

2.6. ANALIZA VARIANCE

- Analizo variance (AVAR oz. ANOVA – Analysis of Variance) uporabimo, kadar želimo:

- Primerjati več kot dve populaciji.

- Ali se vsebnost PCB v vodi iz Krupe, Kolpe in Lahinja razlikuje?

$$H_0: \mu_{\text{Krupa}} = \mu_{\text{Kolpa}} = \mu_{\text{Lahinja}}$$

$$H_1: \mu_{\text{Krupa}} \neq \mu_{\text{Kolpa}} \neq \mu_{\text{Lahinja}}$$

- Ob predhodno načrtovanem vzorčenju ugotoviti, kaj vse vpliva na opazovano izmerjeno vrednost vzorca – kakšni so prispevki posameznih dejavnikov k celotni varianci vzorca.

- Kako na spremenljivost nivoja podtalnice v Sloveniji vpliva za katero vodno telo gre (npr. Dravsko, Mursko, Savsko), letni čas vzorčenja (pomlad, poletje, jesen, zima) in način vzorčevanja (avtomatični odčitek, ročna meritev) ter kolikšen delež variance ostane nepojasnen (analizna napaka ε)?

$$\sigma^2_{\text{celotna}} = \sigma^2_{\text{vodno telo}} + \sigma^2_{\text{letni čas}} + \sigma^2_{\text{vzorčenje}} + \varepsilon$$

2.6. ANALIZA VARIANCE



- Spremenljivost podatkov obravnavamo v obliki popravljenе vsote kvadratov (VK, SS – sum of squares) in ne z neposrednimi izračuni varianc.
- Vsote kvadratov, ki pripadajo posameznemu vplivu (faktorju, nivoju) testiramo s statistiko F, da ugotovimo, ali so razlike resnične (statistično značine na določeni ravni zaupanja) ali naključne.



2.6. ANALIZA VARIANCE

- H_0 , da ni razlik med variancami zavrնemo, kadar izračunana vrednost presega tabelirano kritično vrednost $F_{\alpha, v1, v2}$.
- Obravnavani podatki morajo biti neodvisna opazovanja, izhajajoča iz normalne porazdelitve.

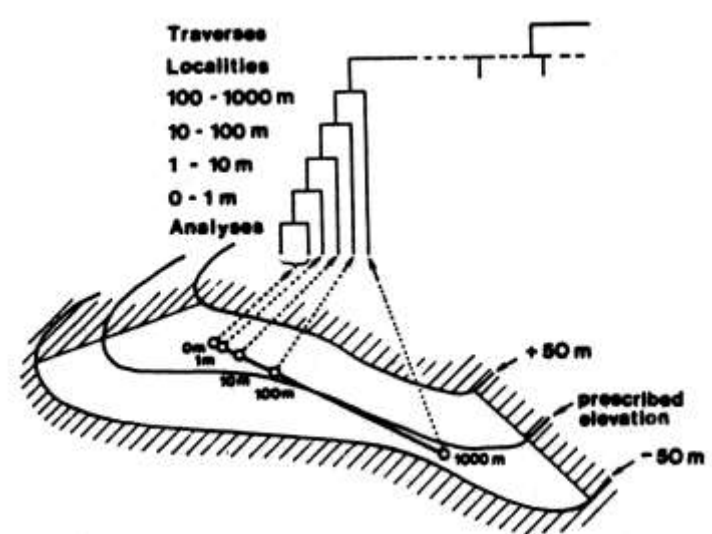
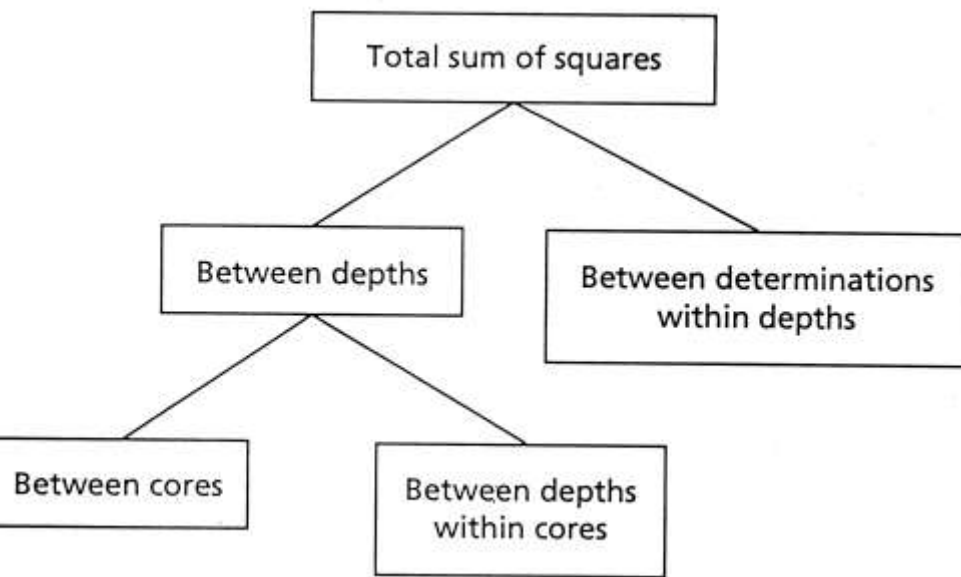
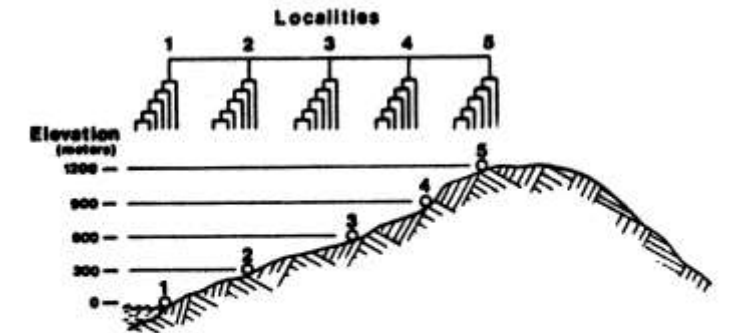
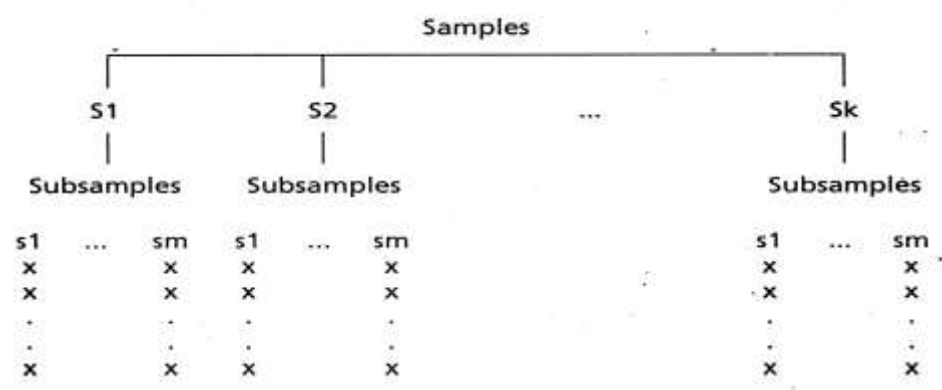
2.6.1. Randomizacija (naključnost) poizkusa

