Testes de hipótese

Intervalos de confiança, formulação de Testes de Hipótese e suas aplicações.

Prof. Raphael Guinâncio Bruce

Propósito

Analisar as aplicações de intervalos de confiança e Testes de Hipótese em inferência estatística a fim de desenvolver a intuição a respeito da incerteza sobre dados, além da capacidade de tirar conclusões a respeito de amostras.

Preparação

Antes de iniciar o conteúdo, tenha papel e lápis em mãos para acompanhar os exemplos e demonstrações. É recomendado algum conhecimento prévio de Cálculo (derivadas, derivadas parciais e integração) e de estimação pontual. Para a resolução de alguns dos exercícios, recomenda-se o uso de calculadora.

Objetivos

- Descrever aspectos conceituais e aplicações de intervalos de confiança.
- Descrever aspectos conceituais e aplicações de Testes de Hipótese.

Introdução

Quando estamos perto de uma eleição importante, os jornais costumam mostrar os resultados de pesquisas de intenção de voto. Em geral, você encontra algo assim:

"O candidato X teria, se a eleição fosse hoje, 21% dos votos, com um intervalo de dois pontos percentuais para mais ou para menos."

Intuitivamente, esse intervalo significa que não podemos ter certeza sobre a votação que o candidato receberia, porque a pesquisa cobre apenas uma parte da população — ou seja, apenas uma amostra. Ainda assim, é improvável que essa votação fique muito distante desses 21%.

Mas o que significa improvável nesse contexto?

Para responder, vamos desenvolver formalmente, neste tema, a teoria por trás dos chamados intervalos de confiança. A partir daí, poderemos testar hipóteses sobre a população de interesse.

A construção de intervalos de confiança

Vamos descrever, neste tópico, um modo de estimar parâmetros chamado de estimação intervalar. Mais especificamente, vamos lidar com um objeto chamado de intervalo de confiança, utilizado para dizer, em números, quão certos estamos sobre uma estimativa. Essa descrição pode parecer abstrata, mas ficará mais clara se partirmos direto para um exemplo.

Seja $X_1, \ldots, X_n$ uma amostra aleatória de uma distribuição com parâmetro θ que queremos estimar. Suponha que observamos as ocorrências $X_1 = x_1, X_2 = x_2, \ldots X_n = x_n$.

Até o momento, sabemos fazer o que chamamos de estimação pontual, ou seja, a partir de uma amostra definir uma estatística $\hat{\theta}$ (É uma função dessa amostra que nos dá um número real. com a qual queremos tentar nos aproximar do valor verdadeiro θ .

Esse estimador pontual, por si só, não nos dá muita informação sobre θ , porém. Em particular, sem nenhuma informação adicional, não temos como saber quão próximo do parâmetro θ está nosso estimador $\hat{\theta}$. Para isso, queremos obter um intervalo que nos dê tal informação, utilizando a abordagem da estimação intervalar.

Por essa abordagem, ao invés de obter somente $\hat{\theta}$ como estimador de θ , também vamos produzir um intervalo em que é provável que esteja o valor verdadeiro θ . Assim, ao invés de dizermos simplesmente que, por exemplo:

$$\hat{\theta} = 34.25$$

diremos que:

$$\left[\hat{\theta}_l, \hat{\theta}_h\right] = [30.69, 37.81]$$

Esperamos encontrar no intervalo acima o valor real de θ . Isto é, produzimos duas estimativas para θ , **uma estimativa superior** $\hat{\theta}_h, e$ **uma estimativa inferior** $\hat{\theta}_l$ **.**

As letras h e l que diferenciam essas estimativas são padrão na literatura de Estatística e representam as palavras em inglês high (alto) e low (baixo).

Vamos formalizar um pouco alguns conceitos. Seja, como no exemplo acima, $X_1, \ldots, X_n$ uma amostra aleatória de uma distribuição com parâmetro θ que queremos estimar. Nosso objetivo é encontrar dois estimadores para θ. São eles:

1. O estimador inferior: $\hat{\theta}_l = \hat{\theta}_l\left(X_1, \ldots X_n\right)$,e
2. O estimador superior: $\hat{\theta}_h = \hat{\theta}_h\left(X_1, \ldots X_n\right)$.

> Lembre-se sempre de que o estimador é uma função da amostra.

O estimador intervalar, ou intervalo de confiança, é dado por $\left[\hat{\theta}_l, \hat{\theta}_h\right]$. Os estimadores $\hat{\theta}_l$ e $\hat{\theta}_h$ são escolhidos de modo que a probabilidade do intervalo $\left[\hat{\theta}_l, \hat{\theta}_h\right]$ ter dentro dele o parâmetro θ é maior que

$1-\alpha$.

Mas o que é $1-\alpha$?

Esse valor é o que chamamos de nível de confiança. Ele diz a probabilidade desse intervalo estimado conter o nosso parâmetro de interesse. Idealmente, queremos que o valor α seja baixo, para que o nível de confiança $1-\alpha$ seja alto.

> Valores comuns para α são 0.01, 0.05 e 0.10 que correspondem, respectivamente, a níveis de confiança de 99%, 95% e 90%.

Assim, quando queremos encontrar o intervalo de confiança de 95% para um parâmetro qualquer θ, precisamos encontrar $\hat{\theta}_l$ e $\hat{\theta}_h$ tais que:

$$P\left(\hat{\theta}_l \leq \theta \leq \hat{\theta}_h\right) \geq 0.95$$

A probabilidade de que o parâmetro θ esteja no intervalo $\left[\hat{\theta}_l, \hat{\theta}_h\right]$ deve ser maior que 95% (P denota probabilidade).

Vamos agora para a definição geral:

- **Definição (estimador intervalar):** Seja $X_1, \ldots, X_n$, uma amostra aleatória de uma distribuição com parâmetro θ a ser estimado. Um **estimador intervalar** com **nível de confiança** $1-\alpha$ consiste em dois estimadores $\hat{\theta}_l = \hat{\theta}_l\left(X_1, \ldots X_n\right)$ e $\hat{\theta}_h\left(X_1, \ldots X_n\right)$ tais que:

$$P\left(\hat{\theta}_l \leq \theta \leq \hat{\theta}_h\right) \geq 1-\alpha$$

Para todo possível valor de θ . De maneira equivalente, podemos dizer que $\left[\hat{\theta}_l, \hat{\theta}_h\right]$ é o intervalo de confiança de $(1-\alpha)100\%$ Note que a condição $P\left(\hat{\theta}_l \leq \theta \leq \hat{\theta}_h\right) \geq 1-\alpha$ pode ser reescrita como:

$$P\left(\hat{\theta}_l \leq \theta e \hat{\theta}_h \geq \theta\right) \geq 1-\alpha \text{ ou } P\left(\theta \in\left[\hat{\theta}_l, \hat{\theta}_h\right]\right) \geq 1-\alpha$$

Mas como podemos obter intervalos de confiança na prática?

Vamos revisar rapidamente um fato simples sobre variáveis aleatórias e suas distribuições. Seja X uma variável aleatória com função de distribuição acumulada $F_X = P(X \leq x)$. Suponha que queremos encontrar dois valores x_h e x_l tais que:

$$P\left(x_l \leq X \leq x_h\right)=1-\alpha$$

Um jeito simples de fazer isso é escolher x_l e x_h tais que:

$$\begin{gathered}
P\left(X \leq x_l\right)=\frac{\alpha}{2} \\
P\left(X \geq x_h\right)=\frac{\alpha}{2}
\end{gathered}$$

Essas duas probabilidades somadas resultam em α . Elas representam as regiões da distribuição em que X está fora do intervalo $\left[x_l, x_h\right]$. De maneira equivalente, temos as distribuições acumuladas, denotadas por F :

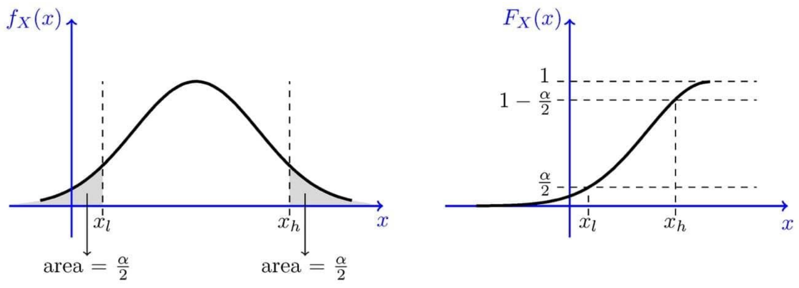
$$\begin{gathered}
F_X\left(x_l\right)=\frac{\alpha}{2} \\
F_X\left(x_h\right)=1-P\left(X \geq x_h\right)=1-\frac{\alpha}{2}
\end{gathered}$$

Podemos obter x_h e x_l como função inversa da distribuição acumulada:

$$\begin{gathered}
x_l=F_X^{-1}\left(\frac{\alpha}{2}\right) \\
x_h=F_X^{-1}\left(1-\frac{\alpha}{2}\right)
\end{gathered}$$

O que o resultado anterior nos diz?

