Curso de Formación de Personal Investigador Usuario de Animales para Experimentación

INFERENCE ON TWO POPULATIONS

Jesús Piedrafita Arilla
jesus.piedrafita@uab.cat
Departament de Ciència Animal i dels Aliments

UAB

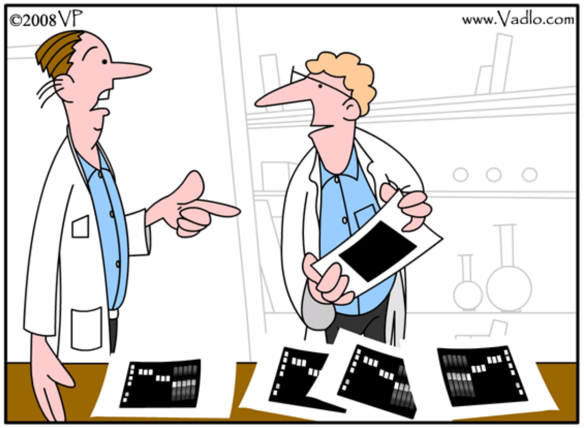
Learning objectives

- Distinguish tests to make comparisons between two means (continuous variables).
- Define test statistics to compare means of independent samples when the within group variances are equal or not.
- Establish a test statistic to compare two variances and describe the *F* distribution.
- Develop some examples with R Commander.
- Establish the conceptual basis to compute the Power of a test.
- Compute Power and Sample size with R.
- Establish the relationship between sample size, power of the test, variability, and size of the difference to be detected.
- Develop a test to compare means for paired data.
- Discuss Statistical significance vs Relevance.

Inferences on two populations

We will present methods to compare **location parameters of two samples** through **parametric procedures**, which assume that the **samples follow a normal distribution**, $y \sim N(\mu, \sigma)$.

	Parametric	Non parametric
Two independent samples	t-test	Wilcoxon Rank-Sum test - U of Mann-Whitney
Paired data	Paired t-test	Wilcoxon Signed-Rank test



Data don't make any sense, we will have to resort to ***statistics***.

Test for the comparison of two means, unknown but equal variance in both groups (T-test), independent samples

Test statistic:

Pooled variance

Sum of squares of sample 2

$$T = \frac{\bar{Y}_1 - \bar{Y}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2} \qquad S_p^2 = \frac{(n_1-1)S_1^2 + \overbrace{(n_2-1)S_2^2}}{n_1+n_2-2}$$

Alternative hypothesis | **Accepted if:**

$$H_1: \mu_1 \neq \mu_2 \quad \rightarrow \quad t > t_{1-\frac{\alpha}{2}}^{n_1+n_2-2} \text{ or } t < t_{\frac{\alpha}{2}}^{n_1+n_2-2}$$

Two-sided test

$$H_1: \mu_1 > \mu_2 \quad \rightarrow \quad t > t_{1-\alpha}^{n_1+n_2-2}$$

$$H_1: \mu_1 < \mu_2 \quad \rightarrow \quad t < t_{\alpha}^{n_1+n_2-2}$$

One-sided test

This test is **little sensitive to the normality assumption**, specially if sample size is large

Test for the comparison of two means, unknown but unequal variances in both groups (T-test), independent samples

Test statistic:

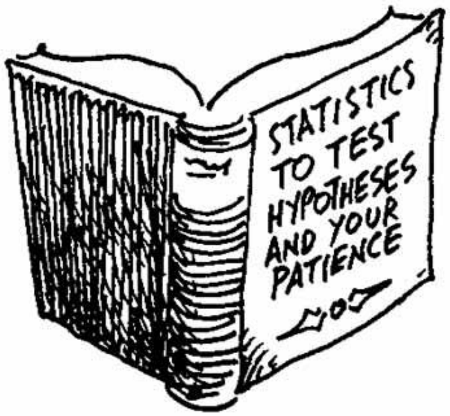
$$T = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim t_\nu \qquad \frac{1}{\nu} = \frac{1}{n_1 - 1}\left(\frac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{\frac{s_2^2}{n_2}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)^2$$

Alternative hypothesis		Accepted if:	
$H_1 : \mu_1 \neq \mu_2$	$\rightarrow$	$t > t^{\nu}_{1-\frac{\alpha}{2}}$ or $t < t^{\nu}_{\frac{\alpha}{2}}$	Two-sided test
$H_1 : \mu_1 > \mu_2$	$\rightarrow$	$t > t^{\nu}_{1-\alpha}$	One-sided test
$H_1 : \mu_1 < \mu_2$	$\rightarrow$	$t < t^{\nu}_{\alpha}$	

If Y_1 and Y_2 are not normal, but sample size is large, this statistic is distributed N(0, 1) and then we can use z instead of t.

