estudar.com.vc

Estatística II

Lista Extra de Exercícios

1. IC da Média com a Variância Populacional Conhecida

Magalhães e Lima, cap.7

Por analogia a produtos similares, o tempo de reação de um novo medicamento pode ser considerado como tendo distribuição normal com desvio padrão igual a 2 minutos (a média é desconhecida). Vinte pacientes foram sorteados, receberam o medicamento e tiveram o seu tempo de reação anotado. Os dados foram os seguintes (em minutos): 2,9; 3,4; 3,5; 4,1;4,6;4,7; 4,5; 3,8; 5,3; 4,9; 4,8; 5,7; 5,8; 5,0; 3,4; 5,9; 6,3; 4,6; 5,5; e 6,2. Obtenha um intervalo de confiança para o tempo médio de reação. Use γ = 96%.

2. Propriedades de Estimadores

Bussab e Morettin, cap. 11

Suponha que X tenha uma distribuição uniforme no intervalo (0; α) onde α é desconhecido. Uma amostra de *n* observações $X_1, X_2, \ldots, X_n$ é retirada. Sabemos que $E(X) = E\ (X_i) = \frac{\alpha}{2}$ e $Var(X) = Var(X_i) = \frac{\alpha^2}{12}$ para todo *i = 1, ..., n.* Logo, se calcularmos a média amostral $\bar{X}$, essa deve estar próxima de $\frac{\alpha}{2}$ e podemos estimar α por $\hat{\alpha} = 2\bar{X}$.

a. Calcule $E(\hat{\alpha})$.
b. Calcule $EQM(\hat{\alpha})$.
c. $\hat{\alpha}$ é consistente? Por quê?

3. IC da Proporção

Magalhães e Lima, cap.7

Numa pesquisa com 50 leitores, o candidato José João obteve a preferência de 17 desses eleitores. Supondo que a eleição ocorresse na época da pesquisa, construa os intervalos de confiança otimista e conservador para a proporção de votos a serem recebidos pelo candidato mencionado. Use um coeficiente de confiança igual 94%.

4. IC da Variância

O peso de componentes mecânicos produzidos por uma determinada empresa é uma variável aleatória que se supõe ter distribuição Normal. Pretende-se estudar a variabilidade do peso dos referidos componentes. Para isso, uma amostra de tamanho 11 foi obtida, cujos valores em grama são:

98 97 102 100 98 101 102 105 95 102 100

Construa um intervalo de confiança para a variância do peso, com grau de confiança igual a 95%.

5. IC da Média com Variância Populacional Desconhecida

Adaptado, Magalhães e Lima, cap.7

Uma amostra de 25 observações de uma Normal foi coletada e forneceu uma média amostral de 8 e um desvio de 4. Construa intervalos de confiança com 80% e 95% para a média populacional. Comente as diferenças.

6. Propriedades de Estimadores e EMV

Fonseca e Tores, vol.2

Sabendo que certa população tem distribuição Binomial, com parâmetro p desconhecido, obtenha o estimador de máxima verossimilhança de p, com base numa amostra aleatória de dimensão n, $(X_1, X_2, \dots, X_n)$, dela obtida. Proceda à análise do estimador.

7. IC da Variância

Em uma fábrica, uma amostra de 30 parafusos apresentou os seguintes diâmetros (em mm):

10 13 14 11 13 14 11 13 14 15 12 14 15 13 14 12 12 11 15 16 13 15 14 14 15 15 16 12 10 15

Supondo que os diâmetros sejam aproximadamente normais, obtenha um intervalo de confiança para a variância do diâmetro de todos os parafusos produzidos nessa fábrica, usando o nível de confiança de 98%. Para facilitar a solução do exercício, você pode usar os seguintes resultados:

$$\sum_{i=1}^{30} X_i = 401 \qquad \sum_{i=1}^{30} {X_i}^2 = 5443$$

8. IC da Proporção

Um teste realizado com 280 pessoas consiste em “adivinhar” em qual das mãos (ambas fechadas) do pesquisador estava uma moeda. Em 44% das tentativas a identificação foi correta da mão selecionada. Com um coeficiente de confiança de 95%, construa um intervalo de confiança para o verdadeiro valor de vezes que foi selecionada a mão correta.

