F_{1-\infty\,(m;n)} = \frac{1}{F\alpha_{(n;m)}}$

APLICAÇÕES
- Análise de Variância (ANOVA)
- Teste de Homogeneidade de Variâncias

DEF.9: DISTRIBUIÇÃO “t” DE STUDENT

A distribuição "*t*" estuda os casos de pequenas amostras (n<30) e/ou quando se desconhece a variância populacional (σ^2). Assim, faz-se necessário estimar σ^2 através da estatística S^2, função da a.a. $X_1;...;X_n$.
O Teorema do Limite Central mostra que a distribuição amostral da média tem distribuição $N(\mu; \sigma^2/n)$, quando n é grande. Logo, a estatística Z

$$Z = \frac{X-\mu}{\sigma/\sqrt{n}} \sim N(0;1)$$

Agora, se usarmos a estatística S^2 e n for pequeno, Z não terá mais distribuição Normal. Qual será será então a nova distribuição?
O estatístico W.S.Gosset ("The Probable Error of Mean" Biometrika, 1.908), sob pseudônimo de Student, obteve uma distribuição exata para nova estatística, denotada por "*t*"de Student:

$$t = \frac{\overline{X}-\mu}{S/\sqrt{n}}$$; S^2 um estimador imparcial de σ^2.

Esta distribuição é definida pela razão entre uma v.a com distribuição Normal Padrão e a raiz quadrada de uma v.a com distribuição Qui-quadrado, ou seja, $X = \frac{Z}{\sqrt{U/\upsilon}}$, com Z e

U independentes. Desta forma X tem distribuição “t” de Student com f.d.p : $f_X(x) = \dfrac{\Gamma[(\upsilon+1)/2]}{\Gamma\left(\dfrac{\upsilon}{2}\right)} \dfrac{1}{\sqrt{\upsilon\pi}} \dfrac{1}{\left(1+\dfrac{x^2}{\upsilon}\right)^{(\upsilon+1)/2}}$

RESULTADO 9: Se Z ~ N(0,1) e U ~ χ^2_ν e são independentes, então X = $\dfrac{Z}{\sqrt{U/\upsilon}}$ ~t_ν.

EXERCÍCIO 9: Prove o resultado 9.

RESULTADO 10: Se X é uma v.a com distribuição t_ν, então:

$$E(X) = 0 \text{ e } V(X) = \frac{\upsilon}{\upsilon - 2}, \ \upsilon > 2.$$

EXERCÍCIO 10: Prove o resultado 10.

Algumas funções geradoras de momento (f.g.m.):

01) Normal
$M_x(t) = e^{\left(\mu . t + ½ \sigma^2 . t^2\right)} \rightarrow N(0;1) \rightarrow M_x(t) = e^{t^2/2} = E(e^{tx})$

02) Qui-quadrado
$M_x(t) = \left(\dfrac{1}{1-2t}\right)^{k/2}, t < ½$

03) *t* de Student
Não existe!!

04) F de Snedecor
Não existe!!

Outros resultados:
ΣBernoulli ~Binomial
ΣPoisson ~Poisson ($\Sigma\ \lambda$)
ΣExponencial ~Gamma (n; λ)

II SUFICIÊNCIA

2.1 INTRODUÇÃO

CONCEITO
Dada uma população com f.p. ou f.d.p pertencente a classe $f_\theta = \{f_X(x \mid \theta) \ ; \ \theta \in \Theta\}$ onde

Θ é o espaço paramétrico, um estimador ou uma estatística, $T = t(X_1,X_2,...,X_n)$, pode ser interpretada como o procedimento destinado a retirar da a.a., $(X_1,X_2,...,X_n)$, informação sobre o valor desconhecido do parâmetro θ. Esta estatística existindo, é chamada de suficiente para θ. Assim uma estatística suficiente permite um resumo das informações trazidas pela amostra, ou seja, resume os dados sem perder nenhuma informação sobre o parâmetro θ. Portanto, conhecida uma estatística suficiente, os dados da amostra passam a ser irrelevantes, pois nada mais dizem sobre θ, ficaram exaustos quando se calculou a estatística T suficiente.

2.2 DEFINIÇÃO: ESTATÍSTICA SUFICIENTE:

Seja $[X_1, X_2, ... ,X_n]$ uma a.a. da população com f.p. ou f.d.p na classe $f_\theta = \{f_X(x \mid \theta); \theta \in \Theta\}$; uma estatística T é suficiente para θ se e somente se a distribuição de $(X_1 , X_2 ,... , X_n)$ condicionada por T=t não depender de θ, para qualquer que seja o valor t do domínio de T.

EXERCÍCIOS :

1) Seja $[X_1, X_2, X_3]$ uma a.a. de uma população de Bernoulli com parâmetro θ.

a) A v.a. X_i é discreta ou contínua? Por quê?

$X_i \sim b(1; \theta) \Rightarrow f_{X_i}(x_i) = \theta^{x_i}(1-\theta)^{1-x_i}$, $I_{(0;1)}(x)$

Discreta, pois seu domínio é um conjunto enumerável finito.

b) Qual o espaço amostral da amostra $X_1; X_2 ;X_3$?

$S = 2^n = 2^3 = 8$

$S = \{(0;0;0); (0;0;1);...;(1;1;1)\}$

c) Qual a distribuição de prob. conjunta de $[X_1; X_2 ;X_3]$?

$$f_{X_1;...;X_n}(\bullet;\theta) = f\left(x_1\right) * ... * f\left(x_n\right) =$$
$$\theta^{x_1}(1-\theta)^{1-x_1}\theta^{x_2}(1-\theta)^{1-x_2}...\theta^{x_n}(1-\theta)^{1-x_n} = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}$$

d) Qual o espaço amostral de θ?

$\theta = \{\theta: 0 < \theta < 1\}$

e) Seja a estatística $T=\Sigma X_i$, representando o n° de sucessos do conjunto de n observações. Qual a distribuição da v.a. T?

$$T=\binom{n}{t}\theta^t(1-\theta)^{n-t}$$

Prove através da f.g.m.
Bern. = $M_x(t) = \theta + (1-\theta)e^t$
Bin.= $M_x(t) = [\theta + (1-\theta)e^t]^n$

f) Qual a prob. de ocorrência do resultado { $x_1;x_2;...;x_n$.} condicionado à T=t?

$$P[\underline{X}/T=t] = P[X_1=x_1;...;X_n=x_n/T=t] = \frac{P[X_1 = x_1,\dots,X_n = x_n / \overbrace{T = t}^{\Sigma x_i}]}{P[T = t]} =$$

$$\frac{\theta^{\sum x_i}(1-\theta)^{n-\sum x_i}}{\binom{n}{t}\theta^t(1-\theta)^{n-t}} = \frac{1}{\binom{n}{t}}$$

g) A Est. T= Σx_i é suficiente para (estimar) θ?