Test for the comparison of two variances, independent samples

Test statistic:



$$F = \frac{s_1^2}{s_2^2}$$

Alternative hypothesis		Accepted if:	
$H_1 : \sigma_1^2 \neq \sigma_2^2$	$\rightarrow$	$F > F_{1-\frac{\alpha}{2}}^{n_1-1,n_2-1}$ or $F < F_{\frac{\alpha}{2}}^{n_1-1,n_2-1}$	Two-sided test
$H_1 : \sigma_1^2 > \sigma_2^2$	$\rightarrow$	$F > F_{1-\alpha}^{n_1-1,n_2-1}$	One-sided test
$H_1 : \sigma_1^2 < \sigma_2^2$	$\rightarrow$	$F < F_{\alpha}^{n_1-1,n_2-1}$	One-sided test

This test is **very sensitive to the non normality** of distributions. Small deviations to normality lead to accept the alternative hypothesis.

F distribution

The F-distribution becomes relevant when we try to calculate the ratios of variances of normally distributed statistics. Suppose we have two samples with n_1 and n_2 observations, the ratio

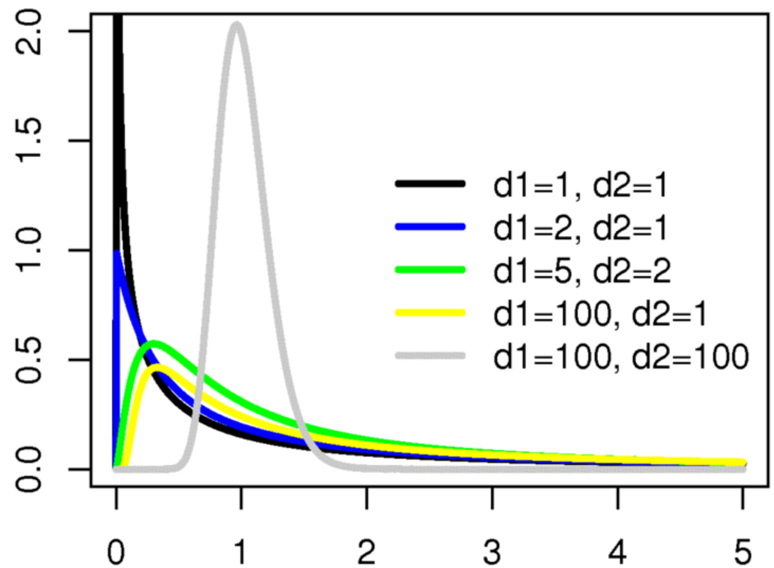
$$F = \frac{s_1^2}{s_2^2}$$

is distributed according to an F distribution (named after R.A. Fisher) with $df_1 = n_1 - 1$ numerator degrees of freedom, and $df_2 = n_2 - 1$ denominator degrees of freedom, with p.d.f.:

$$f(x) = \frac{\sqrt{\frac{(d_1\,x)^{d_1}\,d_2^{d_2}}{(d_1\,x + d_2)^{d_1+d_2}}}}{x\,\mathrm{B}\left(\frac{d_1}{2}, \frac{d_2}{2}\right)}$$

B is the beta-distribution

The F-distribution is skewed to the right, and the F-values can be only positive.



http://en.wikipedia.org/wiki/F-distribution

Comparing two sample means – Descriptives

We want to compare the mean weight of male and female cats. Data of the two samples are assumed to be independent.

