Statistical Analysis

Analysis of Variance (ANOVA)

Ana Cristina Costa

ccosta@novaims.unl.pt

NOVA IMS Information Management School

Topics

- **LU7 – ANOVA**
  - Introduction
  - One-way ANOVA with fixed effects
  - Multiple comparison tests
  - Tests to the equality of k variances
  - Applications in Excel

Objectives

- **At the end of this learning unit students should be able to**
  - Understand the applicability of ANOVA
  - State the assumptions of ANOVA
  - Formulate the hypotheses of the One-way Analysis of Variance
  - Apply the One-way Analysis of Variance and interpret its results
  - Apply multiple comparison tests
  - Test the equality of three or more variances of independent populations
  - Identify alternative tests when the assumptions of ANOVA are not met

Suggested reading

- Newbold, P., Carlson, W. L., Thorne, B. (2013). Statistics for Business and Economics. 8th Edition, Boston: Pearson, pages 645-660 (ch. 15).

- Mariappan, P. (2019). Statistics for Business. New York: Chapman and Hall/CRC, pages 301-315.

- Mendenhall, W., Beaver, R. J., & Beaver, B. M. (2013). *Introduction to probability and statistics*. 14th Edition, Boston: Brooks/Cole, Cengage Learning, pages 425-440.

- Holmes, A., Illowsky, B. & Dean, S. (2019) "F Distribution and One-Way ANOVA". In Introductory Business Statistics. (accessed: August 2021)

- The Pennsylvania State University (2021) "Lesson 13: One-Factor Analysis of Variance". In STAT 415 Introduction to Mathematical Statistics. (accessed: August 2021)

Resources on the internet

- **Does your data violate one-way ANOVA assumptions?**
  https://www.quality-control-plan.com/StatGuide/oneway_anova_ass_viol.htm (accessed: August 2021)

- **Possible alternatives if your data violate one-way ANOVA assumptions**
  https://www.quality-control-plan.com/StatGuide/oneway_anova_alts.htm (accessed: August 2021)

- Costa, A. C. (2019). **One-way ANOVA process and nonparametric counterparts**. NOVA Information Management School, Universidade Nova de Lisboa, https://doi.org/10.13140/RG.2.2.25687.47520/1

Introduction

- Ronald Fisher introduced a technique that allows to analyse data (response variable) that are affected by various external factors (independent variables) that may, or may not, operate at the same time

- It is based on a comparison of the variability among samples means, thus it was named analysis of variance (ANOVA)

- The ANOVA model to be applied depends on the planning of the experiment and the number of factors

Introduction

- **Example**
  - A marketing director plans to re-launch a product to market, so he studied three different marketing campaigns, each of them differently combining factors such as product price, product presentation, associated promotions, etc.. Any of these campaigns is carried out at the point of sale, without any publicity in the media.
  - The objective is to determine if there is a difference between the three campaigns for their effectiveness
  - The total sales are recorded for a period of limited duration in a randomly selected set of stores, for each marketing campaign, forming three samples independent of each other
  - In this case, the observations come from (three) groups classified by a **single factor** (marketing campaign), so you should use **One-way ANOVA**

Introduction

- **Example**
  - Now, suppose that the location of the stores may also influence the sales performance and interact with the marketing campaigns effectiveness. Stores can be divided by two locations: in city centre or elsewhere.
  - Sales can be recorded in a randomly selected set of stores, for each marketing campaign, within each location. As a result, six samples independent of each other are obtained.
  - In this case, the observations come from (six) groups classified by **two factor**s (marketing campaign and location). Hence, the **Two-way ANOVA** will be appropriate to investigate differences in sales performance caused by each of the factors or their interactions

