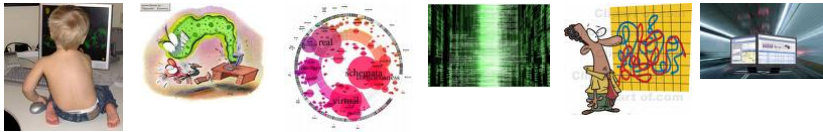


# Kvantitativa metoder: Datainsamling, presentation och analys av data

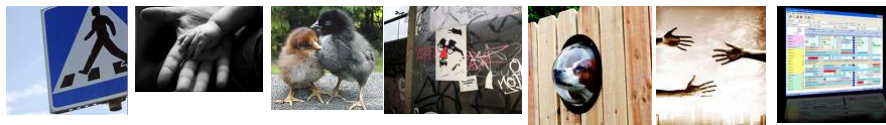
Magnus Hansson  
Handelshögskolan

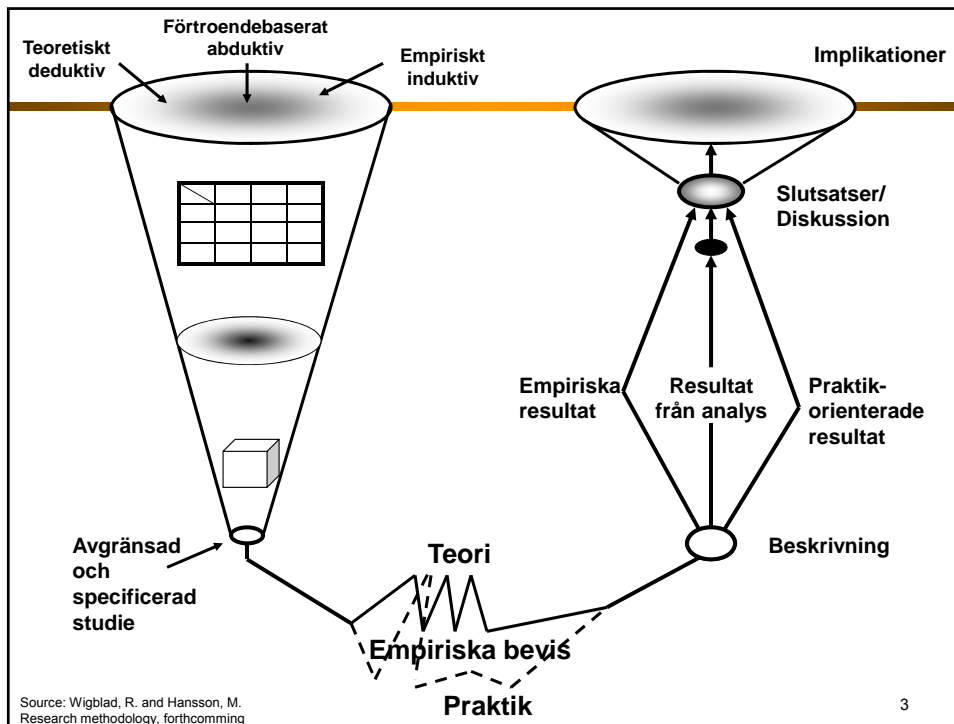
Örebro Universitet



## Föreläsningen i korta drag

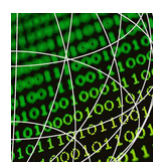
- Surveyundersökningen som datainsamlingsmetod
- Beskrivande statistik
- Presentation av data
- Analys av och resultat utifrån data





## Presentation av kvantitativa data: Beskrivande statistik

- I likhet med andra typer av fallstudier blir det många gånger viktigt att man som forskare tydligt presenterar sin insamlade (bearbetade data).
- Den beskrivande statistiken kan ha olika syften (eg.):
  - Illustrativa
  - Informativa
- Jmf. med det som Sutton & Staw (1995) konstaterar:
  - Grafer, Illustrationer, Listor med variabler, referenser och hypoteser är INTE teori.