9. IC da Média com a Variância Populacional Conhecida

Magalhães e Lima, cap.7

Será coletada uma amostra de uma população Normal com desvio padrão igual a 9. Para uma confiança de γ = 90%, determine a amplitude do intervalo de confiança para a média populacional nos casos em que o tamanho da amostra é 30 ou 100. Comente as diferenças.

10. Estimador de Máxima Verossimilhança

Adaptado, Fonseca e Torres, vol.2

Seja $(X_1, X_2, \dots, X_n)$ uma amostra aleatória extraída de uma população com função densidade da forma:

$$f(x) = \begin{cases} \dfrac{x^2 e^{-\frac{x}{\omega}}}{2\omega^3}; & x > 0, \quad \omega > 0 \\ 0 \quad ; & c.c. \end{cases}$$

Identifique o estimador de máxima verossimilhança do parâmetro desconhecido, com base na amostra aleatória.

11. IC da Média com a Variância Populacional Desconhecida

Adaptado, Magalhães e Lima, cap.7

Uma amostra de 100 cidades brasileiras, de até 20 mil habitantes, indicou que o valor médio da hora-aula para professores do ensino fundamental em escolas municipais é de R$10 com um desvio padrão de R$2. Obtenha um intervalo de confiança para o valor médio nacional da hora-aula em cidades do tipo mencionado. Use γ = 0,95 como coeficiente de confiança.

Gabarito

1.

Passo 1: DEFINIR A VARIÁVEL DE INTERESSE

Para isso precisamos reler o enunciado e descobrir qual a variável que ele quer estudar, fazer inferências a partir dos dados coletados. Logo no final ele fala "Obtenha um intervalo de confiança para o tempo médio de reação", portanto a minha variável de interesse será:

X: Tempo de reação dos pacientes ao novo medicamento.

OBS: NÃO é o tempo médio (tempo médio =$\bar{X}$)

Passo 2: DESCOBRIR QUAL TIPO DE INTERVALO DE CONFIANÇA ESTÁ SENDO PEDIDO

No enunciado ele pede um intervalo do tempo médio de reação, ou seja, ele quer utilizar um intervalo de confiança da média populacional para tentar estimar o verdadeiro valor do tempo de reação dos pacientes que fazem uso do novo medicamento. Ainda no enunciado ele nos informa que a média é desconhecida e não fornece nenhuma informação do tipo sobre o desvio padrão populacional (além do seu valor, 2 minutos), dessa forma podemos assumir que o mesmo é conhecido e estamos falando de um caso de Intervalo de Confiança da Média com Desvio Padrão Conhecido.

Passo 3: APLICAÇÃO DA FÓRMULA E CÁLCULO DO INTERVALO DE CONFIANÇA

A fórmula do intervalo de confiança é a seguinte:

$$IC(\mu, \gamma) = \left[\bar{X} \pm z\frac{\sigma}{\sqrt{n}}\right]$$

Pelo enunciado nós temos o n (número de itens na amostra) que é igual a 20, o desvio padrão σ = 2 minutos e através dos dados da amostra informados, podemos fazer a média dos mesmos e encontrar o nosso $\bar{X}$.

$$\bar{X} = 4{,}475$$

Para encontrar o Z nós recorremos à tabela da Normal Padrão e com um γ = 96%, obteremos:

$$z = 2{,}05$$

Assim com todas as incógnitas podemos calcular o Intervalo pedido:

$$IC(\mu, \gamma = 96\%) = \left[4{,}475 - 2{,}05\frac{2}{\sqrt{20}}\,; 4{,}475 + 2{,}05\frac{2}{\sqrt{20}}\right]$$

$$IC(\mu, \gamma = 96\%) = [3{,}828; 5{,}662]$$

Podemos inferir, com 96% de confiança/certeza que o verdadeiro valor da média está contido no intervalo a cima.

2.