Como $P[\underline{x}/ \Sigma x_i = t] = \dfrac{1}{\binom{n}{t}}$ não depende de θ, então T é uma <u>ESTATÍSTICA SUFICIENTE</u>!!

2) Na situação do exercício anterior, seja $T_0 = X_1 + X_3$.

a) Qual a distribuição de T?

T ~b(2; θ)

b) A estatística T é suficiente?

$$P[\underline{x}/T=t] = \frac{P[X_1 = x_1, X_2 = x_2, X_3 = x_3 / X_1 + X_3 = t]}{P[T = t]} =$$

$$= \frac{\theta^{x_1}(1-\theta)^{1-x_1}\theta^{x_2}(1-\theta)^{1-x_2}\theta^{t-x_1}(1-\theta)^{1-(t-x_1)}}{\binom{2}{t}\theta^t(1-\theta)^{2-t}} =$$

$$= \frac{\theta^{x_2+t}(1-\theta)^{3-(2+t)}}{\binom{2}{t}\theta^t(1-\theta)^{2-t}} = \frac{\theta^{x_2}(1-\theta)^{1-x_2}}{\binom{2}{t}}$$; Portanto T não é suficiente!

3) INSPEÇÃO POR AMOSTRAGEM: Uma máquina produz n itens sucessivamente. Cada item produzido é bom com probabilidade θ e defeituoso com probabilidade 1-θ, onde θ é um parâmetro desconhecido. Suponha que não há dependência entre a qualidade dos itens produzidos e seja $X_i = 1$ se o i-ésimo item é bom e $X_i = 0$ em caso

contrário.

a) Qual a distribuição de probabilidade da v.a X_i ?

b) Qual a distribuição de probabilidade da v.a T = $\sum_{i=1}^{n} X_i$?

c) Determine a distribuição conjunta de $\underline{X}$ ’= [X_1, X_2,..., X_n].

d) Verifique se a estatística T = $\sum_{i=1}^{n} X_i$ é suficiente para θ.

4) No contexto do exercício 1, verifique se a estatística S = $X_1.X_2 + X_3$ é suficiente para estimar θ, seguindo a seqüência de itens seguinte:

a) Qual o contradomínio da v.a S ? Relacione os eventos e o contradomínio de S.

b) Calcule P (X_1=x_1,X_2=x_2,X_3=x_3) para todos os 8 termos [X_1,X_2,X_3]. Construa uma tabela que mostre tudo que é pedido.

c) Calcule P(S=0), P(S=1) e P(S=2), ou seja, determine a distribuição de probabilidade de S, P(S=s). Construa uma tabela que mostre os valores.

d) Calcule a distribuição conjunta de $\underline{X}$ ’ = [X_1,X_2,X_3] condicionada a S = s.

e) Qual a conclusão que se pode tirar sobre S ?

5) Seja [X_1, X_2,..., X_n]uma a.a. de uma população P(θ) onde θ > 0. Mostre diretamente que $\sum_{i=1}^{n} X_i$ é uma estatística suficiente para θ, respondendo os itens :

a) A v.a X_i é discreta ou contínua ? Por quê ?

b) Escreva a f.p. da v.a X_i.

c) Escreva a f.p. conjunta para $\overline{X}$ ’ = [X_1, X_2,..., X_n].

d) Calcule a probabilidade conjunta de $\underline{X}$, condicionada a T = t.

e) O que se conclui sobre T = $\sum_{i=1}^{n} X_i$.

2.3 TEOREMA DE NEYMAN-FISHER OU TEOREMA DA FATORIZAÇÃO

Seja uma a.a. [X_1, X_2, ... ,X_n] de uma distribuição f(x;θ), θ ∈ Θ. A estatística T($\underline{X}$) é suficiente para θ se e somente se existe função g(t, θ), definida para todo t e para todo θ ∈ Θ, e h($\underline{X}$) definida em R^n tal que :

$$P(\underline{X},\theta) = g(T(\underline{X}),\theta).h(\underline{X})$$

Exemplos:

01) X_i ~Bernoulli (θ); amostra = [x_1,x_2,x_3]. T=ΣX é suficiente?

$$f_{\underset{\sim}{X}}\left(\underset{\sim}{x},\theta\right)=\prod_{i=1}^{n=3} f\left(x_1,\theta\right)=\Pi\theta^{x_i}(1-\theta)^{1-x_i}=\theta^{\Sigma x_i}(1-\theta)^{3-\Sigma x_i}=\left(\frac{\theta}{1-\theta}\right)^{\Sigma x_i}.(1-\theta)^3$$

$$\therefore \begin{cases} g(t;\theta)=\left(\frac{\theta}{1-\theta}\right)^{\Sigma x_i}.(1-\theta)^3 \text{depende da amostra apenas atravé s de t} \\ h(x_1;\ldots;x_n)=1 \text{ independe de } \theta\text{, assim T é suficiente} \end{cases}$$

02) $X_i \sim N(\theta;\sigma^2)$, $-\infty < \theta < +\infty$, σ^2 conhecido.

$$f_X(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2}$$

Pelo teorema tem-se:

$$f_{\underset{\sim}{X}}\left(\underset{\sim}{x};\theta\right)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}.e^{-\frac{(X-\theta)^2}{2\sigma^2}}=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n.e^{-\frac{\Sigma(X-\theta)^2}{2\sigma^2}}\overset{*}{=}$$

Mas $\Sigma(X-\theta)^2=\Sigma[(X-\overline{X})-(\theta-\overline{X})]^2=\Sigma[(X-\overline{X})^2-2(X-\overline{X})(\theta-\overline{X})+(\theta-\overline{X})^2]=$

$=\Sigma(X-\overline{X})^2-2\Sigma(X-\overline{X})(\theta-\overline{X})+\Sigma(\theta-\overline{X})^2=\Sigma(X-\overline{X})^2+n(\theta-\overline{X})^2$

$$\overset{*}{=}f_{\underset{\sim}{X}}\left(\underset{\sim}{x};\theta\right)=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n.e^{-\frac{1}{2\sigma^2}[\Sigma(X-\overline{X})^2+n(\theta-\overline{X})^2]}\overset{*}{=}g(t;\theta).h(x_1;\ldots;x_n)=$$

$=$

$\therefore \overline{X}$ é uma estatística suficiente para θ.

