Teste t (de Student)

- Teste não-viesado;
- O teste t (unilateral e bilateral);
- Teste t pareado;
- Teste t para duas amostras;
  - ◇ Variâncias iguais (homogeneidade);
  - ◇ Variâncias proporcionais.
- Propriedades e exemplos;

Em analogia com o conceito de estimador não-viesado, podemos também classificar testes de hipótese em viesados ou não-viesados.

Definição 51 (Teste não viesado)

Suponha que desejamos testar a hipótese

$$H_0 : \theta \in \Omega_0,$$
$$H_1 : \theta \in \Omega_1.$$

através do teste δ. Dizemos que δ é ***não-viesado*** *se (e somente se) para $\theta \in \Omega_0$ e $\theta' \in \Omega_1$, vale*

$$\pi(\theta \mid \delta) \leq \pi(\theta' \mid \delta),$$

ou seja, se a função poder é pelo menos tão grande no espaço onde H_0 é falsa (Ω_1) quanto no espaço em que H_0 é verdadeira (Ω_0).

Teste t: motivação

Suponha que estamos interessados em testar hipóteses sobre a média (μ) de uma distribuição Normal quando a variância (σ^2) é desconhecida. Sabemos que é possível encontrar uma quantidade pivotal $(\bar{X}_n - \mu)/\hat{\sigma}'$ tal que é possível construir um intervalo de confiança para μ da forma $\bar{X}_n - c\hat{\sigma}'/\sqrt{n}, \bar{X}_n + c\hat{\sigma}'/\sqrt{n}$, onde $c = T^{-1}(\gamma; n-1)$ é a f.d.a. inversa de uma distribuição t de Student com $n-1$ graus de liberdade.

Pergunta 5 (Como falar sobre μ quando ambas μ e σ^2 são desconhecidas?)

Suponha que Palmirinha esteja interessada em testar a hipótese de que a média da concentração de amido na sua pamonha seja maior que $\mu_0 = 7mg/L$, ou seja,

$$H_0 : \mu \geq \mu_0,$$
$$H_1 : \mu < \mu_0,$$

mas ela desconhece a variância do processo, σ^2. Como testar hipóteses sobre $\theta = (\mu, \sigma^2)$?

Exemplo

Palmirinha pode começar computando a estatística

$$U := \sqrt{n}\frac{\bar{X}_n - \mu_0}{\hat{\sigma}'},$$

onde $\hat{\sigma}' = \sqrt{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2/(n-1)}$ como já estudado. A partir daí, pode construir um teste δ_c que rejeita H_0 se $U \leq c$, para uma constante real c definida.

Observação 25 (A distribuição de U sob H_0)

Quando H_0 é verdadeira, em particular, quando $\mu = \mu_0$, U tem distribuição t de Student com $n-1$ graus de liberdade, não importando o valor de σ^2. Isto é, U é pivotal.

Definição 52 (Teste t)

Um teste δ_c que rejeita H_0 se $U \geq c$ (equiv. $U \leq c$), com $c = T^{-1}(1 - \alpha_0; n-1)$ é chamado um teste t (unicaudal) de tamanho α_0.

Teorema 30 (Propriedades do teste t)

Suponha que δ_c rejeita H_0 se $U \geq c$. Então

i) $\mu = \mu_0 \implies \pi(\mu, \sigma^2 \mid \delta_c) = \alpha_0$

ii) $\mu < \mu_0 \implies \pi(\mu, \sigma^2 \mid \delta_c) < \alpha_0$

iii) $\mu > \mu_0 \implies \pi(\mu, \sigma^2 \mid \delta_c) > \alpha_0$

iv) $\lim_{\mu \to -\infty} \pi(\mu, \sigma^2 \mid \delta_c) = 0$;

v) $\lim_{\mu \to \infty} \pi(\mu, \sigma^2 \mid \delta_c) = 1$;

vi) *δ_c é não-viesado e tem tamanho α_0.*

Prova: Ver Teorema 9.5.1. de DeGroot.

P-valor para o teste t

Lembre-se de que o p-valor é a probabilidade, **sob** H_0, de observarmos uma estatística tão ou mais extrema do que a que foi observada.