- Nenadzorovana spremenljivost poizkusne snovi lahko vodi do pristranske ocene variance.
- Takšno tveganje preprečimo s homogenizacijo vzorca – primerke pošljemo v analizo v naključnem vrstnem redu, s čimer "napako meritve" enakomerno porazdelimo med vsa opazovanja.
- Natančnost analitke dodatno preverjamo s ponovitvami (replikami) opazovanj – naključno izbran primerek damo v analizo dvakrat pod različnimi zaporednimi številkami.
- Točnost analitike preverimo tako, da v analitiko pošljemo tudi kakega od standardov.

2.6.2. Modeli AVAR

- Glede na to, koliko virov variance obravnavamo in kako jih določimo, ločimo več vrst AVAR:
 - Model določenih učinkov – vnaprej določimo nivoje (vplive, faktorje), ki nas zanimajo.
 - Enojna (one-way) – primerjava po stolpcih
 - Dvojna (two-way) – primerjava po stolpcih in vrsticah
 - Model naključnih učinkov – nivoji so naključni in vseh možnih vplivov na varianco niti ne poznamo.
 - Grajena (hierarhična, nested)
 - Uravnotežena
 - Neuravnotežena

2.6.2. Modeli AVAR





2.6.2.1. Enojna AVAR

■ Primer:

Vsebnost Au smo določali s tremi različnimi metodami. Z vsako metodo smo isti vzorec analizirali po štirikrat. Ali dajejo vse tri enake rezultate (so enako dobre)?

	AA	ES	NA
M1	10	13	33
M2	15	16	40
M3	12	14	32
M4	12	13	30
\bar{x}	12,25	14	33,75
s^2	4,25	2	18,92
Σx	49	56	135
Σx^2	613	790	4613

2.6.2.1. Enojna AVAR

- Matriko podatkov posplošeno zapišemo:

kategorija	1	2	...	k	Skupaj
1	x_{11}	x_{21}	...	x_{k1}	
2	x_{12}	x_{22}	...	x_{k2}	
:	:	:	...	:	
n	x_{1n}	x_{2n}	...	x_{kn}	
$\sum_{i=1}^n x_{ij}$ Vsota stolpcev	$\sum_i^{n_1} x_{1j}$	$\sum_i^{n_2} x_{2j}$		$\sum_i^{n_k} x_{ij}$	$\sum_{j=1}^k \sum_{i=1}^n x_{ij}$
$\sum_{i=1}^n x_{ij}^2$ Vsota kvadratov	$\sum_i^{n_1} x_{1j}^2$	$\sum_i^{n_2} x_{2j}^2$		$\sum_{i=1}^{n_k} x_{ij}^2$	$\sum_{j=1}^k \sum_{i=1}^n x_{ij}^2$
Št. opazovanj	n_1	n_2	...	n_k	$N = nk$



2.6.2.1. Enojna AVAR

- Model zapišemo: $x_{ij} = \mu_j + \alpha_j + \varepsilon_{ij}$
kjer je α razlika od skupne srednje vrednosti, ki je posledica obravnavanega dejavnika in ε naključni člen – napaka.
- $H_0: \mu_1 = \mu_2 = \mu_k$ in $\alpha_1 = \alpha_2 = \alpha_k$ Vse metode dajejo povprečno enake rezultate.
 $H_1: \mu_1 \neq \mu_2 \neq \mu_k$ in $\alpha_1 \neq \alpha_2 \neq \alpha_k$ Vsaj ena od metod daje drugačen rezultat.
- Test izvedemo tako, da iz ustreznih vsot kvadratov (VK, $\sum x^2_{ij}$) izračunamo testni F.
- Običajno delamo z računalniškim programom (Excel, CSS Statistica, SPSS,...), ki rezultate poda v tabeli analize variance.
- **Pozor: Različni računalniški programi zahtevajo različen vnos podatkov!**

