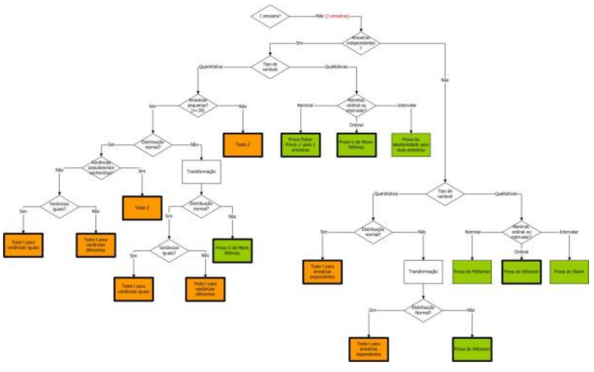
Comparação de duas médias amostrais

Tratamento Paramétrico

Aula de hoje

Testes de Tendência Central (média, mediana, proporção)	Classificação	Variável 1	Variável 2	Número Grupos	Dependência	Premissas
Teste Z	Paramétrico	Quantitativa	-	1	-	Variância pop. * conhecida
Teste t	Paramétrico	Quantitativa	-	1	-	Distribuição normal
Wilcoxon (teste dos sinais, Wilcoxon p/ 1 amostra)	Não-Paramétrico	Quantitativa / Qualitativa ordinal	-	1	-	
Teste t p/ 2 amostras	Paramétrico	Quantitativa	Nominal	2	Independentes	Distribuição normal; variâncias iguais
Teste t p/ 2 amostras com variâncias diferentes	Paramétrico	Quantitativa	Nominal	2	Independentes	Distribuição normal; variâncias diferentes
Teste t pareado	Paramétrico	Quantitativa	Nominal	2	Dependentes	Distribuição normal
ANOVA	Paramétrico	Quantitativa	Nominal	3 ou mais	Independentes	Distribuição normal; variâncias iguais
ANOVA p/ medidas repetidas	Paramétrico	Quantitativa	Nominal	3 ou mais	Dependentes	Distribuição normal; Esfericidade
Mann-Whitney (Wilcoxon não-pareado)	Não-Paramétrico	Quantitativa / Qualitativa ordinal	Nominal	2	Independentes	
Wilcoxon (Wilcoxon pareado, teste dos sinais)	Não-Paramétrico	Quantitativa / Qualitativa ordinal	Nominal	2	Dependentes	
Kruskal-Wallis	Não-Paramétrico	Quantitativa / Qualitativa ordinal	Nominal	3 ou mais	Independentes	
Friedman	Não-Paramétrico	Quantitativa / Qualitativa ordinal	Nominal	3 ou mais	Dependentes	
Teste p/ 1 proporção	Paramétrico	Nominal	-	1	-	
Teste p/ 2 proporções	Paramétrico	Nominal	Nominal	2	Independentes	



Diferenças entre testes paramétricos e não-paramétricos

- Testes paramétricos:
  - Baseados em parâmetros da amostra (média e desvio-padrão).
  - Funcionam melhor se distribuição normal
- Testes não-paramétricos:
  - Baseiam-se em postos (*ranks*) dos dados.
  - Pouco influenciados por valores extremos
  - Não dependem da distribuição dos dados

Transformação

- E quando houver uma clara discrepância dos dados em relação à distribuição Normal?
  - Duas saídas possíveis:
    - Transformar os dados (ex. calculando o logaritmo ou a raiz quadrada) em uma tentativa de obter uma distribuição aproximadamente Normal;
      - Desvantagem: a interpretação dos resultados fica mais complexa.
    - Utilizar um teste não-paramétrico adequado.

Implicações do tamanho da amostra

- Amostras muito pequenas (< 6 observações): Testes de normalidade e variância se tornam pouco confiáveis nessas situações, comprometendo a validação das premissas, e portanto sugere-se utilizar testes não-paramétricos.

