Teste de Hipóteses

tradução do Jay Davore

índice

Exemplo (Morettin)

Uma industria usa como um dos componentes das máquinas que produz um parafuso importado, que deve ser resistente à tração. Suponha dois países A e B que produzem os parafusos, cada um com especificações diferentes:
resistência média à tração de A=145 kg σ_A=12 kg
resistência média à tração de B=155 kg σ_B=20 kg O edital do leilão indica que antes dele ser realizado, será divulgado a resistência média $\bar{x}$ de uma amostra de 25 parafusos do lote. Qual a regra de decisão para se inferir se os parafusos são de A ou de B?

Exemplo (Morettin)

primeira aproximação:

se $\bar{x} \leq 150$ os parafusos são de A, c.c., são de B

Se no dia do leilão somos informados que $\bar{x} = 148$, diremos que os parafusos são de A.

Exemplo (Morettin)

primeira aproximação:

se $\bar{x} \leq 150$ os parafusos são de A, c.c., são de B

Se no dia do leilão somos informados que $\bar{x} = 148$, diremos que os parafusos são de A.
podemos errar nessa conclussão?, i.e, uma amostra de 25 parafusos de B, pode apresentar média de 148?.

Exemplo (Morettin)

primeira aproximação:

se $\bar{x} \leq 150$ os parafusos são de A, c.c., são de B

Se no dia do leilão somos informados que $\bar{x} = 148$, diremos que os parafusos são de A.
podemos errar nessa conclussão?, i.e, uma amostra de 25 parafusos de B, pode apresentar média de 148?.

Exemplo (Morettin)

Podemos cometer dois tipos de erros:

Erro de tipo I: dizer que os parafusos são de A quando na realidade são de B, isto ocorre quando a amostra de B, apresenta média $\bar{x}$ inferior a 150.

Erro de tipo II: dizer que os parafusos são de B, quando na realidade são de A, isto ocorre quando a amostra de A, apresenta média $\bar{x}$ superior a 150.

Exemplo (Morettin)

Podemos cometer dois tipos de erros:

Erro de tipo I: dizer que os parafusos são de A quando na realidade são de B, isto ocorre quando a amostra de B, apresenta média $\bar{x}$ inferior a 150.
Erro de tipo II: dizer que os parafusos são de B, quando na realidade são de A, isto ocorre quando a amostra de A, apresenta média $\bar{x}$ superior a 150.

Exemplo (Morettin)

São definidas duas hipóteses numeradas:

Hipótese H_0: os parafusos são de origem B. Isto equivale a dizer que a resistência X de cada parafuso segue uma distribuição com média $\mu = 155$ e desvio padrão $\sigma = 20$

Hipótese H_1: os parafusos são de A, isto é, $\mu = 145$ com $\sigma = 12$.

Exemplo (Morettin)

São definidas duas hipóteses numeradas:

Hipótese H_0: os parafusos são de origem B. Isto equivale a dizer que a resistência X de cada parafuso segue uma distribuição com média $\mu = 155$ e desvio padrão $\sigma = 20$
Hipótese H_1: os parafusos são de A, isto é, $\mu = 145$ com $\sigma = 12$.

Exemplo (Morettin)

Construimos agora a RC (região crítica) para o teste:

$$RC = \{y \in \mathcal{R} | y \leq 150\}$$

e denotamos a probabilidade de cometer cada um dos erros:

$$P(\text{Erro tipo I}) = P(\bar{X} \in RC | H_0 \text{é verdadeira}) = \alpha$$

$$P(\text{Erro tipo II}) = P(\bar{X} \notin RC | H_1 \text{é verdadeira}) = \beta$$

Exemplo (Morettin)

Construimos agora a RC (região crítica) para o teste:

$$RC = \{y \in \mathcal{R} | y \leq 150\}$$

e denotamos a probabilidade de cometer cada um dos erros:

$$P(\text{Erro tipo I}) = P(\bar{X} \in RC | H_0 \text{é verdadeira}) = \alpha$$

$$P(\text{Erro tipo II}) = P(\bar{X} \notin RC | H_1 \text{é verdadeira}) = \beta$$

Exemplo (Morettin)

Quando H_0 for verdadeira, $\bar{X} \sim N(155, 16)$:

$$\begin{aligned} P(\text{erro I}) &= P(\bar{X} \in RC | H_0 \quad \text{é verdadeira}) \\ &= P(\bar{X} \leq 150 | \bar{X} \sim N(155, 16)) \\ &= P(Z \leq \frac{150 - 155}{4}) \\ &= P(Z \leq -1,25) = 0,10565 = 10,56\% = \alpha \end{aligned}$$

Exemplo (Morettin)

Quando H_1 for verdadeira, $\bar{X} \sim N(145; 5,76)$:

$$
\begin{aligned}
P(\text{erro II}) &= P(\bar{X} \notin RC | H_1 \quad \text{é verdadeira}) \\
&= P(\bar{X} > 150 | \bar{X} \sim N(145; 5,76)) \\
&= P(Z > \frac{150 - 145}{2,4}) \\
&= P(Z > 2,08) = 0,01876 = 1,88\% = \beta
\end{aligned}
$$

Exemplo (Morettin)

Com a regra de decisão adotada estaremos comentendo o erro tipo I com maior probabilidade que o erro tipo II (privilegiando a afirmação que os parafusos são de A).

