

Innehåll

I. Grundläggande begrepp

II. Deskriptiv statistik

III. Statistisk inferens

- Hypotesprövning

Statistiska analyser

Parametriska analyser

- Univariata analyser

- Multivariata analyser

Icke-parametriska analyser

- Univariata analyser

- Multivariata analyser



Forskningsfrågor

Finns det skillnad mellan personer med psykos, depression eller ångest avseende livskvalitet?

Finns det skillnader över tid mellan baseline, tre månader och sex månader avseende god livskvalitet för personer med psykos?

Skiljer sig grupper med olika psykiatri diagnoser och utbildningsnivå åt avseende livskvalitet?

I vilken grad kan högt blodtryck förklaras av saltintag, vikt och fysisk aktivitet?

I vilken grad kan god eller dålig hälsa förklaras av utbildningsnivå, religion och inställning till goda kostvanor?

Statistiska analyser

(tre eller fler grupper/måttillfällen/variabler)

Parametriska

Skillnader mellan tre eller fler grupper vid samma tidpunkt
Oberoende ANOVA

Skillnader inom samma grupp över tid (tre eller fler mätpunkter)
Beroende ANOVA

Samband mellan tre eller fler variabler
Multipel regression

Icke-parametriska

Kruskal Wallis

Friedman

Logistisk regression

Multivariata statistiska analyser

Parametrisk analys:
ANOVA (ANalysis Of Variance)
Variansanalys
Oberoende och beroende mätningar

Icke-parametriska analyser:
Kruskal-Wallis - oberoende mätningar
Friedman - beroende mätningar

Multivariat parametrisk statistisk analys

Varför ANOVA?

- Många t-test ökar risken för Typ I-fel
- Typ I-fel = **förkastar** nollhypotesen trots att den är **sann**

Ex. När forskaren behåller analysresultaten och säger att de är korrekta men egentligen så finns ingen skillnad eller inget samband.

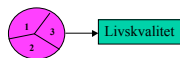
- Fördelen med ANOVA är att endast ett enda test utförs (F-test). ANOVA utnyttjar all information samtidigt

Multivariat parametrisk statistisk analys

ANOVA används när vi vill

1. Undersöka skillnaden mellan 3 eller fler grupper vid ett mätillfälle=oberoende ANOVA.

Forskningsfråga: Finns det skillnad i livskvalitet mellan personer med 1. psykos, 2. depression och 3. ångest?



ANOVA används när vi vill

2. Undersöka skillnaden mellan 3 eller fler mätillfällen av samma grupp eller följa en grupp över tid = beroende ANOVA.



Forskningsfråga: Finns det skillnader över tid mellan baseline, tre månader och sex månader avseende god livskvalitet för personer med psykos?

Baseline 3 månader 6 månader

ANOVA

- Huvudsyftet är att analysera varianser
- Utgångspunkten för ANOVA är beräkningen av materialets samlade varians
- Variansen = standardavvikelsen i kvadrat= S^2

Arbetsordning

Steg 1 ANOVA (F-test)

Steg 2 Post-hoc test (eftertest)

Grupperingsvariabeln, studerar effekten av en variabel på nominal skalnivå (den oberoende variabeln) kategoridata.



Effektvariabeln är en variabel på intervall/kvot skalnivå (den beroende variabeln).



1. Psykos

Grupp 1

2. Depression

Grupp 2

3. Ångest

Grupp 3

Effektvariabel
(livskvalitet)



ANOVA steg 1

- F-testet ger en samtidig jämförelse mellan samtliga grupper medelvärden
- F-testet visar endast om det finns åtminstone en statistiskt signifikant medelvärdesskillnad mellan grupperna

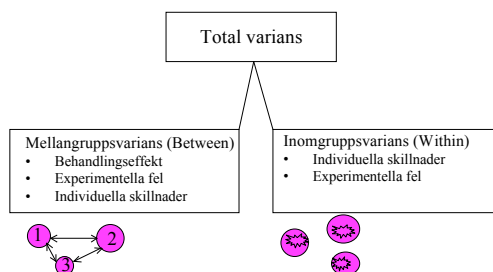
ANOVA steg 2

- Post hoc test (eftertest) används för att ta reda på mellan vilka grupper eller vilka mätningar det finns en signifikant skillnad
- Post hoc test kontrollerar risken för Typ I-fel i högre utsträckning än vanliga t-test (Fisher LSD, Tukey HSD, Scheffé m.fl.)

Variationskällor

- Det intressanta är att studera variationer i variabelvärden och vad dessa variationer beror på
- När skillnader mellan olika grupper undersöks är det storleken på variationerna mellan grupperna som studeras
- För att kunna uttala sig om skillnader mellan grupperna måste variationer inom grupperna också studeras

Envägs oberoende ANOVA



Envägs oberoende ANOVA

Forskningsfråga: Finns det skillnad i livskvalitet mellan personer med 1. psykos, 2. depression och 3. ångest?

- Beroende variabel: livskvalitet
- Oberoende variabel (grupperingsvariabeln): diagnosgrupper, t.ex. 1) psykos, 2) depression, 3) ångest.



