

Hlutfylgni

Fyrirlestur í Aðferðafræði II

© 1998, 2000, 2001, 2003, 2004 Guðmundur 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.

Her- og heilbrigðisútgjöld

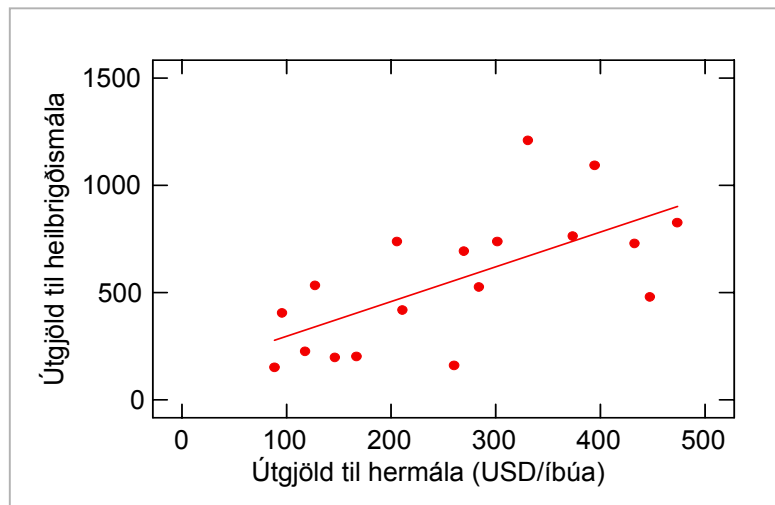
Á myndinni sjást greinilegt tengsl her- og heilbrigðisútgjalda hjá 18 Evrópuþjóðum.

Við gerum ráð fyrir að tengslin séu ekki vegna hreinnar tilviljunar.

Áhrifin eru umtalsverð; útgjöld til heilbrigðismála eykst um \$1,62 fyrir hvern dal sem fer í hermál.

Tengslin eru tiltölulega sterk; herútgjöld skýra 42% af dreifingu heilbrigðisútgjalda.

En eru þetta orsakatengsl?



$$\text{Heilbrigðisútgjöld} = 133,5 + 1,623 \cdot \text{Herútgjöld}$$
$$r(16) = 0,65; r^2 = 0,42$$

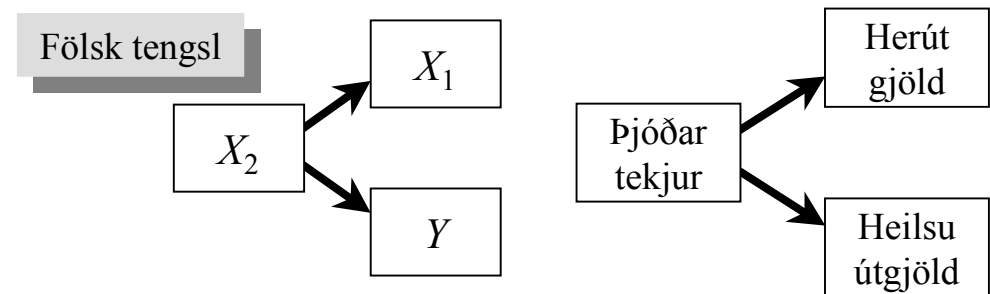
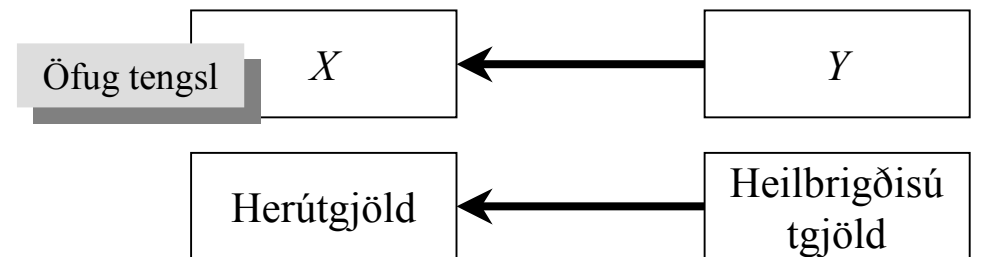
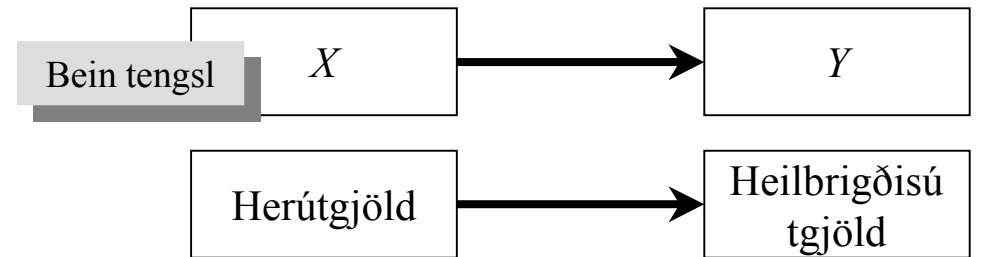
Tvíbreytu orsakatengsl

