1 Elementos da Teoria de Decisão

1 Seja $X_1, \ldots, X_n$ $a.a.N(\mu, \sigma^2)$, $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$,

i) Demonstre que $\overline{X}$ e S^2 são independentes.

ii) Mostre que $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$; $\quad \frac{(n-1)}{\sigma^2}S^2 \sim \chi^2_{n-1}$.

iii) Considere a família de decisões $\mathcal{A} := \{\delta_b(x) : \delta_b(x) = bS^2, b \in \mathbb{R}^+\}$ para estimar σ^2. Determine a regra $\delta_{b^*} \in \mathcal{A}$ que minimiza o risco quadrâtico.

iv) Determine o risco quadrâtico das regras de decisão $\delta_1 \in \mathcal{A}$ e $\delta_{\frac{n-1}{n}} \in \mathcal{A}$ e compare graficamente as funções de risco das decisões δ_{b^*}, δ_1, $\delta_{\frac{n-1}{n}}$, como funções de σ^2.

2 Seja $X_1, \ldots, X_n$ $a.a.$ de media μ e variância σ^2. Mostre que a função de risco da função de perda l

$$l(\sigma^2, a) := \frac{a}{\sigma^2} - 1 - \ln\frac{a}{\sigma^2},$$

considerada na família de decisões $\mathcal{A} := \{\delta_b(x) : \delta_b(x) = bS^2, b \in \mathbb{R}^+\}$ é minimizada em $b = 1$.

3 Seja θ uma v.a. contínua de densidade $f(\cdot)$ e f.d.a. $F(\cdot)$. Se δ é uma regra de decisão adotada para "estimar" θ, determine a decisão ótima δ^* que minimiza a perda media

$$E_\Theta(l(\delta, \Theta))$$

onde l apresenta a seguinte forma, para a e b constantes positivas,

$$l(\theta, \delta) := \begin{cases} a(\delta - \theta) & \text{se} \quad \delta \leq \theta \\ b(\theta - \delta) & \text{se} \quad \theta < \delta \end{cases}.$$

4 Seja θ uma v.a. contínua de densidade $f(\cdot)$ e f.d.a. $F(\cdot)$. Se δ é uma regra de decisão adotada para "estimar" θ, determine a decisão ótima δ^* que minimiza a perda media

$$E_\Theta(l(\delta, \Theta))$$

onde l é uma perda 0-1 e apresenta a seguinte forma, para b constante positiva,

$$l(\theta, \delta) := \begin{cases} 0 & \text{se} \quad |\delta - \theta| \leq b \\ 1 & \text{se} \quad |\delta - \theta| > b \end{cases}.$$

5 Seja θ uma v.a. e seja $l(\theta, \delta)$ uma função perda, construa a função perda $l^*(\theta, \delta) := al(\theta, \delta) + g(\theta)$ e demonstre que l e l^* apresentam a mesma decisão ótima. Onde a é uma constante positiva e $g(\cdot)$ uma aplicação limitada.

6 Assuma que θ é uma v.a.. Se seu interesse for "estimar" uma posição central para θ usando uma função de perda l demonstre que

  i) se $l(\theta, \delta) := \frac{(\delta-\theta)^2}{\theta}$, então a decisão ótima é $\{E(\theta^{-1})\}^{-1}$

  ii) se $l(\theta, \delta) := \frac{(\delta-\theta)^2}{\delta}$, então a decisão ótima é $\{E(\theta^2)\}^{1/2}$

7 Seja (θ, ϕ) um vetor aleatório de densidade conjunta $f(\theta, \phi)$. Usando como função perda media a Medida de Kullback Liebler, decisões do tipo $\delta(\theta, \phi) = \delta_1(\theta)\delta_2(\phi)$ e assumindo a independência entre θ e ϕ, demonstre que a decisão que minimiza a medida de Kullback Liebler é aquela definida por $\delta_1(\theta) = f(\theta)$ (densidade marginal de θ) e $\delta_2(\phi) = f(\phi)$ (marginal de ϕ).

8 Resolva os exercícios da seção 1.3, pag. 75 do Bickel.

9 Desenvolva os detalhes do exemplo 1.4.1 do texto do Bickel (pag. 35). Calcule neste caso para cada estado z (capacidade de funcionamento) o melhor preditor de Y (falhas no funcionamento). Posteriormente, determine o risco de predição.

1.1 Otimalidade (veja o livro do Bickel) (ainda faltam exercícios)

1 Seção 3.2-pag 161: Exercícios pag 197: 1; 2; 3;

2 Resolva o exercício 4 pag 197(sec 3.2)

  2.1 Compare com a solução do exercício 4 pag 139(sec 2.1)

  2.2 Compare com a solução do exercício 16 da pag. 144 (sec. 2.2))

  2.3 Resolva o exercício 4 (pag. 197)

3 Seção 3.3-pag 170: Exercícios pag 199: 3; 4; 5; 7; 12.

4 Seção 3.4-pag 179: Exercícios pag 203: 1; 10; 12; 18; 20; 22.

2 Modelos Bayesianos, princípios e conceitos básicos

1 Considere as observações x_1 e x_2 condicionalmente independentes dado (θ_1, θ_2) com distribuições $N(\theta_i, 1)$, $i = 1, 2$; respectivamente. Suponha que a distribuição a priori para (θ_1, θ_2) é impropria Uniforme (i.e.$f(\theta_1, \theta_2) \propto 1$) sempre que $\theta_1 > \theta_2$. Mostre que

$$f(\theta_1|x_1, x_2) \propto \phi(\theta_1 - x_1)\Phi(\theta_1 - x_2),$$

onde $\phi(\cdot)$ e $\Phi(\cdot)$ representam a f.d.p. e a f.d.a. respectivamente, da distribuição normal padrão.

2 Seja x o número de sucessos em n realizações independentes com probabilidade de sucesso θ em cada realização. A distribuição a priori (impropria) $f(\theta) \propto \theta^{-1}(1-\theta)^{-1}$ é assumida. Demonstre que isto corresponde a considerar uma distribuição a priori uniforme para $\ln(\frac{\theta}{1-\theta})$.

3 Determine a distribuição a posteriori para θ, assumindo que x é $N(\theta, 1)$ dado θ e que a priori para θ é uma exponencial dupla: $f(\theta) \propto \exp(-|\theta|)/2$.

