



# Rökfræði tölfræðilegra prófa: Saga og hugmyndir

**Helgi Tómasson**

`mailto:helgito@hi.is`

**19. maí 2009**

Markmið þessa fyrirlestrar er að rekja sögu og kynna hugmyndafræði við ályktanir með áherslu á þær hugmyndir sem liggja að baki tölfræðilegum kenningprófunum.

Markmið þessa fyrirlestrar er að rekja sögu og kynna hugmyndafræði við ályktanir með áherslu á þær hugmyndir sem liggja að baki tölfræðilegum kenningprófunum.

# 1 Skipulag

- Sögulegt yfirlit

# 1 Skipulag

- Sögulegt yfirlit
- Tölfræðileg hugmyndafræði

# 1 Skipulag

- Sögulegt yfirlit
- Tölfræðileg hugmyndafræði
- Form kenningaprófana

# 1 Skipulag

- Sögulegt yfirlit
- Tölfræðileg hugmyndafræði
- Form kenningaprófana
- Nokkrar tilvitnanir

# 1 Skipulag

- Sögulegt yfirlit
- Tölfræðileg hugmyndafræði
- Form kenningaprófana
- Nokkrar tilvitnanir
- Pre-test aðferðir, gagnaköfun



# 1 Skipulag

- Sögulegt yfirlit
- Tölfræðileg hugmyndafræði
- Form kenningaprófana
- Nokkrar tilvitnanir
- Pre-test aðferðir, gagnaköfun  
, algengur praxís

# 1 Skipulag

- Sögulegt yfirlit
- Tölfræðileg hugmyndafræði
- Form kenningaprófana
- Nokkrar tilvitnanir
- Pre-test aðferðir, gagnaköfun  
, algengur praxís
- Hlutverk líkans

# 1 Skipulag

- Sögulegt yfirlit
- Tölfræðileg hugmyndafræði
- Form kenningaprófana
- Nokkrar tilvitnanir
- Pre-test aðferðir, gagnaköfun  
, algengur praxís
- Hlutverk líkans
- Staðan í dag, tíðnitúlkun og Bayes tölfræði

# 1 Skipulag

- Sögulegt yfirlit
- Tölfræðileg hugmyndafræði
- Form kenningaprófana
- Nokkrar tilvitnanir
- Pre-test aðferðir, gagnaköfun  
, algengur praxís
- Hlutverk líkans
- Staðan í dag, tíðnitúlkun og Bayes tölfræði
- Lokaorð

# 1 Skipulag

- Sögulegt yfirlit
- Tölfræðileg hugmyndafræði
- Form kenningaprófana
- Nokkrar tilvitnanir
- Pre-test aðferðir, gagnaköfun  
, algengur praxís
- Hlutverk líkans
- Staðan í dag, tíðnitúlkun og Bayes tölfræði
- Lokaorð

- Hugmyndir eru m.a. unnar úr Poirier (1995), Spanos (1999), Cox & Hinkley (1974) og Young & Smith (2005)

# Saga

- Tölfræði, ályktunarfræði byggir á líkindafræði sem er stærðfræðigrein

# Saga

- Tölfræði, ályktunarfræði byggir á líkindafræði sem er stærðfræðigrein
- Líkindafræðin er stærðfræðigrein því þarf ekki að túlka hvað líkur mæla.



# Saga

- Tölfræði, ályktunarfræði byggir á líkindafræði sem er stærðfræðigrein
- Líkindafræðin er stærðfræðigrein því þarf ekki að túlka hvað líkur mæla.

Líkindafræði byggir á „measure-theory“, þ.e. stærðir mengja er mæld.

# Saga

- Tölfræði, ályktunarfræði byggir á líkindafræði sem er stærðfræðigrein
- Líkindafræðin er stærðfræðigrein því þarf ekki að túlka hvað líkur mæla.

Líkindafræði byggir á „measure-theory“, þ.e. stærðir mengja er mæld.

Ef að nota á líkindafræði, þá verður að túlka t.d. hvað það þýðir að líkur á einhverju séu  $1/2$ .