Við vitum það eitt að það er fylgni á milli breytanna tveggja. Það segir okkur ekki hvort eða hvers konar orsakasamband sé til staðar.

Það er hugsanlegt að herútgjöld orsaki heilbrigðisútgjöld.

Einnig gætu heilbrigðisútgjöld valdið herútgjöldum.

Þetta geta einnig verið fölsk tengsl (*spurious association*). Einhver þriðja breyta (X_2) hefur áhrif á hinar tvær (X_1 og Y) en X_1 hefur engin áhrif á Y .



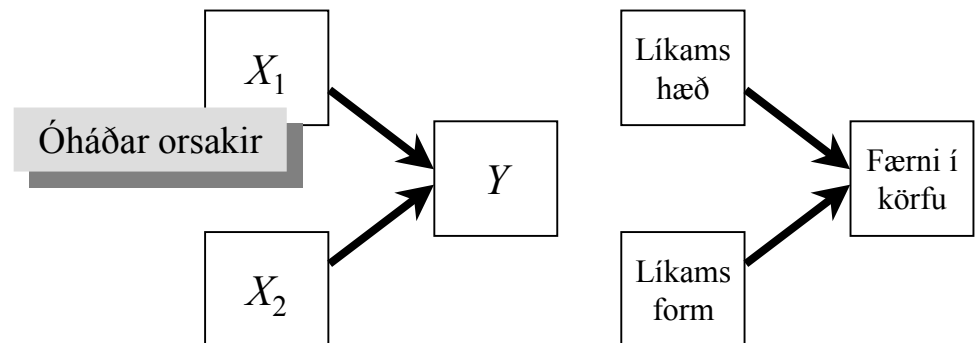
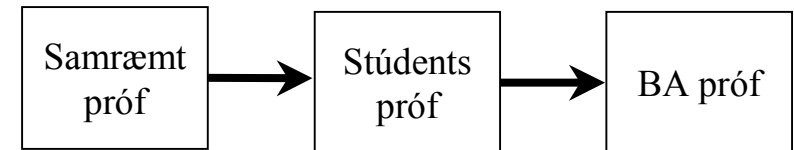
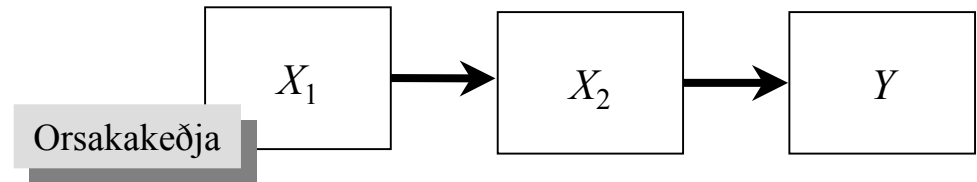
Þrúbreytutengsl

Millibreyta (*intervening variable*) er oft á milli frum- og fylgibreytu. Frumbreytan hefur þá óbein (*indirect*) áhrif í gegnum millibreytuna.

Í orsakakeðju (*chain relationship*) er fylgni milli X_1 og Y þó svo að frumbreytan hafi engin bein áhrif heldur aðeins óbein áhrif (*effect*) í gegnum millibreytuna.

Millibreytan breytir túlkun á tengslunum og því mikilvægt að uppgötva millibreytutengsl.

Fjölmarginar orsakir geta haft óháð (*independent*) áhrif á fylgibreytu.



