

Vetenskaplig metodik

Vilka metoder används?

Vi kan dela in metoder i flera grupper:

- Deduktiva metoder
- Metoder för hantering av experiment
- Metoder för publicering och liknande.

Från föreläsning 3

- Föreläsningen handlade framför allt om den Hypotetiskt-Deduktiva metoden.
- HD-metoden är ett redskap för att relatera hypoteser och teorier till observationer och experiment.
- Den är på sätt och vis en avancerad form av induktion.
- Vissa anser att metoden är kärnan i vetenskaplig metodik.

Generella frågor

- Vad är det du ska göra? Vilket är ditt projekt?
- Varför gör du det? Är det intressant? Är det viktigt?
- Hur ska du göra det? Vilka metoder?
- När ska du göra det? Hur lång tid tar det?

Forskningsämne

- Bör vara klart formulerat
- Bör vara signifikant. Ska t.ex. inte vara upprepning av något tidigare gjort
- Bör ha tydliga gränser
- Bör vara sådant att nödvändiga data går att få
- Bör vara sådant att tydliga slutsatser går att dras

Formen på projektet

- Kan vara i form av en fråga. Är det t.ex. så att funktionell programmering löser matematisk programmering bättre än andra paradigmen?
- Kan vara i form av test av en hypotes. Hypotesen kan t.ex. vara att funktionell programmering löser matematisk programmering bättre än andra paradigmen

Värdet i att ha rätt

En känd forskare sade en gång att det viktiga i forskning är att ha rätt. Man måste kunna välja antaganden som är sanna. Man måste ha intuition!

Vetenskaplig metodik i projektarbete

Vi kan karakterisera arbetet genom att tala om fyra faser:

- Analys
- Hypotesbildning
- Syntes
- Validering

Analysfasen

Målet här är att få en inledande förståelse för vad problemet/projektet går ut på. Kan bestå av momenten:

- Beskriv problemet
- Bestäm mått på framgång
- Gör studier av liknande arbete
- Sätt mål

Hypotesfasen

Handlar om att föreslå en lösning. Här krävs förstås kreativitet. Andra moment kan vara:

- Beskriv hypotesen/lösningen
- Ange kriterier för att avgöra om hypotesen/lösningen är tillräckligt bra

Syntesfasen

Handlar om att realisera en lösning eller att testa en hypotes mot verkligheten. Fasen kan ha momenten:

- Fastställ konsekvenser av hypotesen
- Implementera lösningen
- Designa experiment för att testa hypotesen/lösningen
- Genomför experimenten

Valideringsfasen

Handlar om att utvärdera värdet av hypotesen/lösningen efter att experimenten är genomförda. Fasens moment kan vara:

- Mät lösningens framgång eller hur bra hypotesen stämmer med experimenten
- Dra slutsatser om hypotesens/lösningens värde
- Gör dokumentation, t.ex. genom att skriva en rapport eller artikel
- Utsätt hypotesen/lösningen för kritik av oberoende bedömare

Vetenskapsmäns förhållningsätt

Vi kan också karakterisera vetenskaplig metod efter dess utövares attityd. Enligt Merton är följande typiskt för vetenskapsmän. De fem principerna går under akronymen CUDOS:

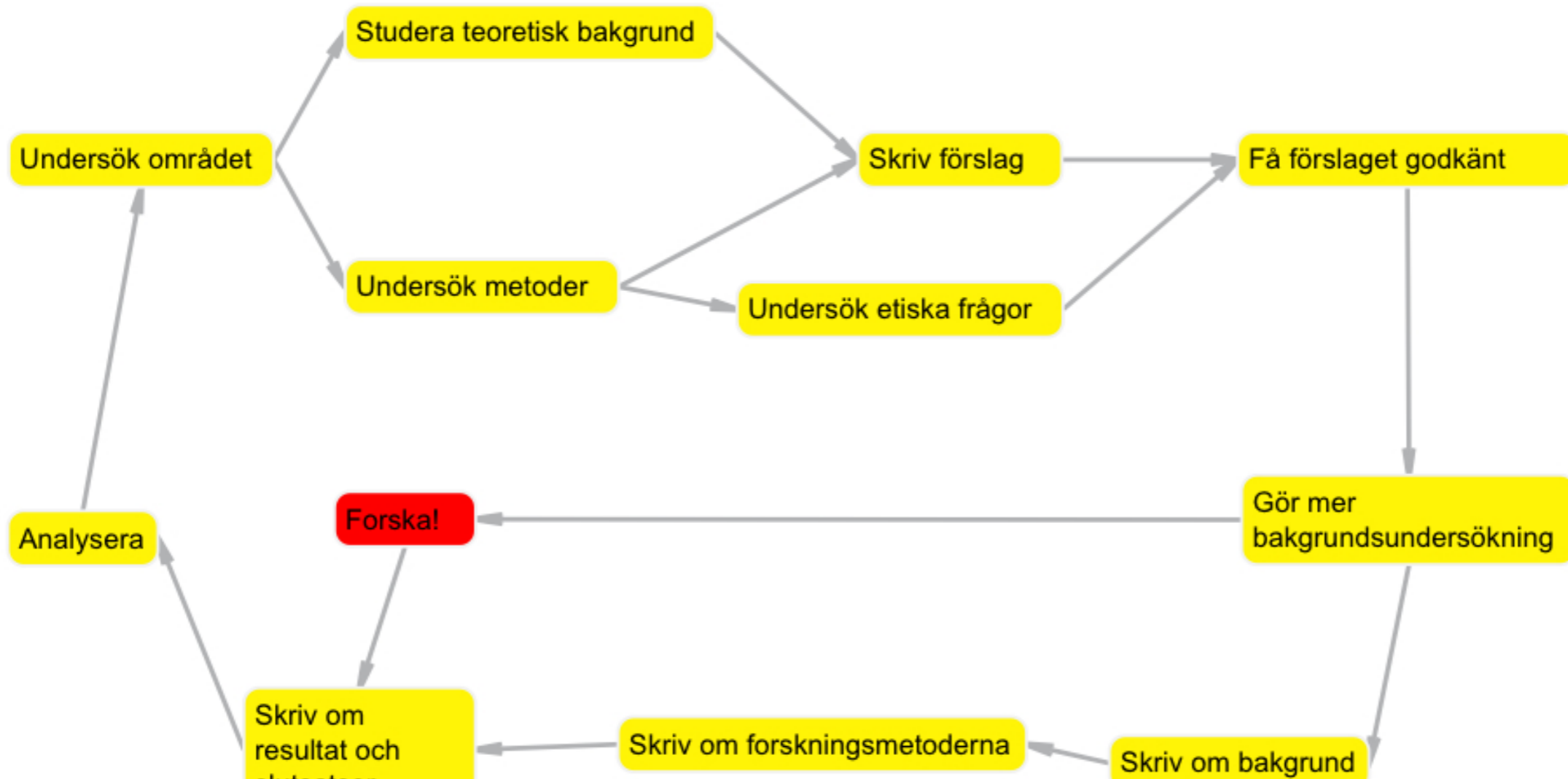
- Communalism - kunskap är allas
- Universalism - alla får bidra
- Disinterestedness - forskning styrs inte av särintressen
- Originality - forskning ska vara ny
- Skepticism - resultat måste ifrågasättas

Data

En del i forskningsprocessen är att hantera data av olika slag. Det visar sig att man kan beskriva data på flera olika sätt.