Origen Real	Decisão por A	Decisão por B
A	Sem Erro	Erro tipo II ($\beta = 1,88\%$)
B	Erro tipo I ($\alpha = 10,56\%$)	Sem erro

Exemplo

Ao escolher um valor de $\bar{x}$ diferente de 150, apenas as probabilidades de α e β mudarão. Se escolhermos $\bar{x} < 150$ α diminui e β aumenta. Logo, deve existir um ponto em que α seja igual a β. Este ponto é $\bar{x} = 148,75$, neste caso $\alpha = \beta = 5,94\%$

Exemplo

Se fixarmos $\alpha = 5\%$:

$$
\begin{aligned}
5\% = P(\text{erro tipo I} &= P(\bar{X} \leq \bar{X}_c | \bar{X} \sim N(155, 16)) \\
&= P(Z \leq -1,645)
\end{aligned}
$$

neste caso $\bar{x}_c = 148,42$ e a Regra de decisão será:
Se $\bar{x}_c$ for inferior a 148,42, dizemos que o lote é de A, caso contrário é de B.
A probabilidade de erro tipo II é:

$$
\begin{aligned}
\beta = P(\text{erro tipo II} &= P(\bar{X} > 148,42 | \bar{X} \sim N(145; 5,76)) \\
&= P(Z > 1,425) = 7,93\%
\end{aligned}
$$

Exemplo

O procedimento é comúm em casos que a decisão que tomamos não é apenas entre duas possíveis populações. Suponha as seguintes hipóteses:

H_0: os parafusos são de origem B ($\mu = 155$ e $\sigma = 20$)
H_1: os parafusos não são de origem B (μ e σ desconhecidos)

os parâmetros sob H_1 não podem ser especificados.
podemos especificar ainda hipóteses como

Exemplo

O procedimento é comúm em casos que a decisão que tomamos não é apenas entre duas possíveis populações. Suponha as seguintes hipóteses:

H_0: os parafusos são de origem B ($\mu = 155$ e $\sigma = 20$)
H_1: os parafusos não são de origem B (μ e σ desconhecidos)

os parâmetros sob H_1 não podem ser especificados.
podemos especificar ainda hipóteses como

H_1: os parafusos não são de origem B ($\mu < 155$ e σ qualquer)

Se fixarmos α, a regra de decisão dependerá apenas das informações de H_0.

Exemplo

O procedimento é comúm em casos que a decisão que tomamos não é apenas entre duas possíveis populações. Suponha as seguintes hipóteses:

H_0: os parafusos são de origem B ($\mu = 155$ e $\sigma = 20$)
H_1: os parafusos não são de origem B (μ e σ desconhecidos)

os parâmetros sob H_1 não podem ser especificados.
podemos especificar ainda hipóteses como

H_1: os parafusos não são de origem B ($\mu < 155$ e σ qualquer)

Se fixarmos α, a regra de decisão dependerá apenas das informações de H_0.

Exemplo

nesse caso, não podemos encontrar β, pois não temos um único parâmetro μ como alternativa e nada sabemos sobre σ. e a regra de decisão será:

Se $\bar{x} > 148,42$, diremos que o lote é de origem B, caso contrário, não é de origem B

Origem real	Decisão Não B	Decisão B
B	Erro tipo I, $\alpha = 5\%$	Sem erro
não B	Sem erro	Erro tipo II, $\beta = ?$

Função característica de operação

a Função característica de operação é definida para o exemplo como:

$$\beta(\mu) = P(\text{aceitar} H_0|\mu) = P(\bar{X} > 148,42|\mu)$$

Isto é, a probabilidade de aceitar H_0 como função de μ.
Usualmente consideramos

$$\pi(\mu) = 1 - \beta(\mu)$$

que é a probabilidade de rejeitar H_0 em função de μ.

esta função é chamada de função poder do teste
consideraremos para isto que σ é o mesmo para todos os valores de μ

Função característica de operação

a Função característica de operação é definida para o exemplo como:

$$\beta(\mu) = P(\text{aceitar} H_0|\mu) = P(\bar{X} > 148,42|\mu)$$

Isto é, a probabilidade de aceitar H_0 como função de μ.
Usualmente consideramos

$$\pi(\mu) = 1 - \beta(\mu)$$

que é a probabilidade de rejeitar H_0 em função de μ.

esta função é chamada de função poder do teste
consideraremos para isto que σ é o mesmo para todos os valores de μ

Exemplo (Morettin)

Se não tivermos idéia sobre os valores das resistências médias dos parafusos dos outros países, a situação nos levaria á seguinte hipótese alternativa:

H_1: os parafusos não são de origem B ($\mu \neq 155$)

Exemplo (Morettin)

Se não tivermos idéia sobre os valores das resistências médias dos parafusos dos outros países, a situação nos levaria á seguinte hipótese alternativa:

H_1: os parafusos não são de origem B ($\mu \neq 155$)
A regra de decisão deverá especificar dois pontos $\bar{x}_{c_1}$ e $\bar{x}_{c_2}$

Exemplo (Morettin)

Se não tivermos idéia sobre os valores das resistências médias dos parafusos dos outros países, a situação nos levaria á seguinte hipótese alternativa:

H_1: os parafusos não são de origem B ($\mu \neq 155$)
A regra de decisão deverá especificar dois pontos $\bar{x}_{c_1}$ e $\bar{x}_{c_2}$
Se $\bar{x}$ estiver entre $\bar{x}_{c_1}$ e $\bar{x}_{c_2}$, diremos que os parafusos são de origem B, fora desse intervalo, não são de origem B

Exemplo (Morettin)

Se não tivermos idéia sobre os valores das resistências médias dos parafusos dos outros países, a situação nos levaria á seguinte hipótese alternativa:

H_1: os parafusos não são de origem B ($\mu \neq 155$)
A regra de decisão deverá especificar dois pontos $\bar{x}_{c_1}$ e $\bar{x}_{c_2}$
Se $\bar{x}$ estiver entre $\bar{x}_{c_1}$ e $\bar{x}_{c_2}$, diremos que os parafusos são de origem B, fora desse intervalo, não são de origem B

Fixado α, existirão muitos valores que satisfazem essa condição. Daremos preferência àquelas soluções $\bar{x}_{c_1}$ e $\bar{x}_{c_2}$, simétricas em relação à média.

Exemplo (Morettin)

Se não tivermos idéia sobre os valores das resistências médias dos parafusos dos outros países, a situação nos levaria á seguinte hipótese alternativa:

H_1: os parafusos não são de origem B ($\mu \neq 155$)
A regra de decisão deverá especificar dois pontos $\bar{x}_{c_1}$ e $\bar{x}_{c_2}$
Se $\bar{x}$ estiver entre $\bar{x}_{c_1}$ e $\bar{x}_{c_2}$, diremos que os parafusos são de origem B, fora desse intervalo, não são de origem B

Fixado α, existirão muitos valores que satisfazem essa condição. Daremos preferência àquelas soluções $\bar{x}_{c_1}$ e $\bar{x}_{c_2}$, simétricas em relação à média.

Exemplo (Morettin)

Se fixarmos α em 5%:

$$\begin{aligned} 0,05 = P(\text{erro tipo I}) &= P(\bar{X} > \bar{x}_{c_1} \text{ ou } \bar{X} > \bar{x}_{c_2} | \bar{X} \sim N(155; 16)) \\ &= P(Z < -1,96 \text{ ou } Z > 1,96) \end{aligned}$$

de onde:

$$-1,96 = \frac{(\bar{x}_{c_1} - 155)}{4} \Rightarrow \bar{x}_{c_1} = 147,16$$

e

$$1,96 = \frac{(\bar{x}_{c_2} - 155)}{4} \Rightarrow \bar{x}_{c_2} = 162,84$$

A região de Rejeição (RC) da hipótese H_0 é:

$$RC = \{\bar{x} \in \mathcal{R} | \bar{x} < 147,16 \text{ ou } \bar{x} > 162,84\}$$

Hipóteses procedimento de teste

- A Hipótese nula, denotada por H_0 é a afirmação que é inicialmente assumida como verdade.
- A Hipótese alternativa denotada por H_a ou H_1, é uma afirmação contrári a H_0.
- as possíveis conclusões do processo de teste de hipóteses são ou rejeitar H_0 ou não rejeitar H_0.

Um teste de hipóteses

- Um teste de hipóteses é um método que mediante o uso de dados amostrais, permite decidir se a hipótese nula deve ou não ser rejeitada.

Procedimento de teste

um procedimento de teste é especificado por:

1. Uma prova estatística: uma função de dados amostrais nas quais a decisão será baseada.
2. Uma região de rejeição: O conjunto de todos os valores do teste estatístico para o qual H_0 será rejeitada (a hipótese nula será rejeitada se e somente se o valor do teste estatístico pertence a essa região).

errores em teste de hipótese

1. O erro do tipo I, consiste em rejeitar a hipótese nula H_0 quando ela é verdadeira.
2. o erro do tipo II, envolve não rejeitar H_0 quando H_0 é falsa.

região de rejeição α e β

Suponha um experimento no qual o tamanho amostral é fixado e o teste estatístico é escolhido.

Ao diminuir o tamanho da região de rejeição de forma a obter um menor valor de α, resulta em aumentar o tamanho de β, para qualquer valor do parâmetro consistente com H_a.

nível de significância

Especifica o maior de α que pode ser tolerado e encontrar a região de rejeição que tenha o valor de α.
Isto resulta em β ter o menor valor possível, sujeito ao limite en α.
O valor resultante de α é referido como o nível de significância.

Nível do teste α

Um teste que corresponde ao nível de significância é chamado de nivel do teste α.
Um teste com nível de significância α é aquel pelo qual a probabilidade do erro de Tipo I é controlada num nível específico.