Teste *t* para 2 amostras

Hipóteses do teste:

- **Hipótese nula**: média das duas populações são iguais.
- **Hipótese alternativa**: média das duas populações são diferentes.

$$H_0 : \mu_1 = \mu_2 \qquad \text{ou} \qquad H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 \neq \mu_2 \qquad\qquad\quad H_1 : \mu_1 - \mu_2 \neq 0$$

Funcionamento do teste t para 2 amostras independentes

- **Variâncias iguais:**

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

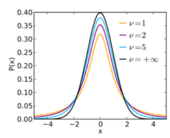
com $n_1 + n_2 - 2$ graus de liberdade

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

s: desvio padrão conjugado

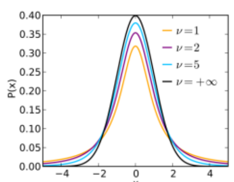
- **Variâncias diferentes:**

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$



Cálculo do valor de p e distribuição t de student

- O valor de p é a probabilidade de obter uma dada diferença entre as médias $\bar{x}_1 - \bar{x}_2$ dado que H0 é verdadeira.



Teste *t* para 2 amostras independentes (e variâncias iguais)

- O teste t para 2 amostras independentes também é conhecido como teste *t* não-pareado
- **Comparação de médias** de **2 grupos** independentes de observações usando amostras representativas.
- Premissas (suposições):
  - indivíduos sorteados **aleatoriamente** da população
  - duas amostras devem ser **independentes**
  - a variável de interesse deve se distribuir de forma **Normal** em cada uma das populações (das quais as amostras foram colhidas)
  - Deve-se saber se as **variâncias** são aproximadamente **iguais** ou **não**

Premissas

- Distribuição normal: Adequado a dados com distribuição simétrica, o que levou à simplificação de que o teste é mais adequado para dados com distribuição normal.
- Variâncias iguais (homocedasticidade) ou variâncias diferentes (heterocedasticidade): Necessário saber se as variâncias das populações estudadas são iguais ou diferentes entre si

Exemplo (Teste *t* – 2 amostras independentes)

- Comparação do peso médio de um grupo de 24 ovelhas que passou por um processo de *flushing* (recebeu nutrição altamente calórica algumas semanas antes do acasalamento) com um grupo-controle de 30 ovelhas.

Teste *t* para 2 amostras independentes (para variâncias iguais)

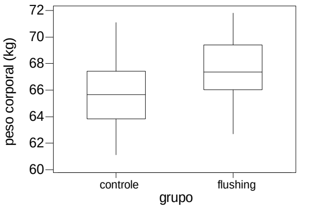
1) Estabelecer as hipóteses do teste

- **Hipótese nula**: os pesos médios dos dois grupos são iguais.
- **Hipótese alternativa**: os pesos médios são diferentes.

$$H_0 : \mu_{controle} = \mu_{flushing} \qquad \text{ou} \qquad H_0 : \mu_{controle} - \mu_{flushing} = 0$$
$$H_1 : \mu_{controle} \neq \mu_{flushing} \qquad \qquad H_1 : \mu_{controle} - \mu_{flushing} \neq 0$$

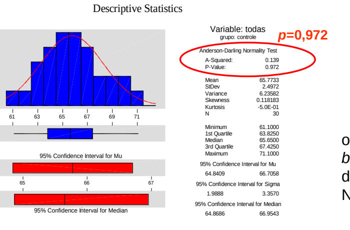
2) Observar **gráficos** referentes a cada uma das amostras. Verificar visualmente se a suposição de distribuição Normal é adequada.

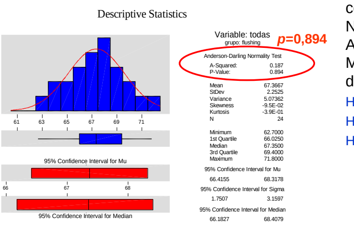
Verifique também se as variâncias são aproximadamente iguais.



Checando as premissas

- Normalidade
- Variâncias





Neste caso, através da observação dos *histogramas e boxplots*, pode-se assumir que os dados sigam a distribuição Normal.

No entanto, podemos confirmar utilizando um teste de Normalidade (como o teste de Anderson-Darling feito pelo Minitab) para confirmar a hipótese de Normalidade.

