

# Hagnýtar hagamælingar: Almennar hugleiðingar

Helgi Tómasson

29-08-2016

## Marmkið og skipulag námskeiðs

- Tilgangur námskeiðs er að þjálfa þátttakendur í að miðla boðskap og upplýsingum.
- Hver nemandi á að fá a.m.k. 3 tækifæri.
- Fyrsti nemandafyrirlestur á að vera mjög fljótlega. Hann á að vera mjög stuttur (5-10 mínútur) þar sem nemandi á að kynna sig, sín áhugamál og hvað hann vill helst sýsla við. Það á að setja upp glærur með titli nafni og nokkrum atriðum.
- Í næsta fyrirlestri á að skilgreina einhverja rannsókn, vísindaspurningu, gagnavinnslu og hugsanlega nálgun á vandanum. Það á að koma fram gróflega hvers vegna þetta er áhugavert, hvað sé vitað og hvernig ætlunin er að nálgast vandann.
- Síðan á að skila ritgerð um rannsóknina. Hún þarf ekki að vera löng en þarf að uppfylla kröfur um uppsetningu, heimildir, númerakerfi á myndum og töflum o.s.frv.
- Hver nemandi heldur síðan lokafyrirlestur um ritgerðina. Hvað í áætlun gekk eftir, niðurstöður, framhald.
- Ekkert próf er en fólk á að mæta stundvíslega í tíma og vera virkt í tímum.

- Í fyrra studdist ég við „Mastering Metrics“ Angrist & Pische (2015). Það er barnaútgáfa af „Mostly Harmless Econometrics“ Angrist & Pische (2009).
- Áður notaði ég „Macroeconomic Patterns and Stories“ Leamer (2008). Lykilfrásinn í þeirri bók er „we are pattern seeking storytelling animals“.
- Í fyrra hafði ég *causality* sem tema í námskeiðinu. Hvað þýðir það að eitthvað orsaki eitthvað? Þetta er heilmikið rætt í Angrist & Pische (2015) og Angrist & Pische (2009).
- Í gagnagreiningu er oft undirliggjandi áhugi á að greina einhvers konar orsakasamhengi. Í ekonómétríunni felst stundum einhvers konar „causal“ ályktun þó svo að *causality* sé ekki nákvæmlega skilgreint.
- Þið getið googlað höfunda eins og t.d. Lauritzen og Pearl um causality og „*Graphical Models*“. Í því umhverfi er reynt að setja fram stærðfræðilegar skilgreiningar.
- Í vetur ætla ég að flytja nokkra fyrirlestra um bayesískar aðferðir og byggja á bókinni „*Bayesian Econometrics*“ Koop (2003). Ég ætla að æfa mig í reiknitæknilegum atriðum, octave/pylab/julia.

## Saga úr Angrist & Pische (2015)

- Af hverju hrökklast fólk úr námi?
- Hafa námslánin eitthvað með þetta að gera?
- Sumir segja að stórar skuldir aftri námsmönnum og að fólki gangi betur á erfðum fjármunun heldu en á lánum.
- Nemendum með uppruna í lægri tekjuhópum er af ýmsum ástæðum hættara við brottfalli.
- Brottfallsástæðan virðist ekki endilega tengd skuldastöðu.
- Menn vilja einhvers konar „*other things equal (ceteris paribus)*“ samanburð.
- Vandinn (í observational stúdium) er að það erfitt að stilla upp *other things equal*.
- Þó að það sé erfitt, er það ekki alltaf ómögulegt. Ekónómetrísk fræði kenna okkur um *omitted variables bias, selection bias* og *instrumental variables* aðferðir.