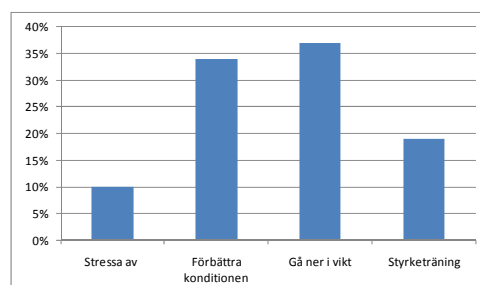
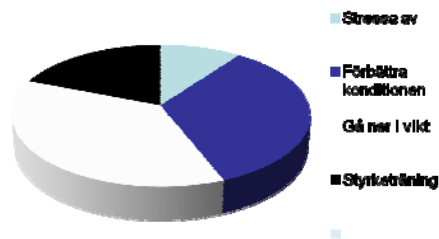
## Univariat analys

- Univariat analys innebär analys/mätning av en variabel i taget
- Exempel på presentationer av univariata analyser:
  - Frekvenstabeller
  - Diagram
    - Stapeldiagram
    - Linjediagram
    - Pie charts
    - Histogram
    - Plotdiagram
    - Ytdiagram

5

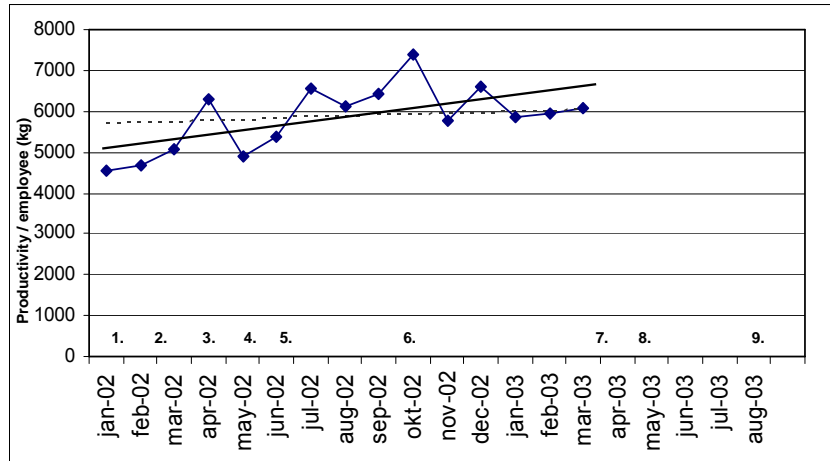
## Några grundläggande diagram: Frekvenstabell, cirkel- och stapeldiagram

Motiv	n	%
Stressa av	9	10%
Förbättra konditionen	31	35%
Gå ner i vikt	33	37%
Styrketräning	17	19%
TOTAL	90	100%



6

## Linjediagram



7

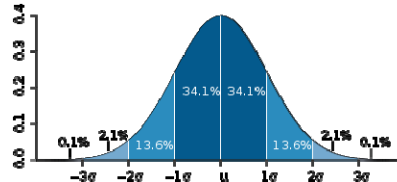
## Centraltendenser

- Aritmetiskt medelvärde ( $\bar{x}$ )
  - $\Sigma$  (alla värden)/antalet värden
  - Bör användas enbart i samband med kvot/intervallvariabler.
- Medianvärde
  - Anger mittpunkten på en fördelning av värden.
  - Känsligt för extremvärden
  - Kan användas i samband med kvot/intervallvariabler och ordinalvariabler
- Modal/typvärde
  - Det värde som förekommer oftast i en fördelning.
  - Kan användas vid alla slags variabler.
    - E.g 1,2,3,2,3,4,5,6,4,3,2,1,5,6,7,3,4,2,2,1,1
      - Modal/Typvärde = 2 (förekommer 5 gånger)