Introduction

- **Definitions**
  - **Experimental units**: objects/subjects on which observations are made
  - **Factor**: independent variable (characteristic) completely controlled in an experience, with ***k*** **levels**
    - The different degrees of intensity, or different categories, of the factor are the **levels**
    - If the factor levels correspond to different intensities measured on a scale, the factor is said to be <u>quantitative</u>
    - If the levels of a factor differ only in some characteristics, the factor is said to be <u>qualitative</u>
  - **Group** or **Treatment**: specific combination of factors' levels
    - In case of one single factor, each group corresponds to a level of the factor

Introduction

- **Previous example**
  - Suppose you want to investigate if differences in sales performance are caused by stores location and/or type of marketing campaign
    - The experimental units are the stores where the sales are recorded
    - There are two factors: location and marketing campaign
    - The 'location' factor has 2 levels
    - The 'marketing campaign' factor has 3 levels
    - The experiment has 6 groups/treatments
      - An independent random sample of experimental units is obtained from each group/treatment

Introduction

- **One-way ANOVA with fixed effects**
  - In this course we will study only the case of the analysis of variance with one factor and fixed effects
  - **Fixed effects model**: case in which the levels of the factor are fixed, which means that groups/treatments are fixed at the outset (in opposition to randomly determined)
  - **One-way ANOVA**
    - The only factor has *k* levels
    - Each group/treatment corresponds to a level of the factor
    - ANOVA allows to compare the equality of *k* populations' means based on the samples obtained for each of the k groups/treatments

Introduction

- **Completely randomized design**
  - In a completely randomized design to compare ***k* groups/treatments**, a set of ***n*** relatively homogeneous experimental units are randomly divided into *k* groups of dimensions n_1, n_2, ... n_k, in which $\mathbf{n_1 + n_2 + ... + n_k = n}$
  - All experimental units in each group receive the same treatment, so that each treatment is applied to exactly one group
  - To each group/treatment is associated a population which consists of all the observations that would be obtained if the treatment was repeatedly applied to all possible experimental units

One-way ANOVA with fixed effects

One-way ANOVA

- Consider k populations $\mathbf{X_1 \sim N(\mu_1,\sigma_1^2), X_2 \sim N(\mu_2,\sigma_2^2), \ldots, X_K \sim N(\mu_k,\sigma_k^2)}$ for which homoscedasticity of variances is verified (i.e., $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 = \sigma^2$)

- Consider a completely randomized experimental design, in which k independent random samples of the populations under study are obtained , with dimensions n_1, n_2, ... n_k, in which $\mathbf{n_1+n_2+ \ldots +n_k=n}$

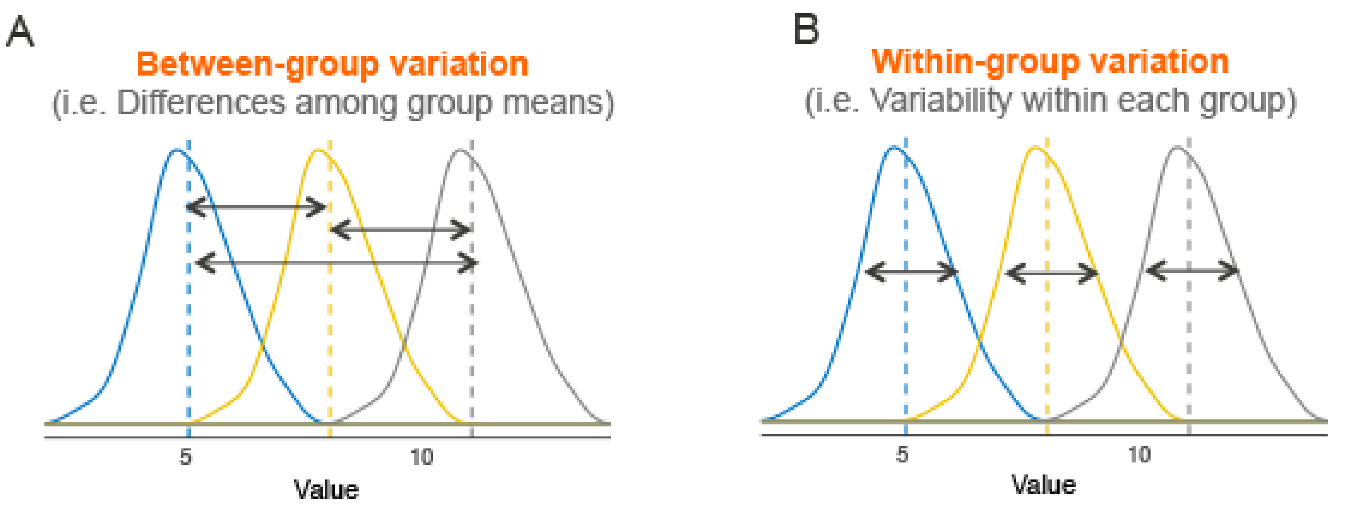
- **Objective**: Test the equality of three or more populations' means, i.e., test if for a given factor the mean is the same in all its levels

