

# Samenvatting TOE Algemeen

---

VOCUS heeft deze samenvatting te danken aan Marlon Renes.

Het gebruik van deze samenvatting is bedoeld als studeerhulp na het lezen van de verplichte literatuur. Gebruik van deze samenvatting is geheel voor eigen risico.

Soms wordt er verwezen naar bladzijden of tabellen in het originele boek.

Succes met studeren!

## Inhoud

Samenvatting TOE .....	1
Algemeen .....	3
Hoorcollege 1 – 20/04/2020 .....	3
TOE Hoorcollege 2 22-04-2020 .....	9
TOE Hoorcollege 3 30-04-2020 .....	13
TOE Hoorcollege 4.....	21
01-05-2020 .....	21

# TOE

## Algemeen

---

### Hoorcollege 1 – 20/04/2020

#### Surveys

Kom overview 'correlationeel onderzoek'

- ➔ Kwantitatieve data  
(Getallen & geen teksten)

Correlationele data is overal.

Bijv. CBS onderzoek/Algemeen onderzoek

#### Verschillende types data

- Tevredenheid
- Polls (political) → beschrijven
- Data voor de overheid → CBS

Digitale tijdperk zorgt voor exponentiële groei van data.

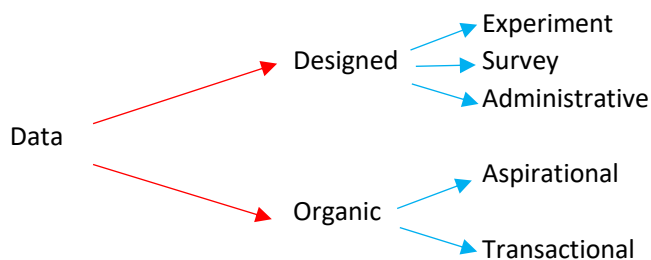
- ➔ Data verzamelen via alle devices

#### Correlationele data

Wordt op 2 manieren geproduceerd:

1. Toevallig (organic)
2. Opzettelijk (designed)

Deze data is ontworpen, onderzoekers zorgen voor deze data.



Organic = Data wordt geproduceerd zonder dat je er als onderzoeker moeite voor hoeft te doen.

- Social media
- Digitale devices

Organic → Aspirational = Mensen doen het eigenlijk opzettelijk, maar niet opzettelijk omdat ze meedoen met een onderzoek. Bijv. Mensen willen met hun tweet iets aan de wereld laten zien. (Jij produceert dan de data)

Organic → Transactional = Wat gebeurt er met je bank/creditcard.

- Bankgegevens
- Energiemaatschappij (energiemeters)

Designed = Opzettelijke manier

Designed → Experiment = experimenten (KOM)

Designed → Survey = survey (KOM)

Designed → Administratieve = Administrative handelingen (belastingdienst)

- Data verzamelen (niet opzettelijk)/mensen dingen gaan vragen.

Vragenlijsten

- Wat gebeurt er als wij een vragenlijst gebruiken om data te verzamelen?
- Hoe kunnen we die data verzamelen?
- Wat voor soort vragenlijsten zijn er?
- Wat zijn de voordelen & nadelen er van?
- Wat voor vertekeningen/fouten spelen een rol bij het gebruiken van zo'n vragenlijst?

Data verzamelen

In onderzoek heb je de onderzoekscyclus

- We hebben een bepaalde populatie.
- Hierover ga je data verzamelen.
- Van die data kun je dingen beschrijven.
- Dan hopelijk conclusie trekken aan de hand van de data die verzameld is in de steekproef.
- Terugkoppelen naar de hele populatie (Inferentie)

Waarom doen we dat?

- ➔ Werkelijkheid beschrijven (beschrijvende statistiek)
- ➔ Relaties kunnen beschrijven (oorzaak-gevolg relaties) (causaliteit)
- ➔ Kunnen generaliseren

Inferentie doelen (generaliseren)

- Beschrijven

Data verzamelen omdat je dingen wilt beschrijven.

VB. Info CBS: de economische positie van de vrouw in NL → beschrijvend (ze kijken niet naar de oorzaak)

- Causaliteit

Wat veroorzaakt nou een verandering in de mening van populatie/mensen.

VB. Vaders die een dochter hebben & wat is de mening van deze mannen over gelijkheid mannen & vrouwen.

- Voorspellen

VB. Voorverkiezingen VS soort van poll over welke mensen gaan kiezen.

Verschillende manieren om vragenlijsten af te nemen & voor/nadelen

1. Face-to-face

Bemoeienis van de interviewer is erg groot.