8

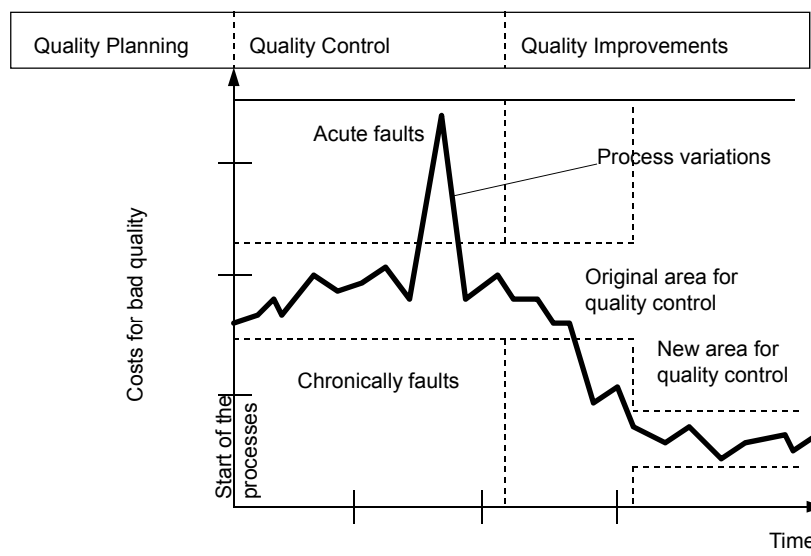
## Spridningsmått

- Graden av variation i ett stickprov kan vara lika intressant som typvärden eller genomsnittliga värden.
- Ange vidd eller omfång
  - Spridningen mellan lägsta och högsta värdet.
- Standardavvikelse  $\sigma = \sqrt{Var(X)}$ 
  - Genomsnittlig variation/spridning kring ett medelvärde
  - +/- 1 std = 68,2%
  - +/- 2 std = 95,4%
  - +/- 3 std = 99,8%
  - → Jmf mått av processkapabilitet



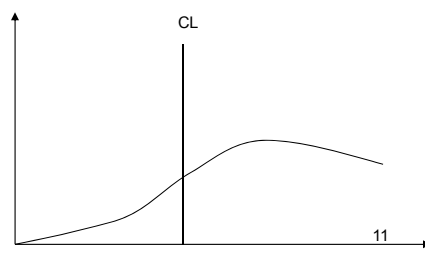
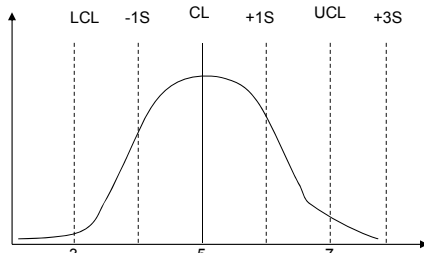
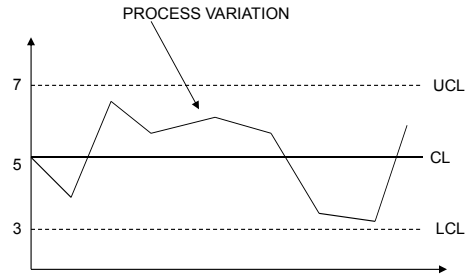
9

## Ett exempel från TQM

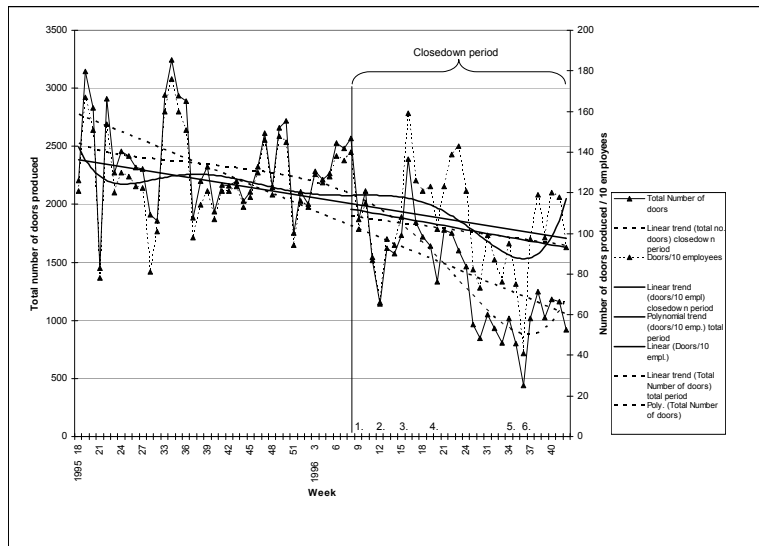


# Processvariationer

	X1	X2	X3	X4
1-2				
3-4				
5-6				
7-8				
9-10				



# Exempel på univariat analys



12

## Exempel på univariat analys (forts)

- Kontinuerliga produktivetsdata (utfall)

