Estatística

XI

M. Cristina Miranda e Anabela Rocha

Testes de hipóteses

Testes de hipóteses

HIPÓTESE ESTATÍSTICA conjuntura sobre a distribuição de uma variável aleatória.

- hipótese simples, se essa conjuntura especifica por completo o parâmetro da distribuição, *e.g.* $H_0 : \mu = 2$;
- hipótese composta, se a conjuntura englobar mais do que um valor para o parâmetro *e.g.* $H_0 : \mu > 2$;

Considera-se uma hipótese sobre o espaço do parâmetro de que depende a população, a que geralmente se chama H_0, a **hipótese nula**. Por oposição temos a **hipótese alternativa**, H_1.

Regra de decisão

Regra de Teste ou Regra de Decisão é uma regra ou um procedimento através do qual se decide se se rejeita ou não a hipótese H_0.

Seja $X_1, ..., X_n$ uma amostra aleatória de uma população $\mathcal{N}(\mu, 25)$.

Para testar $H_0 : \mu = 17 vs H_1 : \mu > 17$, uma possível regra de decisão seria:

Rejeitar H_0 , se $\bar{X} > 17 + \delta$.

A tomada de decisão vai depender dos resultados amostrais, rejeitando-se H_0, no caso de os resultados amostrais o indiciarem.

Ao fazê-lo, é possível cometer um de dois tipos de erros:

Decisão\ Realidade	H_0 verdadeira	H_0 falsa
rejeito H_0	ERRO TIPO I	decisão correcta
não rejeito H_0	decisão correcta	ERRO TIPO II

Dimensão do teste

Dimensão do teste

A tomada de decisão vai depender dos resultados amostrais, rejeitando-se H_0, no caso de os resultados amostrais o indiciarem.

A decisão é então estabelecida fixando a probabilidade de cometer o erro do tipo I.

$\alpha = P(rejH_0|H_0verd)$ a que se chama **nível de significância, dimensão ou tamanho do teste.**

Região crítica ou região de rejeição

O conjunto de valores que X pode assumir divide-se então em duas regiões:

Região crítica:

inclui todos os valores que conduzem à rejeição de H_0 - **RC**;

Região não crítica:

inclui todos os valores que não conduzem à rejeição de H_0 - $\bar{RC}$;

Então, por exemplo, para testar

$H_0 : \mu = \mu_0$ *vs* $H_1 : \mu = \mu_1$

temos

$$\alpha = P(rejH_0|H_0verd) = P((X_1, ..., X_n) \in RC|\mu = \mu_0).$$

Processos para determinar regras de decisão

1. Utilizando um intervalo de confiança (só para testes bilaterais). Por exemplo, para testar:

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu \neq \mu_0$$

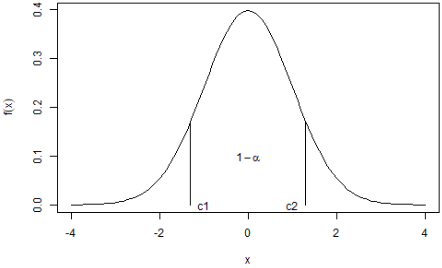
A partir de uma amostra, determina-se um intervalo de confiança (IC) para μ, com uma confiança de $(1-\alpha)x100\%$, e a regra de decisão será:

- Rejeitar H_0, com um nível de significância α, se $\mu \in IC$.

Neste caso RC resulta da reunião de dois intervalos, do tipo

$$]-\infty, c_1[\cup]c_2, \infty[$$

onde c_1e c_2 representam pontos críticos, determinados de acordo com a distribuição em causa.



2. Considerando uma estatística de teste - com comportamento diferente sob as duas hipóteses a testar (H_0 e H_1), e que permita estabelecer uma regra de decisão para esse teste.

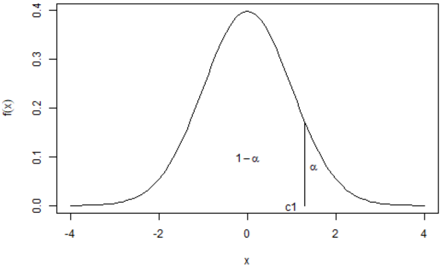
No exemplo anterior, a estatística de teste podia ser a estatística T:

$$T = \frac{\bar{X} - \mu}{S_c \sqrt{n}} \sim t_{n-1}.$$

T tem distribuição t-Student com n-1 g.l., supondo H_0 verdadeira, ou seja, $\mu = \mu_0$ de uma distribuição Normal.

