

Biomeetria praks 7

Illustreeritud (mittetäielik) tööjuhend

Eeltöö

1. Avage *MS Excel*'is oma kursuse ankeedivastuseid sisaldav andmestik,
2. lisage uus tööleht, nimetage see ümber leheküljeks 'Praks7' ja
3. kopeerige kogu 'Andmed'-lehel paiknev andmetabel lehekülje 'Praks7' ülemisse vasakusse nurka.

Ülesanne.

Kas sugu ja auto omamine on seotud, st kas autot omavate tudengite arv on mees- ja naistudengite hulgas erinev?

1. Konstrueerige kahemõõtmeline sagedustabel tunnustele 'SUGU' ja 'AUTO'.
2. Leidke sellesse tabelisse lisaks absoluutsetele sagedustele ka kahed suhtelised sagedused (nii veeru- kui ka reaprotsendid).
3. Kommenteerige tabelit (sõnastades laused vähemalt kahe arvutatud protsendi kohta)!
4. Testimaks tuvastatud seose/erinevuse statistilist olulisust, pange kirja vastav hüpoteeside paar.
5. Tehke uus 2-mõõtmeline sagedustabel tunnustele 'SUGU' ja 'AUTO', mis seekord sisaldab üksnes absoluutseid sagedusi;
6. arvutage viimase alusel nullhüpoteesile vastavad (seost ei ole) sagedused ja
7. viige läbi hii-ruut-test (leidke p -väärtus).
8. Sõnastage lõppjärelendus (kas seos on statistiliselt oluline, mille alusel te seda otsustasite ja milles see seos ikkagi seisneb?).

Tööjuhend

1. Konstrueerige kahemõõtmeline sagedustabel tunnustele 'SUGU' ja 'AUTO'.

- Paigutage kursor andmetabeli suvalisse lahtrisse (leheküljel 'Praks 7').
- *Insert*-sakk → *PivotTable*

Loodav tabel paigutage samale 'Praks7'-lehele:

Create PivotTable

Choose the data that you want to analyze

Select a table or range
Table/Range: Praks7!\$A\$1:\$S\$58

Use an external data source
Choose Connection...
Connection name:

Choose where you want the PivotTable report to be placed

New Worksheet
 Existing Worksheet
Location: Praks7!\$V\$1

OK Cancel

PivotTable Field List

Choose fields to add to report:

- SUGU
- PIKKUS
- MASS
- PEA_P
- JALANR
- ODE_VEND
- MAT_HINNE
- HOMMIK
- PUDER
- LEMMIK
- HAIGE
- SPORT
- SUUSK
- AUTO

Drag fields between areas below:

Report Filter: [Empty]
Column Labels: [Empty]
Row Labels: [Empty]
Values: [Empty]

Drag fields between areas below:

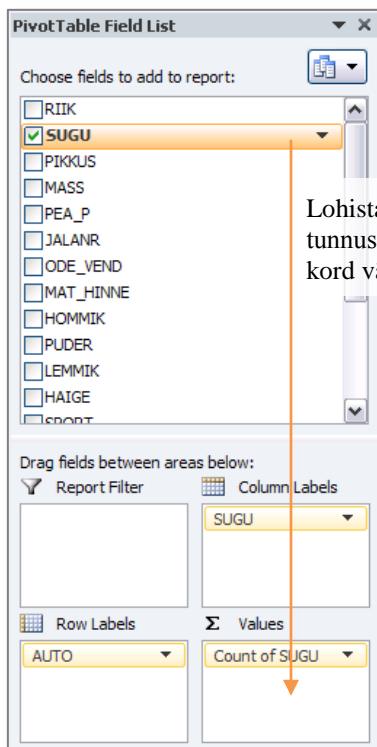
Report Filter: [Empty]
Column Labels: SUGU
Row Labels: AUTO
Values: Count of SUGU

Näiteks. Samas võib lasta Excelil kokku lugeda ka mõne teise **kõigil** tudengitel registreeritud tunnuse väärtused.

Tulemus:

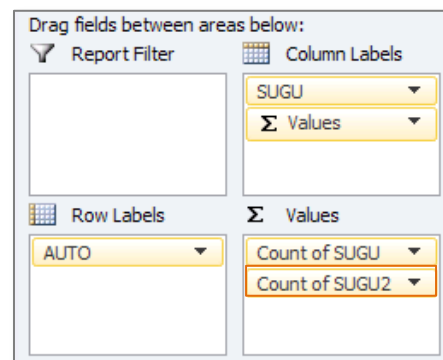
Count of SUGU	Column Labels		
Row Labels	N	M	Grand Total
ei	34	2	36
jah	13	8	21
Grand Total	47	10	57

- Leidke tabelisse lisaks absoluutsetele sagedustele ka kahed suhtelised sagedused (nii rea- kui ka veeruprotsendid).