2. Via de post

Interactie met de respondent is er laag. (Hierdoor kunnen onduidelijkheden niet worden toegelicht)

3. Via de telefoon

Weinig privacy.

4. Via het internet

Veel voordelen, namelijk het kunnen toevoegen van videos of audios. Het is wel nadelig omdat je goed moet nadenken hoe je gaat communiceren met degene die de vraag moet invullen.

5. Mixed mode

<i>Attribute</i>	<b>Face-to-face</b>	<b>Post/mail</b>	<b>Telephone</b>	<b>Internet</b>
<b>Cost</b>	Hoog	Laag	Gemiddeld	Laag
<b>Response rate</b>	Hoog	Laag	Hoog	Gemiddeld
<b>Researcher control over interview</b>	Hoog	Laag	Gemiddeld	Gemiddeld
<b>Interviewer effects</b>	Hoog	Laag	Gemiddeld	Laag

#### Survey modes in NL

- Iedereen staat geregistreerd bij de gemeente
- Meer dan 90% gebruikt internet (minder post/telefoon, meer internet)
- Mix device services (bijv. internet vragenlijsten ook smartphone proef maken)
- Marktonderzoekers (selecte/a-selecte steekproeven)

#### Types of mixed-mode

1. Verschillende methodes bij verschillende groepen respondenten.  
(Bijv. te maken met anonimiteit)
2. Ene methode om mensen voor je te winnen & andere methode om de vragen af te nemen.
3. Ene methode voor dataverzameling & andere voor de reminders.
4. Een methode voor het grootste gedeelte van het interview & eentje voor de 'gevoelige' info.
5. Een methode voor het panel & een andere methode voor de andere keer aan het panel.

#### Panel surveys

(cross-sectional data)

- ➔ Mensen volgen over de tijd heen. (Bijv. elk jaar dezelfde vragenlijst laten invullen,)

Voordelen:

- Veranderingen over de tijd beschrijven.
- 1 keer vragen = niet weten wat het effect is van bepaalde achtergrond factoren.

Nadelen:

- Attrition (uitval)
- Telkens hetzelfde antwoord geven 'omdat je het de vorige keer ook had'

#### Voorbeeld van een vragenlijst

*Wel/geen zomertijd.*

De EU wou beslissing hierover doen.

- ➔ Grote vragenlijst (onder alle mensen die in de EU wonen)
- ➔ 4,6 miljoen antwoorden

Problemen met deze survey:

- Niet representatief
- Niet voor elke burger het goede laten zien.

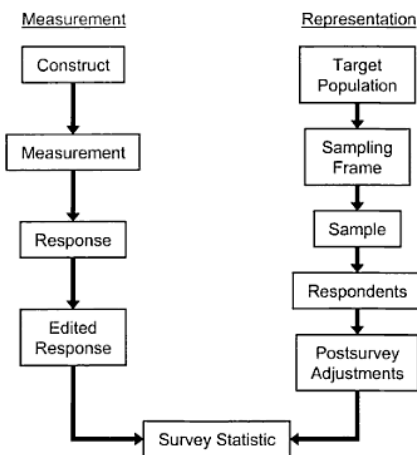
Waarom wordt er dan geld uitgegeven aan zo'n survey?

## Data cyclus

1. Concept/theory
  2. Research questions
  3. Research designs
  4. Hypotheses
  5. Data collection
  6. Data analysis
- ➔ Supporting data streng thens the theory
  - ➔ Non-supporting the data
  - ➔ Revisions of theory and reseersch designs

1. Idee ➔ iets willen weten over de mening van mensen in de EU over zomertijd.
  2. Onderzoeksvraag
  3. Onderzoeksontwerp ➔ Bij wie & hoe gaan we meten?
- ➔ Hier gaat het fout.

## Survey lifecycle



Elk stapje is een bron voor het maken van fouten.

Elke keer dat een fout wordt genegeerd kan er een vertekening komen van de resultaten.

Fouten die gemaakt kunnen worden in deze Cycle

Linker rijtje:

5 ➔ measurement error

6 ➔ proccesing error

Rechter rijtje:

1 ➔ Coverage error

2 ➔ Sampling error

3 ➔ Non respons error

4 ➔ Adjustment error

1. Dekkingsfout
2. Steekproeffout
3. Non response fout
4. (5) Meetfout
5. (6) Procesfout

Dekkingsfout

➤ Doelpopulatie

Steekproefkader → Hieruit trekken we een steekproef

VB:

Doelpopulatie → alle mensen in NL

➔ Lijst met telefoonnummers

(Mensen zonder telefoon staan dus niet op de lijst & dus niet in het steekproefkader)

(Mensen met 2 telefoons hebben 2 keer zoveel kans om in de lijst terecht te komen)

(Ook telefoonnummers van bedrijven, die zijn niet interessant)

➔ Dekkingsfouten

Hoeveel mensen uit je doelpopulatie staan ook daadwerkelijk op de lijst?

