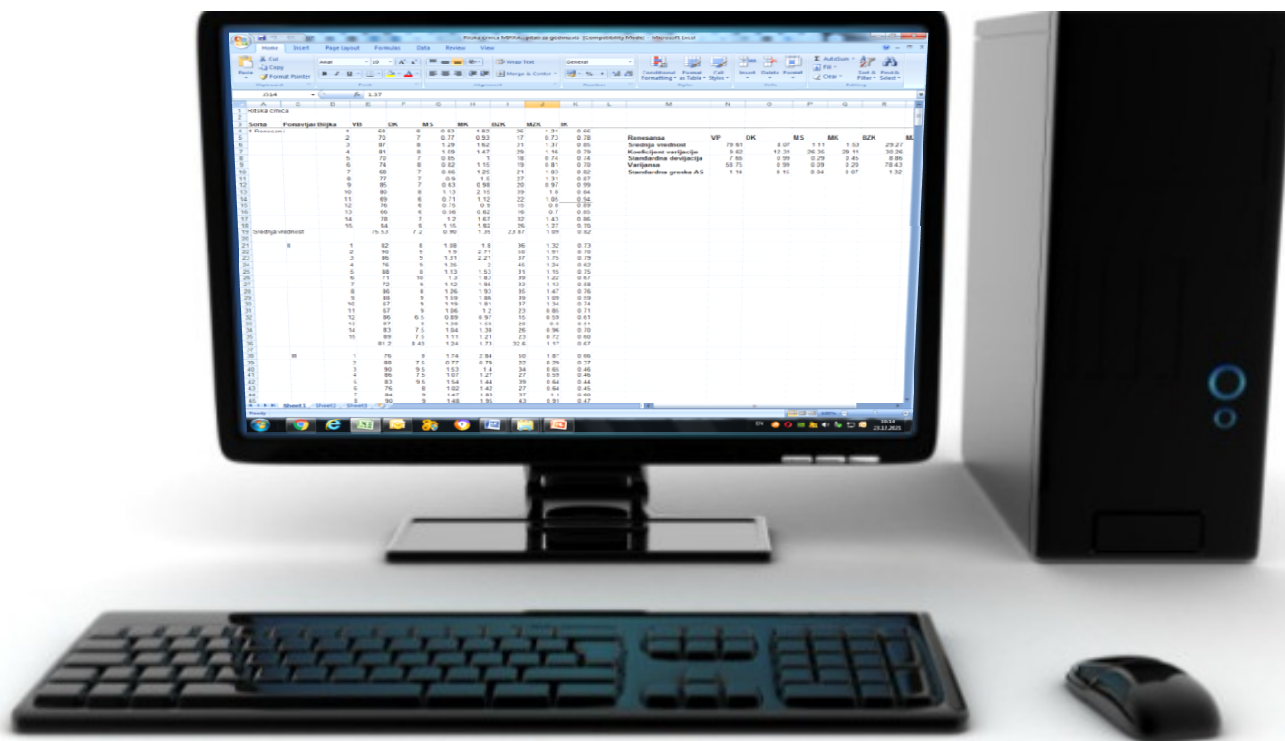


ОБРАДА ПОДАТАКА



Бројчани подаци добијени експериментом су често веома бројни, па је неопходно да се среде, обраде различитим математичким и статистичким моделима анализе.

На овај начин могу да се:

- дају одговарајућа тумачења добијених резултата
- утврде разлике које су настале под дејством примењених третмана
- установе узрочно-последичне повезаности резултата
- на крају, тестира се и постављена радна хипотеза

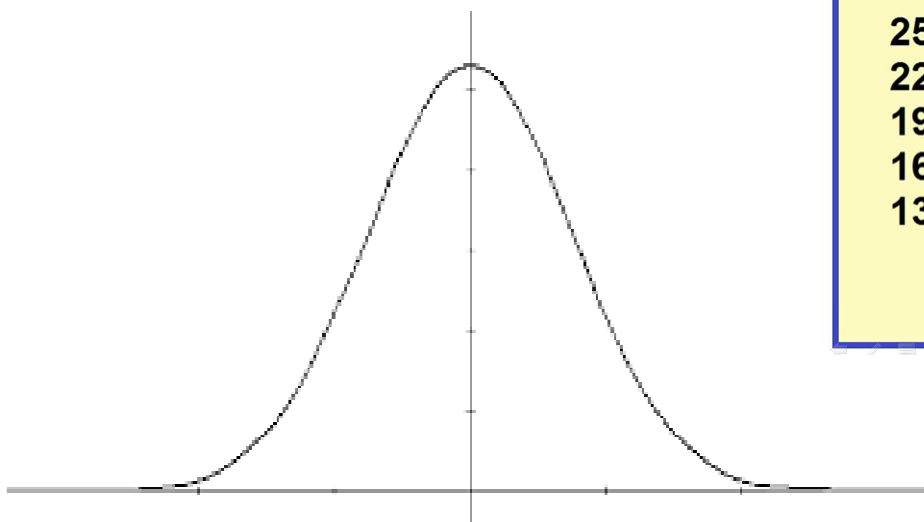
The screenshot displays a Microsoft Excel spreadsheet with the following structure:

- Columns:** A-Z, with data starting from column M.
- Rows:** 4-55, grouped by year and treatment.
- Year Groups:** 2006 KUMANE (rows 6-11, 13-18), 2007 KUMANE (rows 19-24), 2005 KUMANE (rows 25-30), 2006 KUMANE (rows 31-36), 2007 KUMANE (rows 37-42), 2005 KUMANE (rows 43-48), 2006 KUMANE (rows 49-54).
- Treatment Groups:** 1 Renesansa (rows 7, 14, 21, 28, 35, 42, 49), 8. Pobeda (rows 26, 33, 40), 3 EVROPA 90 (rows 44, 51, 54).
- Sub-headers:** 'Broj zrna po klasu' and 'Masa zrna po klasu [g]'.
- Data Tables:** Each sub-header is followed by a table with columns for 'Kontrola' and 'Kumane'. The 'Kontrola' table has 3 columns (1, 2, 3) and the 'Kumane' table has 5 columns (1, 2, 3, 4, 5).
- Footer:** 'Ready' and system tray showing '10:18 23.12.2021'.