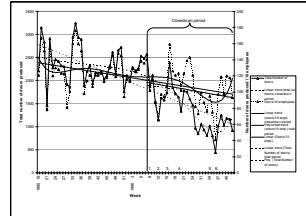
- Trendanalyser

- Linjär

- Räkna linjens ekvation ( $Y=mx+b$ )

- Polynomial

- Olika grader av känslighet (i exemplet 6 grader)
- För att fånga variationer i datamaterialet; Känslighetsanalys

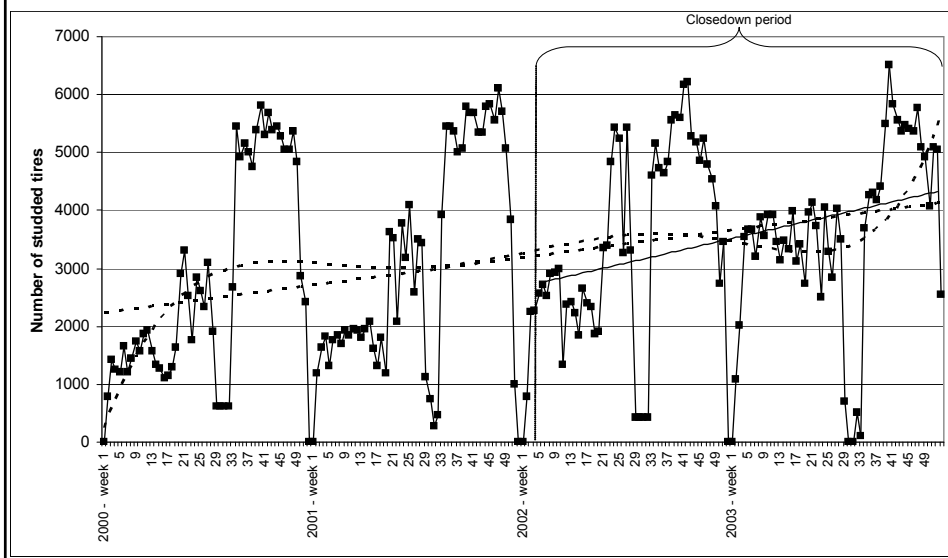


- Andra exempel på trendanalytiska instrument:

- Exponentiell
- Glidande medelvärde
- Logaritmisk

13

## Polynomial trendanalys



**Table 3 Comparative analysis – the studding department**

	2000	2001	△ %	2002	△ %	2003	△ %	2000-2003
<b>Week 1-29</b>								
Average	1720,92	2147,14	24,77%	2703,22	25,90%	3191,97	18,08%	85,48%
Standard deviation	721,6	907,7		1290,93		956,92		
<b>Week 33-52</b>								
Average	4834,84	5106,32	5,62%	4911,49	-3,82%	5005,7	1,92%	3,53%
Standard deviation	1011,06	1154,7		848,24		756,24		
<b>Total</b>								
Average	2859,63	3200,57	11,92%	3451,45	7,84%	3752,61	8,73%	31,23%
Standard deviation	1777,78	1657,41		1729,49		1388,81		