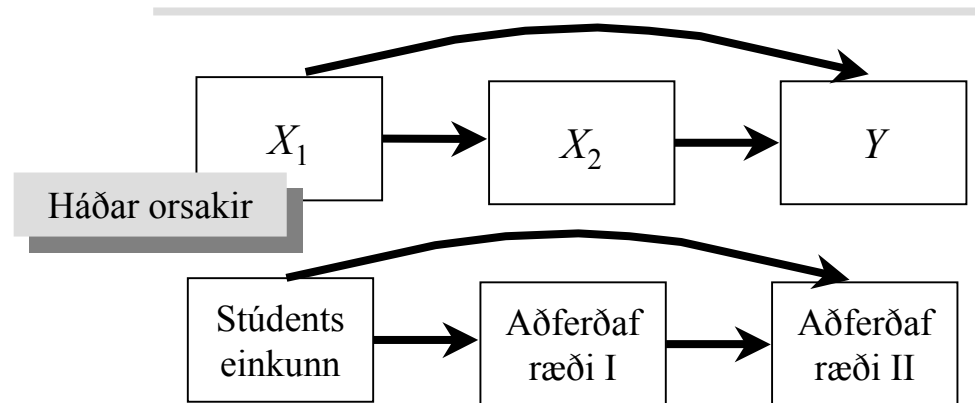
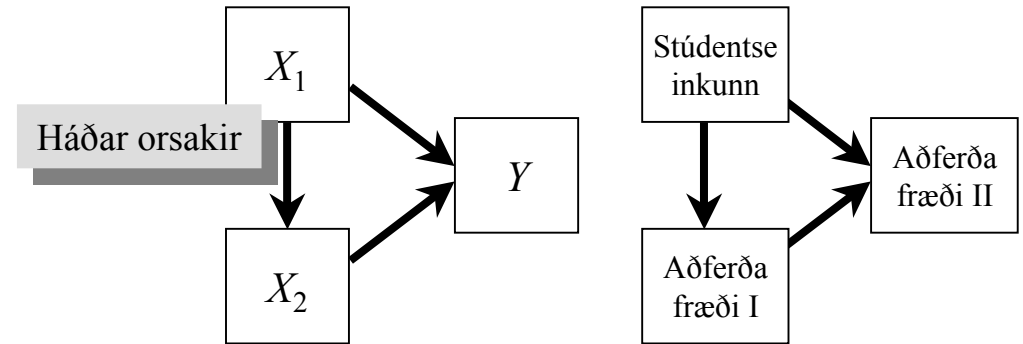
Háðar orsakir

Algennt er að margar innbyrðis háðar frumbreytur hafi áhrif á fylgibreytu.

Það er þægilegt að setja slík sambönd upp svipað og keðju til að átta sig á beinum og óbeinum áhrifum frumbreytanna.

Hér hefur stúdentseinkunn bein áhrif á einkunn í Aðferðafræði II en einnig óbein áhrif í gegnum einkunn í Aðferðafræði I.

Margþætt tengsl beinna og óbeinna orsaka er víða að finna innan félagsvísinda.



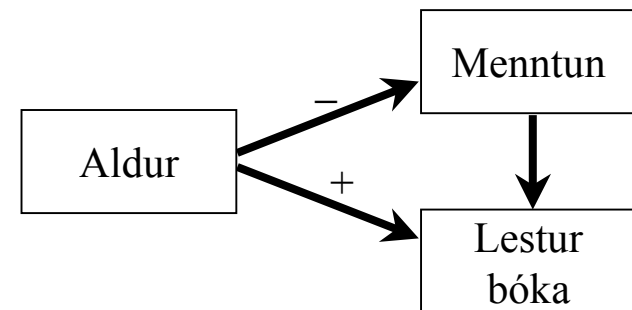
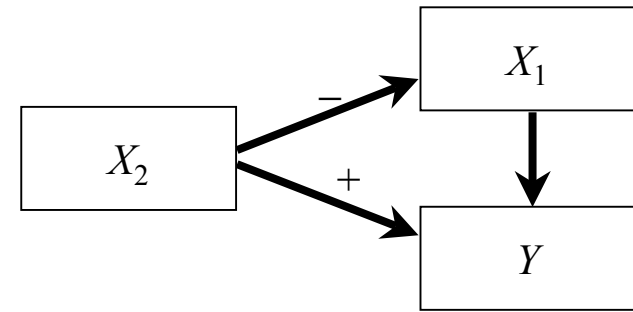
Bæling (*suppression*)

Bæling felst í því að grunnfylgni (*zero-order correlation*) er lág jafnvel þótt frumbreyta hafi umtalsverð áhrif á fylgibreytu.

