PGF5103 Tópicos em tratamento estatístico de dados em fís. experimental.
1ª prova – 11/10/2013. Duração: 100 minutos. SEM CONSULTA

Observação: *Vários itens destes problemas admitem mais de uma resposta, conforme a interpretação que você faça e a importância que atribua às circunstâncias enunciadas. Assim, explique e justifique o procedimento que adotar, a fim de que seja possível determinar a coerência e consistência da resposta. Em outras palavras, a discussão vale tanto quanto o resultado.*

Esta prova consiste de 3 questões, cujos enunciados estão em três páginas diferentes. A última página tem um formulário, com mais fórmulas que as necessárias para resolver as questões.

1) (3,5 pontos) Uma experimentadora ajustou os parâmetros a e b da função

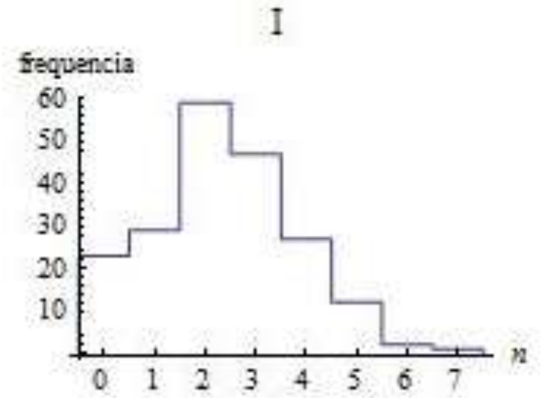
$$z(x, y; a, b) = ax + by$$

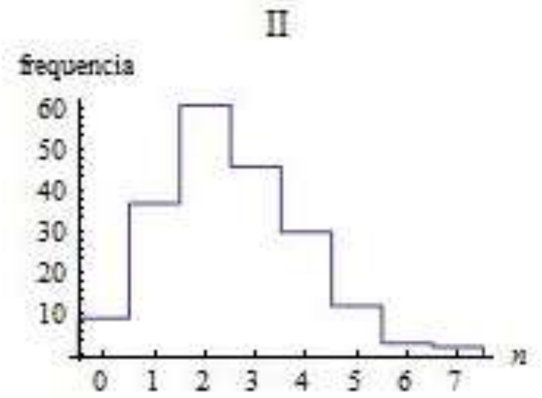
aos seus dados experimentais $\{(x_i, y_i, z_i, \sigma_i,), i = 1..N\}$, onde σ_i é o desvio-padrão de z_i e os erros em x_i e y_i podem ser ignorados. Ela usou o método dos mínimos quadrados e obteve o resultado

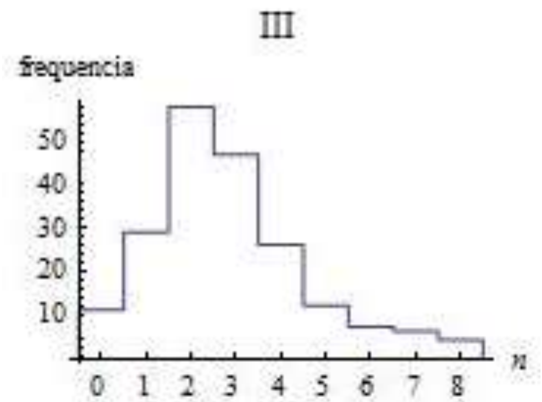
$$(\hat{a}, \hat{b}) = (2{,}0; 3{,}0) \quad \text{com matriz de covariância} \quad \begin{pmatrix} 2{,}0 & 1{,}0 \\ 1{,}0 & 4{,}0 \end{pmatrix} \quad .$$

a) Calcule o valor interpolado de z quando $(x, y) = (0; 2)$, $\hat{z}_{0;2}$, **e sua respectiva incerteza**.
b) Calcule o valor interpolado de z quando $(x, y) = (1; 1)$, $\hat{z}_{1;2}$, **e sua respectiva incerteza**.
c) Calcule a covariância entre $\hat{z}_{0;2}$ e $\hat{z}_{1;1}$.
d) Em linhas gerais, como deveria ser modificado o experimento para diminuir as variâncias dos resultados? Enuncie mais de uma maneira de reduzir as incertezas nos resultados e explique as eventuais implicações de cada escolha.

2. (3 pontos) Uma experiência do laboratório didático consiste na análise das propriedades estatísticas de um jogo de dados especiais, que têm duas das seis faces marcadas; assim, a probabilidade de um dado cair com uma face marcada para cima em um lance é $p = \frac{1}{3}$. Cada grupo recebe 8 dados, lança o conjunto, conta quantos caíram com a face marcada para cima, n, e repete m = 200 vezes o lançamento do conjunto de dados.







