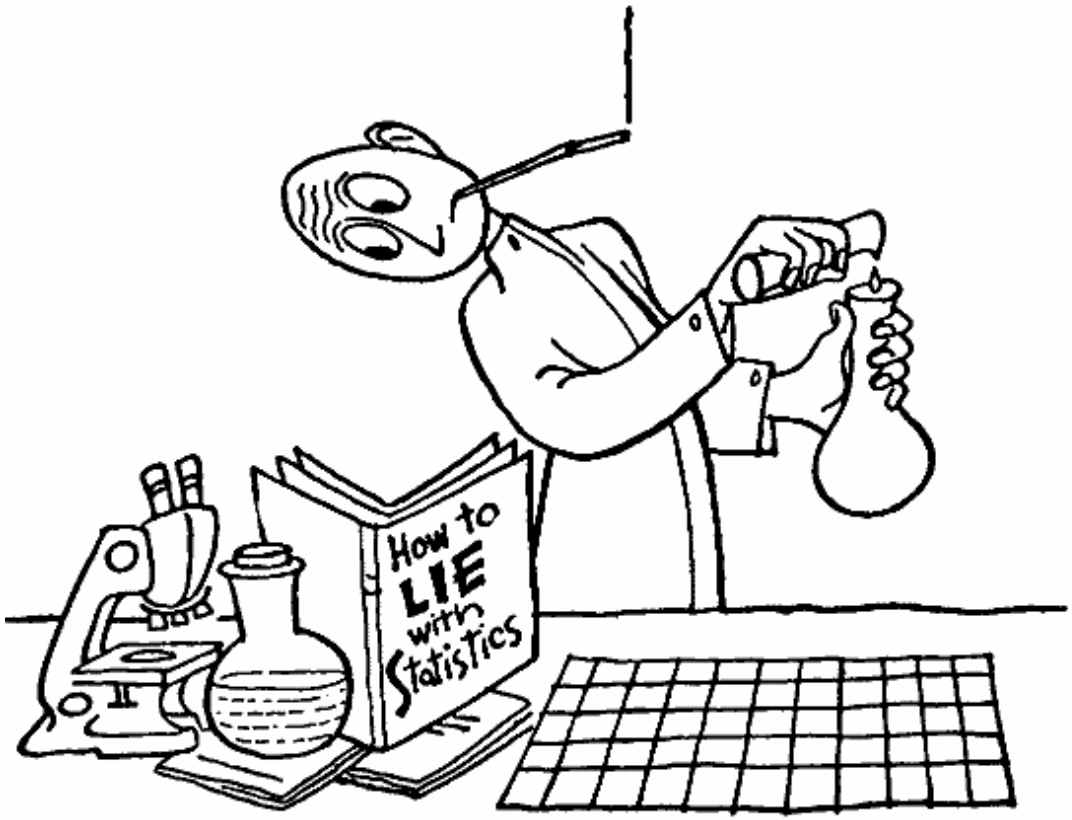


ST-fredag epidemiologi och biostatistik 2017

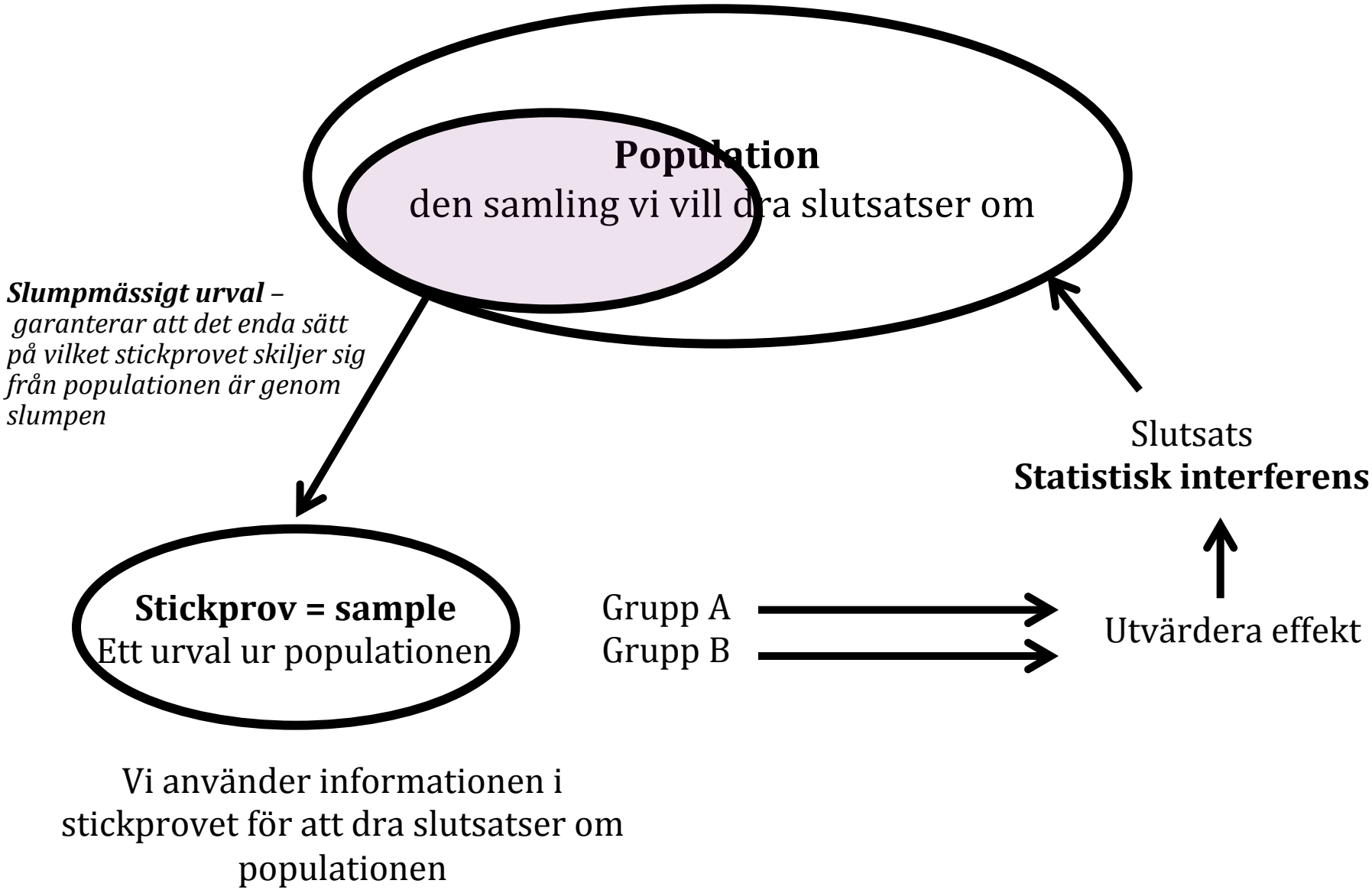
Emma Larsson. ST-läkare, PhD. PMI, KS Solna
Gabriella Jäderling. Överläkare, PhD. PMI KS Solna
Mikael Eriksson. Specialistläkare, doktorand. PMI KS Solna.
Max Bell. Överläkare, Docent. PMI KS Solna