Passo 1: CALCULAR A ESPERANÇA DO ESTIMADOR

Através do enunciado, sabemos que:

$$E(\hat{\alpha}) = E(2X) = 2E(X) = 2\frac{\alpha}{2} = \alpha$$

Passo 2: CALCULAR EQM

O Erro Quadrático Médio consiste na seguinte fórmula: $EQM(\hat{\alpha}) = Var\ (\hat{\alpha}) + Viés^2(\hat{\alpha})$, portanto precisamos calcula-lo em duas partes:

$$Viés\ (\hat{\alpha}) = E(\hat{\alpha}) - \alpha = \alpha - \alpha = 0$$

Assim, podemos afirmar que o parâmetro é não-viesado.

$$Var\ (\hat{\alpha}) = Var\ (2X) = 2^2 Var(X) = 4\frac{\alpha^2}{12} = \frac{\alpha^2}{3}$$

Agora com ambas as partes, calculamos o $EQM(\hat{\alpha})$:

$$EQM(\hat{\alpha}) = \frac{\alpha^2}{3} + 0^2 = \frac{\alpha^2}{3}$$

Passo 3: CONSISTÊNCIA

O estimador proposto para o parâmetro α não pode ser considerado consistente pelo fato de que apesar de ser não-viesado, quando meu n tende a infinito, a $Var(\hat{\alpha})$ não tende a zero:

$$\lim_{n\to\infty} Viés\ (\hat{\alpha}) = 0 \qquad\qquad \lim_{n\to\infty} Var\ (\hat{\alpha}) \neq 0$$

E se ambos não tendem a zero juntos, meu Erro Quadrático Médio não tende a zero, ou seja, no limite, meu estimador não é consistente:

$$\lim_{n\to\infty} EQM(\hat{\alpha}) \neq 0$$

3.

Passo 1: DEFINIR A VARIÁVEL DE INTERESSE

Ao reler o enunciado, percebemos que a variável a qual o mesmo busca estudar é binária (Bernoulli), ou seja, X só assume valores de 1 e 0, sucesso ou fracasso, respectivamente, sendo:

X = 1, o candidato José João obteve a preferência

X = 0, outro candidato obteve a preferência

Passo 2: DESCOBRIR QUAL TIPO DE INTERVALO DE CONFIANÇA ESTÁ SENDO PEDIDO

É pedido claramente no exercício o intervalo de confiança da proporção de votos recebido por determinado candidato.

Passo 3: APLICAÇÃO DA FÓRMULA E CÁLCULO DO INTERVALO DE CONFIANÇA

A fórmula do intervalo de confiança é a seguinte:

$$IC(p,\gamma) = \left[\hat{p} \pm z\sqrt{\frac{p(1-p)}{n}}\right]$$

É mencionado que a amostra utilizada possuía um n = 50 observações de eleitores, e também é pedido para que se explore ambos os cenários, tanto otimista quanto conservador.

$$Otimista: p = \hat{p}$$
$$Conservador: p = 0{,}5$$

Para encontrar o $\hat{p}$ é necessário calcular com os dados fornecidos no enunciado. Sabemos que $\hat{p}$ é a probabilidade de sucesso amostral e definimos como sucesso a intenção/preferência de voto ao candidato João José. Dessa forma, se temos 17 pessoas com tal preferência em um grupo de 50:

$$\hat{p} = \frac{17}{50} = 34\%\ ou\ 0{,}34$$

O último fator necessário é o z o qual encontraremos através da tabela Normal, com um γ = 94%:

$$z = 1{,}88$$

Com todas as incógnitas conseguimos obter:

$$IC_{Otimista}(p,\gamma = 94\%) = \left[0{,}34 - 1{,}88\sqrt{\frac{0{,}34(0{,}66)}{50}}\ ;\ 0{,}34 + 1{,}88\sqrt{\frac{0{,}34(0{,}66)}{50}}\right]$$

$$IC_{Otimista}(p,\gamma = 94\%) = [0{,}214\ ;0{,}466]$$

$$IC_{Conservador}(p,\gamma = 94\%) = \left[0{,}34 - 1{,}88\sqrt{\frac{0{,}5(0{,}5)}{50}}\ ;\ 0{,}34 + 1{,}88\sqrt{\frac{0{,}5(0{,}5)}{50}}\right]$$

$$IC_{Conservador}(p,\gamma = 94\%) = [0{,}207\ ;0{,}473]$$

Podemos inferir, com 94% de confiança/certeza que o verdadeiro valor da proporção de eleitores que prefeririam votar no candidato João José está contido no intervalo a cima.

4.

Passo 1: DEFINIR A VARIÁVEL DE INTERESSE

Ao reler o enunciado, percebemos que a variável que se quer obter informações através de um intervalo de confiança é:

X: O peso de componentes mecânicos produzidos por uma determinada empresa.