Een dekkingsfout speelt pas een rol als:

De mensen die niet op jou lijst staan anders zijn dan de mensen van jou doelpopulatie.

➔ 2 soorten dekkingsfouten

1. Onderschatting

Niet op de lijst, maar wel in de populatie.

2. Overschatting

Wel op de lijst, maar niet in de populatie.

Standaardfout bij gebruik van een kleine steekproef → wordt heel groot.

Standaardfout bij gebruik van een grote steekproef → steeds groter

Wanneer de standaardfout te groot is, kun je verschillen die er evt. wel zijn niet terugvinden

Steekproeffout

'Komt eigenlijk altijd voor'

➤ Maakt niet uit hoe groot je je steekproef ook maakt of hoe goed je je best doet op een steekproefkader/ a-selecte steekproef.

➤ Wat er in de steekproef gebeurd is nooit hetzelfde als wat er in de populatie gebeurd. (altijd iets van verschil)

Wanneer steekproeffout belangrijk?

- Te kleine steekproef
- Geen aselecte steekproef

Foutmarge = een maat om die onzekerheid aan te geven.

Bij een survey = een foutmarge van 3 percentagepunten vaak genoeg. (3% van de waarheid)

(zie afbeeldingen aantekeningen)

Voorbeeld:

'Vragenlijst van 40 minuten'

➤ Veel te lang

➤ Veel te langdradig/uitgebreid

Gevolg:

➤ Mensen vallen uit/beantwoorden de helft niet.

= non-response fout

### Non-response fout

- ➔ Mensen die geselecteerd zijn om de vragenlijst te beantwoorden, niet beantwoorden. Dit wordt een probleem wanneer de mensen die niet antwoorden anders zijn dan de mensen die wel antwoorden.

### 2 soorten non-response

1. Unit-nonresponse  
Mensen weigeren om überhaupt mee te doen.
2. Stem-nonresponse  
Mensen weigeren bepaalde vragen te beantwoorden.

### Redenen:

- Geen interesse meer
- Geen vertrouwen in vertrouwelijkheid
- Gevoelige gegevens

### Non-response bias

Verschil tussen mensen die wel/niet beantwoorden.  
(zie afbeelding aantekeningen)

### Adjustment error

- Speelt een rol wanneer de data die uit een survey komt anders wordt behandeld door de data analisten.
- Data kan verschillen omdat het bijvoorbeeld uit verschillende landen komt.
- Je kunt niet al deze data op 1 hoop gooien.
- Er moet wat gebeuren zodat het gelijk komt.
- Data wordt dus aangepast, maar dit kan elke data analist op een andere manier doen.
- Hierdoor krijg je verschillen → fouten!

Oplossing: Gewogen gemiddelde berekenen.

Wanneer er op een bepaalde vraag geen antwoord was (dit is bijv. fout gegaan) dan doe je imputation.

- ➔ Kijken wat mensen beantwoorden op gerelateerde vragen.

Wanneer data analisten die verschillend doen ontstaan er fouten.

### Meetfout

'Mensen geven het verkeerde antwoord'

- Manier waarop je de vragenlijst doet.
- Face-to-face kan ervoor zorgen dat mensen andere antwoorden gaan geven. (survey-mode effect)

### Vragenlijst EU

1. Dekkingsfout  
Genoeg mensen in Europa die geen internet hebben.
2. Steekproeffout  
Alleen vrijwilligers (a-selecte steekproef kun je generaliseren naar de populatie)
3. Non-response  
Heel hoog, mensen die niet geïnteresseerd waren vulde het niet in.
4. Adjustment error → 80% kwam uit Duitsland (geen gewogen gemiddelde gemeten)



## TOE Hoorcollege 2 22-04-2020

Construct → Conceptual Definition → Operational Defenition → Variabele

Construct

- Duidelijke constructen zoals gewicht/bloeddruk kun je meten.
- Ook constructen zonder duidelijk meetinstrument (bijv. intelligentie)

Conceptual Definition

- Wat bedoeld de onderzoeker met het begrip.

Operationele definitie

- Hoe een wetenschapper een bepaald begrip wil meten.

Variabel

- De uitkomst/waarde van het bepaalde meetinstrument.

Betrouwbaarheid & Validiteit

- Betrouwbaar meetinstrument
  - Metingen herhalen = telkens hetzelfde vinden.
  - Geen variatie.
  - Precisie.