# Saga

- Tölfræði, ályktunarfræði byggir á líkindafræði sem er stærðfræðigrein
- Líkindafræðin er stærðfræðigrein því þarf ekki að túlka hvað líkur mæla.

Líkindafræði byggir á „measure-theory“, þ.e. stærðir mengja er mæld.

Ef að nota á líkindafræði, þá verður að túlka t.d. hvað það þýðir að líkur á einhverju séu  $1/2$ .

- Við höfum tíðnitúlkun (frequentista) sem segja að  $p$  mæli hversu oft atburður eigi sér stað í endurteknum tilraunum.

- Við höfum tíðnitúlkun (frequentista) sem segja að  $p$  mæli hversu oft atburður eigi sér stað í endurteknum tilraunum.
- Bayesistar leyfa einnig að líkur sé mælikvarði á vissu.

- Við höfum tíðnitúlkun (frequentista) sem segja að  $p$  mæli hversu oft atburður eigi sér stað í endurteknum tilraunum.
- Bayesistar leyfa einnig að líkur sé mælikvarði á vissu.

- Tölfræði byrjar óformlega á 19. öld með mönnum eins og Galton, og Pearson, eldri (the biometric tradition, (Spanos, 1999)).

- Tölfræði byrjar óformlega á 19. öld með mönnum eins og Galton, og Pearson, eldri (the biometric tradition, (Spanos, 1999)).
- Nútíma tölfræði byrjar með Fischer snemma á 20. öld.



- Tölfræði byrjar óformlega á 19. öld með mönnum eins og Galton, og Pearson, eldri (the biometric tradition, (Spanos, 1999)).
- Nútíma tölfræði byrjar með Fischer snemma á 20. öld.
- Fischer átti í miklum ritdeilum við ýmsa menn, þar á meðal Pearson eldri. Þeir deildu um hugmyndafræði auk tæknilegra atriða.

- Tölfræði byrjar óformlega á 19. öld með mönnum eins og Galton, og Pearson, eldri (the biometric tradition, (Spanos, 1999)).
- Nútíma tölfræði byrjar með Fischer snemma á 20. öld.
- Fischer átti í miklum ritdeilum við ýmsa menn, þar á meðal Pearson eldri. Þeir deildu um hugmyndafræði auk tæknilegra atriða.
- Pearson áttaði sig ekki á að Fischer breytti leikreglum greinarinnar. Pearson taldi að tölfræði væri til að greina gögn

- Tölfræði byrjar óformlega á 19. öld með mönnum eins og Galton, og Pearson, eldri (the biometric tradition, (Spanos, 1999)).
- Nútíma tölfræði byrjar með Fischer snemma á 20. öld.
- Fischer átti í miklum ritdeilum við ýmsa menn, þar á meðal Pearson eldri. Þeir deildu um hugmyndafræði auk tæknilegra atriða.
- Pearson áttaði sig ekki á að Fischer breytti leikreglum greinarinnar. Pearson taldi að tölfræði væri til að greina gögn en Fischer gekk út frá því að menn væru að meta líkan og alhæfa út frá gögnum í gegnum líkanið.

- Heimspeki Fischer er nú ráðandi hvað þetta varðar.

- Heimspeki Fischer er nú ráðandi hvað þetta varðar.

Það á alltaf að hugsa sér eitthvert líkan þegar álykta á út frá gögnum.

- Heimspeki Fischer er nú ráðandi hvað þetta varðar.

Það á alltaf að hugsa sér eitthvert líkan þegar álykta á út frá gögnum.

- Aðalgreinar ályktunarfræði eru punktmat, bilmat og kenningaprófanir

- Heimspeki Fischer er nú ráðandi hvað þetta varðar.

Það á alltaf að hugsa sér eitthvert líkan þegar álykta á út frá gögnum.

- Aðalgreinar ályktunarfræði eru punktmat, bilmat og kenningaprófanir
- Grunnatriði í heimsspeki Fischers var endurtakanleiki. Ályktun átti að gefa reglu í endurteknum tilraunum

- Heimspeki Fischer er nú ráðandi hvað þetta varðar.

Það á alltaf að hugsa sér eitthvert líkan þegar álykta á út frá gögnum.