Passos recomendados em análise de teste de hipóteses

- Identifique o parâmetro de interesse θ e descreva no contexto do problema atual
- Determine o valor nulo e estabeleça a hipótese nula $H_0 : \theta = \theta_0$
- Estabeleça a hipótese alternativa $H_1 : \theta \neq \theta_0$(ou $H_1 : \theta < \theta_0$, ou $H_1 : \theta > \theta_0$)
- Dê a formula para o valor calculado da estatística de teste

Passos recomendados em análise de teste de hipóteses

- Estabeleça a região de rejeição para o nível de significância selecionado. (Fixe a probabilidade α de cometer o erro do tipo I- Lembre que a região é construida usando os valores do parâmetro hipotetizado em H_0.)
- Calcule qualquer quantidade amostral necessária, substitua na fórmula para o valor do teste estatístico, e calcule esse valor.
- Decida si H_0 será rejeitada e estabeleça as conclusões para o problema atual.

A formulação das hipóteses 2 e 3, devem ser feitas antes de examinar os dados.

Caso I: Populaçao Normal com σ conhecido

- Hipótese Nula:

$$H_0 : \mu = \mu_0$$

- valor do teste estatístico:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

Caso I: Populaçao Normal com σ conhecido

Hipótese alternativa	Região de rejeição para o teste nível α
$H_a : \mu > \mu_0$	$z \geq z_\alpha$
$H_a : \mu < \mu_0$	$z \leq -z_\alpha$
$H_a : \mu \neq \mu_0$	$z \geq z_{\alpha/2}$ ou $z \leq -z_{\alpha/2}$

Exemplo (Morettin)

Uma máquina automática de encher pacotes de café o faz segundo uma N(μ, $\sigma^2 = 400g^2$). INicialmente a máquina está regulada para $\mu = 500g$. Uma amostra é colhida periódicamente para verificar se o processo está sob controle. Se uma amostra apresentar $\bar{x} = 492$ você pararía a produção para regular a máquina?

Exemplo (Morettin)

Seja X o peso de cada pacote. Então $X \sim N(\mu, 40)$, assim:

$$H_0 : \quad \mu = 500g.$$
$$H_1 : \quad \mu \neq 500g$$

(a máquina pode ser regulada para mais ou para menos)

Pela suposição de que $\sigma^2 = 400$ ela será a mesma. Para todo μ, $\bar{X} \sim N(\mu, 400/16)$. Isto é, se H_0 é verdadeiro, $\bar{X} \sim N(500, 25)$

Exemplo (Morettin)

Seja X o peso de cada pacote. Então $X \sim N(\mu, 40)$, assim:

$$H_0 : \quad \mu = 500g.$$
$$H_1 : \quad \mu \neq 500g$$

(a máquina pode ser regulada para mais ou para menos)
Pela suposição de que $\sigma^2 = 400$ ela será a mesma. Para todo μ, $\bar{X} \sim N(\mu, 400/16)$. Isto é, se H_0 é verdadeiro, $\bar{X} \sim N(500, 25)$
Se fixarmos $\alpha = 0,01$, veremos que H_0 deverá ser rejeitada quando $\bar{X}$ for muito pequena ou muito grande (teste bilateral).

Exemplo (Morettin)

Seja X o peso de cada pacote. Então $X \sim N(\mu, 40)$, assim:

$$\begin{aligned} H_0 : &\quad \mu = 500g. \\ H_1 : &\quad \mu \neq 500g \end{aligned}$$

(a máquina pode ser regulada para mais ou para menos)
Pela suposição de que $\sigma^2 = 400$ ela será a mesma. Para todo μ, $\bar{X} \sim N(\mu, 400/16)$. Isto é, se H_0 é verdadeiro, $\bar{X} \sim N(500, 25)$
Se fixarmos $\alpha = 0,01$, veremos que H_0 deverá ser rejeitada quando $\bar{X}$ for muito pequena ou muito grande (teste bilateral).

Exemplo (Morettin)

temos que

$$z_1 = -2,58 \qquad = (\bar{x}_{c_1} - 500)/5 \Rightarrow \bar{x}_{c_1} = 487,1$$
$$z_2 = 2,58 \qquad = (\bar{x}_{c_2} - 500)/5 \Rightarrow \bar{x}_{c_2} = 512,9$$

segue-se a região critica é:

$$RC = \{\bar{x} \in \mathcal{R} | \bar{x} \leq 487,1 \quad ou \quad \bar{x} \geq 512,9\}$$

Exemplo (Morettin)

no caso do exemplo: $\bar{x}_0 = 492$

como $\bar{x}_0$ não pertence à RC, concluimos que não devemos rejeitar H_0.

Exemplo (Morettin)

no caso do exemplo: $\bar{x}_0 = 492$
como $\bar{x}_0$ não pertence à RC, concluimos que não devemos rejeitar H_0.