Met een betrouwbaar meetinstrument zijn je metingen consequent.

- Validiteit van het meetinstrument
  - Meetinstrument = valide als het meet wat het beoogt te meten.
  - Hoe accuraat het meetinstrument bij het construct past (correctness of measurement)

(zie afbeelding aantekeningen)

5 verschillende vormen Validiteit

### 1. Begripsvaliditeit

Vijf verschillende manieren bekeken worden:

1. Face/indruks  
Als een expert beoordeeld of het meetinstrument er goed uitziet.
2. Content/inhouds  
Worden alle aspecten van het theoretische begrip wel gemeten.
3. Convergent  
Ander meetinstrument dat hetzelfde theoretische begrip meet. → Moeten scores van die meetinstrumenten nauw samenhangen
4. Divergent  
Het moet niet of minder samenhangen met iets dat iets anders meet.
5. Criterium  
Er hoort samenhang te zijn met een begrip waarvan we weten dat er samenhang is.

Samenhang meten

Hoe kun je bepalen hoe sterk de samenhang/relatie is? (Tussen 2 scores)

➔ Met behulp van de correlatie

- De sterkte
- De richting

Van een lineaire relatie. (op interval/ratio meetniveau)

Correlatie efficient = r

Waarde tussen -1 en 1

Wat betekend een correlatie van 0?

- Geen lineair verband.

Kan zijn dat er wel een verband is, maar dat het er uitziet als een parabool.

Correlatie valt dicht bij 1 of -1

- Er is een sterk verband.

Correlatie = ook voor het meten van betrouwbaarheid.

$r = .876 \rightarrow$  sterk verband, hoge convergente validiteit.

Vormen van Betrouwbaarheid

1. Test-her-test betrouwbaarheid  
Als je 2 keer het meetinstrument gebruikt, wil je dat er 2 keer ongeveer hetzelfde uitkomt.
2. Interne beoordelaars betrouwbaarheid  
Het meetinstrument moet onafhankelijk zijn van degene die het gebruikt.
3. Interne validiteit  
In hoeverre hangen de vragen/stellingen binnen een vragenlijst samen.  
➔ Correlatie binnen de vragen van de vragenlijst.

Voorbeeld:

Vraag  $\rightarrow$  'Hoe zou je kunnen meten wat mensen eten gedurende dag?'

- Alles opschrijven in een eetdagboek
- App (waar bijv. allemaal foto's ingezet kunnen worden)
- Vragenlijst
- Experiment

Survey (vragenlijst eten) problemen

- Het gaat over dingen die je je moet herinneren (over wat je normaal gesproken doet & niet specifiek nu)
- Niet betrouwbaar
- Meet niet wat we willen meten  $\rightarrow$  Geen geschikte vragenlijst.

Dingen die je moet beoordelen vanuit je herinnering

- Percentage dat niet goed de data meer kan vertellen, nadat het een lange tijd geleden is  $\rightarrow$  wordt hoe langer geleden, hoe slechter.
- Als de tijd lang terug gaat is er bij mensen de neiging dat ze te weinig opgeven.

Observaties in een lab problemen

Als mensen weten dat ze geobserveerd worden gaan ze zich anders gedragen.

(mensen gaan veel gezonder eten, niet valide)

& mensen kun je niet 24 uur lang observeren in een lab.

Fotograferen

'Wekelijkse boodschappen fotograferen'

- Eet jij al je boodschappen op?
- Wie eet wat op?  $\rightarrow$  niet valide.

Eetdagboek

Dit werkt voor mensen die het zelf willen, maar wanneer het voor een onderzoek is dan gaat het mis (niet perse omdat mensen het niet willen).

➔ Mensen vergeten het & gaan het uit het geheugen doen (niet betrouwbaar).

App

Probleem:

- Non-response
- Niet iedereen zou met de app om kunnen gaan ➔ dekkingfout.

Goed meten is erg lastig omdat:

- De validiteit.  
Construct validiteit ➔ invloed op external validiteit  
In welke mate kunnen we de resultaten van de steekproef generaliseren van de populatie.  
Hierdoor ontstaan dekkingfouten & non-response fouten, die invloed hebben op de externe validiteit.
- Betrouwbaarheid  
Het schatten van mensen vanuit het geheugen, zorgt voor veel minder betrouwbaarheid.

Metten

'Vragenlijst werk heel vaak als je geïnteresseerd bent in'

1. Feiten  
Kun je vragen aan een specialist met een vragenlijst.
2. Behaviours/gedrag  
Kan via observatie, maar ook via een vragenlijst waarin mensen kunnen invullen hoe ze zich gedragen onder bepaalde omstandigheden.
3. Opinion/mening  
Vraag het iemand aan de hand van een vragenlijst.