Data: catWeights

Statistics > Summaries > Numerical summaries

Variable: BODYwt; Summarize by groups: Sex; Statistics: select

```
      mean        sd  IQR        cv BODYwt:n
F 2.359574 0.2739879 0.35 0.1161175       47
M 2.900000 0.4674844 0.70 0.1612015       97
```

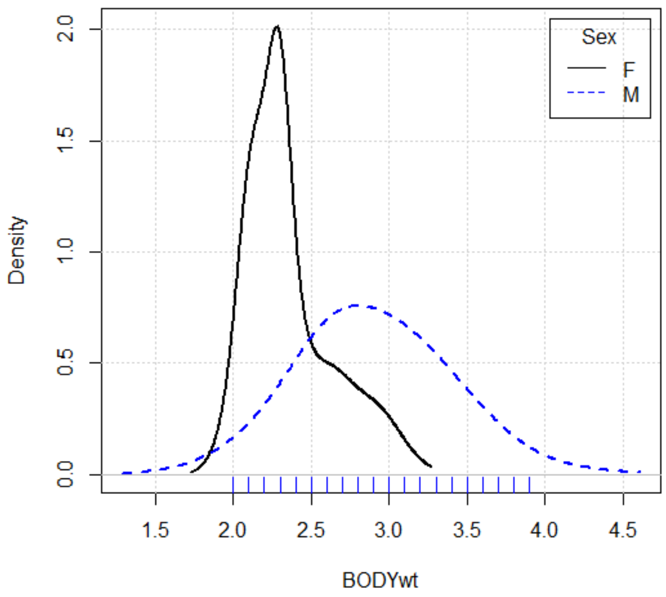
We see numerical differences between sexes in means and sd (variance) regarding BODYwt

Are the differences statistically significant?

Comparing two sample means – Densities and Boxplots

Graphs > Density estimate

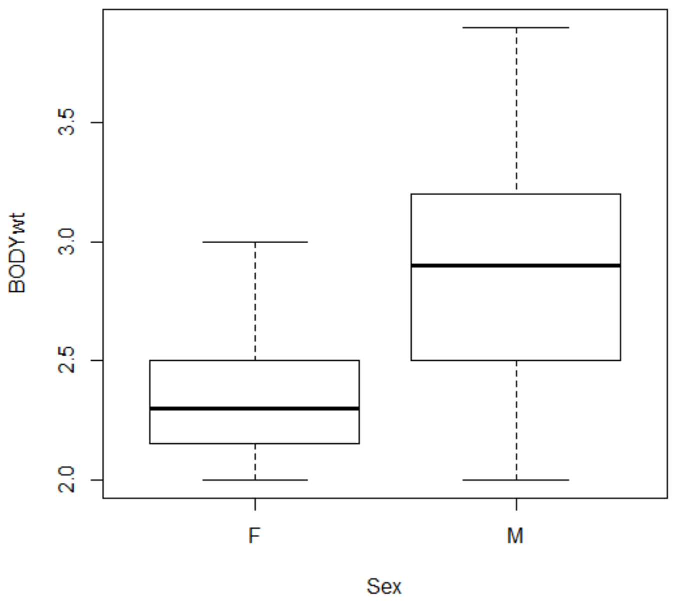
Variable: BODYwt; Plot by: Sex



Graphs > Boxplot

Variable: BODYwt; Plot by: Sex;

Options: Identify outliers no



These graphs have been discussed in a previous lesson

Shapiro test by Sex

Statistics > Data summaries > Test of normality

Variable: BODYwt; Normality test: Shapiro-Wilk; Test by: Sex

```
Sex = F

      Shapiro-Wilk normality test
data:  BODYwt
W = 0.89096, p-value = 0.0003754
```

→ not normal

```
Sex = M

      Shapiro-Wilk normality test
data:  BODYwt
W = 0.97883, p-value = 0.119
```

→ normal

Comparing two sample means – Homog. of variances

Statistics > Variances > Two-variances F-Test

Factors: Sex; Response Variable: BODYwt

```
F          M
0.07506938 0.21854167

       F test to compare two variances

data:  BODYwt by Sex
F = 0.3435, num df = 46, denom df = 96, p-value = 0.0001157
alternative hypothesis: true ratio of variances is not equal
to 1
95 percent confidence interval:
 0.2126277 0.5803475
sample estimates:
ratio of variances
         0.3435015
```