# Dæmi um niðurstöður: Borgar heilbrigðistrygging sig?

Randomized Trials 5

TABLE 1.1  
Health and demographic characteristics of insured and uninsured couples in the NHIS

	Husbands			Wives		
	Some HI (1)	No HI (2)	Difference (3)	Some HI (4)	No HI (5)	Difference (6)
A. Health						
Health index	4.01 [.93]	3.70 [1.01]	.31 (.03)	4.02 [.92]	3.62 [1.01]	.39 (.04)
B. Characteristics						
Nonwhite	.16	.17	-.01 (.01)	.15	.17	-.02 (.01)
Age	43.98	41.26	2.71 (.29)	42.24	39.62	2.62 (.30)
Education	14.31	11.56	2.74 (.10)	14.44	11.80	2.64 (.11)
Family size	3.50	3.98	-.47 (.05)	3.49	3.93	-.43 (.05)
Employed	.92	.85	.07 (.01)	.77	.56	.21 (.02)
Family income	106,467	45,656	60,810 (1,355)	106,212	46,385	59,828 (1,406)
Sample size	8,114	1,281		8,264	1,131	

Notes: This table reports average characteristics for insured and uninsured married couples in the 2009 National Health Interview Survey (NHIS). Columns (1), (2), (4), and (5) show average characteristics of the group of individuals specified by the column heading. Columns (3) and (6) report the difference between the average characteristic for individuals with and without health insurance (HI). Standard deviations are in brackets; standard errors

More often than not, however, such comparisons are misleading.

### *Fruitless and Fruitful Comparisons*

Simple comparisons, such as those at the top of Table 1.1, are often cited as evidence of causal effects. More often than not, however, such comparisons are misleading. Once again the problem is *other things equal*, or lack thereof. Comparisons of people with and without health insurance are not apples to apples; such contrasts are apples to oranges, or worse.

Among other differences, those with health insurance are better educated, have higher income, and are more likely to

Tekið úr Angrist & Pische (2015, bls. 4-5).

## Kenningaprófun er flókin

- Á fyrri hluta 20. aldar deildu R.A. Fisher og J. Neyman/E.S. Pearson um kenningaprófanir og hvað p-gildi þýddu. *The phrase "Errors of the second kind", although apparently only a harmless piece of technical jargon, is useful as indicating the type of mental confusion in which it was coined* (Fisher, 1955).
- Neyman/Pearson dáðust að kröfu Fishers um "model", og því er það síðar varð einn aðalásteitingarsteinn þeirra. Þ.e. hvernig skilgreina ætti hugtakið módel.

- Lindley 1999 segir: *My personal view is that  $p$ -values should be relegated to the scrap heap and not considered by those who wish to think and act coherently.*
- Spanos (1999, page 725) segir um deilur Fishers og Neyman: *Fisher and Neyman would turn in their graves if they were to find out how the modern textbook hybrid on hypothesis testing managed to reconcile what they considered irreconcilable differences!* .
- Hubbard & Lindsay (2008, page 82) on Fisher's  $p$ -values: *his widely misunderstood and defective  $p$ -values blanket the empirical literature.*
- Gigerenzer [The empire of chance, p. 106] segir: Although the debate [Fisher Neyman] continues among statisticians, it was silently resolved in the "cookbooks" written in the 1940s to the 1960s, largely by non-statisticians, to teach students in the social sciences statisticians, "the rules of statistics".
- We call this compromise the "hybrid theory" and it goes without saying that neither Fisher nor Neyman and Pearson would have looked with favor on this offspring of their forced marriage. (Spanos síða 727)



- Bayesistar leyfa að líkur séu notaðar sem mælikvarði á vissu.
- Segjum að samkvæmt  $H_0$  sé dreifing  $X$   $f_0(x)$  og samkvæmt  $H_1$  þá sé dreifing  $X$   $f_1(x)$ . Líkur á að  $H_0$  sé rétt eru  $\pi_0$  og líkur á að  $H_1$  sér rétt séu  $\pi_1$ ,  $\pi_0 + \pi_1 = 1$ . Mælingar á  $X$  eru  $x$  og þá gildir skv. reglu Bayes að:

$$P(H_0\text{sönn}|X = x) = \frac{\pi_0 f_0(x)}{\pi_0 f_0(x) + \pi_1 f_1(x)}$$

- Maður skrifar stundum:

$$\frac{P(H_0\text{sönn}|X = x)}{P(H_1\text{sönn}|X = x)} = \frac{\pi_0 f_0(x)}{\pi_1 f_1(x)}$$

