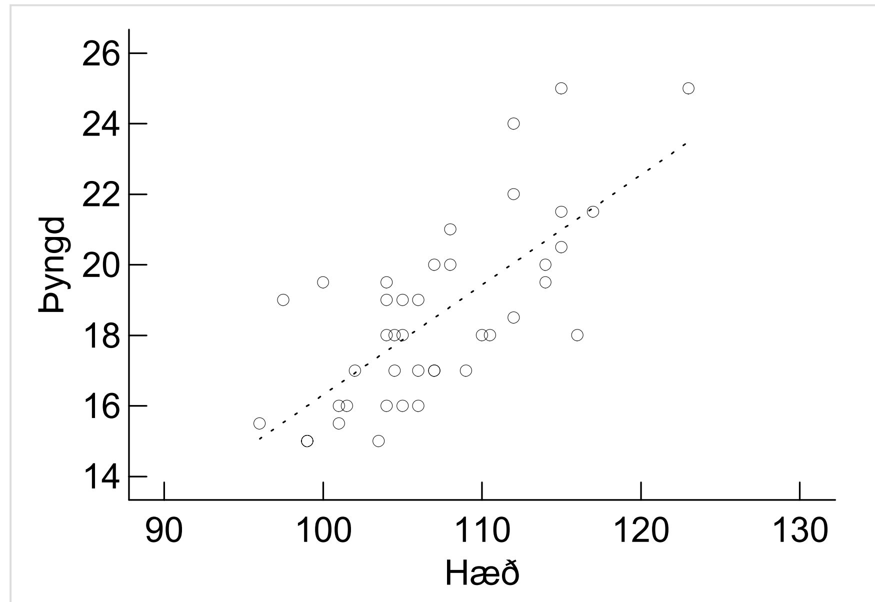
Sveigfylgni

Fyrirlestur í Aðferðafræði II

Beinlínutengsl

Beinlínutengsl eru algengust í félagsvísindum

- Sumar breytur hafa ekki beinlínutengsl
- Beinlína gefur oft fullnægjandi nálgun að formum tengslanna
- Áhrif frumbreytu nokkurn veginn þau sömu óháð því hvaða gildi hún tekur

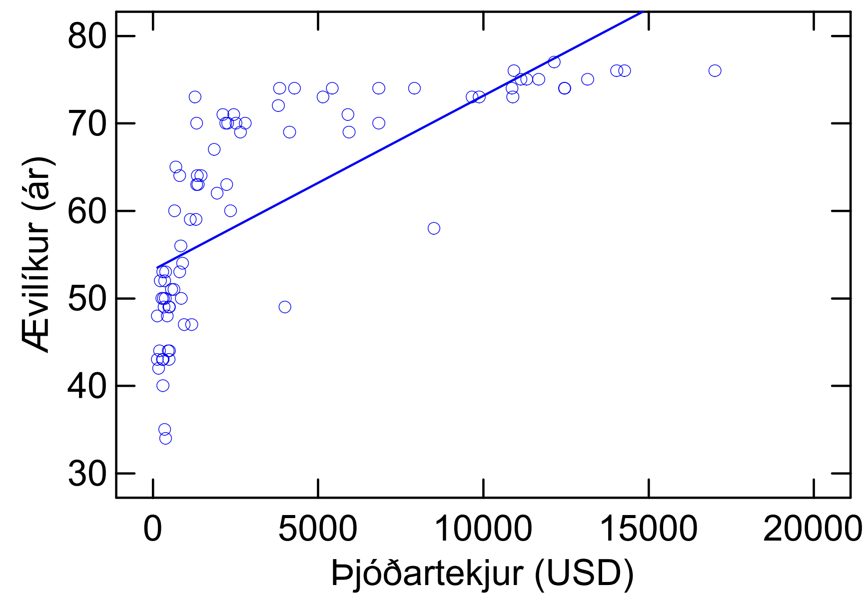


Dæmi um sveigbogatengsl

Aukning ævilíkinda er hröð við lágar en hæg við háar þjóðartekjur. Áhrif þjóðartekna á ævilíkur eru því meiri hjá fátækum en ríkum þjóðum.

