Probabilidade e Estatística - Semana XII

Estimação

Muitas vezes, quando coletamos dados de um experimento temos, de antemão, uma idéia de que tipo de distribuição estes dados vem. Entretanto, a informação que estamos buscando é qual seria o(s) parâmetro(s) θ (ou θ_1, θ_2, ...) que definem univocamente esta distribuição. Por exemplo, se sabemos que os dados vem de uma distribuição exponencial, estaremos interessados no parâmetro $\theta = \beta$ da distribuição. Se os dados vem de uma distribuição normal, gostaríamos de saber os valores de $\theta_1 = \mu$ e $\theta_2 = \sigma^2$. É importante notar que uma vez que os dados são aleatórios, os parâmetros também podem ser tratados como variáveis aleatórias.

Modelo estatístico

Chamamos de modelo estatístico a definição do tipo de variáveis que serão observadas, a escolha de uma classe de funções de probabilidade para os dados sendo coletados, a definição dos parâmetros (θ) desta distribuição e, se possível, uma vaga idéia de como é a função densidade de probabilidade dos parâmetros.

Espaço de Parâmetros

Assim que os parâmetros forem definidos, devemos escolher o espaço de parâmetros Ω, que é o conjunto de todos os possíveis valores que os parâmetros podem ocupar.

Classes de Problemas de Inferência

Os problemas de estimação se dividem em:

- predição: quando estamos interessados em quantidades que não foram medidas. Estas quantidades podem ser possíveis valores futuros das variáveis aleatórias observáveis ou podem ser o valor dos parâmetros das funções probabilidade. Neste último caso o problema é chamado de *estimação*

- tomada de decisão: é basicamente um teste cuja resposta é inferida através das medidas dos dados observados. Acontece quando temos, por exemplo, um medicamento sendo testado em diversas pessoas e devemos decidir se o medicamento é eficaz ou não contra a doença para a qual o teste está sendo aplicado. Este tipo de classe de problemas é definido de *teste de hipótese*.

- desenho de experimento: quano o experimentador tem controle sobre o experimento, ele pode desenhar o experimento de forma a facilitar a sua posterior análise ou pode ser feito de tal maneira a testar somente uma variável de interesse.

Estatística

Definição 1: *Suponha que os observáveis de interesse em determinado experimento sejam $X_1, X_2, ..., X_n$. Seja r uma função arbitrária com imagem em $\mathbb{R}$, então a variável aleatória $T = r(X_1, ..., X_n)$ é chamada de estatística.*

Estimação

Existem diversos métodos para se trabalhar os dados observados de determinado experimento e estimarmos, à partir dos dados, qual o melhor valor possível dos parâmetros da f.d.p. que representa o experimento. Estudaremos um dos mais utilizados, que é o estimador de máxima verossimilhança, ou E.M.V.

Estimador de Máxima Verossimilhança

Suponha que façamos um experimento onde coletamos n variávies aleatórias $X_1, X_2, ..., X_n$. Suponhamos que sabemos que estas variáveis vêm de uma função densidade de probabilidade de apenas um parâmetro θ, ou seja, e uma distribuição $f(x|\theta)$. Caso as variáveis aleatórias sejam independentes, então a f.d.p. conjunta dos n dados será dada por:

$$f_n(x_1, x_2, ..., x_n|\theta) = f(x_1|\theta)\, f(x_2|\theta)...f(x_n|\theta) = [f(x|\theta)^n]$$

A função acima representada é chamada de *função verossimilhança*.

Definição 2 (Estimador de Máxima Verossimilhança): *Para cada possível vetor observado $\vec{x} = \{x_1, x_2, ..., x_n\}$, onde n é o número de observações realizadas, se $\delta(\vec{x}) \in \Omega$ representar o valor de $\theta \in \Omega$ para o qual a função verossimilhança $f_n(\vec{x}|\theta)$ é um máximo, e $\hat{\theta} = \delta(\vec{X})$ ser o estimador de θ definido desta maneira. O estimador $\hat{\theta}$ é chamado de* estimador de máxima verossimilhança *de θ. Depois de $\vec{X} = \vec{x}$ ser observado, o valor de $\delta(\vec{x})$ é chamado de* estimativa de máxima verossimilhança *de θ.*

Exemplo 1

Uma empresa produz lâmpadas. Ela sabe que o tempo de vida destas lâmpadas segue uma distribuição exponencial, mas está interessada em saber qual é o valor do tempo de vida médio. Ela observa o tempo de vida de 3 lâmpadas: $X_1 = 3$, $X_2 = 1,5$ e $X_3 = 2,1$. Qual é a melhor estimativa do tempo de vida médio a partir destes 3 valores?