KAROLINSKA
Universitetssjukhuset

BIOSTATISTIK

Att genomföra en studie…

- Frågeställning = hypotes
- Studiedesign
- Styrkeberäkning (power-beräkning)
- Variabler: utfall resp. prediktorer
- Datainsamling
- Deskriptiv statistik
- Analytisk statistik
- Interferens


Population
den samling vi vill dra slutsatser om
Slumpmässigt urval –
garanterar att det enda sätt på vilket stickprovet skiljer sig från populationen är genom slumpen
Slutsats
Statistisk interferens
Stickprov = sample
Ett urval ur populationen
Grupp A
Grupp B
Utvärdera effekt
Vi använder informationen i stickprovet för att dra slutsatser om populationen

Datanivåer (typer av variabler)

Kvalitativa variabler = kategorivariabler		**Kvantitativa variabler = numeriska variabler**	
Nominala variabler Utfallen är kategorier som ***inte kan rangordnas***	**Ordinala variabler** Utfallen är ***ordnade*** kategorier		
Blodgruppering Kön Yrke	VAS Betyg Rökning (nej, lite, mkt)		

Datanivåer (typer av variabler)

Kvalitativa variabler = kategorivariabler		**Kvantitativa variabler = numeriska variabler**	
Nominala variabler Utfallen är kategorier som ***inte kan rangordnas***	**Ordinala variabler** Utfallen är ***ordnade*** kategorier	**Intervallvariabler** Kan beräkna differenser men inte kvoter	**Kvotvariabler** Kan beräkna differenser och kvoter
Blodgruppering Kön Yrke	VAS Betyg Rökning (nej, lite, mkt)	Temperatur mätt i C och F	Längd Vikt Ålder Rökning (antal cig/ dag)

Ytterligare indelning numeriska (kvantitativa) variabler

Diskret = kan endast anta vissa värden inom ett intervall, oftast heltal

Kontinuerlig = kan i princip anta alla värden inom ett intervall

Deskriptiv statistik - centralmått

Medelvärdet $$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Median (Medianvärde används som centralmått för snedfördelade kontinuerliga / diskreta variabler samt för variabler som mäts med ordinalskalan. Medianvärdet är det värde **som ligger exakt i mitten med lika många mätvärden ovan som under sig**.

Mode = typvärde

Deskriptiv statistik - spridningsmått

Varians $$s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2 = \frac{1}{n-1}\left((x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots\right)$$

Standardavvikelse $$\sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$$

Hur mycket de olika värdena i en population avviker från medelvärdet. Om de olika värdena ligger samlade nära medelvärdet → standardavvikelsen låg.

OBS!!!! Skilj från SEM (standard error of the mean) SEM är ett mått på osäkerheten i skattningen av medelvärdet. SD är ett mått på spridningen av observationerna.
(Standardavvikelsen i stickprovet/roten ur stickprovsstorleken)

$$SE = \frac{s}{\sqrt{n}}$$

Range $X_{\max} - X_{\min}$

Inter Quartile Range = 75:e percentilen – 25:e percentilen $q_3 - q_1$

Varians och Standardavvikelse

Exempel: 5 observationer 24, 27, 28, 31, 34

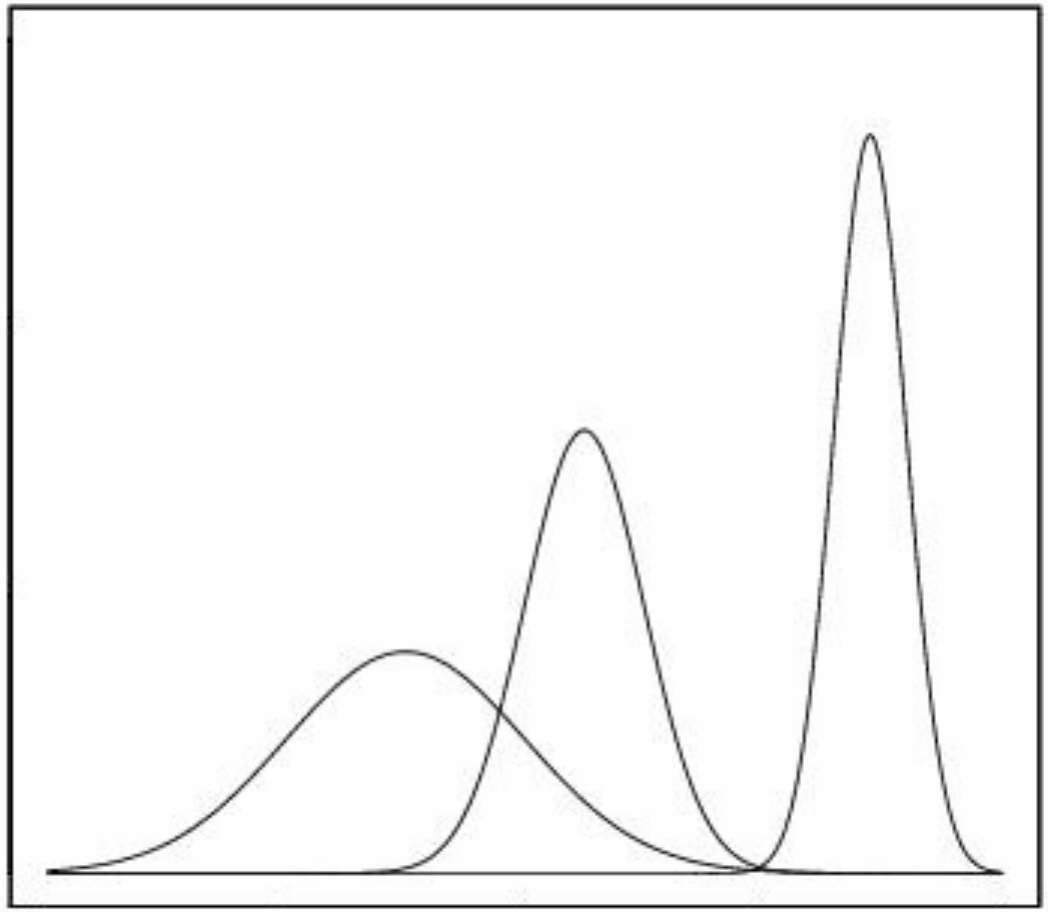
Summan: 144. Medelvärde 28,8.