4 Se x possui distribuição de Poisson com valor de media θ desconhecido e a densidade a priori para θ for uma Gama de parâmetros conhecidos a e b. Demonstre que a distribuição a posteriori de θ resulta uma Gama de parâmetros conhecidos $a+1$ e $b+x$,

$$f(x/\theta) = \frac{\theta^x e^{-\theta}}{x!}, \;\; x = 0, 1, \cdots$$

$$f(\theta) = \frac{a^b}{\Gamma(b)}\theta^{b-1}e^{-a\theta}, \;\theta > 0$$

a e b constantes positivas.

5 Se $x_1, \cdots, x_n$ for uma amostra aleatória proveniente da Poisson com media θ desconhecido e θ for uma variável aleatória com distribuição a priori Gama de parâmetros conhecidos a e b. Demonstre que a posteriori de θ corresponde a uma Gama de parâmetros $n + a$ e $\sum_{i=1}^{n} x_i + b$. Demonstre ainda que $E(\theta/x)$ é media ponderada entre $\overline{x}$ e o valor esperado de θ a priori. Onde $x = (x_1, \cdots, x_n)$.

6 Seja θ de função de distribuição F com densidade f dada por

$$f(\theta) = \frac{m^{m/2} n^{n/2}}{B(m/2, n/2)} \theta^{m/2-1} (n + m\theta)^{-(n+m)/2}$$

demonstre que f possui uma única moda em $\theta_1 = \frac{n(m-2)}{m(n+2)}$ e que os pontos de inflexão da distribuição estão a uma distancia $\frac{n\{(m-2)(n+m)\}^{1/2}}{m(n+2)}$ da moda (para direita e esquerda).

7 Considere um dado x desde a $Exp(\lambda)$ e outro dado y desde $Poisson(\lambda)$, i.e.

$$f(x/\lambda) = \lambda e^{-\lambda x},\ x > 0, \qquad f(y/\lambda) = \frac{\lambda^y e^{-\lambda}}{y!},\ y = 0, 1, \cdots.$$

Sabendo que λ é desconhecido portanto possui uma priori $f(\lambda)$, ache os valores de x e y para os quais se verifica o Principio da Verossimilhança. Interprete.

8 Seção 1.2-pag 12 do Bickel: exercícios da pag 71 do Bickel: 1; 2; 3; 4

3 Decisão (Seção 1.3-pag 16 do Bickel)

Exercícios da pag 75 do Bickel: 1; 2 (excluir item c) por enquanto); 3; 5; 9

4 Métodos de Estimação

Seção 2.1.1-pag 99 do Bickel: Exercícios pag 138 - Bickel: 2; 3; 4; 5 (a),b),c)); 9; 11; 13;
Seção 2.2-pag 107 do Bickel: Exercícios pag 142 - Bickel: 5; 9; 10; 11; 15; 16; 27; 37; 38
Seção 2.3-pag 121 do Bickel: Exercício 27 (pag 91)e Exercícios pag 152 - Bickel: 2; 8; 12.

5 Suficiência e Principios da Inferência

1 Sejam T_1 e T_2 duas estatísticas.

1.1 Se $T_1 = g(T_2)$ e T_1 é suficiente para θ, então T_2 é suficiente para θ.

1.2 Se $T_1 = g(T_2)$, g aplicação bijetora, então T_1 é suficiente para θ se, e somente se T_2 é suficiente para θ.

2 Usando a noção de equivalência introduzida no livro do Lehmann e Casella (TPE), pag 36, mostre que as estatísticas apresentadas no exemplo 6.9 da pag. 36 são equivalentes.

3 Resolva os problemas 6.5 e 6.9 (da pag. 69-70 Lehmann e Casella TPE), citados citados nas paginas 36 e 37 desse texto.

4 Seção 1.5-pag 41 do Bickel: Exercícios da pag 84 do Bickel: 1; 2; 3; 4; 5; 7

5 Demonstre "por definição" que a estatística $T = \sum_{i=1}^{n} X_i$ é suficiente minimal para θ, onde $X_i \sim \text{Ber}(\theta)$, $X_1, \cdots X_n$ a.a.

6 Sejam $X_1, \cdots X_n$ i.i.d. U(θ,θ+1), $\theta \in \mathbb{R}$.

6.1 Mostre que as estatísticas $T_1 := (X_{(1)}, X_{(n)})$ e $T_2 := (X_{(n)} - X_{(1)}, \frac{X_{(n)}+X_{(1)}}{2})$ são suficientes minimais para θ.

6.2 Mostre que $R = X_{(n)} - X_{(1)}$ é ancilar para θ.

6 Formas Exponenciais(Seção 1.6-pag 49 do Bickel)

Exercícios da pag 87 do Bickel: 3; 4; 5; 11; 21; 23

1. Definição: Supondo que x é uma observação desde $f(x/\theta)$. A família $\mathcal{F}$ de prioris $f(\theta)$ é dita fechada sobre amostras desde $f(x/\theta)$ se para toda priori $f(\theta) \in \mathcal{F}$, $f(\theta/x) \in \mathcal{F}$.

   1.1. Demonstre que se x for Binomial ou Binomial Negativa, i.e.

$$f(x/\theta) \propto \theta^x (1-\theta)^{n-x}$$

   Então a família de distribuições Beta ($\mathcal{F}$) é fechada sobre amostras desde $f(x/\theta)$.

1.2. Demonstre que se x for uniformemente distribuida dado θ, ou seja, $f(x/\theta) = \theta^{-1},\ 0 < x < \theta,\ \theta > 0$

$$\mathcal{F} := \{g(\theta; a, r) : 0 \leq a, 0 < r\}, \quad g(\theta; a, r) := r a^r \theta^{-r-1} I_{(a,\infty)}(\theta),$$

então $\mathcal{F}$ é fechada sobre amostras desde $f(x/\theta)$, logo a posteriori $f(\theta/x)$ assume a forma

$$f(\theta/x) = g(\theta; a^*, r+1), \quad a^* = \max\{a, x\}.$$

2. Suponha que $f(x/\theta)$ pertence a família exponencial $k-$paramétrica, com densidade

$$f(x/\theta) = \exp\left\{\sum_{j=1}^{k} A_j(\theta) B_j(x) + C(x) + D(\theta)\right\}.$$

Demonstre que a família $\mathcal{F}$ cujos membros tem a seguinte forma

$$f(\theta; a_1, a_2, \cdots, a_k, d) = \exp\left\{\sum_{j=1}^{k} a_j A_j(\theta) + dD(\theta) + c(a_1, a_2, \cdots, a_k, d)\right\}$$

onde a constante de normalização resulta

$$c(a_1, a_2, \cdots, a_k, d) = -\ln\left\{\int \exp\left\{\sum_{j=1}^{k} a_j A_j(\theta) + dD(\theta)\right\} d\theta\right\}$$

é fechada sobre amostras desde $f(x/\theta)$. Logo se a priori de θ tiver a forma $f(\theta; a_1, a_2, \cdots, a_k, d)$ então a posteriori de θ, $f(\theta/x)$ tera a forma $f(\theta; a_1 + B_1(x), a_2 + B_2(x), \cdots, a_k + B_k(x), d+1)$. Neste caso a família $\mathcal{F}$ é dita *Conjugada Natural* de $f(x/\theta)$.

