

Fylgnistuðullinn fí; \emptyset

Fyrirlestur í Aðferðafræði II

© 1998, 2001, 2003 Guðmundur B. Arnkelsson

All rights reserved. Copying or distribution prohibited without explicit permission. Students in Methodology II at the University of Iceland may print a copy for their own private use.

Þekking á heimilisfangi

Taflan sýnir tengsl kynferðis og þekkingar á heimilisfangi sínu

Þriðjungur pilta veit *ekki* hvar hann á heima samanborið við innan við tíundu hverja stúlku.

Tengslin sjást með því að bera saman skilyrtar dreifingar

Hjá drengjum skiptist þekking á heimilisfangi í 10:5; hjá stúlkunum er skiptingin 13:1.

Svör drengja dreifast því öðru vísi en svör stúlkna. Það má segja að meiri dreifing sé í svörum piltanna.

Tengslin kynferðis og þekkingar eru umtalsverð.

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	10	13	23
Þekkir ekki	5	1	6
Samtals	15	14	29

Uppsetning taflna

Láta frumbreytu skilgreina dálka og fylgibreytu skilgreina línur.

Rita nöfn breyta yfir viðkomandi dálkum.

Hafa gildi breyta í eðlilegri röð

- Raðbreytur hækki frá vinstri til hægri og neðan frá og upp.
- Nafnbreytur séu í eðlilegri röð

Gefa upp heildartölur

Styrkur tengsla í 2x2 töflum

Tengslin eru meiri eftir því sem munurinn á skilyrtu dreifingunum eykst.

Á efstu myndinni er dreifing þekkingar eins lík hjá drengjum og stúlkum og hægt er.

Hér geta skilyrtu dreifingarnar aldrei orðið eins og því verða alltaf einhver tengsl.

Á miðmyndinni eru dreifingarnar orðnar mun ólíkari og tengsl kyns og þekkingar orðin meiri.

Á neðstu myndinni eru skilyrtu dreifingarnar eins ólíkar og þær geta orðið; tengslin eru fullkomin, þ.e. enginn drengur og allar stúlkur þekkja heimilisfang sitt.

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	8	7	15
Þekkir ekki	7	7	14
Samtals	15	14	29

Mjög lítil tengsl

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	3	11	14
Þekkir ekki	12	3	15
Samtals	15	14	29

„Miðlungs“ tengsl

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	0	14	14
Þekkir ekki	15	0	15
Samtals	15	14	29

Fullkomin tengsl

Áhrif skálína á styrk tengsla

Skálínurnar í töflunni er ákveðinn mælikvarði á tengslin í töflunni

Eftir því sem skilyrtu dreifingarnar eru ólíkari eykst munur á skálínunum (ad) og (bc).

Munur skálína er því ákveðinn mælikvarði á hve mikil tengsl eru á milli breytanna tveggja.

Algengast er að nota mismuninn á margfeldunum; þá er jafnan $bc-ad$ mælikvarði á stærð tengslanna.

Ef marka má útreikningana hér til hliðar þá gefur þessi mælikvarði sannfærandi mat á tengslum kynferðis og þekkingar á heimilisfangi.

Fylgibreyta	Frumbreyta		Samtals
	0	1	
1	a	b	a + b
0	c	d	c + d
Samtals	a + c	b + d	N = a + b + c + d

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	8	7	15
Þekkir ekki	7	7	14
Samtals	15	14	29

$$bc - ad = 7 \cdot 7 - 8 \cdot 7 \\ = 49 - 56 = -7$$

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	3	11	14
Þekkir ekki	12	3	15
Samtals	15	14	29

$$bc - ad = 11 \cdot 12 - 3 \cdot 3 \\ = 132 - 9 = 123$$

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	0	14	14
Þekkir ekki	15	0	15
Samtals	15	14	29

$$bc - ad = 14 \cdot 15 - 0 \cdot 0 \\ = 210 - 0 = 210$$

Mælitalan fí

Mælikvarðinn $bc-ad$ er háður heildarfjöldanum í töflunni

Því fleiri einstaklingar sem eru í töflunni, verður mismunurinn $bc-ad$ stærrí jafnvel þótt tengslin séu nákvæmlega þau sömu.

Neðri taflan hefur tífoldan fjölda á við hina; það veldur því að $bc-ad$ verður hundraðfalt hærra í neðri töflunni þótt tengslin séu nákvæmlega þau sömu.

Mælitalan fí leysir þennan vanda.

Fí (ϕ) leiðréttir fyrir fjöldann í töflunni þannig að niðurstaðan verði á bilinu $-1,0$ til $+1,0$.

Í báðum töflunum er $\phi = 0,59$.

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	3	11	14
Þekkir ekki	12	3	15
Samtals	15	14	29

$$bc - ad = 11 \cdot 12 - 3 \cdot 3 \\ = 132 - 9 = 123$$

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	30	110	140
Þekkir ekki	120	30	150
Samtals	150	140	290

$$bc - ad = 110 \cdot 120 - 30 \cdot 30 \\ = 13.200 - 900 \\ = 12.300$$

Fylgibreyta	Frumbreyta		Samtals
	0	1	
1	a	b	a+b
0	c	d	c+d
Samtals	a+c	b+d	N=a+b+c+d

$$\phi = \frac{bc - ad}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

Útreikningur og túlkun á ϕ

Myndirnar sýna útreikning á ϕ fyrir töfluna á upphafstöflunni.