СТАТИСТИЧКА ОБРАДА ПОДАТАКА

Мере централне тенденције

Средња вредност-аритметичка средина



Из негруписаних података

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

Табела дистрибуције фреквенција груписаних података

Интервал	Фреквенција (f)	Средина интервала (x)
25 - 27	5	26
22 - 24	7	23
19 - 21	14	20
16 - 18	11	17
13 - 15	3	14
Укупно:	40	

18 13 16 21
~~20~~ 18 ~~23~~ 17
20 22 24 23
~~20~~ ~~28~~ 17 16
20 ~~28~~ 25 21
~~21~~ 19 24 17
20 22 18 28
~~17~~ 15 ~~20~~ 16
~~19~~ 18 26 23
20 27 15 19

Из груписаних података

$$\bar{X} = \frac{\sum fx}{\sum f}$$

Мере варијације

Статистички параметар	Формула за негруписане податке	Формула за груписане податке
Стандардна грешка средње вредности [$s_{\bar{x}}$]	$s_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$	$s_{\bar{x}} = \frac{\sigma}{\sqrt{\sum f}}$
Стандардна девијација [σ]	$\sigma = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{\sum N}}{\sum N - 1}}$	$\sigma = \sqrt{\frac{\sum fx^2 - \frac{(\sum fx)^2}{\sum f}}{\sum f - 1}}$
Варијанса [σ^2]	$\sigma^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{\sum N}}{\sum N - 1}$	$\sigma^2 = \frac{\sum fx^2 - \frac{(\sum fx)^2}{\sum f}}{\sum f - 1}$
Коефицијент варијације [V]	$V = \frac{\sigma}{\bar{x}} * 100 [\%]$	

Тест хомогености варијанси

Ако оригинални подаци указују на велике разлике између третмана, то значи да се можда ради о:

- неједнаким варијансама
- узорцима који потичу из врло различитих скупова
- сувише хетерогеном материјалу

У тим случајевима се примењује тест хомогености варијанси.

- Када тест покаже да се ради о истим варијансама, примењује се анализа варијансе.
- Када тест укаже на неједнакост варијанси, неопходно је да се уради трансформација оригиналних података.

Тестирање хомогености варијанси може да се уради различитим тестовима, али најчешће се користи

Бартлетов тест (Bartlett)

$$\chi_{\text{raw}}^2 = \left[\sum_{i=1}^{i=a} (n_i - 1) \right] \ln \bar{\sigma}^2 - \sum_{i=1}^{i=a} (n_i - 1) \ln \sigma_i^2$$

$$\text{Correction factor} = 1 + \frac{1}{3(a-1)} \left[\sum_{i=1}^{i=a} \frac{1}{n_i - 1} - \frac{1}{\sum_{i=1}^{i=a} (n_i - 1)} \right]$$

$$\chi_{\text{corrected}}^2 = \frac{\chi_{\text{raw}}^2}{\text{Correction factor}}$$



Maurice Stevenson Bartlett (1910- 2002)

Када је израчуната вредност χ^2 мања од табличне вредности P , за одређени ниво значајности, прихвата се хипотеза о хомогености варијанси.

У случају да је израчуната вредност χ^2 већа од табличне вредности, не прихвата се хипотеза о хомогености варијанси и подаци се трансформишу.

Трансформација података

- у случајевима када нису задовољене претпоставке за анализу варијансе
- трансформација се врши кореновањем вредности (добитених података)

Израчунавање изгубљених података

- ако дође до губљења података експеримента (оштећење огледа)
- да изгубљени податак не би умањио вредност огледа, израчунава се вредност изгубљеног податка (парцеле, индивидуе)
- израчунавање се обавља по одговарајућој формули

$$X = \frac{tT + bB - S}{(t-1)(b-1)}$$

Где је:

X – вредност изгубљене парцеле

t – број третмана

b – број блокова (понављања)

T – сума вредности третмана са изгубљеном парцелом

B – сума вредности блока са изгубљеном парцелом

S – сума вредности свих познатих парцела

Тестирање разлика

- да би се установило дејство примењених третмана, упоређују се и тестирају разлике њихових средњих вредности
- за ово тестирање најчешће се користи t – тест разлика аритметичких средина
- значајност разлика се тестира применом табеле t - дистрибуције

T- тест независних узорака

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}}$$

Разлика аритметичких средина два третмана

Стандардна грешка разлике аритметичких средина

$$S_{\bar{x}_1 - \bar{x}_2} = \sqrt{\left(\frac{N_1 s_1^2 + N_2 s_2^2}{N_1 + N_2 - 2} \right) \left(\frac{N_1 + N_2}{N_1 N_2} \right)}$$

N_1, N_2 – број мерених индивидуа у третманима 1 и 2
 s_1^2, s_2^2 – стандардне девијације (варијансе) третмана 1 и 2

Узрочно-последичне везе

- поред утврђивања вредности одређених третмана и параметара, потребно је да се утврде везе између појава и особина

Рачунарски програми

Statistica, GenStat, Excel, MSTATC