- Aðalgreinar ályktunarfræði eru punktmat, bilmat og kenningaprófanir
- Grunnatriði í heimsspeki Fischers var endurtakanleiki. Ályktun átti að gefa reglu í endurteknum tilraunum
- Gögn  $\neq$  upplýsingar.



- Heimspeki Fischer er nú ráðandi hvað þetta varðar.

Það á alltaf að hugsa sér eitthvert líkan þegar álykta á út frá gögnum.

- Aðalgreinar ályktunarfræði eru punktmat, bilmat og kenningaprófanir
- Grunnatriði í heimsspeki Fischers var endurtakanleiki. Ályktun átti að gefa reglu í endurteknum tilraunum
- Gögn  $\neq$  upplýsingar.

Upplýsingar eru hversu nákvæmlega er hægt að álykta um óþekkt atriði í líkani.

- Heimspeki Fischer er nú ráðandi hvað þetta varðar.

Það á alltaf að hugsa sér eitthvert líkan þegar álykta á út frá gögnum.

- Aðalgreinar ályktunarfræði eru punktmat, bilmat og kenningaprófanir
- Grunnatriði í heimsspeki Fischers var endurtakanleiki. Ályktun átti að gefa reglu í endurteknum tilraunum
- Gögn  $\neq$  upplýsingar.

Upplýsingar eru hversu nákvæmlega er hægt að álykta um óþekkt atriði í líkani.

- Fischer þróaði einnig „optimalitets“ hugtök, þ.e. að hægt væri að sanna að eitthvað væri best.

- Fischer þróaði einnig „optimalitets“ hugtök, þ.e. að hægt væri að sanna að eitthvað væri best.
- Hugmyndafræði Fischers gekk gegn Bayesismanum sem hafð lífað í frekar óformlegri mynd frá dögum Laplace/Lagrange ofl. í kringum árið 1800.

# Kenningaprófanir

- Kenningaprófun er ein megingrein ályktunarfræðinnar.

# Kenningaprófanir

- Kenningaprófun er ein megingrein ályktunarfræðinnar.
- Hinar greinarnar eru punktmat og bilmat.

# Kenningaprófanir

- Kenningaprófun er ein megingrein ályktunarfræðinnar.
- Hinar greinarnar eru punktmat og bilmat.
- Hér er venjulega verið að álykta um óþekktan stika í líkindadreifingu.

- Gangurinn í kenningaprófun er í grófum dráttum sá að sett er fram núllkenning,  $H_0$ , reiknaðar eru líkur á fenginni útkomu eða ótrúlegri, gefið  $H_0$  rétt. Ef ótrúleg útkoma hefur fengist þá er það túlkað sem vísbending gegn  $H_0$ .
- Fischer hafði  $H_0$  sem aðalatriði og hafði ekki frammi neina sérstaka stefnu hvað gera skyldi ef  $H_0$  væri hafnað.



- Gangurinn í kenningaprófun er í grófum dráttum sá að sett er fram núllkenning,  $H_0$ , reiknaðar eru líkur á fenginni útkomu eða ótrúlegri, gefið  $H_0$  rétt. Ef ótrúleg útkoma hefur fengist þá er það túlkað sem vísbending gegn  $H_0$ .
- Fischer hafði  $H_0$  sem aðalatriði og hafði ekki frammi neina sérstaka stefnu hvað gera skyldi ef  $H_0$  væri hafnað.
- Þar lendir hann í ritdeilum við Neyman og Pearson yngri.

- Gangurinn í kenningaprófun er í grófum dráttum sá að sett er fram núllkenning,  $H_0$ , reiknaðar eru líkur á fenginni útkomu eða ótrúlegri, gefið  $H_0$  rétt. Ef ótrúleg útkoma hefur fengist þá er það túlkað sem vísbending gegn  $H_0$ .
- Fischer hafði  $H_0$  sem aðalatriði og hafði ekki frammi neina sérstaka stefnu hvað gera skyldi ef  $H_0$  væri hafnað.
- Þar lendir hann í ritdeilum við Neyman og Pearson yngri.