OBS: Não é a variabilidade (S^2 = Variabilidade/Variância)

Passo 2: DESCOBRIR QUAL TIPO DE INTERVALO DE CONFIANÇA ESTÁ SENDO PEDIDO

No enunciado é mencionado que "pretende-se estudar a variabilidade do peso dos requeridos componentes", portanto, a única forma de estudar tal variabilidade é através de um intervalo de confiança da variância.

Passo 3: APLICAÇÃO DA FÓRMULA E CÁLCULO DO INTERVALO DE CONFIANÇA

A fórmula do intervalo de confiança é a seguinte:

$$IC(\sigma^2,\gamma) = \left[\frac{(n-1)S^2}{q_2};\ \frac{(n-1)S^2}{q_1}\right]$$

Com as informações que nos foram fornecidas, nós sabemos que a amostra tem um número n de 11 observações, e a variância amostral S^2 não é fornecida, porém, com todos os dados da amostra dados no enunciado podemos calcula-la facilmente.

$$\bar{X} = \sum \frac{X_i}{n} = 100$$

$$S^2 = \frac{\sum(X_i - \bar{X})^2}{n-1} = 8$$

Os únicos dados que faltam ser calculados são o q_1 e o q_2, que encontraremos com ajuda da tabela *qui-quadrada* com n-1= 10 graus de liberdade e o γ = 95% fornecido pelo enunciado:

$$q_1 = 3{,}25$$

$$q_2 = 20{,}48$$

Com todas as incógnitas conseguimos obter:

$$IC(\sigma^2, \gamma = 95\%) = \left[\frac{(11-1)\,8}{20{,}48};\ \frac{(11-1)\,8}{3{,}25}\right]$$

$$IC(\sigma^2, \gamma = 95\%) = [3{,}90;\ \ 24{,}61]$$

Podemos inferir, com 95% de confiança/certeza que o verdadeiro valor da variância do peso de componentes mecânicos produzidos por uma determinada empresa está contido no intervalo a cima.

5.

Passo 1: DESCOBRIR QUAL TIPO DE INTERVALO DE CONFIANÇA ESTÁ SENDO PEDIDO

Esse exercício é mais conceitual e, portanto, direto, não tendo uma variável de interesse definida, e, portanto, podemos pular direto para o tipo de intervalo de confiança com qual estamos trabalhando.

Nesse caso ele já especifica no próprio enunciado que quer 2 intervalos de confiança para a média populacional e nos resta saber se o desvio padrão populacional é conhecido ou não, para saber qual fórmula devemos aplicar no cálculo do intervalo.

Ainda no enunciado quando ele fala "foi coletada (uma amostra) e forneceu média amostral de 8 e um desvio de 4" podemos perceber que esse desvio mencionado é proveniente da amostra de 25 observações coletadas e portanto é o que nós chamamos de amostral, caracterizando então o intervalo como de média com desvio padrão populacional conhecido.

Passo 2: APLICAÇÃO DA FÓRMULA E CÁLCULO DO INTERVALO DE CONFIANÇA

A fórmula do intervalo de confiança é a seguinte:

$$IC(\mu,\gamma) = \left[\bar{X} \pm t\frac{s}{\sqrt{n}}\right]$$

O enunciado nos fornece o valor de n (número de observações na amostra) igual a 25, uma média amostral $\bar{X}$ = 8 e o desvio padrão também amostral *s* = 4. Dessa forma, a única incógnita restante é o *t* que para consegui-lo, recorremos à tabela *t-student* com n-1 = 24 graus de liberdade. Um γ = 80% e γ = 95%, nos permite obter, respectivamente:

$$t_{80\%} = \ 1{,}318 \ e \ t_{95\%} = 2{,}064$$

Assim podemos calcular ambos os intervalos pedidos e enfim compará-los:

$$IC\ (\mu,\gamma = 80\%) = \left[8 - 1{,}318\frac{4}{\sqrt{25}}\ ; 8 + 1{,}318\frac{4}{\sqrt{25}}\right]$$

$$IC\ (\mu,\gamma = 80\%) = [6{,}945\ ; 9{,}054]$$

Amplitude = 2,1

$$IC\ (\mu, \gamma = 95\%) = \left[8 - 2{,}064\frac{4}{\sqrt{25}}\ ; 8 + 2{,}064\frac{4}{\sqrt{25}}\right]$$
$$IC\ (\mu, \gamma = 95\%) = [6{,}348\ ; 9{,}651]$$

Amplitude = 3,3

Passo 3: COMPARAÇÃO DE AMBOS OS INTERVALOS

O intervalo com maior nível de confiança (95%) possui uma maior amplitude, ou seja, um maior intervalo. A lógica disso está na interpretação do intervalo de confiança, está sendo "Acredito com x% de nível de significância que o intervalo contém o verdadeiro valor da média". Quanto maior a minha "certeza" de que o intervalo contém o verdadeiro valor da média, maior tem que ser as minhas margens de erro para que 95% das vezes eu esteja "certo", ou seja, eu aumento as minhas margens para que eu erre menos.