## Ett fortsatt exempel på en univariat analys

**Table 2 Productivity analysis**

Measures	Productivity before closure announcement Δ (%)	Productivity during the closedown period	Mean productivity and standard deviation before closure announcement	Mean productivity and standard deviation for closedown period	Productivity and standard deviation ANP	Productivity and standard deviation CDP	Test for parallelism	Significant Closedown effect (p=0,05)
<b>Case</b>								
<b>Fundia Steel Wire Rod</b>	47,22% (*)	48,64% (**)	45,69 [δ = 5,84]	46,53 [δ = 7,34]	41,10 [δ = 4,36]	50,88 [δ = 3,84]	26,532	Yes
Type of data	Calculation based on linear trends of (ton/hour)/employee. The test for parallelism is based on the average of the 44 tonnes/hour assortment. Advance notice period = 3 months, Countdown period = 4 months.							
<b>Gusab Stainless</b>	37,53% (*)	60,13% (**)	5110,50 [δ = 927,1027]	5808,29 [δ = 1201,151]	6632,43 [δ = 536,23]	5955,00 [δ = 799,84]	3110,267	Yes
Type of data	Calculation based on linear trends of (kilos/month)/employee Advance notice period = 2 months, Countdown period = 16 months							
<b>Gislaved Studding</b>	78,26% (*)	64,81% (**)	3204,30 [δ = 1779,03]	3760,84 [δ = 1510,94]	2491,02 [δ = 728,90]	3724,00 [δ = 1603,49]	6510,466	Yes
Type of data	Calculation based on linear trends of total number of (studded tires/month)/employee. Advance notice period = 2 months, Countdown period = 16 months.							

(\*) Refers to the productivity development *until* the announcement of the closedown decision

(\*\*) Refers to the productivity development *from* the announcement of the closedown decision to the final closure



## Exempel på klassificeringsschema för att möjliggöra jämförelse mellan olika fall

Time-frame Type of setting	Short term (0-6 months)		Middle-range (7-12 months)		Long-term (> 12 months)	
	Advanced Notice Period	Countdown period	Advanced Notice Period	Countdown period	Advanced Notice Period	Countdown period
SR driven closures						
Strategic Non-SR and Tactic SR driven closure			Fundia Wire Rod 3 months	Fundia Wire Rod 4 months		
Non-SR driven closure	Gislaved Tire Manuf. 2 months	Gislaved Tire Manuf. 4 months	Cabinet Factory 7,5 months	Cabinet Factory 1 month	Gislaved Studding 2 months	Gislaved Studding 16 months

## Lite mer om univariat analys

Table 3 Case specific outcomes

Outcomes Case	Type of production	Type of new rationalisation	Type of new improvement	Classification and lengths of ANP(*) and CDP(**)	Comparative value of the Closedown effect
Fundia Steel Wire Rod	Process	Rationalisation of work methods	Radical	Middle range: ANP = 3 months, CDP = 4 months	2,864
Gusab Stainless	Process	Increased discipline	Incremental	Long term: ANP = 2 months, CDP = 16 months	3,612
Gislaved Studding	Mass	Enhanced efforts	None	Long term: ANP = 2 months, CDP = 16 months	1,515

(\*) Advance notice period (ANP) (\*\*) Countdown period (CDP)

## Bivariat analys

- Analys av två variabler i taget i syfte att visa hur de två är relaterade till varandra.
- Söka bevis för att variation i den ena variabeln sammanfaller med den andra.
- Flera olika tekniker för att studera relationer
  - Valet av teknik är beroende av de två variabelernas art/typ.
  - Viktigt att notera: man mäter samband (relationer mellan variabler) INTE ett orsak-verkan förhållande mellan dessa variabler.

19

## Bivariata analysmetoder

	Nominal-variabler	Ordinal variabler	Kvot- och intervall-variabler	Dikotoma variabler
Nominal variabler	Contingency tabell + Chi2test + Cramérs V	Contingency tabell + Chi2test + Cramérs V	Contingency tabell + Chi2test + Cramérs V (Om kvot-/intervall variabeln är beroende variabel → jmf medelvärden + eta)	Contingency tabell + Chi2test + Cramérs V
Ordinal variabler	Contingency tabell + Chi2test + Cramérs V	Spearman's rho	Spearman's rho	Spearman's rho
Kvot- och intervall variabler	Contingency tabell + Chi2test + Cramérs V (Om kvot-/intervall variabeln är beroende variabel → jmf medelvärden + eta)	Spearman's rho	Pearson's <i>r</i>	Spearman's rho
Dikotoma variabler	Contingency tabell + Chi2test + Cramérs V	Spearman's rho	Spearman's rho	Phi