Variance estimates

H_0: variances do not differ, is rejected

Comparing two sample means – t test

Statistics > Means > Independent samples t-test

Groups: Sex; Response Variable: BODYwt; Options: Not equal variances

```
        Welch Two Sample t-test

data:  BODYwt by Sex
t = -8.7095, df = 136.84, p-value = 8.831e-15
alternative hypothesis: true difference in means is not equal
to 0
95 percent confidence interval:
 -0.6631268 -0.4177242
sample estimates:
mean in group F mean in group M
       2.359574        2.900000
```

H_0: means do not differ, is rejected

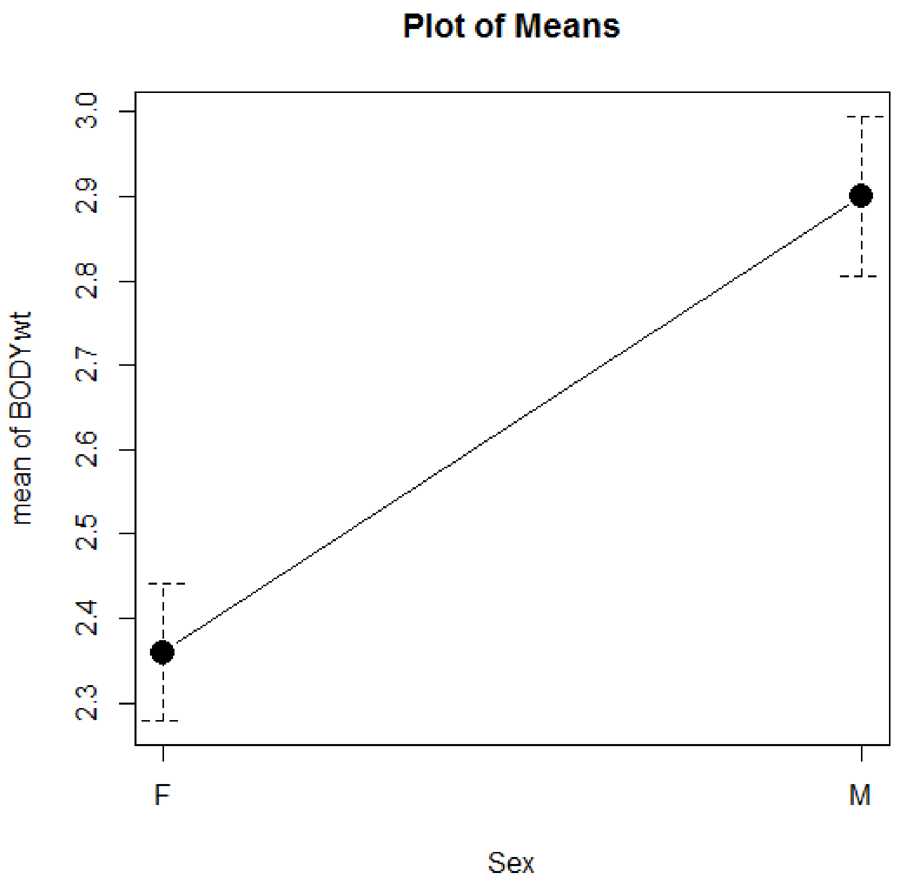
The tests for the equality of variances are sensitive to the non normality of the data. Because of that, Welch test is a more general (unequal variances) and recommended test.

Comparing two sample means – Graph

Graphs > Plot of means

Factors: Sex; Response Variable: BODYwt

Options: Confidence intervals, Top center



Power (1-β) of a one-sided test (1)