6.

Passo 1: FUNÇÃO DE PROBABILIDADE E DE VEROSSIMILHANÇA

Como se trata de uma Bernoulli, sabemos que sua função de probabilidade é igual a:

$$f(x \mid p) = p^x(1-p)^{1-x}$$
$$Com\ x = 0\ ou\ x = 1$$

Dessa forma, nossa função $L(p)$ de máxima verossimilhança, nada mais é do que o produtório da função de probabilidade apresentada a cima. Como não podemos derivar e igualar a zero um produtório, transformaremos a função $L(p)$ em uma $l(p)$ com o uso do logarítmo:

$$L(p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

$$l(p) = \ln\left(\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}\right)$$

Passo 2: SIMPLIFICAÇÃO E DERIVAÇÃO DA FUNÇÃO l

Para Maximizar a Verossimilhança, necessitamos derivar tal função e iguala-la a zero, porém antes disso, simplificá-la através de propriedades de estimadores para facilitar tal tarefa.

$$l(p) = \ln\left(\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}\right)$$

Pelas propriedades de logaritmos: $\log_a^{(b.c)} = \log_a b + \log_a c$, portanto:

$$l(p) = \sum_{i=1}^{n} \ln\,(p^{x_i}\,(1-p)^{1-x_i}) = \sum_{i=1}^{n} [\ln p^{x_i} + \ln(1-p)^{1-x_i}]$$

Também sabemos que: $\log_a b^m = m.\log_a b$

$$l(p) = \sum_{i=1}^{n} [x_i \ln p + \ln(1-p)] - x_i \ln\ (1-p)]$$

Como *ln p* é considerado uma constante (não depende do *i* do somatório, podemos tirar ele do somatório:

$$l(p) = \ln p\left(\sum_{i=1}^{n} X_i\right) + \ln(1-p)\left(\sum_{i=1}^{n} 1 - x_i\right)$$

Simplificada ao máximo a função, podemos derivá-la em função do parâmetro p:

$$\frac{d\,l(p)}{d\,p} = 0 + \frac{1}{p}\sum_{i=1}^{n} x_i + n\frac{-1}{1-p} - \sum_{i=1}^{n} x_i \frac{-1}{1-p}$$

Quando igualamos a zero e eliminamos os denominadores, temos:

$$(1-\hat{p})\sum_{i=1}^{n} x_i - n\hat{p} + \hat{p}\sum_{i=1}^{n} x_i = 0$$

$$\sum_{i=1}^{n} x_i - \hat{p}\sum_{i=1}^{n} x_i - \hat{p}n + \hat{p}\sum_{i=1}^{n} x_i = 0$$

$$\hat{p} = \sum_{i=1}^{n} \frac{X_i}{n}$$

Passo 3: ANÁLISE DO ESTIMADOR – VIÉS

Um estimador é não-viesado, desde que $E(\hat{\theta}) = \theta$. Consequentemente,

$$E(\hat{p}) = E\left(\sum_{i=1}^{n} \frac{X_i}{n}\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} p = \frac{1}{n}\,np = p$$

Assim, pode-se dizer que $\hat{p}$ é um estimador não viesado para p.

Passo 4: ANÁLISE DO ESTIMADOR – CONSISTÊNCIA

Para analisar a consistência é necessário o erro quadrático médio. Dessa forma, já analisamos o viés e agora teremos que analisar a variância.

$$Var(\hat{p}) = Var\left(\sum_{i=1}^{n} \frac{X_i}{n}\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right)$$
$$= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} p(1-p) = \frac{1}{n^2}\, n\, p(1-p) = \frac{p(1-p)}{n}$$

Sendo $\lim\limits_{n\to\infty} \frac{p(1-p)}{n} = 0$,

Por ser não-viesado e por verificar a condição de que no limite, sua variância tende a zero, diz-se que $\hat{p}$ é um estimador coerente p.