20

## Exempel på en contingency tabell

Motiv	Kön			
	Män		Kvinnor	
	Antal	Procent	Antal	Procent
Stressa av	3	7	6	13
Kondition	15	36	16	33
Viktminskning	8	19	25	52
Styrketräning	16	38	1	2
TOTAL	42		48	

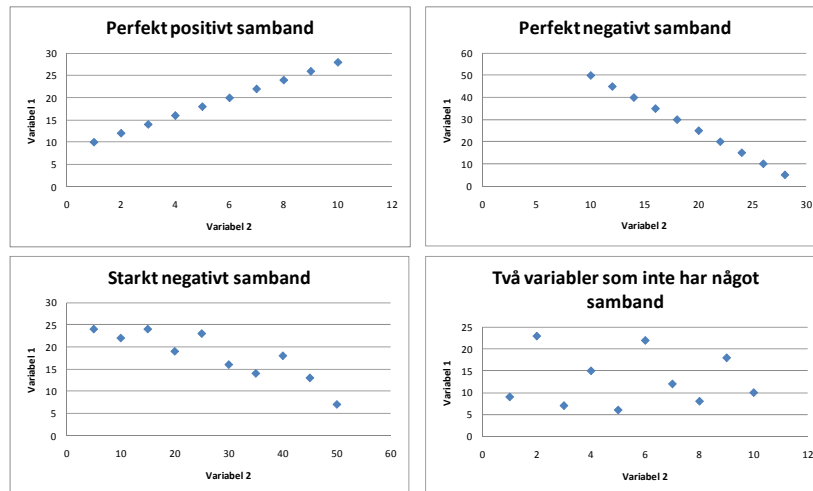
21

## Pearsons $r$

- Metod för att studera relationer mellan intervall- och kvotvariabler:
  - Koefficienten kommer ligga mellan 0 (inget samband) och 1 (perfekt samband) → Visar på styrkan i sambandet mellan variabler.
  - Ju närmare koefficienten ligger 1, desto **starkare** samband, vis-á-vis, ju närmare 0, desto **svagare** samband.
  - Koefficienten är antingen positiv eller negativ, vilket visar riktningen på sambandet.

22

## Exempel på spridningsdiagram och olika typer av samband



23

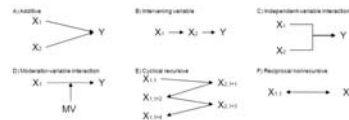
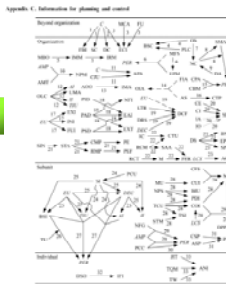
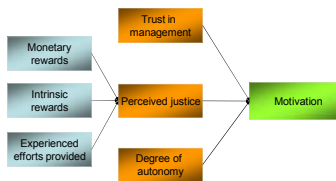
## Multivariat analys

- Innebär att man analyserar tre eller fler variabler samtidigt.
- Används vid tre typer av tillfällen då man har:
  - 1. [Falska samband]
  - 2. Mellanliggande variabler
  - 3. Tre (eller fler) variabler av betydelse
    - Med förklaringsvärde till den beroende variabeln.

24

# Mellanliggande variabler

- Varför finns det ett samband mellan olika variabler?
- Många gånger komplexa förhållande mellan variabler → Komplexa förklaringsmodeller.
- Jmf. tidigare exempel från föreläsning om teori och referensram