Diz que tanto nosso limite inferior x_l quanto nosso limite superior x_h serão função do nível de significância α. As figuras, a seguir, com a função de distribuição marginal $f_X(x)$ e a função distribuição acumulada $F_X(x)$ ajudarão a deixar mais clara essa relação.



A figura à esquerda, com a distribuição marginal de X , nos dá uma ideia mais clara. O intervalo $[x_l, x_h]$ no eixo x define o nosso intervalo de confiança. A área acima desse intervalo é dada por $P(x_l \leq X \leq x_h)$, que por sua vez é igual a $1-\alpha$. A área cinza, dada por $\frac{\alpha}{2}+\frac{\alpha}{2}=\alpha$ nos diz a probabilidade de estarmos fora do intervalo de confiança.

A figura à direita, por sua vez, nos dá o gráfico função de distribuição acumulada F_X. É possível ver que se aumentamos α, o intervalo $[x_l, x_h]$ nessa figura também diminui. Assim, à medida que aumentamos α, "apertamos" o intervalo de confiança ao aumentar a área à direita de x_h e à esquerda de x_l.

Vejamos um exemplo prático envolvendo uma distribuição bem conhecida, a normal. Tome uma amostra aleatória $X_1, \ldots, X_n$ com distribuição normal $X \sim N(\mu, \sigma^2)$ em que μ é desconhecida e σ^2 é conhecida. Sabemos que um bom estimador pontual de μ é a média amostral $\bar{X}$ (se você não recorda ou não estudou, procure um material de estimação pontual, distribuição amostral e lei dos grandes números). Ao mesmo tempo que $\bar{X}$ é um estimador com boas propriedades, sabemos que a estimativa $\bar{x}$ está sujeita a um erro aleatório, uma vez que é obtida por meio de uma amostra. Isso faz com que o processo de estimação apresente o seguinte resultado:

$$\bar{x} = \mu \pm \epsilon$$

Nossa estimativa $\bar{x}$ é igual à média populacional μ somada a certo erro aleatório ϵ. Isso ocorre porque não sabemos quão próxima está a estimativa $\bar{x}$ do verdadeiro parâmetro populacional μ. Assim, queremos um jeito de quantificar esse erro aleatório ϵ. Para isso, iremos usar a estimação intervalar para obter um intervalo de confiança a partir do entendimento que temos a respeito do estimador $\bar{X}$. A partir do Teorema Central do Limite, sabemos que:

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Ou seja, tomamos a média amostral, subtraímos dela seu valor esperado, e dividimos o resultado pelo seu desvio padrão. Lembre-se ainda de que a distribuição normal acumulada é dada pela seguinte expressão (Z é a notação tradicional para uma variável aleatória com distribuição normal padrão):

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

Assim, queremos obter um intervalo de confiança com probabilidade $(1-\alpha)$ de conter μ. Seguindo as definições colocadas anteriormente, queremos obter $z_{\alpha/2}$ e $z_{1-\alpha/2}$, que aqui equivalem a x_l e x_h, respectivamente, de modo que:

$$P\left(z_{\alpha/2} \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq z_{1-\alpha/2}\right) = 1-\alpha$$

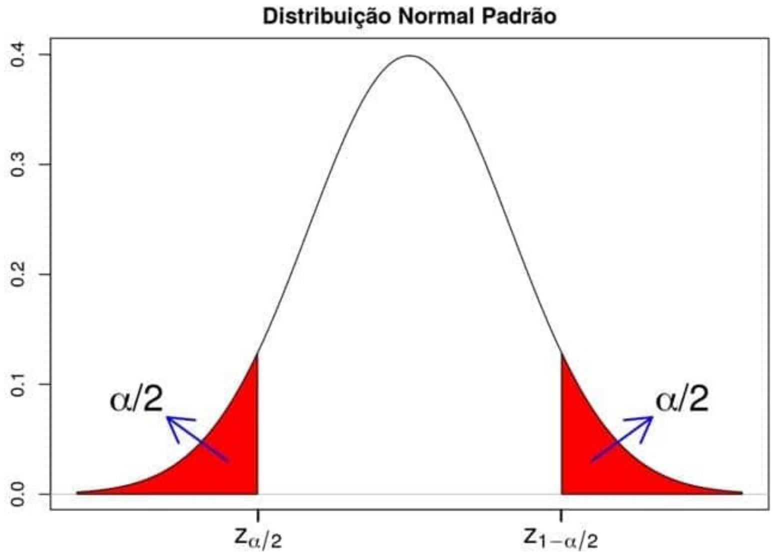
Na qual, para um dado nível de significância α, a distribuição acumulada de Z é dada por Φ. Desse modo, temos:

$$z_{\alpha/2} = \Phi^{-1}\left(\frac{\alpha}{2}\right)$$
$$z_{1-\alpha/2} = \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$$

A função inversa de Φ acima, dada por Φ^{-1}, responde à seguinte pergunta:

"Para a probabilidade acumulada Φ ser igual a $\frac{\alpha}{2}$ ou $\left(1-\frac{\alpha}{2}\right)$, qual deve ser o valor de $z_{\alpha/2}$ ou $z_{1-\alpha/2}$, respectivamente?"

Para entender melhor, veja a figura seguinte com a distribuição normal padrão:



Para um dado α, a área branca acima do intervalo entre os pontos $z_{\alpha/2}$ e $z_{1-\alpha/2}$ é igual $(1-\alpha)$ e representa a probabilidade de se extrair um elemento da população e obter um valor nesse intervalo.

A área vermelha é igual a α , e representa o contrário: a probabilidade de extrair um elemento da população estando ele fora do intervalo $[z_{\alpha/2}, z_{1-\alpha/2}]$.

Note também que, pela simetria da distribuição normal padrão em torno de sua média, temos $z_{\alpha/2} = -z_{1-\alpha/2}$.

Desse modo, vamos supor que desejamos obter o intervalo $[z_{\alpha/2}, z_{1-\alpha/2}]$ que contenha μ com probabilidade 0.95 , ou seja, que tem 95\% de chance de conter a média populacional. Temos, portanto:

$$
\begin{gathered}
P\left(z_{\alpha/2} \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq z_{1-\alpha/2}\right) = 0.95 \\
z_{\alpha/2} = \Phi^{-1}(0.025) = -1.96 \\
z_{1-\alpha/2} = \Phi^{-1}(0.975) = 1.96
\end{gathered}
$$

Logo, temos $IC_{\mu}^{95\%} = [z_{0.025}, z_{0.975}] = [-1.96, 1.96]$ em que $IC_{\theta}^{100\%(1-\alpha)}$ é o intervalo de confiança com probabilidade $(1-\alpha)$ de conter o parâmetro θ . Para a normal padrão, alguns valores de $(1-\alpha)$ são convencionalmente pedidos. São eles:

$$
\begin{gathered}
(1-\alpha) = 0.90 \Rightarrow z_{0.95} = 1.64 \\
(1-\alpha) = 0.95 \Rightarrow z_{0.975} = 1.96 \\
(1-\alpha) = 0.99 \Rightarrow z_{0.995} = 2.57
\end{gathered}
$$

Vamos fazer outro exemplo. Dessa vez, suponha que temos uma amostra aleatória X_1, X_2, X_3, X_4 a partir de uma distribuição normal $N(\mu, 1)$: ou seja, a variância é igual a 1, e a média é um parâmetro μ . Saiba também que $\Phi(2) = 0.9544$. Queremos saber a probabilidade de μ estar contido no intervalo $[\bar{X}-1, \bar{X}+1]$. Nossos limites inferior e superior, respectivamente, não estão normalizados para podermos usar a distribuição normal padrão.

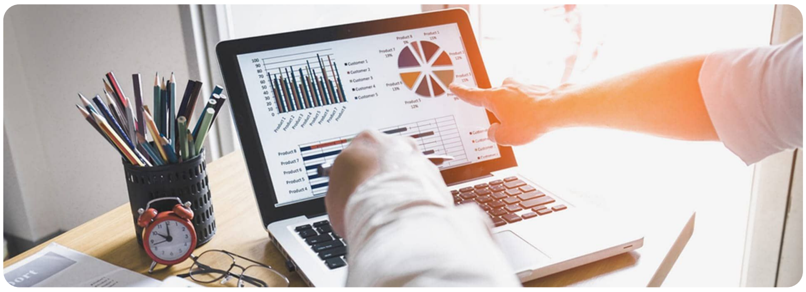
Precisamos, então, fazer isso:

$$
\begin{aligned}
P(\mu \in [\bar{X}-1, \bar{X}+1]) &= P(\bar{X}-1 \leq \mu \leq \bar{X}+1) \\
&= P(-1 \leq \mu - \bar{X} \leq 1) \\
&= P(1 \geq \bar{X} - \mu \geq -1) \\
&= P(-1 \leq \bar{X} - \mu \leq 1) \\
&= P\left(-2 \leq \frac{\bar{X}-\mu}{\sqrt{1/4}} \leq 2\right) \\
&= P(-2 \leq Z \leq 2) \\
&= 0.9544
\end{aligned}
$$

Veja a seguir o passo a passo do exemplo acima:

Da primeira para a segunda igualdade, subtraímos por $\bar{X}$ todos os valores. Da segunda para a terceira, multiplicamos por -1 . Da terceira para a quarta não realizamos nenhuma operação, apenas mudamos a apresentação da igualdade anterior. Por fim, dividimos todos os membros por $\frac{\sigma}{\sqrt{n}} = 1/\sqrt{4}$. O importante desse procedimento é que queremos normalizar nosso parâmetro de interesse ao final dele.

