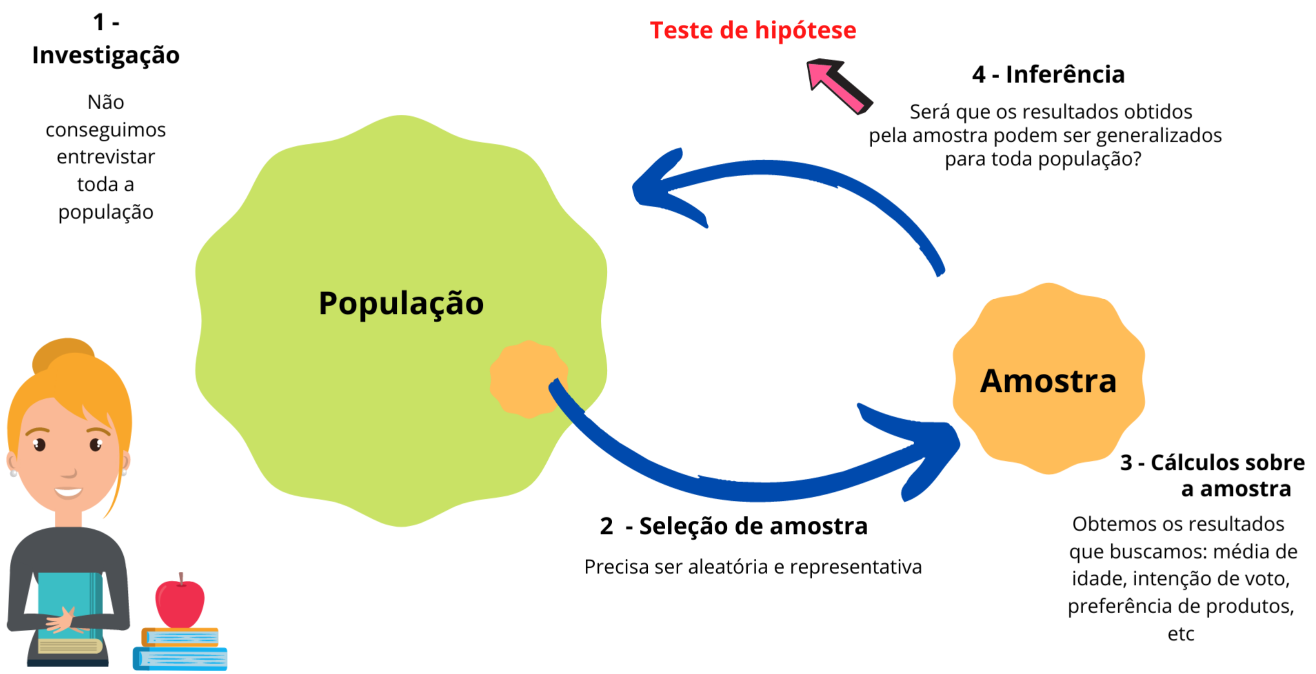
Métodos Quantitativos

André Amorim
Finanças Corporativas

contato@andreamorim.com.br
www.andreamorim.com.br

Anhanguera

11 ª Aula – Testes de hipóteses para a média



Feita determinada afirmação sobre uma população, usualmente sobre um parâmetro dessa, desejamos saber se os resultados experimentais provenientes de uma amostra contrariam ou não tal afirmação.

Muitas vezes, essa afirmação sobre a população é derivada de teorias desenvolvidas no campo substantivo do conhecimento.

A adequação ou não dessa teoria ao universo real pode ser verificada ou refutada pela amostra.

O objetivo do teste estatístico de hipóteses é, então, fornecer uma metodologia que nos permita verificar se os dados amostrais trazem evidências que apoiem ou não uma hipótese (estatística) formulada.

Formulando hipóteses

Considere, por exemplo, a afirmação (A) “vai chover hoje”. Essa afirmação pode ser considerada uma hipótese, cuja negativa é outra hipótese, (B) “não vai chover hoje”.

