Testes de Hipóteses Paramétricos

Elementos de Probabilidades e Estatística

Sumário

1. Noções Básicas

Noções Básicas

- Nos capítulos anteriores vimos como estimar um parâmetro desconhecido a partir de dados observados (estimativas pontuais e intervalos de confiança para o parâmetro).
- Nosso objetivo agora é fazer inferências sobre afirmações relacionadas a aspetos desconhecidos de uma população, utilizando os mesmos dados observados.

Noções Básicas (cont.)

- Uma **hipótese estatística** refere-se a uma suposição sobre aspetos desconhecidos de uma população, podendo envolver:
  - Parâmetros, ou
  - A forma da distribuição.
- Quando a hipótese envolve um parâmetro e assume-se que a forma da distribuição é conhecida, diz-se que é uma **hipótese paramétrica**.
- Se o objetivo é explorar a forma da distribuição ou investigar um parâmetro sem que se tenha conhecimento da distribuição, denomina-se **hipótese não paramétrica**.
- Um **teste de hipóteses** é um procedimento estatístico que permite tomar uma decisão sobre a validade de uma hipótese formulada sobre a população.
- A decisão baseia-se na informação fornecida pelos dados de uma amostra, determinando se há evidências suficientes para suportar ou rejeitar a hipótese.

Noções Básicas (cont.)

Inicialmente, serão abordados os testes de hipóteses voltados para parâmetros de populações, incluindo:

- O valor esperado μ de uma população.
- A variância σ^2 de uma população.
- A proporção de sucessos p em uma população.

Noções Básicas (cont.)

Exemplo 1

Uma máquina de ensacar açúcar está regulada para encher sacos de 8 kg (isto é, $\mu = 8$). Será possível concluir a partir da medição do peso de um certo número de sacos que a máquina está a encher corretamente?

Noções Básicas (cont.)

Hipótese estatística

Uma hipótese estatística é uma conjetura, alegação ou afirmação sobre uma característica da população.

Teste de hipóteses

Um teste de hipóteses é um procedimento estatístico para decidir se os dados sustentam ou não uma hipótese estatística.

Tipos de Hipóteses Estatísticas

Hipótese nula H_0

A **hipótese nula**, H_0, é uma afirmação sobre o valor de um parâmetro populacional, que deve conter a condição de igualdade:

$$H_0 : \theta = \theta_0 \quad \text{ou} \quad H_0 : \theta \leq \theta_0 \quad \text{ou} \quad H_0 : \theta \geq \theta_0$$

Hipótese alternativa H_1

A hipótese alternativa, H_1, é uma afirmação que deve ser considerada se a hipótese nula é falsa. Normalmente H_1 é a negação de H_0.

$$H_1 : \theta \neq \theta_0 \quad \text{ou} \quad H_1 : \theta > \theta_0 \quad \text{ou} \quad H_1 : \theta < \theta_0$$

Testamos a hipótese nula supondo-a verdadeira ao longo do procedimento do teste e extraímos uma amostra da população para chegar a uma conclusão que nos leve a rejeitar H_0 ou não rejeitar H_0.

Ciências ULisboa

Tipos de Teste

Tipos de teste

A hipótese alternativa vai definir o tipo de teste que vamos usar:

- $H_1 : \theta \neq \theta_0$ (teste bilateral)
- $H_1 : \theta > \theta_0$ (teste unilateral à direita)
- $H_1 : \theta < \theta_0$ (teste unilateral à esquerda)

Noções Básicas (cont.)

Estrutura geral de um teste de hipóteses

1. Formular as hipóteses do teste;
2. Fixar o nível de significância;
3. Definir a Estatística de Teste e identificar a respectiva distribuição amostral;
4. Definir a Regra de Decisão;
   - 4.1 Região Crítica (ou Região de Rejeição): regra de decisão em função do valor observado da Estatística de Teste;
   - 4.2 Regra de decisão em função do *p-value* ou *valor-p*;
5. Conclusão: Tomada de decisão a partir dos valores amostrais observados.

Ciências ULisboa

Noções Básicas (cont.)

Exemplo 1 (cont.)

Uma máquina de ensacar açúcar está regulada para encher sacos de 8 kg (isto é, $\mu = 8$). Será possível concluir a partir da medição do peso de um certo número de sacos que a máquina está a encher corretamente?

- Hipótese nula

$$H_0 : \mu = \mu_0 = 8 \quad \text{(máquina regulada)}$$

- Hipótese alternativa

$$H_1 : \begin{cases} \mu \neq \mu_0 = 8 & \text{(compromisso entre ambos)} \quad \text{(teste bilateral)} \\ \mu > \mu_0 = 8 & \text{(ponto de vista do fabricante)} \quad \text{(teste unilateral à direita)} \\ \mu < \mu_0 = 8 & \text{(ponto de vista do consumidor)} \quad \text{(teste unilateral à esquerda)} \end{cases}$$

Noções Básicas (cont.)

A resposta de um teste de hipóteses pode ser expressa de duas maneiras:

- **Rejeitar H_0**: Isso significa que os dados observados fornecem evidências fortes contra a hipótese nula (H_0). Neste caso, conclui-se que a hipótese alternativa (H_1) é mais plausível.
- **Não rejeitar H_0**: Isso indica que não há evidências suficientes nos dados para rejeitar a hipótese nula (H_0). Neste caso, mantemos a hipótese H_0 por falta de evidências contrárias.

No entanto, ao fazer inferências sobre uma população com base em uma amostra, há sempre a possibilidade de cometer erros, ou seja, assumir uma conclusão errada sobre a hipótese em teste.