EXERCÍCIOS:

1)*PROCESSO DE POISSON*

Suponha que as chegadas de n consumidores em um serviço sejam contadas seguindo um Processo de Poisson com parâmetro de chegada θ. Seja X_i o tempo de espera até a chegada do i-ésimo consumidor.

a) Qual a distribuição do tempo de espera X_i ?
b) Escreva de maneira formal a f.d.p da v.a X_i.
c) Escreva a f.d.p da v.a X_i usando a função indicadora I, que vale 1 quando $X > 0$ e 0 em caso contrário (contradomínio de X_i).

d) Qual a função densidade conjunta da a.a. (X_1, X_2,...,X_n) da população X ~ E (θ) ?

e) Verifique se a estatística T = $\sum_{i=1}^{n} X_i$ é suficiente para θ, usando o Teorema de Neyman-Fisher.

2) Seja (X_1, X_2, ... , X_n) a.a. de uma população com f.p. P (θ).

a) Escreva a f.p. da variável aleatória X.
b) Escreva a f.p. da v.a X, usando a função indicadora I (que descreverá o contradomínio de X).
c) Qual a f.p. conjunta da a.a. [X_1, X_2,..., X_n] ?
d) Determine usando o teorema de Neyman-Fischer a estatística suficiente para θ.

3) Seja [X_1, X_2,..., X_n] uma a.a. de uma distribuição N (0, θ^2), $0 < \theta < \infty$.

a) Escreva a f.d.p da variável aleatória X.
b) Escreva a f.d.p da v.a X usando a variável indicadora I para o seu domínio.
c) Qual a função de densidade conjunta da a.a. [X_1,X_2,...,X_n]?
d) Qual a estatística suficiente para θ ?

4) Seja [X_1, X_2,..., X_n] uma a.a. de tamanho n de uma distribuição geométrica com parâmetro θ, G(θ).

a) Escreva a f.p. da variável aleatória X_i.
b) Escreva a f.p. da variável aleatória X_i, usando a variável indicadora para o seu contradomínio.
c) Qual a f.p. conjunta da a.a. (X_1, X_2,..., X_n)?
d) Determine a estatística suficiente para estimar θ.

5) Seja [X_1, X_2,...,X_n] uma a.a. de uma distribuição U [0, θ].

a) Escreva a f.d.p da variável aleatória X.
b) Escreva a f.d.p da v.a X_i usando a variável indicadora I para o seu contradomínio.
c) Qual a função densidade conjunta da a.a. [X_1,X_2,...,X_n]?
d) Determine a estatística suficiente para θ.

6) Seja [X_1,X_2,...,X_n] a.a., onde X_i é uma v.a i.i.d. N(μ,σ^2), com ambos os parâmetros desconhecidos. Seja $\underline{\theta}$'= [μ,σ^2].

a) Escreva a f.d.p da variável aleatória X_i.
b) Escreva a f.d.p da variável aleatória X_i usando a variável indicadora I para o seu contradomínio.
c) Qual a função de densidade conjunta da a.a. (X_1,X_2,...,X_n)?

d) Determine usando o T. de Neyman-Fischer a estatística suficiente para $\underline{\theta}' = [\mu, \sigma^2]$.

2.4 FAMÍLIA EXPONENCIAL UNIPARAMÉTRICA

Uma v.a. em $\mathbb{R}$ possui distribuição da família exponencial uniparamétrica se a sua f.p. ou f.d.p é da forma $f(x|\theta) = \exp\{c(\theta)T(x) + d(\theta) + S(x)\}.I_A(x)$, onde $\theta \in \Theta$, intervalo aberto de $\mathbb{R}$ e o conjunto $A = \{x \mid f(x|\theta) > 0\}$ é independente de θ.

EXEMPLO 1 :A distribuição binomial, b(n,θ), pertence a família exponencial, pois :

$$f(x|\theta) = \binom{n}{x}\theta^x \cdot (1-\theta)^{n-x} = \left\{\exp\left[\ln\binom{n}{x} + x\ln\theta + (n-x)\ln(1-\theta)\right]\right\} I_{N}^{*}(x)$$

$$= \left\{\exp\left[\ln\binom{n}{x} + x\ln\theta + n\ln(1-\theta) - x\ln(1-\theta)\right]\right\} I_{N}^{*}(x)$$

$$f(x|\theta) = \left\{\exp\left[\ln\binom{n}{x} + x\ln\frac{\theta}{1-\theta} + n\ln(1-\theta)\right]\right\} I_{N^*}(x)$$

$$f(x|\theta) = \left\{\exp\left[\underbrace{x\ln\frac{\theta}{1-\theta}}_{T(x)} + \underbrace{n\ln(1-\theta)}_{d(\theta)} + \underbrace{\ln\binom{n}{x}}_{S(x)}\right]\right\} I_{N^*}(x)$$

EXEMPLO 2: A distribuição U [0, θ] não pertence a família exponencial, pois :

$$f_X(x) = \frac{1}{\theta} \qquad 0 \le X \le \theta$$

$f_X(x) = \exp\{-\ln\theta\}$ não corresponde a forma de família exponencial, uma vez que o limite na extremidade b depende de θ.

Se $[X_1, X_2,...,X_n]$ é uma a.a. de uma população com f.p. ou f.d.p da família exponencial, então a estatística $T(\underline{X})$ é uma estatística suficiente para θ, pelo Teorema de Fatorização.

EXERCÍCIOS:

1) Seja $[X_1, X_2,...,X_n]$ uma a.a. de uma distribuição normal $N(\mu,\theta)$ com μ fixo e conhecido.
   a) Escreva a função densidade de probabilidade conjunta.
   b) Escreva a f.d.p conjunta na forma da família exponencial.
   c) Identifique as funções $c(\theta)$, $T(\underline{x})$, $d(\theta)$ e $S(\underline{x})$.
   d) Qual a estatística suficiente para θ?