Bein lína gerir ráð fyrir sömu áhrifum óháð þjóðartekjum og gróflega rangmetur því áhrif þeirra á ævilíkur.

Bein lína vanmetur einnig styrk tengslanna, þ.e. skýrða dreifingu, þar sem dreifingin er meiri í kringum beinu línuna en í kringum samsvarandi sveigbogalínu.



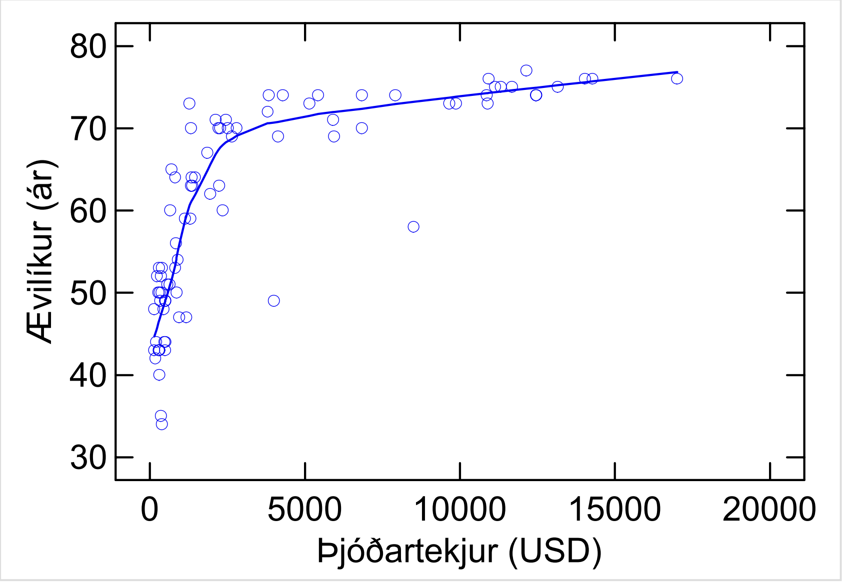
Ævilíkur = 53,26 + 0,002 · Þjóðartekjur
$r(80) = 0{,}72$; $r^2 = 0{,}52$

Sveigbogalína

Sveigbogalína gefur betri mynd af áhrifum og styrk tengsla; áhrif eru breytileg og minni dreifing í kringum línuna.

Eta (η) er mælitala sem metur styrk tengsla

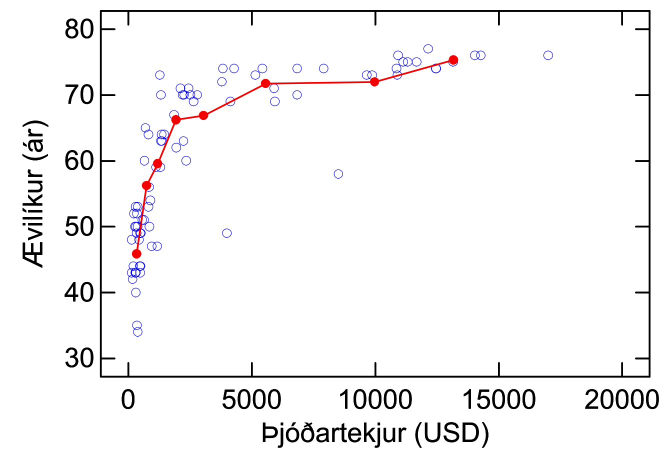
- Er túlkuð svipað og r, þ.e. η^2 gefur til kynna skýrða dreifingu
- Miðast við rofna frumbreytu og samfellda fylgibreytu
- Því þarf að skipta samfelldri frumbreytu í flokka áður en η er reiknað



Tengsl sem röð meðaltala

Línan sem felld er að gögnunum gefur upp sennilegustu spágildin fyrir samsvarandi gildi frumbreytu

- Við getum því litið á línuna sem felld er að gögnunum sem röð eins konar meðaltala þar sem meðaltölin gefa sennilegustu gildi fylgibreytunnar.
- Röng lína, t.d. bein lína þar sem sveigbogi er réttari, leiðir því til rangra spágilda og því rangrar spár.
- Við vitum sjaldnast hvaða lína er rétt en getum ráðið í sennilegasta form línunnar með því að skoða hvernig ólíkar línur falla að gögnunum. Bein lína er t.d. ósennileg fyrir tengsl ævilíkinda og þjóðartekna.