> Dito isso, temos uma chance de mais de 95% de cobrir o parâmetro desconhecido μ com nosso estimador intervalar. Ao abrir mão de alguma precisão em nossa estimativa, saindo de um ponto para um intervalo, o resultado foi um aumento da confiança que nossa estimação está correta.



Vamos tentar outro exemplo. Suponha que temos uma amostra aleatória $X_1, \ldots, X_{100}$ de uma distribuição desconhecida, com média amostral dada por $\bar{X} = 15.6$ e variância amostral dada por $S^2 = 8.4$. Vamos construir um intervalo de confiança de 99% para o parâmetro $\theta = E\left[X_i\right]$, ou seja, para a esperança populacional. Assim:

$$
\begin{gathered}
P(\theta \in [\bar{X} - \epsilon, \bar{X} + \epsilon]) = 0.99 \\
P(\bar{X} - \epsilon \leq \theta \leq \bar{X} + \epsilon) = 0.99 \\
P(-\epsilon \leq \bar{X} - \theta \leq +\epsilon) = 0.99 \\
P\left(-\epsilon \frac{\sqrt{100}}{\sqrt{8.4}} \leq \frac{\bar{X} - \theta}{\sqrt{8.4}/\sqrt{100}} \leq +\epsilon \frac{\sqrt{100}}{\sqrt{8.4}}\right) = 0.99 \\
P\left(-\epsilon \frac{\sqrt{100}}{\sqrt{8.4}} \leq Z \leq +\epsilon \frac{\sqrt{100}}{\sqrt{8.4}}\right) = 0.99
\end{gathered}
$$

Note que parece haver uma imprecisão da segunda para a terceira linha: o termo entre as duas desigualdades não está com a ordem trocada? Não há problema, na verdade: na segunda linha, podemos multiplicar todos os termos nas duas desigualdades (dentro da probabilidade) por -1 , e depois somamos $-\bar{X}$ a todos esses termos. Lembre-se de que ao multiplicar uma desigualdade por um valor negativo, trocamos a sua direção!

Como $z_{0.995} = 2.57$, temos:

$$
\epsilon \frac{\sqrt{100}}{\sqrt{8.4}} = 2.57 \Rightarrow \epsilon = 0.74
$$

Então, temos:

$$
\begin{aligned}
P(\bar{X} - \epsilon \leq \theta \leq \bar{X} + \epsilon) &= P(15.6 - 0.74 \leq \theta \leq 15.6 + 0.74) = 0.99 \\
&= P(14.86 \leq \theta \leq 16.34) = 0.99
\end{aligned}
$$

Logo, segue que o intervalo de confiança é dado por $IC_{E[X_i]}^{99\%} = [14.86, 16.34]$. Isso quer dizer que há 99% de chance de encontrarmos o verdadeiro valor do parâmetro $E\left[X_i\right]$ dentro desse intervalo. Conseguimos chegar a esse intervalo com os seguintes elementos: número de observações n, variância S^2 e média amostral $\bar{X}$. Esse é o jeito intuitivo de obter o intervalo.

Existe um modo mais rápido, dado pela definição da variância conhecida e desconhecida. Veja a seguir:

Intervalo de confiança para a média populacional – variância conhecida

Seja $X_1, \ldots, X_n$ uma amostra aleatória com variância conhecida dada por $\operatorname{Var}\left[X_i\right] = \sigma^2 lt; \infty$, em que n é suficientemente grande. O intervalo de confiança com a probabilidade $(1-\alpha)$ de conter o parâmetro que representa a esperança populacional $\theta = E\left[X_i\right]$ é dado por:

$$< br > IC_\theta^{100\%(1-\alpha)} = \left[\bar{X} - z_{\frac{\alpha}{2}\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}\sqrt{n}}\right] < br >$$

Intervalo de confiança para a média populacional – variância desconhecida

Seja $X_1, \ldots, X_n$ uma amostra aleatória com variância desconhecida dada por $\operatorname{Var}\left[X_i\right] = \sigma^2 lt; \infty$, em que né suficientemente grande. O intervalo de confiança com a probabilidade $(1-\alpha)$ de conter o parâmetro que representa a esperança populacional $\theta = E\left[X_i\right]$ é dado por:

$$< br > IC_\theta^{100\%(1-\alpha)} = \left[\bar{X} - z_{\frac{\alpha}{2}\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}\sqrt{n}}\right] < br >$$

Note que no resultado visto acima, utilizamos o fato de “*n* ser grande o suficiente” para utilizarmos o Teorema Central do Limite, o qual permite que, ao normalizamos o estimador amostral, possamos dizer que esse estimador segue uma distribuição normal padrão.

Pequenas amostras

Um detalhe desse exercício é que se pode notar que, pelo fato de usarmos o Teorema Central do Limite para obter intervalos de confiança, não precisamos saber muitos detalhes sobre a distribuição do parâmetro que queremos estimar. Para obter o intervalo de confiança, precisamos somente de estatísticas como $\bar{X}$ e S^2, mas, para isso, precisamos ter um número grande de observações.

O que podemos fazer se não temos isso?

Nesse caso, não podemos utilizar um resultado assintótico como o Teorema Central do Limite, então precisamos saber mais detalhes a respeito da distribuição de probabilidade da qual obtemos nossa amostra.

Como podemos obter intervalos de confiança quando não temos muitas observações?

Primeiro, vamos ver o caso especial quando uma amostra possui distribuição normal. Antes disso, precisamos discutir rapidamente duas distribuições de probabilidade intimamente ligadas à distribuição normal que serão úteis para a obtenção de intervalos de confiança para a média e variância da normal $N(\mu, \sigma^2)$.

Veja a seguir a primeira distribuição:

Distribuição qui-quadrado

Sejam $Z_1, \ldots, Z_n$ variáveis aleatórias que seguem uma distribuição normal padrão. A variável aleatória Y definida por $Y = Z_1^2 + Z_2^2 + \ldots + Z_n^2$. Segue uma distribuição qui-quadrado com n graus de liberdade, definida por $Y \sim \chi^2(n)$. Essa distribuição possui as seguintes propriedades:

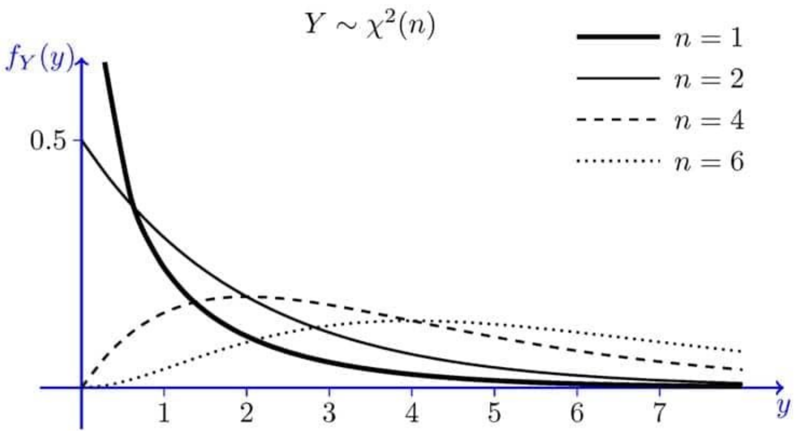
$\Rightarrow$ A distribuição qui-quadrado é um caso especial da distribuição gama. Mais especificamente, temos $Y \sim \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$. Logo, $f_Y(y) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} y^{\frac{n}{2}-1} e^{-\frac{y}{2}}, \text{ para } ygt;0.$

$\Rightarrow E[Y] = n$

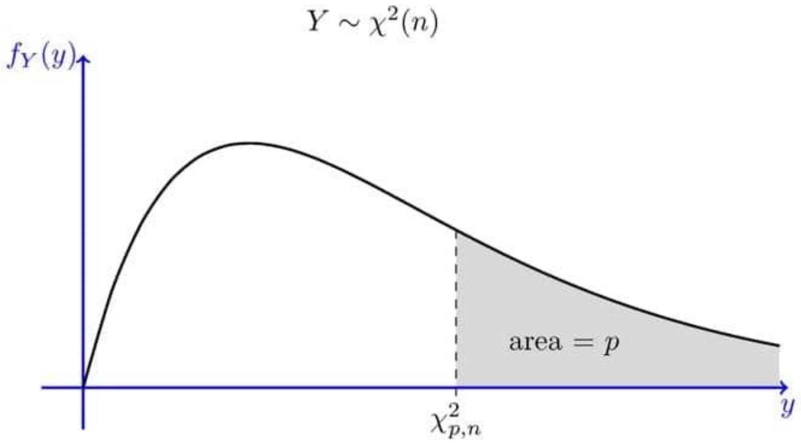
$\Rightarrow \mathrm{Var}[Y] = 2n$

$\Rightarrow$ Para qualquer $p \in [0, 1]$ e $n \in N$, definimos $\chi^2_{p,n}$ como o valor para o qual $P\left(Ygt; \chi^2_{p,n}\right) = p$, em que $Y \sim \chi^2(n)$.