7.

Passo 1: DEFINIR A VARIÁVEL DE INTERESSE

Ao reler o enunciado, percebemos que a variável que se quer obter informações através de um intervalo de confiança é:

X: O diâmetro dos parafusos produzidos em uma fábrica.

OBS: Não é a variabilidade (S^2 = Variabilidade/Variância)

Passo 2: DESCOBRIR QUAL TIPO DE INTERVALO DE CONFIANÇA ESTÁ SENDO PEDIDO

No enunciado é pedido objetivamente um intervalo de confiança para a variância do diâmetro de todos os parafusos produzidos nessa fábrica.

Passo 3: APLICAÇÃO DA FÓRMULA E CÁLCULO DO INTERVALO DE CONFIANÇA

A fórmula do intervalo de confiança é a seguinte:

$$IC(\sigma^2, \gamma) = \left[\frac{(n-1)S^2}{q_2}; \frac{(n-1)S^2}{q_1}\right]$$

Com as informações que nos foram fornecidas, nós sabemos que a amostra tem um número n de 30 observações, e a variância amostral S^2 não é fornecida, porém, com todos os dados da amostra dados no enunciado podemos calcula-la facilmente.

$$S^2 = \sum_{i=1}^{30} \frac{X_i^2 - \bar{X}^2}{n-1} = 2{,}86$$

Os únicos dados que faltam ser calculados são o q_1 e o q_2, que encontraremos com ajuda da tabela *qui-quadrada* com n-1 = 29 graus de liberdade e o γ = 98% fornecido pelo enunciado:

$$q_1 = 14{,}25$$
$$q_2 = 49{,}58$$

Com todas as incógnitas conseguimos obter:

$$IC(\sigma^2, \gamma = 98\%) = \left[\frac{(30-1)\ 2{,}86}{49{,}58}; \frac{(30-1)\ 2{,}86}{14{,}25}\right]$$
$$IC(\sigma^2, \gamma = 98\%) = [1{,}67;\ \ 5{,}82]$$

Podemos inferir, com 98% de confiança/certeza que o verdadeiro valor da variância do diâmetro dos parafusos produzidos pela fábrica está contido no intervalo a cima.

8.

Passo 1: DEFINIR A VARIÁVEL DE INTERESSE

Ao reler o enunciado, percebemos que a variável a qual o mesmo busca estudar é binária (Bernoulli), ou seja, X só assume valores de 1 e 0, sucesso ou fracasso, respectivamente, sendo:

X = 1, a pessoa escolheu a mão correta (com a moeda)
X = 0, a pessoa escolheu a mão errada.

Passo 2: DESCOBRIR QUAL TIPO DE INTERVALO DE CONFIANÇA ESTÁ SENDO PEDIDO

Ele pede no enunciado o intervalo para o verdadeiro valor de vezes que foi selecionada a mão correta. Como sabemos que tal variável se comporta de forma binária, e no enunciado ele me fornece uma proporção amostral de sucesso (x=1) é seguro afirmar que para estudar o verdadeiro valor de vezes que a mão correta foi selecionada usaremos um intervalo de confiança da proporção.

Passo 3: APLICAÇÃO DA FÓRMULA E CÁLCULO DO INTERVALO DE CONFIANÇA

A fórmula do intervalo de confiança é a seguinte:

$$IC(p,\gamma) = \left[\hat{p} \pm z\sqrt{\frac{p(1-p)}{n}}\right]$$

Dentre as informações fornecidas no enunciado temos que, o número n de observações na amostra é igual a 280, e a proporção de sucessos $\hat{p}$ encontrada através da mesma é de 44%.

O *z* será encontrado com ajuda da tabela da normal com um γ = 95%:

$$z = 1{,}96$$

O que nos resta decidir é se seremos otimistas ou conservadores na construção do intervalo. Nesse caso, por simples preferência, optaremos pelo cenário otimista.

$$IC_{Otimista}(p, \gamma = 95\%) = \left[0{,}44 - 1{,}96\sqrt{\frac{0{,}44(1 - 0{,}44)}{280}}\ ;\ 0{,}44 + 1{,}96\sqrt{\frac{0{,}44(1 - 0{,}44)}{280}}\right]$$

$$IC_{Otimista}(p, \gamma = 95\%) = [0{,}381\ ;0{,}498]$$

Podemos inferir, com 95% de confiança/certeza que o verdadeiro valor da proporção de pessoas que adivinhavam corretamente em qual mão estava a moeda está contido no intervalo a cima.