- Neyman-Pearson leiddu (með stærðfræði) út setningu um öflugasta mögulega próf.

- Neyman-Pearson leiddu (með stærðfræði) út setningu um öflugasta mögulega próf.
- Neyman-Pearson lögðu áherslu á að kenningaprófun væri ákvörðunarvandi, þ.e. að ef  $H_0$  væri hafnað þá yrði að hafna henni gegn einhverjum valkost og þá nota valkostinn. Öflugasta mögulega tölfræði próf væri öflugt gegn tilteknum valkostum.

- Neyman-Pearson leiddu (með stærðfræði) út setningu um öflugasta mögulega próf.
- Neyman-Pearson lögðu áherslu á að kenningaprófun væri ákvörðunarvandi, þ.e. að ef  $H_0$  væri hafnað þá yrði að hafna henni gegn einhverjum valkost og þá nota valkostinn. Öflugasta mögulega tölfræði próf væri öflugt gegn tilteknum valkostum.
- Fischer og Neyman-Pearson deildu hart um túlkanir á kenningaprófunum.

- Neyman-Pearson leiddu (með stærðfræði) út setningu um öflugasta mögulega próf.
- Neyman-Pearson lögðu áherslu á að kenningaprófun væri ákvörðunarvandi, þ.e. að ef  $H_0$  væri hafnað þá yrði að hafna henni gegn einhverjum valkost og þá nota valkostinn. Öflugasta mögulega tölfræði próf væri öflugt gegn tilteknum valkostum.
- Fischer og Neyman-Pearson deildu hart um túlkanir á kenningaprófunum.

- Þeir myndu áreiðanlega snúa sér nokkra hringi í gröfinni ef þá óraði fyrir hvernig hugmyndir þeirra voru bræddar saman í einhvers konar „hybrid-theory” í kokkabókskennslubókum í tölfræði, sem oft eru skrifaðar af fólki sem ekki er með fagmenntun í tölfræði. „To teach students in social science: the rules of statistics”.

- Þeir myndu áreiðanlega snúa sér nokkra hringi í gröfinni ef þá óraði fyrir hvernig hugmyndir þeirra voru bræddar saman í einhvers konar „hybrid-theory” í kokkabókskennslubókum í tölfræði, sem oft eru skrifaðar af fólki sem ekki er með fagmenntun í tölfræði. „To teach students in social science: the rules of statistics”.
- Þeir yrðu einnig ósáttir við hvernig það sem þeir töldu óásættanlegar andstæður er notað í nútímavísindagreinum, p-gildi síða upp og síða niður.



# Bayesískar nálganir

- Jeffrey's reyndi að finna lendingu með því að bræða bayesískar hugmyndir við kenningar.

# Bayesískar nálganir

- Jeffreys reyndi að finna lendingu með því að bræða bayesískar hugmyndir við kenningar.
- Þekktar eru útfærslur af Jeffeys paradox. Þ.e. það eru tilfelli þar sem hefðbundin „sampling-theory” kenningapróf og Bayes kenningprófi ber ekki saman.

- Bayes kenningprófun er í grófum dráttum að sett er fram núllkenning og síðan eru settar fram líkur,  $\pi_0$ , á að núllkenning sé rétt og líkur,  $\pi_1$ , á að valkenning sé rétt. Síðan er gögnum safnað, og reiknað

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

Posterior odds = Prior odds  $\times$  Bayes factor

- Bayes kenningprófun er í grófum dráttum að sett er fram núllkenning og síðan eru settar fram líkur,  $\pi_0$ , á að núllkenning sé rétt og líkur,  $\pi_1$ , á að valkenning sé rétt. Síðan er gögnum safnað, og reiknað

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

Posterior odds = Prior odds  $\times$  Bayes factor

- Jeffreys stakk upp á að nota BF (Bayes factor) sem vísbendingar. T.d. að ef  $BF > 1$  þá túlkast það sem stuðningur við  $H_0$  en  $BF < 0.01$  túlkast sem gegn  $H_0$ .