Realidade	**Decisão**	
	Rejeitar H_0	Não rejeitar H_0
H_0 verdadeira	Erro de tipo I	**Não há erro**
H_0 falsa	**Não há erro**	Erro de tipo II

Noções Básicas (cont.)

O **erro do tipo I** é geralmente considerado o mais grave, motivo pelo qual se atribui a ele uma baixa probabilidade de ocorrência. O **nível de significância** do teste, denotado por α, é definido como:

$$\alpha = P(\text{erro do tipo I}) = P(\text{rejeitar } H_0 | H_0 \text{ verdadeira}).$$

Comumente, α é fixado em 0.05 ou 0.01. Vale destacar que, ao reduzir a probabilidade de cometer um erro do tipo I, ou seja, rejeitando menos frequentemente H_0, há um aumento correspondente na chance de cometer um erro do tipo II.

Noções Básicas (cont.)

Estatística de Teste

Uma **estatística de teste** é uma variável aleatória, que é função da amostra aleatória, $T \equiv T(X_1, X_2, \ldots, X_n)$, que é usada para tomar a decisão acerca de rejeitar, ou não, a hipótese nula.

De modo geral, as variáveis fulcrais utilizadas na construção de intervalos de confiança tornam-se uma estatística de teste quando o parâmetro populacional em questão é substituído pelo valor hipotetizado sob a hipótese nula (H_0).

Por exemplo, se quisermos construir um intervalo de confiança a 95% para o valor médio (μ) de uma população $X \sim N(\mu, \sigma)$ com σ conhecido, usamos a seguinte estatística amostral

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

Agora, se quisermos fazer um teste de hipóteses para o valor médio μ, com um nível de significância de 5% consideramos como estatística de teste a mesma estatística amostral anterior sob a validade da hipótese nula, isto é, tomamos $\mu = \mu_0$ na expressão anterior

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \overset{sob\, H_0}{\sim} N(0, 1)$$

Região Crítica (ou Região de Rejeição) RC

A **região crítica** (ou **região de rejeição**) é o conjunto de valores da estatística de teste avaliada sob H_0, que nos levam a rejeitar a hipótese nula. Vai ser determinada com base em H_1 e fixando α:

$$P(T \in RC | H_0 \text{ verdadeira}) = \alpha$$

Decisão

Seja t_{obs} o valor observado da estatística de teste.

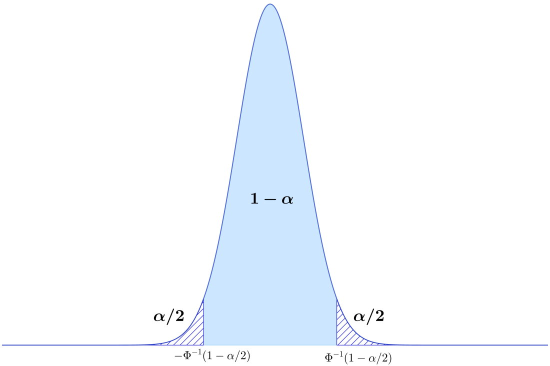
- Se $t_{obs} \in RC$ então rejeita-se H_0 ao nível de significância $\alpha \times 100\%$.
- Se $t_{obs} \notin RC$ então não se rejeita H_0 ao nível de significância $\alpha \times 100\%$, ou melhor, não há evidência estatística suficiente para rejeitar H_0.

2. Teste de hipóteses para o valor esperado, variância conhecida

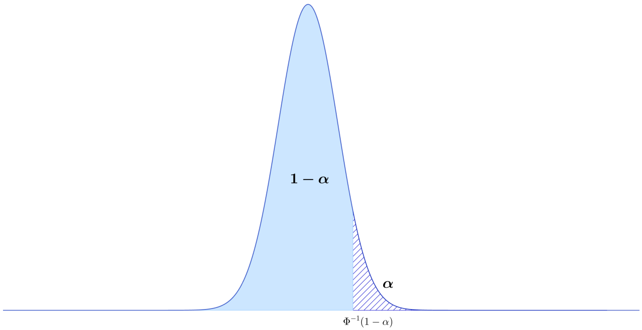
Teste de hipóteses para o **valor esperado** de uma população **Normal** com **variância CONHECIDA**

1. **Definir** H_0 **e** H_1: $H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu \neq, <, > \mu_0$
2. **Estatística de teste**:
$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \overset{sob\, H_0}{\sim} N(0,1)$$
3. **Região Crítica**:
$$H_1 : \mu \neq \mu_0 \quad RC = \{z \in \mathbb{R} : z < -c \vee z > c\}, \text{com } c = \Phi^{-1}(1 - \alpha/2)$$
$$H_1 : \mu > \mu_0 \quad RC = \{z \in \mathbb{R} : z > c\}, \text{com } c = \Phi^{-1}(1 - \alpha)$$
$$H_1 : \mu < \mu_0 \quad RC = \{z \in \mathbb{R} : z < c\}, \text{com } c = \Phi^{-1}(\alpha)$$

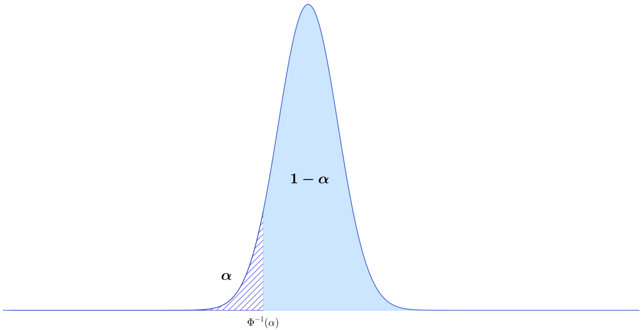
Região Crítica: Teste Bilateral



Região Crítica: Teste Unilateral à Direita



Região Crítica: Teste Unilateral à Esquerda



Exemplo

Exemplo 1 (cont.): Uma máquina de ensacar açúcar está regulada para encher sacos de açúcar de $8kg$. Para controlar o funcionamento da máquina escolheram-se ao acaso 25 sacos da produção e observou-se $\bar{x} = 8.5$ kg. Se o peso do saco é normalmente distribuído com desvio padrão de 1 kg, teste com com um nível de significância de 5% se a máquina está regulada.