Om vi tar medelvärdet och räknar hur varje värde skiljer sig från medelvärdet (-4,8; -1,8; -0,8; 2,2; 5,2) och summeras deviationerna → summan = 0

Kvadrera deviationerna, summera → 58,8

58,8/(5-1) = variansen = 14,7

Standardavvikelsen = roten ur variansen = 3,8



Lågt medelvärde. Hög varians

Högt medelvärde. Låg varians

Sammanfattning centralmått och spridningsmått

Normalfördelad variabel

→ medelvärdet och standardavvikelse

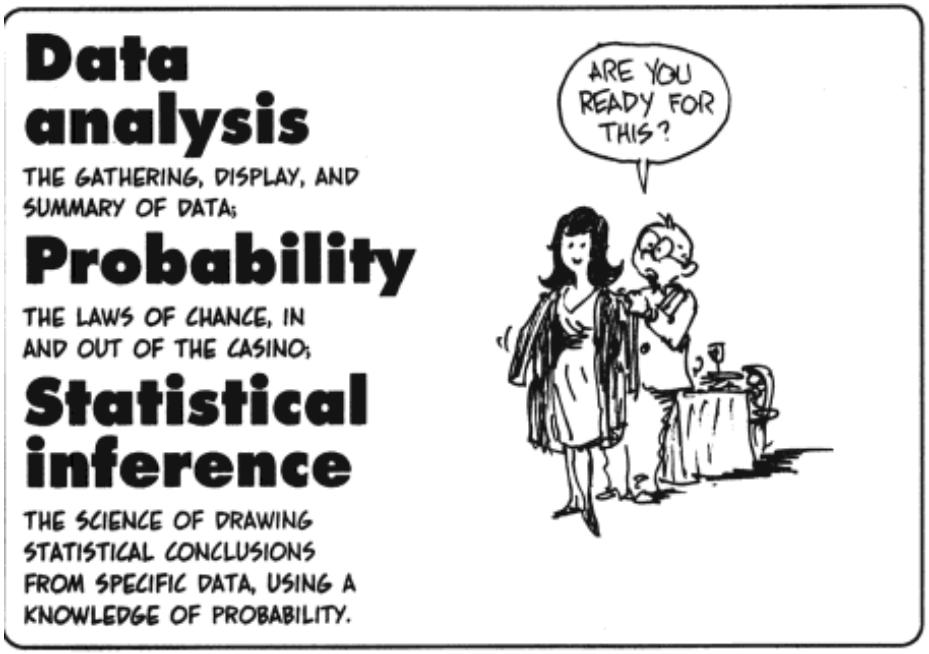
Ordinala data, kvantitativa data som ej är symmetriska (skevt fördelad variabel, extremvärden)

→ median och IQR

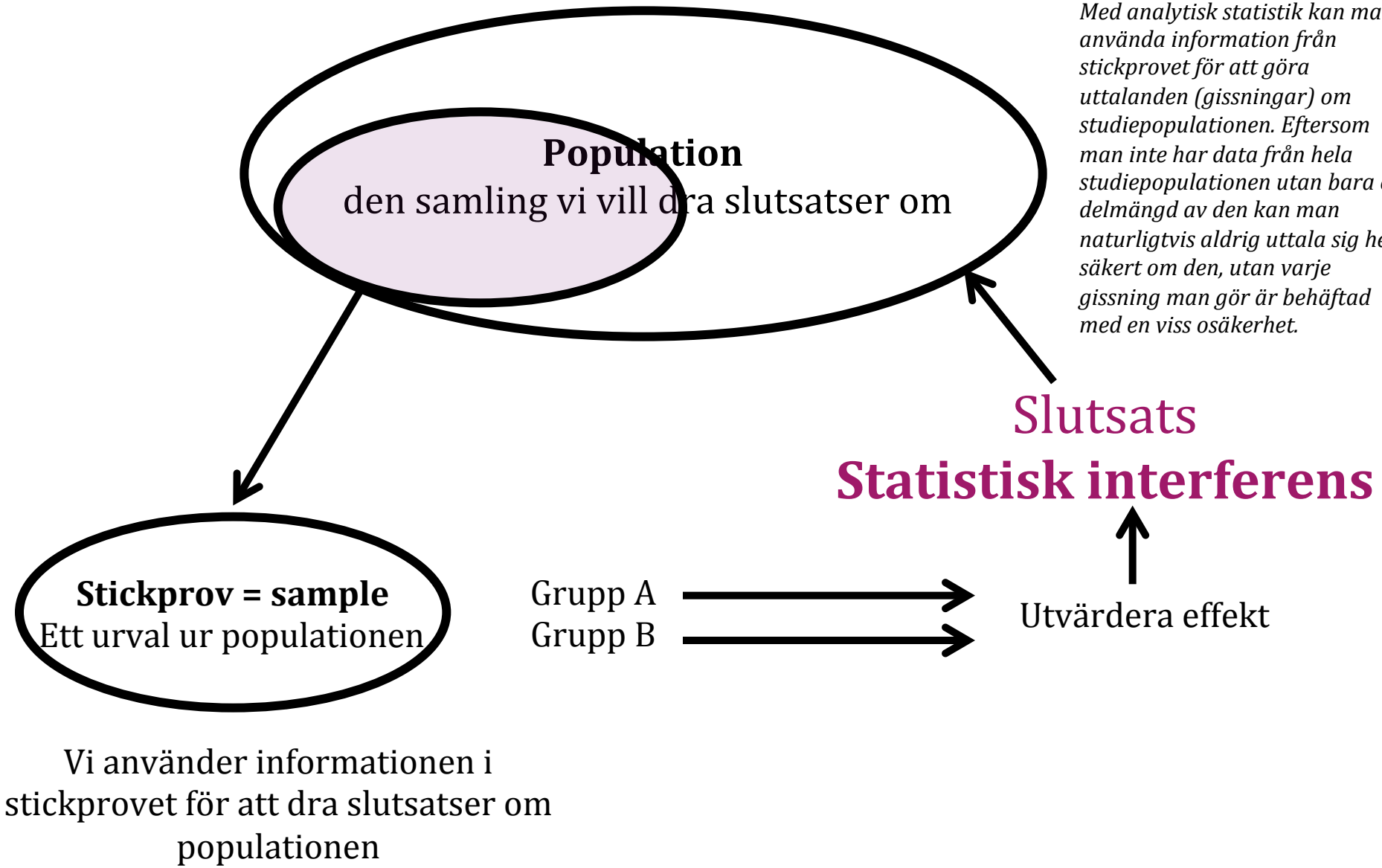
Analytisk statistik

Statistisk interferens

→de metoder som används för att utifrån ett *stickprov* dra slutsatser om en hel *population*

