Hoorcollege week 4. 21-09-15

Toetsen en associatiematen

Associatiematen: zeggen iets over de *sterkte* van relatie tussen twee (of meer) variabelen. Het gaat dus niet over wie wat voorspelt. Is er een sterke relatie, ja of nee?

Toetsen: zeggen iets over de *generaliseerbaarheid* van de resultaten (van steekproef naar populatie) Hierbij ga je verder; is het significant/is het verband sterk genoeg?

Toetsen: drie onderdelen

Dit moet je rapporteren:

1. Toetsstatistiek (χ^2, t, F): getal wat iets zegt over het verband. Als je alleen de p-waarde geeft, is dat niet genoeg.
2. Vrijheidsgraden (df): het aantal onafhankelijke observaties op grond waarvan die statistiek is berekend. Zonder vrijheidsgraden zijn de toetswaarden nutteloos.
3. *p*-waarde: de kans dat een bepaald resultaat op toeval berust. Wordt gebaseerd op toetsstatistiek en de vrijheidsgraden.

Toetsen

- *Kans*-termen

"gesteld dat de nul-hypothese (de "niets aan de hand" hypothese) waar is, hoe groot is dan de kans op de gevonden samenhang?" de kans dat wat je vindt, berust op toeval.

- Hypothese/propositie vs. H0/H1

Hypothese/propositie: woordelijke versie van wat jij verwacht, bijvoorbeeld: motivatie en tevredenheid hangen positief met elkaar samen.

(correlatievoorbeeld)

Wat je verwacht te vinden in de test:

H0: $r = 0$ geen effect, er is geen verband tussen de variabelen

H1: $r > 0$ er is een positief effect

Pearson correlatie

- Maat voor de sterkte van de lineaire relatie tussen twee *interval* variabelen uitgedrukt in *r.* Maat die je kan gebruiken als je variabelen allebei in interval gemeten zijn.
- $-1 \leq r \leq 1$ -1 = sterk negatief   +1 = sterk positief
- Absolute waarde stijgt => sterkte relatie stijgt ook

(*r* zegt niets over "slope"/helling)

Hoe dichter de puntjes van correlaties bij de virtuele lijn is, die je kan trekken, hoe sterker de associatie maat en hoe sterker r.

Het teken (+ of -) van de correlatie zegt *an sich* niets over interpretatie. Kijk goed naar je vragen: wat betekent het nu?

Pearson is speciaal gemaakt voor lineaire correlaties. Scatterplot, en kijk of het klopt.

Pearson correlatie: assumpties

- Beschrijvende statistiek (als maat van lineaire samenhang in bepaalde steekproef): geen enkele assumpties
- Voor toetsen bij correlaties (van steekproef naar populatie) wordt een bivariate normaalverdeling aangenomen. Dit is moeilijk toetsbaar. De toets is gelukkig redelijk robuust t.a.v. schendingen hiervan.

T-test

Wanneer?: Als *twee groepen* (k = 2) worden vergeleken op een *interval*-variabele.

Bijvoorbeeld: mannen/vrouwen, ja/nee

Hypothese: Mannen zijn banger voor spinnen dan vrouwen

H0: μ(man) = μ(vrouw)

H1: μ(man) > μ(vrouw)

t-test

- Betrokken parameters
  - t-waarde: volgt t-verdeling
  - df

SPSS zoekt het significantieniveau voor je op (p-waarde)

- → Resultaat: Bijv. $t(38) = 4.5$, $p < .001$

Parameter altijd cursief, 38: vrijheidsgraden, p-waarde kun je nooit vanuit gaan → kijk wat zijn de gemiddelden (in dit voorbeeld).

t-test: assumpties

1. Afhankelijke variabele is in beide populaties normaal verdeeld
2. Varianties zijn in beide deelpopulaties gelijk (homogeniteit van varianties)
3. Observaties (errors) zijn onafhankelijk van elkaar

t-test: assumpties

Ad 1. Als N1 en N2 voldoende groot, dan benadert de verdeling van steekproefgemiddelde "vanzelf" de gewenste normaalverdeling

Ad 2. Belangrijk bij ongelijke groepsgrootte. Max:min ≥ 1.5 : 1) Neem bij schending (blijkt uit Levene toets) de *separate variance estimate* van t.

Ad 3. Kwestie van goede onderzoeksopzet (bijv. vragenlijsten afzonderlijk laten invullen)

Wanneer je de output afleest:

Levene's test/variances is een assumptie check → is dus geen t-test

De means → de echte t-test

Als de levene's test significant is dus <0.05 dan pak je dus de onderste.

Als de levene's test niet significant is dus >0.05 dan pak je dus de bovenste.

Eenweg variantie-analyse (one-way ANOVA)

Wanneer?: Als *meer dan twee groepen* (k > 2) worden vergeleken op een *interval*-variabele

Waarom niet meerdere t-toetsen? Het aantal t-toetsen (h) neemt snel toe met het aantal groepen (k): h = k (k - 1)/2. De kans op fouten neemt toe, op het moment dat je t-testen herhaald.