Á neðri myndinni eykst bóklestur með aukinni menntun. Samt getur grunnfylgni menntunar og bóklesturs verið lítil eða engin.

Aukinn aldur er samfara minni menntun og auknum bóklestri og því verður breytileiki í aldri til að hylja áhrif menntunar á bóklestur.

Ef tengslin væru skoðuð fyrir hvert aldursár myndu tengslin koma í ljós.



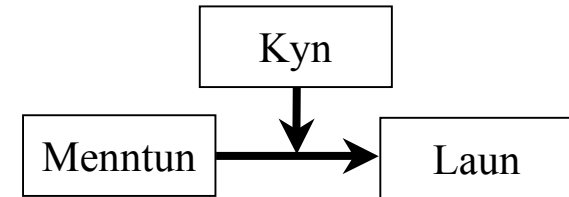
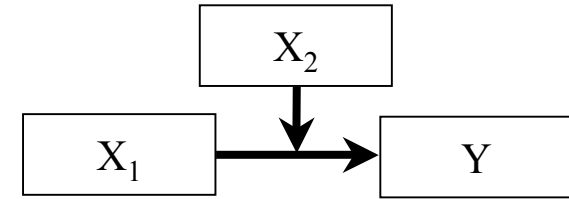
Samvirkni (*interaction*)

Ef áhrif frumbreytu á fylgibreytu fara eftir gildi þriðju breytu er talað um samvirkni.

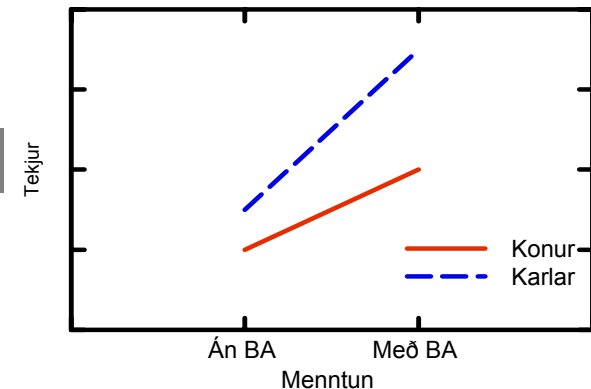
Samvirkni getur falist í því að:

- tengsl séu aðeins fyrir hendi í vissum tilfellum en ekki öðrum
- áhrifin séu stundum jákvæð og stundum neikvæð
- áhrifin séu mismikil.

Neðsta myndin sýnir mismikil áhrif menntunar á laun eftir kynjum, þ.e. samvirkniáhrif menntunar og kyns á tekjur.



Tilbúið dæmi



Orsakagreining

Vísindi leitast við að skýra tengsl breyta. Í því felst m.a. að finna orsakatengsl breyta.

Ýmis viðmið eru notuð til að ákvarða um orsakatengsl. Öll eftirfarandi viðmið þurfa að vera til staðar til að sýna fram á orsakasamband:

- Tengsl séu til staðar
- Orsök sé á undan afleiðingu í tíma
- Annarra skýringa hefur verið leitað og tekist að hafna þeim

Lykilatriði í orsakagreiningu er að kanna áhrif „þriðju breytu.“

Fölsk tengsl: Fylgni samfara engum tengslum.

Bæling: Engin fylgni samfara tengslum.

Orsakakeðja: Fylgni samfara tengslum í gegnum millibreytu.

Háðar orsakir: Fylgni en hluta tengsla í gegnum millibreytu.

Samvirkni: Fylgni en tengsl háð gildum þriðju breytu.

