DEPARTAMENTO DE ESTATÍSTICA - UFMG
PROVA DE ESTATÍSTICA E PROBABILIDADE
SELEÇÃO - MESTRADO/ UFMG - 2013/2014

Instruções:

1. Cada questão respondida corretamente vale **1 (um) ponto**.
2. Cada questão respondida incorretamente vale **-1 (menos um) ponto**.
3. Cada questão deixada em branco vale 0 (zero) pontos
4. Pelo menos 9 (nove) questões devem ser respondidas pelo aluno
5. A nota final será a soma dos pontos (negativos e positivos) de todas as questões
6. As opções escolhidas devem ser assinaladas na folha de respostas no final da prova.

A prova tem duração de 3 horas.

É proibido: usar celular; consultar referências bibliográficas diferentes das que estão estabelecidas no edital de seleção; emprestar calculadoras e/ou livros para consulta de outros candidatos durante a prova

DEPARTAMENTO DE ESTATÍSTICA - UFMG
PROVA DE ESTATÍSTICA E PROBABILIDADE
SELEÇÃO - MESTRADO/ UFMG - 2013/2014

Nome do Candidato(a):

Questão 1. Seja X uma variável aleatória com distribuição gama, $\Gamma(5,2)$. Dado $X = x$, uma variável aleatória Y é escolhida uniformemente no intervalo $[0, x]$. A densidade da distribuição gama, $\Gamma(n, \alpha)$, é dada por

$$f(x) = \frac{1}{\Gamma(n)} \alpha^n x^{n-1} e^{-\alpha x} \quad x > 0, \quad \alpha > 0, \quad n > 0.$$

Indique o valor de $\mathbb{E}(Y)$.

a. 5/2

b. 5/4

c. 5

d. 5/3

Questão 2. O par de variáveis aleatórias discretas (X, Y) tem a função de probabilidade conjunta

$$\mathbb{P}(X = i, Y = j) = \theta^{i+j+1} \text{ se } i, j = 0, 1, 2,$$

para algum $\theta > 0$ e zero em caso contrário. Indique se são falsas ou verdadeiras as afirmações seguintes:

1. $\theta + 2\theta^2 + 3\theta^3 + 2\theta^4 + \theta^5 = 1$
2. $\mathbb{E}(XY) = \theta^3 + 4\theta^4 + 4\theta^5$
3. $\mathbb{E}(X) = \theta^2 + 3\theta^3 + 3\theta^4 + 2\theta^5$

a. VFF

b. VVF

c. VVV

d. VFV

Questão 3. Considere a variável aleatória contínua X com distribuição exponencial e densidade dada por

$$f(x) = \lambda e^{-\lambda x} \quad x > 0, \quad \lambda > 0.$$

A função geradora de momentos de uma variável aleatória, se existir, é dada por $\mathbb{E}[e^{tX}]$, onde $t \in \mathbb{R}$. Considere $s > 0$ e $\mathbb{E}[e^{-sX}] = 0,25$, indique o valor de $\mathbb{E}[e^{-2sX}]$.

a. 1

b. 1/3

c. 1/5

d. $1/7$

Questão 4. Sejam U e V variáveis aleatórias normais independentes com média 0 e variância 1. Definimos

$$X = U, \qquad Y = \rho U + (\sqrt{1-\rho^2}\,)V, \quad |\rho| < 1.$$

Indique o valor de $Var(X+Y)$.

a. $2(1-\rho)$

b. $2(1+\rho)$

c. $2+\rho$

d. $2-\rho$

Questão 5. Sejam $X_1, \cdots, X_n$ uma amostra aleatória da variável aleatória X com distribuição uniforme no intervalo $[0, \theta]$, com $\theta > 0$. Assuma $T = \max\{X_1, \cdots, X_n\}$ e considere um estimador para θ, definido por $\delta = aT$, onde a é uma constante real. Encontre o valor de a que minimiza a função risco deste estimador. A função risco é dada por

$$R(\theta, \delta) = \mathbb{E}_\theta(aT - \theta)^2.$$

a. $\dfrac{n+2}{n+1}$

b. $\dfrac{n+2}{n+3}$

c. $\dfrac{n+1}{n+2}$

d. $\dfrac{n}{n+1}$

Questão 6. Se U é uma variável aleatória uniforme no intervalo $[0,1]$, indique a fórmula certa para simular uma variável aleatória Pareto de tipo I. A densidade da variável Pareto de tipo I é dada por

$$f(x) = \begin{cases} \dfrac{\alpha}{\theta}\left(\dfrac{\theta}{x}\right)^{\alpha+1} & \text{se} \quad x \geq \theta, \quad \alpha > 0, \quad \theta > 0, \\ 0 & \text{se} \quad x < \theta, \end{cases}$$