Sabemos da distribuição exponencial que ela tem um único parâmetro β. O valor médio da distribuição é $1/\beta$. O estimador de máxima verossimilhança neste caso é:

$$f_n(\vec{x}|\theta) = \theta^n \exp(-\theta \sum_{x=1}^{n} x_i)$$

A fim de minimizarmos o E.M.V., podemos maximizar $L = \ln(\text{E.M.V.})$, já que o logaritmo é uma função monótona crescente. Assim, definindo $y = \sum_{x=1}^{n} x_i$:

$$L = n \ln(\theta) - (\theta\, y)$$

Maximizando L temos:

$$\frac{dL}{d\theta} = 0 \;\rightarrow\; \frac{n}{\theta} - y$$

que resulta em:

$$\theta = \frac{n}{\sum_{x=1}^{n} x_i} = \frac{1}{\bar{X}}$$

No presente caso temos que $\hat{\theta} = 3/6.6 = 0,455$.

Exemplo 2

Suponha que X_1, X_2, ..., X_n são amostras aleatórias provenientes de uma distribuição normal com média desconhecida e variância conhecida σ^2. Encontre a melhor estimativa para a média desta istribuição

A função verossimilhança é, para o parâmetro desconhecido $\theta = \mu$, dada por:

$$f_n(\vec{x}|\theta) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta)^2\right]$$

Podemos ver facilmente que esta expressão será máxima quando o argumento do expoente for o máximo possível. Para tanto temos que minimizar Q, definido como:

$$Q(\theta) = \sum_{i=1}^{n}(x_i - \theta)^2$$

Assim:

$$\frac{dQ}{d\theta} = 0 = -\sum_{i=1}^{n} 2(x_i - \theta) = 0 \;\rightarrow (\sum_{i=1}^{n} x_i) = n\theta$$

que resulta em $\theta = (\sum_{i=1}^{n} x_i)/n = \bar{x}$. É importante notar neste exemplo que o valor de $\hat{\theta} = \hat{\mu}$ não é afetado pelo valor de σ^2.

Exemplo 3

Considerando o mesmo exemplo anterior, mas onde é desconhecido tanto μ quanto σ^2 da distribuição normal de onde vieram os dados. Encontrar a melhor estimativa destes parâmetros.

Neste caso devemos encontrar a melhor estimativa para ambos os parâmetros simultaneamente. Uma possível estratégia é maximizarmos um dos parâmetros para um valor fixo do segundo, e em seguida maximizarmos o segundo parâmetro utilizando a melhor estimativa encontrada para o primeiro e, finalmente, encontrarmos a solução através dos dois estimadores encontrados.

Na prática, podemos utilizar o resultado do exemplo anterior e ver que a melhor estimativa para $\theta_1 = \mu$ é $\bar{X}$. Novamente, é melhor trabalharmos com o logaritmo da função verossimilhança, assim:

$$\begin{aligned} L(\theta_2) &= \ln\left(f_n(\vec{x}|\theta_2)\right) \\ &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 \end{aligned}$$

Assim:

$$\frac{dL}{d\theta_2} = -\frac{1}{n}\frac{1}{\theta^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 0$$

Resolvendo, temos:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Este é um ponto de máximo pois a segunda derivada é negativa no ponto crítico.

Exemplo 4

Suponha que seja desconhecida a proporção de compradoras mulheres de determinado produto em um supermercado, em relação ao número de compradores homens. Se fez uma pesquisa, onde foi avaliado o sexo de 70 compradores e se notou que 58 foram mulheres, enquanto que 12 foram homems. Encontre a melhor estimativa desta proporção.