Teorema 31 (P-valor para um teste t unicaudal)

Suponha que observarmos $U = u$ *e seja* $T(\cdot; n-1)$ *a d.f.a. de uma distribuição t de Student com* $n-1$ *graus de liberdade. Para a hipótese*

$$H_0 : \mu \geq \mu_0,$$
$$H_1 : \mu < \mu_0,$$

o p-valor vale $T(u; n-1)$, *enquanto para a hipótese*

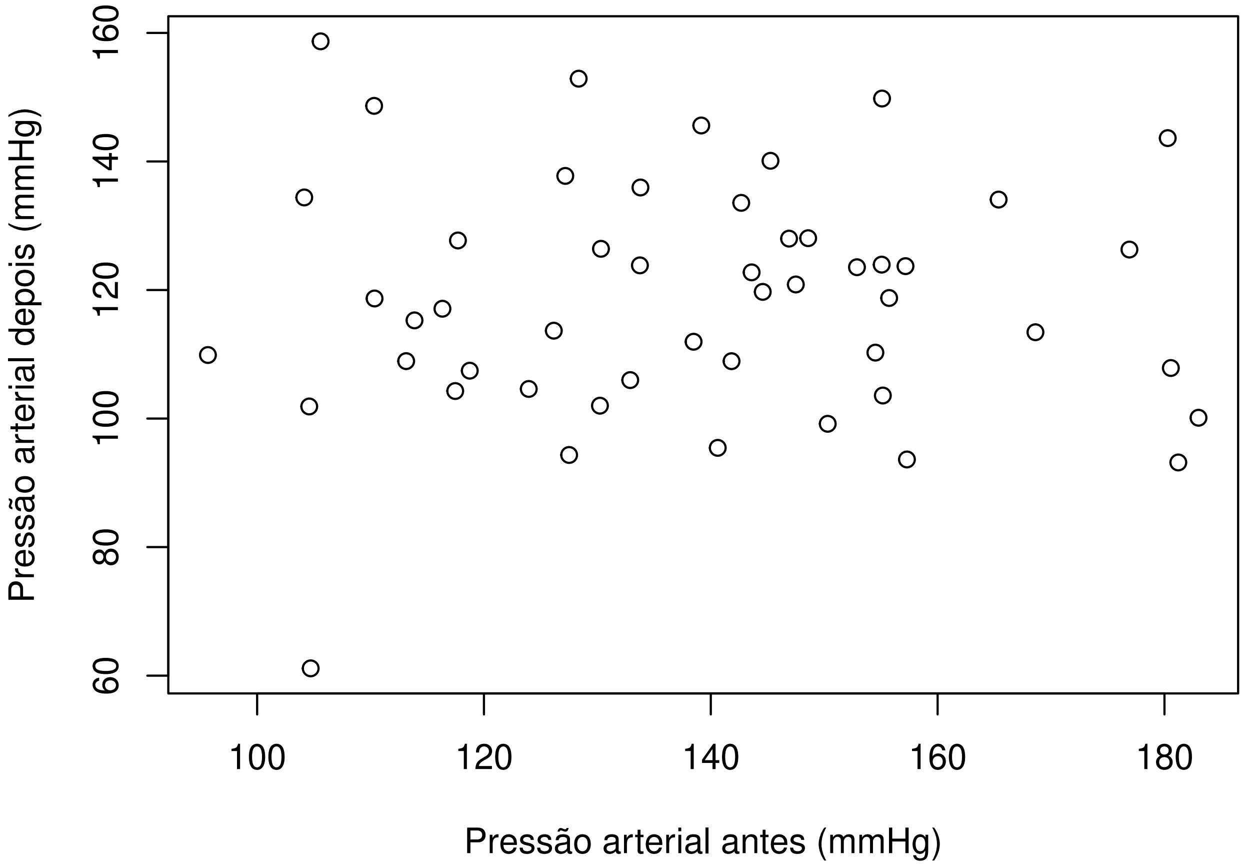
$$H_0 : \mu \leq \mu_0,$$
$$H_1 : \mu > \mu_0,$$

o p-valor vale $1 - T(u; n-1)$.

Prova: Notar que δ_c depende de $c = T^{-1}(1-\alpha_0; n-1)$. Ver Teorema 9.5.2 em DeGroot.

Teste t pareado: motivação

Suponha que estamos interessados em medir o efeito de uma droga sobre a pressão arterial sistólica de um grupo de pacientes. Suponha que medimos as pressões arteriais de n pacientes antes ($\boldsymbol{X}$) e depois ($\boldsymbol{Y}$) de administrar a droga. Vamos supor que $X_i \sim \text{Normal}(\mu_{\textbf{antes}}, \sigma^2)$ e $Y_i \sim \text{Normal}(\mu_{\textbf{depois}}, \sigma^2)$.