Observe que as duas hipóteses levantadas são complementares, isto é, ocorre a primeira ou ocorre a segunda, não há outra possibilidade.

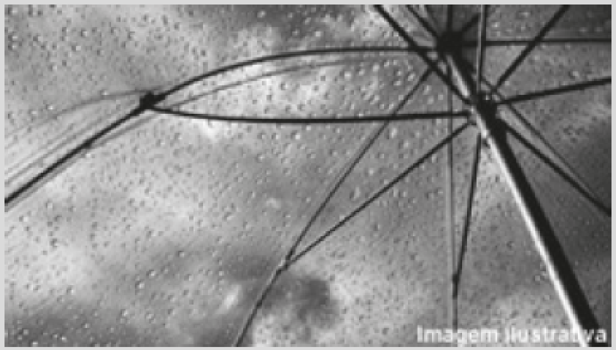
Formulando hipóteses

Como verificar a veracidade da hipótese (A)? É possível ter certeza absoluta da ocorrência de (A) ou (B)?

Para respondermos a essas perguntas, observe a Figura 3.9.

Figura 3.9 | Previsão do Tempo em Natal – RN para o dia 18/06/2015

Temperatura mínima: 21° C
Temperatura máxima: 29° C
Probabilidade de precipitação: 60%
Sol e aumento de nuvens de manhã.
Pancadas de chuva à tarde e à noite.



Fonte: Climatempo

Formulando hipóteses

Veja que a previsão do tempo para Natal traz uma informação muito importante, a probabilidade de precipitação, ou seja, a chance de chover.

Para facilitar nossa discussão, vamos denotar as hipóteses A e B como a seguir:

H_0: vai chover hoje

H_1: não vai chover hoje

Formulando hipóteses

A hipótese H_0, denominada hipótese nula, geralmente é afirmativa ou, no caso de uma variável quantitativa, uma hipótese de igualdade.

Ela é nossa principal hipótese, o foco da nossa análise e a que queremos pôr à prova.

A hipótese H_1, denominada hipótese alternativa, é aquela que será aceita se rejeitarmos a hipótese nula.

Formulando hipóteses

Em relação a nossa decisão de aceitar ou rejeitar H_0, podemos ter quatro resultados possíveis, elencados na Tabela 3.2.

Decisão	Possibilidades para H_0	
	Verdadeira	Falsa
Não rejeitar H_0	Decisão correta	Erro tipo II
Rejeitar H_0	Erro tipo I	Decisão correta

Formulando hipóteses

Para o nosso exemplo, a ocorrência do **erro tipo I** seria rejeitar a hipótese "vai chover hoje" e, ao final do dia, constatarmos que choveu.

Denotamos por α a probabilidade de ocorrência desse erro e, nesse caso, α = **60%**.
A ocorrência do **erro tipo II**, nesse caso, seria não rejeitar a hipótese "vai chover hoje" e, no final do dia, constatarmos que não choveu.

Formulando hipóteses

A probabilidade de ocorrência desse erro é **β = 40%.** Podemos escrever P(erro tipo I) =α e P(erro tipo II) = β.

Portanto, respondendo às perguntas feitas anteriormente, para verificar a veracidade da hipótese (A) temos de realizar um teste de hipóteses.

Contudo, nunca teremos certeza absoluta da ocorrência de uma hipótese, pois sempre estamos sujeitos a cometer um dos erros apresentados na Tabela 3.2.

Exemplificando

Uma variável $X \sim N(\mu, 18)$ é estudada em determinada população.

Parte dos pesquisadores suspeita que $\mu = \mu_1 = 55$ e outros que $\mu = \mu_2 = 50$

No intuito de pôr à prova essas suspeitas eles decidiram fazer testes para identificar qual delas é a correta.

Para isso foi retirada uma amostra da população, a qual é apresentada a seguir.

49 – 50 – 48 – 51 – 47 – 48 – 55 – 50 – 55 – 49 – 51 – 53

Com 95% de confiança, qual é a verdadeira média da população, $\mu_1 = 55$ ou $\mu_2 = 50$?