2) Suponha $[X_1, X_2,...,X_n]$ uma a.a. de uma população com uma das seguintes densidades. Ache a estatística suficiente para θ, com **a** fixo, e identifique $c(\theta)$, $T(x)$, $d(\theta)$ e $S(\underline{x})$.

   a) Beta(θ,1)
   b) Weibull(θ,a)
   c) Paretto (θ,a)

2.5 FAMÍLIA EXPONENCIAL K-PARAMÉTRICA

A família de distribuições $\{P_\theta;\ \theta \in \Theta\}$ é dita família exponencial com k parâmetros ou k-paramétrica, se existem as funções de valor real $c_1, c_2, \ldots, c_K$ e $d(\theta)$, e ainda $T_1, T_2, \ldots, T_K$ funções de variável real, e também S, definidas em R^n e um conjunto $A \subset R^n$ tal que a f.d.p (ou f.p.) de P_θ pode ser escrita na forma :

$$p(\underline{x},\theta) = \left\{\exp\left[\sum_{i=1}^{K} c_i(\theta)T_i(\underline{x}) + d(\theta) + S(\underline{x})\right]\right\} I_A(x)$$

e pelo Teorema da Fatorização o vetor $\underline{T}(\underline{x}) = [T_1(\underline{x}), \ldots, T_K(\underline{x})]'$ é suficiente para $\underline{\theta}' = (\theta_1, \theta_2, \ldots, \theta_K)$.

EXERCÍCIOS :

1) Seja ($X_1, X_2, \ldots, X_n$) uma a.a. de uma distribuição $N(\mu,\sigma^2)$, onde $\underline{\theta} = \begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}$.

a) Qual o espaço paramétrico do parâmetro θ ?
b) Escreva a f.d.p de X_i.
c) Escreva a f.d.p conjunta da amostra.
d) Escreva a f.d.p conjunta na forma da família exponencial.
e) Qual a estatística suficiente para θ ?

2) Seja ($X_1, X_2, \ldots, X_n$) uma a.a. de uma distribuição $\Gamma(\alpha,\beta)$, onde $\underline{\theta} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ é o vetor de parâmetros.

a) Qual o espaço paramétrico de θ ?
b) Escreva a f.d.p da v.a X_i.
c) Escreva a f.d.p conjunta da amostra.
d) Escreva a f.d.p conjunta na forma da família exponencial e identifique $c_i(\theta)$, $T_i(\underline{x})$, $d(\theta)$ e $S(\underline{x})$.
e) Qual a estatística suficiente para θ ?

III MODELOS PARAMÉTRICOS

3.1 CONCEITO

Quando se usa a designação *paramétrico*, o significado do termo é o de que a forma da f.p. ou f.d.p da v.a. foi especificada a *priori* e não é posta *em questão*. Além disto tem-se que:

- as inferências dizem respeito a um número finito de parâmetros;
- as inferências dependem da forma especificada para a f.d.p. ou f.p.

3.2 DEFINIÇÕES

1ª.) **Espaço paramétrico** Θ é o conjunto de todos os valores que o parâmetro θ pode assumir.

2ª.) Uma parametrização é chamada **identificável** quando para parâmetros diferentes tem-se f.d.p's ou f.p's diferentes ou seja para $\theta_1 \neq \theta_2$ tem-se $P_{\theta_1} \neq P_{\theta_2}$.

EXEMPLO 1

No modelo $X \sim N(\theta,\sigma^2)$, o parâmetro θ possui o espaço paramétrico $\Theta = \{\theta \mid \theta \in R\}$.

EXEMPLO 2

O modelo $X \sim N(\mu,\sigma^2)$ tem o parâmetro $\underline{\theta}' = [\mu,\sigma^2]$ com espaço paramétrico $\Theta = \{(\mu,\sigma^2) \mid -\infty < \mu < -\infty, \sigma > 0\}$

EXERCÍCIOS

1) Uma geólogo mede os diâmetros de um grande número, n, de seixos em um leito de um riacho. Considerações teóricas o conduz a acreditar que o logaritmo do diâmetro da pedra possui distribuição $N(\mu,\sigma^2)$. Ele deseja usar suas observações para obter alguma informação sobre μ e σ^2, mas não tem conhecimento nenhum da magnitude dos dois parâmetros.

   a) Escreva se o modelo a ser usado pelo geólogo é paramétrico ou não e o porquê.
   b) Escreva o espaço paramétrico do parâmetro $\underline{\theta}' = (\mu,\sigma^2)$.
   c) Sabendo-se que uma parametrização é chamada de <u>identificável</u> quando para $\theta_1 \neq \theta_2$ tem-se $f(x, \theta_1) \neq f(x, \theta_2)$, escreva se a parametrização do modelo é identificável.

2) Seja $[X_1, X_2,...,X_n]$ n determinações de uma constante μ (física). Um modelo sugerido para a v.a. X_i é $X_i = \mu+\varepsilon_i$, i = 1, 2,..., n, onde ε_i pode ser considerado como erro de medida (instrumento, pessoa, etc.). A constante física é conhecida na Física como o verdadeiro valor. As condições para ε_i são :
   1ª) A distribuição de ε_i é independente de μ ;
   2ª) A distribuição de ε_i é simétrica e contínua em torno de 0 ;
   3ª) Os ε_i são identicamente distribuídos ;
   4ª) ε_i é independente de ε_j para $i \neq j$;

5ª) $\varepsilon_i \sim N(0,\sigma^2)$;
6ª) σ^2 é conhecida (variância da população).

Pede-se:
a) O modelo é paramétrico?
b) A parametrização é identificável ?
c) Supondo que o instrumento de medida seja viciado para o lado positivo por 0,1. A parametrização é identificável ?
d) Supondo que o instrumento de medida seja viciado, mas com vício b desconhecido. O modelo é identificável ?
e) Qual o espaço paramétrico de θ, desconsiderando as suposições dos itens c e d ; desconsiderando somente o item d e não desconsiderando item nenhum ?

3) O número de ovos postos por um inseto segue uma distribuição de Poisson com media λ desconhecida. Uma vez botado, cada ovo tem uma chance desconhecida p de chocar e o ato de chocar um ovo é independente do chocamento dos outros. Um entomologista estuda o conjunto de n de tais insetos, observando ambos o número de ovos botados o número de ovos chocados para cada ninho.

a) modelo é paramétrico, por quê?
b) Escreva o espaço paramétrico do parâmetro (λ,p).
c) A parametrização é identificável ?
d) Qual a expressão da função distribuição de probabilidade correspondente a cada uma das v.a's envolvidas no estudo ?

IV ESTIMAÇÃO

4.1 ESTIMAÇÃO POR PONTO

A estimação pontual (por ponto) consistirá simplesmente em, à falta de melhor informação, adotar a estimativa disponível como sendo o valor do parâmetro. A idéia é, em sua essência, extremamente simples, porém a qualidade dos resultados irá depender fundamentalmente da conveniente escolha do estimador. Assim, dentre os vários estimadores razoáveis que poderemos imaginar para um determinado parâmetro, devemos ter a preocupação de escolher aquele que melhor satisfaça às propriedades de um bom estimador.