Neste exemplo, vamos trabalhar com os três tipos de testes a seguir:

Hipótese nula

$$H_0 : \mu = \mu_0 = 8 \qquad \text{(máquina regulada)}$$

Hipótese alternativa

$$H_1 : \begin{cases} \mu \neq \mu_0 = 8 & \text{(compromisso entre ambos)} \quad \text{(teste bilateral)} \\ \mu < \mu_0 = 8 & \text{(ponto de vista do consumidor)} \quad \text{(teste unilateral à esquerda)} \\ \mu > \mu_0 = 8 & \text{(ponto de vista do fabricante)} \quad \text{(teste unilateral à direita)} \end{cases}$$

Ciências ULisboa

Teste bilateral

Situação

X - peso de um saco de açúcar em kg
$X \sim N(\mu, \sigma)$
$\sigma = 1$ (conhecido)
$n = 25$, $\bar{x} = 8.5$

Hipóteses

$$H_0 : \mu = \mu_0 = 8 \qquad v.s. \qquad H_1 : \mu \neq 8 \qquad \text{(teste bilateral)}$$

Nível de significância: $\alpha = 0.05$

Estatística de teste

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

Exemplo: Teste Bilateral (cont.)

Região de rejeição

$$\begin{aligned} RC &= \{z : z < -\Phi^{-1}(1-\alpha/2) \quad \text{ou} \quad z > \Phi^{-1}(1-\alpha/2)\} \\ &= \{z : z < -\Phi^{-1}(0.975) \quad \text{ou} \quad z > \Phi^{-1}(0.975)\} \\ &= \{z : z < -1.96 \quad \text{ou} \quad z > 1.96\} \\ &=]-\infty, -1.96[\, \cup \,]1.96, +\infty[\end{aligned}$$

Valor observado da estatística de teste

$$z_{obs} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{8.5 - 8}{\frac{1}{\sqrt{25}}} = 2.5 \in RC$$

Decisão: Rejeita-se H_0 ao nível de significância de $\alpha = 0.05$.

Conclusão: Há evidência estatística suficiente do peso médio ser diferente de 8kg (máquina desregulada).

Teste unilateral à esquerda

Hipóteses

$$H_0 : \mu = \mu_0 = 8 \qquad v.s. \qquad H_1 : \mu < 8 \qquad \text{(teste unilateral à esquerda)}$$

Nível de significância: $\alpha = 0.05$

Estatística de teste

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

Região de rejeição

$$\begin{aligned} RC &= \left\{z : z < \Phi^{-1}(\alpha)\right\} \\ &= \left\{z : z < \Phi^{-1}(0.05)\right\} \\ &= \{z : z < -1.645\} \\ &=]-\infty, -1.645[\end{aligned}$$

Exemplo: Teste Unilateral à Esquerda (cont.)

Valor observado da estatística de teste

$$z_{obs} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{8.5 - 8}{\frac{1}{\sqrt{25}}} = 2.5 \notin RC$$

Decisão: Não se rejeita H_0 ao nível de significância de $\alpha = 0.05$.

Conclusão: Não há evidência estatística suficiente suportando que o peso médio é inferior a 8kg.

Exemplo: Teste Unilateral à Direita

Teste unilateral à direita

Hipóteses

$$H_0 : \mu = \mu_0 = 8 \qquad v.s. \qquad H_1 : \mu > 8 \qquad \text{(teste unilateral à direita)}$$

Nível de significância: $\alpha = 0.05$

Estatística de teste

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

Região crítica

$$\begin{aligned} RC &= \left\{z : z > \Phi^{-1}(1-\alpha)\right\} \\ &= \left\{z : z > \Phi^{-1}(0.95)\right\} \\ &= \{z : z > 1.645\} \\ &=]1.645, +\infty[\end{aligned}$$

Exemplo: Teste Unilateral à Direita (cont.)

Valor observado da estatística de teste

$$z_{obs} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{8.5 - 8}{\frac{1}{\sqrt{25}}} = 2.5 \in RC$$

Decisão: Rejeita-se H_0 ao nível de significância de $\alpha = 0.05$.

Conclusão: Há evidência estatística suficiente do peso médio ser superior a 8kg (máquina desregulada).

