

STOCKHOLM UNIVERSITY
Statistiska institutionen
Hans Nyquist

STATISTISK VETENSKAPSTEORI OCH METOD
TENTAMEN
2022-06-03

Tid: 14.00 - 19.00

Tentamen består av sju problem med totalt 40 poäng. För att få full poäng på en uppgift krävs detaljerade och väl motiverade svar. Miniräknare är tillåtet hjälpmedel.

Problem 1 (5 poäng)

Vad är skillnaden mellan rationalism och empiricism?

I samband med en modellbyggnad kan det finnas flera olika modeller att välja mellan. Är modellbyggandet baserat på rationalism eller empiricism när modellbyggaren väljer modell enbart utifrån data, t.ex. med hjälp av något hypotestest eller mått som AIC (Akaikes Information Criterion) eller BIC (Bayesian Information Criterion)? Motivera ditt svar.

Problem 2 (5 poäng)

I en del läroböcker står det att parametrarna betraktas som stokastiska i Bayesiansk inferens. Förklara vad som avses och ge en tolkning.

Problem 3 (5 poäng)

Förklara vad som menas med att observationer är teoriberoende (som diskuteras i bl.a. Hansson (2007)).

Problem 4 (5 poäng)

Klassisk inferensteori uppfyller inte den starka likelihoodprincipen. Vad är det i den klassiska inferensteorin som gör att den starka inferensprincipen inte uppfylls?

Problem 5 (5 poäng)

Vad menas med en kausal relation? Vilka möjligheter och begränsningar finns det att dra slutsatser om kausalitet i ett experiment respektive en observationsstudie?

Problem 6 (5 poäng)

Ange några argument för att stödja ett antagande om normalfördelning i en modell. Ange också hur ett argument för ett normalfördelningsantagande också kan vara ett argument mot ett normalfördelningsantagande.

Problem 7 (10 poäng)

Freedman skriver i sin artikel i kursmaterialet att "In general, social scientists who run regressions are not doing data analysis. Instead, they are mechanically applying regression models in situations where the assumptions do not hold. The computer outputs - the parameter estimates, the standard errors, the t-tests - are usually devoid of scientific meaning. Rather than facing up to this issue, however, the investigators just label the outputs as "descriptive statistics". This is swindle. If the assumptions of the regression model do not hold, the computer outputs do not describe anything: they are mere numerical artifacts." Ge din egen tolkning till detta uttalande och relatera det till begreppet vetenskaplig metod. Kan Freedmans argument generaliseras till annat än regressionsanalys, t.ex. till tidsserieanalys och maskininlärning?

Referenser: Friedman, D.A. Statistics and the Scientific Method, i Cohort Analysis in Social Research eds. Mason, W.M. och Fienberg, S.E.

Hansson, S.O. (2007) The Art of Doing Science, KTH, Stockholm

LYCKA TILL!