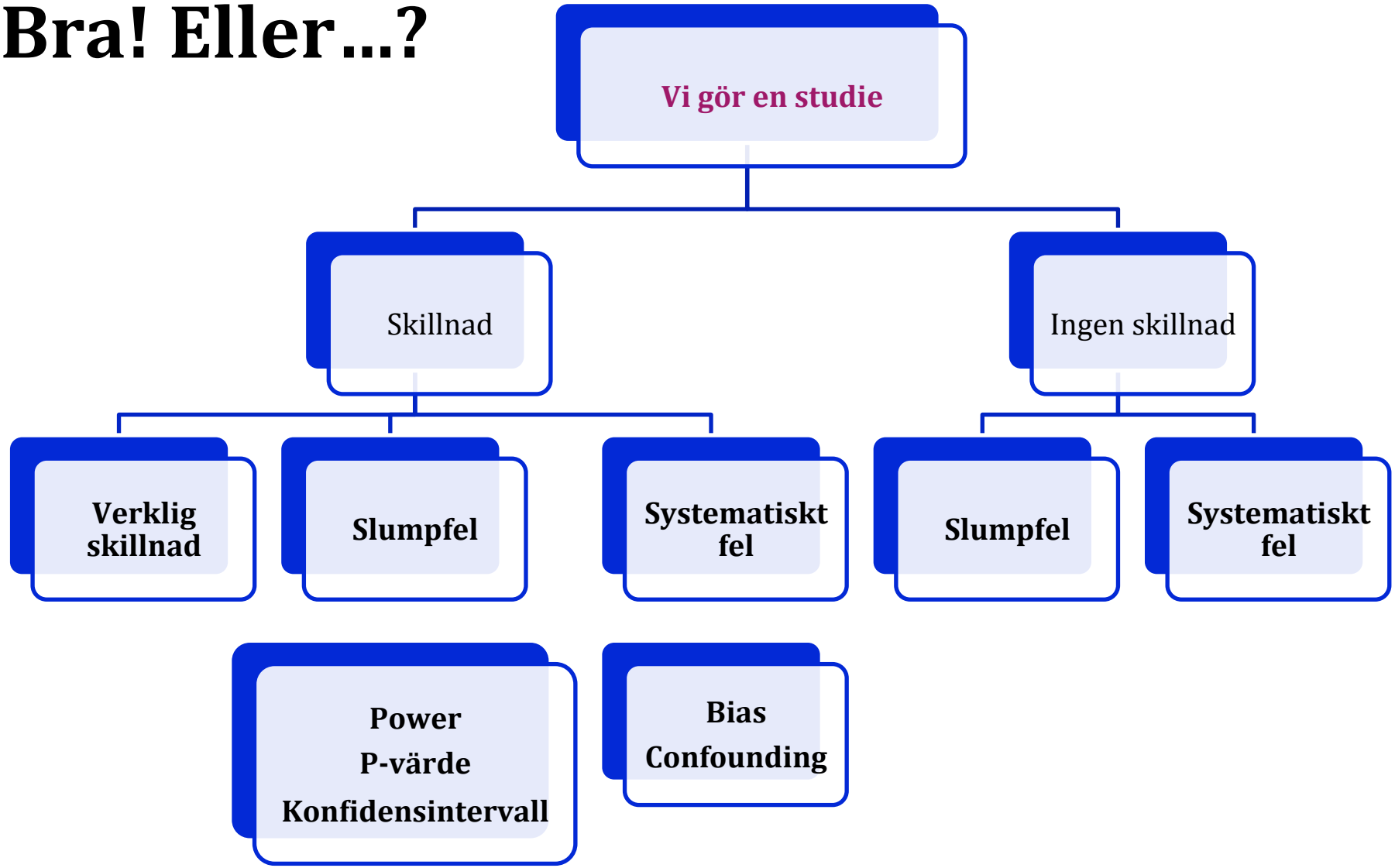


Anledningen till att man studerar ett mindre stickprov i stället för hela studiepopulationen kan vara att studiepopulationen är svår att begränsa eller definiera, eller att den är för stor för att studeras.



Med analytisk statistik kan man använda information från stickprovet för att göra uttalanden (gissningar) om studiepopulationen. Eftersom man inte har data från hela studiepopulationen utan bara en delmängd av den kan man naturligtvis aldrig uttala sig helt säkert om den, utan varje gissning man gör är behäftad med en viss osäkerhet.

Vi hittar en skillnad!
Bra! Eller…?



Statistisk interferens (slutledning)

- Skatta samband mellan exponering och utfall genom att beräkna ett eller flera riskmått
- Skatta inverkan av slumpmässiga fel (konfidensintervall och p-värden)
- Tolka resultatet (riskmåttet) med hänsyn till inverkan av slumpmässiga fel
- Tolka resultaten med hänsyn till inverkan av systematiska fel

Associationsmått

Samband mellan exponering (t.ex. en behandling) och förekomst av ett visst utfall

Absoluta och relativa mått

- Riskdifferens
- Relativ risk
  - Ex – om risken för utfall är 40% hos exponerade och 20% hos oexponerade är den absoluta riskskillnaden 20 procentenheter och den relativa risken är 2
- Odds Ratio
  - Odds = sannolikheten för en händelse/sannolikheten för ej händelse
- Hazard Ratio

Punktskattningar och konfidensintervall

Punktskattning (*eng: point estimate*) information från stickprovet används för att skatta en parameter i studiepopulationen. En punktskattning kan vara ett enskilt värde eller andel, men det kan också vara en skillnad i värden eller andelar mellan olika grupper.

Exempel: Vi väljer slumpmässigt ut 100 studenter på KI (=vårt stickprov). Vi registrerar ålder för de 100 studenterna och använder deras medelålder för att gissa – skatta – medelåldern för alla studenter vid KI. Detta är en punktskattning av medelåldern vid KI

Punktskattningar och konfidensintervall

För att ange osäkerheten i en punktskattning kan man använda ett **konfidensintervall**

Om man väljer 100 stickprov ur en studiepopulation och genomför samma mätning på alla stickprov och beräknar ett 95 konfidensintervall för varje stickprov kan man förvänta sig att 95 av konfidensintervallen täcker det sanna värdet i studiepopulationen.

→ **”ett 95 konfidensintervall täcker med 95% sannolikhet det sanna värdet i studiepopulationen”**

KI = punktskattningen ± konstant · standardfelet

Konstanten beror på konfidensintervallets konfidensgrad. Ett 90 konfidensintervall har konfidensgraden 90% och till detta använder man konstanten 1,64. Ett 95 konfidensintervall har konfidensgraden 95% och till detta använder man konstanten 1,96. Standardfelet = Standardavvikelsen i stickprovet/roten ur stickprovsstorleken

Exempel konfidensintervall

100 studenter

Medelålder 31.3

Standardavvikelsen 9.7 år $\sqrt{\frac{\sum (x-\bar{x})^2}{n-1}}$

95 % konfidensintervall för denna punktskattning

$SE = \frac{s}{\sqrt{n}}$

31.3 ± 1.96 x (9.7/√100) = 31.3 ± 1.9 → **29.4 – 33.2**

→ **Medelåldern i studiepopulationen (= alla KI studenter) med 95% sannolikhet ligger mellan 29.4 år och 33.2 år**

Hypotesprövning

Antag att man har beräknat ett konfidensintervall och därmed kunnat konstatera att det sanna värdet i studiepopulationen med 95% sannolikhet ligger mellan två specifika gränsvärden

→ ***vad kan man använda denna information till? ??***

Med analytisk statistik kan man aldrig bevisa något. Däremot kan man avfärda en **teori – hypotes** – som mindre trolig. För att göra detta använder man sig av något som kallas **hypotesprövning.**

Nollhypotes och alternativ hypotes

Nollhypotes (*eng: null hypothesis)* = det finns ***ingen*** effekt, ingen skillnad mellan grupperna,

Alternativhypotes = det **finns en verklig** skillnad, det som ”finns kvar” om man förkastar nollhypotesen (kan vara enkelsidig eller dubbelsidig)

Exempel: Vi vill veta om det finns lika många män som kvinnor på KI. Vår nollhypotes, d.v.s. hypotesen om ingen skillnad, blir då ”andelen kvinnor är lika stor som andelen män”, ”skillnaden mellan andelen män och andelen kvinnor=0” eller ”50% av alla studenter är kvinnor”.