  H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$

  H_1: $\exists_{i,j\ (i \neq j)}: \mu_i \neq \mu_j$

One-way ANOVA with fixed effects

- **Idea underlying the ANOVA**
  - The whole idea behind the analysis of variance is to compare the ratio of between group variance to within group variance
  - If the variance caused by the interaction between the samples is much larger when compared to the variance that appears within each group, then it is because the means are not the same



One-way ANOVA with fixed effects

- **Idea underlying the ANOVA**
  - The procedure for testing H_0 is based on two independent estimators of the populations variance σ^2
    - The 1^{st} will be a valid estimator either H_0 is true or not
    - The 2^{nd} is a valid estimator when H_0 is true
  - The proposed statistic for the test corresponds to the ratio between the 2^{nd} and 1^{st} estimators
    - If H_0 is true, the two estimators tend to produce similar estimates, therefore the test statistic will be approximately equal to 1
    - If H_0 is false, the 2^{nd} estimator tends to overestimate σ^2, i.e. the test statistic tends to take values higher than 1

One-way ANOVA with fixed effects

- **Idea underlying the ANOVA**
  - The 1st estimator consists in calculating the sample variance for each population (S_i^2: i=1, ..k) and take the mean of the estimates that are obtained
  - If H_0 is true, the mean values are all equal, thus it is obtained a set of *k* samples withdrawn from the same population
    - The variance of the global sample mean is equal to σ^2/n
    - We can get a "sample" of the *k* sample means and calculate the variance of these values (variability between the groups)
    - Thus, obtaining an estimate of σ^2/n (by multiplying by *n* an estimate of σ^2 is determined)

One-way ANOVA with fixed effects

- **Mathematical model**
  - The random variables X_{ij} are independent and verify (for each experimental unit $j=1,\ldots,n_i$ from each group $i=1,\ldots,k$)

$$X_{ij} = \mu_i + \varepsilon_{ij} \quad \Leftrightarrow \quad X_{ij} = \mu + \alpha_i + \boldsymbol{\varepsilon_{ij}}$$

    - $\mu_i = \mu + \alpha_i$ → mean value of the level *i* of the factor in the population

      μ → mean value of the population

      α_i → effect of the factor

    - $\boldsymbol{\varepsilon_{ij}} \sim N(0, \sigma^2)$ → random residual

  - To check whether there are significant differences between groups / treatments, the following null hypothesis is formulated:

$$H_0: \quad \alpha_1 = \alpha_2 = \ldots = \alpha_k$$

    - I.e., there are no differences between the means of k populations
    - Or, the effects of the application of the k treatments are not statistically significant

  - The alternative hypothesis is then: the effects of the treatments are significant