2.6.2.1. Enojna AVAR

- Splošna shema tabele enojne analize variance je:

Vir variance	Vsota kvadratov (VK, SS)	Stopnje prostosti (v, d.f.)	Varianca - povprečje VK (\overline{VK}) VK/ v	F	$F_{\alpha, k-1, N-k}$
Med stolpci (povprečji)	VK_M (SS_B)	k-1	$s^2_M = VK_M / (k-1)$	s^2_M / s^2_Z	tabele
Znotraj stolpcev	VK_Z (SS_Z)	N-k	$s^2_Z = VK_Z / (N-k)$		
Skupaj	VK_C (SS_T)	N-1			

2.6.2.1. Enojna AVAR

■ Izračun vsote kvadratov:

$$VK_M = \frac{\sum_{j=1}^k \left(\sum_{i=1}^n x_{ij} \right)^2}{n_i} - \frac{\left(\sum_{j=1}^k \sum_{i=1}^n x_{ij} \right)^2}{N}$$

$$VK_Z = \sum_{j=1}^k \sum_{i=1}^n x_{ij}^2 - \frac{\sum_{j=1}^k \left(\sum_{i=1}^n x_{ij} \right)^2}{n_i}$$

$$VK_C = \sum_{j=1}^k \sum_{i=1}^n x_{ij}^2 - \frac{\left(\sum_{j=1}^k \sum_{i=1}^n x_{ij} \right)^2}{N} = VK_M + VK_Z$$

2.6.2.1. Enojna AVAR

	AA	ES	NA	
M1	10	13	33	
M2	15	16	40	
M3	12	14	32	
M4	12	13	30	
Σx	49	56	135	240
Σx^2	613	790	4613	6016

$$VK_M = \frac{\sum_{j=1}^k \left(\sum_{i=1}^n x_{ij} \right)^2}{n_i} - \frac{\left(\sum_{j=1}^k \sum_{i=1}^n x_{ij} \right)^2}{N}$$

$$VK_Z = \sum_{j=1}^k \sum_{i=1}^n x_{ij}^2 - \frac{\sum_{j=1}^k \left(\sum_{i=1}^n x_{ij} \right)^2}{n_i}$$

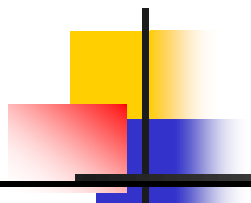
$$VK_C = \sum_{j=1}^k \sum_{i=1}^n x_{ij}^2 - \frac{\left(\sum_{j=1}^k \sum_{i=1}^n x_{ij} \right)^2}{N} = VK_M + VK_Z$$

$$VK_M = \frac{49^2 + 56^2 + 135^2}{4} - \frac{240^2}{12} = 5940,5 - 4800 = 1140,5$$

$$VK_Z = 6016 - 5940,5 = 75,5$$

$$VK_C = 6016 - 4800 = 1216$$

2.6.2.1. Enojna AVAR



Vir variance	VK	ν	Varianca - VK/ν	F	$F_{\alpha,k-1,N-k}$
Med stolpci	VK_M	$k-1$	$s^2_M = VK_M/(k-1)$	s^2_M/s^2_Z	tabele
Znotraj stolpcev	VK_Z	$N-k$	$s^2_Z = VK_Z/(N-k)$		
Skupaj	VK_C	$N-1$			

Vir variance	VK	ν	Varianca	F	$F_{\alpha,k-1,N-k}$
Med	1140,5	2	570,2	67,94***	4,26 8,02 16,39
Znotraj	75,5	9	8,39		
Skupaj	1216	11			



Arial 10 B I

E26

A B C

1		AA	ES
2	meritev 1	10	13
3	meritev 2	15	16
4	meritev 3	12	14
5	meritev 4	12	13

- Spelling... F7
- Research... Alt+Click
- Error Checking...
- Speech
- Shared Workspace...
- Share Workbook...
- Track Changes
- Euro Conversion...
- Compare and Merge Workbooks...
- Protection
- Online Collaboration
- Solver...
- Goal Seek...
- Scenarios...
- Formula Auditing
- Macro
- Add-Ins...
- AutoCorrect Options...
- Customize...
- Options...
- Conditional Sum...
- Lookup...
- Data Analysis...

Σ 100%

2

0.00 0.00

H I J K L

26

27

28

	A	B	C	D	E	F	G	H	I	J
1		AA	ES	NA						
2	meritev 1	10	13	33						
3	meritev 2	15	16	40						
4	meritev 3	12	14	32						
5	meritev 4	12	13	30						
6										
7										
8										
9										

Data Analysis

Analysis Tools

Anova: Single Factor
Anova: Two-Factor With Replication
Anova: Two-Factor Without Replication
Correlation
Covariance
Descriptive Statistics
Exponential Smoothing
F-Test Two-Sample for Variances
Fourier Analysis
Histogram

OK

Cancel

Help



	A	B	C	D	E	F	G	H	I	J
1		AA	ES	NA						
2	meritev 1	10	13	33						
3	meritev 2	15	16	40						
4	meritev 3	12	14	32						
5	meritev 4	12	13	30						
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										

Anova: Single Factor

Input

Input Range:

\$B\$1:\$D\$5

Grouped By:

 Columns Rows Labels in First Row

Alpha: 0,05

Output options

 Output Range: New Worksheet Ply: New Workbook

OK

Cancel

Help

E12 67,976821192053

	A	B	C	D	E	F	G	H	I	
1	Anova: Single Factor									
2										
3	SUMMARY									
4	<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>					
5	AA	4	49	12,25	4,25					
6	ES	4	56	14	2					
7	NA	4	135	33,75	18,91667					
8										
9										
10	ANOVA									
11	<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>			
12	Between Groups	1140,5	2	570,25	67,97682	3,7E-06	4,256495			
13	Within Groups	75,5	9	8,388889						
14										
15	Total	1216	11							
16										
17										
18										
19										
20										
21										
22										


 Resume... Ctrl+R

Add to Report

 Arial 10

Data: avar pwp 1t....

	1 metoda	2 Au
meritev 1az	AA	
meritev 2aa	AA	
meritev 3aa	AA	
meritev 4aa	AA	
meritev 1es	ES	
meritev 2es	ES	
meritev 3es	ES	
meritev 4es	ES	
meritev 1na	NA	
meritev 2na	NA	
meritev 3na	NA	
meritev 4na	NA	30

ANOVA

Multiple Regression

Nonparametrics

Distribution Fitting

Advanced Linear/Nonlinear Models

Multivariate Exploratory Techniques

Industrial Statistics & Six Sigma

Power Analysis

Data-Mining

Statistics of Block Data

STATISTICA Visual Basic

Probability Calculator

Data: avar pwp 1t....

	1 metoda	2 Au
meritev 1az	AA	10
meritev 2aa	AA	15
meritev 3aa	AA	12
meritev 4aa	AA	12
meritev 1es	ES	13
meritev 2es	ES	16
meritev 3es	ES	14
meritev 4es	ES	13
meritev 1na	NA	33
meritev 2na	NA	40
meritev 3na	NA	32
meritev 4na	NA	30

ANOVA/MANOVA One-Way ANOVA: avar pwp 1t.sta

Quick | Options

Variables

Dependent variables: Au

Categorical factor: metoda

Factor codes: selected

Between effect: metoda

OK

Cancel

Options

Syntax editor

File Edit View Insert Format Statistics Graphs Tools Data Workbook Window Help

Add to Workbook ▾ Add to Report ▾

Arial 10 **B** *I* U

Data: avar pwp 1t.sta

	1 metoda	2 Au
meritev 1a	AA	10
meritev 2a		
meritev 3a		
meritev 4a		
meritev 1b		
meritev 2b		
meritev 3b		
meritev 4b		
meritev 1c		
meritev 2c		
meritev 3c		
meritev 4c		

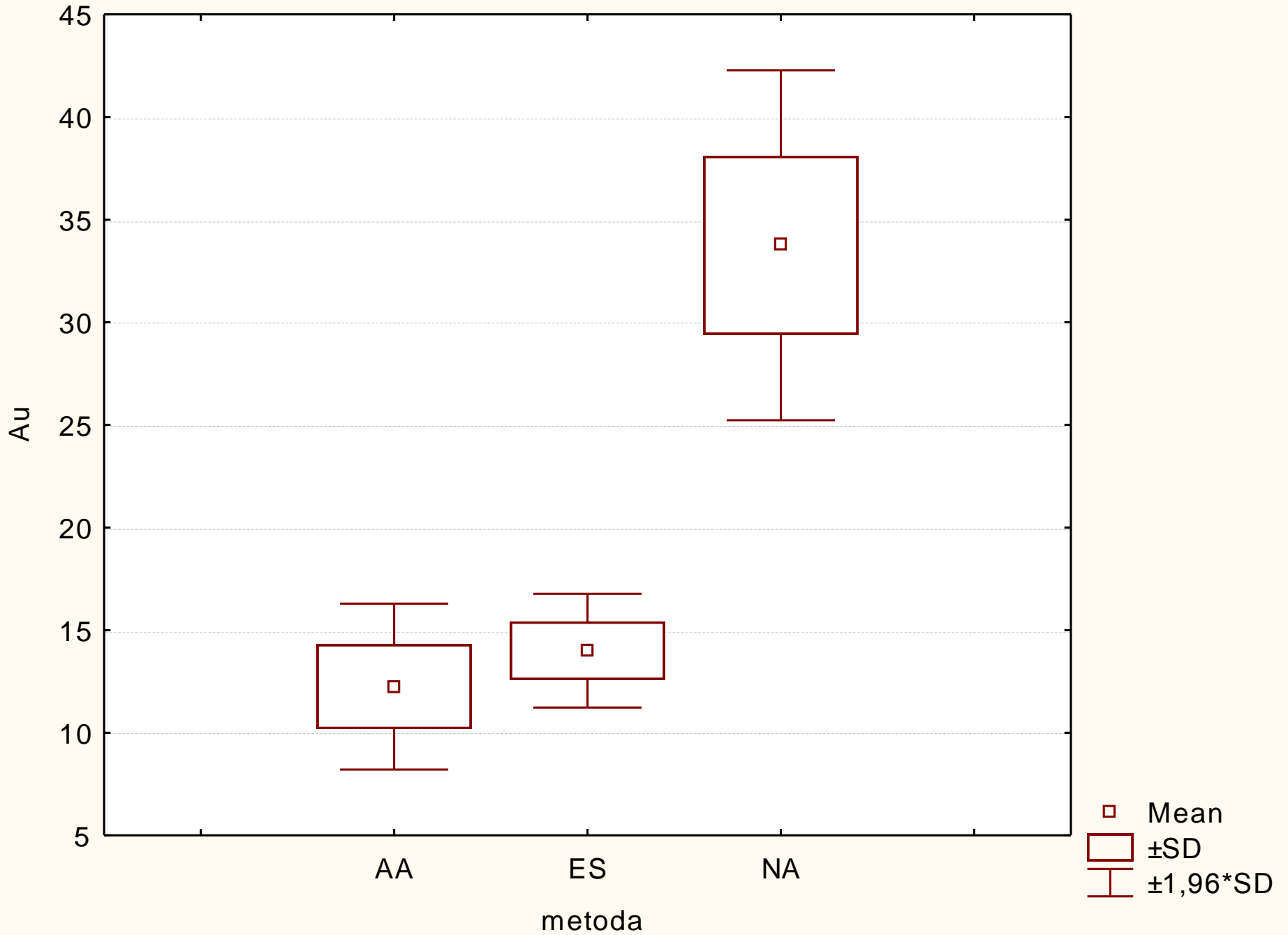
Workbook1* - Univariate Tests of Significance for Au (avar pwp 1t.sta)