→Vi vill kunna förkasta nollhypotesen!

Hypotesprövning med konfidensintervall

När man gör hypotesprövning med konfidensintervall beräknar man konfidensintervallet för den punktskattning man vill testa. **Enligt definitionen av konfidensintervall skall det sanna värdet i studiepopulationen med 95% sannolikhet ligga inom konfidensintervallets gränser.**

→ Om nollhypotesen ligger inom konfidensintervallets gränser kan alltså nollhypotesen mycket väl vara det sanna värdet, och nollhypotesen kan med stor sannolikhet stämma.

→ Om nollhypotesen däremot inte ligger inom intervallets gränser kan man förkasta nollhypotesen, d.v.s. göra uttalandet är nollhypotesen är mindre trolig, till förmån för alternativhypotesen. **När man har förkastat nollhypotesen säger man att resultatet är statistiskt signifikant, eller statistiskt säkerställt.**

Hypotesprövning med konfidensintervall

Exempel:

- *Nollhypotesen - andelen kvinnor vid KI är 50%.*
- *Vi beräknar ett 95% konfidensintervall för andelen kvinnor = 49%- 59%.*
- *Nollhypotesen ligger innanför konfidensintervallets gränser, vi kan därför inte förkasta nollhypotesen, d.v.s. andelen kvinnor kan vara 50%.*

Vidden på konfidensintervallet speglar studiens precision, snävare intervall = högre precision

Statistisk signifikans

- En ”gräns” (signifikansnivå) där man anser att inverkan av slumpmässiga fel är så liten att resultaten inte kan förklarats av slumpen
- Inom medicin traditionellt till 5 procent
- Godtycklig gräns!
  - Vad är skillnad på p = 0.07 och p = 0.47?
  - Vad är skillnad på p = 0.000001 och p 0.01?

P-värde

Förutom att testa nollhypoteser med konfidensintervall kan man även använda p-värde

P-värde = sannolikheten (måste ligga mellan 0 och 1) för att få det observerade utfallet (om studien upprepades under samma förhållanden) givet att nollhypotesen är sann (dvs ingen skillnad)

≈ Sannolikheten att resultatet beror på slumpen

Om p-värdet är tillräckligt litet anser man att det är orimligt att nollhypotesen är sann, och alltså förkastar man den. Definitionen av "tillräckligt liten" kan förstås variera, men den gräns man sätter upp kallas för testets signifikansnivå, eller risknivå. Vanliga nivåer är 1% (0,01), 5% (0,05) och 10% (0,10). Det är viktigt att man redan innan man börjat med de statistiska analyserna har bestämt sig för vilken signifikansnivå man skall använda!

Statistisk signifikans är ej lika med klinisk relevans!!!

P-värde vs konfidensintervall

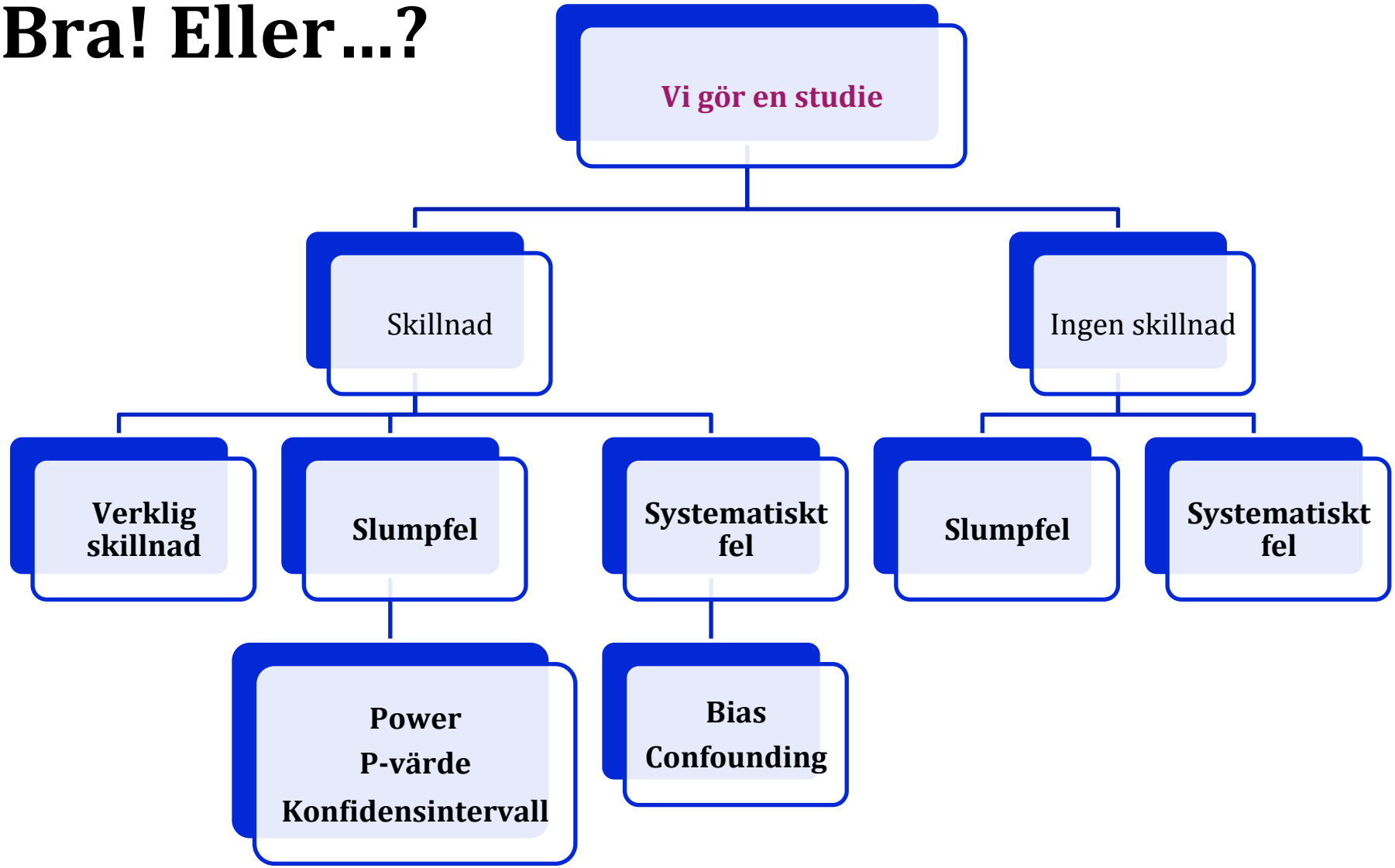
P-värde

- Används vid hypotesprövning
- Olika metoder beroende på om data är normalfördelade eller ej

Konfidensintervall

- Används vid hypotesprövning
- Uttrycker osäkerhet i en skattning
- Normalfördelade data

Vi hittar en skillnad!
Bra! Eller…?