One-way ANOVA with fixed effects

Notation

k	Number of levels of the factor (no. of populations)
n_i	Number of observations of the *i* (population) level (i=1, .. ,k)
X_{ij}	Response variable of the *i* (population) level for the experimental unit (individual) *j*
$n = \sum_{i=1}^{k} n_i$	Total number of observations
$X_{i\bullet} = \sum_{j=1}^{n_i} X_{ij}$	Sample total corresponding to the *i* (population) level
$\overline{X}_{i\bullet} = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij} = \frac{X_{i\bullet}}{n_i}$	Sample mean corresponding to the *i* (population) level
$\overline{X} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij} = \frac{1}{n}\sum_{i=1}^{k} X_{i\bullet} = \frac{1}{n}\sum_{i=1}^{k} n_i \overline{X}_{i\bullet}$	Global sample mean

One-way ANOVA with fixed effects

Notation

$S^2 = \frac{1}{n-1}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij}-\bar{X}\right)^2 = \frac{1}{n-1}\left(\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^2 - n\bar{X}^2\right)$	Global sample variance
$S_i^2 = \frac{1}{n_i-1}\sum_{j=1}^{n_i}\left(X_{ij}-\bar{X}_{i\bullet}\right)^2 = \frac{1}{n_i-1}\left(\sum_{j=1}^{n_i}X_{ij}^2 - n_i\bar{X}_{i\bullet}^2\right)$	Sample variance of the *i* (population) level
$SST = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij}-\bar{X}\right)^2 = (n-1)S^2$	Sum of squares of the total deviations around the global mean (**total variability of responses**)
$SSTr = \sum_{i=1}^{k} n_i(\bar{X}_{i.}-\bar{X})^2 = \sum_{i=1}^{k}\frac{\bar{X}_{i.}^{\ 2}}{n_i} - n\bar{X}^2$	Sum of squares of the deviations between the levels of the factor (**variation due to the treatments**; **variation between groups**)
$SSE = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij}-\bar{X}_{i.}\right)^2 = \sum_{i=1}^{k}(n_i-1){S_i}^2$	Sum of squares of the deviations within the levels of factor (**variation due to error**; **variation within groups**)

One-way ANOVA with fixed effects

Notation

SST = SSTr + SSE	**Total variation = Variation explained by the treatments + Variation due to error**
$MST = \dfrac{SST}{n-1}$	Mean squares of the total deviations
$MSTr = \dfrac{SSTr}{k-1}$	Mean squares of the deviations between the levels of factor **If H_0 is true, it is an unbiased estimator of the population variance with k–1 degrees of freedom**
$MSE = \dfrac{SSE}{n-k}$	Mean squares of the deviations within the levels of factor **It is an unbiased estimator of the population variance with n–k degrees of freedom**
S_i^2	**It is an unbiased estimator of the population variance with n_i–1 degrees of freedom**

One-way ANOVA with fixed effects

- ## ANOVA table

Source of variation	Degrees of freedom (df)	Sum of squares	Mean squares	*F*
Treatments (*Between* groups)	k – 1	SSTr	$MSTr = \frac{SSTr}{k-1}$	$F_{obs} = \frac{MSTr}{MSE}$
Error (*Within* groups)	n – k	SSE	$MSE = \frac{SSE}{n-k}$	
Total	n – 1	SST		

One-way ANOVA with fixed effects

- **Hypotheses**

$H_0\colon \mu_1 = \mu_2 = \cdots = \mu_k$

$H_1\colon \exists_{i,j(i \neq j)}\colon \mu_i \neq \mu_j$

- **Test statistic**

$$F_{obs} = \frac{MST_r}{MSE} \sim F_{(k-1,n-k)}$$

- **Decision rule** (right-sided test)

Reject H_0 if $F_{obs} > F_{(k-1,\, n-k;\, 1-\alpha)}$

- ***p-value***

$P[F_{(k-1,\, n-k)} \geq F_{obs}]$

One-way ANOVA with fixed effects

Assumptions of application

- The violation of the **independence condition of the observations** has, in general, serious consequences about the validity of the ANOVA

