1 Distribuições Amostrais

1 Determine a distribuição amostral de uma a.a. de tamanho n proveniente de a) $N(\mu, \sigma^2),\ \mu \in R,\ \sigma^2 \in R_+$; b) $exp(\lambda),\ \lambda > 0$; c) $Gama(\alpha, \beta),\ \alpha > 0, \beta > 0$.
Calcule em cada caso o valor esperado e a varianância amostrais. Gera 100 a.a. desde uma distribuição Normal de media 2 e variância 0.25 e calcule os momentos amostrais de ordem 1 e 2, centrados em 0 e em $\overline{x}_n$.

2 Seja $X_1, \cdots, X_n$ uma a.a. desde $f(\cdot)$ e M_2^a o momento amostral de ordem 2 centrado em a. Determine o valor a^* tal que $M_2^{a^*}$ é mínimo.

3 Seja $X_1, \cdots, X_n$ uma a.a. desde $f(\cdot)$ de variância σ^2 e momentos $E(X^r) = \mu_r$ e sejam

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2, \quad M_2^{\overline{X}_n} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

a variância amostral e o segundo momento amostral centrado em $\overline{X}_n$. Demonstre que

$$E\left\{S_n^2\right\} = \sigma^2,\ Var\left\{S_n^2\right\} = \frac{1}{n}\left\{\mu_4 - \frac{n-3}{n-1}\sigma^4\right\}$$

Calcule $E\left\{M_2^{\overline{X}_n}\right\}$ e $Var\left\{M_2^{\overline{X}_n}\right\}$.

4 Seja $X_1, \cdots, X_n$ uma a.a. desde $f(\cdot)$ de variância σ_x^2 e media μ_x. E seja outra a.a. independente da anterior $Y_1, \cdots, Y_n$ uma a.a. desde $g(\cdot)$ de variância σ_y^2 e media μ_y. Definamos $Z_i := X_i - Y_i,\ i = 1, \cdots, n$. Então

$$E(\overline{Z}_n) = \mu_x - \mu_y,\ \ Var(\overline{Z}_n) = \frac{1}{n}\left\{\sigma_x^2 + \sigma_y^2\right\}$$

5 Seja $X_1, \cdots, X_n$ uma a.a. desde $f(\cdot)$. Calcule a distribuição exacta de $\overline{X}_n$, para as seguintes distribuições a) $Bernoulli(p)$, b) $Poisson(\lambda)$, c) $Exp(\theta)$, d) $U(0, 1]$.

6 Demonstre que se $\overline{X}_n$ é a media amostral desde uma a.a. de tamanho n proveniente de uma distribuição $N(\mu, \sigma^2)$ então $\overline{X}_n$ possui distribuição $N(\mu, \frac{\sigma^2}{n})$.

7 Demonstre que se $X_1, \cdots, X_n$ são normalmente distribuidas e independentes de medias $\mu_1, \cdots, \mu_n$ e variâncias $\sigma_1^2, \cdots, \sigma_n^2$, respectivamente, então

$$U := \sum_{i=1}^{n} \left\{ \frac{X_i - \mu_i}{\sigma_i} \right\}^2$$

possui distribuição χ_n^2.

8 Seja $Z_1, \cdots, Z_n$ uma a.a. desde $N(0,1)$. a)Demonstre que Z_1+Z_2 e Z_1-Z_2 são independentes. b) Demonstre que $\sum_{i=1}^n (Z_i - \overline{Z}_n)^2$ possui distribuição χ_{n-1}^2, assumindo que $\overline{Z}_n$ e $\sum_{i=1}^n (Z_i - \overline{Z}_n)^2$ são independentes.

9 Seja $X_1, \cdots, X_n$ uma a.a. desde $N(\mu, \sigma^2)$.a) Determine a distribuição de $\sum_{i=1}^n \left\{ \frac{X_i - \mu}{\sigma} \right\}^2$ e de $\sum_{i=1}^n \left\{ \frac{X_i - \overline{X}_n}{\sigma} \right\}^2$. b) Determine a distribuição de $\frac{n-1}{\sigma^2} S_n^2$, onde $S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X}_n)^2$.