Podemos observar como a função densidade de probabilidade $\chi^2(n)$ muda à medida que aumentamos n no gráfico abaixo. Veja:



Se fixamos um n e queremos saber $\chi^2_{p,n}$ para uma probabilidade p , como definido anteriormente, temos que p é a área abaixo da densidade de $\chi^2(n)$ e à direita do valor $\chi^2_{p,n}$. Para visualizar mais facilmente, veja a figura a seguir:



Mas afinal, qual será o uso da qui-quadrado na obtenção de intervalos de confiança?

O motivo ficará claro no teorema a seguir, o qual usaremos para estimar a variância de variáveis aleatórias com distribuição normal.

- Teorema: Seja $X_1, \ldots, X_n$ uma amostra independente e identicamente distribuída com distribuição $N(\mu, \sigma^2)$. Seja também S^2 a variância amostral dessa amostra. Assim, a variável Yé definida como:

$$Y = \frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2$$

Essa variável aleatória segue uma distribuição qui-quadrado com $(n-1)$ graus de liberdade, ou seja, $Y \sim \chi^2(n-1)$. Além disso, $\bar{X}$ e S^2 são variáveis aleatórias independentes.

Utilizaremos esse resultado quando a variância σ^2 for desconhecida e queremos estimá-la. Se quando procuramos estimar a variância de uma normal utilizamos a qui-quadrado, para a estimação da média utilizamos a distribuição t (também conhecida como "t de Student"). Vamos defini-la a seguir.

Distribuição t

Seja $Z \sim N(0,1)$, e $Y \sim \chi^2(n)$, em que $n \in N$. Assuma também que Z e Y são independentes. A variável aleatória T é definida como:

$$< br > T = \frac{Z}{\sqrt{Y/n}} < br >$$

A variável aleatória T tem distribuição t com n graus de liberdade tal que $T \sim T(n)$. A distribuição t tem as seguintes propriedades:
A densidade da distribuição t de Student tem um formato de sino centrada em 0 (zero), semelhante ao formato de uma normal padrão. A diferença entre as duas é que a distribuição té mais achatada que a densidade de uma normal. Isso pode ser observado na próxima figura.

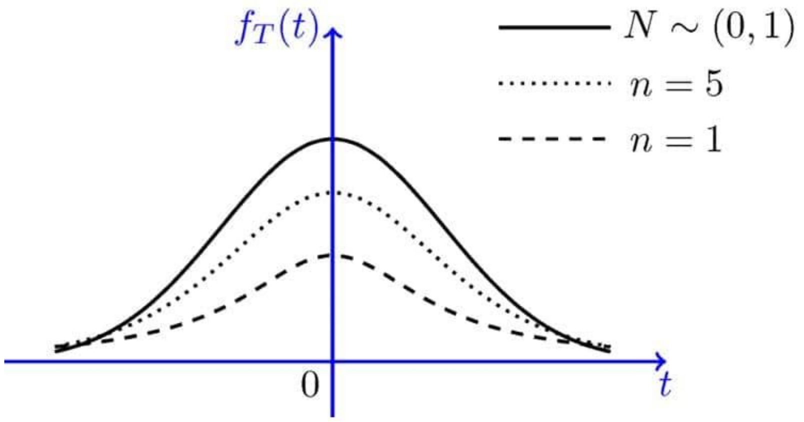
$\Rightarrow E[T] = 0$ para $ngt;0$, porém $E[T]$ é indefinida se $n = 1$.

$\Rightarrow \text{Var}[T] = \frac{n}{n-2}$ para $ngt;2$, porém $\text{Var}[T]$ é indefinida para $n \leq 2$. Repare que nesses casos de indefinição, a variância será negativa ou terá denominador igual a zero, o que é absurdo.

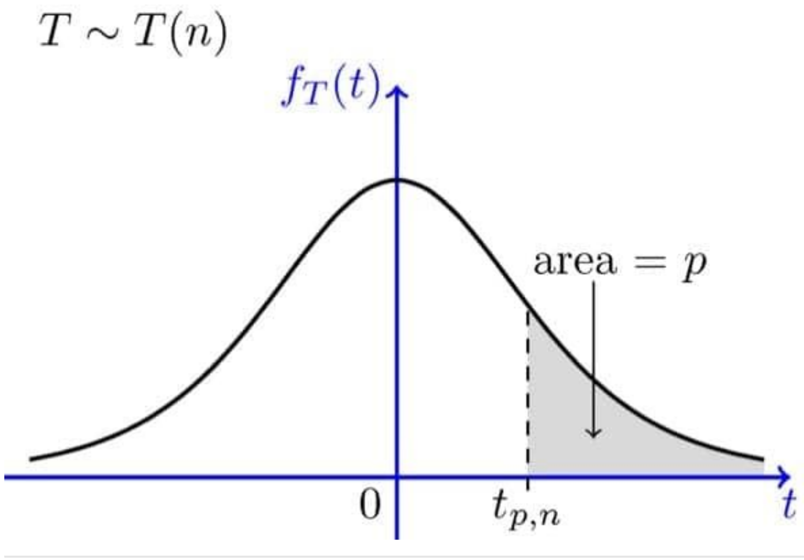
$\Rightarrow$ À medida que n aumenta, a densidade da t de Student se aproxima da densidade da normal padrão $N(0,1)$.

$\Rightarrow$ Para qualquer $p \in [0,1]$ e $n \in N$, definimos $t_{p,n}$ como o valor para o qual $P(T > t_{p,n}) = p$. Como a distribuição t tem densidade simétrica, temos $t_{1-p,n} = -t_{p,n}$.

A ilustração a seguir mostra a densidade de uma distribuição *t* de Student para alguns valores de *n* e a compara com a densidade de uma distribuição normal padrão. Como podemos ver, a densidade da distribuição *t* é mais **achatada** ou **espalhada** que a da normal padrão.



Se fixamos um n e queremos saber $t_{p,n}$ para uma probabilidade p , como foi definido, temos que p é a área abaixo da densidade de $T(n)$ e à direita do valor $t_{p,n}$. Para visualizar mais facilmente, veja a próxima figura.



Mas, para que precisamos utilizar a t de Student na obtenção de intervalos de confiança?

O teorema seguinte mostra que, quando não podemos utilizar o teorema central do limite, podemos recorrer a ela para a obtenção de estimadores intervalares para a média de uma normal $N\left(\mu, \sigma^2\right)$.

Teorema: Seja $X_1, \ldots, X_n$ uma amostra independente e identicamente distribuída com distribuição $N\left(\mu, \sigma^2\right)$. Seja também S^2 a variância amostral dessa amostra. Assim, a variável T é definida como:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

Segue uma distribuição t de Student com $(n-1)$ graus de liberdade, ou seja, $T \sim T(n-1)$.

Vejamos como podemos utilizar esses resultados a respeito da t de Student e qui-quadrado para obter intervalos de confiança para a média μ e variância σ^2 de uma normal, respectivamente.

Primeiro, vamos mostrar como obter a média. Aqui, assumimos que $X_1, \ldots, X_n$ é uma amostra aleatória retirada de uma normal $N\left(\mu, \sigma^2\right)$. Nosso objetivo é encontrar um estimador intervalar para μ . Como dissemos, não precisamos mais que n seja grande, ou seja, né qualquer número inteiro positivo. Existem dois cenários possíveis, dependendo de σ^2 ser conhecido ou não. Se sabemos σ^2 , o procedimento para encontrar μ é igual ao utilizado para um n grande. Mais especificamente, temos o seguinte cenário:

Hipóteses

A variável aleatória $X_1, \ldots, X_n$ é obtida a partir de uma distribuição $N\left(\mu, \sigma^2\right)$, em que $Var\left[X_i\right] = \sigma^2$ é conhecida.

Parâmetro a ser estimado

$\mu = E\left[X_i\right]$

Intervalo de confiança

$\left[\bar{X} - z_{\frac{\alpha}{2}\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}}^{\frac{S}{n}}\right]$ é um intervalo de confiança com $(1-\alpha)100\%$ de chance de conter μ.

O caso mais interessante é quando temos uma situação semelhante à anterior, porém não sabemos a variância σ^2. Pelo último teorema visto, vemos que a variável aleatória T é definida por:

$$T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$$

Essa variável segue uma distribuição t com $(n-1)$ graus de liberdade, ou seja, $T \sim T(n-1)$.