Univariate Tests of Significance for Au (avar pwp 1t.sta)
Sigma-restricted parameterization
Effective hypothesis decomposition

Effect	SS	Degr. of Freedom	MS	F	p
Intercept	4800,000	1	4800,000	572,1854	0,000000
metoda	1140,500	2	570,250	67,9768	0,000004
Error	75,500	9	8,389		

Univariate Tests of Significance for Au (avar pwp 1t.sta)

Categ. Box & Whisker Plot: Au



2.6.2.1. Enojna AVAR

- Porazdelitev variance med posamezne nivoje:

$$\sigma_{celotna}^2 = \sigma_{med}^2 + \sigma_{znotraj}^2$$

$$\sigma_{znotraj}^2 = \sigma_{napaka}^2 = \sigma_{error}^2$$

$$s_e^2 = s_Z^2$$

$$s_M^2 = s_e^2 + n \cdot s_m^2$$

$$s_m^2 = \frac{s_M^2 - s_e^2}{n}$$

$$s_c^2 = s_m^2 + s_e^2 = 100\%$$

2.6.2.1. Enojna AVAR

$$s_e^2 = s_Z^2 = 8,39$$

$$s_M^2 = s_e^2 + n \cdot s_m^2 = 570,2$$

$$s_m^2 = \frac{s_M^2 - s_e^2}{n} = \frac{570,2 - 8,39}{4} = 140,45$$

$$s_c^2 = s_m^2 + s_e^2 = 148,84 = 100\% = 94,4\%^{***} + 5,6\%$$

- Končna interpretacija primera je, da smo 99,9% prepričani, da dajejo tri različne analitske metode različne rezultate. Preko 94% celotne variance odpade na razlike med njimi, 5,6 pa na naključno napako. Iz grafov je razvidno, da od ostalih dveh metod odstopa NA.

2.6.2.2. Dvojna AVAR

- Model zapišemo: $x_{ij} = \mu_j + \alpha_j + \beta_j + \varepsilon_{ij}$
kjer je α vpliv, ki ga ima na srednjo vrednost nivo (dejavnik), ki je zapisan v stolpcih, β spremenljivost med vrsticami. ε je naključni člen – napaka.
- $H_0: \mu_1 = \mu_2 = \mu_k$ in $\alpha_1 = \alpha_2 = \alpha_k$ in $\beta_1 = \beta_2 = \beta_k$
 $H_1: \mu_1 \neq \mu_2 \neq \mu_k$ in $\alpha_1 \neq \alpha_2 \neq \alpha_k$ in $\beta_1 \neq \beta_2 \neq \beta_k$
- Celotno varianco razdelimo na del, ki so ga prispevali stolpci (vsota kvadratov med stolpci – VK_M) in del, ki je posledica spremenljivosti znotraj stolpcev, torej med vrsticami (vsota kvadratov med vrsticami – VK_Z). Del variance, ki ostane nepojasnen je napaka.

2.6.2.2. Dvojna AVAR

■ Primer: Vsebnost Au smo določali s tremi različnimi metodami. Z vsako metodo smo isti vzorec analizirali po štirikrat. Ali dajejo vse tri enake rezultate in koliko se med seboj razlikujejo štirje vzorci?

	AA	ES	NA	Σx_j	Σx_j^2
M1	10	13	33	56	1358
M2	15	16	40	71	2081
M3	12	14	32	58	1364
M4	12	13	30	55	1213
Σx_i	49	56	135	240	
Σx_i^2	613	790	4613		6016

2.6.2.2. Dvojna AVAR

- Matriko podatkov posplošeno zapišemo:

Kategorija I Kategorija II	1	2	...	s	Vsota vrstic $\sum_{j=1}^k x_{ij}$	Vsota kvadratov vrstic $\sum_{j=1}^k x_{ij}^2$
1	x_{11}	x_{21}	...	x_{s1}	$\sum_j^{k_1} x_{i1}$	$\sum_j^{k_1} x_{i1}^2$
2	x_{12}	x_{22}	...	x_{s2}		
:	:	:	...	:		
v	x_{1n}	x_{2n}	...	x_{sv}	$\sum_j^{kn} x_{ij}$	$\sum_{j=1}^{k_n} x_{ij}^2$
Vsota stolpcev $\sum_{i=1}^n x_{ij}$	$\sum_i^{n_1} x_{1j}$	$\sum_i^{n_2} x_{2j}$		$\sum_{i=1}^{n_k} x_{ij}$	$\sum_{j=1}^k \sum_{i=1}^n x_{ij}$	
Vsota kvadratov stolpcev $\sum_{i=1}^n x_{ij}^2$	$\sum_i^{n_1} x_{1j}^2$	$\sum_i^{n_2} x_{2j}^2$		$\sum_{i=1}^{n_k} x_{ij}^2$		$\sum_{j=1}^k \sum_{i=1}^n x_{ij}^2$