Áhrif þriðju breytu

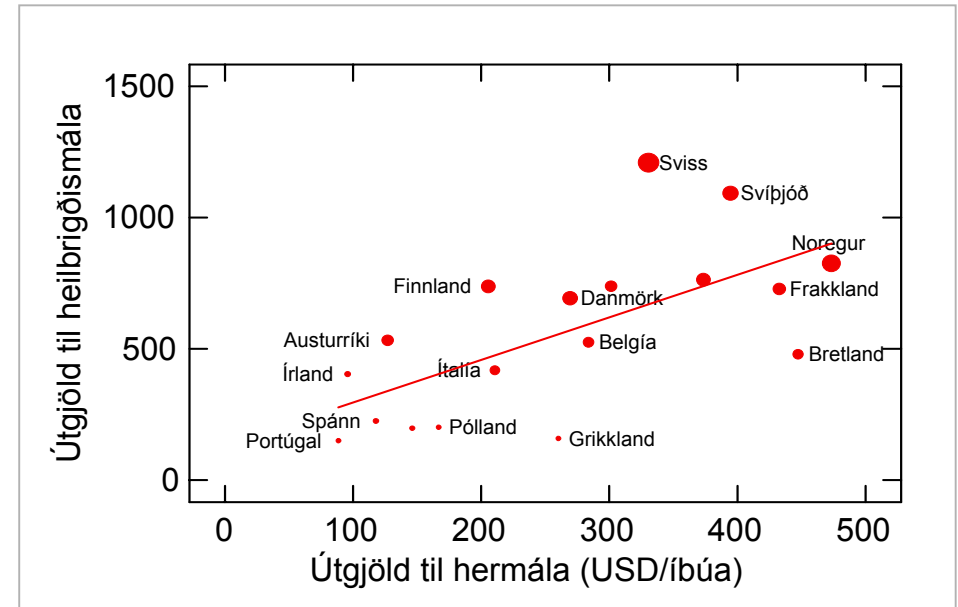
Í okkar dæmi kemur til álita að þjóðartekjur skapi fölsk tengsl her- og heilbrigðisútgjalda.

Á myndinni eru punktarnir stærri eftir því sem þjóðartekjur aukast.

Ríkar þjóðir (stórir punktar) virðast vera hægra megin og ofarlega á myndinni.

Fylgnifylkið staðfestir að þjóðartekjur hafa fylgni við bæði her- og heilbrigðisútgjöld þjóða.

Því koma fölsk tengsl til álita, þ.e. að engin orsakatengsl séu milli her- og heilbrigðisútgjalda.



	Heilbrigði	Hermál	Tekjur
Heilbrigði	1,0		
Hermál	0,65	1,0	
Tekjur	0,94	0,68	1,0

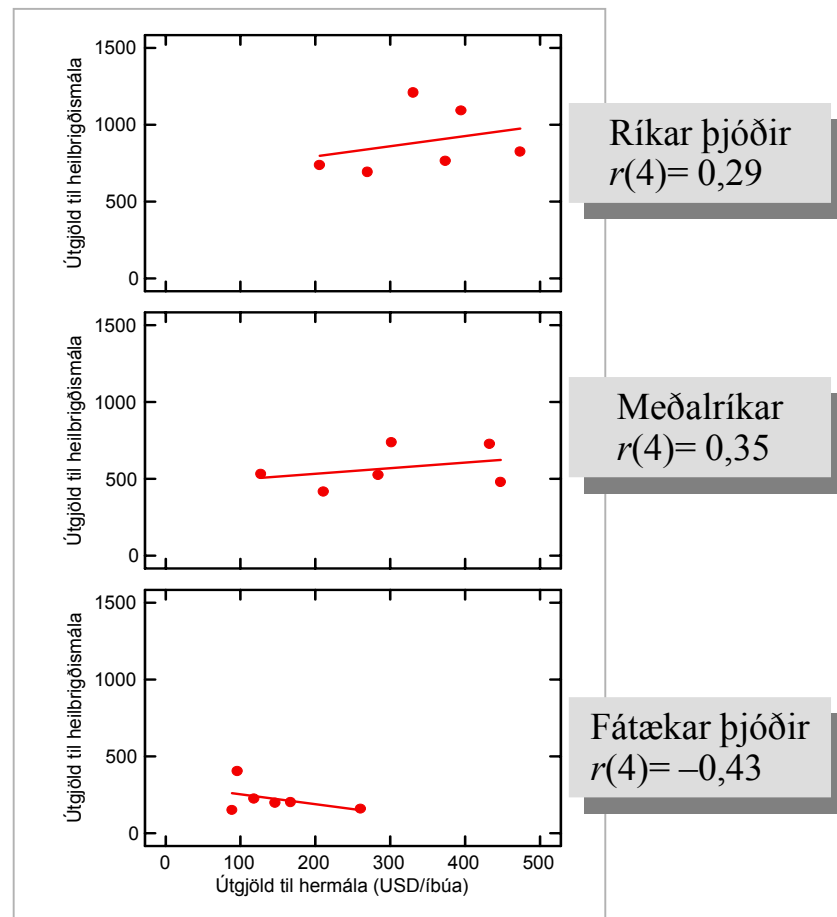
Leiðrétt fyrir þriðju breytu

Ef þjóðartekjur valda fylgni her- og heilbrigðisútgjalda ætti hún að hverfa fyrir jafnríkar þjóðir. Þetta er kallað að halda þriðju breytu „fastri“ eða „leiðrétt“ fyrir hana.