Note que, agora, como não sabemos a variância, utilizamos o estimador S^2. Vamos obter o intervalo de confiança com probabilidade $(1-\alpha)$ de conter μ, passo a passo, como fizemos no caso assintótico. Utilizando a definição de $t_{p,n}$, temos:

$$\begin{gathered}
P\left(-t_{\frac{\alpha}{2},n-1} \leq T \leq t_{\frac{\alpha}{2},n-1}\right) = 1-\alpha \\
P\left(-t_{\frac{\alpha}{2},n-1} \leq \frac{\bar{X}-\mu}{S/\sqrt{n}} \leq t_{\frac{\alpha}{2},n-1}\right) = 1-\alpha \\
P\left(\bar{X} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}\right) = 1-\alpha
\end{gathered}$$

Vemos então o seguinte cenário:

Hipóteses

A variável aleatória $X_1, \ldots, X_n$ é obtida a partir de uma distribuição $N\left(\mu, \sigma^2\right)$, em que $Var\left[X_i\right] = \sigma^2$ é desconhecida.

Parâmetro a ser estimado

$\mu = E\left[X_i\right]$

Intervalo de confiança

$\left[\bar{X}-t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}, \bar{X}+t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}\right]$ é um intervalo de confiança com $(1-\alpha)100\%$ de chance de conter μ .

Por fim, suponha agora que queiramos obter um estimador intervalar para a variância σ^2 a partir de uma amostra aleatória $X_1,\ldots,X_n$ obtida de uma normal $N\left(\mu,\sigma^2\right)$. Suponha que μ também é desconhecida. Novamente, temos uma amostra que não é necessariamente grande, isto é, n é qualquer número inteiro positivo. Pelo resultado obtido após abordarmos a qui-quadrado anteriormente, temos:

$$Q=\frac{(n-1)S^2}{\sigma^2}=\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$$

Segue uma distribuição qui-quadrado com $n-1$ graus de liberdade, isto é, $Q\sim\chi^2(n-1)$. Utilizando a definição de $\chi^2_{p,n}$, um intervalo de confiança com $(1-\alpha)100\%$ de chance de conter σ^2 pode ser obtido da seguinte maneira:

$$\begin{gathered}
P\left(\chi^2_{1-\frac{\alpha}{2},n-1}\leq Q\leq\chi^2_{\frac{\alpha}{2},n-1}\right)=1-\alpha \\
P\left(\chi^2_{1-\frac{\alpha}{2},n-1}\leq\frac{(n-1)S^2}{\sigma^2}\leq\chi^2_{\frac{\alpha}{2},n-1}\right)=1-\alpha \\
P\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2},n-1}}\leq\sigma^2\leq\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}\right)=1-\alpha
\end{gathered}$$

Desse modo, concluímos que $\left[\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2},n-1}},\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2},n-1}}\right]$ é um intervalo de confiança com $(1-\alpha)100\%$ de chance para σ^2 .

Note que, no cálculo acima, da segunda para a terceira igualdade, o que fazemos é elevar todos os termos dentro da probabilidade a -1 (e, portanto, fazendo com quem era maior, que agora seja menor e vice-versa) e depois multiplicar esses termos por $(n-1)S^2$. Desse modo, temos o seguinte cenário:

Hipóteses

A variável aleatória $X_1,\ldots,X_n$ é obtida a partir de uma distribuição $N\left(\mu,\sigma^2\right)$, em que $E\left[X_i\right]=\mu$ e $Var\left[X_i\right]=\sigma^2$ são desconhecidos.

Parâmetro a ser estimado

$\sigma^2=\operatorname{Var}\left[X_i\right]$

Intervalo de confiança

$\left[\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}, \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2},n-1}}\right]$ é um intervalo de confiança com $(1-\alpha)100\%$ de chance de conter σ^2 .

Para complementar seus estudos, assista à explicação de como usar a tabela da normal para resolver uma ilustração numérica dos conceitos desenvolvidos no tópico **A construção de intervalos de confiança.**

Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Verificando o aprendizado

Questão 1

Para estimar a proporção de eleitores que votarão no Candidato A em uma eleição, uma agência de pesquisa de opinião chamada YBOPY escolhe uma amostra aleatória de tamanho n da população. A amostragem é feita com reposição. Seja θ a proporção de eleitores que planejam votar no Candidato A entre todos os eleitores. Quão grande n precisa ser para que obtenhamos um intervalo de confiança de 90\% com margem de erro de 3\%? Em outras palavras, quão grande n precisa ser para que valha a seguinte desigualdade:

$$< br > P(\bar{X} - 0.03 \leq \theta \leq \bar{X} + 0.03) \geq 0.90 < br >$$

Em que $\bar{X} = \frac{1}{2}$ é a proporção de pessoas em nossa amostra que diz que votará no Candidato A . Assinale a alternativa que corresponde à resposta correta.

A

357

B

487

C

747

D

574

A alternativa C está correta.

Note que, pelo enunciado, cada indivíduo vota no Candidato A com probabilidade p ou não vota nele com probabilidade $(1-p)$. Estamos falando, nesse caso, de uma distribuição de probabilidade $\text{Bernoulli}(p)$, com $E[X_i]=p$ e $\text{Var}[X_i]=p(1-p)$. Assim, queremos estimar a média da distribuição dada por p.

Em um exercício desse tipo, no qual é perguntado o tamanho mínimo da amostra, estamos em uma situação em que podemos explorar o Teorema Central do Limite.

Desse modo, podemos nos valer do resultado do qual supomos ter uma amostra grande. Assim:

$$
\begin{aligned}
\left[\bar{X}-z_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}},\bar{X}+z_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right] &= (1-\alpha) \\
\left[\bar{X}-z_{0.05}\frac{\sqrt{X(1-X)}}{\sqrt{n}},\bar{X}+z_{0.05}\frac{\sqrt{X(1-X)}}{\sqrt{n}}\right] &= 0.90 \\
\left[\frac{1}{2}-1.64\frac{\sqrt{\frac{1}{2}\left(1-\frac{1}{2}\right)}}{\sqrt{n}},\bar{X}+1.64\frac{\sqrt{\frac{1}{2}\left(1-\frac{1}{2}\right)}}{\sqrt{n}}\right] &= 0.90 \\
\left[\frac{1}{2}-1.64\frac{1}{2\sqrt{n}},\frac{1}{2}+1.64\frac{1}{2\sqrt{n}}\right] &= 0.90
\end{aligned}
$$

Como a margem de erro é de 3%, temos:

$$
\begin{aligned}
1.64\frac{1}{2\sqrt{n}} &= 0.03 \\
1.64\frac{1}{0.06} &= \sqrt{n} \\
n &= \left(\frac{1.64}{0.06}\right)^2 \\
n &= 747
\end{aligned}
$$

Logo, devemos ter no mínimo 747 indivíduos em nossa amostra para chegar ao intervalo de confiança de 90% com margem de erro de 3%.

Questão 2

Uma amostra aleatória $X_1,\dots X_{16}$ é obtida de uma distribuição normal com média desconhecida $\mu=E[X_i]$ e variância desconhecida $\text{Var}[X_i]=\sigma^2$. Para essa amostra, $\bar{X}=16.7$ e $S^2=7.5$. Ache os intervalos de confiança para $\alpha=0.05$ para μ e σ^2. Para a resolução desse exercício, saiba que $t_{0.025,15}=2.13, \chi^2_{0.025,15}=27.49$ e $\chi^2_{0.975,15}=6.26$. Assinale a alternativa correta:

A

[1.09, 6.97]

B

[12.09, 17.97]

C

[4.09, 11.97]

D

[4.09, 17.97]

A alternativa D está correta.

Para obter o intervalo de confiança para a média populacional μ, uma vez que não sabemos a média, tampouco a variância, e visto que a amostra é pequena, temos que utilizar a distribuição t. O intervalo, nesse caso, é:

$$\left[\bar{X} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}, \bar{X} + t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}\right]$$

Uma vez que $n = 16.7, \alpha = 0.05$ e $t_{0.025,15} = 2.13$ temos:

$$\left[16.7 - 2.13\frac{\sqrt{7.5}}{\sqrt{4}}, 16.7 + 2.13\frac{\sqrt{7.5}}{\sqrt{4}}\right] = [15.24, 18.16]$$

Logo, $[15.24, 18.16]$ é um intervalo de confiança de 95% para μ. Vejamos agora qual o intervalo de confiança de 95% para σ^2. Uma vez que $\chi^2_{0.025,15} = 27.49$ e $\chi^2_{0.975,15} = 6.26$. O intervalo é:

$$\left[\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}, \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2},n-1}}\right] = \left[\frac{15\times7.5}{27.49}, \frac{15\times7.5}{6.26}\right]$$

Assim, [4.09, 17.97] é um intervalo de confiança de 95% para μ.