25

# Exempel på multivariat analys, med tre eller fler variabler (Häsenen, Hellgren & Hansson, 2011)

**Table 4.** Results of hierarchical regression analysis with standardized regression coefficients

Independent variables T1	Dependent variables T2				
	Job performance	Goal commitment	OCB	Job satisfaction	Job-induced tension
<i>Background variables</i>					
1 Sex (F = 1/M = 2)	.10	-.20*	.00	-.03	-.10
2 Age	-.16	-.05	.03	-.02	-.04
<i>Goal setting variables</i>					
3 Supervisor support/participation	-.09	.09	.16	.34**	-.06
4 Goal stress	-.12	-.15	-.06	-.03	.33**
5 Goal efficacy	.11	-.02	.33**	.11	.12
6 Goal rationale	.39*	.15	-.06	.20	-.14
7 Use of goal setting in performance appraisal	-.14	-.04	-.13	-.08	.08
8 Tangible rewards	.13	.28**	-.04	.11	-.22*
9 Goal conflict	.16	-.04	-.12	-.01	.01
10 Organizational facilitation of goal achievement	-.04	.02	.02	-.03	-.05
11 Dysfunctional effects of goals	-.17	-.13	.13	-.01	.26*
12 Goal clarity	-.19	-.13	.10	-.04	.27*
R <sup>2</sup>	.14*	.27***	.17**	.29***	.24***
R <sup>2</sup> adjusted	.09*	.20***	.10**	.23***	.17***

n = 136-151, after listwise deletion. \* p < .05, \*\* p < .01, \*\*\* p < .001.

26

## Chi2test

- Tillämpas på contingencytabeller och tekniken gör att man kan slå fast hur säker man kan vara på att det finns ett samband mellan två variabler i populationen.
- Beräkning av det förväntade värdet eller frekvensen för de olika cellerna i contingencytabellen.
- Uträkning av det förväntade och faktiska värdet och summering av skillnaderna.
- Chi2värdet är statistiskt signifikant inte enbart på värdets storlek, utan även på antalet kategorier som ingår i de två variablerna som analyseras.



Degrees of Freedom	0.01	0.05	0.10	0.20	0.50	1.00	2.00	5.00	10.00	20.00	50.00	100.00
1	6.635	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
2	5.991	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
3	6.251	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
4	6.479	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
5	6.626	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
6	6.757	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
7	6.871	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
8	6.965	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
9	7.043	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
10	7.107	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
11	7.167	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
12	7.224	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
13	7.277	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
14	7.327	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
15	7.375	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
16	7.421	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
17	7.464	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
18	7.506	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
19	7.547	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
20	7.587	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
21	7.626	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
22	7.664	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
23	7.701	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
24	7.737	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
25	7.772	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
26	7.807	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
27	7.841	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
28	7.874	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
29	7.907	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
30	7.939	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
31	7.971	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
32	7.999	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
33	8.027	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
34	8.054	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
35	8.081	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
36	8.107	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
37	8.133	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
38	8.158	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
39	8.183	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
40	8.207	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
41	8.231	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
42	8.255	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
43	8.278	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
44	8.301	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
45	8.324	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
46	8.347	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
47	8.369	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
48	8.391	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
49	8.413	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501
50	8.435	3.841	2.706	2.071	1.642	1.385	1.108	0.838	0.716	0.616	0.554	0.501

27

## Statistisk signifikans

- Avser den risknivå man är villig att acceptera då man drar slutsatser.
- $p < 0,05$  = att man i 5 av 100 fall drar en felaktig slutsats
- Handlar om att bedöma säkerheten i att resultatet från ett slumpmässigt urval kan generaliseras till den population från vilket stickprovet dragits.

28

## Variansanalyser

- Analys av den totala variationen i beroende variabeln
  - Två slags variationer studeras
    - Inom (sub)grupper
      - Felvariansen
    - Mellan (sub)grupper
      - Den förklarade variansen
- ANCOVA
  - Analys av covarians.
- MANCOVA
  - Multipel analys av covarians

29

## Analys, resultat och slutsatser

- Vid en forskningsstudies början formuleras forskningsfråga/or, syften och avgränsningar.
- I det stadiet är inte forskaren säker på utkomsten av undersökningen.
- Analys och slutsatser formas av den information som samlas in.
- Analysen kan göras löpande, samtidigt som man samlar in data, för att man skall kunna (iterativt) skapa, återskapa eller omskapa det teoretiska ramverket och/eller avgränsningar.
- Analysen kan också göras efter det att all data har samlats in, skrivits ut och sammanställts.