Considere a família das f.d.p. ou f.p. $\{f_X(x,\theta)|\theta\in\Theta\}$. Pode ocorrer do experimentador necessitar selecionar um membro desta família como sendo a f.d.p. ou f.p. da variável aleatória X. Isto é, ele necessita de uma estimativa por ponto do parâmetro θ. Seja $[X_1, X_2,...,X_n]$ uma a.a. da distribuição que tem como f.d.p. (f.p.) um membro da família $\{f_X(x,\theta)|\theta\in\Theta\}$. Nosso problema é definir uma *estatística* $T(\underline{X})$ de maneira que $[X_1, X_2,...,X_n]$ são os valores experimentais observados, então o número $T(\underline{X})$ é uma boa estimativa pontual de θ .

EXEMPLO :

Seja $[X_1, X_2,...,X_n]$ uma a.a. da distribuição com f.d.p. .

$$f(x) = \theta^{x}(1-\theta)^{1-x} \qquad x = 0,1$$

A estimativa $T(\underline{X}) = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ é um bom *estimador* do parâmetro θ .

4.2- PROPRIEDADES DOS ESTIMADORES PARAMÉTRICOS

4.2.1- Consistência

DEF. 1 Um estimador $T_n(\underline{X})$ é *consistente* para estimar o parâmetro θ quando, a medida que, se aumenta o tamanho n da a.a. $\underline{x}$, consegue-se uma maior precisão na estimativa ou seja, para $T_n(\underline{X}) = T(x_1,x_2, ... ,x_n)$ e um ε, real positivo arbitrariamente pequeno, n > n' tem-se $P(\theta-\varepsilon < T_n < \theta+\varepsilon)>P(\theta-\varepsilon < T_{n'} < \theta-\varepsilon)$. Assim um estimador $T_n(\underline{X})$ é dito *consistente* se,

$$\lim_{n\to\infty} P(|T_n-\theta| < \varepsilon)= 1 \qquad \forall\varepsilon > 0$$

DESIGUALDADE DE TCHEBYCHEV

$$P[|X–E(X)| \geq \lambda] \leq \frac{V(X)}{\lambda^2} \qquad \forall \lambda > 0$$

$$P[|X – E(X)| < \lambda] \geq 1-\frac{V(X)}{\lambda^2} \qquad \forall \lambda > 0$$

EXERCÍCIO 1:

Prove que a média amostral, $\bar{x}$, de uma a.a. de uma população (distribuição) com média μ e variância σ^2 é estimador consistente do parâmetro μ..

TEOREMA DAS CONDIÇÕES PARA CONSISTÊNCIA DE UM ESTIMADOR

As condições gerais para consistência de um estimador são:

- $\lim_{n\to\infty} E[T_n(\underline{X})] = \theta$
- $\lim_{n\to\infty} V[T_n(\underline{X})] = 0$

são suficientes para que a estatística $T_n(\underline{x})$ seja estimador consistente do parâmetro θ.

EXERCÍCIO 2

Verifique aplicando as condições dadas no teorema anterior e também o *conceito de convergência em probabilidade* se os estimadores seguintes são consistentes .

a) $X \sim N(\mu,\sigma^2)$, $T_n(\underline{X}) = \bar{x}$ como estimador de μ (parâmetro média populacional)

b) $X \sim N(\mu,\sigma^2)$, $T_n(\underline{X}) = s^2 = \dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$ como estimador de σ^2 (parâmetro variância populacional).

c) X ~R(θ^2) Distribuição Rayleig, $T_n(\underline{x}) = \dfrac{\sum_{i=1}^{n} x_i^2}{2n}$ para o parâmetro θ^2

d) X~ $\Gamma(4,1/\theta)$, $T_n(\underline{x}) = \dfrac{\sum_{i=1}^{n} x_i}{4n}$ como estimador de θ.

4.2.2.- Estimadores Não-viciados

DEF. **2**: Quando estimamos o parâmetro θ ou uma função do parâmetro, $q(\theta)$, usando como estimador a estatística $T(\underline{X})$, onde $\underline{X}' = [X_1, X_2,...,X_n]$ é uma a.a., temos que uma medida do desempenho do estimador pode ser dada pelo *erro médio quadrático*:

$$R(\theta,T) = E[T(\underline{X}) - q(\theta)]^2$$

Que pode ser colocada na forma $R(\theta,T) = V(T(\underline{X})) + b^2(\theta,T)$, onde $b^2(\theta,T)$ é o quadrado do vício do estimador para o parâmetro θ.

EXERCÍCIO 3
Prove que o erro médio quadrático, definido acima como a esperança do quadrado do desvio da estatística em relação ao parâmetro, pode ser decomposto na forma acima especificada.

É importante observar que:
- A variância $V(T(\underline{X}))$ mede como os valores de $T(\underline{X})$ estão dispersos em torno do centro $E(T(\underline{X}))$, isto define a *precisão* da estatística.
- O vício $b(T,\theta) = E(T(\underline{X}) - q(\theta)$ mede quanto e em qual direção o centro da distribuição de $T(\underline{X})$ (valor esperado) está fora do verdadeiro valor $q(\theta)$, isto define a *acurácia* da estatística.

DEF. 3 Se a estatística $T(\underline{X})$ que estima o parâmetro $q(\theta)$ é tal que $R(\theta,T) = V(T(\underline{X}))$, então $T(\underline{X})$ é chamado de *estimador não viciado* de $q(\theta)$.

DEF. **4** Dados os estimadores $T(\underline{X})$ e $S(\underline{X})$ do parâmetro $q(\theta)$, $S(\underline{X})$ será estimador inadmissível de $q(\theta)$ se $R(\theta,T) < R(\theta,S)\ \forall\theta \in \Theta$.

EXERCÍCIO 4:
Suponha que $\underline{X}$ seja uma a.a. de tamanho n de uma população $N(\mu,\sigma^2)$ e considere os estimadores dos parâmetros μ e σ^2, respectivamente $\bar{x}$ e $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$.

a) Calcule o valor da esperança e variância de $\bar{x}$.
b) Calcule o vício e o erro médio quadrático de $\bar{x}$.
c) Calcule o valor da esperança e variância de $\hat{\sigma}^2$.
d) Calcule o vício e o erro médio quadrático de $\hat{\sigma}^2$.

EXERCÍCIO 5:
Suponha que $[X_1,X_2, ...,X_n]$ seja uma a.a. de uma população $P(\theta)$.
a) Verifique se o estimador $\bar{x}$ do parâmetro θ é não viciado.
b) Existe algum outro estimador não viciado para θ ? Qual ?