- Margir þekkja BIC

=Bayesian-Information-Criterion. Ef tvö líkön eru borin saman með kenningaprófi þá er munurinn á BIC í þeim:

$$\Delta\text{BIC} = -2\log\left(\frac{\max_{M_1} f(x|M_1)}{\max_{M_2} f(x|M_2)}\right) - (k_2 - k_1)\log(n)$$

þar sem  $k_2$  og  $k_1$  flækjustig(fjöldi stika) í líkönum 1 og 2. Þá er:

$$\text{Bayes-factor} \simeq \exp(-\Delta \text{BIC}/2)$$

- Ef binomial líkan er gefið og prófa á t.d.  
 $H_0 : \theta = 1/2$ , þegar  $n = 100$  og  $x=60$ , leiðir það til þess að frequentisti hafnar  $H_0$ . Ef Bayesisti lætur óvissuna undir  $H_1$  vera jafna dreifingu  $(0,1)$  þá túlkar hann útkomuna  $H_0$  í hag.
- Sumir Bayesistar vilja ekki prófa svona.

# Nokkrar tilvitnanir

- "Is it more serious to convict an innocent man or to acquit a guilty? That will depend on the consequences of the error; is the punishment death or fine; what is the danger to the community of released criminals; what are the current ethical views on punishment? From the point of view of mathematical theory all that we can do is to show how the risk of errors can be controlled and minimized. The use of these statistical tools in any given case, in determining just how the balance should be struck, must be left to the investigator." (Neyman and Pearson 1933)





- "The emphasis given to formal tests of significance throughout [R.A. Fisher's] Statistical Methods...has caused scientific research workers to pay undue attention to the results of the tests of significance they perform on their data, particularly data derived from experiments, and too little to the estimates of the magnitude of the effects they are investigating."... "The emphasis on tests of significance and the consideration of the results of each experiment in isolation, have had the unfortunate consequence that scientific workers have often regarded the execution of a test of significance on an experiment as the ultimate objective." (Yates 1951)



- Overemphasis on tests of significance at the expense especially of interval estimation has long been condemned...."(Cox 1977)
- "...There are considerable dangers in overemphasizing the role of significance tests in the interpretation of data."(Cox 1977)

**Meira**

# Meira

- In a survey of papers published in the American Economic Review, the authors found that "59% use the word 'significance' in ambiguous ways at one point meaning 'statistically significantly different from the null,' at another 'practically important' or 'greatly changing our scientific opinion,' with no distinction." (McCloskey and Ziliak 1996)

- "Most statisticians are all too familiar with conversations [that] start:

Q: What is the purpose of your analysis?

A: I want to do a significance test.

Q: No, I mean what is the overall objective?

A (with puzzled look): I want to know if my results are significant.

And so on...." (Chatfield 1991)

- "It is desirable to report the observed values of the test statistics and not just the p values.

- "It is desirable to report the observed values of the test statistics and not just the p values.

The quantitative results being tested, such as mean values, proportions, or correlation coefficients, should be given whether the test was significant or not." ...



- "It is desirable to report the observed values of the test statistics and not just the p values.

The quantitative results being tested, such as mean values, proportions, or correlation coefficients, should be given whether the test was significant or not." ...

Even if there is a large real effect a non-significant result is quite likely if the number of observations is small.

- "It is desirable to report the observed values of the test statistics and not just the p values.

The quantitative results being tested, such as mean values, proportions, or correlation coefficients, should be given whether the test was significant or not." ...

Even if there is a large real effect a non-significant result is quite likely if the number of observations is small.

Conversely, if the sample size is very large a statistically significant result may occur when there is only a small real effect.

The statistical significance should not be taken as synonymous with clinical importance." (Altman, S.M. et al. 1983)

- "The result is that non-statisticians tend to place undue reliance on single 'cookbook' techniques, and it has for example become impossible to get results published in some medical, psychological and biological journals without reporting significance values even if of doubtful validity.