Á myndinni hafa þjóðartekjur verið flokkaðar niður og reiknað meðaltal hvers flokks. Meðaltölin ná að sýna tengsl breytanna nokkuð vel og mun betur en t.d. bein lína.

Eta fyrir rofna frumbreytu

Myndbandaleiga með þrjár tegundir fríðinda

- Engin fríðindi
- Sjöunda hver spóla ókeypis
- Að *meðaltali* 7. hver ókeypis

Engin fríðindi	Sjöunda hver ókeypis	Að meðaltali 7. hver ókeypis
2	7	10
4	6	7
3	6	9
0	9	8
1	7	12
2	7	8

Upplýsingar um fríðindi minnka óvissu í spá

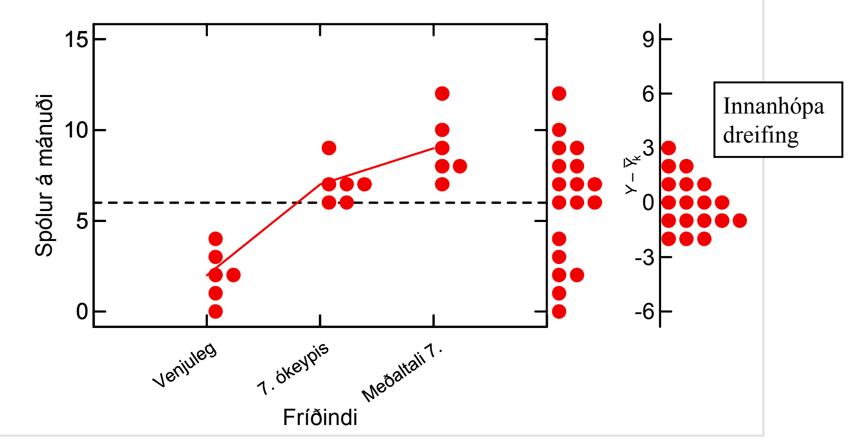
- Án upplýsinga er óvissan mikil, sbr. dreifitölu heildar
- Innan hvers hóps er dreifingin minni: Óvissan minnkar ef fríðindin eru þekkt.

Mælitala	Engin fríðindi	Sjöunda hver ókeypis	Að meðaltali 7. hver ókeypis	Í heild
Meðaltal	2,0	7,0	9,0	6,0
Dreifitala	2,0	1,2	3,2	11,1
Fjöldi	6	6	6	18

Hlutfallsleg skýring frumbreytu

Við höfum þrjár spurningar

- marktekt; lærum það síðar
- áhrif; munur meðaltala
- skýrða dreifingu (styrk tengsla); fer eftir því hvað dreifing kringum meðaltöl hópanna er hlutfallslega minni en dreifingin kringum heildarmeðaltalið



Áhrif felast í hér í mismuni meðaltala. Skýrða dreifingu má meta með η^2; hversu mikið dreifing (þ.e. óvissa) minnkar við að þekkja frumbreytuna (þ.e. hópmeðaltölin).

Heildardreifing er mun meiri en dreifing innan hópanna þriggja. Það samsvarar því að spáóvissa minnki við það að vita hvaða fríðindi viðskiptavinur hefur.

K: VI.1, bls. 197–199

Eta sem skýrð dreifing

Við getum metið skýrða dreifingu á tvo vegu:

- Athuga hve mikið hlutfall af breytileikanum má skýra með meðaltölum hópanna (millihópadreifingunni).
- Athuga hve mikið hlutfall af breytileikanum er í kringum meðaltöl hópanna.