Introdução a Testes de Hipótese

Pesquisadores testam suas hipóteses o tempo todo:

- Esse remédio reduz uma doença?
- Colocar mais moeda na economia aumenta a inflação?
- Esse novo material resiste a temperaturas maiores que 500°C?

Todas essas perguntas podem ser formuladas como testes de hipóteses.

Vamos ilustrar o procedimento de testagem com um exemplo mais trivial:

Digamos que você seja um pesquisador, acorda certo dia, coloca seu jaleco branco, pega seu café e, inspirado por este tema, deseja saber se uma das moedas em sua carteira é honesta, ou seja, se a probabilidade de cair cara ou coroa é igual, e tem valor ½ . Para verificar, você joga a moeda 10 vezes.

Para quais dos possíveis resultados abaixo você consideraria que a moeda não é honesta?

A. 5 caras e 5 coroas

B. 4 caras e 6 coroas

C. 6 caras e 4 coroas

D. 2 caras e 8 coroas

E. 1 cara e 9 coroas

F. 0 caras e 10 coroas

Apenas olhando, podemos suspeitar das opções D, E e F, pois nelas a moeda parece mostrar coroa muito mais vezes. Não queremos, porém, apenas suspeitar, mas, sim, testar a hipótese de a moeda ser honesta ou não. Existe um procedimento específico para isso na Estatística e que será visto neste módulo.

Antes, vamos formalizar os conceitos:

Hipótese

Uma hipótese é uma declaração sobre um parâmetro da população.

> **Essa definição é genérica, mas o aspecto importante é que uma hipótese é uma declaração sobre um parâmetro de uma população, não de uma amostra.**

O objetivo de um teste de hipótese é decidir, com base em uma amostra de uma população, qual de duas hipóteses complementares é verdadeira.

No exemplo acima, como temos n lançamentos de uma moeda, e queremos testar se ela é de fato honesta, ou seja, que obtemos cara com probabilidade $1/2$ ou coroa com probabilidade ½, estamos testando o parâmetro p de uma distribuição $\text{Binomial}(x, n = 10, p)$, em que n é o número de total de tentativas, x é o número de sucessos (no nosso caso, o número de caras) e pé a probabilidade de sucesso.

Hipóteses complementares

As duas hipóteses complementares em um problema envolvendo um teste de hipóteses são chamadas de hipótese nula e hipótese alternativa, denotadas H_0 e H_1 , respectivamente.

A hipótese nula H_0 se refere à hipótese que queremos testar sobre o parâmetro populacional. No exemplo anterior, queremos testar se a moeda é honesta. Em termos estatísticos, isso significa que queremos saber se $p = 0,5$. Ou seja, nossa hipótese nula pode ser expressa da seguinte maneira: $H_0 : p = 0,5$.

A hipótese que queremos testar é:

> A moeda não **rouba** para uma das faces.

Essa situação comum, na qual H_0 declara que algum fenômeno não ocorre, é o motivo de chamarmos essa hipótese de hipótese nula. A definição deixa claro que a hipótese nula e a alternativa são complementares. Logo, a hipótese alternativa é $H_1 : p \neq 0,5$.

Em um problema envolvendo um teste de hipótese, depois de observar a amostra, o pesquisador deve decidir se rejeita ou não H_0 e decidir se H_1 é verdadeira. Vamos elucidar o significado de rejeitar, não rejeitar e aceitar abaixo, após definirmos formalmente o procedimento de teste de hipótese.

Teste de Hipótese

Procedimento para testar uma hipótese, ou um Teste de Hipótese, é uma regra que especifica:

1. Para quais valores amostrais decidimos não rejeitar H_0 como verdadeira.
2. Para quais valores amostrais decidimos rejeitar H_0 , e H_1 é aceita como verdadeira.

O subconjunto do espaço amostral para o qual H_0 será rejeitada é chamado de região de rejeição ou região crítica. O complemento da região de rejeição é chamado de região de não rejeição.

Note que não utilizamos o termo aceitar a hipótese nula, mas existe um motivo para isso, e entendê-lo é importante para interpretar resultados estatísticos de maneira correta.

Vamos deixar mais claro com exemplos:

Neste primeiro exemplo, suponha que desejamos medir o efeito de matricular uma criança em uma creche sobre seu salário quando adulta.

Será que investimento em educação na primeira infância tem impacto em longo prazo, proporcionando maior capacidade de geração de renda no futuro?

Considere que esse efeito é capturado pelo parâmetro θ . Se testamos a hipótese $H_0 : \theta = 0$, estamos testando a hipótese nula de não ter efeito algum de creches sobre salários futuros, ou seja, pessoas que foram para creches e pessoas que não foram para creches têm aproximadamente o mesmo salário. Se não rejeitamos essa hipótese, estamos apenas dizendo que, para nossa amostra, não conseguimos distinguir o efeito das creches sobre salários de zero. Isso não significa aceitar que ele é positivo ou negativo.

Agora veja esse outro exemplo: Em um tribunal, é razoável presumir que o réu é inocente até que se prove o contrário. É a chamada presunção de inocência.

Um promotor tem que se esforçar bastante até conseguir provas suficientes para convencer um juiz ou júri sobre a culpa do réu. Caso ele consiga algumas provas, porém não o suficiente para fazer uma acusação que convença o júri, isso não prova de maneira alguma que o réu era inocente. Apenas que o promotor não pode provar que ele era culpado.

Talvez o promotor tenha sido preguiçoso na sua investigação, talvez o réu tenha escondido seu rastro na cena do crime. Não sabemos. O que se pode dizer é que, para aquele corpo de evidência apresentado, ele não pode ser considerado culpado.

Vejamos agora um exemplo clássico do mundo da lógica e da filosofia. Considere que somos pesquisadores e queremos efetuar o seguinte teste de hipótese:

H_0 : Todos os cisnes da espécie cygnus olor são brancos
H_1 : Existe pelo menos um cisne da espécie cygnus olor que não é branco

Não há a possibilidade de observar os membros da população de cisnes dessa espécie, do mundo inteiro. Para testar a validade dessa hipótese, assuma que foi desenhado um experimento no qual observam-se esses cisnes em *n* locais escolhidos por sorteio.

Considere que $C_1, C_2, \ldots, C_n$ são os resultados das observações dos $ngt;0$ locais. Se foi observado algum cisne que não seja branco no local i , teremos $C_i = 1$. Se todos os cisnes observados no local i forem brancos, temos $C_i = 0$.

Existem dois resultados possíveis para esse experimento, dada a evidência que recolhemos:

Evidência 1

$\sum_{i=1}^{n} C_i gt; 0$, ou seja, foi observado pelo menos 1 cisne que não é branco nas observações realizadas. A conclusão possível é: H_0 foi rejeitada (a evidência é conclusiva e independe do valor de observações n). H_1 é aceita.

Evidência 2

$\sum_{i=1}^{n} C_i = 0$, ou seja, todas as observações realizadas constataram somente cisnes brancos. A conclusão possível é: H_0 não foi rejeitada (a evidência não foi conclusiva). H_1 também não foi aceita.

Note que observar apenas 1 cisne que não é branco já permite dizermos que H_0 é falsa, e devemos rejeitá-la; logo, H_1 é verdadeira, e podemos aceitá-la. O contrário, quando apenas observamos cisnes brancos, não nos permite dizer muito. Pode ser que apenas não tenhamos feito observações em locais suficientes (pode ser que todos eles se concentrem em um lago que não foi sorteado). Pode ser que, caso tivéssemos observado os mesmos locais em outra época do ano, pudéssemos observar cisnes negros. Do mesmo modo, não conseguimos dizer se existe pelo menos um cisne negro, como diz a H_1 , pois apenas observamos cisnes brancos em nosso experimento. Essa situação deixa claro, mais uma vez, que não rejeitar é diferente de aceitar. Apenas podemos rejeitar ou não a hipótese nula, nunca a aceitar.

Com os três exemplos anteriores, fica clara a distinção entre não rejeitar e aceitar uma hipótese.

Hipóteses alternativas podem ser unilaterais (ou unicaudais), ou seja, serem da forma $H_1 : \theta gt; \theta_0$ ou $H_1 : \theta lt; \theta_0$, ou também podem ser bilaterais (ou bicaudais), da forma $H_1 : \theta \neq \theta_0$.

Hipóteses alternativas unilaterais costumam surgir quando o parâmetro que queremos testar assume valores não negativos ou não positivos, isto é, $H_0 : \theta \geq \theta_0$ ou $H_0 : \theta \leq \theta_0$, enquanto hipóteses alternativas bilaterais costumam surgir nos casos exemplificados, quando testamos se a hipótese nula é igual a zero.

> Testes bicaudais são mais comuns em aplicações do que testes unicaudais, porém testes unicaudais são mais fáceis de serem analisados.