Probabilidade tipo II

hipótese alternativa	probabilidade Tipo II $\beta(\mu^`)$
$H_a : \mu > \mu_0$	$\Phi\left(z_\alpha + \frac{\mu_0 - \mu^`}{\sigma/\sqrt{n}}\right)$
$H_a : \mu < \mu_0$	$1 - \Phi\left(-z_\alpha + \frac{\mu_0 - \mu^`}{\sigma/\sqrt{n}}\right)$
$H_a : \mu \neq \mu_0$	$\Phi\left(z_{\alpha/2} + \frac{\mu_0 - \mu^`}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\alpha/2} + \frac{\mu_0 - \mu^`}{\sigma/\sqrt{n}}\right)$

Tamanho da amostra

O tamanho da amostra n para o qual o nível do teste α e também o $\beta(\mu^{`}) = \beta$ no valor alternativo $\mu^{`}$ é:

$$n = \begin{cases} \left[\frac{\sigma(z_\alpha + z_\beta)}{\mu_0 - \mu^{`}}\right]^2 & \text{teste unicauldal} \\ \left[\frac{\sigma(z_{\alpha/2} + z_\beta)}{\mu_0 - \mu^{`}}\right]^2 & \text{teste bicauldal} \end{cases}$$

Amostras grandes

Quando o tamanho amostral é grande, os testes z para o caso I são modificados para estabelecer procedimentos de teste válidos sem necessidade de uma população normal ou um σ conhecido.

Amostras grandes ($n > 40$)

para n grande, s é proximo a σ

$$\text{Teste estatístico:} \qquad Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

o uso da região de rejeição para o caso I, resulta em um procedimento de teste para o qual o nível de significância é aproximadamente α.

População normal

Se $X_1, ..., X_n$ é uma a.a. de uma distribuição normal, a variável padronizada:

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

tem distribuição t com $n-1$ graus de liberdade

Teste t

$$\begin{aligned}\text{Hipótese Nula :} \qquad & H_0 : \mu = \mu_0 \\ \text{Valor da estatística do teste:} \qquad & t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}\end{aligned}$$

Teste t para uma amostra

Hipótese alternativa	Região de rejeição para o teste nível α
$H_a : \mu > \mu_0$	$t \geq t_{\alpha-1}$
$H_a : \mu < \mu_0$	$t \leq -t_{\alpha,n-1}$
$H_a : \mu \neq \mu_0$	$t \geq t_{\alpha/2,n-1}$ ou $t \leq -t_{\alpha/2,n-1}$

Uma típica curva β para o teste t

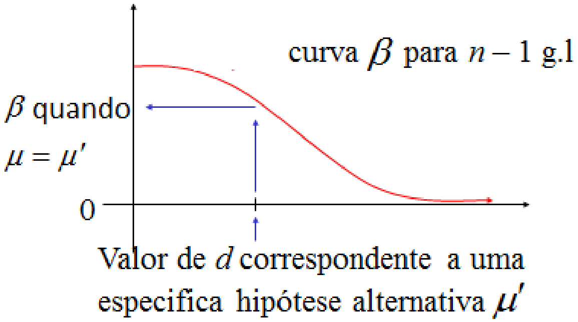


Figura: curva t

Testes para grandes amostras sobre p

Seja p a proporção de individuos ou objetos em uma população que possuem uma propriedade especifica.
Testes para amostras grandes sobre p, são um caso especial de procedimentos gerais para amostras grandes sobre um parâmetro θ.

Testes para grandes amostras sobre p

- Hipótese Nula:

$$H_0 : p = p_0$$

- Valor do teste estatístico:

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

Testes para grandes amostras sobre p

Hipótese alternativa	Região de rejeição para o teste nível α
$H_a : p > p_0$	$z \geq z_\alpha$
$H_a : p < p_0$	$z \leq -z_\alpha$
$H_a : p \neq p_0$	$z \geq z_{\alpha/2}$ ou $z \leq -z_{\alpha/2}$

deve se provar que é válido:

$$np_0 \geq 10 \quad \text{e} \quad n(1 - p_0) \geq 10$$

Exemplo (Morettin)

Uma estação de TV afirma que 60 % das tvťs estavam ligados no seu programa especial do fim de semana. Uma rede competidora quer mostrar que esta afirmação não é verdadeira e usa uma amostra de 200 famílias para um teste. Qual deve ser o procedimento adotado para avaliar a veracidade da informação?

Exemplo (Morettin)

Colocamos à prova a afirmação:

$$H_0: \ p = 0,60$$

se a afirmação não for verdadeira, espera-se uma proporção menor. A hipótese alternativa seria:

$$H_1: \ p < 0,60$$

a estatística a ser usada é $\hat{p}$, da teoria sabemos que

$$\hat{p} \sim N(p, \frac{p(1-p)}{200})$$

Exemplo (Morettin)

Fixamos o $\alpha = 0,05$ teremos:

$$\hat{p} \sim N(0,60, 0,24/200)$$

que fornece a RC:

$$RC = \{\hat{p} \in \mathcal{R} | \hat{p} \leq 0,544\}$$

Exemplo (Morettin)

Devemos achar $\hat{p}_c$ tal que $P(\hat{p} \leq \hat{p}_c) = 0,05$, usando a aproximação normal teremos

$$P(Z \leq \frac{\hat{p}_c - 0,60}{\sqrt{0,24/200}}) = 0,05$$

o que implica

$$\frac{\hat{p}_c - 0,60}{\sqrt{0,24/200}} = -1,645$$

Segue que $\hat{p}_c = 0,544$
Admitindo que das 200 pessoas pesquisadas, 104 dizeram que assistiram o programa, isto é, $\hat{p} = 0,52$
Vemos que $0,52 \in \mathcal{R}$, portanto somos levados a rejeitar H_0.