Ciências ULisboa

3. Teste de hipóteses para o valor esperado, variância desconhecida

Teste de hipóteses para o **valor esperado** de uma população **Normal** com **variância DESCONHECIDA**

1. **Definir H_0 e H_1**: $H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu \neq, <, > \mu_0$
2. **Estatística de teste**:

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \overset{sob\ H_0}{\sim} t_{(n-1)}$$

3. **Região Crítica**:

$$H_1 : \mu \neq \mu_0 \quad RC = \{t \in \mathbb{R} : t < -c \vee t > c\}, \text{ com } c = F^{-1}_{t_{(n-1)}}(1 - \alpha/2)$$

$$H_1 : \mu > \mu_0 \quad RC = \{t \in \mathbb{R} : t > c\}, \text{ com } c = F^{-1}_{t_{(n-1)}}(1 - \alpha)$$

$$H_1 : \mu < \mu_0 \quad RC = \{t \in \mathbb{R} : t < c\}, \text{ com } c = F^{-1}_{t_{(n-1)}}(\alpha)$$

Ciências ULisboa

Exemplo

Exemplo 2: Numa investigação experimental em Física dos Materiais, está a ser estudada a densidade média de uma liga metálica utilizada em dispositivos de alta precisão. Com base em especificações industriais de ligas similares, acredita-se que a densidade média do material seja 2.70 g/cm^3. Para verificar esta especificação, um grupo de físicos experimentais recolheu uma amostra de 10 medições da densidade da liga metálica (em g/cm^3):

$$2.65,\ 2.72,\ 2.70,\ 2.68,\ 2.71,\ 2.73,\ 2.69,\ 2.74,\ 2.67,\ 2.68$$

Com base na amostra fornecida, é possível concluir, com um nível de significância de 5%, que a densidade média da liga metálica difere de 2.70 g/cm^3?
Assuma que a densidade da liga é normalmente distribuída.

Exemplo

Situação

X - densidade de uma liga metálica em g/cm³
$X \sim N(\mu, \sigma)$
μ - desconhecido
σ - desconhecido
$n = 10 < 30 \quad , \bar{x} = 2.697 \quad s = 0.028$

Hipóteses

$$H_0 : \mu = \mu_0 = 2.70 \qquad v.s. \qquad H_1 : \mu \neq 2.70$$ (teste bilateral)

Nível de significância: $\alpha = 0.05$

Ciências ULisboa

Exemplo (cont.)

Estatística de teste

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

Região de rejeição

$$\begin{aligned} RC &= \left\{ t : t < -F^{-1}_{t_{(n-1)}}(1-\alpha/2) \quad \text{ou} \quad t > F^{-1}_{t_{(n-1)}}(1-\alpha/2) \right\} \\ &= \left\{ t : t < -F^{-1}_{t_{(9)}}(0.975) \quad \text{ou} \quad t > F^{-1}_{t_{(9)}}(0.975) \right\} \\ &= \{ t : t < -2.262 \quad \text{ou} \quad t > 2.262 \} \\ &= \,]-\infty, -2.262[\, \cup \,]2.262, +\infty[\end{aligned}$$

Ciências ULisboa

Exemplo (cont.)

Valor observado da estatística de teste

$$t_{obs} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{2.697 - 2.70}{\frac{0.028}{\sqrt{10}}} = -0.3388 \notin RC$$

Decisão: Como $t_{obs} \notin RC$, não se deve rejeitar H_0 ao nível de significância de $\alpha = 0.05$.

Conclusão: Não há evidência estatística suficiente suportando que a densidade média da liga metálica seja diferente de 2.70 g/cm³.

Exemplo (cont.)

No R:

```
> amostra <- c(2.65, 2.72, 2.70, 2.68, 2.71, 2.73, 2.69, 2.74, 2.67, 2.68)
> t.test(amostra, mu = 2.70, alternative = "two.sided")

 One Sample t-test

data:  x
t = -0.33518, df = 9, p-value = 0.7452
alternative hypothesis: true mean is not equal to 2.7
95 percent confidence interval:
 2.676753 2.717247
sample estimates:
mean of x
    2.697
```

4. Teste de hipóteses para a variância de uma população normal

Teste de hipóteses para a **variância** de uma população normal

1. **Definir** H_0 **e** H_1: $H_0 : \sigma^2 = \sigma_0^2 \quad vs \quad H_1 : \sigma^2 \neq, <, > \sigma_0^2$
2. **Estatística de teste**:
$$Q = \frac{(n-1)S^2}{\sigma_0^2} \overset{sob\ H_0}{\sim} \mathcal{X}^2_{(n-1)}$$
3. **Região Crítica**:
$$H_1 : \sigma^2 \neq \sigma_0^2 \quad RC = \{q \in \mathbb{R}^+ : q < a \vee q > b\}, \text{ com } a = F^{-1}_{\mathcal{X}^2_{(n-1)}}(\alpha/2)$$
$$\text{e } b = F^{-1}_{\mathcal{X}^2_{(n-1)}}(1-\alpha/2)$$
$$H_1 : \sigma^2 > \sigma_0^2 \quad RC = \{q \in \mathbb{R}^+ : q > c\}, \text{ com } c = F^{-1}_{\mathcal{X}^2_{(n-1)}}(1-\alpha)$$
$$H_1 : \sigma^2 < \sigma_0^2 \quad RC = \{q \in \mathbb{R}^+ : q < c\}, \text{ com } c = F^{-1}_{\mathcal{X}^2_{(n-1)}}(\alpha)$$

Ciências ULisboa

Exemplo

Exemplo 3: Um engenheiro biomédico deseja verificar se a variância no diâmetro de partículas administradas por um inalador médico é diferente de um valor padrão conhecido. Estudos anteriores indicam que a variância ideal do diâmetro das partículas emitidas pelo dispositivo é 0.15 mm^2, de modo a garantir uma distribuição eficiente no trato respiratório.
Para testar a estabilidade de um novo modelo de inalador, o engenheiro recolhe uma amostra com 15 medições do diâmetro de partículas (em mm), medidas em diferentes momentos de funcionamento do aparelho:

0.90, 1.10, 0.95, 0.85, 0.98, 1.12, 0.89, 0.94, 1.02, 0.96, 1.05, 0.87, 1.08, 0.93, 1.00

Com base nestas medições, pode-se concluir, ao nível de significância de 1%, que a variância do diâmetro das partículas difere do valor padrão de 0.15 mm^2?
Assuma que o diâmetro das partículas segue uma distribuição normal.