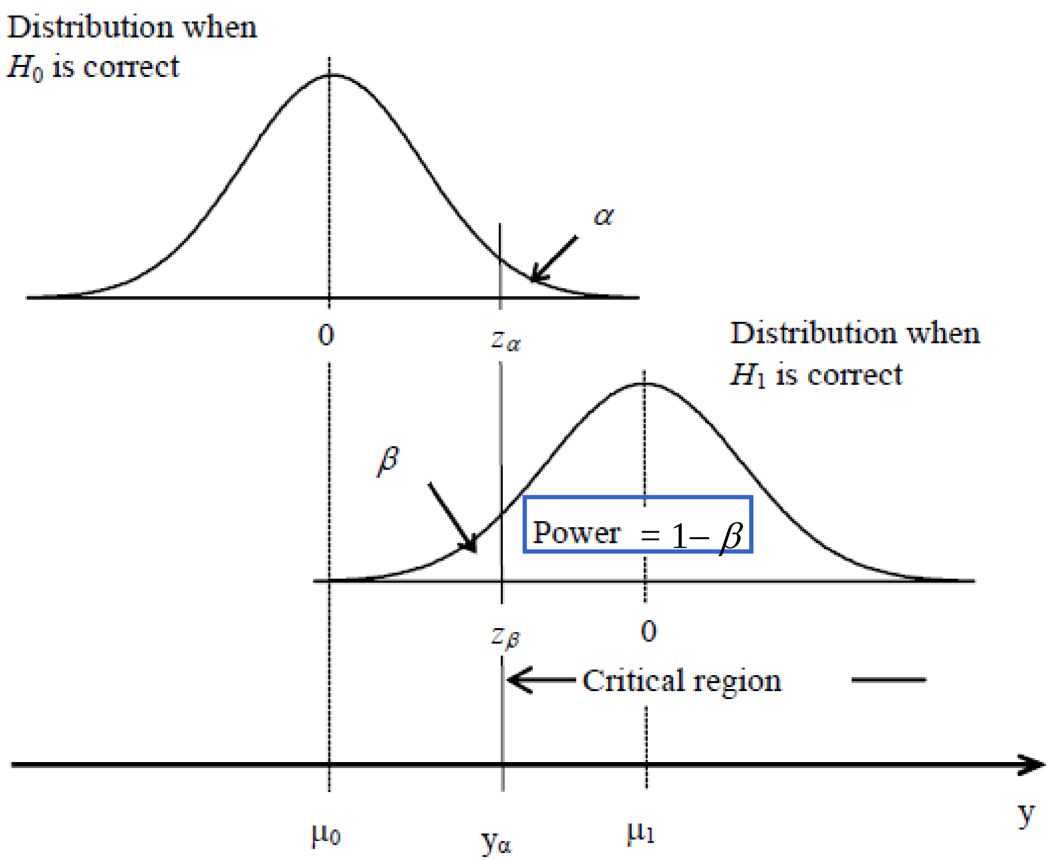


Figure 6.9 Standard normal distributions for H_0 and H_1. The power, type I error (α) and type II error (β) for the one-sided test are shown. On the bottom is the original scale of variable *y*

$$z_\beta = \frac{y_\alpha - \mu_1}{\sigma_{D1}}$$

$$y_\alpha = (\mu_0 + z_\alpha \sigma_{D0})$$

$$z_\beta = \frac{(\mu_0 + z_\alpha \sigma_{D0}) - \mu_1}{\sigma_{D1}}$$

z_β is expressed in terms of z_α and the means and SD of the distributions

Power

$\mu_0 < \mu_1 : P[z > z_\beta]$

$\mu_0 > \mu_1 : P[z < z_\beta]$

Kaps M., Lamberson W.R. 2004. *Biostatistics for the Animal Sciences*. CABI, Wallinford.

Power of a one-sided test (2)

For specific tests, the **appropriate standard deviation** must be specified.

For the test hypothesis $\mu_1 > \mu_0$:

$$z_\beta = \frac{(\mu_0 + z_\alpha \, \sigma_0 / \sqrt{n}) - \mu_1}{\sigma_1 / \sqrt{n}} = \frac{(\mu_0 - \mu_1)}{\sigma / \sqrt{n}} + z_\alpha$$