Fel i hypotesprövningar

Naturligtvis kan man, som sagts tidigare, aldrig bevisa något med analytisk statistik, utan man löper hela tiden risk att göra fel av något slag – antingen förkastar man en nollhypotes som inte borde ha förkastats, eller så förkastar man inte en nollhypotes som borde ha förkastats. ***Dessa två felen betecknas α och β.***

→ ***Typ I fel: man hittar en effekt som egentligen inte finns.***

→ ***Typ II fel (β): man inte hittar en effekt som faktiskt finns (sample size)***

Statistisk styrka (eng: power) betecknar sannolikheten att hitta en effekt som finns, och beräknas genom 1-β.

		VERKLIGHET	
		Behandling har effekt	Behandling har ingen effekt
STUDIEN VISAR	Behandling har effekt	Bra!	α
	Behandling har ingen effekt	β	Bra!

”Power – beräkning” (styrkeberäkning)

Sannolikheten att en studie ska kunna påvisa ett förväntat resultat

→Hur stor studien måste vara för att upptäcka ett

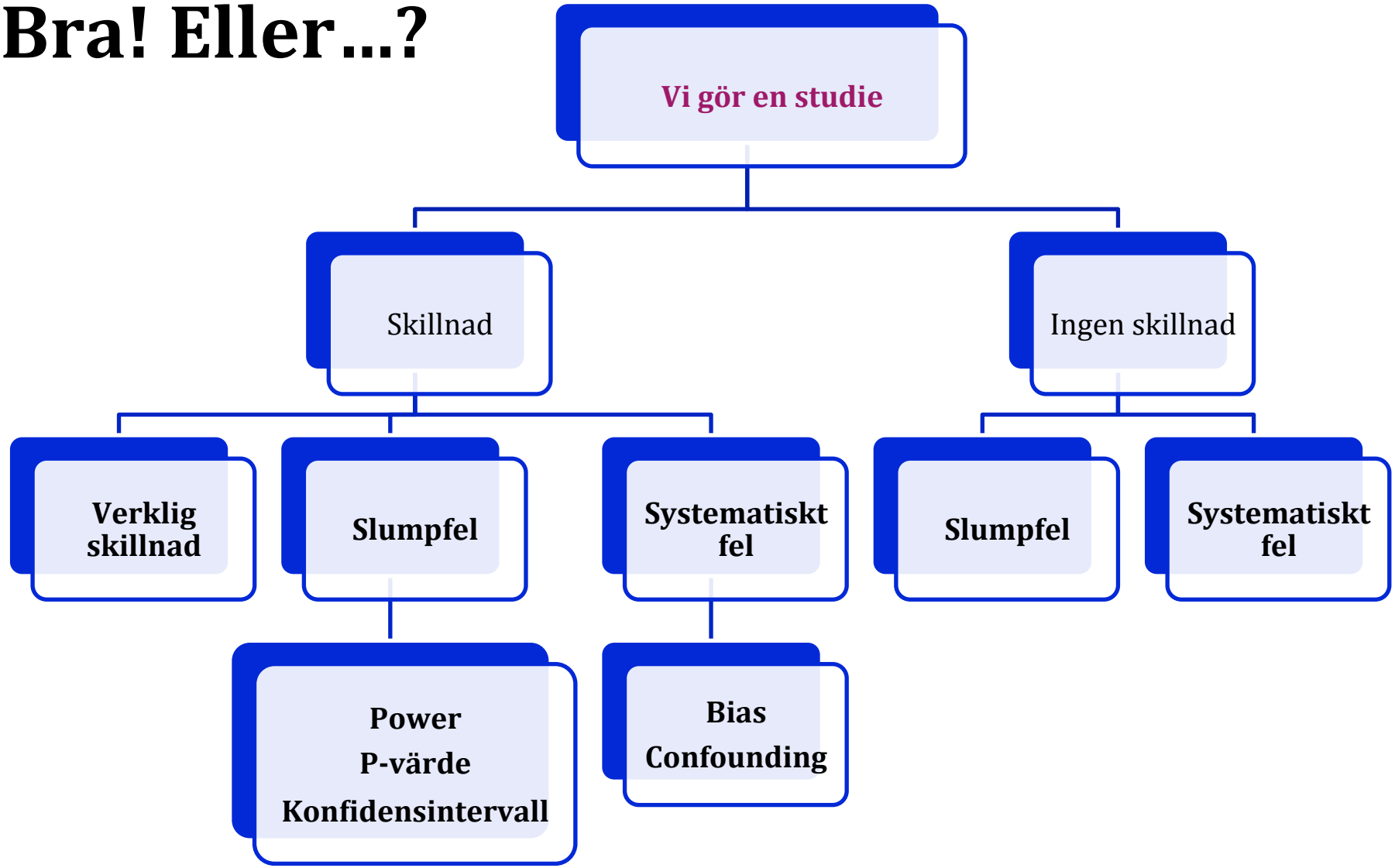
förväntat resultat (ex skillnad i sjukdomsrisk) hos exponerade

individer och oexponerade individer

80 % anses vara rimlig statistisk styrka − bestäms av:

- Studiens storlek
- Sambandets styrka (effektstorlek)
- Signifikansnivå

Vi hittar en skillnad!
Bra! Eller…?



Bias och confounding

Systematiska fel (*eng: bias) uppstår när mätresultaten har en tendens att avvika från det sanna värdet på ett systematiskt sätt* (dvs alla mätvärden avviker från det sanna värdet på samma sätt).

- Selection bias – fel vid identifiering av studiepopulationen
  - Ex man undersöker enbart de som frivilligt anmäler sig till undersökning
- Observation or information bias – fel vid mätning av exponering eller utfall
  - Ex s.k. recall bias i falkontrollstudier, ”lättare” att kontroller erinrar sig att de varit utsatta för misstänka exponeringsfaktorer

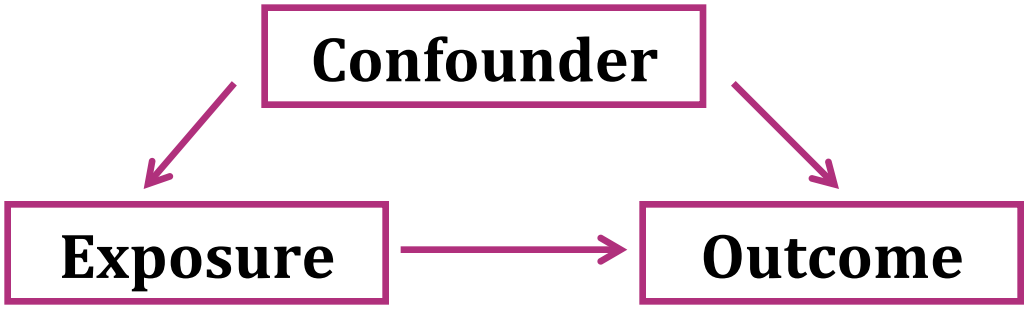
Bias och confounding

Confounding (störfaktor): exponerade och oexponerade kan skilja sig med avseende på förekomsten av någon annan faktor som i sig påverkar risken att insjukna.