- Kans op toevalstreffers (type I fouten) neemt eveneens toe
- Liever één F-toets dan reeks t-toetsen

Eenweg variantie-analyse (one-way ANOVA)

- Betrokken parameters
- F-waarde
- df (within) en df (between) (twee vrijheidsgraden nodig bij een F-test)
  - degrees of freedom within (aantal proefpersonen die in totaal hebben meegedaan)
  - degrees of freedom between (aantal groepen)

SPSS zoekt weer het significantieniveau voor je op (p-waarde)

Resultaat: tenminste twee groepsgemiddeldes verschillen van elkaar, maar welke? Kijk weer naar de gemiddelden.

One-way ANOVA: assumpties (zelfde als t-toets)

1. Afhankelijke variabele is in beide populaties normaal verdeeld
2. Varianties zijn in beide deelpopulaties gelijk (homogeniteit van varianties)
3. Observaties (errors) zijn onafhankelijk van elkaar

One-way ANOVA: assumpties

Ad 1. Als N1 en N2 voldoende groot, dan benadert de verdeling van steekproefgemiddelde "vanzelf" de gewenste normaalverdeling

Ad 2. Belangrijk bij ongelijke groepsgrootte. Max:min ≥ 1.5 : 1) *Oplossingen schending lastiger dan t-toets* (transformatie data, Kruskal-Wallis variantie-analyse)

Ad 3. Kwestie van goede onderzoeksopzet (bijv. vragenlijsten afzonderlijk laten invullen)

Mann-Whitney U-test

Wanneer? Als *twee groepen* (k = 2) worden vergeleken op een *ordinale* variabele (of als assumpties t-test niet voldaan)

Mann-Whitney U-test

Voorbeeld:

Hypothese: Mannen hebben hoger opleidingsniveau dan vrouwen

H0: Mediaan(man) = Mediaan(vrouw)

H1: Mediaan(man) > Mediaan (vrouw)

Betrokken parameter:

- Mann-Whitney U/ Wilcoxon W/ Z
- Geen df!

Resultaat:

Bijv. MWU = 345, $p < .001$

MWU: assumpties

- Onafhankelijke steekproeven uit de populatie
- Meetniveau is ten minste ordinaal. Kan daarom ook gebruikt worden wanneer de data niet

aan alle assumpties van t-test voldoet.

Regressie

Wanneer?: Een *interval*-variabele (Y) zo goed mogelijk wilt *voorspellen* uit één (enkelvoudige regressie) of meerdere (multipele regressie) *interval*-variabele(n) (X_1, X_2, ... X_k). X voorspelt Y.

Bijvoorbeeld: Hypothese: Hoe ouder, hoe wijzer.

H0: B = 0

H1: B > 0

B betekent de helling/sloop.

- Lineaire enkelvoudige regressie

(a = constante, y-dakje= voorspelde waarde van y, b = afhankelijk variabele)

- Multipele regressive (meerdere voorspellers → meerdere b's en meerdere x'en)
- *Voorspellen* van Y uit X

    X → Y

- Enkelvoudige lineaire regressie vs. correlatie

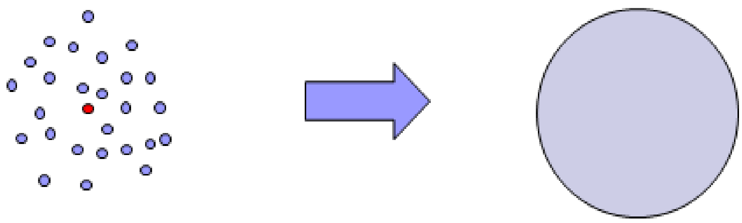
    X → Y   X1 ←verband→ X2

Waarom geen (reeks) correlaties?

- Geen "richting" in correlaties
- Multipele regressie: unieke en gezamenlijke bijdrage van alle onafhankelijke variabelen, kan je bekijken op je centrale concept.

Regressie: bijdrage

Venn-diagrammen: variantie



• = μ • = observaties

μ = populatie gemiddelde

Hoe groter de cirkel (variaties in observaties), hoe groter de variantie.

Zie slide 36 en 37 voor de ven-diagrammen.

Correlatie/enkelvoudige regressie: variantie van X is gele bolletje variantie Y is het rode bolletjes. Hoe groter de overlap, hoe groter meer x y voorspelt, hoe groter het verband is tussen beide.

Multipele regressive: je hebt variantie van x1, maar hebt daarnaast meerdere voorspellers. De overlap is hier de gezamenlijke bijdrage. Er wordt getest welk deel ervan gezamenlijk verklaart wordt door je x'en.

Er zijn nog meer stukken; unieke bijdragen.