2.6.2.2. Dvojna AVAR

- Splošna shema tabele dvojne analize variance je:

Vir variance	Vsota kvadratov	Stopnje prostosti	Varianca - povprečje VK	F	F _{tabeliran}
Stolpci (I)	VK _I (SS _C)	k-1	$s^2_I = VK_I / (k-1)$	s^2_I / s^2_O	$F_{\alpha, k-1, (k-1)(n-1)}$
Vrstice (II)	VK _{II} (SS _R)	n-1	$s^2_{II} = VK_{II} / (n-1)$	s^2_{II} / s^2_O	$F_{\alpha, n-1, (k-1)(n-1)}$
Ostanek	VK _O (SS _R)	(k-1)(n-1)	$s^2_O = VK_O / ((n-1)(k-1))$		
Skupaj	VK _C (SS _T)	N-1			

2.6.2.2. Dvojna AVAR

■ Izračun vsote kvadratov:

$$VK_I = \frac{\sum_{j=1}^k \left(\sum_{i=1}^n x_{ij} \right)^2}{n_i} - \frac{\left(\sum_{j=1}^k \sum_{i=1}^n x_{ij} \right)^2}{N}$$

$$VK_{II} = \frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} \right)^2}{k_j} - \frac{\left(\sum_{j=1}^k \sum_{i=1}^n x_{ij} \right)^2}{N}$$

$$VK_O = VK_C - VK_I - VK_{II}$$

$$VK_C = \sum_{j=1}^k \sum_{i=1}^n x_{ij}^2 - \frac{\left(\sum_{j=1}^k \sum_{i=1}^n x_{ij} \right)^2}{N} = VK_I + VK_{II} + VK_O$$

2.6.2.2. Dvojna AVAR

$$VK_I = \frac{49^2 + 56^2 + 135^2}{4} - \frac{240^2}{12} = 5940,5 - 4800 = 1140,5$$

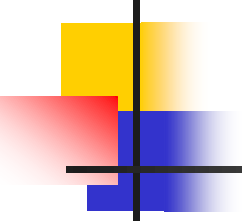
$$VK_{II} = \frac{56^2 + 71^2 + 58^2 + 55^2}{3} - \frac{240^2}{12} = 4855,3 - 4800 = 55,3$$

$$VK_C = 6016 - 4800 = 1216$$

$$VK_O = 1216 - 1140,5 - 55,3 = 20,2$$

	AA	ES	NA	Σx_j	Σx_j^2
M1	10	13	33	56	1358
M2	15	16	40	71	2081
M3	12	14	32	58	1364
M4	12	13	30	55	1213
Σx_i	49	56	135	240	
Σx_i^2	613	790	4613		6016

2.6.2.2. Dvojna AVAR



Vir variance	Vsota kvadratov	Stopnje prostosti	Varianca - povprečje VK	F	F _{tabeliran}
Stolpci (I)	1140,50	2	570,25	169,66***	5,14
Vrstice (II)	55,33	3	18,44	5,49*	4,76
Ostanek	20,17	6	3,36		
Skupaj	1216	11			

2.6.2.2. Dvojna AVAR

- Porazdelitev variance med posamezne nivoje:

$$\sigma_{celotna}^2 = \sigma_{stolpci}^2 + \sigma_{vrstice}^2 + \sigma_{napaka}^2$$

$$s_e^2 = s_O^2$$

$$s_I^2 = s_e^2 + n \cdot s_{stolpci}^2$$

$$s_{stolpci}^2 = \frac{s_I^2 - s_O^2}{n}$$

$$s_{II}^2 = s_e^2 + k \cdot s_{vrstice}^2$$

$$s_{vrstice}^2 = \frac{s_{II}^2 - s_O^2}{k}$$

$$s_c^2 = s_{stolpci}^2 + s_{vrstice}^2 + s_e^2 = 100\%$$

2.6.2.2. Dvojna AVAR

$$s_e^2 = s_o^2 = 3,361$$

$$s_{stolpci}^2 = \frac{s_I^2 - s_o^2}{n} = \frac{570,25 - 3,361}{4} = 141,72$$

$$s_{vrstice}^2 = \frac{s_{II}^2 - s_o^2}{k} = \frac{18,444 - 3,361}{3} = 5,03$$

$$s_c^2 = s_{stolpci}^2 + s_{vrstice}^2 + s_e^2 = 100\% = 150,11$$

$$s_c^2 = 94,4^{***} + 3,4^* + 2,2$$

- Zaključimo lahko, da smo 99,9% prepričani, da 94,4% vseh razlik povzroči izbrana analitska metoda. Med seboj se 95% statistično značilno razlikujejo tudi posamezne meritve, vendar je delež variance, ki nastane zaradi njih majhen – 3,4%. Analitiko ocenjujemo kot dobro, saj na analitsko napako oz. ostanek odpadeta le 2,2% celotne variance.