1 - associerad med exponeringen

2 - självständig riskfaktor för sjukdomen

3 - inte ett mellansteg i kausalkedjan från exponering till effekt (ej E → C → O)



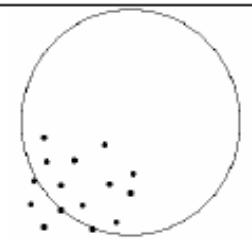
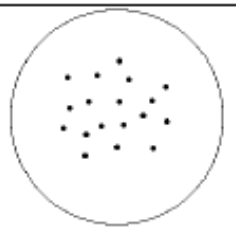
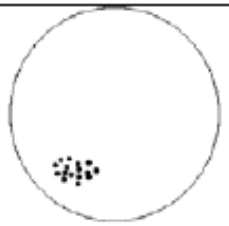
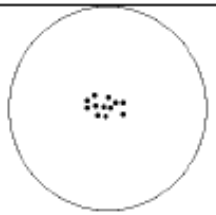
Hur hantera confounding?

- Stratifiering
- Matchning
- Regressionsmodell (flera förväxlingsfaktorer)

*Skilj från **effektmodifiering** = sambandet mellan en exponering och ett utfall varierar beroende på värdet av en annan faktor (effektmodifieraren)*

Exempel olika risker i olika åldersgrupper. Olika risker för män och kvinnor. Effektmodifiering är ett fynd att redovisa

Felkällor

	Låg validitet	Hög validitet
Låg precision		
Hög precision		

Fullständig **precision** – frånvaro av slumpmässiga fel.

Åtgärd → exempelvis öka antalet individer, bättre mätinstrument

Fullständig **validitet** – frånvaro av systematiska fel

Normalfördelning

En **teoretisk fördelning** (≈ definierad av en matematisk formel).

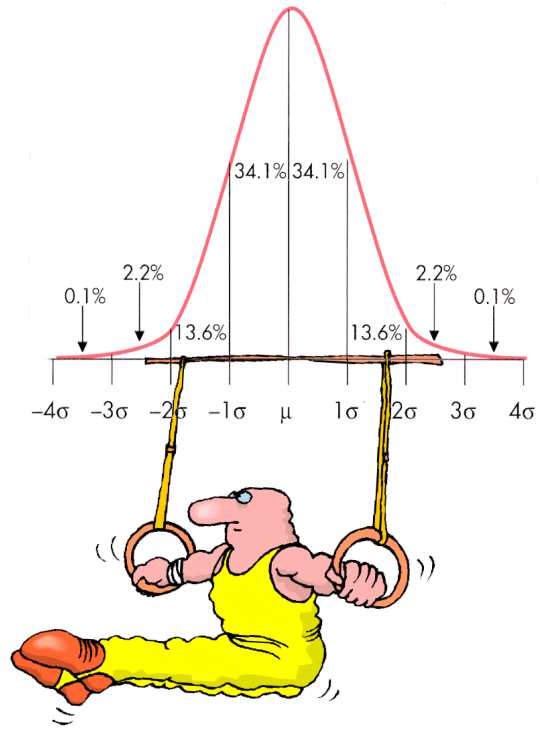
Många fenomen i naturen följer normalfördelningen, exempelvis födelsevikter.

Går från minus till plus oändligheten.

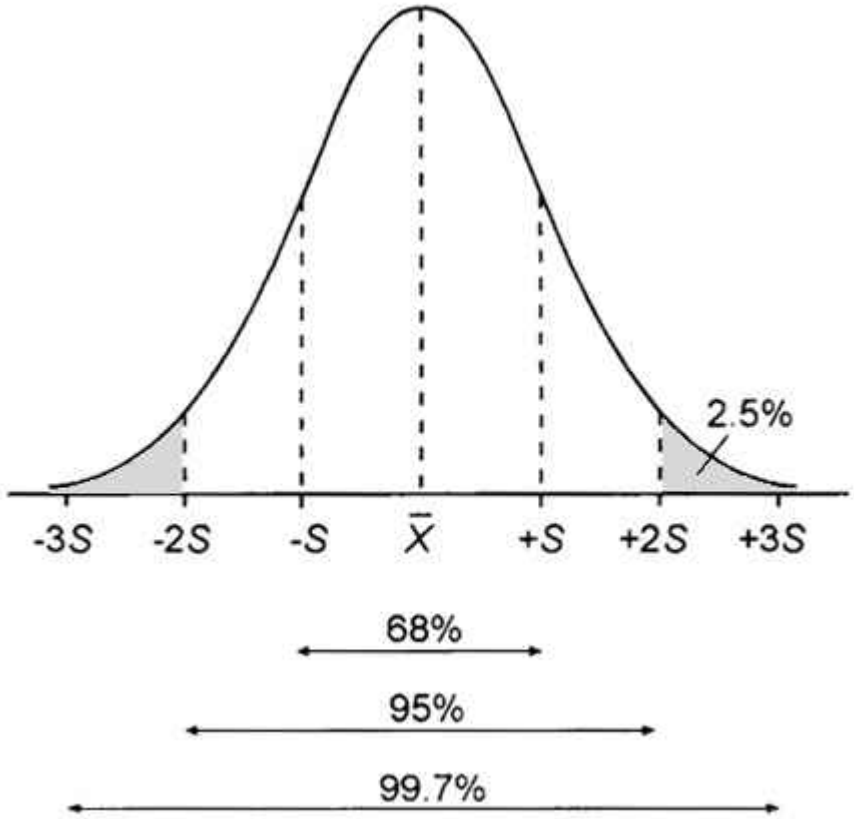
Symmetrisk. Ytan under kurvan n är 1

Definieras utifrån **två faktorer**:

- medelvärdet
- standardavvikelsen (= roten ur variansen)



Normalfördelning forts.



Inom intervallet som går från 1,96 standardavvikelser under medelvärdet till 1,96 standardavvikelser över medelvärdet ligger 95% av observationerna

Sjukdomsförekomst

Incidence – *"Hur många blir sjuka?"*

→ mått på hur många som insjuknar under en given tidsperiod

Kumulativ incidence (=risk): andelen av en population som insjuknar under en given tidsperiod. Antalet insjuknade/antal friska vid start

Ex årligen insjuknar 5 % av Sveriges befolkning i influensa

Incidence rate: antal nya sjukdomsfall under en given tidsperiod/ total risktid.

Ex incidensen av Mb Chron är 5 per 100.000 personår

Prevalens – *"Hur många är sjuka?"*

"Disease status" "Disease burden in a population"

→proportionen av populationen som har en sjukdom (eller egenskap) vid ett givet tillfäll3

Sensitivitet/specificitet

Sensitivitet = sannolikheten att en sjuk individ blir klassificerad som sjuk

Specificitet = sannolikheten att en frisk individ blir klassificerad som frisk

Positivt och negativt prediktivt värde

Sannolikheten för sjukdom, givet ett testresultat, kallas för testets prediktiva värde.