Response proces

Een vragenlijst moet zo valide/betrouwbaar mogelijk zijn, zonder allemaal fouten.

Dit kan wanneer er wordt beantwoord in deze 4 stappen:

Comprehension ➔ Retrieval ➔ Judgement ➔ Response

Stap 1: Comprehension

Begrip, begrijpen wat de stelling of de vraag betekend.

Stap 2: Retrieval

Informatie die je daar voor nodig hebben terug vinden (opgraven uit je geheugen).

Stap 3: Judgement

Beoordelen, klopt de info ➔ is het de juiste info & beantwoordt het de vraag.

Stap 4: Response

Beantwoord

Het gaat niet altijd helemaal volgens het stappenplan hierboven.

- Welke week wordt er nou precies bedoeld?
- Kun je je dat wel herinneren?
- Kun je dat nog wel beoordelen?
- Wil je dat antwoord wel geven?

Wat voor invloed heeft dat op de resultaten?

Het antwoord dat er door de meeste mensen wordt gegeven zal lager zijn dan de werkelijkheid.

➔ Het gemiddelde gaat dus omhoog.

Wat gebeurt er met de spreiding?

Voorbeeld:

'Hoeveel minuten douche jij? '

- Sensor
- Als we het vragen. → veel meer spreiding.

Als je het zelf aan geeft is de betrouwbaarheid iets lager & de spreiding ligt dan dus juist hoger.

Als onderzoeker moet je er dus dan rekening mee houden dat die spreiding groter is dan de werkelijkheid.

Questioning wording

1. Hoe stel je de vraag op? Moeilijke woorden? Is de vraag te lang? Is de zinsopbouw te ingewikkeld?  
De vraag moet: kort, duidelijke en eenduidig zijn.
2. Leading questions  
Een vraag/inleiding van een vraag waarmee je iemand in een bepaalde richting trekt.
3. Double-barred questions  
Dubbele vraag, vragen moeten duidelijk zijn/er moet eenduidig gevraagd worden. Vragen naar 1 ding.
4. Telescoping effects  
Hoe verder terug, hoe moeilijker het is om een schatting te maken, hoe meer vertekening er komt.
5. Sensitive question  
Gevoelige onderwerpen zorgen voor vertekening.
6. Volgorde effect  
De volgorde waarop je vragen stelt heeft invloed op de antwoorden die mensen geven.

Bij matrixvragen wordt het response proces overgeslagen.

Optimizing

Alle stappen van het response proces netje doorlopen.

Satisficing

Stappen van het response proces overslaan.

Gedrag van de respondent kan verschillende bias veroorzaken

- Social desirability  
Antwoord geven omdat je denkt dat iemand dat wil horen.
- Acquiescence  
Iemand heeft de neiging om ja te zeggen/het ergens mee eens te zijn.
- Primacy/recency effects  
Heleboel antwoord categorieën zorgt ervoor dat mensen de laatste of eerste antwoordopties kiezen.
- Fence sitting  
De neiging om extreme categorieën te mijden (bijv. kiezen tussen 0 & 10, zijn mensen geneigd om 5/6/7 te kiezen)
- Straight lining/non differentiation  
Alle bolletjes onder elkaar invullen (geen valide meting).

## TOE Hoorcollege 3 30-04-2020

Het meten van een theoretisch begrip

Construct (het begrip) → Conceptual definition (definitie) → Operation definition (hoe ga je dat in het onderzoek meten) → Variabele (hoe maak je de variabele)

Operation definition : Vragenlijst

- ➔ Antwoorden worden gecodeerd
- 1. Not at all
- 2. A little
- 3. Rather much
- 4. Very much

Nu kun je gaan kijken naar:

- Item score  
= de score van 1 persoon op 1 vraag. (Bijv. 1 persoon op de 1<sup>e</sup> vraag 'not at all' dan is de item score 1)

Je wilt uiteindelijk 1 score van alle vragen van die persoon. Dit kun je op verschillende manier doen:

- Manier 1: De som van alle itemscores  
Dit werkt niet wanneer er sprake is van ontbrekende waarden.
- Manier 2: Gemiddelde score nemen
- Manier 3: Gewogen gemiddelde  
Wanneer het ene onderdeel belangrijker is dan het andere onderdeel.

Kijk uit:

Bijv. Hoe hoger de score, hoe meer intelligentie.