3. Determine a família Conjugada Natural para $f(x/\theta)$ quando $f(x/\theta)$ corresponde a uma $Ber(\theta)$, $Bin(n, \theta)$, $Exp(\theta)$, $N(\mu, \sigma^2)$; neste último caso considere $\theta = (\mu, \sigma^2)$.

4. Determine a familia conjugada natural associada com as seguintes verossimilhanças:

   - $f(x|\theta)$ dada por uma distribuição $\mathrm{N}(\theta, \sigma^2)$, σ^2 conhecido;
   - $f(x|\theta)$ dada por uma distribuição Poisson(θ);

- $f(x|\theta)$ dada por uma distribuição Gama(ν, θ);
- $f(x|\theta)$ dada por uma distribuição NegBin(m, θ);
- $f(x|\theta)$ dada por uma distribuição Multinom$(\theta_1, \cdots, \theta_k)$;
- $f(x|\theta)$ dada por uma distribuição N$(\mu, 1/\theta)$.

5. Assuma $f(x|\theta, \phi) = \theta^\phi x^{\phi-1} e^{-\theta x}/\Gamma(\phi)$ onde $\theta > 0, \phi > 0$. Demonstre que na familia conjugada natural associada a (θ, ϕ) a densidade marginal de ϕ apresenta a forma $f(\phi) \propto c^\phi \Gamma(1 + d\phi)/(\Gamma(\phi))^d$, onde c e d representam valores conhecidos.

6. A distribuição a posteriori derivada desde uma priori impropria, pode resultar numa posteriori impropria:
   Se n e p são as quantidades aleatorias de interesse e x representa o numero de sucessos em n ensaios independentes (com probabilidade de sucesso p). Assumindo ainda, que a priori $\pi(n, p) = \pi(n)\pi(p)$ e adicionalmente, que $\pi(p)$ é dada como sendo uma $U(0, 1)$. Por outro lado, considerando $\pi(n) \propto 1$, $\forall n \in \mathbb{R}$ (ou seja uma priori impropria). Demonstre que a posteriori para n é impropria.

6.1 Estimação e Modelo Normal

Dica: Sugestão de referência: Christian P. Robert (1994). The Bayesian Choice: A Decision - Theoretic Motivation. Springer-Verlag. New York.

1. Verifique os seguintes resultados, considerando uma perda quadratica sob a quantidade de interesse θ.

2. Considere uma amostra $x_1, \cdots, x_n$ i.i.d. N(θ, σ^2) onde (θ, σ^2) é a quantidade de interesse. Assuma a priori não informativa de Jeffreys π e determine,

   - A distribuição de $\theta|\sigma, \overline{x}, s^2$.
   - A distribuição de $\sigma^2|\overline{x}, s^2$.
   - A distribuição de $\theta|\overline{x}, s^2$.

   Onde, $\overline{x} = \frac{\sum_{i=1}^n x_i}{n}$, $s^2 = \sum_{i=1}^n (x_i - \overline{x})^2$.

Tabela 1: Conjugada Natural e Estimador de Bayes

Verossimilhança	Conjugada	Estimador
$\mathrm{N}(\theta,\sigma^2)$	$\mathrm{N}(\mu,\tau^2)$	$\frac{\mu\sigma^2+\tau^2 x}{\sigma^2+\tau^2}$
Poisson(θ)	Gama(α,β)	$\frac{\alpha+x}{\beta+1}$
Gama(ν,θ)	Gama(α,β)	$\frac{\alpha+\nu}{\beta+x}$
Bin(n,θ)	Beta(α,β)	$\frac{\alpha+x}{\alpha+\beta+n}$
NegBin(n,θ)	Beta(α,β)	$\frac{\alpha+n}{\alpha+\beta+x+n}$
Multi$(n;\theta_1,\cdots,\theta_k)$	Dirichlet$(\alpha_1,\cdots,\alpha_k)$	$\frac{\alpha_i+x_i}{(\sum_j\alpha_j)+n}$
$\mathrm{N}(\mu,1/\theta)$	Gama$(\alpha/2,\beta/2)$	$\frac{\alpha+1}{\beta+(\mu-x)^2}$

6.2 Priori-Posteriori

1. Demonstre que se $x\sim\mathrm{N}(\theta,\theta^2)$. A familia de distribuições $\mathcal{F}$, $\mathrm{IN}(\alpha,\mu,\tau)$ é conjugada. Onde a familia $\mathcal{F}$ é dada por $\mathcal{F}=\left\{\pi(\theta)=|\theta|^{-\alpha}\exp\left\{-(\frac{1}{\theta}-\mu)^2/(2\tau^2)\right\}\right\}$. Onde α,μ,τ representam quantidades fixas e conhecidas.

2. Seja $x=(x_1,\cdots,x_n)$ uma amostra aleatoria desde $\mathrm{N}(\mu,\sigma^2)$, onde μ e σ^2 são desconhecidos. Considere $\theta=(\mu,\sigma^2)$.

   2.1. Demonstre que $f(x|\theta)=(2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}(s^2+n(\overline{x}-\mu)^2)\right\}$ onde $s^2=\sum_{i=1}^n(x_i-\overline{x})$.

   2.2. Demonstre que se considera uma densidade a priori $f(\theta)=k(a,b,w)(\sigma^2)^{-(b+3)/2}\exp\left\{-\frac{1}{2\sigma^2}(a+(\mu-m)^2/w)\right\}$ onde $k(a,b,w)=a^{b/2}2^{-(b+1)/2}(\pi w)^{-1/2}\left\{\Gamma(\frac{1}{2}b)\right\}^{-1}$ é uma constante, assim como a,b e w. Então a distribuição a posteriori, resulta: $f(\theta|x)=k(a_1,b_1,w_1)(\sigma^2)^{-(b_1+3)/2}\exp\left\{-\frac{1}{2\sigma^2}(a_1+(\mu-m_1)^2/w_1)\right\}$, onde $w_1=\frac{w}{1+nw}$, $m_1=\frac{m+nw\overline{x}}{1+nw}$, $b_1=b+n$, $a_1=a+s^2+\frac{n(\overline{x}-m)^2}{1+nw}$.