a. $\theta(1-U)^{1/\alpha}$

b. $\theta^{-1}(1-U)^{-1/\alpha}$

c. $\theta(1-U)^{-1/\alpha}$

d. $\theta^{-1}\left(\dfrac{\alpha}{\theta U}\right)^{1/(\alpha+1)}$

Questão 7. Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória da distribuição geométrica com função distribuição de probabilidade $P(X = x) = \theta\,(1-\theta)^{x-1}$ para $x = 1, 2, \ldots$, sendo $0 < \theta < 1$. O limite inferior da desigualdade de Cramer-Rao para a variância de estimadores não viciados de $1-\theta$ é dado por:

a. $\dfrac{n}{\theta(1-\theta)^2}$

b. $\dfrac{\theta(1-\theta)^2}{n}$

c. $\dfrac{\theta^2(1-\theta)}{n}$

d. $\dfrac{n}{\theta^2(1-\theta)}$

Questão 8. Para responder esta questão, a seguinte informação pode ser útil:

- Z_δ = quantil da $N(0,1)$ com área δ à esquerda.
  $Z_{0,99} = 2,32$; $Z_{0,95} = 1,64$; $Z_{0,90} = 1,28$.
- $t_{\delta;\nu}$ = quantil da t-Student com ν graus de liberdade (área δ à esquerda).
  $t_{0,99;20} = 2,52$; $t_{0,95;20} = 1,72$; $t_{0,90;20} = 1,32$.
- $\chi^2_{\delta;\nu}$ = quantil da Qui-Quadrado com ν graus de liberdade (área δ à esquerda).
  $\chi^2_{0,99;20} = 37,56$; $\chi^2_{0,95;20} = 31,41$; $\chi^2_{0,90;20} = 28,41$;
  $\chi^2_{0,01;20} = 8,26$; $\chi^2_{0,05;20} = 10,85$; $\chi^2_{0,10;20} = 12,44$.
- $F_{\delta;\nu_1,\nu_2}$ = quantil da F_{ν_1,ν_2} com área δ à esquerda.
  $F_{0,99;1,20} = 8,09$; $F_{0,95;1,20} = 4,35$; $F_{0,90;1,20} = 2,97$;
  $F_{0,01;1,20} = 0,0001$; $F_{0,05;1,20} = 0,0040$; $F_{0,10;1,20} = 0,0162$.

Suponha que $X_1, X_2, \ldots, X_n$ seja uma amostra aleatória proveniente da distribuição $N(\mu, \sigma^2)$ com média μ conhecida e variância σ^2 desconhecida. Desejamos testar as hipóteses: $H_0 : \sigma^2 \geq 11$ contra $H_1 : \sigma^2 < 11$. Qual é o procedimento de teste uniformemente mais poderoso assumindo que $n = 20$ e $\alpha = 0,01$?

a. Rejeitar H_0 se $\sum_{i=1}^{20}(X_i - \mu)^2 < 119,35$.

b. Rejeitar H_0 se $\sum_{i=1}^{20}(X_i - \mu)^2 < 90,86$.

c. Rejeitar H_0 se $\sum_{i=1}^{20} X_i^2 < 9,53$.

d. Rejeitar H_0 se $\sum_{i=1}^{20} X_i^2 < 10,92$.

Questão 9. Considere que: $X_1, X_2, \ldots, X_n$ é uma amostra aleatória da distribuição exponencial com média $\theta_1 > 0$, e $Y_1, Y_2, \ldots, Y_n$ é uma amostra aleatória da exponencial com média $\theta_2 > 0$. Assuma também que X_i e Y_j são independentes para todo $i, j = 1, 2, \ldots, n$. Desejamos testar as hipóteses: $H_0 : \theta_1 = \theta_2$ contra $H_1 : \theta_1 \neq \theta_2$. O teste da razão de verossimilhanças, para este caso, rejeitará H_0 se:

a. $\dfrac{\bar{X}\bar{Y}}{(\bar{X}+\bar{Y})^{1/2}} < c$ onde c é uma constante.

b. $\dfrac{\bar{X}\bar{Y}}{\bar{X}+\bar{Y}} < c$ onde c é uma constante.

c. $\dfrac{(\bar{X}\bar{Y})^2}{\bar{X}+\bar{Y}} < c$ onde c é uma constante.

d. $\dfrac{\bar{X}\bar{Y}}{(\bar{X}+\bar{Y})^2} < c$ onde c é uma constante.

Questão 10. Para responder esta questão, a seguinte informação pode ser útil:

- Z_δ = quantil da $N(0,1)$ com área δ à esquerda.
  $Z_{0,99} = 2,32$; $Z_{0,95} = 1,64$; $Z_{0,90} = 1,28$.
- $t_{\delta;\nu}$ = quantil da t-Student com ν graus de liberdade (área δ à esquerda).
  $t_{0,99;14} = 2.62$; $t_{0,95;14} = 1.76$; $t_{0,90;14} = 1.34$.
- $\chi^2_{\delta;\nu}$ = quantil da Qui-Quadrado com ν graus de liberdade (área δ à esquerda).
  $\chi^2_{0,99;14} = 29,14$; $\chi^2_{0,95;14} = 23,68$; $\chi^2_{0,90;14} = 21,06$.
- $F_{\delta;\nu_1,\nu_2}$ = quantil da F_{ν_1,ν_2} com área δ à esquerda.
  $F_{0,99;2,12} = 6,92$; $F_{0,95;2,12} = 3,88$; $F_{0,90;2,12} = 2,81$;
  $F_{0,99;3,11} = 6,21$; $F_{0,95;3,11} = 3,58$; $F_{0,90;3,11} = 2,66$.