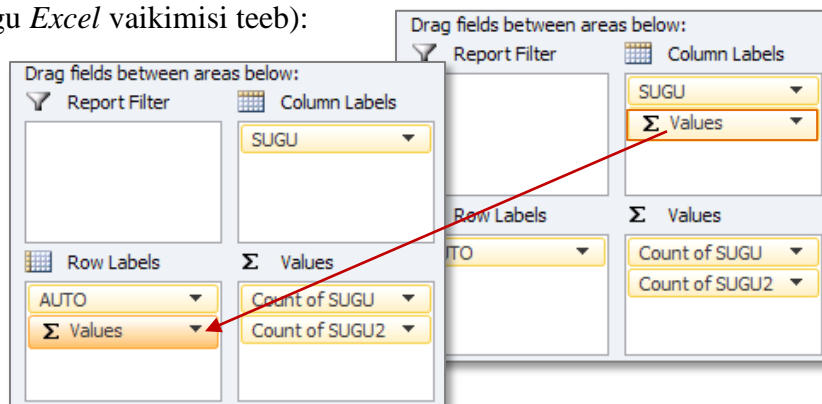
Lohistage tunnus 'SUGU' ka teine kord väärtuste lahtrisse.

Tulemus:



	N		M		Total Count of SUGU	Total Count of SUGU2
Row Labels	Count of SUGU	Count of SUGU2	Count of SUGU	Count of SUGU2		
ei	34	34	2	2	36	36
jah	13	13	8	8	21	21
Grand Total	47	47	10	10	57	57

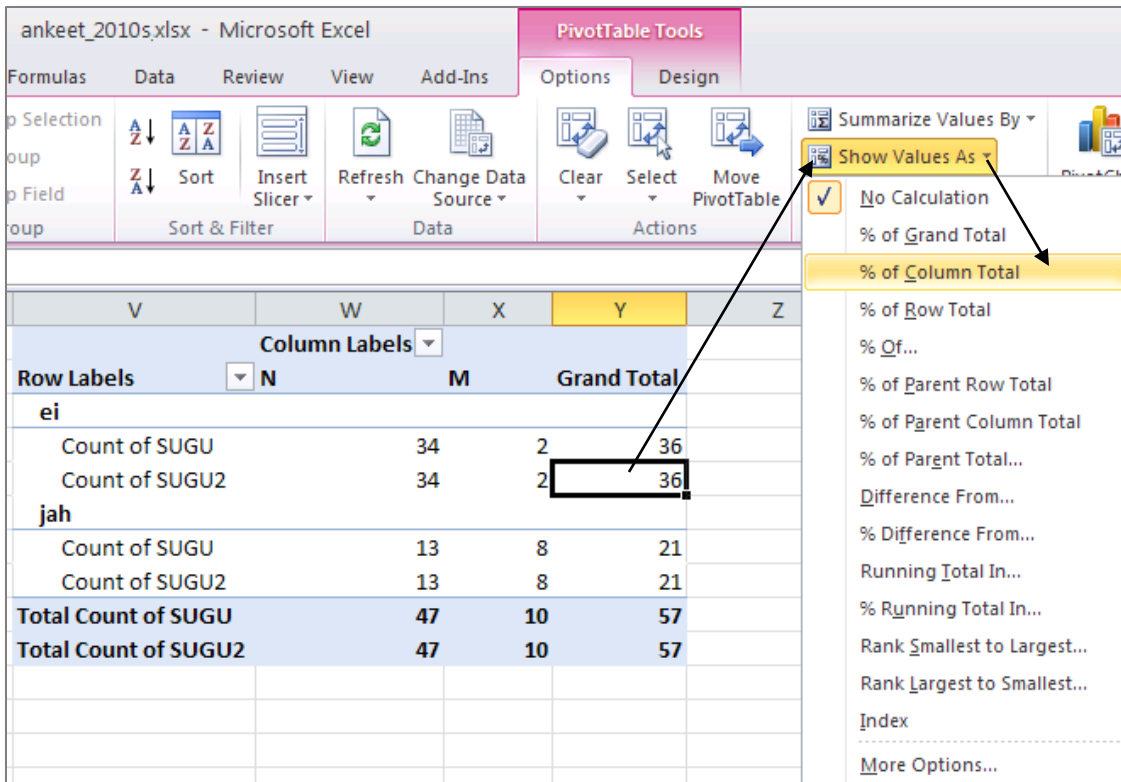
- Tabeli selgema esituse huvides jagage tabel arvutatavate väärtuste järgi ridadeks, mitte veergudeks (nagu Excel vaikimisi teeb):