Utilizamos a distribuição de Bernoulli para a função verossimilhança:

$$f_n(\vec{x}|\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}$$

O logaritmo desta função é:

$$L = \ln\left(f_n(\vec{x}|\theta)\right) = \ln(\theta)\left[\sum_{i=1}^{n}(x_i)\right] + \ln\left((1-\theta)\right)\left[n - \sum_{i=1}^{n}(x_i)\right]$$

Igualando isto a derivada de L a zero e definindo $y = \left[\sum_{i=1}^{n}(x_i)\right]$, temos:

$$\frac{dL}{d\theta} = \frac{y}{\theta} - \frac{1}{1-\theta}(n - y) = 0$$

que resulta em $\theta = y/n = \sum_{i=1}^{n}(x_i)/n$ que é 58/70.

Distribuição de χ^2

Como vimos anteriormente, o M.L.E. da variância de uma amostra de n variáveis aleatórias retiradas de uma distribuição normal com média μ e variância desconhecida é dada por:

$$\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2.$$

Como visto anteriormente, se as variáveis X_i forem retiradas de uma distribuição normal, então as variáveis $Z_i = (X_i - \mu)/\sigma$ seguem uma distribuição normal padrão. Por conta disso, a variável $\sum_{i=1}^{n} Z_i^2$ segue uma distribuição do tipo χ^2 com n graus de liberdade. Podemos ver também que:

$$\begin{aligned}\sum_{i=1}^{n} Z_i^2 &= \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}\\ &= \frac{n}{\sigma^2}\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{n}\\ &= \frac{n\hat{\sigma^2}}{\sigma^2}\end{aligned}$$

Assim, a variável $\frac{n\hat{\sigma^2}}{\sigma^2}$ também segue uma distribuição do tipo χ^2 com n graus de liberdade.

Exemplo 5

Suponha que um ponto com coordenadas (X, Y) seja escolhido aleatoriamente no plano xy, onde X e Y são variáveis aleatórias independentes retiradas de uma distribuição normal padrão. Seja um círculo de raio R, centrado na origem, qual deve ser o raio R deste círculo de forma que a probabilidade do ponto estar dentro do círculo seja maior do que 0,99?

A distância D do ponto com coordenadas (X, Y) até a origem é dada por $D^2 = X^2 + Y^2$. Para que o ponto esteja dentro do círculo, temos que $D^2 < R^2$. Assim, o que queremos na verdade é que:

$$\Pr\left(X^2 + Y^2 = D^2 < R^2\right) \leq 0,99$$

Como D^2 é a soma de variáveis aleatória que seguem a distribuição normal padrão, então D^2 segue a distribuição de χ^2 com 2 graus de liberdade. Olhando na tabela vemos que $R^2 = 9,210$ é um raio tal que α (a área à direita de R) é igual a 0,01. Ou seja $R = \sqrt{R^2} = 3,035$.

Distribuição de S^2

Devemos lembrar que S^2 é igual a:

$$S^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Vamos considerar inicialmente que:

$$\begin{aligned}\sum_{i=1}^{n}(X_i-\mu)^2 &= \sum_{i=1}^{n}[(X_i-\bar{X}) + (\bar{X}-\mu)]^2\\ &= \sum_{i=1}^{n}[(X_i-\bar{X})^2 + 2(X_i-\bar{X})(\bar{X}-\mu) + (\bar{X}-\mu)^2]\\ &= \left\{\sum_{i=1}^{n}(X_i-\bar{X})^2\right\} + n(\bar{X}-\mu)^2\end{aligned}$$

Isto porque no termo do meio temos $\sum_{i=1}^{n}(X_i - \bar{X}) = 0$. Assim, dividindo todos os termos acima por σ^2 e usando $S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)$ temos:

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 = \frac{(n-1)S^2}{\sigma^2} + \frac{\bar{X}-\mu}{\sigma^2/n}$$

O termo na esquerda representa uma distribuição do tipo χ^2 com n graus de liberdade. O último termo na direita representa uma distribuição do tipo χ^2 com 1 grau de liberdade. O que acontece é que o primeiro termo na direita é uma distribuição do tipo χ^2 com $n-1$ graus de liberdade.