Kruskal-Wallis

- Den icke-parametriska motsvarigheten till oberoende variansanalys
- Jämför "mean ranks" för varje grupp, medelvärdena rangordnas
- Eftertest (parvisa jämförelser) kan göras för hand eller med Mann-Whitney

Friedman testet

- Den icke-parametriska motsvarigheten till envägs variansanalys för upprepade mätningar.
- Jämför ”mean ranks”.

Tvåvägs ANOVA (variansanalys)

- Två oberoende variabler (två förklarande grupperingsvariabler). Ex. diagnos och utbildning
- Vi kan statistiskt pröva effekterna av de båda faktorerna separat för att se om huvudeffekterna är signifikanta.

Forskningsfråga: Skiljer sig grupper med olika psykiatri diagnoser och utbildningsnivå åt avseende livskvalitet?

Tvåvägs ANOVA (variansanalys)

- Vi kan dessutom pröva faktorerna tillsammans för att se om de samvarierar (om en signifikant interaktionseffekt föreligger).
- Två faktorer (variabler) sägs interagera (det finns en interaktionseffekt) om skillnader i nivåer hos den ena faktorn beror på (påverkas av) nivån hos den andra faktorn.

En tvåvägs variansanalys besvarar tre frågor:

- Finns det någon huvudeffekt (main effect) av faktor A? Ex. diagnos: psykos, depression och ångest.
- Finns det någon huvudeffekt (main effect) av faktor B? Ex. utbildningsnivå: högskola, gymnasium, folkskola.
- Finns det någon interaktionseffekt (interaction effect) av faktor A och B?

Korrelation (samband) och regression sid 105 i E

Statistiska analyser

Parametriska

Pearson's produktmoment-korrelation

Regressionsanalys

Multipel regressionsanalys

Logistisk regressionsanalys

Icke-parametrisk

Samband mellan 2 variabler

Spearman's rangkorrelation

Samband mellan >2 variabler

Sambandsanalyser

Uppgiften är att undersöka:

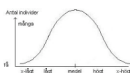
1. Om det finns något samband mellan två variabler.
2. Och om ja, beskriva hur sambandet ser ut.

För sambandsanalyser finns två tekniker:

1. **Korrelation (r):** talar om sambandet är positivt eller negativt och hur starkt sambandet är mellan variablerna.
Ex. $r = .63$ eller $r = -.42$. (Spridningsdiagram)
2. **Regression:** används för att förutsäga (predicera) värdet på en variabel utifrån kunskapen om en annan

Förutsättningar

- Oberoende mätningar
- Materialet ska vara normalfördelat
- Intervall- eller kvotskala



- Även om vi i en studie finner en hög korrelation (ett starkt samband) är det inte säkert att det därför finns ett orsakssamband (kausalsamband).
- **Skensamband, varning!** (sid 113 E)
- **Finns anledning att varna för okritisk användning av metoder för att studera samband.**
- Man bör vara försiktig vid tolkningen av sambandsmått.

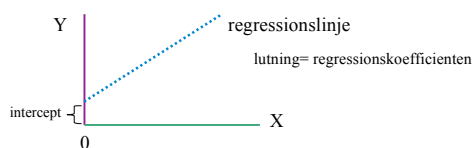


Linjär regressionsanalys

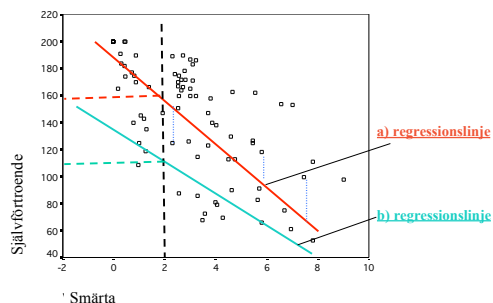
- Pearsons r uttrycker sambandets styrka.
- Den säger inget om hur snabbt y ökar med x
- y är den beroende variabeln
- x är den oberoende variabeln (förklaringsvariabel)



- Med hjälp av regressionssambandet kan vi beräkna förväntade värden för y på givna värden på x .
- Vi gör en prediktion eller förutsägelse.



Spridningsdiagram



Multipel regression

- Resultatet av multipel regression är en ekvation som representerar den bästa prediktionen av en förklarad variabel utifrån flera kontinuerliga eller dikotoma förklarande variabler.
- Att särskilja effekten av enskilda förklaringsvariabler.

- Inga kausala/orsaks samband analyseras.
- Regression fungerar bäst när varje förklarande variabel har starkt samband med den förklarade variabeln men inte med de andra förklarande variablerna.
- OBS! Extrema värden (outliers) kan påverka regressionslösningen i hög grad och måste därför hanteras.

Logistisk regression, sid 222 i E

- Samband mellan mer än två variabler.
- Icke-parametriskt test.
- Om den beroende variabeln (y) är dikotom.
- Logistisk regression kan även användas när det finns fler än två nivåer på den beroende variabeln.
- De oberoende variablerna kan vara kontinuerliga, diskreta, dikotoma eller en blandning.



Forskningsfråga: I vilken grad kan god eller dålig hälsa förklaras av utbildningsnivå, religion och inställning till goda kostvanor?