$\sigma_0 / \sqrt{n}, \sigma_1 / \sqrt{n}$ Standard errors

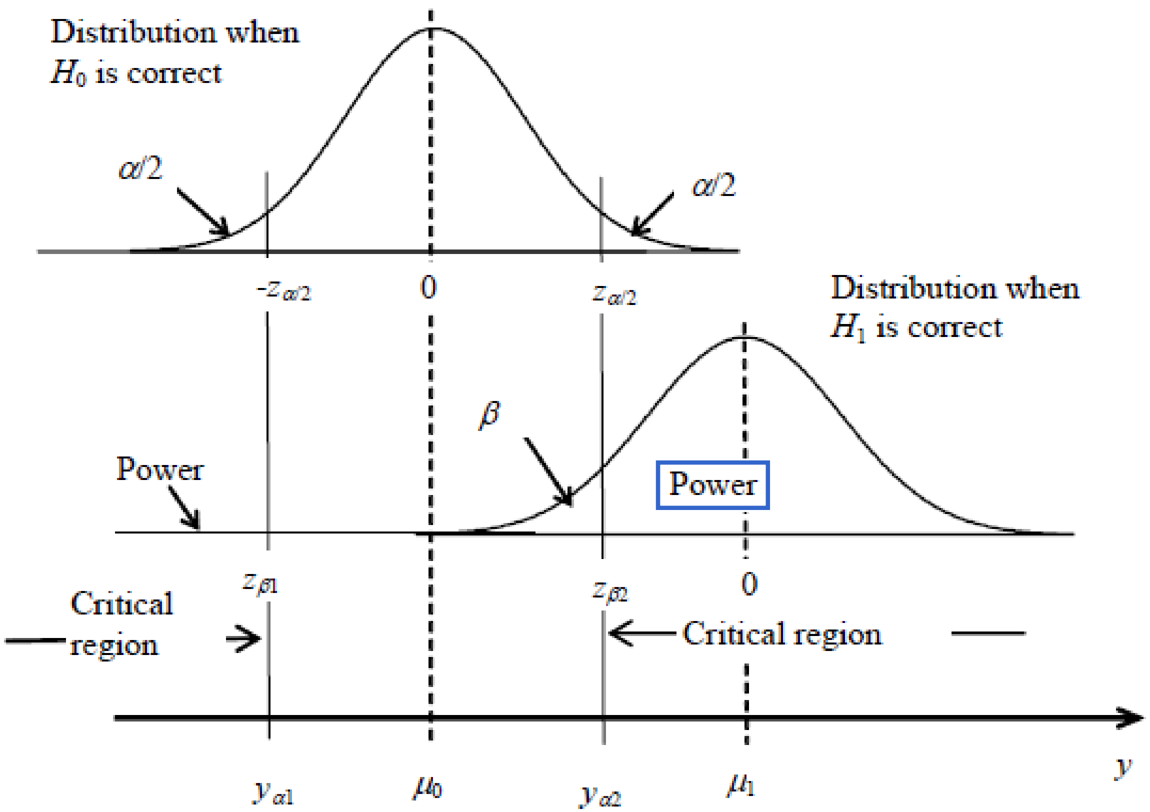
$$\sigma_0 = \sigma_1 = \sigma$$

Testing the hypothesis of the **difference of two means**:

$$z_\beta = \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} + z_\alpha$$

We shall see later that from these formulas we can arrive at calculating the sample size of an experiment

Power of a two-sided test



$$z_{\beta 1} = \frac{(\mu_0 - z_{\frac{\alpha}{2}} \sigma_{D0}) - \mu_1}{\sigma_{D1}}$$

$$z_{\beta 2} = \frac{(\mu_0 + z_{\frac{\alpha}{2}} \sigma_{D0}) - \mu_1}{\sigma_{D1}}$$

Power:

$P[z < z_{\beta 1}] + P[z > z_{\beta 2}]$,

using the H_1 distribution

Figure 6.10 Standard normal distributions for H_0 and H_1. The power, type I error (α) and type II error (β) for the two-sided test are shown. On the bottom is the original scale of the variable *y*

(Kaps and Lamberson, 2004)

Power calculations in a *t*-test with R

We take the BODYwt results from slide 8.

Note that d = relevant difference / common sd

```
> library(pwr)
> pwr.t2n.test(n1=47, n2=97, d=0.54/0.2739879,
sig.level=0.05, alternative="two.sided")

     t test power calculation

             n1 = 47
             n2 = 97
              d = 1.97089
      sig.level = 0.05
          power = 1
    alternative = two.sided
```

We have a probability of 1 (or 100%) of rejecting the null hypothesis (H_0) when false

In this case we have used as common sd the sd of females, as females act as the "control" to calculate the difference between the means of males and females: *Glass* delta. With equal variances, the square root of Pooled variance should be used.