Exemplo (cont.)

Situação

X - diâmetro das partículas, em mm.
$X \sim N(\mu, \sigma)$
μ - desconhecido
σ - desconhecido
$n = 15$, $\bar{x} = 0.976$, $s^2 = 0.0071$

Hipóteses

$$H_0 : \sigma^2 = \sigma_0^2 = 0.15 \qquad v.s. \qquad H_1 : \sigma^2 \neq 0.15 \qquad \text{(teste bilateral)}$$

Nível de significância: $\alpha = 0.01$

Ciências ULisboa

Exemplo (cont.)

Estatística de teste

$$Q = \frac{(n-1)S^2}{\sigma_0^2} \overset{sob\ H_0}{\sim} \mathcal{X}^2_{(n-1)}$$

Região de rejeição

$$\begin{aligned} RC &= \left\{ q : q < F^{-1}_{\mathcal{X}^2_{(n-1)}}(\alpha/2) \quad \text{ou} \quad q > F^{-1}_{\mathcal{X}^2_{(n-1)}}(1-\alpha/2) \right\} \\ &= \left\{ q : q < F^{-1}_{\mathcal{X}^2_{(14)}}(0.005) \quad \text{ou} \quad q > F^{-1}_{\mathcal{X}^2_{(14)}}(0.995) \right\} \\ &= \{ q : q < 4.074675 \quad \text{ou} \quad q > 31.31935 \} \\ &=]0, 4.074675[\cup]31.31935, +\infty[\end{aligned}$$

Ciências ULisboa

Exemplo (cont.)

Valor observado da estatística de teste

$$q_{obs} = \frac{(n-1)s^2}{\sigma_0^2} = \frac{14 \times 0.0071}{0.15} = 0.6626 \in \mathrm{RC}$$

Decisão: Como $q_{obs} \in RC$, deve-se rejeitar H_0 ao nível de significância de $\alpha = 0.01$.

Conclusão: Há evidência estatística suficiente suportando que a variância do diâmetro das partículas é diferente de 0.15mm².

Ciências ULisboa

5. Teste de hipóteses para uma proporção (probabilidade de sucesso)

Teste de hipóteses para uma proporção (probabilidade de sucesso)

1. **Definir** H_0 **e** H_1: $H_0 : p = p_0 \quad vs \quad H_1 : p \neq, <, > p_0$
2. **Estatística de teste (via TLC)**:

$$Z = \frac{\bar{X} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \overset{a}{\sim} N(0,1)$$

3. **Região Crítica**:

$$H_1 : p \neq p_0 \quad RC \approx \{z \in \mathbb{R} : z < -c \vee z > c\}, \text{com } c = \Phi^{-1}(1-\alpha/2)$$
$$H_1 : p > p_0 \quad RC \approx \{z \in \mathbb{R} : z > c\}, \text{com } c = \Phi^{-1}(1-\alpha)$$
$$H_1 : p < p_0 \quad RC \approx \{z \in \mathbb{R} : z < c\}, \text{com } c = \Phi^{-1}(\alpha)$$

Ciências ULisboa

Exemplo

Exemplo 4: Num inquérito nacional para avaliar o apoio a uma nova política de saúde proposta pelo governo, foram entrevistados 600 eleitores. Destes, 330 mostraram-se favoráveis à política. Analistas políticos sugerem que, em geral, mais de 50% da população apoia políticas semelhantes. Determine se existe evidência estatística que suporte a suposição existente. Considere um nível de significância de 1%.

Exemplo (cont.)

Situação

X - eleitor apoia a nova política de saúde
$X \sim \text{Bernoulli}(p)$
p - proporção de eleitores que apoiam a nova política de saúde
$n = 600$, $x = 330$, $\bar{x} \equiv \hat{p} = 330/600 = 0.55$

Hipóteses

$H_0 : p \leq p_0 = 0.5 \quad vs \quad H_1 : p > 0.5$ (teste unilateral à direita)

Nível de significância
$\alpha = 0.01$

Exemplo (cont.)

Estatística de teste

$$Z = \frac{\bar{X} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \overset{a}{\sim} N(0,1)$$

Região de rejeição

$$\begin{aligned} \text{RC} &= \{z \in \mathbb{R} : z > \Phi^{-1}(1-\alpha)\} \\ &= \{z \in \mathbb{R} : z > \Phi^{-1}(0.99)\} \\ &= \{z \in \mathbb{R} : z > 2.326\} \\ &=]2.326, +\infty[\end{aligned}$$

Valor observado da estatística de teste

$$z_{obs} = \frac{\bar{x} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.55 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{600}}} = \frac{0.05}{0.02041} = 2.449 \in \text{RC}$$

Exemplo (cont.)

Decisão

Como $z_{obs} \in$ RC então deve-se rejeitar H_0 ao nível de significância de 1%.

Conclusão

Há evidência estatística suficiente suportando que a proporção de eleitores que apoiam a nova política é superior a 50%.

Ciências ULisboa

Exemplo (cont.)