2.6.2.3. Grajena analiza variance



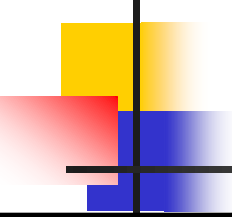
- V grajeni (hierarhični, nested) shemi vsak nivo prvega (najpomembnejšega) dejavnika razdelimo glede na drugi kriterij v različne nivoje podrejenega (drugega) faktorja.
- Shema je lahko uravnotežena (vedno so z meritvami zapolnjena vsa možna mesta v shemi) ali neuravnotežena (meritve v shemi lahko manjkajo).

2.6.2.3.1. Dvojna analiza variance s ponovitvami



- Najenostavnejša oblika hierarhične AVAR je dvojna analiza variance z n ponovitvami v vsaki celici.
- Ponovitve so lahko analitske, lahko pa na tak način uvedemo v shemo dodaten – tretji nivo.
- Običajno velja, da kadar je na končnem nivoju – nivoju napake, zbrano več kot 30% celotne variance, analitika (metoda, postopek) ni ustrezna. Takšne spremenljivke iz nadaljnih obravnav izvržemo.

2.6.2.3.1. Dvojna analiza variance s ponovitvami



	AA		ES		NA	
	Analiza 1	Analiza 2	Analiza 1	Analiza 2	Analiza 1	Analiza 2
M1	10	11	13	13	33	32
M2	15	12	16	15	40	39
M3	12	12	14	13	32	31
M4	12	11	13	14	30	30

- Matriko podatkov in tabelo analize varaince lahko posplošeno zapišemo:

Generalized computational scheme for a two-way analysis of variance with n repeated observations in each cell. T_{ij+} is the sum of the n observations for the i th category 1 and the j th category 2. x_{ijk} is the k th observation of the i th category 1 and j th category 2.

Category 2	Category 1					c	Row Totals
	1	2	3	.			
1	T_{11+}	T_{21+}	T_{31+}	...	T_{c1+}	T_{+1+}	
2	T_{12+}	T_{22+}	T_{32+}	...	T_{c2+}	T_{+2+}	
3	T_{13+}	T_{23+}	T_{33+}	...	T_{c3+}	T_{+3+}	
.	
r	T_{1r+}	T_{2r+}	T_{3r+}	...	T_{cr+}	T_{+r+}	
Column Totals	T_{1++}	T_{2++}	T_{3++}	...	T_{c++}	T_{+++}	

TWO-WAY ANALYSIS OF VARIANCE: MULTIPLE OBSERVATIONS

Source of variation	Sum of Squares	Degrees of freedom	Variance or mean square	F-ratio*
Col means	$S_c = \sum (T_{+j+}^2 / nc) - T'$	$c - 1$	$s_c^2 = S_c / (c-1)$	$F_c = s_c^2 / s_w^2$
Row means	$S_r = \sum (T_{i++}^2 / nr) - T'$	$r - 1$	$s_r^2 = S_r / (r-1)$	$F_r = s_r^2 / s_w^2$
Interaction	$S_I = S_S - S_c - S_r$	$(c-1)(r-1)$	$s_I^2 = S_I / (c-1)(r-1)$	$F_I = s_I^2 / s_w^2$
Sub-total	$S_S = \sum \sum (T_{ij+}^2 / n) - T'$	$rc - 1$		
Within	$S_w = S_T - S_S$	$rc(n-1)$	$s_w^2 = S_w / rc(n-1)$	
Total	$S_T = \sum \sum \sum x_{ijk}^2 - T'$	$nrc - 1$		

where $T' = T_{+++}^2 / nrc$

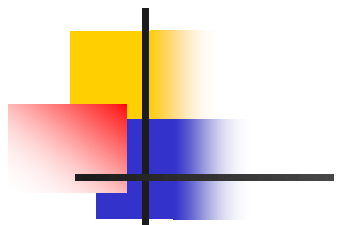
(*) S_I and S_w can be pooled to give $s_R^2 = (S_I + S_w) / (nrc - r - c + 1)$, if F_I is not significant. The F-ratios then become $F_c' = s_c^2 / s_R^2$ and $F_r' = s_r^2 / s_R^2$ with $nrc - r - c + 1$ d.f. for the denominator.

2.6.2.3.2. Hierarhična analiza variance



- Bolj zapletene sheme AVAR lahko rešujemo le s statističnimi računalniškimi programi kot so npr. CSS Statistica, SPSS ali Statpack.
- Potreben je ustrezen zapis podatkov ter pravilna izbira modela.
- Vedno model preverimo s podatki, za katere vemo, kakšen mora biti izid AVAR!