   2.3. É possível determinar a familia conjugada de distribuições associada a forma 2.1.?

   2.4. Demonstre que $f(\mu|x)=k_\mu(t_1,b_1)\left\{1+(\mu-m_1)^2/(b_1t_1)\right\}^{-(b_1+1)/2}$ onde $t_1=w_1a_1/b_1$ e $k_\mu(t_1,b_1)=(t_1b_1)^{-1/2}\left\{B(\frac{1}{2},\frac{1}{2}b_1)\right\}^{-1}$.

Ou seja $(\mu - m_1)/\sqrt{t_1} \sim$ t-Student (b_1 graus de liberdade).
Dica: Considere a substituição $y = \frac{1}{2\sigma^2}(a_1 + (\mu - m_1)^2/w_1)$, na forma $f(\theta|x)$ em 2.2.

2.5. Demonstre que $f(\sigma^2|x) = k_{\sigma^2}(a_1, b_1)(\sigma^2)^{-(b_1+2)/2}\exp(-\frac{a_1}{2\sigma^2})$ onde $k_{\sigma^2}(a_1, b_1) = (\frac{1}{2}a_1)^{b_1/2}\left\{\Gamma(\frac{1}{2}b_1)\right\}^{-1}$.
Ou seja $a_1/\sigma^2 \sim$ Qui-Quadrado (b_1 graus de liberdade).

7 Teste de Hipótese

1 Denotemos por μ a verdadeira média de nível de radioatividade (picocuries por litro). O valor 5 pCi/L é considerado como linha divisória entre água segura e não segura. Qual dos seguintes testes recomenda conduzir?
$H_0 : \mu = 5$ vs $H_1 : \mu > 5$
$H_0 : \mu = 5$ vs $H_1 : \mu < 5$
Explique seu raciocínio em termos dos erros tipo I e II.

2 Os dados correspondem a uma distribuição $Bin(n, p)$. Conduza o seguinte teste
$H_0 : p = 0.75$ vs $H_1 : p < 0.75$. Assuma $n = 150$.
a)Se a $p^* = 0.72$ for uma estimativa pontual de p. Determine a força da evidência contida nos dados (π-value).
b)Para o nível de significância $\alpha = 0.01$ determine a região de rejeição do teste. Verifique se a entimativa p^* apresenta evidência suficiente para rejeitar H_0 ao nível α. Verifique se p^* pertence à região de rejeição do teste para o nível α.

3 Suspeita-se da honestidade de um dado de 6 faces. Procurando suporte para tal afirmação considera-se o número de vezes que a face 2 é obtida numa seqüência de n lançamentos independentes.
a)Determine a hipótese nula H_0 e a alternativa H_1.
b)Em $n = 20$ lançamentos independentes obtem-se 2 vezes a face 2. Calcule a força da evidência contida nos dados e responda: Para qué níveis de significância α, a hipótese H_0 é rejeitada?. Interprete.
Calcule o π-value utilizando a aproximação normal e responda: Para qué níveis de significância α, a hipótese H_0 é rejeitada?. Compare.
c)Em $n = 20$ lançamentos independentes obtem-se 6 vezes a face 2.

Calcule a força da evidência contida nos dados e determine se os dados resultam significantes ao nível $\alpha = 0.10$.

4 Membros de uma associação profissional desejam provar que menos da metade dos eleitores apoiam as medidas tomadas pela equipe econômica do governo para enfrentar a crise financeira internacional. Seja p a proporção de eleitores que apoiam as medidas.
a)Determine a hipótese nula e a alternativa de um teste que permita avaliar a situação.
b)Se uma pesquisa com 500 eleitores selecionada ao acaso revela que 228 apoiam as medidas econômicas, podemos dizer que os dados são significantes ao nível $\alpha = 0.05$?.

5 Suponha que um processo de produção é considerado fora de controle se mais do 3% dos seus produtos resultam defeituosos. Para controlar o processo, de 4 em 4 horas uma amostra ao acaso de 100 produtos é inspecionada.
a)Quantos produtos defeituosos precisamos encontrar numa inspeção para poder concluir que há evidência ao nível $\alpha = 0.05$, do que o processo esta fora de controle?. Esta é a região crítica do teste para $\alpha = 0.05$.
b)Qual sería a região crítica do teste para $\alpha = 0.05$, se no lugar de 100 produtos fossem inspecionados somente 10 produtos?.

6 Membros de uma associação patronal desejam demonstrar que mais do 60% dos seus associados apoiam a política de privatização do governo. Determine a região crítica do teste de hipótese para essa situação, para um nível de significância $\alpha = 0.05$, supondo que os dados são colhidos desde uma amostra com 80 associados selecionados ao acaso.

7 Seja X uma variável aleatória de distribuição Bernoulli, onde
$P(X = 1) = \theta = 1 - P(X = 0)$.
a)Considere uma amostra de tamanho $n = 10$ e o teste
$H_0 : \theta \leq \frac{1}{2}$ vs $H_1 : \theta > \frac{1}{2}$. Assuma a região critica $\left\{6 \leq \sum_{i=1}^{n} x_i\right\}$.

i) Calcule a função poder do teste

ii) Qual o tamanho do teste?

b)Para uma amostra aleatória de tamanho $n = 10$.

i) Calcule o teste mais poderoso de tamanho α ($\alpha = 0.0547$) de $H_0 : \theta = \frac{1}{2}$ vs $H_1 : \theta = \frac{1}{4}$

ii) Calcule o poder do teste mais poderoso em $\theta = \frac{1}{4}$

c)Para uma amostra aleatória de tamanho $n = 10$ e considerando o teste $H_0 : \theta = \frac{1}{2}$ vs $H_1 : \theta = \frac{1}{4}$.

i) Calcule o teste minimax associado à função de perda l $0 = l(a_1, \theta_0) = l(a_2, \theta_1)$, $l(a_1, \theta_1) = 1719$, $l(a_2, \theta_0) = 2241$ onde a_1 representa a decisão de aceitar H_0 e a_2 representa a decisão de rejeitar H_0. θ_0 corresponde ao valor na hipótese nula e θ_1 representa o valor de θ na alternativa.

ii) Compare o risco máximo do teste minimax com o risco máximo do teste mais poderoso obtido na parte b).