- **Normal populations**
  - The ANOVA shows robustness relatively to small deviations from the normality condition. The distribution of the population should be approximately symmetric and mesokurtic.
  - The normality condition must be tested by means of adjustment tests, such as the **Shapiro-Wilk test**
  - If the normality cannot be assumed for <u>all samples</u>, the nonparametric **Kruskal-Wallis test** can be used instead of the ANOVA’s *F*-test, and it is just slightly less powerful than ANOVA

One-way ANOVA with fixed effects

Kruskal-Wallis test

- Must be used when at least one of the *k* samples does not come from a normal population, or with ordinal variables. The samples' distributions should have an identical shape.
- For k=2, the Kruskal-Wallis test is identical to the Wilcoxon-Mann-Whitney test (a.k.a. Wilcoxon rank sum test a.k.a. Mann-Whitney U test)
- Consists on applying the ANOVA to the ranks of the observations (*Wilcoxon scores*). The test statistic has a $\chi^2_{(k-1)}$ distribution.
- **Hypotheses**
  - $\mathbf{H_0}$: The *k* samples come from identical populations
  - $\mathbf{H_1}$: Not all of the *k* samples come from identical populations

One-way ANOVA with fixed effects

- **Assumptions of application**
  - **Populations with the same variance**
    - The ANOVA shows, in general, reduced robustness relatively to small deviations from the homoscedasticity condition. Greater robustness is achieved when the **design** is **balanced** ($\mathbf{n_1 = n_2 = ... = n_k = b}$)
    - The hypothesis of homoscedasticity should be tested using
      - **Bartlett's test**
      - **Levene's median test**: it is one of the most robust against the violation of the assumption of normality, and it is one of the most powerful
    - If the homoscedasticity cannot be assumed, but normality does, the **Welch's test** can be used instead of the ANOVA's *F*-test
    - If both nonnormality and unequal variances are present, employing a **transformation** may be preferable (*e.g.* log, square root or Box-Cox)

One-way ANOVA with fixed effects

- **Example 1**
  - A governmental department is concerned with the increased costs occurring within the R&D projects that are commissioned to institutes A, B, C and D. It was then decided to analyse the costs associated with different projects, calculating for each one of them the ratio between the incurred final cost and the cost initially indicated in the budget. For each project, the two costs were expressed on a constant basis. Do the four institutes have a distinct global behaviour in relation to the increasing of the costs?

A	B	C	D
1.0	1.7	1.0	3.8
0.8	2.5	1.3	2.8
1.9	3.0	3.2	1.9
1.1	2.2	1.4	3.0
2.7	3.7	1.3	2.5
	1.9	2.0	

Data in the **Example1** sheet of the LU7_Examples Excel file

One-way ANOVA with fixed effects

- **Example 1**
  - Let $X_1 \sim N(\mu_1,\sigma^2)$, $X_2 \sim N(\mu_2,\sigma^2)$, $X_3 \sim N(\mu_3,\sigma^2)$ and $X_4 \sim N(\mu_4,\sigma^2)$ be the random variables corresponding to the ratio between the final cost and the initial cost for the institutes A, B, C and D, respectively
  - Hypotheses: H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$

    H_1: $\exists_{i,j\ (i \neq j)}: \mu_i \neq \mu_j$
  - Sample statistics:

$$n = 22, \qquad \bar{x} = 2.1227, \qquad s^2 = 0.79136(36)$$

	A	B	C	D
Mean	1.5	2.5	1.7	2.8
Standard deviation	0.791	0.746	0.805	0.696
Variance	0.625	0.556	0.648	0.485
Sample size	5	6	6	5