Um experimento foi conduzido para investigar o efeito da cor do papel (azul, verde ou laranja) sobre a proporção de questionários respondidos. Quinze supermercados foram selecionados em uma cidade e formou-se aleatoriamente 3 grupos de cinco estabelecimentos. Em cada grupo utilizou-se uma das cores no papel do questionário. Os dados (em %) são mostrados na tabela abaixo.

		Supermercado				
	Cor	1	2	3	4	5
1	Azul	28	26	31	27	35
2	Verde	34	29	25	31	29
3	Laranja	31	25	27	29	28

O modelo assumido para este conjunto de dados é: $Y_{ij} = \mu_i + \epsilon_{ij}$ com μ_i representando a média do tratamento $i = 1,2,3$ e ϵ_{ij} representando o erro da observação $j = 1,2,3,4,5$ do tratamento i. Outras informações sobre este conjunto de dados são fornecidas a seguir:

Fonte de variação	Soma de quadrados
Tratamento	7,60
Erro	116,40
Total	124,00

Assumindo um nível de significância de 10%, desejamos testar se há diferença entre as porcentagens médias de respostas de cada cor. Este teste terá a seguinte conclusão:

a. Não rejeitamos H_0, pois temos: estatística de teste $= 0,239$ menor que o ponto crítico.

b. Não rejeitamos H_0, pois temos: estatística de teste $= 0,392$ menor que o ponto crítico.

c. Rejeitamos H_0, pois temos: estatística de teste $= 2,55$ maior que o ponto crítico.

d. Rejeitamos H_0, pois temos: estatística de teste $= 4,17$ maior que o ponto crítico.

Questão 11. Seja Z_n uma variável aleatória com distribuição de Poisson com média n, com $n \in \mathbb{N}$. O limite em distribuição de $Y_n = (Z_n - n)/\sqrt{n}$, quando $n \to \infty$,

a. é uniforme no intervalo $(-1,1)$.

b. é t-student com média 0 e variância 1.

c. é normal com média 0 e variância 1.

d. não existe.

Questão 12. Sejam $X_1, \ldots, X_n$ variáveis aleatórias independentes e identicamente distribuídas uniformemente no intervalo $[0, \theta]$, $\theta > 0$. Os estimadores de máxima verossimilhança e do método dos momentos de θ são respectivamente

a. $\max\{X_1, \ldots, X_n\}$ e $2\sum_{i=1}^{n} X_i/n$.

b. $\max\{X_1, \ldots, X_n\}$ e $\sum_{i=1}^{n} X_i/n$.

c. $\sum_{i=1}^{n} X_i/n$ e $\max\{X_1, \ldots, X_n\}$.

d. $\prod_{i=1}^{n} X_i$ e $2\sum_{i=1}^{n} X_i/n$.

Questão 13. Seja Y uma variável aleatória com distribuição binomial com parâmetros $n = 300$ e $p \in (0, 1)$. Se o valor observado de Y é $y = 75$, um intervalo de confiança para p com cobertura aproximada de 90% é

a. $(0,209; 0,291)$.

b. $(0,229; 0,271)$.

c. $(0,188; 0,313)$.

d. $(0,178; 0,323)$.

Questão 14. Considere uma fábrica que produz um certo tipo de peça e defina p como sendo a proporção de peças defeituosas. Suponha que 500 peças sejam testadas na fabricação e 10 sejam defeituosas. Suponha que estamos interessados em testar a hipótese $H_0 : p = 0,03$ contra $H_1 : p < 0,03$. O p-valor neste caso fica dado por

a. 0,012.

b. 0,051.

c. 0,065.

d. 0,095.

Questão 15. Sejam $X_1, \ldots, X_n$ uma amostra aleatória de uma distribuição com função de densidade dada por $f(x) = \theta x^{\theta-1}$, $0 < x < 1$, $\theta > 0$ e $f(x) = 0$, caso contrário. O estimador de máxima verossimilhança e uma estatística suficiente para θ são respectivamente

a. $\max\{X_1, \ldots, X_n\}$ e $\prod_{i=1}^{n} X_i$.

b. $-\sum_{i=1}^{n} \log X_i/n$ e $\sum_{i=1}^{n} X_i$.

c. $\prod_{i=1}^{n} X_i$ e $-n/\sum_{i=1}^{n} \log X_i$.

d. $-n/\sum_{i=1}^{n} \log X_i$ e $\prod_{i=1}^{n} X_i$.