8 Seja X uma variável aleatória de densidade $f(x;\theta) = \theta x^{\theta-1} I_{(0,1)}(x)$.

a) Considere o teste $H_0 : \theta \leq 1$ vs $H_1 : \theta > 1$, uma amostra de tamanho 2 e uma região critica $C^* = \left\{(x_1, x_2) : 3/4x_1 \leq x_2\right\}$. Calcule a função poder e o tamanho do teste.

b) Para uma amostra aleatória de tamanho 2 calcule o teste mais poderoso de tamanho $\alpha = \frac{1}{2}(1 - \ln(2))$ de $H_0 : \theta = 1$ vs $H_1 : \theta = 2$.

c) Algum dos testes antes obtidos (a), b)) é não viciado?.

d) Para uma amostra aleatória de tamanho 2 e considerando o teste $H_0 : \theta = 1$ vs $H_1 : \theta = 2$, calcule o teste minimax associado à função de perda l onde $l(d, \theta)$ assume os seguintes valores: $0 = l(a_1, \theta_0) = l(a_2, \theta_1)$, $l(a_1, \theta_1) = 1 - \ln(2)$, $l(a_2, \theta_0) = \frac{1}{2} + \ln(2)$ onde a_1 representa a decisão de aceitar H_0 e a_2 representa a decisão de rejeitar H_0. θ_0 corresponde ao valor na hipótese nula e θ_1 representa o valor de θ na alternativa.

e) Considere o teste $H_0 : \theta = 1$ vs $H_1 : \theta = 2$ e uma amostra de tamanho 2. Seja $\alpha = P(I)$ e $\beta = P(II)$. Calcule o teste que minimize simultaneamente α e β.

9 Seja X uma observação simples proveniente da densidade $f(x;\theta) = \theta x^{\theta-1} I_{(0,1)}(x), \ \theta > 0$.

a) Considere o teste $H_0 : \theta \leq 1$ vs $H_1 : \theta > 1$, calcule a função poder e o tamanho do teste que rejeita H_0 se e somente se $\frac{1}{2} \leq X$.

b) Calcule o teste mais poderoso de tamanho α de $H_0 : \theta = 2$ vs $H_1 : \theta = 1$.

c) Considerando o teste $H_0 : \theta = 2$ vs $H_1 : \theta = 1$, calcule o teste minimax associado à função de perda l onde $l(d, \theta)$ assume os seguintes valores: $0 = l(a_1, 2) = l(a_2, 1)$, $l(a_1, 1) = 1 = l(a_2, 2)$ onde a_1 representa a decisão de aceitar H_0 e a_2 representa a decisão de rejeitar H_0.

d) Achar o teste UMP(α) para $H_0 : 2 \leq \theta$ vs $H_1 : \theta < 2$.

e) Considere o espaço dos testes "quociente de verossimilhanças" de $H_0 : \theta = 2$ vs $H_1 : \theta = 1$ e determine aquele que minimize $\alpha + \beta$, onde $\alpha = P(I)$ e $\beta = P(II)$.

10 Seja $X_1, \cdots, X_n$ uma amostra aleatória desde $f(x, \theta) = \frac{1}{\theta} x^{\frac{1-\theta}{\theta}} I_{(0,1)(x)}$. Considere o teste $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$.

a) Para a amostra de tamanho n calcule o teste uniformemente mais poderoso de tamanho α.

b) Considere $n = 2$, $\theta_0 = 1$ e $\alpha = 0.05$, calcule a função poder associada ao teste UMP(α) do item a).

11 Para cada situação apresentada a seguir, verifique se os dados apresentam evidência suficiente para rejeitar a hipótese nula.
a)População normal, $n = 15$, $\overline{X} = 83.9$, $s = 18.2$, $\alpha = 10\%$, para o teste $H_0 : \mu = 85$ vs $H_1 : \mu < 85$.
b)População normal, $n = 15$, $\overline{X} = 79.1$, $s = 11.8$, $\alpha = 10\%$, para o teste $H_0 : \mu = 76$ vs $H_1 : \mu \neq 76$.
c)$n = 36$, $\overline{X} = 80.4$, $s = 16.2$, $\alpha = 5\%$, para o teste $H_0 : \mu = 76$ vs $H_1 : \mu \neq 76$.
Onde s denota a desviação padrão amostral.

12 Sabendo que a resistência à tensão, de uma peça de algodão possui distribuição normal.
a)A resistência é medida em 15 peças selecionadas ao acaso, observando-se uma média amostral igual a 39.3 e um desvío padrão amostral igual a 2.6. Verifique se os dados são significantes ao nível $\alpha = 10\%$, para o

teste $H_0 : \mu = 40$ vs $H_1 : \mu \neq 40$.
b)Determine a região crítica dos teste enunciado em a) para $\alpha = 10\%$.
c)A resistência é medida em 54 peças selecionadas ao acaso, observando-se uma média amostral igual a 42.4 e uma desviação padrão amostral igual a 3.1. Calcule a força da evidência contida nos dados e determine para qué níveis de significância H_0 é rejeitada.
d)Melhorias implementadas no tratamento da fibra de algodão permitem suspeitar que a resistência tem aumentado. Perante esta afirmação reformule o teste. Se essa resistência foi medida em 15 peças observando-se uma média amostral de 41.3 com uma desviação padrão amostral igual a 2.6. Verifique se os dados são significantes ao nível $\alpha = 0.05$. Determine a região crítica do teste para $\alpha = 0.05$.

13 A demanda biológica de oxigênio (DBO) é um índice de poluição controlado nas industrias papeleras (para preservar o equilibrio ambiental toda industria papelera debe consumir uma quantidade de oxigênio que não supere um "valor limite"). Em 43 medidas coletadas numa industria no peíodo: Setembro 1999-Fevereiro 1999, a média e a desvío padrão dos dados observados foram 3242 ppd e 757 ppd, respectivamente. Aquela empresa tinha estabelecido como valor limite 3000 ppd para o DBO medio. Julgaria que os dados amostrais suportam que a meta foi atingida ao nível $\alpha = 5\%$?.

14 Uma empresa mineira acredita que a exploração de urânio é possível numa certa região , isto é, na região a concentração média de urânio é superior a 10. Admitindo-se que a distribuição desta concentração é normal e que as medições em 13 pontos selecionados ao acaso na região são
7.92, 10.29, 19.89, 17.73, 10.36, 13.50, 8.81, 6.18, 7.02, 11.71, 8.33, 9.32, 14.61
a)Verifique se há evidência suficiente contra a hipótese de abandono da área .
b)Qual sería a região crítica do teste ao nível de significância $\alpha = 2\%$?

15 Em 18 condenas por "possessão de drogas" num tribunal norteamericano as condenas atribuídas tiveram média de 38 meses e desvío padrão amostral de 4 meses. Considerando que as condenas são normalmente distribuídas, verifique se os dados suportam ao nível de significância $\alpha = 5\%$ a suspeita de que nesse tribunal a condena por "possessão de

drogas" é em média maior do que 36 meses. Conduzir o mesmo teste considerando $\alpha = 1\%$. Interprete.