Exemplo 6

Suponha que $X_1, ..., X_n$ sejam n variáveis aleatórias retiradas de uma distribuição com média μ e variância σ^2. Assumindo que o tamanho desta amostra seja $n = 16$, determine o valor das seguintes probabilidades:

a) **Pr** $\left(\frac{1}{2}\sigma^2 \leq \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 \leq 2\sigma^2\right)$

b) **Pr** $\left(\frac{1}{2}\sigma^2 \leq \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \leq 2\sigma^2\right)$

No primeiro caso podemos rearranjar o argumento da probabilidade, multiplicando por n e dividindo por σ^2:

$$\frac{1}{2}\sigma^2 \leq \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 \leq 2\sigma^2 =$$
$$= \frac{n}{2} \leq \frac{n\hat{\sigma^2}}{\sigma^2} \leq 2n$$

Assim, temos Pr $\left(8 \leq \frac{n\hat{\sigma^2}}{\sigma^2} \leq 32\right)$, com $\frac{n\hat{\sigma^2}}{\sigma^2}$ tendo uma distribuição χ^2 com 16 graus de liberdade. Esta probabilidade é igual a:

$$\text{Pr}\left(8 \leq \frac{n\hat{\sigma^2}}{\sigma^2} \leq 32\right) = \text{Pr}\left(\frac{n\hat{\sigma^2}}{\sigma^2} \leq 32\right) - \text{Pr}\left(\frac{n\hat{\sigma^2}}{\sigma^2} \leq 8\right)$$
$$\sim 0,99 - 0,05 = 0,94$$

No segundo caso temos praticamente a mesma situação. Entretanto $\sum_{i=1}^{n}(X_i - \bar{X})^2/\sigma^2$ tem a distribuição χ^2 com $n-1$ graus de liberdade. Neste caso vemos que:

$$\text{Pr}\left(8 \leq \frac{n\hat{\sigma^2}}{\sigma^2} \leq 32\right) = \text{Pr}\left(\frac{n\hat{\sigma^2}}{\sigma^2} \leq 32\right) - \text{Pr}\left(\frac{n\hat{\sigma^2}}{\sigma^2} \leq 8\right)$$
$$=\sim 0,995 - 0,5 = 0,945$$

Estimadores Viesados e Não-Viesados

O método da máxima verossimilhança nos dá uma estimativa para determinada estatística. Entretanto, podemos ter outras estimativas para as mesmas estatísticas. Surge então a dúvida de qual qual dentre os possíveis candidados é o melhor. Para isso utiliza-se o conceito de estimadores viesados e não-viesados.

Definição 3 (Viés de Estimadores): *Um estimador $\hat{\theta}$ de um parâmetro θ é dito não-viesado quando o valor esperado deste estimador é igual ao parâmetro θ, ou seja, quando:*

$$E\left[\hat{\theta}\right] = \theta$$

$\bar{X}_n$ como estimador de μ

Podemos ver facilmente que a média amostral $\bar{X}_n$ é um estimador não viesado da média da distribuição pois:

$$\begin{aligned}
\text{E}\left[\bar{X}_n\right] &= \text{E}\left[\frac{X_1 + X_2 + ... + X_n}{n}\right] \\
&= \frac{1}{n}(\text{E}\left[X_1\right] + ... + \text{E}\left[X_n\right]) \\
&= \frac{n\mu}{n} \\
&= \mu
\end{aligned}$$

Estimador de σ^2

Como vimos anteriormente, um estimador da média da distribuição μ é o valor da média amostra $\bar{X}_n$. Poderíamos pensar, num primeiro momento, que poderíamos substituir μ por $\bar{X}_n$ na equação de $\hat{\sigma^2}$ a fim de encontrarmos o melhor estimador da variância da distribuição σ^2. Podemos ver se o estimador resultante contém viés ou não.