Hipóteses do teste de Normalidade:

H_0: Distribuição é Normal

H_1: Distribuição não é Normal

***p* >> 0,05 → podemos assumir que os dados sigam a distribuição Normal**

Teste F para variâncias

- O **teste F**, também conhecido como teste da razão de variâncias, pode ser utilizado para testar se dois conjuntos de dados apresentam a mesma variância.

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

- A estatística do teste é

$$F = \frac{s_1^2}{s_2^2}, \quad s_1^2 > s_2^2, \text{ onde } s_1^2 \text{ e } s_2^2 \text{ são as variâncias dos 2 grupos}$$

graus de liberdade do numerador: $n_1 - 1$

graus de liberdade do denominador: $n_2 - 1$

Ainda o Teste F

- No caso do nosso exemplo, os resultados obtidos no Minitab são:

```
Test for Equal Variances

Level1    controle
Level2    dieta
ConfLvl   95,0000

F-Test (normal distribution)

Test Statistic: 1,229
P-Value       : 0,617

Levene's Test (any continuous distribution)

Test Statistic: 0,238
P-Value       : 0,628
```

Para um nível de significância $\alpha = 0,05$:

Como $p>0,05$, não há evidência de desigualdade entre as variâncias e a hipótese de igualdade das variâncias permanece válida.

Voltando ao nosso exemplo...

3) Calcular a estatística do teste *t*. (fórmula)

Nesse caso: *t*=2,43

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

com $n_1 + n_2 - 2$ graus de liberdade

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

s: desvio padrão conjugado

4) Obter o valor de p: *p*=0,018

Há uma chance de 1,8% de obter uma diferença entre os pesos médios de 1,59 kg ou superior, se a hipótese nula for verdadeira.

Two-Sample T-Test and CI: dieta; controle

Two-sample T for dieta vs controle

	N	Mean	StDev	SE Mean
dieta	24	67,37	2,25	0,46
controle	30	65,77	2,50	0,46

Difference = mu dieta - mu controle
Estimate for difference: 1,593
95% CI for difference: (0,279; 2,908)
T-Test of difference = 0 (vs not =): T-Value = 2,43 P-Value = 0,018
DF = 52
Both use Pooled StDev = 2,39

5) Decidir se rejeita ou não H_0:

É pouco provável que a hipótese nula – que é a hipótese de que não há diferença entre os pesos – seja verdadeira. Assim, rejeitamos a hipótese nula em favor da hipótese alternativa, de que há diferença entre os pesos médios. Além disso, o peso das ovelhas que passaram pelo processo de *flushing* é, em média, 1,59kg superior ao das ovelhas do grupo-controle.

6) Intervalo de confiança de 95% para a diferença entre as médias:

IC 95% para a diferença: (0,279; 2,908)

IC 95% não inclui o valor 0 (zero). Portanto, a diferença entre as médias não é compatível com 0, o que confirma a rejeição da hipótese nula.

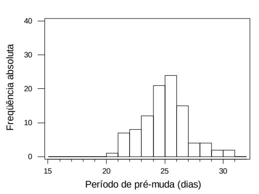
Teste *t* para 2 amostras independentes (variâncias diferentes)

- Nesse caso, utiliza-se um teste *t* modificado, com a seguinte estatística:

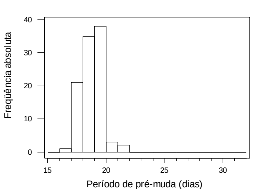
$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- Como esse teste não segue uma distribuição *t*, o cálculo do valor de *p* não é direto. No entanto, os pacotes estatísticos (como o Minitab) incluem essa opção de teste, e fazem a estimativa de *p*.

Histogramas



23 ºC



25 ºC

Período de pré-muda de ninfa de carrapatos

(Dados hipotéticos)

Exemplo – Teste t para amostras independentes (variâncias desiguais)

- Exemplo: Comparar os tempos médios de pré-muda (em dias) de ninfa do carrapato *Amblyomma cajennense*, em laboratório, nas temperaturas de 23°C e 25°C.