Zie slide 36 en 37

Regressie: bijdrage

- Gezamenlijke bijdrage: R^2 getoetst d.m.v. F-toets ("ANOVA" tabel in SPSS)
- Unieke bijdrage: b/β voor iedere X getoetst d.m.v. t-waarde ("coefficients" tabel in SPSS) hierin staan de beta's en de constante.
- In theorie kan R^2 significant zijn, zonder significante b/β's: dat gebeurt als die overlap behoorlijk groot is.

Regressie

Betrokken parameters:

- b of β: positief of negatief
- a: constante (vaak geen inhoudelijke interpretatie)
- R^2: proportie verklaarde variantie in het geheel

Neiging tot overschatting R^2 in populatie: adjusted R^2

- B vs. β

B = *ongestandaardiseerd*: als X een eenheid toeneemt, neemt Y (na correctie voor alle andere X-en; als aanwezig) b eenheden toe.

β = *gestandaardiseerd*: als X een standaarddeviatie toeneemt, neemt Y (na correctie voor alle andere X-en; als aanwezig) β eenheden toe.

Er niet per se een voorkeur, maar wees je ervan bewust dat er verschil in zit.

Voordeel van β: onafhankelijk van schaling X en Y, dus vaak beter interpreteerbaar. Als je een 1 tot 7 schaal hebt en een inkomen schaal, kan je ze gemakkelijker interpreteren.

Nadeel van β: geijkt op standaarddeviaties in steekproef, dus bij vergelijking met andere steekproeven problematisch.

Regressie: assumpties en uitbijters

Uitbijters (lees: outlier) (let op: komt een tentamenvraag over uitbijters):

Een *uitbijter* is een persoon met scorepatroon dat sterk afwijkt van de rest in de steekproef. Kan onevenredig veel invloed op uitkomst hebben.

- *Univariaat:* afwijkend op een variabale (X of Y) hele hoge X waarde en normale Y waarde (of andersom)
- *Multivariaat:* afwijkende combinatie van variabelen (bijv. 15-jarige met jaarinkomen van 50.000 euro: Wizkid)

Regressie: uitbijters

Wat doe je ermee?

- *Codeerfout?* Verbeteren of verwijderen
- Lid van andere populatie? Is het een bijzonder geval? Zo nodig verwijderen, maar beperkt generaliseerbaarheid
- Persoon gedraagt zich slechts in afwijking van hypothese? Altijd erin houden!

Cronbach's alpha

Waarom?: Is niet zo zeer de toets, maar kan de onderlinge betrouwbaarheid van aantal items bekijken wanneer je somscore wilt berekenen. Geldt alleen voor interval vragen.

- Eén algemene maat (gebaseerd op correlaties) voor hoe goed de items voor hetzelfde concept samenhangen

- 0 ≤ α ≤ 1 (precies 0 en 1 wil je niet als resultaat hebben)
- Alleen mogelijk als *alle* vragen op interval/rationiveau!

Alpha < .6 => Betrouwbaarheid niet oke:

- o kunnen er nog items uit? Te smal meten is ook niet goed; haal er niet teveel uit.
- o aparte analyses voor X1, X2, etc (volgend college)

Alpha > .6 => Betrouwbaarheid is oké: somvariabele kan gemaakt worden

(in wetenschappelijke literatuur houden is het .7 voor Alpha)

Hou de tabellen van slide 48 en 49 (tabel 2 in de bijlage) bij het maken van opdracht 2. Het geeft over een overzicht van de testen als voorbereiding op het tentamen.

Regressie: assumpties

1. X en Y interval-variabelen

2. X-en *fixed*, Y *random* getrokken voor elke lineaire combinatie van X-en

Dankzij deze assumptie alleen error in Y, niet in X-en.

Controle: geen. MR robuust tegenover schending

3. Lineaire samenhang tussen X en Y

Controle: plot van residuen (Y- tegen voorspelde waarden (). Vormt idealiter horizontale band rondom Y- = 0 (en geen curvilineair verband)

4. Errors zijn onafhankelijk

- ➢ *Controle:* statistieken voor speciale gevallen (Durbin-Watson statistiek)
- ➢ *Meestal niet nodig:* onafhankelijkheid vooral kwestie van goede onderzoeksopzet en dataverzameling

5. Errors hebben constante variantie voor lineaire combinatie van X-en (homoscedasticiteit)

- ➢ *Controle:* zelfde plot als bij 3. Hoedt u voor "toeters" en diabolo's.

6. Errors zijn normaal verdeeld voor elke lineaire combinatie van X-en

- ➢ *Controle:* voor elke apart problematisch. SPSS geeft wel normaalplotjes voor alle errors gezamenlijk.
- ➢ Meestal controle niet nodig. MR is redelijk robusst t.o.v. afwijkingen van normaliteit

7. Uitbijters ("outliers")

Uitbijter: Persoon met scorepatroon dat sterk afwijkt van de rest in de steekproef.