16 Um fabricante de aparelhos de TV afirma que são necessários no máximo 250 microamperes (μA) para atingir um certo grau de brilhantismo num tipo de TV. Uma amostra de 20 aparelhos produz um promedio amostral de $\overline{X} = 257.3\ \mu A$. Denotemos por m ao verdadeiro promedio de μA necessário para atingir o grau de brilhantismo desejado e suponhamos que m é a média de uma população normal com σ conhecido e igual a 15.
a) Calcule a força da evidência contida nos dados para o nível $\alpha = 0.05$ conduzindo o teste cuja hipótese nula especifica que m é no máximo 250 μA.
b)Calcule a região crítica do teste para o nível $\alpha = 0.05$.
c)Se $m = 260$, Qual é a probabilidade de cometer um erro tipo II?.
d)Para qué valor de n (tamanho amostral) a probabilidade de cometer o erro tipo II resulta igual a 0.01.

17 O ponto de desvanecimento de cada uma de 16 amostras de uma certa marca de vegetais hidrogenados foi determinado, resultando numa média amostral $\overline{X} = 94.32$. Considerando que o ponto de desvanecimento possui distribuição normal de desvío conhecido $\sigma = 1.20$.
a)Verifique se a amostra apresenta evidência suficiente para rejeitar H_0 ao nível $\alpha = 0.01$, calculando o π-value. Onde $H_0 : \mu = 95$ vs $H_1 : \mu \neq 95$.
b)Se $\alpha = 0.01$ e $\mu' = 94$. Qual é a probabilidade de cometer erro tipo II?
c)Se $\alpha = 0.01$. Qué valor de n (tamanho amostral) é necessário para obter uma probabilidade de cometer erro tipo II, com $\mu' = 94$ igual a 0.1?.

18 O promedio desejado de SiO_2 em certo tipo de cemento aluminoso é de 5.5. Para provar se o verdadeiro promedio da porcentagem numa planta de produção em particular é 5.5, foram coletadas 16 amostras. Supondo que a porcentagem de SiO_2 numa amostra está normalmente distribuída com desvío conhecido e igual a $\sigma = 0.3$ e sabendo que na amostra selecionada obteve-se $\overline{X} = 5.25$, responda
a)os dados indicam de forma conclusiva que o verdadeiro promedio de porcentagem não é $\mu = 5.5$?.

b)Se o verdadeiro promedio da porcentagem é $\mu = 5.6$ e se utiliza uma prova de nível $\alpha = 0.01$, qual é a probabilidade de detectar essa desviação desde a hipótese nula? (i.e. os dados são significantes ao nível α?).
c)Se $\alpha = 0.01$ qué valor de n (tamanho amostral) é necessário para obter uma probabilidade de cometer erro tipo II, com $\mu' = 5.6$ igual a 0.01?.

19 Um experimento para comparar a resistência de coesão à tensão do morteiro modificado de látex de polímeros, com a resistência do morteiro não modificado, supondo que os dados tem distribuição Normal; resultou em $\overline{X} = 18.12\ kfg/cm^2$ para o morteiro modificado e em $\overline{Y} = 16.87\ kfg/cm^2$ para o morteiro não modificado. Sejam μ_1 e μ_2 as verdadeiras resitências de coesão à tensão para os morteiros modificado e não modificado respectivamente. Verifique se os dados suportam a rejeição de H_0. Onde $H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 > 0$ com nível de significância $\alpha = 0.01$, nas seguintes situações:

a) Se para o morteiro modificado foi utilizada uma amostra de tamanho $m = 40$ e para o não modificado foi utilizada uma amostra de tamanho $n = 32$. Os valores dos desvios são conhecidos σ_1 e σ_2 (associados respectivamente ao morteiro modificado e ao não modificado). $\sigma_1 = 1.6$ e $\sigma_2 = 1.4$. Proponha uma estatística para conduzir o teste e verifique se os dados indicam a rejeição de H_0.

b) Se para o morteiro modificado foi utilizada uma amostra de tamanho $m = 40$ e para o não modificado foi utilizada uma amostra de tamanho $n = 32$. Os valores dos desvios são conhecidos σ_1 e σ_2 (associados respectivamente ao morteiro modificado e ao não modificado). $\sigma_1 = \sigma_2 = 1.6$. Proponha uma estatística para conduzir o teste e verifique se os dados indicam a rejeição de H_0.

c) Sabendo que

$$m = 30, \sum_{i-1}^{m}(X_i - \overline{X})^2 = 40.1,\ n = 22, \sum_{j=1}^{n}(Y_j - \overline{Y})^2 = 53.22.$$

Assumindo que os desvios σ_1 e σ_2 são desconhecidos e iguais, proponha uma estatistica para o teste e determine se os dados indicam a rejeição de H_0.

20 Os estudantes universitários homens entediam-se mais facilmente que as estudantes mulheres?. Esta pergunta foi examinada pelo artigo “Boredom in Young Adults Gender and Cultural Comparisons”(*J. of Cross Cultural Psych.* pp. 209-223). Os autores aplicaram uma escala denominada *Escala Proneness de tedio* a 97 estudantes homens e a 148 estudantes mulheres, todos eles de universidades norteamericanas. Assumindo que a classificação fornecida pela escala Proneness possui distribuição normal verifique se a seguinte informação apoia a hipótese da investigação.
Conduzir o teste adequado utilizando um nível de significância $\alpha = 0.05$ e os dados da seguinte tabela

Gênero	Tamanho amostral	Média am.	Desvío verdadeiro (σ)
Homens	97	10.40	4.83
Mulheres	148	9.26	4.86

21 Denotemos por μ_1 e μ_2 aos verdadeiros promedios de durações de superficies de rodagem para duas marcas competidoras de medida FR78-15 de pneus radiais. Conduzir o seguinte teste de hipótese assumindo que a duração das superficies de rodagem possui distribuição normal
$H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 \neq 0$ com nível de significância $\alpha = 0.05$, usando a seguinte informação: $m = 40$, $\overline{X} = 36500$, $\sigma_1 = 2200$ (valor verdadeiro do desvío) e $n = 40$, $\overline{Y} = 33400$, $\sigma_2 = 1900$ (valor verdadeiro do desvío).

22 Conforme a herança mendeliana, a descendência de certo cruzamento deveria ser vermelha, preta ou branca na seguinte proporção: 9/16, 3/16, 4/16. Se um experimento mostrou 74,32 e 38 descendentes nessas categorias, a teoria está confirmada?.