One-way ANOVA with fixed effects

- **Example 1**

Source of variation	df	Sum of squares	Mean squares	F
Treatments	$k - 1 = 3$	**SSTr** = SQT – SQE = 6.1586	**MSTr** = 2.0529	$\mathbf{F_{Obs} = 3.5327}$
Error	$n - k = 18$	**SSE** = $4\times0.625 + 5\times0.556 + 5\times0.648 + 4\times0.485 = 10.46$	**MSE** = 0.5811	
Total	$n - 1 = 21$	**SST** = (21) (0.79136) = 16.6186		

- $F_{Obs} = 3.5327 > F_{(3,\ 18;\ 0.95)} = 3.1599 \ \Rightarrow$ reject H_0 at the 5% significance level
- p-value = $P(F \geq 3.5327) = 0.0359$
- There is evidence that the four institutes do not have an equal global behaviour, in mean value, in relation to the increasing of costs

One-way ANOVA with fixed effects

- **Confidence intervals**
  - Notation: **S^2 = MSE** (unbiased estimator of the population variance σ^2)
  - Confidence interval for the mean value (μ_i) of the *i* treatment

$$\bar{X}_{i\bullet} \pm t_{(n-k);1-\alpha/2} \frac{S}{\sqrt{n}}$$

  - Confidence interval for the difference of mean values $\mu_i - \mu_j$

$$(\bar{X}_{i\bullet} - \bar{X}_{j\bullet}) \pm t_{(n-k);1-\alpha/2} S \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$$

    - $T_{(n-k)\,;\,1-\alpha/2}$ is the percentile of order $1-\alpha/2$ of the distribution $T_{(n-k)}$

Multiple comparison tests

- ***Post-hoc* tests**
  - The rejection of H_0 using the *F* test of the ANOVA only allows to conclude the non equality between the mean values of the *k* groups, but it does not identify which groups
  - The multiple comparison tests (or *post-hoc* tests) have the same assumptions as ANOVA and allow to compare each pair of mean values: $\mathbf{H_0: \mu_i = \mu_j}$ $(i \neq j)$
  - The application of *t*-tests simultaneously does not allow to control the global significance level: **P[correct joint decision] = $(1-\alpha)^m < 1-\alpha$**
    - When we conduct many pairs of comparisons at the alpha significance level, we may conclude that two treatments are different, although they are not (inflation of type I error).

Multiple comparison tests

- **Tukey's HSD test** (*Honestly Significant Difference*)
  - Only applicable in balanced designs: $\mathbf{n_1 = n_2 = ... = n_k = b}$
  - **Test statistic**

$$W = \frac{\left|\overline{X}_{i\bullet} - \overline{X}_{j\bullet}\right|}{\sqrt{\frac{S^2}{b}}} \sim q_{(k;n-k)}$$

$S^2 = MSE$

$q_{(k;\, n-k)}$ → ***Studentized Range*** distribution with (k; n–k) degrees of freedom

  - **Decision rule**
    - Reject H_0 when $W_{Obs} \geq q_{(k;\, n-k;\, 1-\alpha)}$

Multiple comparison tests

- **Tukey-Kramer test**
  - Extension of Tukey's HSD test for unbalanced designs
  - **Test statistic**

$$W = \frac{\left|\overline{X}_{i\bullet} - \overline{X}_{j\bullet}\right|}{\sqrt{\dfrac{S^2}{2}\left(\dfrac{1}{n_i} + \dfrac{1}{n_j}\right)}} \sim q_{(k;n-k)}$$

$S^2 = MSE$

$q_{(k;\, n-k)}$ → ***Studentized Range*** distribution with (k; n–k) degrees of freedom

  - **Decision rule**
    - Reject H_0 when $W_{Obs} \geq q_{(k;\, n-k;\, 1-\alpha)}$