Erros em Testes de Hipótese

Uma decisão a respeito de um teste pode ser correta ou incorreta. Uma decisão incorreta é, logicamente, um erro. Porém, no contexto de Testes de Hipótese existem dois tipos de erros, chamados, não muito criativamente, de Erro Tipo 1 e Erro Tipo 2, definidos a seguir:

Tipos de erro em Testes de Hipótese

Existem dois tipos de erro que podemos cometer em Testes de Hipótese. O Erro Tipo 1 (ET1) ocorre se rejeitamos a hipótese nula H_0 quando H_0 é verdadeira. O Erro Tipo 2 (ET2) ocorre se não rejeitamos a hipótese nula H_0 quando H_0 é falsa.

Essa definição fica mais fácil de ser compreendida observando a tabela a seguir:

		Realidade	
		H_0 é verdadeira	H_0 é falsa
Testes	Não rejeitar H_0	Decisão correta	Erro Tipo II (β)
	Rejeitar H_0	Erro Tipo I (α)	Decisão correta

No exemplo dos cisnes brancos, caso não rejeitássemos a hipótese nula de todos os cisnes serem brancos, mas, eventualmente, tivessem achado um cisne negro em um lago escondido da Sibéria, teríamos cometido o Erro Tipo 2, pois não rejeitamos H_0 quando H_0 era falsa.

Suponha agora que tivéssemos rejeitado H_0, pois julgamos ter visto pelo menos um cisne que não era branco. Anos depois, porém, um cientista prova que, por um motivo genético, todos os cisnes devem ser brancos e, logo, aquelas aves de outras cores eram apenas patos. Nesse caso, teríamos cometido Erro Tipo 1, pois rejeitamos H_0 quando H_0 era verdadeira.

Obviamente, não temos como saber se cometeremos um erro ou outro no momento que realizamos o teste. Como saberíamos do cisne negro escondido na Sibéria ou da descoberta que só seria feita anos depois, afinal?

No entanto, não ter certeza de algo não significa que não podemos quantificar. Iremos, assim, atribuir probabilidades para os nossos erros, de modo que seja possível minimizar a chance de incorrer em um deles. Como pesquisadores, nos perguntamos: Dada a nossa hipótese e a nossa amostra, qual a probabilidade de cometer Erro Tipo 1 ou Erro Tipo 2 quando testamos nossas hipóteses?

Vamos definir a seguir três conceitos muito importantes:

Probabilidade de ET1

A probabilidade de cometer um Erro Tipo 1 em nossa tomada de decisão em um teste de hipótese é dada por:

$$< br > \alpha = P(\text{ Erro Tipo 1}) = P\text{ (rejeitar } H_0 \mid H_0 \text{ é verdadeira) } < br >$$

A probabilidade de cometer Erro Tipo I dada por α é chamada de nível de significância ou tamanho do teste.

Probabilidade de ET2

A probabilidade de cometer um Erro Tipo 2 em nossa tomada de decisão em um teste de hipótese é dada por:

$$< br > \beta = P(\text{ Erro Tipo 2}) = P\text{ (não rejeitar } H_0 \mid H_0 \text{ é falsa) } < br >$$

Poder do teste

O poder de um teste de hipóteses é obtido a partir da probabilidade de cometermos um Erro Tipo 2, e é dado por:

$$< br > 1 - \beta = 1 - \text{ (não rejeitar } H_0 \mid H_0 \text{ é falsa) } = P\text{ (rejeitar } H_0 \mid H_0 \text{ é falsa) } < br >$$

A probabilidade α é definida antes do teste de hipótese. Imagine uma prova de salto em altura. Quanto maior $1 - \alpha$, maior a altura da barra para o nosso teste. Isto é, quanto menor o nível de significância que impomos ao nosso teste, mais forte deverá ser nossa evidência amostral para dizermos com confiança "o nosso parâmetro θ é diferente de θ_0 ".

Por exemplo, para dizermos que um parâmetro θ é estatisticamente significante (i.e., que ele é estatisticamente diferente do parâmetro real θ_0) para $\alpha = 0,01$, deveremos ter uma evidência mais forte do que para $\alpha = 0,05$. Quando rejeitamos H_0 com $\alpha = 0,01$, estamos $1 - \alpha = 99\%$ confiantes que não incorremos em Erro Tipo I. Quando rejeitamos $H_0 \, \text{com} \, \alpha = 0,05$, estamos $1 - \alpha = 95\%$ confiantes que não incorremos em Erro Tipo I. E assim por diante.

> Pode parecer complicado, mas basta lembrar que a "altura da barra", ou seja, quão confiantes queremos estar de nosso resultado caso rejeitemos H_0, é dada por $1 - \alpha$.

Enquanto α nos permite colocar a "altura da barra" em certo ponto para podermos confiar melhor em nossos resultados, o poder do teste $1 - \beta$ nos diz a nossa capacidade como "atletas" de pular essa barra. Saindo da metáfora, dizemos poder nos dar a probabilidade de rejeitar H_0 quando H_0 é de fato falsa.

Estatística de teste

Mas o que utilizamos para realizar testes? Sabemos no que consiste um teste, o que é uma hipótese e como iniciar o procedimento? Será que no exemplo da moeda no início deste módulo, julgamos na base do “olhômetro”? Existem, na verdade, diversos modos, porém todos eles são do mesmo tipo: estatísticas de teste.

Veja a definição abaixo:

Estatística de teste

Uma estatística de teste é uma variável aleatória $W\left(X_1, \ldots, X_n\right)$ que é uma função da amostra, ou seja, que depende das observações. Trata-se de uma estatística na qual podemos efetuar um teste de hipótese.

Essa definição parece vaga e deixa a possibilidade para centenas de estatísticas serem estatísticas de teste. Vamos focar naquelas utilizadas para as situações mais comuns em Testes de hipótese: comparações de média ou de variâncias.

Primeiramente devemos executar Testes de Hipótese a respeito da média. Uma boa estatística a ser considerada nesse caso é:

$W\left(X_1, \ldots, X_n\right)=\frac{\bar{X}-\mu_0}{\sigma / \sqrt{n}}$, se sabemos que $\operatorname{Var}\left[X_i\right]=\sigma^2$ $W\left(X_1, \ldots, X_n\right)=\frac{\bar{X}-\mu_0}{S / \sqrt{n}}$, se não sabemos $\operatorname{Var}\left[X_i\right]$

Digamos que temos uma amostra aleatória $X_1, \ldots, X_n$ obtida de uma normal $N\left(\mu, \sigma^2\right)$, em que μ é desconhecida, mas σ^2 é conhecida. Queremos construir um teste de nível α para poder avaliar as seguintes hipóteses:

$$
\begin{aligned}
& H_0: \mu=\mu_0 \\
& H_1: \mu \neq \mu_0
\end{aligned}
$$

Note que, se assumimos H_0 como verdadeira, temos que $W\left(X_1, \ldots, X_n\right)=\frac{\bar{X}-\mu_0}{\sigma / \sqrt{n}}$ segue uma normal $N(0,1)$. Escolheremos um valor c que nos dá os intervalos para os quais rejeitaremos ou não H_0. Se $|W| \leq c$, não rejeitamos H_0, se $|W|>c$, rejeitamos H_0. Para achar esse valor c, temos:

$$
P\left(|W|>c \mid H_0 \text { verdadeira }\right)=\alpha
$$

Como a densidade da normal padrão é simétrica em torno de 0, temos:

$$
P\left(|W|>c \mid H_0 \text { verdadeira }\right)=2 P\left(W>c \mid H_0 \text { verdadeira }\right)
$$

Logo, $P(W gt; c \mid H_0$ verdadeira $) = \frac{\alpha}{2}$. Segue que:

$$c = z_{\frac{\alpha}{2}}$$

Assim, não rejeitamos H_0 se:

$< br > \left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| \leq z_{\frac{\alpha}{2}} < br >$

Para essa discussão ficar mais clara, vamos lembrar do conteúdo de intervalo de confiança. A expressão acima, que define a região de não rejeição, nada mais é que um intervalo de confiança definido por:

$$\mu_0 \in \left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

Essa expressão deve ser familiar se você leu com atenção o módulo anterior. É o intervalo de confiança com $(1-\alpha)100\%$ de probabilidade de conter o parâmetro μ_0 .

Desse modo, se o intervalo de confiança para certo parâmetro não incluir o valor dele sob a validade da hipótese nula, a hipótese nula será rejeitada. Se esse intervalo incluir o valor, ela não será rejeitada.