30

## Analys, resultat och hypoteser

- Det teoretiska ramverkets roll.
  - Om det teoretiska ramverket inte (på ett tillfredsställande sätt) har utvecklats (deduktiv ansats) och fungerar som en bas från vilken data samlats in kommer det vara väldigt svårt att genomföra en analys.
- Behovet av att ha ett omfattande dataset, som är relaterat till den specifika och definierade forskningsuppgiften.
  - Genom att ha ett omfattande dataset kan det bli möjligt för forskaren att uppnå empiriska mått.
  - Empirisk mått innebär att flera (många) respondenter har samma/liknande uppfattningar om ett specifikt objekt, händelse, etc.
- Vid enkätundersökningar söker man (oftast) svar på hypoteserna.
  - Hypoteserna testas (verifieras eller falsifieras) med hjälp av det insamlade datamaterialet och (statistiska) analysmetoder.
  - I slutsatsdiskussionen skall det tydligt framgå huruvida hypoteserna är verifierade eller falsifierade.

31

## Ett par praktiska rekommendationer

- Tydlig koppling mellan uppsatsens alla delar.
- Tydlighet i motiv till olika val och steg genom uppsatsen
- I analys och slutsatsdiskussionerna, se till att ta diskussionen längre, problematisera och med stöd i den teoretiska referensramen underbygga argumentationen.
- Tydliggöra resultaten och utkomsten av undersökningen.
  - Explicitgöra resultaten: En slutsats vi kan dra... En annan slutsats vi har dragit... etc.
    - Teoretiska resultat
    - Empiriska resultat
    - Praktiska rekommendationer (Policy implications)
- Tydliggöra kopplingen mellan resultaten och problemområdet och hur resultaten svarar på problemställningen.

32

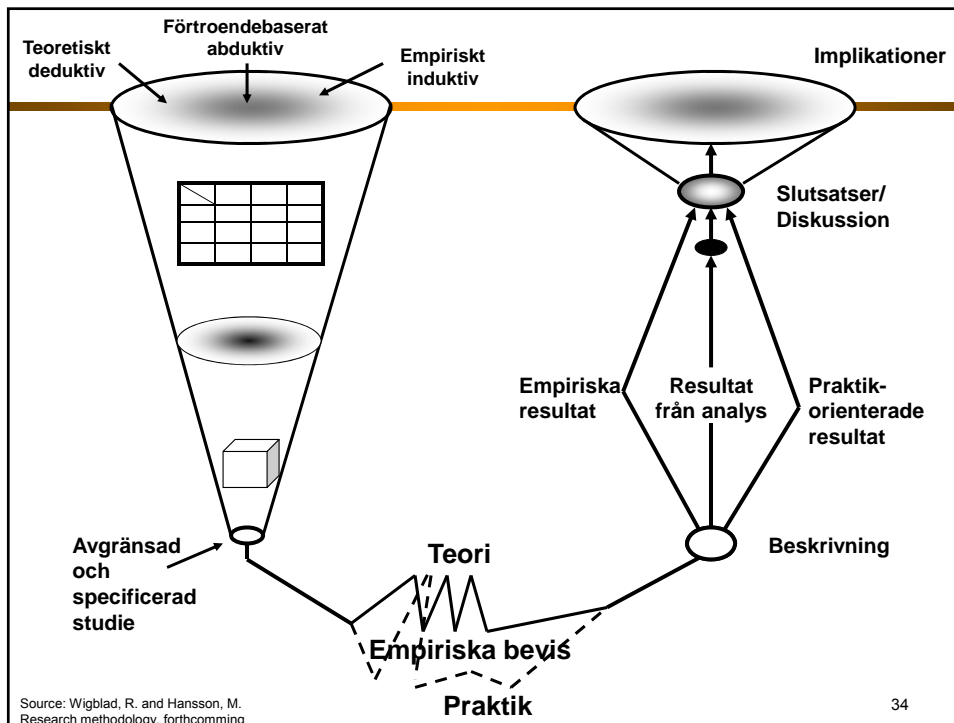


# Sagostund...

Nasses taxi  
av Sven Nordqvist



33



34