Os histogramas de 3 turmas de estudantes estão ao lado, identificados como I, II e III. “Frequencia” é o número de vezes em que um lance resultou em n dados com a face marcada para cima.

a) Calcule a probabilidade de nenhum dado cair com a face marcada para cima em um único lance e a dos 8 dados caírem com a face marcada para cima em um único lance. *Se efetuar os cálculos sem calculadora, aproxime* $3^2 \approx 10$*, para facilitar as contas.*

b) Calcule o número esperado de lances com n = 0 e n = 8 faces para cima, bem como seus desvios-padrão, quando os estudantes repetem m = 200 vezes o lançamento dos 8 dados.

c) Das três turmas, quais histogramas parecem corresponder a medições realmente efetuadas e quais parecem ter sido inventados? Justifique sua resposta.

3) (3,5 pontos) Um pêndulo é formado por uma esfera de aço presa por um fio de massa muito pequena, com 991 mm de comprimento, de forma que o período **para pequenas oscilações** é 2,000(2) s. Um grupo de estudante faz quatro medidas do período para uma certa amplitude de oscilação e obteve os seguintes resultados:

{1,95; 2,50; 2,30; 2,25}, em s.

.

a) Adote que esses dados obedecem a uma mesma distribuição gaussiana, de desvio-padrão desconhecido e, com a estatística $t = \frac{\bar{x}-x_0}{\sigma_m}$, teste a hipótese do período do pêndulo valer x_0 = 2,000 s e conclua se a rejeita ou não. Explique porque a incerteza desse valor teórico (= 0,002 s) não tem importância nessa aplicação particular do teste *t*.
b) Os estudantes tomam agora um total de 49 dados na mesma situação, e o professor da disciplina sabe que o valor verdadeiro do período, com a amplitude usada, que não é pequena, vale 2,100(2) s e o desvio-padrão *dos dados* é 0,280 s. Determine a probabilidade do teste estatístico levar à rejeição da hipótese *T*=2,000 s nessa medida com 49 dados.

Tabela. Valores críticos t_α para alguns valores comuns do nível de significância, α, em função do número de graus de liberdade, ν, em um teste **bicaudal**. Os níveis de significância estão discriminados na primeira linha, e a primeira coluna dá o número de graus de liberdade a que se referem os valores críticos apresentados em cada linha.

ν \ α	0,05	0,01	0,001
1	12,7	63,7	636
2	4,30	6,97	31,6
3	3,18	5,84	12,9
4	2,78	4,60	8,61
5	2,57	4,03	6,87
10	2,23	3,17	4,59
20	2,09	2,85	3,85
30	2,04	2,75	3,65
∞	1,96	2,58	3,29

Formulário

Método dos mínimos quadrados

As estimativas dos parâmetros de uma função

$$z = af(\vec{x}) + bg(\vec{x})$$

são obtidas, pelo método dos mínimos quadrados, da solução do sistema linear

$$\vec{D} = \mathbf{M}\hat{A}$$

em que essas matrizes são definidas por

$$\begin{pmatrix} \sum \frac{z_i f(\vec{x}_i)}{\sigma_i^2} \\ \sum \frac{z_i g(\vec{x}_i)}{\sigma_i^2} \end{pmatrix} = \begin{pmatrix} \sum \frac{\left(f(\vec{x}_i)\right)^2}{\sigma_i^2} & \sum \frac{f(x_i)g(\vec{x}_i)}{\sigma_i^2} \\ \sum \frac{f(\vec{x}_i)g(\vec{x}_i)}{\sigma_i^2} & \sum \frac{\left(g(\vec{x}_i)\right)^2}{\sigma_i^2} \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$$

Função de probabilidade Binomial

$$P_{N,p}(n) = \binom{N}{n} p^n (1-p)^{N-n}$$

$$\langle n \rangle = Np \quad \text{e} \quad \text{var}(n) = Np(1-p)$$

Propagação de incerteza

$$\text{var}\big(f(\vec{x})\big) \cong \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial f}{\partial x_i}\bigg|_{\vec{x}=\vec{\bar{x}}} \frac{\partial f}{\partial x_j}\bigg|_{\vec{x}=\vec{\bar{x}}} \text{cov}\big(x_i, x_j\big)$$

$$\text{cov}\big(f(\vec{x}), h(\vec{x})\big) \cong \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial f}{\partial x_i}\bigg|_{\vec{x}=\vec{\bar{x}}} \frac{\partial h}{\partial x_j}\bigg|_{\vec{x}=\vec{\bar{x}}} \text{cov}\big(x_i, x_j\big)$$

Quando f é o produto ou quociente dos x_i, é muito vantajoso, em termos de cálculo dos valores numéricos, trabalhar com as variâncias e covariâncias relativas.

Integral da função densidade de probabilidade normal.

y	0	0,5	1	1,5	2	2,5	3
$\int_0^y N(y'; 0{,}1)dy'$	0	0,19	0,34	0,43	0,48	0,49	0,498

Estimativa da variância

A estimativa da variância em uma medida $\{x_i,\ i=1..N\}$, em que os dados obedecem a uma mesma f.d.p. normal, têm distribuição semelhante à de qui-quadrado com N-1 graus de liberdade e desvio-padrão $\sigma_\sigma \approx \frac{\sigma}{\sqrt{2(N-1)}}$.