

TÖLVUTÍMAR Í AÐFERÐAFRÆÐI II

6. Verkefni í SPSS: Klasasrit og krosstöflur

Í þessu verkefni verður líkt og í fimmta verkefni unnið með niðurstöður úr TIMSS (*Third International Mathematics and Science Study*). Þetta er alþjóðleg rannsókn á námsárangri í stærðfræði og náttúrufræði sem var framkvæmd árið 1995. Við munum vinna með upplýsingar um kynferði 9 ára barna og hversu miklum tíma þau eyða í tölvuleiki á venjulegum skóladegi. Í úrtakinu eru 109 nemendur.

Skoðun gagna

Fyrst skoðum við helstu eiginleikum gagnanna, leitum að afbrigðum í gögnunum, skilgreinum breytur og ákvörðum brottfallsgildi.

1. Sæktu gagnaskrána Timss3 sem er á K-drifi, í möppunni K:/gba/spssdata.

Gagnaskráin inniheldur þrjár breytur: `itsex` gefur upplýsingar um kynferði, `aimatscr` sýnir frammistöðu á stærðfræðiprófi og `asbgday2` gefur fjölda klukkustunda sem spilað er í tölvuleikum á venjulegum skóladegi. Í þessu verkefni munum við skoða tengsl kynferðis og fjölda klukkustunda sem varið er í tölvuleiki.

2. Skýrðu `itsex` sem Kynferði og gildin sem 1 = stúlkur og 2 = piltar í valmynd eða með skipuninni (litlir eða stórir stafir skipta ekki máli).

Skipanir

Hægt er að fá skipanir fyrir aðgerðir í SPSS með því að smella á *paste* hnappinn í valmynd. Samsvarandi skipun birtist þá neðst í skipanaglugganum.

Ef þú manst ekki hvernig þetta er gert, skaltu fletta því upp í fyrri verkefnum. Í Verkefni 1 var sýnt hvernig þetta er gert með valmynd; í Verkefni 3 gerðirðu þetta með skipunum.

3. Skýrðu `asbgday2` sem Klukkustundir í tölvuleikum og gildin sem 1 = ekkert, 2 = < 1 klst, 3 = 1-2 klst, 4 = 3-4 klst og 5 = > 4 klst.

Þetta er einfaldast að gera með skipunum en þú mátt nota valmyndir ef þú vilt.

Að lokum hyggjum við að brottfallsgildum. Þar sem um er að ræða tvær flokkabreytur fáum við hvað skýrasta mynd með því að skoða tíðni (*frequencies*) mæligilda.

4. Skoðaðu tíðnitölur fyrir `itsex` og `asbgday2`.

Skilgreindu óvenjuleg gildi sem brottfall. Einu brottfallsgildin ættu að vera 8 og 9.

Klasasúlurit

Við skulum gera ráð fyrir að okkur hafi tekist að skilgreina breytur og rekið augun í öll afbrigðileg mæligildi. Næsta skref er að skoða gögnin myndrænt.

Okkur langar til að vita hvort það sé kynjamunur í ástundun tölvuleikja. Í gagnasafninu er tölvuleikjanotkun flokkuð í fimm mismunandi flokka eftir því hversu löngum tíma er varið í þá á venjulegum skóladegi. Einföld leið til að meta kynjamun er að teikna súlurit (*bar graph*) þar sem tölvuleikjanotkun er sýnd sérstaklega fyrir annars vegar stúlkur og hins vegar pilta.

Ef dreifing ástundunar er ólík eftir kynjum ber það vott um kynjamun í ástundun.

1. Veldu **Graphs / Bar** og **Clustered** í fyrstu valmynd og smelltu á **Define**. Í næstu valmynd skalt þú færa **Kynferði** í reitinn við **Category Axis** og **Klukkustundir** í tölvuleikjum við **Define Clusters by**. Að lokum skalt þú smella á **Options** og taka af hakið við **Display groups defined by missing values**.

Á myndinni sjást tveir klasar súlna. Svona mynd er því nefnd klasarit (*cluster bar*) eða klasað súlurit (*clustered bar graph*). Annar klasinn gefur upp skilyrta dreifingu á ástundun tölvuleikja fyrir stúlkur og hinn skilyrta dreifingu fyrir piltar. Best fer á því að frumbreytan (kynferði í okkar tilviki) skilgreini klasanna, þ.e. einn klasi sé fyrir hvert gildi frumbreytunnar.

Myndin ber með sér að skilyrta dreifingin sé mjög mismunandi eftir kynjum. Langalgengast er að stúlkur sé ekki í tölvuleik á venjulegum skóladegi og mjög fáar stúlkur verja meira en klukkustund í tölvuleiki. Á hinn bóginn eru það tiltölulega fáir piltar sem stunda ekki tölvuleiki og algengast að þeir séu allt að tvær klukkustundir í tölvuleikjum á venjulegum skóladegi.