10 Demonstrar que se U e V são variáveis aleatorias independentes de distribuição χ_m^2 e χ_n^2 respectivamente, então $X := \frac{U/m}{V/n}$ possui distribuição F com m e n graus de liberdade.

11 Demonstrar que se Z e U são variáveis aleatorias independentes de distribuição $N(0,1)$ e χ_k^2 respectivamente, então $W := \frac{Z}{\sqrt{(U/k)}}$ possui distribuição t com k graus de liberdade.

2 Estimação Pontual

2.1 Métodos de estimação pontual

1 Suponha que X é uma v.a. Normal de média 2 e variância desconhecida σ^2. Qual o EMV e pelo método dos momentos para σ^2, baseado numa amostra aleatória de tamanho n de observações de X?.

2 Suponha que X é uma v.a. $Poisson(\lambda)$, com λ desconhecido. Qual o EMV e pelo método dos momentos para λ, baseado numa amostra aleatória de tamanho n de observações de X?.

3 Suponha que X é uma v.a. $geom(p)$, com p desconhecido. Qual o EMV e pelo método dos momentos para p, baseado numa amostra aleatória de tamanho n de observações de X?.

4 Suponha que X é uma v.a. $\exp(\lambda)$, com λ desconhecido. Qual o EMV e pelo método dos momentos para λ, e para $1/\lambda$ baseado numa amostra aleatória de tamanho n de observações de X?.

2.2 Suficiência

1 Seja $X_1, \cdots, X_n$ a.a. $Ber(p)$, calcule uma estatística suficiente para p.

2 Seja $X_1, \cdots, X_n$ a.a. $U(a, b)$, calcule uma estatística suficiente para (a, b).

3 Seja $X_1, \cdots, X_n$ a.a. $N(4, \sigma^2)$, calcule uma estatística suficiente para σ.

2.3 Família Exponencial

1 Demonstre que as seguintes distribuições pertencem à família exponencial
a)Bernoulli de parâmetro p,
b)Binomial de parâmetros n e p (analise em função de p)
c)Binomial Negativa de parâmetros r e p (analise em função de p)
d)Poisson de parâmetro λ
e)Exponencial de parâmetro λ
f)Gama de parâmetros r e λ (analise em função de λ)
g)Beta de parâmetro (a, b).

2 As distribuições Cauchy e Uniforme pertencem à família exponencial?

2.4 Propriedades

1 Seja $X_1, \cdots, X_n$ a.a. desde $N(\mu, \sigma^2)$, σ^2 desconhecido. Compare os erros quadáticos médios (EQM) dos estimadores apresentados em cada caso.
a)Supondo μ conhecido:

$$\sigma^{*2} = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{n}, \; S_n^2 = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{n-1}.$$

Qual deles é melhor?.
b)Supondo μ desconhecido:

$$\sigma^{*\,2} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}, \; S_n^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}.$$

Qual deles é melhor?.

2 Demonstre que se $\theta_1^*, \cdots, \theta_n^*$ é uma seqüência de estimadores de θ, não viciados e que verificam a consistência quadrática, então eles verificam a consistência simples.

3 Gere 100 valores desde uma normal de media 2 e variância 1, construa os estimadores $\theta_j^* = \frac{1}{j}\sum_{i=1}^{j} X_i$, qual de todos vc escolheria para estimar a media da distribuição ?.

4 Suponha que $X_1, \cdots, X_n$ seja uma amostra aleatória de uma $Poisson(\lambda)$. demonstre que a estatística $T := \sum_{i=1}^{n} X_i$ é completa.

5 Considere uma a.a. de tamanho n proveniente da distribuição $Ber(p)$. Demonstre que $E\{X_1|S=s\}$ é ENVUMV para p, onde $S = \sum_{i=1}^{n} X_i$.