Determination of Sample size through power analysis (1)

- The effect size (minimum difference) of biological interest (relevance), δ
- The standard deviation, σ
- The significance level, α (5%)
- The desired power of the experiment, 1-β (80%)
- The alternative hypothesis (i.e. a one- or two-sided test)
- The sample size

Fix any five of these variables and a mathematical relationship can be used to estimate the sixth

Determination of Sample size (2)

After some algebra from the formulas of power, it can be arrived at:

One-sided test of a population mean

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma^2}{\delta^2}$$

One-sided test of the difference of two population means

$$n = \frac{(z_\alpha + z_\beta)^2}{\delta^2} 2\sigma^2$$

For a **two-sided test**, replace z_α with $z_{\alpha/2}$ in the above expressions.

Remember that z_α and z_β are obtained from a table of percentile of a standard normal distribution:

$z_{\alpha/2}(0.05) = 1.96$; $z_\alpha(0.05) = 1.65$

$z_\beta(0.8) = -0.85$; $z_\beta(0.9) = -1.28$

```
> qnorm(0.975)
[1] 1.959964
```

Determination of Sample size with R

We take again the BODYwt results from slide 8.

```
> pwr.t.test(d=0.54/0.2739879, power=0.8,
sig.level=0.05, alternative="two.sided")
```

Note that this sentence uses **pwr.t.test**. It is included the desired **power=0.8**, and **n1** and **n2** were removed.

```
     Two-sample t test power calculation

              n = 5.201801
              d = 1.97089
      sig.level = 0.05
          power = 0.8
    alternative = two.sided

NOTE: n is number in *each* group
```

We need only 5 individuals in each group for detecting the relevant difference (δ) of 0.54

The students could resort also to GRANMO: http://www.imim.cat/ofertadeserveis/granmo.html

Test for the comparison of two means, paired data

Test statistic:

$$T = \frac{\bar{d}}{S_d / \sqrt{n}} \sim t_{n-1}$$

Alternative hypothesis		Accepted if:	
$H_1: \mu_1 \neq \mu_2$	$\rightarrow$	$t > t^{n-1}_{1-\frac{\alpha}{2}}$ or $t < t^{n-1}_{\frac{\alpha}{2}}$	Two-sided test
$H_1: \mu_1 > \mu_2$	$\rightarrow$	$t > t^{n-1}_{1-\alpha}$	One-sided test
$H_1: \mu_1 < \mu_2$	$\rightarrow$	$t < t^{n-1}_{\alpha}$	

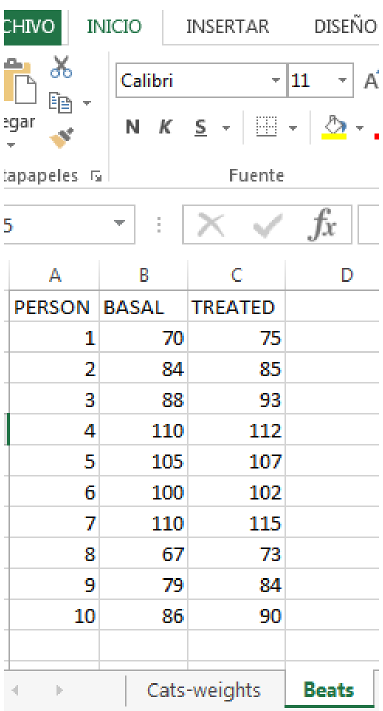
An example of comparison of two means, paired data

Data of **hearth rate of people** before and after taking a pill.

Data: Beats

Statistics > Means > Paired t-test

First variable: TREATED; Second variable: BASAL



PERSON	BASAL	TREATED
1	70	75
2	84	85
3	88	93
4	110	112
5	105	107
6	100	102
7	110	115
8	67	73
9	79	84
10	86	90

```
        Paired t-test

data:  TREATED and BASAL
t = 6.6217, df = 9, p-value = 0.00009683
alternative hypothesis: true difference in
means is not equal to 0
95 percent confidence interval:
 2.435978 4.964022
sample estimates:
mean of the differences
                    3.7
```

H_0 is rejected

A final remark

By increasing the sample size of the experiment (number of replicates of each group) we can always arrive at a STATISTICALLY SIGNIFICANT difference.

This difference, however, must be also RELEVANT in any sense (biological, economical, etc.).

This reasoning must be applied to any other statistic test we apply in our research.