Skv. útreikningunum er $\phi = 0,32$.

Mælitöluna ϕ má túlka svipað og Pearson r .

ϕ í öðru veldi gefur skýrða dreifingu.

Í þessu tilfalli skýrir kynferði 10% ($0,32^2 = 0,10$) af dreifingu þekkingar.

Niðurstaðan $\phi = 0,32$ er ekki mjög upplýsandi; skv. henni skilar eins staðalfráviks breyting á kynferði sér sem 0,32 staðalfrávik í þekkingu.

Vandinn er að þekking getur aðeins breyst úr vanþekkingu í þekkingu en ekki um tiltekinn hluta staðalfráviks.

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	10	13	23
Þekkir ekki	5	1	6
Samtals	15	14	29

$$\phi = \frac{bc - ad}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} = \frac{13 \cdot 5 - 10 \cdot 1}{\sqrt{(23)(6)(15)(14)}} = \frac{65 - 10}{\sqrt{28980}} = \frac{55}{170,24} = 0,323 \approx 0,32$$

Fí fyrir nokkrar töflur

Töflurnar til hliðar eru þær sömu og í upphafi

Mismunur á margfeldi skálína gaf mælikvarða á tengsl sem reyndist erfitt að túlka.

Hér eru töflurnar birtar með \bar{f}_i mælitölunni á tengsl.

Taflan með minnstu mögulegu tengslum er með \bar{f}_i sem er nánast 0,0.

Taflan mið miðlungstengslum er með \bar{f}_i sem er 0,59. Taflan með mestu mögulegu tengslum fær \bar{f}_i sem er 1,0.

Tengslin eru því metin á kunnuglegum mælikvarða.

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	8	7	15
Þekkir ekki	7	7	14
Samtals	15	14	29

$$\phi = -0,03$$

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	3	11	14
Þekkir ekki	12	3	15
Samtals	15	14	29

$$\phi = 0,59$$

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	0	14	14
Þekkir ekki	15	0	15
Samtals	15	14	29

$$\phi = 1,00$$

Túlkun á stefnu sambandsins

Fí getur verið jákvætt eða neikvætt

Að jafnaði tákna neikvæð mælitala á tengsl að tengslin séu neikvæð.

Sama á við um fí ef eftirfarandi forsendur eru fyrir hendi:

- Gildum beggja breyta má raða.
- Gildum beggja breyta er raðað rétt í töfluna.

Töflurnar tvær til hliðar eru eins nema búið er að víxla kynjunum í annarri.

Kynferði er eigindleg breyta; því má raða gildum hennar í hvaða röð sem er.

Því hefur það enga merkingu í þessu tilviki hvort mælitalan er jákvæð eða neikvæð.

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	10	13	23
Þekkir ekki	5	1	6
Samtals	15	14	29

$$\varphi = 0,32$$

Heimilisfang	Kyn		Samtals
	Drengir	Stúlkur	
Þekkir	13	10	23
Þekkir ekki	1	5	6
Samtals	14	15	29

$$\varphi = -0,32$$

Tvíkostabreytur sem raðbreytur

Tvískiptar eigindlegar breytur má túlka sem raðbreytur

Eigindleg breyta gefur til kynna hvort e–ð sé til staðar eða ekki; meginleg breyta sýnir hve mikið er af e–u.

Tvískipt eigindleg breyta sýnir hvort e–ð sé eða sé ekki til staðar; það er *meira* af hafa e–ð heldur en að hafa það ekki. Það má því líta á þetta sem meginlegar upplýsingar.

- Þeir sem eiga tölvu hafa tölvuna fram yfir aðra
- Þeir sem hafa bílpróf hafa ökuréttindin fram yfir aðra
- Konur hafa kvenleikann fram yfir karla (og karlar karlleikann fram yfir konur)



0: Ekki kona

1: Karl



1: Kona

0: Ekki karl

Kynferði má kóða á ýmsa vegu; við getum litið svo á að þannig sé magnbundin tilvist ýmist karl- eða kvenleika.

∅ sem Pearson fylgni

Mælitalan fí er í rauninni Pearson r fyrir tvískiptar breytur

Í gagnafylkinu eru drengir kóðaðir sem „1“ og stúlkur sem „2;“ að þekkja heimilisfang sitt er kóðað „1“ og að þekkja það ekki er kóðað „0.“

Með því móti eru fengnar tvær meginlegar breytur og hægt að reikna Pearson fylgni á hefðbundin hátt.

Í ljós kemur að $r(27)= 0,32$ eða sama niðurstaða og fæst með fí.

Þessi samsvörun er alltaf fyrir hendi— fí samsvarar því að reikna Pearson r fyrir kóðaðar tvískiptar breytur.

Fí er því í raun reikniformúla fyrir r notuð fyrir tvær tvískiptar breytur.

Barn	Kyn	Heimilisfang
1	1	1
2	1	1
3	1	1
...
10	1	1
11	1	0
12	1	0
...
15	1	0
16	1	0
17	2	1
18	2	1
...
27	2	1
28	2	1
29	2	0