- Beskrivning efter abstraktionsnivå
- Uppdelning i primära och sekundära data
- Kvantitativa och kvalitativa data
- Mätning efter olika typer av skala

Forskningsprocessen



Abstraktionsnivåer

- Teori - De viktigaste sambanden uttryckta i abstrakta termer
- Koncept - De grundläggande begreppen.
Teorin uttrycker ofta samband mellan koncept
- Indikatorer - Något mätbart som är relaterat till koncepten eller påvisar existens av dem
- Variabler - Den mätbara komponenten av indikatorerna
- Värderna - Resultaten av mätningar av variablerna

Primära och sekundära data

- Primära data - Är direkta mätningar eller observationer av något. Kan också vara rapporter av personer som själva varit med om något
- Sekundära data - Är ofta sammanställning av primära data som har behandlats i någon form. De kan t.ex. förekomma i rapporter, artiklar eller böcker

Kvantitativa och kvalitativa data

- Kvantitativa data - Sådana data som ges i form av tal
- Kvalitativa data - Data som inte enkelt kan ges i form av tal. Det kan vara åsikter, berättelser eller beskrivningar av situationer

Vetenskapliga data och observationer

- De tidigare observationerna har varit sådana att deras värden är sant/falskt.
- I andra sammanhang tilldelar man observationerna ett numeriskt värde: observationerna E_1, E_2, \dots får värdena $f(E_1), f(E_2), \dots$
- Funktionen f definierar vilken typ av *skala* man använder.
- Vilken typ av skala man använder definierar vilken typ av information man objektivt kan läsa ut ur observationerna.

Typer av skala

- Ordinalskala: Relationer av typen $f(E1) < f(E2)$ är objektiva. En ordinalskala är bara en rangordning av observationerna.
- Kvotskala: Värdet $f(E2)/f(E1)$ är objektivt. En kvotskala förutsätts mäta en verklig storhet.
- Intervallskala: Värdet $(f(E2)-f(E0))/(f(E1)-f(E0))$ är objektivt. Typiskt exempel är temperaturskalor (utom Kelvins).

Samla och analysera sekundära data

- Typiska exempel kan vara skrivet material som romaner, rapporter, biografier, tidningar e.t.c.
- Det kan vara tv-program, filmer, intervjuer e.t.c.
- Det kan vara statistiska undersökningar gjorda i något annat syfte

Problem i analysen

- Att hitta data
- Att autentisera källorna
- Att bedöma trovärdigheten
- Att avgöra hur representativa data är
- Att välja metod för att tolka data

Tre speciella analysmetoder

- Content analysis - Vi räknar helt enkelt förekomster av något t.ex. ord eller typer av bilder i dokument och använder förekomsterna som indikatorer
- Data mining - Vi använder program för att hitta mönster i data
- Metaanalys - Vi gör en analys av flera andra analyser samtidigt och försöker se mönster i dem

Samla primära data

Mängden metoder är stor. Vi kan dock nämna tre huvudområden

- Statistiska undersökningar. Sampling
- Intervjumetoder. Detta kommer att tas upp i en senare föreläsning
- Experiment. Några generella metoder för experiment är att man jämför den riktiga experimentgruppen också har en kontrollgrupp. Man bör också om möjligt använda så kallade dubbelblinda test

Kvantitativ dataanalys

Vi nämner lite om statistiska metoder för analys av data. Metoderna är parametriska d.v.s de bygger på att vi kan göra antaganden om fördelningarna av det vi mäter, t.ex. att vi har normalfördelning. Två metoder som är bra att känna till är:

- Hypotesprövning
- Maximum Likelihood-metoden

Hypotesprövning

Antag att vi har en hypotes H som vi vill pröva. Låt oss jämföra den med en nollhypotes H_0 . Vi designar ett test som ger ett testvärde t . Vi definierar en mängd M och säger att vi förkastar H_0 om t tillhör M . (Det innebär att vi accepterar H .) I så fall Vi säger dessutom att testet är signifikant på α -nivå om sannolikheten för att t skall tillhöra M *givet att H_0 är sann* är mindre än eller lika med α .