Exemplificando

Vamos inicialmente testar se $\mu = \mu_1 = 55$.

Passo 1 (elaborar as hipóteses): precisamos estipular duas hipóteses, a nula e a alternativa. Como a hipótese nula é sempre de igualdade, como foi descrito anteriormente, determinamos:

H_0: $\mu = 55$ H_1: $\mu \neq 55$

Exemplificando

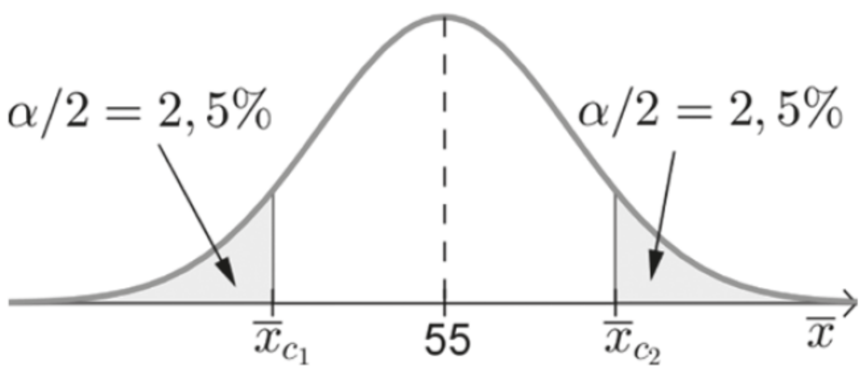
Passo 2 (determinar a estatística de teste): Como nosso objetivo é testar a média populacional da variável $X \sim N(\mu, 18)$, pelo TLC nossa estatística de teste será $\bar{x} \sim N(55, 18/12)$ ou $\bar{x} \sim N(55; 1{,}5)$, caso a hipótese nula seja verdadeira.

Exemplificando

Passo 3 (fixar o nível de significância): Como queremos 95% de confiança, a probabilidade de cometermos o erro tipo I deve ser $\alpha = 100\% - 95\% = 5\%$. Essa probabilidade também é denominada **nível de significância**.

Rejeitaremos a hipótese H_0 caso o valor $\bar{x}$ obtido a partir da amostra seja muito maior ou muito menor que $\mu_1 = 55$ ou, ainda, quando $\bar{x}$ pertencer à **região crítica** (RC), ilustrada na Figura 3.10.

Figura 3.10 | Região crítica para H_0: $\mu = 55$ e H_1: $\mu \neq 55$, com $\alpha = 5\%$



Fonte: O autor (2015)

Exemplificando

A região crítica pode ser denotada por $RC = \{\bar{x} \in \mathbb{R} | \bar{x} \leq \bar{x}_{c_1}$ ou $\bar{x} \geq \bar{x}_{c_2}\}$. Observando a tabela Z e lembrando que $z = \frac{\bar{x}-\mu}{\sqrt{\sigma^2/n}}$, temos:

$$z_1 = -1{,}96 = \frac{\bar{x}_{c_1} - 55}{\sqrt{1{,}5}} \Rightarrow -1{,}96\sqrt{1{,}5} = \bar{x}_{c_1} - 55 \Rightarrow \bar{x}_{c_1} = 55 - 1{,}96\sqrt{1{,}5} \cong 52{,}6;$$

$$z_2 = 1{,}96 = \frac{\bar{x}_{c_2} - 55}{\sqrt{1{,}5}} \Rightarrow 1{,}96\sqrt{1{,}5} = \bar{x}_{c_2} - 55 \Rightarrow \bar{x}_{c_2} = 55 + 1{,}96\sqrt{1{,}5} \cong 57{,}4;$$

$$RC = \{\bar{x} \in \mathbb{R} | \bar{x} \leq 52{,}6 \text{ ou } \bar{x} \geq 57{,}4\}.$$

Exemplificando

Passo 4 (calcular a estatística a partir da amostra): a média amostra é $\bar{x}_0 = 50{,}5$.