Tulemus:

	Column Labels		
Row Labels	N	M	Grand Total
ei			
Count of SUGU	34	2	36
Count of SUGU2	34	2	36
jah			
Count of SUGU	13	8	21
Count of SUGU2	13	8	21
Total Count of SUGU	47	10	57
Total Count of SUGU2	47	10	57

- Esitamaks *Exceli* poolt uuesti kokku loetud tudengite arvu veeruprotsendina:



Tulemus:

Row Labels	N	M	Grand Total
ei			
Count of SUGU		34	2
Count of SUGU2		34	2
jah			
Count of SUGU		13	8
Count of SUGU2		13	8
Total Count of SUGU		47	10
Total Count of SUGU2		47	10

- Analoogselt arvutage ka reaprotsendid ... ().

(alustage sarnaselt eelmise lehekülje alguses kirjeldatule)

Row Labels	N	M	Grand Total
ei			
Count of SUGU		34	2
Count of SUGU2	72,34%	20,00%	63,16%
Count of SUGU3	94,44%	5,56%	100,00%
jah			
Count of SUGU		13	8
Count of SUGU2	27,66%	80,00%	36,84%
Count of SUGU3	61,90%	38,10%	100,00%
Total Count of SUGU		47	10
Total Count of SUGU2	100,00%	100,00%	100,00%
Total Count of SUGU3	82,46%	17,54%	100,00%

- Kommenteerige tabelit**
(sõnastades laused vähemalt ühe rea- ja ühe veeruprotsendi kohta)!

4. Seose statistilise olulisuse testimiseks **pange** esmalt **kirja hüpoteeside paar** (et oleks ikka selge, mida testite).

5. Tehke uus 2-mõõtmeline sagedustabel tunnustele 'SUGU' ja 'AUTO', mis seekord sisaldab üksnes absoluutseid sagedusi.

Column Labels			
Row Labels	N	M	Grand Total
ei			
Count of SU	34	2	36
Count of SU	72,34%	20,00%	63,16%
Count of SU	94,44%	5,56%	100,00%
jah			
Count of SU	13	8	21
Count of SU	27,66%	80,00%	36,84%
Count of SU	61,90%	38,10%	100,00%
Total Count of S	47	10	57
Total Count of S	100,00%	100,00%	100,00%
Total Count of S	82,46%	17,54%	100,00%
Kommentaariid.			
H ₀ : Sugu ja auto omamine ei ole seotud.			
H ₁ : Sugu ja auto omamine on seotud.			
Count of SUGU Column Labels			
Row Labels	N	M	Grand Total
ei	34	2	36
jah	13	8	21
Grand Total	47	10	57

6. Tehke uus, viimati loodud tabeliga analoogne sagedustabel, kuhu arvutage nullhüpoteesile vastavad sagedused.

Kuidas? Vt järgnevaid alapunkte!

- Tehke koopia *Pivot Table*'i abil konstrueeritud tabeli struktuurist ja väärtustest:

Count of SUGU Column Labels				
Row Labels	N	M	Grand Total	
ei	Copy	34	2	36
jah		13	8	21
Grand Total		47	10	57

Count of SUGU Column Labels				
Row Labels	N	M	Grand Total	
ei	Paste Values	34	2	36
jah		13	8	21
Grand Total		47	10	57

- seejärel kustutage ära kopeeritud tabeli sisu

Count of SUGU	Column Labels		
Row Labels	N	M	Grand Total
ei	34	2	36
jah	13	8	21
Grand Total	47	10	57

- ning arvutage sinna asemele sõltumatuse juuhule (nullhüpoteesile) vastavad sagedused (valemist $n_{ij} = n_i \times n_j / n$).