Í sjálfu sér eru niðurstöður myndarinnar eins skýrar og niðurstöður geta orðið. Níu ára stúlkur eru greinilega lítið í tölvuleikjum meðan piltar á þessum aldri stunda þessa iðju af nokkru kappi. En þar sem við höfum ekki áhuga á þessum 109 börnum sem úrtakið nær til heldur öllum íslenskum níu ára börnum, þurfum við að prófa niðurstöðuna tölfræðilega.

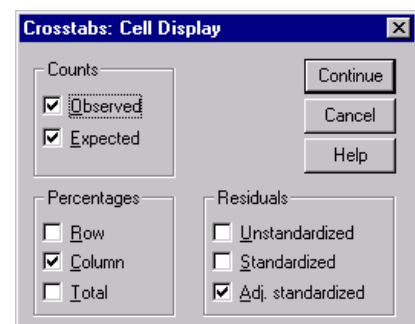
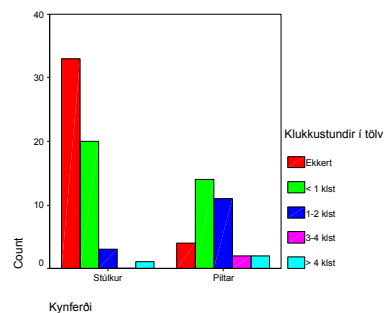
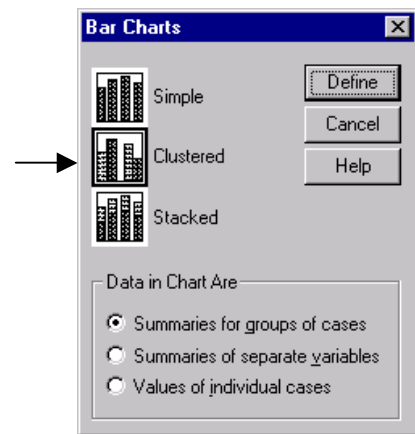
Kíkvaðrat (χ^2)

Við viljum vita hvort munur sé á skilyrtum dreifingum eftir kyni í þýði. Myndin byggðist á aðeins 109 einstaklingum. Þótt skilyrtu dreifingarnar séu eins í þýði er hugsanlegt að fá ólíkar dreifingar í úrtaki dregnu úr slíku þýði. Við viljum því prófa tölfræðilega hvort dreifingarnar tvær séu eins eða ólíkar í þýði allra íslenskra 9 ára barna.

Við höfum tvær flokkabreytur og getum því myndað krosstöflu. Við getum síðan prófað hvort skilyrtu dreifingarnar séu eins í þýði með kíkvaðrat tölfræðiprófinu. Núlltilgátan er að skilyrtu dreifingarnar séu báðar eins. Tölfræðiforritið reiknar væntitíðni (*expected frequency*) en það er sá fjöldi sem búast má við í hverjum og einum reit töflunnar ef skilyrtu dreifingarnar eru eins í þýði. Allur mismunur á fjölda (rauntíðninni) og væntitíðni ber vott um frávik frá núlltilgátunni.

Tvö skilyrði eru fyrir því að nota kíkvaðratpróf í krosstöflu. Fyrri skilyrðið er að ekkert væntigildi í krosstöflunni sé minna en 1,0. Síðara skilyrðið er að hlutfall reita í töflunni með væntigildi undir 5,0 verður að vera innan við 20%.

1. Veldu **Analyse / Descriptive Statistics / Crosstabs**. Settu **Kynferði** í reitinn undir **Columns** og **Klukkustundir** í tölvuleikjum undir **Rows**. Smelltu á **Statistics** og merktu við **Chi-square**. Smelltu að lokum á **Cells** og merktu við **Expected** undir **Counts**, **Column** undir **Percentages** og **Adj. standardized** undir **Residuals**.



Krosstaflan mun nú sýna væntitiðni (*expected counts*), hlutfall drengja og stúlkna í hverjum flokki (*column percentages*) og leiðrétta leif (*adjusted standardized residuals*).

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Klukkustundir í tölvuleikjum á venjulegum skóladegi * Kynferði	90	86,5%	14	13,5%	104	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	26,153 ^a	4	,000
Likelihood Ratio	28,503	4	,000
Linear-by-Linear Association	21,426	1	,000
N of Valid Cases	90		

a. 4 cells (40,0%) have expected count less than 5. The minimum expected count is ,73.

Klukkustundir í tölvuleikjum á venjulegum skóladegi * Kynferði Crosstabulation

	Ekkert	Count	Kynferði		Total
			Stúlkur	Piltar	
Klukkustundir í tölvuleikjum á venjulegum skóladegi		Count	33	4	37
		Expected Count	23,4	13,6	37,0
		% within Kynferði	57,9%	12,1%	41,1%
		Adjusted Residual	4,3	-4,3	
< 1 klist		Count	20	14	34
		Expected Count	21,5	12,5	34,0
		% within Kynferði	35,1%	42,4%	37,8%
		Adjusted Residual	-,7	,7	
1-2 klist		Count	3	11	14
		Expected Count	8,9	5,1	14,0
		% within Kynferði	5,3%	33,3%	15,6%
		Adjusted Residual	-3,5	3,5	
3-4 klist		Count	0	2	2
		Expected Count	1,3	,7	2,0
		% within Kynferði	,0%	6,1%	2,2%
		Adjusted Residual	-1,9	1,9	
> 4 klist		Count	1	2	3
		Expected Count	1,9	1,1	3,0
		% within Kynferði	1,8%	6,1%	3,3%
		Adjusted Residual	-1,1	1,1	
Total		Count	57	33	90
		Expected Count	57,0	33,0	90,0
		% within Kynferði	100,0%	100,0%	100,0%

Uppfyllir prófið forsendur kíkvaðrats? Rökstyddu svarið og berðu undir kennara.