Passo 5 (tomar uma decisão): como $\bar{x}_0 \in RC$, decidimos rejeitar H_0,

isto é, há indícios suficientes que nos permitem refutar a possibilidade de a média populacional ser $\mu_1 = 55$.

Exemplificando

Vamos testar agora se $\mu = \mu_2 = 50$.

Passo 1 (elaborar as hipóteses):

H_0: $\mu = 50$ H_1: $\mu \neq 50$

Passo 2 (determinar a estatística de teste): $\bar{x} \sim N(50{,}18/12)$ ou $\bar{x} \sim N(50; 1{,}5)$, caso a hipótese nula seja verdadeira.

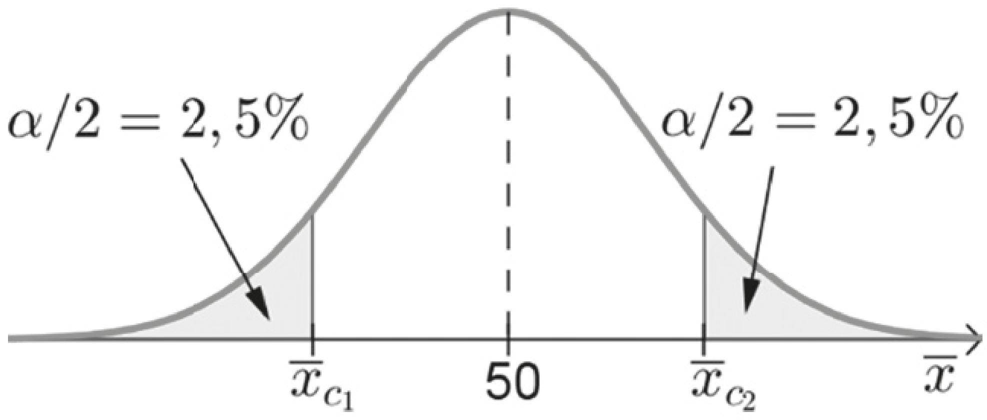
Exemplificando

Passo 3 (fixar o nível de significância): $\alpha = 5\%$

Rejeitaremos a hipótese H_0 caso o valor $\bar{x}$ obtido a partir da amostra pertença à **região crítica** (RC), ilustrada na Figura 3.11.

Observando a tabela Z e lembrando que $z = \frac{\bar{x}-\mu}{\sqrt{\sigma^2/n}}$, temos:

Figura 3.11 | Região crítica para H_0: $\mu = 50$ e H_1: $\mu \neq 50$, com $\alpha = 5\%$



Fonte: O autor (2015)

Exemplificando

$$z_1 = -1{,}96 = \frac{\bar{x}_{c_1} - 50}{\sqrt{1{,}5}} \Rightarrow$$

$$\Rightarrow -1{,}96\sqrt{1{,}5} = \bar{x}_{c_1} - 50 \Rightarrow \bar{x}_{c_1} = 50 - 1{,}96\sqrt{1{,}5} \cong 47{,}6;$$

$$z_2 = 1{,}96 = \frac{\bar{x}_{c_2} - 50}{\sqrt{1{,}5}} \Rightarrow 1{,}96\sqrt{1{,}5} = \bar{x}_{c_2} - 50 \Rightarrow \bar{x}_{c_2} = 50 + 1{,}96\sqrt{1{,}5} \cong 52{,}4;$$

Exemplificando

$RC = \{\bar{x} \in \mathbb{R} | \bar{x} \leq 47{,}6 \text{ ou } \bar{x} \geq 52{,}4\}$.

Passo 4 (calcular a estatística a partir da amostra): a média amostral é $\bar{x}_0 = 50{,}5$.

Passo 5 (tomar uma decisão): como $\bar{x}_0 \notin RC$, não podemos rejeitar H_0, isto é, não há indícios suficientes que nos permitam refutar a possibilidade de a média populacional ser $\mu_2 = 50$.

Desse modo, em concordância com o problema apresentado, devemos concluir que a verdadeira média da população é $\mu_2 = 50$.

FIM