Expressões gerais para $\beta(p^`)$

Hipótese alternativa	$\beta(p^`)$
$H_a : p > p_0$	$\Phi\left(\dfrac{p_0 - p^` + z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p^`(1-p^`)/n}}\right)$
$H_a : p < p_0$	$1 - \Phi\left(\dfrac{p_0 - p^` - z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p^`(1-p^`)/n}}\right)$

Expressões gerais para $\beta(p^{`})$

Hipótese alternativa	$\beta(p^{`})$
$H_a : p \neq p_0$	$\Phi\left(\dfrac{p_0 - p^{`} + z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p^{`}(1-p^{`})/n}}\right) - \Phi\left(\dfrac{p_0 - p^{`} - z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p^{`}(1-p^{`})/n}}\right)$

tamanho amostral

O tamanho amostral n para o qual o nível α do teste têm $\beta(p^`) = p$ é:

$$n = \begin{cases} \left[\dfrac{z_\alpha\sqrt{p_0(1-p_0)}+z_\beta\sqrt{p^`(1-p^`)}}{p^`-p_0}\right]^2 & \text{teste unicaudal} \\ \left[\dfrac{z_{\alpha/2}\sqrt{p_0(1-p_0)}+z_\beta\sqrt{p^`(1-p^`)}}{p^`-p_0}\right]^2 & \text{teste bicaudal} \end{cases}$$

Testes para a variância de uma normal

- Usaremos a distribuição Qui-quadrado
- Considere $\bar{X}$ e S^2, obtidas de uma amostra $(X_1, ..., X_n)$ de $X \sim N(\mu, \sigma^2)$. A soma:

$$\left(\frac{X_1 - \mu}{\sigma}\right)^2 + .. + \left(\frac{X_n - \mu}{\sigma}\right)^2 \quad \sim \chi^2(n)$$

Desde que cada $\left(\frac{X_i - \mu}{\sigma}\right) \sim N(0, 1)$

Testes para a variância de uma normal

- Se definirmos:

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

teremos que

$$Y = \frac{n\hat{\sigma}_*^2}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{(X_i - \mu)}{\sigma} \right)^2 \sim \chi^2(n)$$

Observe que $\hat{\sigma}_*^2$ é muito parecido com $\hat{\sigma}^2$ com μ em lugar de $\bar{X}$

Testes para a variância de uma normal

É necessário conhecer a distribuição de $\sum_{i=1}^{n}(X_i - \bar{X})^2$ para ter a distribuição de S^2
Note que:

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

dado que $\sum_{i=1}^{n}(X_i - \bar{X}) = 0$

Testes para a variância de uma normal

Dividindo por σ^2 e reescrevendo de maneira conveniente, temos:

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

O primeiro termo tem distribuição $\chi^2(n)$, o último $\chi^2(1)$ é razoável pensar que o segundo tenha $\chi^2(n-1)$.

Propriedades importantes

- Seja $(Z_1, ..., Z_n)$ a.a.s de uma $N(0,1)$, então:
  a) $\bar{Z} \sim N(0, 1/n)$
  b) $\bar{Z}$ e $\sum_{i=1}^{n}(Z_i - \bar{Z})^2$ são independentes
  c) $\sum_{i=1}^{n}(Z_i - \bar{Z})^2 \sim \chi^2(n-1)$
- $\chi^2(p) + \chi^2(q) = \chi^2(p+q)$

Testes para a variância de uma normal

Queremos testar:

$$H_0: \ \sigma^2 = \sigma_0^2$$

$$H_1: \ \sigma^2 \neq \sigma_0^2$$

supondo que $X_i \sim N(\mu, \sigma^2)$ $(i = 1, ..., n)$ e os X_i são independentes.
A estatística do teste será sob H_0:

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$$

Testes para a variância de uma normal

como o teste é bilateral, temos a região crítica:

$$P(\chi^2 \in RC|H_0) = P(\chi^2 < \chi_1^2 \quad ou \quad \chi^2 > \chi_2^2) = \alpha$$

sendo α fixado a priori.

observado o valor de s_0^2 de S^2, obtemos $\chi_0^2 = \frac{(n-1)S_0^2}{\sigma_0^2}$

Se $\chi_0^2 \in RC$ rejeitamos H_0, caso contrário, não rejeitamos H_0

Exemplo (Morettin)

Mantendo o controle da variabilidade:
Uma maquina de encher pacotes de café está regulada para uma média de $500g$ e $\sigma = 10g$. O peso de cada pacote $X \sim N(\mu, \sigma^2)$. Colheu-se uma amostra de $n = 16$ pacotes e foi observado que $S^2 = 169g^2$. Você diria que a máquina está desregulada em relação à variância?