No R:

```
> prop.test(x=330,n=600,p=0.5,alternative="greater",conf.level=0.99)

 1-sample proportions test without continuity correction

data:  330 out of 600, null probability 0.5
X-squared = 6, df = 1, p-value = 0.007153
alternative hypothesis: true p is greater than 0.5
99 percent confidence interval:
 0.5025142 1.0000000
sample estimates:
   p
0.55
```

Relação entre intervalos de confiança e testes de hipóteses

Relação entre intervalos de confiança e testes de hipóteses - apenas testes bilaterais

1. Identifique o parâmetro desconhecido θ.
2. Formule as hipóteses
$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta \neq \theta_0$$
3. Dada uma amostra $(x_1, \ldots, x_n)$, construa um IC a $(1-\alpha) \times 100\%$ para θ.
4.
$$\begin{aligned} &\text{Se } \theta_0 \in IC_{(1-\alpha)\times 100\%}(\theta) \Leftrightarrow \text{Não se rejeita } H_0 \text{ ao n.s. } \alpha. \\ &\text{Se } \theta_0 \notin IC_{(1-\alpha)\times 100\%}(\theta) \Leftrightarrow \text{Rejeita-se } H_0 \text{ ao n.s. } \alpha. \end{aligned}$$

- O IC fornece uma região de não rejeição de H_0.
- Variável fulcral e estatística de teste precisam ser da mesma forma.

6. Valor-p

Valor-p

Em vez de fixarmos o nível de significância do teste, de identificarmos a região de rejeição de e verificar se o valor observado da estatística de teste pertence ou não a tal região, podemos olhar diretamente para o valor observado da estatística de teste e determinar para que nível de significância a decisão muda.

Valor-p

Dado o valor observado da estatística de teste, o **valor-p** (p-value) é o menor nível de significância que levaria à rejeição da hipótese nula. Decisão:

- Rejeitar H_0 para valores $\alpha \geq$ valor-p;
- Não rejeitar H_0 para valores $\alpha <$ valor-p.

Cálculo do Valor-p

Seja T uma estatística de teste de interesse e t_{obs} o valor observado desta estatística de teste.

Cálculo do valor-p

- Teste unilateral à esquerda: valor-p $= P(T < t_{obs})$
- Teste unilateral à direita: valor-p $= P(T > t_{obs})$
- Teste bilateral: valor-p $= 2\min\{P(T < t_{obs}), P(T > t_{obs})\}$

Ciências ULisboa

Teste de hipóteses com base no valor-p

Estrutura geral de um teste com base no valor-p

1. Formular as hipóteses do teste;
2. Definir a Estatística de Teste;
3. Calcular o valor observado da estatística de teste;
4. Calcular o valor-p do teste;
5. Rejeitar H_0 para valores de $\alpha \geq$ valor-p; Não rejeitar H_0 para valores de $\alpha <$ valor-p.

Ciências ULisboa

Exemplo

Exemplo 5: Uma máquina de ensacar açúcar está regulada para encher sacos de açúcar de $8kg$. Para controlar o funcionamento da máquina escolheram-se ao acaso 25 sacos da produção e observou-se $\bar{x} = 8.5$ kg e $s = 1$ kg. Se o peso do saco é normalmente distribuído teste com um nível de significância de 5% se a máquina está regulada. Tome a decisão com base no valor-p.

Exemplo (cont.)

Hipóteses: $H_0 : \mu = \mu_0 = 8$ vs $H_1 : \mu \neq 8$ (teste bilateral)

Estatística de teste: $T = \frac{\bar{X}-8}{1/\sqrt{25}} \overset{sob\, H_0}{\sim} T_{(24)}$

Valor observado da estatística de teste:

$$t_{obs} = \frac{8.5 - 8}{1/\sqrt{25}} = 2.5$$

Cálculo do valor-p: Como o teste é bilateral,

$$\begin{aligned} valor - p &= 2\min\{P(T < 2.5), P(T > 2.5)\} \\ &= 2P(T > 2.5) \quad \text{(simetria da T-Student)} \\ &= 2(1 - F_{T_{(24)}}(2.5)) \\ &= 0.019654 \end{aligned}$$

Exemplo (cont.)

Decisão: Rejeita-se H_0 para $\alpha \geq$ valor-p $= 0.019654$; Não se rejeita H_0 para $\alpha <$ valor-p $= 0.019654$.
Assim, a hipótese nula H_0 é rejeitada para os níveis de significância usuais de 5% e 10% mas não é rejeitada para 1%.

No R:

```
> valor_p <- 2*(1-pt(2.5,24))
> valor_p
[1] 0.01965418
```

Teste unilateral à esquerda

Hipóteses: $H_0 : \mu \geq \mu_0 = 8$ vs $H_1 : \mu < 8$ (teste unilateral à esquerda)

Cálculo do valor-p: Como o teste é unilateral à esquerda,

$$\begin{aligned} valor - p &= P(T < 2.5) \\ &= 0.9901729 \end{aligned}$$

Decisão: Como $\alpha = 0.05 < 0.9901729 = valor - p$, não se deve rejeitar H_0 ao n.s. de 5%.

Teste unilateral à direita

Hipóteses: $H_0 : \mu \leq \mu_0 = 8$ vs $H_1 : \mu > 8$ (teste unilateral à direita)

Cálculo do valor-p: Como o teste é unilateral à direita,

$$\begin{aligned} valor - p &= P(T > 2.5) \\ &= 0.009827088 \end{aligned}$$

Decisão: Como $\alpha = 0.05 > 0.009827088 = valor - p$, rejeita-se H_0 ao n.s. de 5%.