Maximum likelihood-metoden

Metoden kan illustreras med två specialfall:

Låt oss anta att vi vill ha reda på värdet på en parameter f . Vi gör ett experiment som ger ett utfall t . Vi undersöker vilket värde f_0 som maximerar $P(t \mid f = f_0)$. Vi antar att f har detta värde f_0 .

Låt oss anta att vi har gjort en observation E . Vi har olika hypoteser H_1, H_2, \dots, H_n som skulle kunna vara möjliga *orsaker* till E . Vi väljer den hypotes H_i som maximerar $P(E \mid H_i)$. (Obs detta är inte samma sak som $P(H_i \mid E)$.)

Kvalitativ dataanalys

Används ibland inom humanistiska och samhällsvetenskapliga områden. Kommer att beskrivas i en senare föreläsning

Metoder inom matematik och logik

- Metoder som man använder är oftast starkt kopplade till ett givet område.
- Vissa generella metoder kan kanske ges.
- Om man ställs inför något man vill bevisa kan man fråga sig: "Påminner det här om något jag redan vet hur man bevisar?"
- Mer formellt kan det uttryckas som *störningsräkning*. Vi tänker oss ett problem vi redan vet hur man hanterar och försöker i någon mening *approximera* avvikelserna.
- Om vi vill bevisa ett påstående $A \Rightarrow B$ kan vi beskriva några knep som är naturliga.

Att försvaga

- En försvagning av $A \Rightarrow B$ är att ersätta påståendet med något som är svagare och därmed kanske lättare att bevisa.
- Det svagare beviset kan sedan kanske förstärkas till det vi vill visa.
- Två möjligheter finns. Antag att $A' \Rightarrow A$. Då är $A' \Rightarrow B$ en försvagning.
- Ett annat alternativ är om $B \Rightarrow B'$. Då är $A \Rightarrow B'$ en försvagning.

Att förstärka

- Ibland kan det vara en god idé att försöka förstärka påståendet som skall bevisas. Paradoxalt nog kan det ibland vara lättare att visa ett starkare påstående.
- Två naturliga former finns också här. Om A
- $A \Rightarrow A'$ så är $A' \Rightarrow B$ en förstärkning.
- Om $B' \Rightarrow B$ så är också $A \Rightarrow B'$ en förstärkning.

Ett exempel: Sexfärgssatsen

- Fyrfärgssatsen säger att alla planära grafer kan fyrfärgas.
- En försvagning av satsen skulle kunna vara: "Det finns ett tal N så att alla planära grafer kan N -färgas"
- Om man försöker bevisa detta påstående kommer man snart fram till att man kan bevisa en förstärkning av påståendet.
- Det är att alla planära grafer kan sexfärgas.
- Denna sats kan ganska enkelt förstärkas till satsen att alla planära grafer kan femfärgas.

Fyrfärgssatsen

- Beviset för fyrfärgssatsen bygger på precis samma teknik som bevisen för de föregående satserna.
- Metoden bygger på induktion över grafens storlek. Man visar att grafen innehåller en mängd av *oundvikliga* delgrafer. Grafen måste innehålla minst en av dessa delgrafer.
- Man försöker sedan visa att de är vad som kallas *reducerbara*. Det innebär att om delgrafan kan fyrfärgas kan denna färgning modifieras och utvidgas till hela grafen.
- Enligt induktionsantagandet kan delgrafan fyrfärgas och därmed är saken klar!

Datorbevis

- Men enkelt är det inte. Problemet är att hitta en mängd av (tillsammans) oundvikliga delgrafer som alla är reducerbara. I sex- och femfärgssatsen var det enkelt.
- Appel och Haken lyckades hitta en mängd med 1936 delgrafer som tillsammans var oundvikliga och de förmodade att de alla var reducerbara.
- För att bevisa att de var reducerbara använde de ett dataprogram som för var och en av delgraferna lyckades skapa en sökt omfärgning av grafen.
- Beviset blev mycket omdiskuterat och väckte frågor om vad vi egentligen kan acceptera som bevis.