Se fôssemos testar o mesmo conjunto de hipóteses, mas caso não soubéssemos μ nem σ^2 , teríamos que utilizar outra estatística. Nesse caso, teríamos:

$$W(X_1, \ldots, X_n) = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$$

Como visto no módulo sobre intervalo de confiança, essa variável aleatória segue uma distribuição t de Student com $n-1$ graus de liberdade, ou seja, $W \sim T(n-1)$. Assim, repetimos a análise do exemplo anterior. As diferenças estão no fato de que precisamos substituir σ por $\mathrm{Se}\, z_{\frac{\alpha}{2}}$ por $t_{\frac{\alpha}{2}, n-1}$. Assim, não rejeitamos H_0 se:

$$\left|\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\right| \leq t_{\frac{\alpha}{2}, n-1}$$

E rejeitamos caso contrário. Podemos resumir esse conteúdo na seguinte tabela:

Teste	Caso	Estatística de teste	Região de não rejeição
$amp; H_0 : \mu = \mu_0$ $amp; H_1 : \mu \neq \mu_0$	$X_i \sim N(\mu, \sigma^2), \sigma$ conhecido	$W = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$\lvert W \rvert \leq z_{\frac{\alpha}{2}}$

Teste	Caso	Estatística de teste	Região de não rejeição
$amp; H_0 : \mu = \mu_0$ $amp; H_1 : \mu \neq \mu_0$	n grande, X_i não segue normal	$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$\lvert W \rvert \leq z_{\frac{\alpha}{2}}$
$amp; H_0 : \mu = \mu_0$ $amp; H_1 : \mu \neq \mu_0$	$X_i \sim N\left(\mu, \sigma^2\right), \sigma$ desconhecido	$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$\lvert W \rvert \leq t_{\frac{\alpha}{2}, n-1}$

Os resultados acima são para testes bilaterais (ou bicaudais). Caso realizemos testes unilaterais (ou unicaudais), podemos apenas repetir o procedimento anterior, removendo o módulo da estatística de teste e utilizando a desigualdade relevante. Assim, temos:

Teste	Caso	Estatística de teste	Região de não rejeição
$< br > \; < br > amp; H_0 : \mu \leq \mu_0 \; < br > amp; H_1 : \mu \quad gt; \mu_0 < br > \; < br >$	$X_i \sim N\left(\mu, \sigma^2\right), \sigma$ conhecido	$W = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$W \geq -z_{\alpha}$
$< br > \; < br > amp; H_0 : \mu \leq \mu_0 \; < br > amp; H_1 : \mu \quad gt; \mu_0 < br > \; < br >$	n *grande,* X_i *não segue normal*	$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$W \geq -z_{\alpha}$
$< br > \; < br > amp; H_0 : \mu \leq \mu_0 \; < br > amp; H_1 : \mu \quad gt; \mu_0 < br > \; < br >$	$X_i \sim N\left(\mu, \sigma^2\right), \sigma$ desconhecido	$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$W \leq t_{\alpha, n-1}$

Note que, como estamos olhando para apenas um lado da distribuição, não precisamos distribuir α por 2 nas estatísticas. Para o teste de hipótese unilateral em que $H_0 : \mu \geq \mu_0$ e $\mu lt; \mu_0$, temos:

Teste	Caso	Estatística de teste	Região de não rejeição
$< br > \; < br > amp; H_0 : \mu \leq \mu_0 \; < br > amp; H_1 : \mu \quad gt; \mu_0 < br > \; < br >$	$X_i \sim N\left(\mu, \sigma^2\right), \sigma$ conhecido	$W = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$W \leq z_{\alpha}$
$< br > \; < br > amp; H_0 : \mu \leq \mu_0 \; < br > amp; H_1 : \mu \quad gt; \mu_0 < br > \; < br >$	n *grande,* X_i *não segue normal*	$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$W \leq z_{\alpha}$
$< br > \; < br > amp; H_0 : \mu \leq \mu_0 \; < br > amp; H_1 : \mu \quad gt; \mu_0 < br > \; < br >$	$X_i \sim N\left(\mu, \sigma^2\right), \sigma$ desconhecido	$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$W \leq t_{\alpha, n-1}$

Neste módulo focamos no caso de comparação de média. Há também outros testes de hipótese para proporções em uma população e igualdade de variâncias, por exemplo. Deixaremos esses casos indicados para a seção “Explore +”

P-Valor

Por fim, iremos abordar rapidamente o conceito de p-valor.

Na discussão anterior apenas rejeitamos ou não rejeitamos uma hipótese nula. Porém, é possível dizer mais informação a respeito de nosso teste de hipótese utilizando o chamado p-valor. Ele nos dá a resposta para perguntas do tipo: Se eu consegui rejeitar H_0 nesse teste para $\alpha = 0.05$, eu conseguiria rejeitar também para $\alpha = 0.01$?

> Em outras palavras, **o p-valor nos dá o menor nível de significância α para o qual conseguimos rejeitar a hipótese nula**.

Intuitivamente, se temos um p-valor muito baixo, isso quer dizer que os dados observados são muito difíceis de terem ocorrido sobre H_0 verdadeira, então podemos estar mais confiantes em rejeitar H_0. Formalmente, temos:

P-Valor

Considere o teste de hipótese entre H_0 e H_1. Seja Wa estatística de teste e $w_1 \circ$ valor observado de W. Assumindo que $\circ H_0$ é verdadeira, o p-valor é dado por $P($ Erro tipo $I \mid c = w_1)$, em que cé o valor crítico acima do qual rejeitamos a hipótese nula.

Antes de finalizar seus estudos, assista ao vídeo que explica como usar a tabela da normal para resolver uma ilustração numérica dos conceitos desenvolvidos no tópico **Estatística de teste.**

Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Verificando o aprendizado

Questão 1

Sejam $X_1 = 3.58, X_2 = 10.03, X_3 = 4.77, X_4 = 14.66$ uma amostra aleatória retirada de uma normal $N(\mu, \sigma^2)$ em que μ e σ^2 são desconhecidas. Saiba também que $z_{0.05} = 1.64$ et $_{0.05,3} = 2.35$. Queremos decidir entre $H_0 : \mu \geq 10$ ou $H_1 : \mu lt; 10$. Assumindo nível de significância de 0.05 , assinale a alternativa correta:

A

$< br > W gt; -t_{\alpha,n-1}$, por isso não rejeitamos $H_0 < br >$

B

$< br > W \leq -t_{\alpha,n-1}$, por isso rejeitamos $H_0 < br >$

C

$< br > W gt; -z_{\alpha/2}$, por isso não rejeitamos $H_0 < br >$

D

$< br > W \leq -z_{\alpha}$, por isso não rejeitamos $H_0 < br >$

A alternativa A está correta.

Como não sabemos a média nem a variância de uma normal na qual queremos realizar um teste unilateral a respeito da média, verificaremos se $W gt; -t_{\alpha,n-1}$. Como, nesse caso,

$$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}} = \frac{8.26-10}{5.10/\sqrt{4}} = -0.68 \text{ e } t_{\alpha,n-1} = t_{0.05,3} = 2.35,$$

temos $W gt; -t_{\alpha,n-1}$, logo não rejeitamos H_0 .

Questão 2

Sejam $X_1, \ldots, X_{81}$ retirados de uma amostra aleatória de uma distribuição desconhecida. A média amostral e a variância amostral são, respectivamente, $\bar{X} = 8.25$ e $S^2 = 14.6$. Calcule o p-valor para o seguinte teste de hipótese: $H_0 : \mu = 9$ e $H_1 : \mu lt; 9$. Assinale a alternativa correta:

A

P-valor = 0.03869

B

P-valor = 0.09869

C

P-valor = 0.01869

D

P-valor = 0.10869

A alternativa A está correta.

Aqui temos uma amostra que não segue necessariamente uma normal, na qual $n = 81$ (ou seja, razoavelmente grande). Como discutimos previamente, para testar essa hipótese temos que saber se $W = -z_{\alpha}$. Nesse caso,

$$W = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{8.25 - 9}{\sqrt{14.6}/\sqrt{81}} = -1.767$$
.

Como vimos, o p-valor é dado por $P($ Erro Tipo 1 $)$ quando o ponto de corte que define se existe significância estatística ou não (i.e., se rejeitamos ou não a hipótese nula) é dado por $c = w_1 = -1.767$. Como a estatística de teste para esse teste é dada por $-z_\alpha$, temos $-z_\alpha = -1.767$.

Note que, por definição, $z_\alpha = \Phi^{-1}(1 - \alpha)$, assim obtemos a probabilidade de Erro Tipo 1 como:

$$\alpha = 1 - \Phi(1.767) = 0.0386$$

Segue que o p-valor é 0.0386.

Considerações finais

Vimos como obter estimadores intervalares e como realizar testes de hipótese. Estimadores intervalares são extremamente úteis quando levamos em conta a incerteza contida na amostra em nossa interpretação das estatísticas. Já testes de hipótese são parte do dia a dia da pesquisa estatística.

Quando temos dados em nossas mãos retirados de uma amostra — quase sempre é o caso —, desejamos dizer algo sobre a população inteira. Neste tema, vimos como fazer isso.

Podcast

Para encerrar, ouça sobre **Teste de hipóteses.**

Conteúdo interativo

Acesse a versão digital para ouvir o áudio.

Explore+

Para saber mais sobre os assuntos tratados neste tema, leia:

- **Bioestatística Básica,** de Sergio Miranda Freire. Disponível no site L@mpada do Departamento de Tecnologias da Informação e Educação em Saúde – DTIES da Universidade do Estado do Rio de Janeiro.

Referências

MORETTIN, P. A.; BUSSAB, W. O. **Estatística básica**. São Paulo: Saraiva Educação SA, 2017.