Teste t pareado: execução

Estamos, portanto, interessados na hipótese[16]

$$
\begin{aligned}
H_0 &: \mu_{\textbf{antes}} \leq \mu_{\textbf{depois}},\\
H_1 &: \mu_{\textbf{antes}} > \mu_{\textbf{depois}}.
\end{aligned}
$$

Podemos modelar a variável $Z_i = X_i - Y_i$ e sabemos que $Z_i \sim \text{Normal}(\mu_Z = \mu_{\textbf{antes}} - \mu_{\textbf{depois}}, 2\sigma^2)$. Desta forma, estamos interessados em testar hipóteses sobre μ_Z a partir de $\boldsymbol{Z}$. Em particular, a hipótese acima se traduz em

$$
\begin{aligned}
H_0 &: \mu_Z \leq 0,\\
H_1 &: \mu_Z > 0,
\end{aligned}
$$

uma hipótese que podemos testar utilizando um teste t unicaudal como já discutido.

[16] Note que, neste caso, esta é a hipótese razoável a ser testada, porque estamos interessados apenas em rejeitar a hipótese de que a droga **não aumenta** a pressão arterial dos pacientes. Drogas que não tem efeito ou causam aumento da pressão não costumam ser aprovadas pela ANVISA.

Considere agora a situação em que dispomos de dois conjuntos de dados, $\boldsymbol{X} = \{X_1, X_2, \ldots, X_m\}$ e $\boldsymbol{Y} = \{Y_1, Y_2, \ldots, Y_n\}$ e queremos estudar diferenças nas médias. Novamente, vamos modelar os processos como distribuições normais: $X_i \sim \text{Normal}(\mu_1, \sigma_1^2), i = 1, 2, \ldots, m$ e $Y_j \sim \text{Normal}(\mu_2, \sigma_2^2), j = 1, 2, \ldots, n$. Sob a premissa de homogeneidade $\sigma_1^2 = \sigma_2^2 = \sigma^2$, podemos testar a hipótese

$$
\begin{aligned}
H_0 &: \mu_1 \leq \mu_2, \\
H_1 &: \mu_1 > \mu_2,
\end{aligned}
$$

computando a estatística

$$
U = \frac{\sqrt{m+n-2}(\bar{X}_m - \bar{Y}_n)}{\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)\left(S_X^2 + S_Y^2\right)}},
$$

onde $\bar{X}_m = (1/m)\sum_{i=1}^m X_i$, $\bar{Y}_n = (1/n)\sum_{j=1}^n Y_j$, $S_X^2 = \sum_{i=1}^m (X_i - \bar{X}_m)^2$ e $S_Y^2 = \sum_{j=1}^n (Y_j - \bar{Y}_n)^2$. **O teste procede analogamente ao que já foi discutido**.

Relaxando a premissa de homogeneidade

Até aqui assumimos variâncias iguais, tanto no teste pareado quanto no teste para duas amostras. Podemos relaxar a premissa de igualdade das variâncias um pouco se assumirmos que $\sigma_2^2 = k\sigma_1^2$, isto é, que a razão entre as variâncias é conhecida. Neste caso, a estatística teste vale