23 Num determinado estudo, pediu-se a cada um de 52 indivíduos que anotasse o número de seus batimentos cardíacos em um minuto. Por hipótese (se H_0 for verdadeira), supõe-se que esses valores vêm de uma distribuição Normal com média 85 batidas por minuto (bat/min) e desvio padrão 12 bat/min. Teste esta hipótese, isto é, $H_0 : X$ é $N(85; 12^2)$, onde X é o número de batimentos cardíacos por minuto . Utilize os dados a seguir.

N. batimentos cardíacos(bat/min)	N. de indivíduos
$[54, 62)$	0
$[62, 70)$	6
$[70, 78)$	6
$[78, 86)$	16
$[86, 94)$	14
$[94, 102)$	4
$[102, 110)$	4
$[110, 118)$	2
$[118, 126)$	0

24 Seja X uma observação simples proveniente da densidade $f(x;\theta) = (2\theta x + 1 - \theta)I_{[0,1]}(x), \ \ -1 \leq \theta \leq 1$.

a) Calcule o teste mais poderoso de tamanho α de $H_0 : \theta = 0$ vs $H_1 : \theta = 1$.

b) Considere o teste: $H_0 : \theta \leq 0$ vs $H_1 : \theta > 0$, que rejeita H_0 se $X > \frac{1}{2}$. Calcule a função poder e o tamanho do teste.

c) Considere o teste: $H_0 : \theta \leq 0$ vs $H_1 : \theta > 0$. No espaço dos testes quociente de verossimilhanças, existe o teste UMP(α)?.

d) Calcule o teste quociente de verossimilhanças generalizado para $H_0 : \theta = 0$ vs $H_1 : \theta \neq 0$.

25 Seja uma amostra aleatória $X_1, \cdots, X_n$ proveniente da distribuição log-normal $LN(\mu, \sigma^2)$, definida para $-\infty < x < \infty, \sigma^2 > 0$ como

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right\} I_{(0,\infty)}(x)$$

a) Suponha $\sigma^2 = \sigma_0^2$ conhecida e calcule o teste UMP(α) de $H_0 : \mu \leq \mu_0$ vs $H_1 : \mu > \mu_0, \ \mu_0$ conhecido.

b) Suponha σ^2 desconhecida. Apresente uma estatistica que permita conduzir o teste $H_0 : \mu \leq \mu_0$ vs $H_1 : \mu > \mu_0$, μ_0 conhecido. Determine a região critica do teste ao nível α.

c) Suponha σ^2 desconhecida e calcule se for possível o teste quociente de verossimilhanças para
$H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$, μ_0 conhecido. Nesse caso determine a região critica do teste ao nível α.

d) Suponha $\mu = \mu_0$ conhecida e calcule o teste UMP(α) de
$H_0 : \sigma^2 \leq \sigma_0^2$ vs $H_1 : \sigma^2 > \sigma_0^2$, σ_0^2 conhecido.

e) Suponha μ desconhecida. Apresente uma estatistica que permita conduzir o teste $H_0 : \sigma^2 \leq \sigma_0^2$ vs $H_1 : \sigma^2 > \sigma_0^2$, σ_0^2 conhecido. Determine a região critica do teste ao nível α.

8 Estimação por Intervalo

1 Seja a a.a. de tamanho 20
13.736, 14.579, 14.025, 13.542, 14.294, 13.815, 13.615, 13.633, 13.893, 14.105, 14.129, 15.029, 13.814, 14.516, 13.982, 14.174, 13.900, 14.319, 13.822, 13.728
desde uma distribuição Normal de média desconhecida μ e variância σ^2.
a)Calcule o $100\gamma\%$ I.C. para μ sabendo que $\sigma^2 = 0.36$, para os níveis de confiança $\gamma = 0.9, 0.95, 0.99$. Calcule o cumprimento de cada intervalo de confiança. Evidência alguma relação entre o cumprimento do intervalo e o nível de confiança?.
b)Calcule o $100\gamma\%$ I.C. para μ supondo σ^2 desconhecido, para os níveis de confiança $\gamma = 0.9, 0.95, 0.99$.

2 Seja a a.a. de tamanho 15
5.055, 6.916, 5.812, 5.044, 4.914, 5.665, 4.772, 5.502, 3.841, 5.782, 4.579, 5.477, 7.158, 5.254, 5.276
desde uma distribuição Normal de média μ e variância desconhecida σ^2.
a)Calcule o $100\gamma\%$ I.C. para σ sabendo que $\mu = 5$, para os níveis de confiança $\gamma = 0.9, 0.95, 0.99$. Calcule o cumprimento de cada intervalo de confiança.

b)Calcule o $100\gamma\%$ I.C. para σ supondo μ desconhecido, para os níveis de confiança $\gamma = 0.9, 0.95, 0.99$. Calcule o cumprimento de cada intervalo de confiança.

3 Sejam $\{X_i\}_{i=1}^n$ uma seqüência de realizações $Ber(p)$ independentes, denotemos por k ao número de sucessos registrados nas n realizações.
a) Supondo que $3 \leq np(1-p)$ e dados ϵ e γ, determine o tamanho amostral mínimo para poder garantir que k/n é uma estimativa de p desconhecido com grau de confiança $P(\epsilon) = P\left\{|\frac{k}{n} - p| \leq \epsilon\right\}$ maior o igual que γ. Considere os seguintes valores de ϵ e γ

ϵ	γ
0.052	0.999
0.04	0.99
0.05	0.95
0.1	0.9

b)Supondo que $3 \leq np(1-p)$ calcule o I.C. conservador para p, para os níveis de confiança $\gamma = 0.9, 0.95, 0.99$, quando se sabe que em uma seqüência de 85 realizações se verificaram 23 sucessos.
c)Supondo que $3 \leq np(1-p)$ calcule o I.C. por estimativa pontual para p, para os níveis de confiança $\gamma = 0.9, 0.95, 0.99$, quando se sabe que em uma seqüência de 85 realizações se verificaram 23 sucessos.
d)Compare os I.C. obtidos pelas metodologías de b) e c).

4 Seja X uma observação desde a densidade $f(x;\theta) = \theta x^{\theta-1} I_{(0,1)(x)},\ \theta > 0$.
a)Calcule a quantidade pivotal e determine o intervalo de confiança que a mesma define para θ.
b)Mostre que $(Y/2, Y)$ é um intervalo de confiança para θ e calcule o coeficiente de confiança. Sendo $Y = \frac{-1}{ln(X)}$.