6 Considere uma a.a. de tamanho n proveniente da $Poisson(\lambda)$.
a) Obtenha o estimador de λ pelo método dos momentos e pelo método de Máxima Verossimilhança. Calcule o vicio em ambos os casos.
b)Demonstre que $S = \sum_{i=1}^{n} X_i$ é estatística suficiente minimal e completa para λ. Demonstre que $e^{-S/n}$ é estimador de Máxima Verossimilhança para $e^{-\lambda}$ e que $e^{-S/n}$ não é ENVUMV para $e^{-\lambda}$.
c) Demonstre que $E\left\{I_{\{X_1=0\}}(X_1)|\sum_{i=1}^{n} X_i = s\right\}$ é ENVUMV para $e^{-\lambda}$ onde $I_{\{X_1=0\}}(X_1) = 1$ se $X_1 = 0$ e 0 em qualquer outro caso.
d) Calcule a variância de $E\left\{I_{\{X_1=0\}}(X_1)|\sum_{i=1}^{n} X_i = s\right\}$ e verifique que é menor que a variância do estimador não viciado de $e^{-\lambda}$, $\frac{1}{n}\sum_{i=1}^{n} I_{\{X_1=0\}}(X_i)$.

7 Seja $X_1, \cdots, X_n$ uma a.a. desde $f(\cdot)$ de variância σ^2 e media μ.
a) Demonstre que $\sum_{i=1}^{n} a_i X_i$ é um estimador não viciado de μ para qualquer conjunto de constantes $a_1, \cdots, a_n$ satisfazendo $\sum_{i=1}^{n} a_i = 1$.
b)Se $\sum_{i=1}^{n} a_i = 1$ demonstre que $Var\{\sum_{i=1}^{n} a_i X_i\}$ é minimizada quando $a_i = 1/n$, $i = 1, \cdots, n$.

8 Sejam $X_1, \cdots, X_n \cdots$ v.a.i.i.d. de alguma distribuição tal que os quatro primeiros momentos centrais existem. Seja $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ verifique se $\{S_n^2\}_{n=1}^\infty$ é consistente (quadrático) de σ^2.

9 Seja $X_1, \cdots, X_n$ uma a.a. desde $\theta x^{\theta-1} I_{(0,1)}(x)$ sendo $\theta > 0$.
a) Calcule o estimador de Máxima Verossimilhança de $\mu = \frac{\theta}{1+\theta}$.
b) Calcule a estatística sufieciente e verifique a completitude. É $\sum_{i=1}^n X_i$ uma estatística suficiente?.

10 Seja $X_1, \cdots, X_n$ uma a.a. desde $Exp(\lambda)$, apresente ENVUMV para

$$\lambda, \ 1/\lambda, \ e^{-k\lambda}, \ k > 0.$$

3 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.

4 Teste de Hipótese

1 Denotemos por μ ao verdadeiro promedio de nível de radioatividade (picocuries por litro). O valor 5 pCi/L é considerado como linha divisoria 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 $Bin(n, p)$. Conduzca 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 fases. Procurando suporte para tal afirmação considera-se o número de vezes que a fase 2 é obtida numa seqüência de n lanzamentos independentes.
a)Determine a hipótese nula H_0 e a alternativa H_1.
b)Em $n = 20$ lanzamentos independentes obtem-se 2 vezes a fase 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 $n = 20$ lanzamentos independentes obtem-se 6 vezes a fase 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 $\{6 \leq \sum_{i=1}^{n} x_i\}$.

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 $C^* = \{(x_1, x_2) : 3/4x_1 \leq x_2\}$. 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 $H_1 : \mu > \mu_0$, μ_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.

5 Inferência Bayesiana

1 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.

2 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 $x = (x_1, \cdots, x_n)$.

3 Seja θ 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 unica 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).

4 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.

5 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 $f(\theta) \in \mathcal{F}, \;\; f(\theta/x) \in \mathcal{F}$.

a) 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 $f(x/\theta)$.

b) 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

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

6 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 $f(x/\theta)$.

7 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 $\theta = (\mu, \sigma^2)$.