EXERCÍCIO 6:
Suponha que $[X_1,X_2, ...,X_n]$ seja uma a.a. de uma população b(1,θ).
a) Verifique se o estimador $\bar{x}$ do parâmetro θ (proporção de sucessos) é não-viciado.
b) Verifique se o estimador $s^2 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$ do parâmetro $\sigma^2 = \theta(1-\theta)$ (variância populacional) é não-viciado.

EXERCÍCIO 7:
Suponha que $[X_1,X_2, ...,X_n]$ seja uma a.a. de uma população $N(\mu,\sigma^2)$. Determine um estimador suficiente, não-viciado e consistente para cada um dos parâmetros μ e σ^2 (considere os casos de μ conhecido ou não).

EXERCÍCIO 8:
Suponha que $[Y_1,Y_2, ... ,Y_n]$ seja uma a.a. de uma população b(1,θ). Determine um estimador suficiente, não-viciado e consistente para cada um dos parâmetros $\mu = n\theta$ e $\sigma^2= n\theta(1-\theta)$ que correspondem a média e a variância da v.a. Y que conta o número de sucessos nessas n provas Bernoulli.

OBS. Uma propriedade fundamental dos estimadores é a suficiência que foi abordada anteriormente e outra fundamental é a da eficiência ou variância mínima que será vista adiante.

EXERCÍCIO 9:
Descreva as quatro propriedades fundamentais dos estimadores: suficiência, consistência, não-tendenciosidade e eficiência (estimador de variância mínima).

4.3 MÉTODOS DE ESTIMAÇÃO

A evolução da Estatística através do tempo provocou o aparecimento de várias metodologias para construção de estimadores de parâmetros.

4.3.1 – Método da Máxima Verossimilhança (Fisher)

Este método foi desenvolvido por Fisher a partir de uma idéia original de Gauss.

Seja a a.a. $\underline{X}' = [X_1,X_2, ...,X_n]$ de uma população (distribuição), com X_i i.i.d., e com f.p. ou f.d.p. pertencente a família $p(x_i,\theta)$ $\theta \in \Theta$ e se deseja estimar o parâmetro θ. A função conjunta $p(\underline{X},\theta)$ representa a probabilidade (ou a f.d.p.) de ocorrer o vetor $\underline{X}' = [X_1,X_2, ...,X_n]$, quando o valor do parâmetro é θ. Assim poderíamos fixar a amostra $\underline{X}$ e procurar o valor de θ que maximiza a probabilidade de ocorrer esta amostra $\underline{X}$.

DEF. 5: O estimador $\hat{\theta}(\underline{X})$ é o estimador de máxima verossimilhança (EMV) de θ, quando $p[\underline{X}, \hat{\theta}(\underline{X})] = \max p(\underline{X},\theta) \;\; \forall\theta \in \Theta$.

EXEMPLO 2

Suponha que θ pode assumir os valores 1 ou 0 e que p(x,θ) é dada pelo quadro abaixo:

x \ θ	0	1
0	0,3	0,6
1	0,7	0,4
∑	1,0	1,0

- O estimador EMV de θ quando x = 0 é $\hat{\theta}(\underline{x}) = 1$, porque maximiza a probabilidade de x = 0 ocorrer (0,6);
- O estimador EMV de θ quando x = 1 é $\hat{\theta}(\underline{x}) = 0$, porque maximiza a probabilidade de x = 1 ocorrer (0,7).

DEF.6: FUNÇÃO DE VEROSSIMILHANÇA

A função $p(\theta,\underline{X})$ é chamada de função de verossimilhança quando na função da distribuição fixamos o valor de $\underline{X}$ e fazemos variar θ, de forma que $L(\underline{X},\hat{\theta}(\underline{X})) = p(\underline{X},\hat{\theta}(\underline{X})) = \max\{p(\underline{X},\hat{\theta}(\underline{X}); \forall\theta \in \Theta\} = L(\underline{X},\hat{\theta}(\underline{X},\theta))$.

CÁLCULO DO ESTIMADOR DE MÁXIMA VEROSSIMILHANÇA

Para determinar o EMV do parâmetro θ basta achar o valor de θ que maximiza a função $p(\underline{X},\theta)$, fixado $\underline{X}$. Como a função $\ell n(p(\underline{X},\theta))$ é não decrescente (monótona crescente), maximizar $p(\underline{X},\theta)$ é o mesmo que maximizar $\ell n[p(\underline{X},\theta)]$. Assim, se o EMV $\hat{\theta}$ existe, deve verificar $\frac{\partial \ell n[p(\underline{x},\theta)]}{\partial\theta} = 0$, uma vez que se supõe que exista derivada parcial em relação a θ.

EXERCÍCIO 9

Seja $\underline{x}' = [X_1, X_2, ..., X_n]$ uma a.a. de uma distribuição b(1,θ).

a) Escreva a f.p. da v.a. X_i ;
b) Escreva a distribuição (conjunta) da amostra;
c) Escreva a função de verossimilhança da a.a.;
d) Qual o valor de θ que maximiza a função de verossimilhança ?

OBS. Nem sempre é possível achar o EMV por derivada, então temos que analisar a função (função não-crescente).

EXERCÍCIO 10

Seja a v.a. X ~ U(0,θ), onde θ é um parâmetro positivo e desconhecido. Determine o EMV de θ.

EXERCÍCIO 11:

Seja X o número de consumidores que chegam em um Serviço e que são observados por hora, em n horas. Se as chegadas formam um Processo de Poisson, então X ~ P(θ), onde θ representa o número esperado de chegadas em uma hora ou equivalentemente, a taxa de chegadas. Na prática θ é desconhecido e nós desejamos estima-lo, usando os valores observados de X (amostra). Determine o EMV de θ.

EXERCÍCIO 12

Seja X uma característica física de interesse que segue a distribuição $N(\mu,\sigma^2)$. Suponha que não se conhece os parâmetros μ e σ^2, e se deseja estima-los com base no conjunto de observações $\underline{X}' = [X_1, X_2, ..., X_n]$. Determine os estimadores EMV destes parâmetros.