5 Seja $f(x,\theta) = \theta \exp(-\theta x) I_{(0,\infty)}(x),\ \theta > 0$ a distribuição exponencial de parâmetro θ, considere as seguintes situações e calcule os intervalos de confiança requeridos.

a)Se $X_1, \cdots, X_n$ é uma a.a. desde $f(\cdot, \theta)$ calcule a quantidade pivotal para θ e o $100\gamma\%$ I.C. para θ.
b)Se $X_1, \cdots, X_n$ é uma a.a. desde $f(\cdot, \theta)$ calcule o $100\gamma\%$ I.C. para $1/\theta$.
c)Se $X_1, \cdots, X_n$ é uma a.a. desde $f(\cdot, \theta)$ calcule o $100\gamma\%$ I.C. para $1/\theta^2$.
d)Se $X_1, \cdots, X_n$ é uma a.a. desde $f(\cdot, \theta)$ calcule o $100\gamma\%$ I.C. para $e^{-\theta}$.
e)Calcule uma quantidade pivotal para θ dependente unicamente de $Y_1 = min\{X_1, \cdots, X_n\}$, determine o $100\gamma\%$ I.C. para θ sugerido pela quantidade pivotal. Apresente o 95% I.C. para θ, assumindo que $y_1 = 0.9$ e $n = 10$.
f)Considere a simple observação X proveniente de $f(\cdot, \theta)$, demonstre que $(X/2, X)$ é um I.C. para $1/\theta$ e calcule seu coeficiente de confiança.

6 Seja $X_1, \cdots, X_n$ uma a.a. desde $f(x, \theta) = \frac{1}{\theta} x^{\frac{1-\theta}{\theta}} I_{[0,1]}(x),\ \theta > 0$. Apresente uma quantidade pivotal para θ e calcule o $100\gamma\%$ I.C. para θ induzido pela mesma.

7 Suponha que as proporções de quatro tipos de sangue de certa raça sejam iguais a 0.16, 0.48, 0.2, 0.16. Dadas as frequências observadas de 180, 360, 130 e 100 para uma outra raça. verifique se ela possui a mesma distribuição quanto aos tipos de sangue.

8 Os números de acidentes automobilísticos semanais de determinada comunidade foram os seguintes: 11, 8, 17, 5, 13, 9, 17, 6, 10, 4. Tais frequências estaão de acordo com a crença de que as condições dos acidentes são as mesmas, neste período de 10 semanas ?

9 Uma amostra de cinquenta peças produzidas por uma máquina forneceu a distribuição de comprimentos das peças dada a seguir, com valores em mm. A especificação de produção indica que o comprimento das peças tem distribuição normal de média 500 mm e desvío padrão 10 mm. Através dos valores observados, concordamos ou discordamos dessa especificação (use o nível descritivo)? Se a especificação não for a postulada, você consegue explicar o que está errado?

10 Os dados seguintes representam os resultados de uma investigação da distribuição do sexo das crianças de 32 famílias possuindo cada uma

Tabela 2: Especificação

Comprimento (mm)	Frequências
$[480, 485)$	1
$[485, 490)$	5
$[490, 495)$	11
$[495, 500)$	14
$[500, 505)$	9
$[505, 510)$	5
$[510, 515)$	4
$[515, 520]$	1
Total	50

delas 4 crianças. Use a distribuição Binomial com $n = 4$ e $p = 0.5$ para calcular as frequências esperadas. Aplique depois o teste Qui Quadrado para verificar se este modelo de distribuição Binomial é satisfatório neste caso.

Tabela 3: Dados

Número de filhos	0	1	2	3	4
Número de famílias	4	10	8	7	3

11 Uma pesquisa eleitoral foi realizada com o objetivo de estudar a influência da idade na preferência por dois candidatos presidenciais A e B. A população de eleitores foi dividida em três faixas de iddade sendo que de cada uma delas foi obtida uma amostra de 200 indivíduos. Em seguida, os eleitores selecionados em cada amostra tiveram suas opiniões registradas. Os resultados obtidos foram os seguintes: Estabeleça uma hipótese apropriada e efetue o teste correspondente.

12 Numa hipotética epidemia 900 crianças contraíram a doença. Das 450 que não receberam tratamento, 104 sofreram efeitos posteriores. Das 450 restantes que efetivamente receberam tratamento, 166 sofreram efeitos posteriores. Teste a hipótese de que a chance de sofrer efeitos

Tabela 4: Dados

Intervalo	Preferem A	Preferem B	Indecisos
$[16, 30)$	67	117	16
$[30, 50)$	109	74	17
$[50, \infty)$	118	64	18

posteriores é a mesma para indivíduos que receberam e não receberam tratamento.

13 Os resultados da classificação de 100 pessoas segundo a cor dos olhos e do cabelo foram os seguintes: Você diria que a cor dos olhos independe

Tabela 5: Dados

	Castanhos	Azuis	Cinza
Claro	13	18	9
Escuro	37	12	11

da cor do cabelo?

14 No exame final de um curso compareceram 105 estudantes do sexo masculino e 40 do sexo feminino. O professor anotou se o estudante escolhia uma carteira nas filas da frente ou nas de trás. Os resultados dessa observação estão apresentados na tabela abaixo,

Tabela 6: Dados

	Filas da Frente	Filas de trás
Homens	35	70
Mulheres	20	20

A escolha da fila é a mesma para homens e mulheres?

15 Uma amostra de 200 adultos foi entrevistada a respeito de certo projeto de lei. Os resultados são os que seguem,

Tabela 7: Dados

	Favoráveis	Contrários
Homens casados	56	24
Homens solteiros	15	25
Mulheres casadas	24	16
Mulheres solteiras	13	27

Verifique se a opinião independe do sexo; além disso, verifique se a opinião independe do estado civil.

16 Uma agência de turismo classifica uma amostra de 200 clientes com menos de 40 anos segundo o sexo e o passeio preferido (praias ou serras). Após a aplicação de um teste Qui Quadrado apropriado, obteve-se um $p - valor = 0.451$. A mesma pesquisa é repetida para uma amostra de 200 clientes com 40 anos ou mais e o valor do $p - valor = 0.001$. Os dados são apresentados a seguir.

Tabela 8: Menos de 40 anos

	Praias	Serras
Masculino	70	30
Feminino	65	35

Tabela 9: 40 anos ou mais

	Praias	Serras
Masculino	50	50
Feminino	70	30

a. Formule as hipóteses para cada caso

b. Que tipo de teste Qui Quadrado foi aplicado?. Justifique.

c. Interprete os resultados.