- "The result is that non-statisticians tend to place undue reliance on single 'cookbook' techniques, and it has for example become impossible to get results published in some medical, psychological and biological journals without reporting significance values even if of doubtful validity. It is sad that students may actually be more confused and less numerate at the end of a 'service course' than they were at the beginning, and more likely to overlook a descriptive approach in favour of some inferential method which may be inappropriate or incorrectly executed."  
"(Chatfield 1985)

- Estimates and measures of variability are more valuable than hypothesis tests." (R.M. Cormack, in discussion of Chatfield 1985)

# Sökudólgarnir?

# Sökudólgarnir?

- "Statistics is intimately connected with science and technology, and few mathematicians have experience or understand of methods of either.



# Sökudólgarnir?

- "Statistics is intimately connected with science and technology, and few mathematicians have experience or understand of methods of either. This I believe is what lies behind the grotesque emphasis on significance tests in statistics courses of all kinds;

# Sökudólgarnir?

- "Statistics is intimately connected with science and technology, and few mathematicians have experience or understand of methods of either. This I believe is what lies behind the grotesque emphasis on significance tests in statistics courses of all kinds; a mathematical apparatus has been erected with the notions of power, uniformly most powerful tests, uniformly most powerful unbiased tests, etc. etc., and this is taught to people, who, if they come away with no other notion, will remember that statistics is about significant differences.

# Sökudólgarnir?

- "Statistics is intimately connected with science and technology, and few mathematicians have experience or understand of methods of either. This I believe is what lies behind the grotesque emphasis on significance tests in statistics courses of all kinds; a mathematical apparatus has been erected with the notions of power, uniformly most powerful tests, uniformly most powerful unbiased tests, etc. etc., and this is taught to people, who, if they come away with no other notion, will remember that statistics is about significant differences.

If they then become scientists they will be searching for uniformity, invariance, and repeatability, not differences; if technologists, they will primarily want to measure things and assess their accuracy, not to test theories.

If they then become scientists they will be searching for uniformity, invariance, and repeatability, not differences; if technologists, they will primarily want to measure things and assess their accuracy, not to test theories.

The apparatus on which their statistics course has been constructed is often worse than irrelevant—it is misleading about what is important in examining data and making inferences." (J.A. Nelder, in discussion of Chatfield 1985)

# Gagnaköfun, pre-test, AIC eða BIC

## Gagnaköfun, pre-test, AIC eða BIC

- Ef F-gildi við stækkun líkans er stærra en tveir á að nota stærra líkanið samkvæmt AIC. (Í línulegum normal líkönum)

## Gagnaköfun, pre-test, AIC eða BIC

- Ef F-gildi við stækkun líkans er stærra en tveir á að nota stærra líkanið samkvæmt AIC. (Í línulegum normal líkönum)
- Ef F-gildi við stækkun líkans er stærra en  $\log(n)$  þá á að nota stærra líkanið samkvæmt BIC.



# Gagnaköfun, pre-test, AIC eða BIC

- Ef F-gildi við stækkun líkans er stærra en tveir á að nota stærra líkanið samkvæmt AIC. (Í línulegum normal líkönum)
- Ef F-gildi við stækkun líkans er stærra en  $\log(n)$  þá á að nota stærra líkanið samkvæmt BIC.
- BIC gefur consistent mat á fjölda mældra skýristærða í líkani, AIC hefur tilhneigingu til að ofmeta fjölda skýristærða í líkani.

# Gagnaköfun, pre-test, AIC eða BIC

- Ef F-gildi við stækkun líkans er stærra en tveir á að nota stærra líkanið samkvæmt AIC. (Í línulegum normal líkönum)
- Ef F-gildi við stækkun líkans er stærra en  $\log(n)$  þá á að nota stærra líkanið samkvæmt BIC.
- BIC gefur consistent mat á fjölda mældra skýristærða í líkani, AIC hefur tilhneiginu til að ofmeta fjölda skýristærða í líkani.
- Þetta samsvarar því einhvers konar pre-test hegðun þar sem marktæknistigið er látið vera fall af fjölda mælinga

- Gagnaköfun er stunduð, og byggir oft á einhvers konar kenningaprófunum. Dæmi um slíkt er GETS (Hendry) og RETINA(White)

- Gagnaköfun er stunduð, og byggir oft á einhvers konar kenningaprófunum. Dæmi um slíkt er GETS (Hendry) og RETINA(White)
- Er hægt að finna sannleikann?