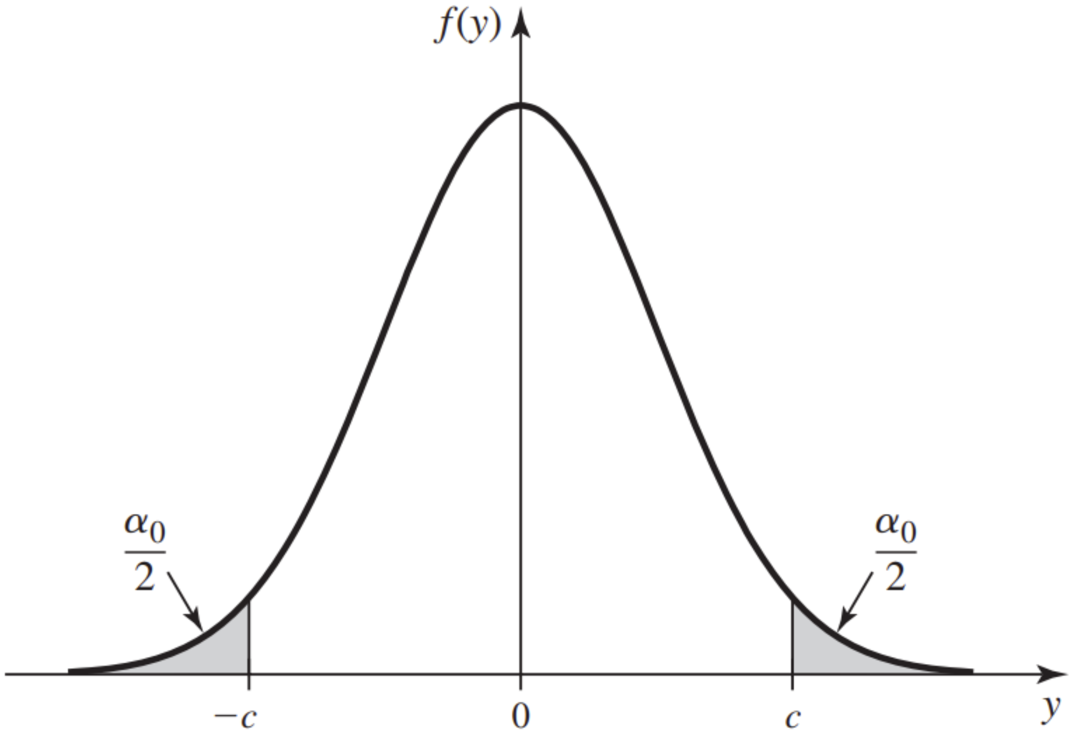
$$U_k = \frac{\sqrt{m+n-2}(\bar{X}_m - \bar{Y}_n)}{\sqrt{\left(\frac{1}{m} + \frac{k}{n}\right)\left(S_X^2 + \frac{S_Y^2}{n}\right)}}.$$

Quando as variâncias são diferentes e desconhecidas e não conhecemos k, temos o problema de Behrens-Fisher[17] que é muito mais difícil de tratar.

[17] Em homenagem ao químico alemão Walter-Ulrich Behrens (1902–1962) e ao biólogo e estatístico britânico Ronald Aylmer Fisher (1890-1962).

O teste t bicaudal (bilateral)

No caso do teste t pareado, podemos estar interessados apenas em testar $\mu_{\text{antes}} = \mu_{\text{depois}}$, o que levaria a uma hipótese alternativa composta e um teste bicaudal (bilateral). Situação parecida acontece no caso de duas amostras quando queremos testar $\mu_1 = \mu_2$. Nesses casos, podemos facilmente adaptar os testes discutidos para acomodar a hipótese bilateral. Em ambos os casos, podemos fazer o teste "rejeite H_0 se $|U| \geq T^{-1}(1 - \alpha_0/2; n-1)$", e este terá tamanho α_0.



O Teste t como um TRV (LRT)

Podemos também entender o teste t como um teste de razão de verossimilhanças. Em particular, temos para um teste t unicaudal,

$$
\begin{aligned}
\Lambda(\boldsymbol{x}) &= \frac{\sup_{(\mu,\sigma^2):\mu>\mu_0} f_n(\boldsymbol{x} \mid \mu, \sigma^2)}{\sup_{(\mu,\sigma^2)} f_n(\boldsymbol{x} \mid \mu, \sigma^2)}, \\
&= \begin{cases} \left(\frac{\hat{\sigma^2}}{\hat{\sigma_0^2}}\right)^{n/2} & \text{se } \bar{x}_n > \mu_0, \\ 1, \text{ caso contrário,} \end{cases}
\end{aligned}
$$

onde $\hat{\sigma^2}$ é o EMV da variância e $\hat{\sigma_0^2} = \frac{1}{n}\sum_{i=1}^n (x_i - \mu_0)^2$. Onde o teste t tradicional rejeita H_0 se $U \geq c$, sua formulação TRV rejeita H_0 se $\Lambda(\boldsymbol{x}) \leq k$. A relação entre c e k é

$$
c = \sqrt{\left[\left(\frac{1}{k^2}\right)^{1/n} - 1\right](n-1)},
$$

o que estabelece que o teste t é um teste de razão de verossimilhanças.

O que aprendemos?

- O teste t permite comparar a média de um conjunto de dados com um valor postulado μ_0;
- Permite também comparar as médias de duas amostras, pareadas ou independentes;
- O teste t é não-viesado e pode ser escrito como um teste de razão de verossimilhanças;

Leitura recomendada

- De Groot seções 9.5 e 9.6;
- * Casella & Berger (2002), seção 8.
- Próxima aula: De Groot, seção 9.7;
- **Exercícios recomendados**
  - De Groot Seção 9.5: exercício 8.