Multiple comparison tests

Alternative parametric tests

TEST	Type I Error (α) (False positive rate)	Type II Error (β) (False negative rate)	Comments
Scheffé	P(Type I Error)$\leq\alpha$	High	Used in complex comparisons. Has some robustness regarding normality and homoscedasticity.
Hochberg (GF2)	Low	Moderate	Generally more conservative than the Tukey-Kramer's test, and always more conservative than the Tukey's HSD
Tukey HSD Tukey-Kramer	P(Type I Error)$=\alpha$ P(Type I Error)$\leq\alpha$	Moderate	**Most recommended**; more powerful than the Bonferroni or Šidák corrections, and the Scheffé test
Dunnett	P(Type I Error)$=\alpha$	Moderate	Used to compare the samples with a control group
SNK: Student-Newman-Keuls	P(Type I Error)$=\alpha$	Moderate	Increasingly less used
Duncan	Moderate	Low	Little used at present; generally, it produces identical results to independent t-tests

See also: Hilton, A., & Armstrong, R. A. (2006). Statnote 6: post-hoc ANOVA tests. *Microbiologist*, *2006*, 34-36.

Multiple comparison tests

- **Alternative nonparametric tests**
  - **Wilcoxon-Mann-Whitney test** (i.e. Mann-Whitney U test) **with the Bonferroni correction**, or with the Dunn-**Sidak correction**, to adjust for the inflation of type I error. This approach is conservative and less powerful than alternative tests, thus should only be used for a small set of comparisons.
    - **Bonferroni correction** to the p-values of multiple tests: multiply the number of multiple comparisons by each p-value. If the resulting p-value is greater than 1 then round to 1.
  - **Dunn's test** is one of the least powerful and it can be very conservative, especially for larger numbers of comparisons

Multiple comparison tests

- **Alternative nonparametric tests**
  - **Hodges-Lehmann test for independent samples**
  - **Steel-Dwass test** (or Dwass-Steel-Critchlow-Fligner test)
  - **Nemenyi test** (or Nemenyi-Damico-Wolfe-Dunn test)
  - **Dunnet's T3 or C procedures** recommended if the df is small or large, respectively
  - **Games-Howell test** does not rely on equal variances and sample sizes nor on normality (it is often recommended)

Multiple comparison tests

- **Example 1**
  - H_0: $\mu_i = \mu_j$ *vs* H_1: $\mu_i \neq \mu_j$ $(i \neq j)$
  - S^2=MSE=0.5811 ; $q_{(K;\, n-k;\, 1-\alpha)} = q_{(4;\, 18;\, 0.95)} = 3.997$ (table)
  - Decision rules:
    - A *vs* B: Wobs = 3.064 < 3.997 $\Rightarrow$ not reject H_0
    - A *vs* C: Wobs = 0.613 < 3.997 $\Rightarrow$ not reject H_0
    - A *vs* D: Wobs = 3.813 < 3.997 $\Rightarrow$ not reject H_0
    - B *vs* C: Wobs = 2.571 < 3.997 $\Rightarrow$ not reject H_0
    - B *vs* D: Wobs = 0.919 < 3.997 $\Rightarrow$ not reject H_0
    - C *vs* D: Wobs = 3.370 < 3.997 $\Rightarrow$ not reject H_0
  - There is no evidence of differences between the mean behaviour of the institutes regarding the increasing of costs [***A vs D*: p-value=0,065**]

Tests to the equality of k variances

- **Tests to the equality of *k* variances (independent samples)**
  - $K \geq 2$ mutually independent random samples with sizes n_i (i=1,...,k) such that $n_1+n_2+ \dots +n_k=n$, from populations with **Normal** distribution
  - **Hypotheses**
    - H_0: $\boldsymbol{\sigma_1^2 = \sigma_2^2 = \dots = \sigma_K^2 = \sigma^2}$
    - H_1: $\exists_{i,j\ (i \neq j)}\ \boldsymbol{\sigma}_i^2 \neq \boldsymbol{\sigma}_j^2$
  - **Bartlett's test**: should not be applied if there are doubts about the normality
  - **Levene's test**: less sensitive to deviations from the normality condition
    - It consists in applying the analysis of variance to a new variable Z_{ij} that corresponds to the absolute deviations between the X_{ij} study variable and the mean, or median, from the respective group