Eta í öðru veldi (η^2) gefur til kynna skýrða dreifingu

- Minnkuð óvissa við að þekkja hópmeðaltölin

$$\eta^2 = \frac{\sum^{k} n_k \left(\bar{Y}_k - \bar{Y}_{Heild}\right)^2}{\sum^{n} \left(Y - \bar{Y}_{Heild}\right)^2}$$

Dreifing meðaltala

Heildardreifing

$$\eta^2 = 1 - \frac{\text{Dreifing innan hópa}}{\text{Heildardreifing}} = \frac{\text{Dreifing milli hópa}}{\text{Heildardreifing}} = \frac{\text{Summa kvaðrata milli hópa}}{\text{Heildarsumma kvaðrata}}$$

K: VI.2, bls. 201–202

Útreikningur á η

$$\eta^2 = \frac{\sum^{k} n_k \left(\overline{Y}_k - \overline{Y}_{Heild}\right)^2}{\sum^{n} \left(Y - \overline{Y}_{Heild}\right)^2}$$

$$= \frac{156}{188} = 0{,}829 \approx 0{,}83$$

Hópur	Spólur	$Y - \overline{Y}_{Heild}$	$\left(Y - \overline{Y}_{Heild}\right)^2$	$\overline{Y}_k$	$\overline{Y}_k - \overline{Y}_{Heild}$	$\left(\overline{Y}_k - \overline{Y}_{Heild}\right)^2$
1	2	–4	16	2	–4	16
1	4	–2	4	2	–4	16
1	3	–3	9	2	–4	16
1	0	–6	36	2	–4	16
1	1	–5	25	2	–4	16
1	2	-4	16	2	-4	16
2	7	1	1	7	1	1
2	6	0	0	7	1	1
2	6	0	0	7	1	1
2	9	3	9	7	1	1
2	7	1	1	7	1	1
2	7	1	1	7	1	1
3	10	4	16	9	3	9
3	7	1	1	9	3	9
3	9	3	9	9	3	9
3	8	2	4	9	3	9
3	12	6	36	9	3	9
3	8	2	4	9	3	9
Summa	108	0	188	108	0	156
Meðaltal	6		10,4	6	0	8,7

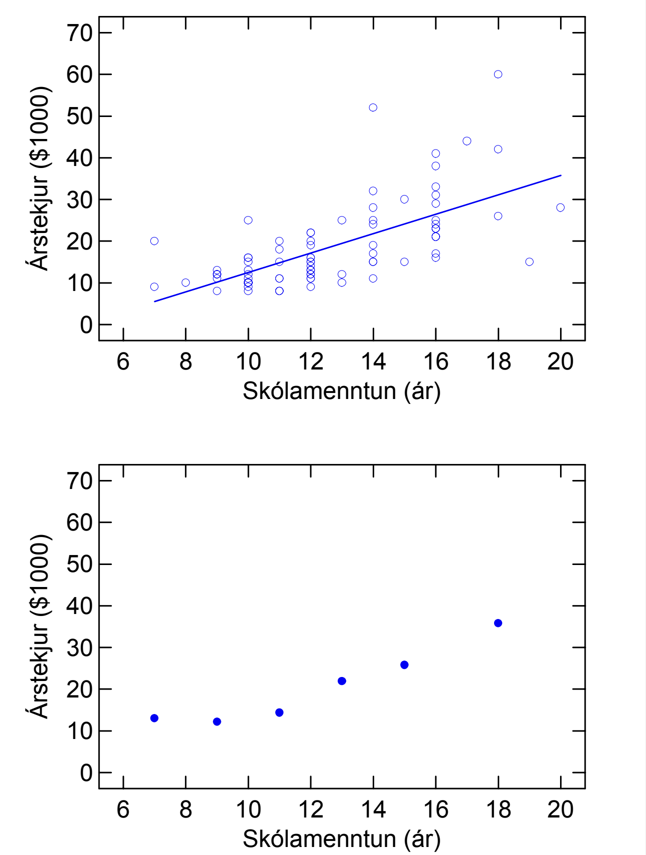
Notkun eta á sveigbogatengsl

Við getum notað eta til að kanna sveigbogatengsl samfelldra breyta

- Með því að skipta frumbreytu í flokka má reikna meðaltal fylgibreytunnar í hverjum flokki og fá þannig mynd af tengslum breytanna.

Dæmi: Tengsl skólagöngu og tekna

- Flokkum frumbreytuna (ár í skóla) í flokka; 6-8, 8-10, …. Síðasti flokkurinn er hafður breiðari vegna lítilla gagna.
- Reikna meðaltöl fyrir hvern flokk og síðan η.
- Niðurstaða: $r(78)= 0,65$; $r^2= 0,42$; $\eta= 0,69$; $\eta^2= 0,48$
- $\eta^2 >> r^2$ og meðaltölin lýsa línunni því lítillega betur en bein lína.