9.

Passo 1: DESCOBRIR QUAL TIPO DE INTERVALO DE CONFIANÇA ESTÁ SENDO PEDIDO

Não é necessário definir a variável de interesse nesse caso pois o exercício é objetivo e não especifica o que está sendo estudado. Nesse caso, partimos direto para a definição do tipo de intervalo de confiança e podemos perceber que o próprio enunciado já sinaliza que é um intervalo da média porém é necessário um olhar mais crítico para ver se o desvio padrão populacional é conhecido ou não.

Como o enunciado em nenhum momento associa o desvio padrão mencionado à uma amostra, podemos assumir que o mesmo é populacional e portanto utilizaremos a fórmula do intervalo de confiança da média com desvio padrão populacional conhecido nos próximos passos.

Passo 2: APLICAÇÃO DA FÓRMULA E CÁLCULO DO INTERVALO DE CONFIANÇA

A fórmula do intervalo de confiança é a seguinte:

$$IC(\mu, \gamma) = \left[\bar{X} \pm z\frac{\sigma}{\sqrt{n}}\right]$$

Pelo enunciado percebemos que ele não quer efetivamente o intervalo de confiança, mas sim a amplitude do mesmo, ou seja qual o tamanho do mesmo (Amplitude = Limite superior – Limite inferior).

$$Amplitude = \left(\bar{X} + z\frac{\sigma}{\sqrt{n}}\right) - \left(\bar{X} - z\frac{\sigma}{\sqrt{n}}\right) = \bar{X} + z\frac{\sigma}{\sqrt{n}} - \bar{X} + z\frac{\sigma}{\sqrt{n}} = \boldsymbol{2z\frac{\sigma}{\sqrt{n}}}$$

Agora sabemos quais variáveis necessitamos e quais estão presentes no enunciado. O desvio padrão populacional σ apresentado é igual a 9 e o exercício pede que calculemos dois cenários, um com um n = 30 e outro com n = 100. A única informação que falta é o z, que recorreremos à tabela da normal para encontra-lo. Com um nível de confiança γ = 90%, temos:

$$Z = 1{,}64 \ ou \ Z = 1{,}65$$

Com todas as incógnitas conseguimos obter:

$$Amplitude_{n=30} = 2 \,.\, 1{,}64 \, \frac{9}{\sqrt{30}} = 5{,}389$$

$$Amplitude_{n=100} = 2 \,.\, 1{,}64 \, \frac{9}{\sqrt{100}} = 2{,}952$$

Passo 3: COMPARAÇÃO DE AMBAS AS AMPLITUDES

Podemos perceber através de ambas as amplitudes que, quanto maior o número de observações na amostra, menor é a amplitude. Isso se dá devido ao fato de que, quanto maior o tamanho da amostra, mais próxima do todo a mesma se encontra, ou seja, eu erro menos e analogamente, preciso de margens de erro menores, diminuindo assim a amplitude de tal intervalo.

10.

Passo 1: FUNÇÃO DE VEROSSIMILHANÇA

Precisamos obter, em primeiro lugar a função de verossimilhança "ajustada" com base nos logaritmos naturais do produtório das funções densidade.

$$l(\omega) = \ln\left[\prod_{i=1}^{n} \frac{{x_i}^2 e^{-\frac{x_i}{\omega}}}{2\omega^3}\right]$$

Passo 2: SIMPLIFICAÇÃO E DERIVAÇÃO DA FUNÇÃO l

Pelas propriedades de logaritmos: $\log_a^{(b.c)} = \log_a b + \log_a c$ e $\log_a^{(b/c)} = \log_a b - \log_a c$ portanto:

$$l(\omega) = \sum_{i=1}^{n} \ln \frac{{x_i}^2 e^{-\frac{x_i}{\omega}}}{2\omega^3}$$

$$= \sum_{i=1}^{n} [\ln {x_i}^2 e^{-\frac{x_i}{\omega}} - \ln 2\omega^3] = \sum_{i=1}^{n} [\ln x_i^2 + \ln e^{-\frac{x_i}{\omega}} - \ln 2 - \ln \omega^3]$$