Tests to the equality of k variances

- **Bartlett's test**
  - **Test statistic**

$$Q=\frac{(n-k)\ln S^2-\sum_{i=1}^{k}(n_i-1)\ln S_i^2}{1+\frac{1}{3(k-1)}\left[\sum_{i=1}^{k}\left(\frac{1}{n_i-1}\right)-\frac{1}{n-k}\right]}\sim\chi^2_{(k-1)}$$

S^2 = MSE

  - **Decision rule**
    - Reject H_0 when $Q_{Obs} \geq \chi^2_{(k-1;\,1-\alpha)}$

Tests to the equality of k variances

- ## Levene's test

  - **Test statistic**

$$F = \frac{\sum_{i=1}^{k} n_i \left(\bar{Z}_{i\bullet} - \bar{Z}\right)^2 \Big/ (k-1)}{\sum_{i=1}^{k}\sum_{j=1}^{n_i} \left(Z_{ij} - \bar{Z}_{i\bullet}\right)^2 \Big/ (n-k)} \sim F_{(k-1;n-k)}$$

$$Z_{ij} = \left|X_{ij} - \bar{X}_{i\bullet}\right|,\ \bar{X}_{i\bullet} \rightarrow \text{mean of the i group}$$

$$\bar{Z}_{i\bullet} = \sum_{j=1}^{n_i} \frac{Z_{ij}}{n_i} \rightarrow \text{mean of the i group}$$

$$\bar{Z} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} Z_{ij} = \frac{1}{n}\sum_{i=1}^{k} n_i \bar{Z}_{i\bullet} \rightarrow \text{global mean}$$

    - If there is a suspicion that the population distribution is not Normal:

$$Z_{ij} = \left|X_{ij} - \tilde{X}_{i\bullet}\right|,\ \tilde{X}_{i\bullet} \rightarrow \text{median of the i group}$$

  - **Decision rule**: Reject H_0 when $F_{Obs} \geq F_{(k-1,\ n-k;\ 1-\alpha)}$

Tests to the equality of k variances

- **Example 1**
  - **Bartlett's test**

S^2 = MSE = 0.5811

$$Q_{obs} = \frac{(22-4)\ln(0.5811) - \left[4\ln(0.625) + 5\ln(0.556) + 5\ln(0.648) + 4\ln(0.485)\right]}{1 + \frac{1}{3(4-1)}\left[\frac{1}{4} + \frac{1}{5} + \frac{1}{5} + \frac{1}{4} - \frac{1}{22-4}\right]}$$

  - **Decision**
    - $Q_{Obs} = 0.0985$ **<** $\chi^2_{(3;\,0.95)} = 7.815 \Rightarrow$ Not reject H_0 at the 5% significance level
    - p-value = $P[\chi^2_{(3)} > 0.0985\,] = 0.992$
    - There is no evidence that the variances of the groups are different

Applications in Excel

- **Example 2**
  - Solve Example 1 using the Real Statistics Resource Pack
    - Compute descriptive statistics and the Shapiro-Wilk test for normality
    - Compute the Levene's test for the equality of k=4 variances
    - Compute the ANOVA
    - Compute the Tukey-Kramer test for multiple comparisons

Applications in Excel

- **Example 3**
  - A sample of five law firms has been selected from each of three major cities for the purpose of obtaining a quote for legal services on a relatively routine business contract. The quotes provided by the fifteen firms are in the **Example3** sheet of the LU7_Examples Excel file. Are the population mean bids from the cities' firms identical?

    1. What is the factor? How many levels does it have? What are the experimental units?
    2. Test the normality assumption
    3. Test the homoscedasticity assumption
    4. Conclude based on appropriate test(s)

Analysis of Variance (ANOVA)

Do the homework*!*

Ana Cristina Costa, ccosta@novaims.unl.pt