- Wij koppelen als mensen hoge score aan veel van het te meten product. Dus je spreekt af dat je dit met de conclusie doet.

Maar: Hoe zit het met de losse vragen?

Bijv. PTSD (mensen kunnen zich niet goed hechten), Vraag: Ik voel me gehecht tot anderen.

- Wanneer mensen hier hoog op scoren komt dit juist niet overeen met PTSD.
- Deze is dus eigenlijk verkeerd om gesteld.
- ➔ Je moet deze vraag ompolen/recoderen.

Het wordt dan dus:

1. Very much
2. Rather much
3. A little
4. Not at all

Wanneer je dit omgedraaid hebt kun je beginnen met het maken van een schaalscore.

Berekenen schaalscore

- Gemiddelde score per persoon  
Over al deze scores kunnen dan beschrijvende statistieken berekend worden.

Hoe weten we of een bepaald meetinstrument een goed meetinstrument is?

1. Begripsvaliditeit
2. Betrouwbaarheid

Dit zijn 2 verschillende metingen & de correlatie tussen deze 2 = de convergente validiteit.

Je hebt 3 soorten betrouwbaarheid

1. Test-her-test
2. Interbeoordelaars
3. Interne

Interne betrouwbaarheid

Dit wordt gemeten met Cronbach's Alpha.

- Die meet hoever de items binnen een vragenlijst met elkaar correleren.

Wat meet je de Cronbach's Alpha

- Je kunt een vragenlijst in twee helften verdelen.
- Je kunt dan de scores bepalen gebaseerd op de ene helft & gebaseerd op de andere helft.
- Hierna ga je kijken of die twee helften met elkaar correleren.

Vraag die hierbij komt kijken:

Hoe verdeel je de helften?

- Even/oneven
- Random vraag
- Boven/onderste

Hoe bepaal je de helften?

- Cronbach's alfa kijkt naar het gemiddelde van hoe je al die manieren van in 2 helften splisten gebruikt.
- Hiervoor kun je de computer gebruiken

Belangrijk

Wanneer je een betrouwbaarheidsanalyse gaat uitvoeren op de computer moet je eerst:

- Alle vragen in dezelfde richting hebben staan (ompolen/hercoderen)

Hoe interpreteer je een Cronbach alpha?

Dit hangt af van waarvoor je de vragenlijst gebruikt.

- Ga je een groep ondervragen?
- Gaat het over een individu?
- Ga je beslissingen meenemen, om te bepalen of iemand bijv. bepaalde behandeling moet krijgen.

Globaal

< 0,7 = geen goede betrouwbaarheid

> 0,8 = goede betrouwbaarheid

Tabel om de betrouwbaarheid van de schaal te verbeteren

- Het kan bijv. zijn dat 1 vraag alles een beetje in de war schopt & dan is het beter om die vraag aan te passen.

In de tabel staat er dan een kolom, in deze kolom staat wat Cronbach's alpha zijn, door de vraag te verwijderen kun je zien wat er veranderd aan de Cronbach's alpha.

Er staat ook een kolom met de item-rest correlatie:

Ook wel de rit-waarde.

- ➔ Het is de correlatie tussen dat ene item (stelling/vraag) en de schaal gemaakt zonder dat item.

Waarvoor is deze tabel? (hierboven ook genoemd)

Final version (after removal of Q2)

Reliability Statistics		Item-Total Statistics				
Cronbach's Alpha	N of Items	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted	
.803	9					
		Q1	9.24	4.896	.578	.772
		Q3	9.41	5.308	.457	.789
		Q4	9.39	4.815	.678	.762
		Q5	9.45	5.399	.564	.783
		Q7	9.35	4.655	.861	.781
		Q8	9.35	4.612	.644	.782
		Q9	9.44	5.482	.416	.794
		Q11	9.40	5.498	.324	.803
		Q12_R	9.26	4.792	.340	.803

35

- Schaal zo betrouwbaar mogelijk maken.

Stap 1: Kijken naar de ritwaarde/correlatie tussen een bepaald item & de rest van de schaal.

- ➔ Is die correlatie onder de 0.2 dan zou dat item mogelijk verwijderd kunnen worden.
- ➔ Dat item & de rest van de schaal hangen dan dus niet samen

Daarnaast kijken:

Als je het item verwijderd, wat gebeurt er dan met Cronbach's alfpa?

- Als deze goed omhoog gaat, dan moeten we dat item verwijderen.
- Kanttekening ➔ Je kunt niet eeuwig door items verwijderen. Dan blijft er niks meer over. Dus max. 1 of 2 items verwijderen.

Deel 2

Voorspellen

Kunnen we een bepaalde variabele voorspellen?