Positivt prediktivt värde = sannolikheten att en individ med ett "sjukt" provsvar verkligen är sjuk

Negativt prediktivt värde = sannolikheten att en individ med ett "friskt" provsvar verkligen är frisk

Verklighet

Test

	Sjuk	Frisk
Sjuk	A	B
Frisk	C	D

Pos pred värde
A/(A+B)

Neg pred värde
D/(C+D)

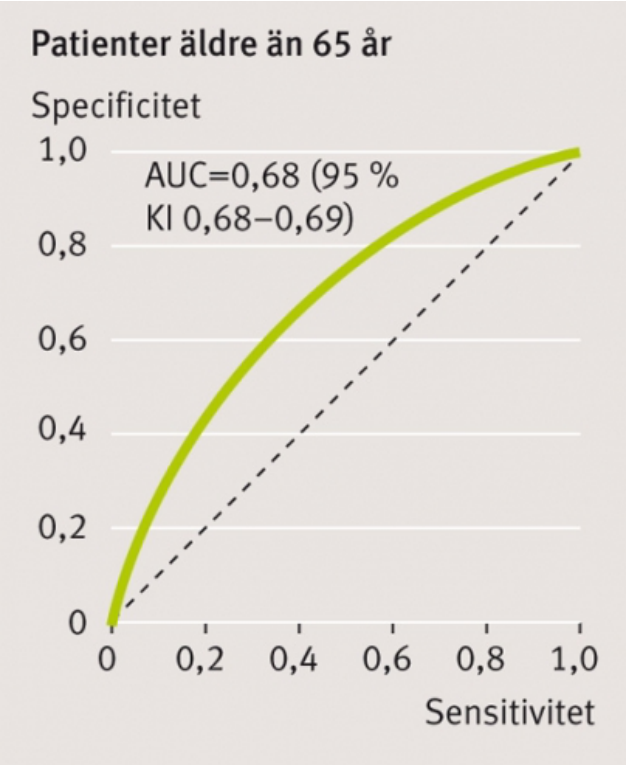
Sensitivitet
A/(A+C)

Specificitet
D/(B+D)


That's all Folks!

ROC-kurva

ROC-diagram. Kurvan visar hur sensitivitet och specificitet ändras då det diskriminativa värdet successivt ändras från det mest patologiska värdet till det minst patologiska. Ju närmare det övre vänstra hörnet kurvan går desto bättre diagnostisk förmåga.



Exempel statistiska tester/modeller

Parametriska tester/modeller *(bygger på antaganden om hur data fördelar sig (oftast normalfördelningen), skattar en effekt (ex skillnad i medelvärde) för vilken man kan beräkna konfidensintervall och p-värde)*

- Two sample t-test
- Paired t-test
- Regressionsanalys
- Variansanalys (ANOVA)

Icke-parametriska: *(kan anv oberoende av hur data fördelar sig, baseras på observationernas ranger, beräknar endast p-värde)*

- Mann-Whitney two sample test
- Chi-2 test
- Fischers exakta test
- Wilcoxons matched pairs test

4. STATISTISKA TEST – EN ÖVERBLICK

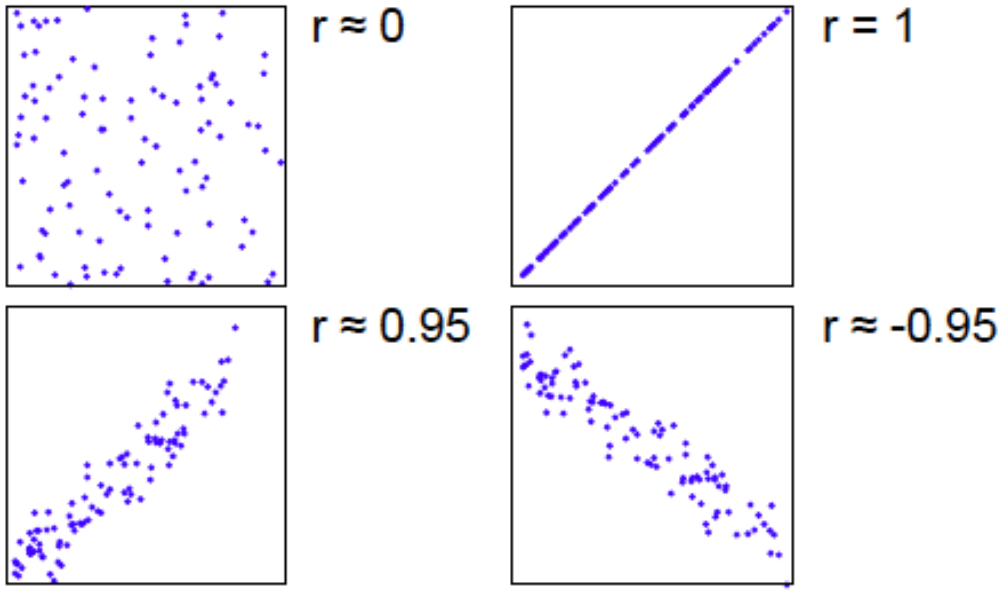
	Nominaldata	Ordinaldata	Kvantitativa data – icke-parametriskt test	Kvantitativa data – parametriskt test
En grupp vs ett förutbestämt värde	Binomialtest	Teckentest	Teckentest	One-sample t-test
Två oberoende grupper	Chi-två-test Fishers exakta test	Mann-Whitney Wilcoxons rangsummetest	Mann-Whitney Wilcoxons rangsummetest	Two-sample t-test (Linjär regression)
Fler än två oberoende grupper	Chi-två-test	Kruskal-Wallis	Kruskal-Wallis	ANOVA (Linjär regression)
Före-efter-mätningar	McNemars test	Wilcoxons teckenrangtest	Wilcoxons teckenrangtest	Parat t-test
Association mellan två variabler	Kontingenskoefficient	Spearmans korrelations-koefficient	Spearmans korrelations-koefficient	Pearsons korrelations-koefficient Linjär regression

Logistisk regression: Om man har en utfallsvariabel som bara kan anta två olika värden kan man använda logistisk regression för att skatta en oddskvot.

Konfidensintervall: Om data uppfyller förutsättningarna kan man också beräkna konfidensintervall för andel, skillnad i andel, medelvärde, och skillnad i medelvärde.

Korrelationskoefficient

Korrelation är ett begrepp inom statistiken som anger styrkan eller riktningen av ett samband mellan 2 variabler. Det kallas även korrelationskoefficient. Uttrycks som ett samband mellan 1 och -1 där 0 anger inget samband,
1 anger maximalt pos samband
och -1 maximalt neg samband



Regressionsanalys

- Studera sambandet mellan olika variabler

- Centrala inom epidemiologisk forskning, multipel regressionsanalys kan användas för att kontrollera för förväxling (confounder) och för att skapa prediktionsmodeller

- Data anv för att bestämma en ekvation eller matematisk modell som beskriver hur en utfallsvariabel varierar som en funktion av en eller flera förklarande variabler. Utifrån modellen kan sedan sambandet mellan variablerna kvantifieras.