Também sabemos que: $\log_a b^m = m.\log_a b$

$$l(\omega) = \sum_{i=1}^{n} \left[2\ln x_i - \frac{x_i}{\omega} \ln e - \ln 2 - 3\ln \omega\right]$$

Como nem todos os itens dentro do somatório dependem de *i*, podemos retirar alguns deles de dentro do somatório:

$$l(\omega) = 2\sum_{i=1}^{n} \ln x_i - \frac{1}{\omega}\sum_{i=1}^{n} x_i - n\ln 2 - n3\ln \omega$$

Simplificada ao máximo a função, podemos derivá-la em função do parâmetro ω. E igualamos a zero em seguida para maximizar a verssimilhança.

$$\frac{d\ l(\omega)}{d\ \omega} = 0 + \frac{1}{\omega^2}\sum_{i=1}^{n} x_i - 0 - \frac{3n}{\omega}$$

$$\frac{1}{\widehat{\omega}^2}\sum_{i=1}^{n} x_i - \frac{3n}{\widehat{\omega}} = 0$$

Agora isolando o parâmetro, temos:

$$\widehat{\omega} = \sum_{i=1}^{n} \frac{x_i}{3n}$$

Um estimador gerado pelo EMV é sempre consistente.

11.

Passo 1: DEFINIR A VARIÁVEL DE INTERESSE

A variável de interesse, a que ele busca estudar através do intervalo de confiança, podemos identificar através do que ele fala no enunciado que é:

X: Valor da hora-aula para professores do ensino fundamental em escolas municipais em cidades brasileiras de até 20 mil habitantes.

OBS: NÃO é o valor médio ($\bar{X}$ = Valor médio)

Passo 2: DESCOBRIR QUAL TIPO DE INTERVALO DE CONFIANÇA ESTÁ SENDO PEDIDO

No enunciado é mencionado que ele quer um intervalo para o valor médio nacional da hora-aula, ou seja, um intervalo de confiança para a média populacional do valor da hora-aula em determinada cidades. Agora precisamos reler o enunciado para saber se o desvio padrão fornecido é proveniente de uma amostra ou não.

Ao se referir à amostra ele fala “Uma amostra de 100 cidades brasileiras, ... , indicou que o valor médio da hora-aula para professores do ensino fundamental em escolas municipais é de R$10 com um desvio padrão de R$2”, dando a entender que tal desvio é amostral e portanto nos indicando que a fórmula a ser utilizada para as demais inferências é a do intervalo de confiança da média com desvio padrão populacional desconhecido.

Passo 3: APLICAÇÃO DA FÓRMULA E CÁLCULO DO INTERVALO DE CONFIANÇA

A fórmula do intervalo de confiança é a seguinte:

$$IC(\mu, \gamma) = \left[\bar{X} \pm t\frac{s}{\sqrt{n}}\right]$$

No enunciado ele nos fala que na amostra coletada de n = 100 cidades brasileiras, a média amostral, ou seja $\bar{X}$ foi igual à R$10 e que o desvio padrão amostral *s* equivalia a R$2. Dessa forma, o único dado faltante para o cálculo do nosso intervalo seria o *t,* o qual recorreremos à tabela *t-student* com n-1 = 99 graus de liberdade e ao γ = 0,95 sinalizado para encontra-lo. Assim, teremos:

$$t = 1{,}984$$

Com todas as incógnitas, conseguimos obter:

$$IC(\mu, \gamma = 95\%) = \left[10 - 1{,}984\frac{2}{\sqrt{100}}\ ;\ \ 10 + 1{,}984\frac{9}{\sqrt{100}}\right]$$

$$IC(\mu, \gamma = 95\%) = [9{,}603; 10{,}397]$$

Podemos inferir, com 95% de confiança/certeza que o verdadeiro valor da média está contido no intervalo a cima.

Bibliografia

BUSSAB, Wilton. O.; MORETTIN, Pedro. A. *Estatística Básica*. 5. São Paulo: Saraiva, 2004.

MAGALHÃES, M. N.; LIMA, A. C. P. *Noções de Probabilidade e Estatística*. 7. São Paulo: Edusp, 2013.

FONSECA, J.; TORRES, D. *Exercícios de Estatística.* 2. Lisboa: Sílabo, 2014.