4.3.2 – Métodos dos Momentos (K. Pearson)

Seja uma a.a. $\underline{X}' = [X_1, X_2, ..., X_n]$ de uma população com f.p. ou f.d.p. dependendo de k parâmetros $\theta_1, \theta_2, ... \theta_k$. Os momentos ordinários da população $m_j = \int_{-\infty}^{\infty} x^j f(x, \theta_1, \theta_2, ... \theta_k)dx$, se existirem, são funções dos k parâmetros $m_j = f(\theta_1, \theta_2, ... \theta_k)$ $j = 1,2,3, ...$

Considere, também, os momentos ordinários da amostra, $M_j = \frac{\sum_{i=1}^{n} x^j}{n}$ $j = 1,2,3, ...,$ forme o sistema de equações:

$$M_j = m_j = f(\theta_1, \theta_2, ... \theta_k) \quad j = 1,2,3, ...$$

e admita que tem solução única, $\hat{\theta}_j (X_1, X_2, ..., X_n)$ $j = 1,2,3, ...,k$. Então, estes k estimadores, solução do sistema de equações, são os estimadores dos parâmetros pelo Método dos Momentos.

OBS. Os estimadores obtidos pelo Método dos Momentos são em geral consistentes e possuem distribuição assintótica Gaussiana, porém não são assintoticamente mais *eficientes* do que os estimadores de máxima verossimilhança.

EXERCÍCIO 13:

Seja $\underline{X}' = [X_1, X_2, ..., X_n]$ amostra aleatória de uma população Bernoulli, $b(1,\theta)$. Determine o estimador de θ pelo Método dos Momentos.

EXERCÍCIO 14

Seja $\underline{X}' = [X_1, X_2, ..., X_n]$ amostra aleatória de uma população $\Gamma(\alpha,\beta)$. Determine os estimadores dos parâmetros pelo Método dos Momentos.

EXERCÍCIO 15

Considere n sistemas com tempos de falha $\underline{X}' = [X_1, X_2, ..., X_n]$. Assumindo que as v.a’s X_i são i.i.d. com distribuição $\varepsilon(\theta)$, determine:

a) O estimador de θ pelo Método dos Momentos;
b) Um estimador de θ diferente do anterior;
c) Um estimador da probabilidade $P(X_i > 1)$;
d) Um estimador da probabilidade $P(X_i \leq 1)$.

4.3.3. – Método dos Mínimos Quadrados (Gauss)

Toda observação aleatória pode ser escrita na forma do modelo

$$Y_i = g_i(\theta_1, \theta_2, ..., \theta_k) + \varepsilon_i \quad i = 1,2, ..., n,$$

onde a primeira parcela corresponde a *parte sistemática* do modelo e a segunda parcela é a parte *estocástica*. As funções g_i são conhecidas e os números reais $\theta_1, \theta_2, ... ,\theta_k$ são desconhecidos (parâmetros), podendo variar livremente no conjunto $\Theta \subset \Re^k$. A parte estocástica ε_i deve satisfazer pelo menos aproximadamente as seguintes restrições:

- ε_i é uma v.a. com $E(\varepsilon_i) = 0$ $\quad$ i = 1,2, ... n
- ε_i é uma v.a. com variância constante $V(\varepsilon_i) = \sigma^2$ $\quad$ i = 1,2, ... ,n
- Os erros ε_i são não-correlacionados, $cov(\varepsilon_i, \varepsilon j) = 0$.

O método consiste em estimar θ pelo estimador que minimiza a Soma dos Quadrados do Erros (resíduos): $SQR = \sum_{i=1}^{n} (Y_i - g_i(\theta_1, \theta_2, ..., \theta_k))^2 = \sum_{i=1}^{n} \varepsilon_i^2$

Os *estimadores de mínimos quadrados* são obtidos resolvendo o sistema de equações:

$$\frac{\partial}{\partial \theta} \sum_{i=1}^{n} [y_i - g_i(\theta)]^2 = 0 \qquad j = 1,2, ... ,k$$

Estas equações são denominadas de *equações normais.*

EXERCÍCIO 16

Seja o modelo linear especificado por $Y_i = \theta_1 + \theta_2 X_i + \varepsilon_i$ Determinar os estimadores de mínimos quadrados dos parâmetros θ_1 e θ_2, com base em n observações (Y_i, X_i).

EXERCÍCIO 17

Resolver o exercício anterior considerando um modelo genérico com p parâmetros e n > p observações. Assuma o tratamento matricial.

EXERCÍCIO 18

Foram tomadas n = 9 amostras de um solo (canteiros) que foram tratadas com diferentes quantidades X de fósforo. Seja Y a quantidade de fósforo presente em sementes de plantas, de 38 dias, crescidas nos diferentes canteiros. Os dados estão abaixo.

a) Determine as estimativas dos parâmetros θ_1 e θ_2 da reta de mínimos quadrados com que se pretende fazer a regressão de Y para X;
b) Determine a estimativa de máxima verossimilhança do coeficiente de correlação ρ que mede a associação entre X e Y, assumindo que Y e X têm distribuição Gaussianas.
c) Desenhe um esboço do Diagrama de Dispersão entre X e Y.

4.4 ESTIMADORES NÃO-VICIADOS UNIFORMEMENTE DE MÍNIMA VARIÂNCIA (UMVU) E INTERVALOS DE CONFIANÇA

4.4.1 – Estatísticas Completas

A uma estatística $T_n(\underline{X})$, nós podemos acrescentar ao conceito de *suficiência* (devido a Fisher) o conceito de *completitude* (devido a Lehmann e Scheffé) e afirmar que ela é uma *estatística ótima*.

DEF.7: **Família completa de densidades**: Seja $[X_1,X_2, \ldots ,X_n]$ uma a.a. da densidade $f(.,\theta)$ com espaço paramétrico Θ, e seja $T = t(\underline{X})$ uma estatística. A família de densidades de T é definida como completa se e somente se $E_\theta[z(T)] \equiv 0 \ \forall \ \theta \in \Theta$ implica em $P_\theta[z(T)=0] = 1 \ \forall \ \theta \in \Theta$, onde z(T) é uma estatística. E, também, uma estatística é dita ser completa se e somente se sua família de densidades é completa.
DEF.8: Uma estatística $T_n(\underline{X})$ é dita ser completa se a função de valor real $g(T_n(\underline{X})) = g(T)$ definida na variação de T que satisfaz a equação $E_\theta[g(T)] = 0 \ \forall \ \theta \in \Theta$ é a função $g(T) = 0$.

EXEMPLO: Seja $[X_1,X_2, \ldots ,X_n]$ uma a.a. de uma distribuição Bernoulli, $b(1,\theta)$. A estatística $T = X_1 – X_2$ **não é completa**, pois $E_\theta[X_1-X_2] = 0 \ \forall \ \theta \in \Theta$, mas $X_1 – X_2$ não é 0 com probabilidade 1.