Ef prófið uppfyllir ekki forsendur kíkvaðrats er hægt að fella brott eða leggja saman flokka. Þó ber að hafa í huga að ávallt tapast einhverjar upplýsingar við það að leggja saman flokka þar sem flokkunin verður grófarei.

- Legðu saman minnstu flokka Klukkustunda í tölvuleikjum þangað til forsendur kíkvaðrats eru uppfylltar. Þetta gerir þú með því að endurkóða breytuna. Í valmyndum má fara leiðina Transform / Recode / Into Same Variables eða með skipunum.

Ertu búin að gleyma hvernig breytur eru kóðaðar? Skoðu verkefni 3 þar sem þessi skipun er útskýrð. Mundu líka að endurskýra gildi breytunnar!

- Endurtaktu kíkvaðratrófið eftir að þú hefur endurkóðað breytuna. Er forsendur prófsins uppfylltar núna?
- Eftir að þú hefur fengið gilda niðurstöðu kíkvaðrats skaltu birta niðurstöðuna myndrænt eins og í upphafi.

Kóðun sömu eða nýrrar breytu

Þú getur ýmist endurkóðað breytuna sem þú ert að vinna með eða búið til nýja breytu. Það er fljótlegt og þægilegt að nota sömu breytu en þá þarftu að gæta þess að vista ekki breytingarnar. Ef þú býrð til nýja breytu geturðu kóðað hana eins og þú vilt, vistað skrána og samt haft aðgang að bæði upprunalegu breytni og nýju endurkóðuðu breytni.

Túlkaðu niðurstöðurnar: _____

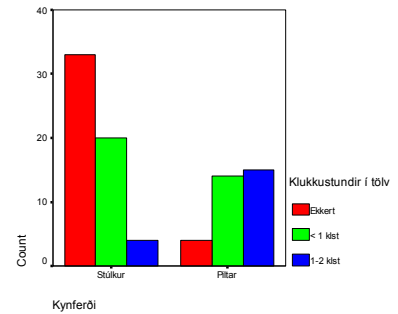
Úr SPSS í Microsoft Word

Að lokum skulum við færa klasasúluritið sem þú gerðir síðast yfir í Word ritvinnsluforritið.

- Smelltu á myndina í niðurstöðuglugganum með hægri músarhnappinum og veldu **Copy** eða veldu **Edit / Copy** á valröndinni. Opnaðu Word og veldu **Edit / Paste Special** á valröndinni. Veldu að taka myndina inn í Word sem **Picture (Enhanced Metafile)** og smelltu á **OK**.

Nú ætti myndin að birtast í Word. Hana má snyrta til á ýmsa vegu, t.d. stækka eða minnka, setja ramma í kringum hana, o.s.frv. Við skulum setja dálítinn texta með myndinni.

- Skrifaðu stutta málsgrein í Word þar sem þú lýsir niðurstöðunum efnislega, skýrir frá niðurstöðu kíkvaðratrófsins og túlkar.



Gilt kíkvaðratpróf

Svona ættu niðurstöðurnar að lita út eftir að þú hefur endurkóðað fylgibreytuna.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Klukkustundir í tölvuleikjum á venjulegum skóladegi * Kynferði	90	86,5%	14	13,5%	104	100,0%

Klukkustundir í tölvuleikjum á venjulegum skóladegi * Kynferði Crosstabulation

			Kynferði		Total
			Stúlkur	Piltar	
Klukkustundir í tölvuleikjum á venjulegum skóladegi	Ekkert	Count	33	4	37
		Expected Count	23,4	13,6	37,0
		% within Kynferði	57,9%	12,1%	41,1%
		Adjusted Residual	4,3	-4,3	
	< 1 klst	Count	20	14	34
		Expected Count	21,5	12,5	34,0
		% within Kynferði	35,1%	42,4%	37,8%
		Adjusted Residual	-,7	,7	
	1-2 klst	Count	4	15	19
		Expected Count	12,0	7,0	19,0
		% within Kynferði	7,0%	45,5%	21,1%
		Adjusted Residual	-4,3	4,3	
Total	Count	57	33	90	
	Expected Count	57,0	33,0	90,0	
	% within Kynferði	100,0%	100,0%	100,0%	

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	25,576 ^a	2	,000
Likelihood Ratio	27,314	2	,000
Linear-by-Linear Association	25,173	1	,000
N of Valid Cases	90		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 6,97.