Exemplo

Exemplo 6: Em uma mina de carvão, a granulometria das partículas é uma característica importante que afeta diretamente a qualidade e o processamento do carvão. Um engenheiro de minas deseja testar se a variância do tamanho das partículas de carvão em uma nova área de mineração é diferente de uma variância padrão conhecida de 0.02 mm², que foi determinada em estudos anteriores na mesma região. A amostra coletada de 20 partículas de carvão da nova área apresenta os seguintes tamanhos (em mm):

$$2.1, 2.3, 2.2, 2.5, 2.1, 2.4, 2.6, 2.2, 2.3, 2.7, 2.4, 2.1, 2.5, 2.3, 2.6, 2.2, 2.4, 2.3, 2.1, 2.5$$

Será que a variância do tamanho das partículas de carvão da nova área de mineração é significativamente diferente da variância padrão de 0.02 mm²? Assuma que o tamanho das partículas segue uma distribuição normal. Use um nível de significância de 10%.

Exemplo (cont.)

Situação

$X-$ tamanho da partícula de carvão em mm
$X \sim N(\mu, \sigma)$
$\mu-$ desconhecido
σ^2- desconhecida
$n = 20$, $s^2 = 0.03410526$

Hipóteses: $H_0 : \sigma^2 = \sigma_0^2 = 0.02$ vs $H_1 : \sigma^2 \neq 0.02$ (teste bilateral)

Nível de significância: $\alpha = 0.10$

Estatística de teste: $Q = \frac{(n-1)S^2}{\sigma_0^2} \overset{sob\ H_0}{\sim} \mathcal{X}^2_{(n-1)}$

Valor observado da estatística de teste: Note que $s^2 = 0.03410526$, assim

$$q_{obs} = \frac{19 \times 0.03410526}{0.02} = 32.4$$

Exemplo (cont.)

Cálculo do valor-p: Como o teste é bilateral,

$$\begin{aligned} valor - p &= 2\min\{P(Q < 32.4), P(Q > 32.4)\} \\ &= 2\min\{0.9717649, 0.0282351\} \\ &= 2 \times (0.0282351) \\ &= 0.0564702 \end{aligned}$$

Decisão: Como $\alpha = 0.1 > 0.0564 = valor - p$, rejeita-se H_0 ao n.s. de 10%.

7. Teste de hipóteses para dois parâmetros populacionais

Testes de hipóteses para um parâmetro e dois parâmetros (amostras independentes)

Parâmetros a estimar	σ^2 conhecido?	Tipo de população	Dimensão da amostra	Estatística de teste e correspondente distribuição amostral
μ	Sim	Normal	Qualquer	$Z_0 = \frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0;1)$
		Outra	$n > 30$	$Z_0 = \frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} \dot\sim N(0;1)$
	Não	Qualquer	$n > 30$	$Z_0 = \frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}} \dot\sim N(0;1)$
		Normal	Qualquer	$T_0 = \frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$
p	- - -	Bernoulli	$n > 30$	$Z_0 = \frac{\hat{P}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \dot\sim N(0;1)$
σ^2	- - -	Normal	Qualquer	$Q_0 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$
$\mu_1 - \mu_2$	σ_1^2 e σ_2^2 conhecidas	Normais	Quaisquer	$Z_0 = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0;1)$
		Outras	$n_1 > 30$ e $n_2 > 30$	$Z_0 = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \dot\sim N(0;1)$
	σ_1^2 e σ_2^2 desconhecidas	Quaisquer	$n_1 > 30$ e $n_2 > 30$	$Z_0 = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \dot\sim N(0;1)$
	σ_1^2 e σ_2^2 desconhecidas $(\sigma_1^2 = \sigma_2^2)$	Normais	Quaisquer	$T_0 = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}$
	σ_1^2 e σ_2^2 desconhecidas $(\sigma_1^2 \neq \sigma_2^2)$	Normais	Quaisquer	$T_0 = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \sim t_r$ sendo r o número natural mais próximo de $r^* = \frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{S_1^2}{n_1}\right)^2+\frac{1}{n_2-1}\left(\frac{S_2^2}{n_2}\right)^2}$
$p_1 - p_2$	- - -	Bernoulli	$n_1 > 30$ e $n_2 > 30$	$Z_0 = \frac{(\hat{P}_1-\hat{P}_2)-(p_1-p_2)_0}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1}+\frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}} \dot\sim N(0;1)$

Exemplo

Exemplo 7: Uma equipa de físicos está a estudar a concentração média de impurezas metálicas em duas amostras de silício provenientes de diferentes laboratórios de fabricação de semicondutores, denominadas regiões A e B. Estas impurezas podem afetar significativamente o desempenho de dispositivos eletrónicos sensíveis.
Foram recolhidas amostras de silício em ambas as regiões, e a concentração de impurezas (em %) foi medida. O objetivo é determinar se existe uma diferença significativa entre as concentrações médias das duas regiões. Os dados recolhidos foram os seguintes:

- **Região A**: Amostra de 25 medições com média $\bar{x}_A = 3.8\%$ e desvio padrão $s_A = 0.6\%$.
- **Região B**: Amostra de 30 medições com média $\bar{x}_B = 4.1\%$ e desvio padrão $s_B = 0.8\%$.

Assuma que as populações seguem distribuições normais e que as variâncias podem ser consideradas iguais. Use um nível de significância de 5% e $P(T_{53} < -1.54) = 0.06434$.

Exemplo (cont.)