EXERCÍCIO 19

a) Considere o exemplo anterior e a estatística $T = \sum_{i=1}^{n} X_i$. Seja z(T) uma estatística que é função de T e $E_\theta[z(T)] \equiv 0 \ \forall \ \theta \in \Theta$. Prove que T é completa, mostrando que z(T) = 0 para t = 0,1, ... ,n.

b) Seja $[X_1,X_2, \ldots ,X_n]$ uma a.a. da distribuição $U(0,\theta)$, onde $\Theta = \{\theta \mid \theta > 0\}$. Mostre que a estatística $Y_{(n)}$ é completa.

TEOREMA DE RAO-BLACKWELL

Seja $[X_1,X_2, \ldots ,X_n]$ uma a.a. da densidade $f(.,\theta)$, e seja $S_1 = s_1(\underline{X})$, $S_2 = s_2(\underline{X})$, ... ,$S_k = s_k(\underline{X})$ um conjunto de estatísticas conjuntamente suficientes. Seja a estatística $T = t(\underline{X})$ um estimador não-viciado de $q(\theta)$. Defina T’ por $T' = E[T|S_1,S_2, \ldots ,S_k]$, então:

- T’ é uma estatística e é uma função de estatísticas suficientes $S_1,S_2, \ldots S_k$;
- $E_\theta[T'] = q(\theta)$, isto é T’ é um estimador não-viciado de $q(\theta)$;
- $V_\theta[T'] \leq V(T) \ \forall \ \theta \in \Theta$ e $V_\theta[T'] < V(T)$ para algum θ a menos que T seja igual a T’ com probabilidade 1.

EXERCÍCIO 20
Prove o resultado enunciado acima.

EXERCÍCIO 21
Seja $[X_1,X_2, ... ,X_n]$ uma a.a. da densidade Bernoulli, b(1,θ). E, seja X_1 um estimador de $q(\theta)= \theta$ que será assumido como $T = t(\underline{X})$ no Teorema de Rao-Blackwell. $\sum_{i=1}^{n} X_i$ é uma estatística suficiente para θ, assim usar-se-á $S = \sum_{i=1}^{n} X_i$ como o conjunto de estatísticas suficientes (de um elemento). Então, de acordo com o Teorema de Rao-Blackwell T' = $E[T|S] = E[X_1|\sum_{i=1}^{n} X_i]$ é um estimador não-viciado de θ com variância não maior do que a de $T = X_1$. Mostre que a variância de T' é menor que a variância de $T = X_1$.

TEOREMA DA FAMÍLIA EXPONENCIAL PARA ESTATÍSTICAS SUFICIENTES E COMPLETAS

Seja $\{P_\theta|\theta \in \Theta\}$ uma família exponencial k-paramétrica dada por $p(\underline{x},\theta) = \{\exp[\sum_{i=1}^{n} c_i(\theta)T_i(\underline{x}) + d(\theta) + S(\underline{x})]\}I_A(\underline{x})$. Suponha que a variação de $\underline{c} = [c_1(\theta),c_2(\theta), ... ,c_k(\theta)]$ tenha um interior não-vazio. Então $\underline{T}(\underline{x}) = [T_1(\underline{x}), T_2(\underline{x}), ... ,T_k(\underline{x})]$ é uma estatística suficiente e completa .

EXERCÍCIO 22
Seja $[X_1,X_2, ... ,X_n]$ uma a.a. de uma população $N(\mu,\sigma^2)$ onde os parâmetros são desconhecidos. Mostre que a estatística $\underline{T}(\underline{x}) = [\sum_{i=1}^{n} X_i , \sum_{i=1}^{n} X_i^2]$ é suficiente e completa.

4.4.2 – Estimadores UMVU

TEOREMA DE LEHMANN-SCHEFFÉ

Se $T(\underline{X})$ é uma estatística suficiente e completa e $S(\underline{X})$ é um estimador não-viciado de q(θ), então $T^*(\underline{X}) = E[S(\underline{X})|T(\underline{X})]$ é um estimador UMVU de q(θ). Se $V_\theta(T^*(\underline{X}) < \infty \ \forall \theta \in \Theta$, $T^*(\underline{X})$ é o único estimador UMVU de q(θ).

Podemos usar o Teorema de Lehmann-Scheffé na busca de estimadores UMVU de dois modos:

- Se nós podemos achar uma estatística uma estatística da forma $h[T(\underline{X})]$ tal que $h[T(\underline{X})]$ seja um estimador não-viciado de q(θ), então $h[T(\underline{X})]$ é UMVU para q(θ). Porquê, isto é verdade ?

- Se nós podemos achar algum estimador não-viciado $S(\underline{X})$ de q(θ), então $E[S(\underline{X})|T(\underline{X})]$ é UMVU para q(θ).

EXERCÍCIO 23
Seja $[X_1,X_2, \ldots ,X_n]$ uma a.a. de uma população $N(\mu,\sigma^2)$ onde os parâmetros são desconhecidos. Determine os estimadores UMVU de μ e σ^2.

EXERCÍCIO 24
Seja $[X_1,X_2, \ldots ,X_n]$ uma a.a. de uma população $\varepsilon(\theta)$ onde o parâmetro θ é desconhecido.

a) Escreva a f.d.p da v.a. X_i ;

b) Escreva a f.d.p. da v.a. X_i na forma da família exponencial;

c) Escreva a f.d.p. da a.a. ;

d) Escreva a f.d.p. da a.a. na forma da família exponencial;

e) Identifique alguma estatística suficiente e completa para estimar θ;

f) Determine um estimador UMVU para θ.

EXERCÍCIO 25
Seja $[X_1,X_2, \ldots ,X_n]$ uma a.a. de indicadores de provas binomiais com probabilidade θ de sucesso.

a) Escreva a f.d.p da v.a. X_i ;

b) Escreva a f.d.p. da v.a. X_i na forma da família exponencial;

c) Escreva a f.d.p. da a.a. ;

d) Escreva a f.d.p. da a.a. na forma da família exponencial;

e) Identifique alguma estatística suficiente e completa para estimar θ;

f) Determine um estimador UMVU para θ.

EXERCÍCIO 26
Seja $[X_1,X_2, \ldots ,X_n]$ uma a.a. de uma população $P(\theta)$ com $\theta > 0$.

a) Escreva a f.d.p da v.a. X_i ;

b) Escreva a f.d.p. da v.a. X_i na forma da família exponencial;

c) Escreva a f.d.p. da a.a. ;

d) Escreva a f.d.p. da a.a. na forma da família exponencial;

e) Identifique alguma estatística suficiente e completa para estimar θ;

f) Determine um estimador UMVU para θ.