A região de rejeição é determinada pela hipótese alternativa, H_1. Por exemplo, Para testar:

$H_0 : \mu = \mu_0$ vs $H_1 : \mu > \mu_0$, teríamos um teste unilateral, sendo $RC]c_1, +\infty[$.



p-value

Regra de decisão baseada no valor-p ou *p-value*

p-value ou valor de prova p

é o menor valor para o nível de significância que conduz à rejeição de H_0, para a amostra em questão.
A regra de decisão é a seguinte: "Rejeita-se H0 , ao nível de significância α, se $p < \alpha$.

p-value ou valor-p

O p-value calcula-se de acordo a hipótese alternativa H_1:

1. para o teste $H_0 : \mu = \mu_0$ vs $H_1 : \mu > \mu_0$ será $p = P(Z > z_{obs})$;
2. para o teste $H_0 : \mu = \mu_0$ vs $H_1 : \mu < \mu_0$, teremos $p = P(Z < z_{obs})$;
3. para o teste $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$, o p-value será $p = \min(P(Z > z_{obs}), P(Z < z_{obs}))$.

Testes para a média

Testes para a média em populações Normais.

Uma das distribuições mais usadas no contexto paramétrico é a distribuição Normal, que depende de dois parâmetros:

a média (μ) e a variância ($sigma^2$).

Suponhamos uma a.a. $(X_1, X_2, ..., X_n)$ e $X_i \sim N(\mu, \sigma^2)$

- No caso de variância conhecida (σ^2), a estatística Z tem distribuição Normal:

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

A estatística Z pode ser usada para testar qualquer uma das seguintes hipóteses:

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu < \mu_0,$$

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu > \mu_0,$$

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu \neq \mu_0.$$

Por exemplo, o teste

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu < \mu_0$$

tem uma região de rejeição RC dada pela regra de decisão:

rejeitar H_0 se o valor observado de Z, $z_{obs} \leq z_\alpha$,

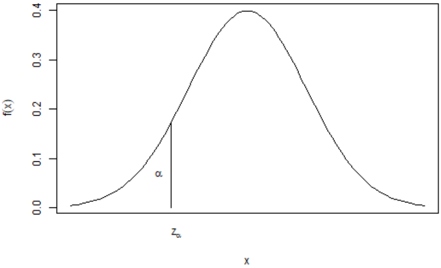
sendo z_α o quantil de uma Normal padrão.

Ou seja, $RC =]-\infty, z_\alpha[$

A regra de rejeição pode ser obtida com o p-value, comparando-o com o nível de significância escolhido para o teste:

Sendo $H_1 = \mu < \mu_0$, o *p-value*$= P(Z < z_{obs})$,

e rejeita-se H_0 se *p-Value*$< \alpha$.



- No caso da variância ser desconhecida, a estatística de teste será

$$T = \frac{\bar{X} - \mu}{S_c/\sqrt{n}} \sim t_{n-1},$$

que permite obter a região de rejeição para cada um dos testes relativos ao valor médio de uma população Normal, fixado um determinado nível de significância, α.

Exemplo

Admitindo a distribuição Normal subjacente aos dados, com um desvio padrão populacional de 5000 unidades, pretende-se testar se o valor médio dos ordenados e salários (Dados_Estat: folha consumo público) é superior a 23000 euros, com um nível de significância $\alpha = 0,04$.

X-Ordenados e salários

$X \sim N(\mu, 5000)$

$H_0 : \mu = 23000$ vs $H_1 : \mu > 23000$

Neste caso, o valor observado da estatística é:

$$z_{obs} = \frac{2.3697893 \times 10^4 - 2.3 \times 10^4}{2.5 \times 10^4 \sqrt{54}}$$

$$z_{obs} = 0.2051$$

O *p-value* relativo ao teste unilateral considerado é, tendo em atenção a hipótese alternativa, $P(Z > 0.2051)$ `r, ou seja, round(teste$p.value,4)`>0,04, logo,

não se rejeita a hipótese nula, ou seja,

o valor médio dordenados/salários não é significativamente superior a 23000 unidades monetárias.

Considerando a mesma hipótese H_0 mas admitindo σ^2 desconhecido, teríamos um valor observado da estatística igual a:

$$t_{obs} = \frac{2.3697893 \times 10^4 - 2.3 \times 10^4}{2.4741164 \times 10^4 \sqrt{54}}$$
$$t_{obs} = 0.2073$$

concluindo, de igual forma, pela não rejeição de H_0 pois o *p-value*=0.4183 é superior ao nível de significância considerado, de 0,04.

Testes para o valor médio em grandes amostras

Dispondo de uma a.a., é possível aplicar o Teorema do Limite Central e podemos testar o valor médio em populações não normais recorrendo à estatística de teste,

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

Teste para uma proporção

Teste para uma proporção (amostra de dimensão elevada)

Ou para testar o valor de uma proporção:

utilizando a média amostral como estimador para o parâmetro p - proporção de sucessos numa população, sendo $P(X_i = 1) = p$ e X uma v.a. com distribuição Bernoulli,

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