K: VI.1, bls. 194–198; VI.2, bls. 203

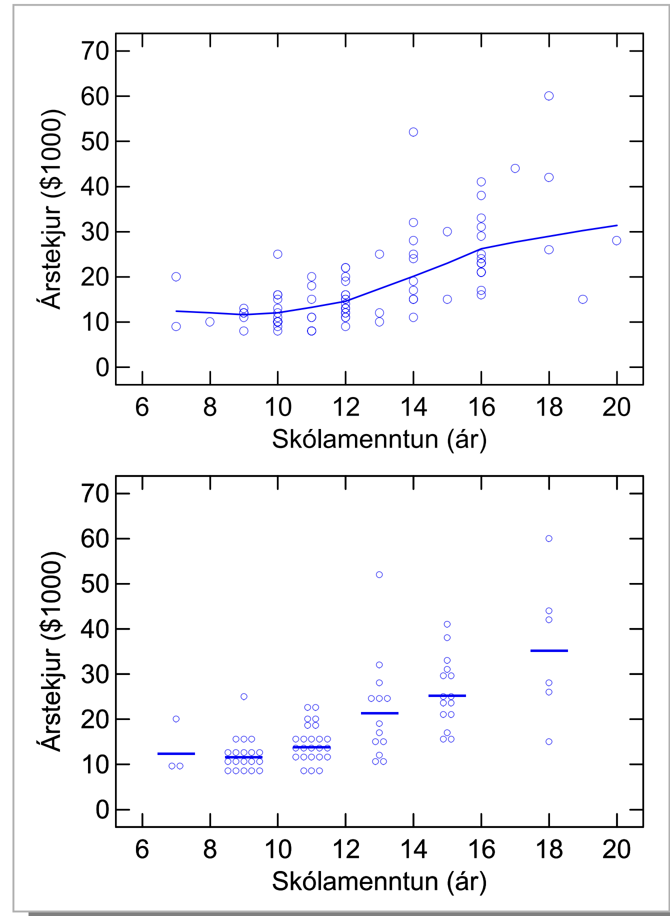
Meðaltöl sem spágildi

Meðaltölin í hverjum flokki gefur bestu spá, dreifingin gefur óvissuna

- η^2 gefur til kynna styrkleika tengslanna—hlutfallslega skýringu

Gæta þarf að flokkastærð

- Ef flokkarnir eru jafnmargir mæligildum verður $\eta^2 = 1{,}0$
- Kohout mælir með a.m.k. 5–6 mæligildum í hverjum flokki
- Almennt séð er þó rétt að velja flokkana þannig að meðaltölin lýsi tengslunum vel.
- Þótt töluvert huglægt mat fylgir þessari aðferð leiðir hún oft til sennilegri tengsla heldur en bein lína og því til sennilegri spár um fylgibreytuna.



K: VI.1, bls. 195–196; VI.2, bls. 202

Almennt ályktanir um η

Notkun við mat sveigbogatengsla

Munur á r^2 og η^2 gefur vísbendingu um frávik frá beinlínutengslum samfelldra breyta.

- $\eta^2 \geq r^2$; eta í öðru veldi er aldrei lægri en r^2
- $\eta^2 - r^2$ er mælikvarði á frávik frá beinlínutengslum
  - $\eta^2 - r^2 \approx 0$; bein lína lýsir tengslunum vel
  - $\eta^2 - r^2 >> 0$; bein lína lýsir tengslunum illa

Þótt bera megi eta (η) saman við r, hefur það (ólíkt η^2) enga skýra túlkun.

Notkun við samanburð hópa

Etu má nota þegar frumbreytan er rofin og fylgibreytan nær samfelld.

- Gefur til kynna hve mikið af heildardreifingunni er vegna ólíkra hópmeðaltala.
- Eta tekur aðeins tillit til eigindlegra upplýsinga í frumbreytu
  - Við getum notað etu þótt frumbreytan sé megindleg (t.d. aldurshópar)
  - Þá lítum við framhjá þeim upplýsingum

K: VI.2, bls. 203