Exemplo (Morettin)

Testaremos então:

$$H_0: \ \sigma^2 = 100$$

$$H_1: \ \sigma^2 \neq 100$$

fixado α em 5% e n=16 temos que a região critica é dada por:

$$RC = \{\chi^2 : \chi^2 \leq 6,262 \quad ou \quad \chi^2 \geq 27,488\}$$

o valor observado da estatística é:

$$\chi_0^2 = \frac{(n-1)s_0^2}{\sigma_0^2} = \frac{(15)(169)}{100} = 25,35$$

Como $\chi_0^2 \notin RC$ não rejeitamos H_0

Intervalo de confiança para σ^2

Da expressão:

$$P(\chi_1^2 \leq \frac{(n-1)S^2}{\sigma^2} < \chi_2^2) = \gamma$$

obtemos a igualdade:

$$\frac{(n-1)S^2}{\chi_2^2} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi_1^2}$$

Teste sobre a média de uma normal com variância desconhecida

Consideramos agora a estatística:

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

dividindo por σ em ambos os membros da fração:

$$\frac{(\sqrt{n}(\bar{X} - \mu)/\sigma)}{(S/\sigma)}$$

Teste sobre a média de uma normal com variância desconhecida

o numerador

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

o quadrado do denominador pode ser escrito como:

$$\frac{(n-1)S^2}{\sigma^2}/(n-1) = \frac{Y}{n-1}$$

onde se os X_i são normais, $Y \sim \chi^2(n-1)$
Logo a estatística é o quociente entre uma $N(0, 1)$ e a raíz quadrada de uma $\chi^2(n-1)$ dividida pelo número de graus de liberdade, portanto:

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

Teste sobre a média de uma normal com variância desconhecida

Testamos as hipóteses:

$$H_0 : \quad \mu = \mu_0$$

$$H_1 : \quad \mu \neq \mu_0$$

a hipótese alternativa poderia ser do tipo unilateral à esquerda ou à direita.
A estatística a ser usada é:

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$$

Teste sobre a média de uma normal com variância desconhecida

fixado o valor de α e colhida a amostra de n individuos, calcularemos $\bar{x}_0$ e s^2 e também o valor $t_0 = \frac{\bar{x}_0 - \mu_0}{s_0/\sqrt{n}}$ de T. dependendo do valor encontrado (estar ou não na RC), rejeitaremos ou não H_0.

IC para a média

temos que

$$P\left(-t_\gamma < \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_\gamma\right) = \gamma$$

O IC será:

$$IC(\mu, \gamma) = \bar{X} \pm t_\gamma \frac{S}{\sqrt{n}}$$

Exemplo (Morettin)

Um fabricante afirma que seus cigarros contêm não mais que 30mg de nicotina. Uma amostra de 25 cigarros fornece média de $31,5$mg. e desvio padrão 3 mg. No nível de 5%, os dados refutam a afirmação do fabricante?

Exemplo (Morettin)

As hipóteses são:

$$H_0 : \quad \mu = 30$$

$$H_1 : \quad \mu > 30$$

Supondo que a quantidade de nicotina X por cigarro $X \sim N(\mu, \sigma^2)$ a estatística :

$$T = \frac{\sqrt{25}(\bar{X} - 30)}{S} \sim t(24)$$

Exemplo (Morettin)

Procuramos o valor tal que

$$P(T > t_c) = 0,05$$

obtendo da tabela que $t_c = 1,711$, ou seja a região critica para T é:

$$RC = \{1,711, +\infty\}$$

O valor observado da estatística é:

$$t_0 = \frac{5(31,5 - 30)}{3} = 2,5$$

como $t_0 \in RC$ Rejeitamos H_0

definição

Procedimentos de teste quando o tamanho da amostra n é pequena estão diretamente baseados na distribuição binomial (antes do que na aproximação normal).

$$P(tipoI) = 1 - bin(c-1; n, p_0)$$
$$\beta(p^{`}) = bin(c-1; n, p^{`})$$

Probabilidade e significância

O método de construção de testes de hipóteses parte da fixação do nível de significância α. Outra maneira de proceder consiste em apresentar a probabilidade de significância ou nível descritivo ou p-valor do teste.

Exemplo

No exemplo da estação de televisão, que afirma que 60% dos televisores estavam ligados no seu programa especial, e utilizando uma amostra de 200 famílias, deseja testar:

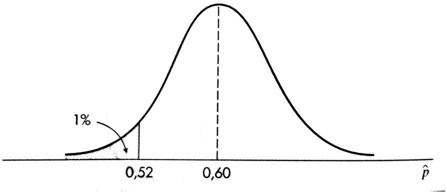
$$H_0 : \quad p = 0,60$$

admitindo que H_0 é verdadeira, temos que
$\hat{p} \sim N(p, \frac{p(1-p)}{n}) = N(0,6; 0,24/200)$.
Colhida a amostra, obtivemos $\hat{p}_0 = 104/200 = 0,52$. Portanto, podemos calcular qual a probabilidade de ocorrerem valores de $\hat{p}$ mais desfavoráveis para H_0 do que esse.
É evidente que quanto menor for $\hat{p}$, maior será a evidência contra
$H_0 : \quad p = 0,6$

Exemplo

$$\begin{aligned} P(\hat{p} < 0,52 | p = 0,6) &= P(Z < \frac{\sqrt{200}(0{,}52 - 0{,}6)}{\sqrt{0{,}24}}) \\ &= P(Z < -2,3) = 0,01 = 1\% \end{aligned}$$

O resultado mostra que se a audiência do programa fosse realmente de 60%, a probabilidade de encontrarmos uma amostra de 200 famílias com 52% ou menos de audiência é de 1%. Fato raro de acontecer, portanto, somos levados a rejeitar a H_0.