Aqui precisamos de um passo intermediário. Podemos ver que:

$$\begin{aligned}
\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 &= \sum_{i=1}^{n}\left[(X_i - \mu) + (\mu - \bar{X}_n)\right]^2 \\
&= \sum_{i=1}^{n}\left[(X_i - \mu)^2 + 2\left(X_i - \mu\right)\left(\mu - \bar{X}_n\right) + \left(\mu - \bar{X}_n\right)^2\right]
\end{aligned}$$

O termo do meio resulta em:

$$\left(X_i - \mu\right)\left(\mu - \bar{X}_n\right) = \mu X_i - \bar{X}_n X_i - \mu^2 + \mu\bar{X}_n$$

e, portanto:

$$\begin{aligned}
\sum_{i=1}^{n} 2\left(X_i - \mu\right)\left(\mu - \bar{X}_n\right) &= 2\mu\sum_{i=1}^{n}X_i - 2\bar{X}_n\sum_{i=1}^{n}X_i + \\
&\quad - 2n\mu^2 + 2n\mu\bar{X}_n \\
&= 2n\mu\bar{X}_n - 2n\bar{X}_n^2 - 2n\mu^2 + 2n\mu\bar{X}_n \\
&= -2n(\mu - \bar{X}_n)^2
\end{aligned}$$

Voltando à equação inicial, temos:

$$\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 = \left[\sum_{i=1}^{n}\left(X_i - \mu\right)^2\right] - n(\mu - \bar{X}_n)^2$$

Com este resultado em mãos, podemos calcular o valor esperado desejado:

$$\begin{aligned}
\text{E}\left[\hat{\sigma^2}\right] &= \text{E}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right] \\
&= \frac{1}{n}\text{E}\left[\left[\sum_{i=1}^{n}\left(X_i - \mu\right)^2\right] - n(\mu - \bar{X}_n)^2\right] \\
&= \frac{1}{n}\left[\sum_{i=1}^{n}\text{E}\left[(X_i - \mu)^2\right]\right] - \text{E}\left[(\mu - \bar{X}_n)^2\right] \\
&= \frac{n\sigma^2}{n} - \frac{\sigma^2}{n} \\
&= \frac{(n-1)\sigma^2}{n}
\end{aligned}$$

Logo vemos que este seria um estimador viesado. Entretanto, podemos ver facilmente que se utilizarmos S^2 como um estimador da variância da distribuição, a média deste estimador será igual à σ^2. Portanto S^2 é um bom candidato a ser estimador desta estatística.

DISTRIBUIÇÃO t-STUDENT

Imagine que tenhamos coletado n variáveis aleatórias de um experimento e calculamos o valor médio destas variáveis, como sendo um estimador do valor médio $\bar{X}$ da distribuição da qual estas variáveis foram obtidas. Imagine que estamos interessados em estimar o quão perto este valor está da média μ verdadeira da distribuição. Isto pode ser feito através do uso da distribuição t-Student.

Definição 4 (Distribuição t-Student): *Seja Z uma variável aleatória proveniente de uma distribuição normal padrão e Y uma variável aleatória proveniente de uma distribuição χ^2 com $n-1$ graus de liberdade, então a variável aleatória U:*

$$U = \frac{Z}{\left(\frac{Y}{n-1}\right)^{1/2}},$$

terá densidade de probabilidade dada por:

$$f(u) = \frac{\Gamma(n/2)}{\Gamma\big((n-1)/2\big)\,\sqrt{\pi(n-1)}} \left(1 + \frac{u^2}{(n-1)}\right)^{-n/2},$$

com $-\infty < u < \infty$,chamada de distribuição de t-Student *com $n-1$ graus de liberdade.*

Do teorema do limite central, sabemos que quando temos n variáveis aleatórias coletadas de uma distribuição com média μ e variância finita σ^2, o $\bar{X}$ destas n variáveis segue uma distribuição normal e $Z = (\bar{X} - \mu)/(\sigma/\sqrt{n})$ segue uma distribuição normal padrão.