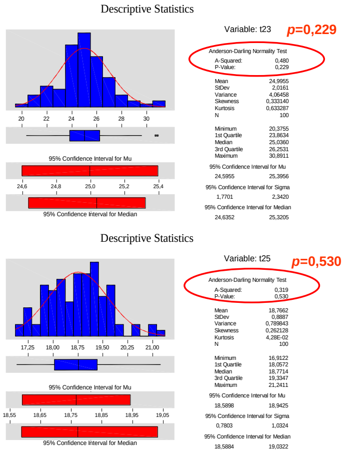
1) Hipóteses:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

Descriptive Statistics: t25; t23

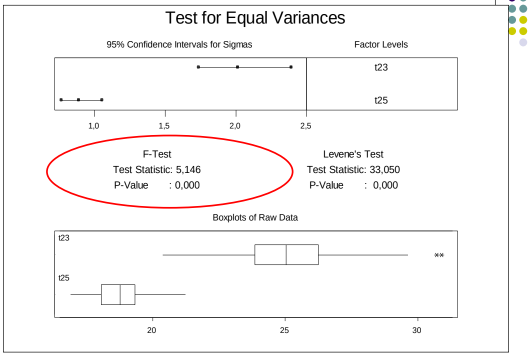
Variable	N	Mean	Median	TrMean	StDev	SE Mean
t25	100	18,766	18,771	18,742	0,889	0,089
t23	100	24,996	25,036	24,943	2,016	0,202

Variable	Minimum	Maximum	Q1	Q3
t25	16,912	21,241	18,057	19,335
t23	20,376	30,891	23,863	26,253



Confirmando a Normalidade dos dados

2) Verificando se as variâncias são iguais (teste F)



p < 0,001 ⇒ **variâncias desiguais**

3) Resultados do teste t para 2 médias amostrais considerando variâncias desiguais

Two-Sample T-Test and CI: t23; t25

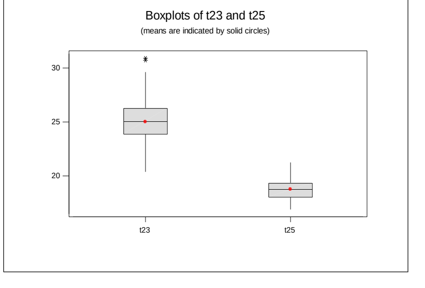
Two-sample T for t23 vs t25

	N	Mean	StDev	SE Mean
t23	100	25,00	2,02	0,20
t25	100	18,766	0,889	0,089

Difference = mu t23 - mu t25
Estimate for difference: 6,229
95% CI for difference: (5,794; 6,665)
T-Test of difference = 0 (vs not =): T-Value = 28,27 P-Value = 0,000 DF = 136

4) Decidir se rejeita ou não a hipótese nula:

O valor de *p* é muito pequeno (*p*<0,001), e, portanto, rejeitamos a hipótese nula de igualdade. Ou seja, os tempos médios de pré-muda para as temperaturas de 23°C e 25 °C são significativamente diferentes, com base nas informações dessas amostras.



Teste *t* pareado

- O teste *t* pareado é utilizado quando selecionamos duas amostras com observações **dependentes** ou **pareadas**.
  - auto-pareamento: cada animal selecionado da população é seu próprio controle;
  - pareamento natural (filhotes da mesma ninhada, gêmeos);
  - pareamento de animais idênticos.

- É baseado na hipótese de que diferenças entre pares de observações se distribuem de forma aproximadamente Normal, embora as observações originais nos grupos possam não apresentar distribuição Normal.
  - Porém, nos casos em que se suspeita que as diferenças não sigam a Normal, podem ser utilizados: transformação dos dados; teste não-paramétrico.
  - Para validar esta premissa, é possível testar a normalidade das amostras separadamente, ao invés de testar as diferenças

Exemplo (teste *t* pareado)

- Um grupo de pesquisadores (Nelson *et al.*, 1998) fez uma comparação de duas diferentes dietas em 11 cães diabéticos, medindo o nível sérico de glicose como uma variável indicadora da qualidade do controle de diabetes. As dietas ou continham fibra pouco insolúvel (LF) ou fibra altamente insolúvel (HF). Os cães foram alocados de modo aleatório para receber uma das dietas primeiro.
- Esse tipo de delineamento é conhecido como “cross-over” (*randomized cross-over trial*).