Á myndinni sést að fylgnin minnkar mikið ef hún er skoðuð sérstaklega fyrir ríkar, meðalríkar og fátækar þjóðir.

Í hverjum myndreit eru þó enn misríkar þjóðir og þjóðartekjum því ekki haldið alveg föstum. Það er því ekki fyllilega leiðrétt fyrir þjóðartekjur.

Ef nákvæmlega jafnríkar þjóðir væru í hverjum hópi fyrir sig, væri leiðrétt fullkomlega fyrir þjóðartekjur með þessari framsetningu.



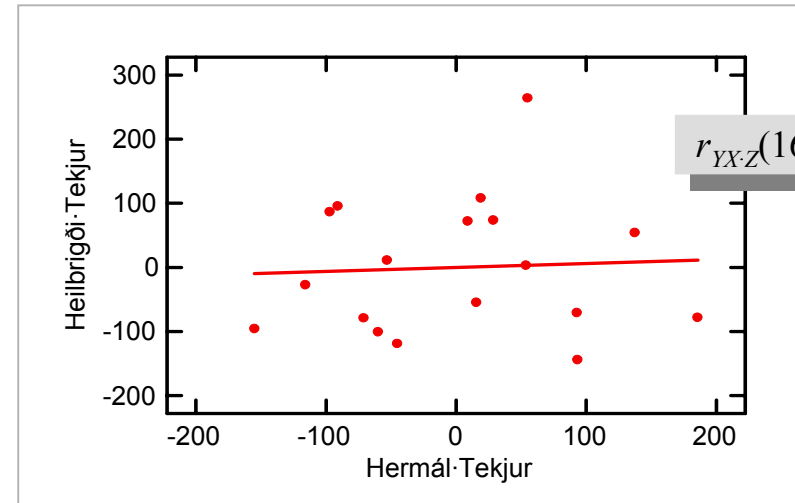
Tölfræðileg leiðrétting

Tölfræðileg leiðrétting leiðrétir að fullu fyrir áhrif þriðju breytu. Það má einnig gera með því að finna leif (*residual*) beggja breyta í tveimur aðfallsgreiningum þar sem þriðja breytan er frumbreyta.

Myndin sýnir að fylgni slíkra leifa er lítil sem engin, sem bendir til þess að skilyrt tengsl breytanna séu lítil.

Einfaldara er að reikna hlutfylgni ($r_{YX \cdot Z}$) þar sem fylgnin er leiðrétt fyrir áhrif þriðju breytunnar.

Báðar leiðir gefa sömu niðurstöðu. Við gerum í báðum tilvikum ráð fyrir að öll tengslin fylgi beinum línunum.



$$r_{xy \cdot z} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{(1 - r_{xz}^2)(1 - r_{yz}^2)}}$$

Útreikningar

Við notum fylgnitölurnar til aðfylla inn í formúluna.

Þegar leiðrétt er tölfræðilega fyrir áhrif þjóðartekna verða tengsl útgjalda til heilbrigðis- og hermála lítil sem engin.

Þetta bendir til falskra tengsla, þ.e. að þjóðartekjur hafi áhrif á báðar útgjaldategundir en áhrif þeirra innbyrðis séu engin.

$$\begin{aligned} r_{\text{Her Heilbrigði-Þjóðartekjur}} &= \frac{r_{yx} - r_{xz} \cdot r_{yz}}{\sqrt{(1 - r_{xz}^2)(1 - r_{yz}^2)}} \\ &= \frac{0,650 - 0,676 \cdot 0,943}{\sqrt{(1 - 0,676^2)(1 - 0,943^2)}} \\ &= \frac{0,650 - 0,676 \cdot 0,943}{\sqrt{(1 - 0,457)(1 - 0,889)}} \\ &= \frac{0,650 - 0,6375}{\sqrt{0,543 \cdot 0,111}} = \frac{0,0125}{\sqrt{0,060}} \\ &= \frac{0,0125}{0,245} = 0,051 \approx 0,05 \end{aligned}$$