Por outro lado, vimos que a variável $Y = \sum_{i=1}^{n}(X_i - \bar{X})^2/\sigma^2$ segue uma distribuição do tipo χ^2 com $n-1$ graus de liberdade. Assim, podemos encontrar a variável U neste caso como sendo:

$$U = \frac{(\bar{X} - \mu)/(\sigma/\sqrt{n})}{\left(\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2/\sigma^2}{n-1}\right)^{1/2}}$$

que pode ser re-escrita como sendo:

$$U = \frac{\sqrt{n}(\bar{X} - \mu)}{(S^2)^{1/2}}, \qquad (1)$$

onde S^2, já definido, é dado por $S^2 = (1/(n-1))\sum_{i=1}^{n}(X_i - \bar{X})^2$. Podemos ver que U não depende de σ, o que é bastante importante para esta análise.

Exemplo 7

Uma amostra aleatória de 9 elementos é coletada de uma distribuição normal com média 20 e variância desconhecida. É provável que a média destes elementos tenha média 24 e desvio padrão de amostra de 4,1?

Que conclusões se pode tirar?

Como não temos a variância da distribuição normal, então este é um caso bastante propício para que se use a distribuição t-Student. O valor de U é dado por:

$$U = \frac{\sqrt{n}(\bar{X} - \mu)}{(S^2)^{1/2}} = \frac{\sqrt{9}(24-20)}{4,1} \sim 2,927$$

À partir deste valor, podemos verificar em uma tabela que a c.d.f. para este valor de U para este valor e 8 graus de liberdade é igual a $\approx$ 0,99, o que significa que estas combinações de valores são bastante improváveis. Isto significa que 24 é um péssimo estimador da média da distribuição.

Intervalos de Confiança

A distribuição de t-Student é bastante útil em outro aspecto. Imaginemos que quando calculamos uma média $\bar{X}$ de n variáveis aleatórias, queremos saber quão boa é a estimativa de que $\mu = \bar{X}$. A distribuição de t-Student é bastante útil neste caso pois não precisamos saber a variância da distribuição original.

Vamos supor que calculamos a variável U para as variáveis aleatórias observadas. A partir deste valor queremos saber a probabiliade γ de que U esteja dentro de um intervalo $-c < U < +c$. Se quisermos saber se $\bar{X} - \mu$ dentro de determinado intervalo de confiança γ (por exemplo, 50%), então queremos achar o valor de c tal que:

$$\Pr(-c < U < +c) = \gamma$$

Utilizando a equação (1), e com um pouco de álgebra elementar, acabamos com:

$$\Pr\left(\bar{X} - \frac{c\,S}{\sqrt{n}} < \mu < \bar{X} + \frac{c\,S}{\sqrt{n}}\right) = \gamma \qquad (2)$$

Se T for a f.d.c. de U, temos então que:

$$\Pr(-c < U < +c) = \gamma = T(c) - T(-c) = T(c) - (1 - T(c)) = 2T(c) - 1,$$

onde usou-se o fato de que U tem uma f.d.p. simétrica. A partir da equação acima podemos encontrar o valor de c, pois:

$$T(c)^{-1} = \frac{\gamma + 1}{2}.$$

Uma vez tendo c, podemos voltar na equação (2) e ver que este valor de c corresponde à probabilidade γ de μ estar entre $\bar{X} \pm \frac{c\,S}{\sqrt{n}}$.

Exemplo 8

Considere que uma série de 8 observações foram feitas de uma distribuição normal com média e variância desconhecidas. Os valores observdos foram: 3,1; 3,5; 2,6; 3,4; 3,8; 3,0; 2,9 e 2,2. Encontre o menor intervalo de confiança para o valor de μ de acordo com os seguintes coeficientes de confiança γ: $a)$ 0,9, $b)$ 0,95 e $c)$ 0,99.

Primeiramente vamos encontrar o valor de c. Para o ítem a temos que $T(c)^{-1} = (0.90+1)/2 = 0,95$. Olhando na tabela dos valores de t-Student vemos que para 7 graus de liberdade temos $c = 1,415$. Calculando o valor da média e do desvio padrão da amostra temos, respectivamente, $\bar{X} = 3,0625$ e $S = 0,5125$. Assim vemos que $\mu = 3,0625 \pm 0,256$ com 90% de confiança. No caso do item b temos $c = 1,895$ e consequentemente $\mu = 3,0625 \pm 0,343$ com 95% de confiança. No ítem c temos $c = 2,998$ e $\mu = 3,0625 \pm 0,543$ com 99% de confiança.