- Gagnaköfun er stunduð, og byggir oft á einhvers konar kenningaprófunum. Dæmi um slíkt er GETS (Hendry) og RETINA(White)
- Er hægt að finna sannleikann?

Pötscher Impossibility theorem segir nei!, nema að leitinni séu settar tilteknar skorður.

- Gagnaköfun er stunduð, og byggir oft á einhvers konar kenningaprófunum. Dæmi um slíkt er GETS (Hendry) og RETINA(White)
- Er hægt að finna sannleikann?  
Pötscher Impossibility theorem segir nei!, nema að leitinni séu settar tilteknar skorður.
- Ljóst er að ef „the rules of statistics“ eru til og hægt að hanna einhvers konara sjálfvirka ályktunarfræði að þá er hún flókin (kíkið á GETS eða RETINA)

# Hlutverk líkans

- Líkanið er hornsteinn allrar tölfræðilegrar vinnu.  
Hvaða ætti að hafa í hug við val á líkani?

# Hlutverk líkans

- Líkanið er hornsteinn allrar tölfræðilegrar vinnu. Hvaða ætti að hafa í hug við val á líkani?
- Poirier (1995) talar um realism versus instrumentalist. Þar er átt við að realisti vilji skilja raunveruleikann en instrumentalisti líti á líkanið sem verkfæri sem geti gefið góðar spár.



# Hlutverk líkans

- Líkanið er hornsteinn allrar tölfræðilegrar vinnu. Hvaða ætti að hafa í hug við val á líkani?
- Poirier (1995) talar um realism versus instrumentalist. Þar er átt við að realisti vilji skilja raunveruleikann en instrumentalisti líti á líkanið sem verkfæri sem geti gefið góðar spár.

- Draper & Smith (1966) eru á svipuðum nótum og hugsa sér þrjár megin nálganir í líkansmíð vísindamanns:
  1. The functional model
  2. The control model
  3. The predictive model

## Staðan í dag

- Tölfræðingar nútímans ekki mjög strangtrúaðir, Bayesismi/tíðintúlkun.
- MCMC aðferðirnar hafa gert bayesískar aðferðir reiknanlegar og margir sem ekki aðhyllast Bayes-tölfræði geta vel hugsað sér að nota MCMC af því það býður upp á að fá reiknanlegar útkomur úr flóknum líkönum.
- Menn telja að grundvallaratriða að reyna að skilja sitt líkan, tölvutæknin er verkfæri. Sumum þykir þó að það verkfæri bjóði upp á hættur

- "It is reassuring to see such a persuasive attempt to bury the 'cookbook,' technique-oriented approach to Statistics." ... "It is fair to say that for many data analysts, computers should be avoided; they provide the means only of getting the wrong answer more quickly and without any thought!" C. Beaumont, discussing (Chatfield 1985)

# Lokaorð

- Margir tölfræðingar ósáttir við hugsunina og praxís í kenningaprófunum

# Lokaorð

- Margir tölfræðingar ósáttir við hugsunina og praxís í kenningaprófunum
- Berger(2002) skrifar grein um hvort Fischer og Neyman-Pearson hefðu einhvern tíma getað orðið sammála ef þeim hefði enst aldur.

# Lokaorð

- Margir tölfræðingar ósáttir við hugsunina og praxís í kenningaprófunum
- Berger(2002) skrifar grein um hvort Fischer og Neyman-Pearson hefðu einhvern tíma getað orðið sammála ef þeim hefði enst aldur.
- Ljóst að kenningaprófun verður stunduð lengi enn.

# Heimildir

Cox, D. & Hinkley, D. (1974). Theoretical Statistics. Chapman and Hall.

Draper, N. & Smith, H. (1966). Applied Regression Analysis. John Wiley & Sons.

Poirier, D. J. (1995). Intermediate statistics and econometrics. MIT press.

Spanos, A. (1999). Probability Theory and Statistical Inference: Econometric Modelling with Observational Data. Cambridge University Press.

Young, G. & Smith, R. (2005). Essentials of Statistical Inference. Cambridge University Press.