VPIS94	Y	X	KM25	KM5	KM1	KM02	PON	HG-MULJ	HG-TLA
1.	5392.50	5125.70	1.	1.	1.	1.	1.	215.	170.
2.	5387.46	5130.72	1.	2.	1.	1.	1.	80.	395.
3.	5388.50	5130.70	1.	2.	2.	1.	1.	90.	180.
4.	5388.12	5131.12	1.	2.	2.	2.	1.	90.	155.
5.	5392.50	5100.70	2.	1.	1.	1.	1.	370.	475.
6.	5392.50	5095.70	2.	2.	1.	1.	1.	170.	165.
7.	5393.50	5095.70	2.	2.	2.	1.	1.	100.	435.
8.	5393.70	5095.70	2.	2.	2.	2.	1.	165.	340.
9.	5402.50	5080.70	3.	2.	1.	1.	1.	80.	815.
10.	5404.50	5078.70	3.	2.	2.	1.	1.	90.	130.
11.	5404.30	5078.48	3.	2.	2.	2.	1.	135.	120.
12.	5404.30	5078.48	3.	2.	2.	2.	2.	150.	
13.	5417.50	5125.70	4.	1.	1.	1.	1.	155.	575.
14.	5427.50	5135.70	4.	2.	1.	1.	1.	95.	130.
15.	5429.65	5138.80	4.	2.	2.	1.	1.	20.	125.
16.	5429.12	5139.10	4.	2.	2.	2.	1.	30.	180.
17.	5417.50	5100.71	5.	1.	1.	1.	1.	510.	655.
18.	5422.50	5110.72	5.	2.	1.	1.	1.	1785.	850.
19.	5423.48	5112.74	5.	2.	2.	1.	1.	185.	345.
20.	5423.70	5113.08	5.	2.	2.	2.	1.	145.	130.
21.	5417.50	5075.71	6.	1.	1.	1.	1.	145.	185.
22.	5407.53	5085.72	6.	2.	1.	1.	1.	65.	185.
23.	5407.50	5084.70	6.	2.	2.	1.	1.	175.	595.
24.	5407.50	5084.30	6.	2.	2.	2.	1.	165.	125.
25.	5417.52	5050.71	7.	1.	1.	1.	1.	85.	325.
26.	5427.52	5050.71	7.	2.	1.	1.	1.	80.	200.
27.	5426.51	5050.70	7.	2.	2.	1.	1.	100.	230.
28.	5426.51	5050.70	7.	2.	2.	1.	2.	180.	
29.	5426.12	5051.10	7.	2.	2.	2.	1.	75.	370.



ONE-WAY HIERARCHICAL
ANALYSIS OF VARIANCE
USGS STATPAC

COLUMN NO 1: AS

SIGNIFICANCE LEVEL $\alpha = .0500$

ANALYSIS OF VARIANCE				ESTIMATION OF VARIANCE COMPONENTS			TEST OF HYPOTHESIS			
SOURCE	BETWEEN	WITHIN		MEAN	LEVEL	UNIT	VARIANCE	F-RATIO	EST. α	SIGNI-
			SUM OF	SQUARE		SIZE	COMPONENT			FICANCE
			SQUARES							
				DEGREES OF						
				FREEDOM						
1	0		1591.85876	28	56.85210	1	29	4.22958	1.93286	.043 *
2	1		823.57818	28	29.41351	2	57	5.09598	1.23560	.298 NS
3	2		595.12501	25	23.80500	3	82	11.12460	2.80138	.004 *
4	3		246.43001	29	8.49759	4	111	7.45197	12.02760	.006 *
5	4		3.51500	5	.70300	5	116	.70300		
TOTAL			3260.50695	115				28.60513		

VARIANCE COMPONENTS AS PERCENTAGES OF THE TOTAL VARIANCE

14.79 * 17.81 NS 38.89 * 26.05 * 2.46

COEFFICIENTS OF VARIANCE COMPONENTS IN ESTIMATES OF THE MEAN SQUARE

1.0000 1.0719 1.6397 2.5534 3.9975
 1.0000 1.0202 1.2119 1.4964
 1.0000 1.0400 1.3800
 1.0000 1.0460
 1.0000

* 25 x 25 km
 NS 5 x 5 km
 * 1 x 1 km
 * 200 x 200 m
 ANALITKA

2.6. Analiza variance

- Podatke za celoten niz spremenljivk običajno podamo tabelarično:

ANALIZA VARIANCE - geokemija tal Slovenije

n=123		odstotki variance					
Element		Varianca	25km	5km	1km	200m	lab
Al	N	3.54	0	37*	18*	0	45
Ca	L	0.329	10*	40	41***	8***	1
Fe	N	1.89	0	36*	22**	0	42
K	L	0.034	27**	7	31***	0	35
Mg	L	0.083	0	70**	18**	4	8
Na	L	0.049	17*	12	45***	0	26
P	L	0.049	13*	5	64***	5	13
Tl	N	0.02	17*	17	52***	7	7
Ag	L	0.096	12	0	19	3	66
As	L	0.046	6	3	53**	7	31
Ba	L	0.043	44***	0	38***	0	18
Be	N	0.746	0	21	27	44*	8
Bi	L	0.037	0	12	0	0	88
Cd	L	0.161	28**	0	36*	26*	10
Co	L	0.039	18*	0	41*	24	17
Cr	N	1610.4	2	58**	22***	0	18
Cu	L	0.069	2	52*	25**	0	21
La	L	0.033	15*	36	26**	11	12
Mn	L	0.083	46***	1	31**	10	12
Mo	L	0.072	40***	11	8	41	0
Nb	L	0.100	4	0	43*	32	21
Ni	L	0.066	35**	18	31***	5	11
Pb	L	0.048	31**	0	30*	36***	3
Sb	N	5.56	16*	0	0	83***	1
Sc	L	0.027	0	45*	29**	0	26
Sn	L	0.069	22*	20	0	30	28
Sr	L	0.039	16*	0	60***	7	17
Th	L	0.034	5	31p	2	14	48
V	L	0.038	0	49*	31***	0	21
W	L	0.249	30**	5	39**	24***	2
Y	L	0.067	21*	8	53***	12*	6
Zn	N	1033.9	20*	2	39*	14	25
Zr	L	0.084	40***	0	50***	0	10