& wat voor soort relatie verwacht je tussen die twee variabelen?

Correlatiecoëfficiënt

Hier kun je een relatie aan/mee beoordelen.

- ➔ De p-waarde komt hier ook bij kijken. Wanneer je de 0-hypothese kan verwerpen is de relatie significant.

Regressie

Correlatie is voor de sterkte en de richting van de relatie.

Regressieanalyse is:

Beschrijven van de lijn die door de puntenwolk gaat (met behulp van een vergelijking).

Deze vergelijking kun je dan gebruiken om voorspellingen mee te doen.

- Zo kun je bijvoorbeeld gaan voorspellen of een nieuw persoon PTSD heeft.

## Onderzoeksvraag

- Als er in de vraag iets staat over voorspellen → leidt altijd in de richting van een regressieanalyse.  
Bijv. Kunnen we de PTSD schaal gebruiken om de productiviteit op werk te voorspellen?
- Als het gaat over 'kunnen we bekijken of er een relatie is & is die relatie significant?' → correlatie.

## Variabelen

Met een regressieanalyse moet je een onderscheid gaan maken tussen de variabelen.

(bij correlatie kijk je puur of er een relatie is)

Je hebt te maken met:

1. Onafhankelijke → Letter X  
Gaan we gebruiken om voorspellingen mee te doen.
2. Afhankelijke variabele → Letter Y  
Dat is de variabele die we graag willen voorspellen.

Vb.

'Kun je de productiviteit op het werk voorspellen met behulp van de PTSD scores?'

Y = de productiviteit

X = PTSD score

## Regressielijn

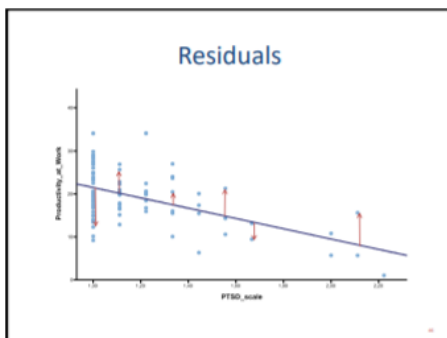
Wordt bepaald met: Least Squares Regression → Kijken naar iets kleins in het kwadraat.

Hoe kunnen we meten welke lijn door de puntenwolk het beste is?

- Aan de hand van de residuen.

Wat zijn residuen?

- Als we een lijn hebben kunnen we de afstand in de richting van de afhankelijke variabele (staat op de Y-as) (dit geval dus de productiviteit)
- De afstand tussen de lijn & de waarden = de residu.



46

Voor elk puntje kun je een residu uitrekenen

'Stel je hebt een lijn die niet bij al deze punten past.'

- Punten liggen niet dicht bij de lijn → grote residuen.
- Waardes die onder de lijn liggen → negatieve residuen
- Waardes die boven de lijn liggen → positieve residuen.

Als je positieve en negatieve bij elkaar op gaat tellen, valt het tegen elkaar weg.

Oplossing: kwadrateren.

- Alle residuen kwadrateren.



Coefficients <sup>a</sup>					
Model		Unstandardized Coefficients		Standardized	Sig.
		B	Std. Error	Coefficients	
1	(Constant)	33,478	2,372		,000
	PTSD score	-11,986	1,972	-.533	,000

a. Dependent Variable: Productivity\_at\_Work

What is the regression equation?

Wanneer je lijn niet goed bij de puntenwolk past

- ➔ Heb je grote residuen.
- ➔ Wanneer je die gaat kwadrateren zijn ze nog groter.
- ➔ Als je ze dan bij elkaar gaat optellen = het een hele grote som.
- ➔ Geeft dus duidelijk aan dat de lijn niet bij de puntenwolk past.

Lijn die wel goed past:

- Grote & kleine residuen
- Veel kleinere som

Best passende lijn:

Opzoek gaan naar de lijn met de kleinst gekwadrateerde som van de residuen.

Wat weten we over de residuen?

1. Puntenwolk met weinig spreiding ➔ kleine residuen.
2. Deze lijn gebruiken zorgt voor nauwkeurige voorspellingen.
3. Residuen geven aan hoeveel we er naast zitten.
4. Puntenwolk met veel spreiding ➔ grotere residuen.
5. Minder nauwkeurige voorspellingen.

Standaardschattingsfout

'Nauwkeurigheid meten'

Standaarddeviatie van de residuen.

- De gemiddelde fout die we maken als de vergelijking gebruiken om voorspellingen mee te doen.