1%
0,52
0,60
$\hat{p}$

definição

O P-valor é o menor nível de significância para o qual H_0 seria rejeitado quando um procedimento de teste específico é utilizado em conjunto de dados.

1. $P - valor \leq \alpha$

$$\Rightarrow \text{rejeitar } H_0 \text{ ao nível } \alpha$$

2. $P - valor > \alpha$

$$\Rightarrow \text{não rejeitar } H_0 \text{ ao nível } \alpha$$

$P - valor$

O $P - valor$ é a probabilidade calculada assumindo H_0 como verdadeira, de obter um valor do teste estatístico tão contrário a H_0 como o valor atual.
quanto menor for o $P - valor$, os dados mostram-se mais contrários à hipótese H_0 (reforçam a idéia de rejeição de H_0).

$P - valores$ para um teste z

$P - valor$:

$$P = \begin{cases} 1 - \Phi(z) & \text{teste para cauda superior} \\ \Phi(z) & \text{teste para cauda inferior} \\ 2\left[1 - \Phi(|z|)\right] & \text{teste bi-caudal} \end{cases}$$

área do $p - valor$

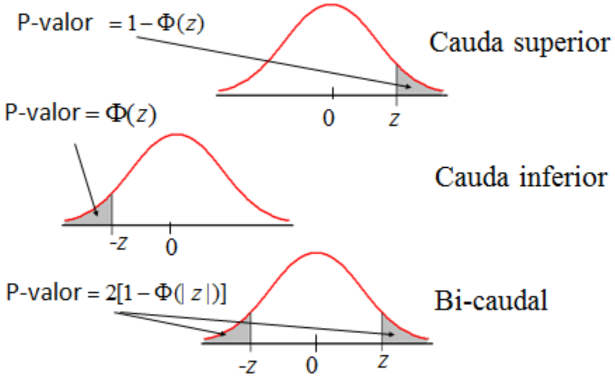


Figura: $p - valores$ para os distintos testes

Exemplo

A espessura-alvo dos wafers de silício usados em um tipo de circuito integrado é 245μ m. Uma amostra de 50 wafers é obtida, e a espessura de cada uma é determinada, resultando em uma espessura média amostral de $246,8\mu$m e um desvio padrão da amostra de 360μm. Esses dados sugerem que a espessura média real do wafer é diferente do valor alvo?
Para este exemplo, o teste é sobre a média μ : espessura média real do wafer, indicando em

$$H_0 : \quad \mu = 245$$

$$H_a : \quad \mu \neq 245$$

a fórmula da estatística do teste é $z = \frac{\bar{x}-245}{s/\sqrt{n}}$
Para os dados da amostra, temos que $z = \frac{246{,}18-245}{3{,}6/\sqrt{50}} = 2,32$
O valor de P em virtude do teste ser bicaudal é

$$valor \quad P = 2(1 - \Phi(2,32)) = 0,0204$$

Conclusão: H_0 não seria rejeitara para um nível de significância de $0,01$.

P-valores para testes t

O P-valor para um teste t será a área de uma curva t.

O número de graus de liberdade para o teste t de uma amostra é $n-1$.

Construindo um teste

1. especifique um teste estatístico
2. decida a forma geral da região de rejeição
3. selecione o valor crítico ou valores críticos específicos que separarão a região de rejeição da região de aceitação.

questões a serem consideradas

1. Quais são as implicações práticas de escolher um particular nível de significância, uma vez que outros aspectos do teste tem sido determinados?
2. Existe algum princípio geral que possa ser utilizado para obter o melhor ou bons testes?
3. Quando existem dois ou mais testes que são apropriados para um situação dada, como estes testes podem ser comparados para decidir finalmente o teste a ser usado?
4. Se um teste é derivado considerando suposições específicas sobre a distribuição da população que foi amostrada, quão bem servirá o teste quando as suposições são violadas?

Estatística vs significância prática

Seja cuidadoso ao interpretar evidências quando o tamanho da amostra é grande, desde que qualquer pequeno afastamento de H_0 quase com certeza será detetado pelo teste (significância estatística), ainda que tal afastamento possa ter pouca significância prática.

O princípio da razão de verossimilhança

1. Encontre o maior valor da verossimilhança para qualquer $\theta \in \Omega_0$
2. Encontre o maior valor da verossimilhança para qualquer $\theta \in \Omega_a$
3. construa a razão:

$$\lambda(x_1, ..., x_n) = \frac{\text{máxima verossimilhança para } \theta \in \Omega_0}{\text{máxima verossimilhança para } \theta \in \Omega_a}$$

Rejeite H_0 quando esta razão for pequena