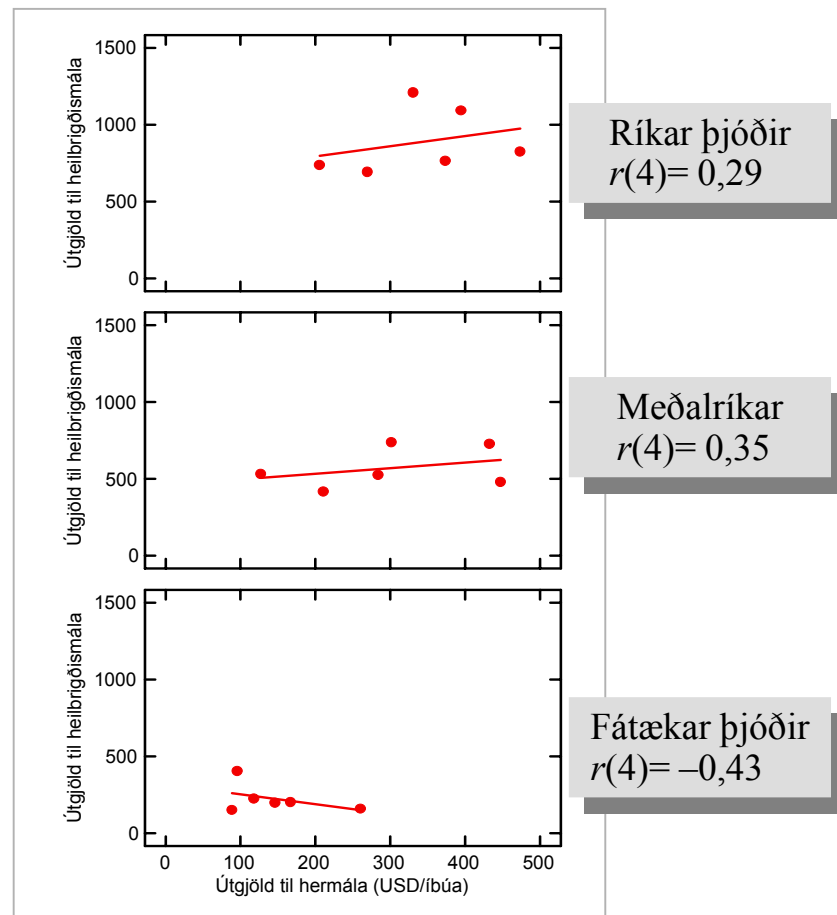
Samvirkni og hlutfylgni

Ef myndin er skoðuð nánar virðist sem þjóðartekjur og herútgjöld gætu haft samvirkniáhrif á heilbrigðisútgjöld.

Ef samvirknin er raunveruleg, þ.e. ekki eingöngu vegna tilviljunar, og umtalsverð, gefur hlutfylgnin villandi mynd af tengslunum.

Hér eru neikvæð tengsl milli útgjalda til heilbrigðis- og hermála hjá fátækum þjóðum en jákvæð hjá efnadri þjóðum. Hlutfylgnin $r_{YX \cdot Z} = 0,05$ er því eins konar meðaltal sem gefur villandi mynd af tengslum breytanna þriggja.

Í stað þess að engin tengsl séu milli her- og heilbrigðisútgjalda að teknu tilliti til efnahags, er hugsanlega samvirkni.


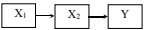





Hlutfylgni og helstu orsakalíkön

Hlutfylgnireikningar geta leiðbeint um þau þau tengsl orsaka sem liggja að baki tengslum breyta.

Með því að bera saman hlut- og grunnfylgni og hafa hliðsjón af þekkingu sinni af eðli breytanna er hægt að ráða í hvers eðlis orsakatengslin eru.

Þriðju breytur eru þó fjölmargar og flest tengsl eru margbreytu- fremur en þríbreytutengsl og því er orsakagreining viðvarandi úrlausnarefni innan aðferðarfræði.

Niðurstaða	Túlkun	Líkan
$r_{xy \cdot z} \approx 0$	fölsk tengsl (spurious relationship)	
- eða -	leiðrétt með millibreytu	
$r_{xy \cdot z} \approx r_{xy}$	tengsl óháð þriðju breytunni	
$r_{xy \cdot z} > r_{xy}$	bæliáhrif (suppression) þriðju breytu	
$r_{xy \cdot z} < r_{xy}$	minnkuð tengsl vegna „leiðréttingar“	

Táknun

X_1 eða X : Frumbreyta

X_2 eða Z : Þriðja breyta

Y : Fylgibreyta