Situação

X_A- concentração de impurezas metálicas na região A.
X_B- concentração de impurezas metálicas na região B.
$X_A \sim N(\mu_A, \sigma_A)$
$X_B \sim N(\mu_B, \sigma_B)$
$\mu_A - \mu_B$ (desconhecido)
$\sigma_A^2 = \sigma_B^2$ (desconhecidas e iguais)

Hipóteses

$$H_0 : \mu_A = \mu_B \quad \Leftrightarrow \quad H_0 : \mu_A - \mu_B = (\mu_A - \mu_B)_0 = 0$$
$$H_1 : \mu_A \neq \mu_B \quad \Leftrightarrow \quad H_1 : \mu_A - \mu_B \neq 0 \text{ (teste bilateral)}$$

Exemplo (cont.)

Nível de significância: $\alpha = 0.05$

Estatística de teste: $T = \dfrac{(\bar{X}_A - \bar{X}_B) - (\mu_A - \mu_B)_0}{\sqrt{\frac{(n_A-1)S_A^2 + (n_B-1)S_B^2}{n_A+n_B-2}\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}} \sim t_{n_A+n_B-2}$

Valor observado da estatística de teste:

$$t_{obs} = \frac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)_0}{\sqrt{\frac{(n_A-1)s_A^2 + (n_B-1)s_B^2}{n_A+n_B-2}\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}} = \frac{(3.8-4.1)-0}{\sqrt{\frac{(25-1)(0.6)^2+(30-1)(0.8)^2}{25+30-2}\left(\frac{1}{25}+\frac{1}{30}\right)}} = -1.5464$$

Cálculo do valor-p: Como o teste é bilateral,

$$\begin{aligned} valor - p &= 2\min\{P(T < -1.5464), P(T > -1.5464)\} \\ &= 2\min\{0.06434, 0.9356\} \\ &= 2 \times (0.06434) \\ &= 0.12868 \end{aligned}$$

Exemplo (cont.)

Decisão: Como $\alpha = 0.05 < 0.12868 =$ valor-p, então não se deve rejeitar H_0 ao nível de significância de 5%.

Conclusão: Não há evidência estatística suficiente suportando que as concentrações médias de impurezas metálicas nas regiões são diferentes.

Exemplo

Exemplo 8: Uma equipa de engenheiros biomédicos está a estudar a eficácia de dois tipos de sensores implantáveis utilizados para deteção de sinais elétricos anómalos no tecido cardíaco. Foram conduzidos testes independentes em dois grupos distintos de pacientes, utilizando sensores do Tipo A e do Tipo B. O objetivo é verificar se existe diferença significativa entre os dois tipos de sensores quanto à proporção de deteção de sinais anómalos.

- **Sensor A**: Foram testados 150 pacientes, dos quais 45 apresentaram deteção positiva de sinais anómalos.
- **Sensor B**: Foram testados 180 pacientes, dos quais 30 apresentaram deteção positiva de sinais anómalos.

Pretende-se saber, ao nível de significância de 1%, se a proporção de deteção positiva difere significativamente entre os dois sensores.

Exemplo (cont.)

Situação

X_A- sinal elétrico anómalo no sensor A
X_B- sinal elétrico anómalo no sensor B
$X_A \sim \text{Bernoulli}(p_A)$
$X_B \sim \text{Bernoulli}(p_B)$
$p_A - p_B$ (desconhecido)
$n_A = 150, x_A = 45, \hat{p}_A = \bar{x}_A = \frac{x_A}{n_A} = \frac{45}{150} = 0.3$
$n_B = 180, x_B = 30, \hat{p}_B = \bar{x}_B = \frac{x_B}{n_B} = \frac{30}{180} = 0.167$

Hipóteses

$$H_0 : p_A = p_B \quad \Leftrightarrow \quad H_0 : p_A - p_B = (p_A - p_B)_0 = 0$$
$$H_1 : p_A \neq p_B \quad \Leftrightarrow \quad H_1 : p_A - p_B \neq 0 \text{ (teste bilateral)}$$

Exemplo (cont.)

Nível de significância: $\alpha = 0.01$

Estatística de teste

$$Z = \frac{\left(\hat{P}_A - \hat{P}_B\right) - (p_A - p_B)_0}{\sqrt{\frac{\hat{P}_A(1-\hat{P}_A)}{n_A} + \frac{\hat{P}_B(1-\hat{P}_B)}{n_B}}} \overset{a}{\sim} N(0,1)$$

Valor observado da estatística de teste

$$z_{obs} = \frac{(\hat{p}_A - \hat{p}_B) - (p_A - p_B)_0}{\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}}} = \frac{(0.3 - 0.167) - 0}{\sqrt{\frac{0.3(1-0.3)}{150} + \frac{0.167(1-0.167)}{180}}} = 2.85$$

Exemplo (cont.)

Cálculo do valor-p: Como o teste é bilateral, temos

$$\begin{aligned}\text{valor-p} &= 2\min\{P(Z > z_{obs}), P(Z < z_{obs})\}\\ &= 2\min\{0.0022, 0.9978\}\\ &= 2 \times 0.0022\\ &= 0.0044\end{aligned}$$

Decisão: Como $\alpha = 0.01 > 0.0044 =$valor-p, então deve-se rejeitar H_0 ao nível de significância de 1%.

Conclusão: Há evidência estatística suficiente suportando que as proporções são diferentes.

Referências

- Manuel Cabral Morais (2020): Probabilidade e Estatística: Teoria, Exemplos & Exercícios. IST Press, 1a edição.