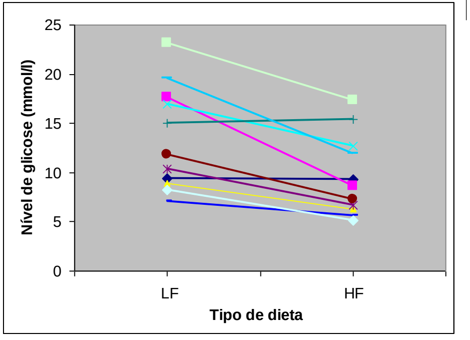
1) Estabelecer as hipóteses do teste

- Hipótese nula: a diferença média do nível de glicose (em mmol/l) entre as duas dietas é zero
- Hipótese alternativa: a diferença média não é zero

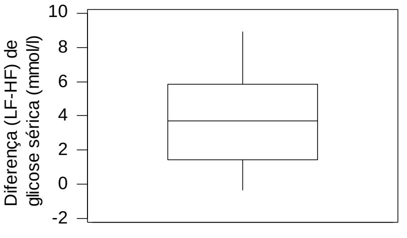
$$H_0 : \mu_{dieta1} = \mu_{dieta2} \quad \text{ou} \quad \mu_d = 0$$
$$H_1 : \mu_{dieta1} \neq \mu_{dieta2} \quad \text{ou} \quad \mu_d \neq 0$$

onde o índice d significa *diferença*

2) Observar um **gráfico** (por exemplo, diagrama de pontos) dos dois grupos que estão sendo comparados.

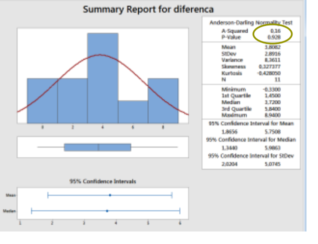


Boxplot da diferença de nível de glicose nas duas dietas



Checando a premissa

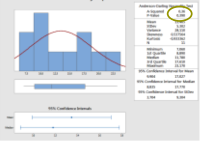
- Teste de normalidade das diferenças:



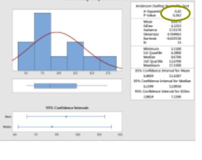
P = 0,928

Checando a premissa

- Alternativamente, e mais simples, teste de normalidade de cada amostra:



P = 0,390



P = 0,262

Ambas amostras possuem distribuição normal

3) Calcular a estatística do teste: ***t* = 4,37**

$$t = \frac{\bar{d}}{s_d / \sqrt{n}}$$

$\bar{d}$ é a média das diferenças

s_d é o desvio padrão das diferenças

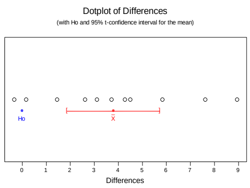
4) Obter o valor de *p*: ***p* = 0,001**

Paired T-Test and CI: LF; HF

Paired T for LF - HF

	N	Mean	StDev	SE Mean
LF	11	13,47	5,30	1,60
HF	11	9,66	4,13	1,24
Difference	11	3,808	2,892	0,872

95% CI for mean difference: (1,866; 5,751)
T-Test of mean difference = 0 (vs not = 0): T-Value = 4,37 P-Value = 0,001



5) Decidir se rejeita ou não H_0:

Se a hipótese nula for verdadeira, há uma chance de apenas 0,1% (p=0,001) de observarmos uma diferença média tão grande quanto 3,81 mmol/l. Como a diferença média é significativamente diferente de zero, rejeitamos H_0. A dieta com fibra altamente insolúvel reduz de modo significativo o nível de glicose em relação à dieta com fibra pouco insolúvel.

6) Intervalo de confiança de 95% para a diferença média:

IC 95% para a diferença: (1,866; 5,751)

IC 95% não inclui o 0 (zero), o que confirma a rejeição de H_0.