	V	W	X	Y
29	Count of SUGU	Column Labels		
30	Row Labels	N	M	Grand Total
31	ei	=Y31*W33/Y33		36 $n_{1.}$
32	jah			21 $n_{.2}$
33	Grand Total	47 $n_{.1}$	10	57 n

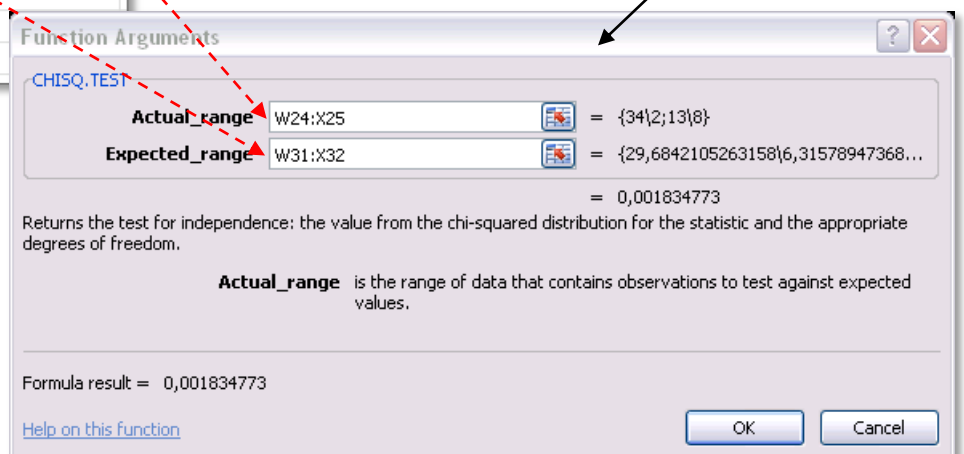
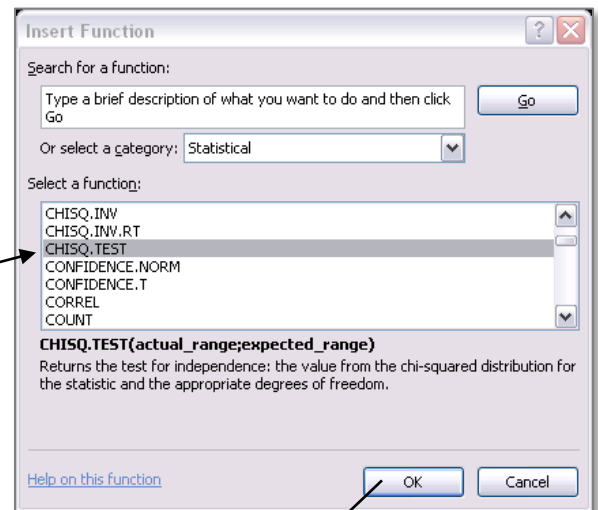
	V	W	X	Y
29	Count of SUGU	Column Labels		
30	Row Labels	N	M	Grand Total
31	ei	29,68421053	6,315789474	36 $n_{1.}$
32	jah	17,31578947	=Y32*X33/Y33	21 $n_{.2}$
33	Grand Total	47 $n_{.1}$	10	57 n

7. Viige läbi hii-ruut test (leidke p -väärtus) – funktsioon CHISQ.TEST.

Et oleks lihtsam aru saada, milliste arvude võrdlemisel χ^2 -test baseerub (ehk siis millised tabelite osad tuleb Excel'i funktsioonile CHISQ.TEST ette anda), võib vastavad lahtrid selguse mõttes näiteks ära värvida.

Järgnevalt, nagu funktsioonide puhul ikka, tuleb kursor panna lahtrisse, kuhu soovitakse tulemust saada (ja juurde võiks enne ka kirjutada, mida arvutama hakatakse).

	V	W	X	Y
21				
22	Count of SUGU	Column Lab		
23	Row Labels	N	M	Grand Total
24	ei	34	2	36
25	jah	13	8	21
26	Grand Total	47	10	57
27				
28				
29	Count of SUGU	Column Labels		
30	Row Labels	N	M	Grand Total
31	ei	29,68421053	6,315789474	36
32	jah	17,31578947	3,684210526	21
33	Grand Total	47	10	57
34				
35				
36	Hii-ruut-test			



NB! Vanemais Exceli versioonides testib sama hüpoteesi funktsioon CHITEST.

Tulemus:

H ₀ : Sugu ja auto omamine ei ole seotud.			
H ₁ : Sugu ja auto omamine on seotud.			
Count of SUGU Column Lab			
Row Labels	N	M	Grand Total
ei	34	2	36
jah	13	8	21
Grand Total	47	10	57
Count of SUGU Column Labels			
Row Labels	N	M	Grand Total
ei	29,68421053	6,315789474	36
jah	17,31578947	3,684210526	21
Grand Total	47	10	57
Hii-ruut-test	0,001834773		

8. Sõnastage lõppjärelus

(Kas seos on statistiliselt oluline? **Miks te nii otsustasite?** Milles see seos ikka seisneb?).