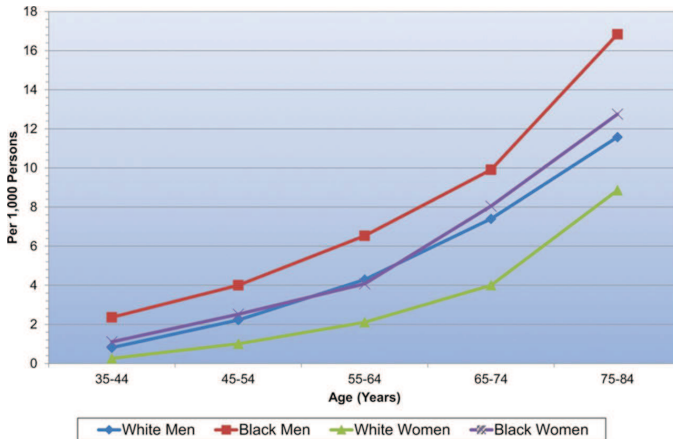
- $\frac{\pi_0}{\pi_1}$  = prior-odds,  $B = \frac{f_0(x)}{f_1(x)}$  = Bayes-factor.
- Ef  $B > 1$  þá hefur trú mín á  $H_0$  aukist.

$$X \sim B(n, \theta), \quad H_0 : \theta = \theta_0, \quad H_1 : \theta \neq \theta_0.$$

Verð að skilgreina a priori dreifingu undir  $H_1$ . Segjum að undir  $H_1$ , sé  $\theta \sim U(0, 1)$ .

- Segjum að  $n = 100$ ,  $\theta_0 = 1/2$  og að  $x = 60$ , þ.e.  $\hat{p} = 0.6$ . Með venjulegri fyrsta árs tölfræði er  $H_0$  hafnað, en  $B = 1.09$ . Þ.e. samkvæmt Bayes aðferðum hefur stuðningur við  $H_0$  aukist en frequentisti segir  $\theta$  er marktækt frábrugðið  $1/2$ . Ef  $n$  er stækkað eykst stuðningur Bayesistans við  $H_0$ . (Young & Smith, 2005)
- Þetta er kennslubókardæmi. Í aðalfyrirlestri á Norræni tölfræðiráðstefnu varaði Gelman við svona í túlkunum á heilbrigðisrannsóknum. Hann tók t.d. dæmi þar sem  $OR=3$ , og 95% öryggismörk t.s. (1.2-8.5).
- Prívat sá ég nýlega fyrirlestur þar sem sagt var frá því að neysla á lýsi yki hættu á blöðruhálskrabbameini. Dæmið hans Gelmans var þar í ýktri mynd. Í fyrirlestrinum var vitnað í varasaman hlut, „Meta-analýsu“, sem varað er við í nýlegri kennslubók eftir Boos og Stefanski (Boos & Stefanski, 2013).

## Incidence of heart attack or fatal coronary heart disease by age, sex, and race



Atherosclerosis Risk in Communities Surveillance: 2005–2011.



Mozaffarian D et al. *Circulation*. 2015;131:e29-e322

Copyright © American Heart Association, Inc. All rights reserved.

- Angrist, J. & Pische, J.-S. (2009). *Mostly Harmless Econometrics*. Princeton University Press.
- Angrist, J. & Pische, J.-S. (2015). *Mastering Metrics*. Princeton University Press.
- Boos, D. & Stefanski, L. (2013). *Essential Statistical Inference: Theory and Methods*. Springer New York Heidelberg.
- Fisher, R. (1955). Statistical methods and scientific induction. *Journal of the Royal Statistical Society. Series B (Methodological)*, 17(1), pp. 69–78.
- Hubbard, R. & Lindsay, R. M. (2008). Why p values are not a useful measure of evidence in statistical significance testing. *Theory & Psychology*, 18(1), 69–88.
- Koop, G. (2003). *Bayesian Econometrics*.
- Leamer, E. (2008). *Macroeconomic Patterns and Stories*. Business and Economics. Springer.
- Spanos, A. (1999). *Probability Theory and Statistical Inference*. Cambridge University Press.
- Young, G. & Smith, R. (2005). *Essential of Statistical Inference*. Cambridge University Press.