Regressie vergelijking

'De som of squared residuals uitrekenen'

- Voorspellingen
- $\hat{Y} = b_0 + b_1 x$   
 $b_0$  ➔ snijpunt met de Y-as  
 $b_1$  ➔ richtingscoëfficiënt (hoe scheef de lijn loopt)

Regressieanalyse doe je altijd met de computer

Output (regressieanalyse) bestaat uit 4 tabellen:

Tabel 1: Variables

Met welke variabele zijn we bezig.

Tabel 2: Coëfficiënten

Coefficients <sup>a</sup>					
Model		Unstandardized Coefficients		Standardized	Sig.
		B	Std. Error	Coefficients	
1	(Constant)	33,478	2,372		,000
	PTSD score	-11,986	1,972	-.533	,000

a. Dependent Variable: Productivity\_at\_Work

What is the regression equation?

In deze tabel staat:

Onderaan → afhankelijke variabele

In de tabel → onafhankelijke variabele

- Kolom 1:  
Hierin staan de richtingscoëfficiënt & het snijpunt met de Y-as.  
1<sup>e</sup> = B 0 (snijpunt)  
2<sup>e</sup> = B 1 (richtingscoëfficiënt)

Hiermee kun je een regressievergelijking opstellen.

( $\hat{y}$  → betekend schatting)

Voorbeeld:

B 0 = 33.479

B 1 = -11.995

$\hat{Y} = 33.5 - 12 X$

→ Hoe groter de PTSD score, hoe kleiner de productiviteitsscore.

Hoe doe je nou een voorspelling?

De waarde van X stop je in de vergelijking.

Voorbeeld:

PTSD score van John → 1,6

Voorspelling van de productiviteitsscore is dan:

$\hat{Y} = 33.5 - 12 \times 1.6 = 14.3$

Tabel 3:

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.533*	.285	.277	5.2938

\* Predictors: (Constant), PTSD\_scale

Samenvatting Model

- Kolom 1: correlatie  
Tussen  $y$  &  $\hat{y}$  = altijd positief
- Kolom 4: Standaards schattingsfout  
(gem. waarde van de residuen)  
Voorbeeld: 5.2938  
Als we deze vergelijking gaan gebruiken, om de productiviteit op het werk te voorspellen, dan zitten we gem. 5,3 punten er naast.
- Kolom 3: Kwadraat van de correlatie (R<sup>2</sup>)  
Het is een maat om aan te geven hoe goed de lijn bij de puntenwolk past.  
→ Ook wel verklaarde variantie.
- Kolom 2: R<sup>2</sup> als percentage  
Hoeveel van die spreiding in de afhankelijk variabele wordt verklaard door de lineaire relatie met de onafhankelijke variabele.  
→ Welke percentage van de spreiding wordt verklaard door dit model (zijn ook andere bronnen die invloed hebben op de productiviteit)  
→ 0,285 = 28,5 % van de spreiding in de productiviteitsscores wordt verklaard door PTSD.  
→ Meer dan 70% door andere bronnen, die niet in het model voorkomen.

Ten slotte: Significantie

We willen toetsen of dit model wel significant is.

Voorbeeld: 'Helpt PTSD wel om de productiviteit te voorspellen?'

- Productiviteit voorspellen zonder PTSD

Als we info hebben over andere variabele (PTSD) kunnen we dan beter voorspellen?

Is dit model wel significant?

Kan op 2 manieren:

1. Toets uitvoeren voor de richtingscoëfficiënt.
2. Toets uitvoeren voor de  $R^2$

Hypothese toetsen.

Stap 1: 2 hypothesen opstellen

- 1  $H_0$  (nulhypothese) → Geen effect/relatie/verschil
  - 1  $H_a$  (andere hypothese) → Wel relatie/effect (dat wat de onderzoeker graag wil aantonen)
- De onderzoeker kiest ook het significantieniveau → waarde van alpha → welke kans wil ke toestaan dat je type 1 fout maakt ( $H_0$  hypothese verwerpen terwijl die eigenlijk waar is) → kleinere kans = meer zekerheid over verwerpen.

Stap 2: Data verzamelen

Stap 3:

Er worden berekeningen gedaan & daar komt een p-waarde uitrollen (de kans dat je data vind die net zo extreem of nog extremer dan de data die je geobserveerd hebt.)

Voorbeeld: P-waarde

Correlatie vinden van 0.3 → wat is de kans dat we dezelfde of nog extremere waarden gaan vinden als we het herhalen (als we de  $H_0$ -hypothese waar zal zijn).

Dus als er echt geen relatie is, wat is dan de kans dat je toch een correlatie vind van 0.3 of hoger.

Stap 4:

Met de p-waarde kun je een keuze maken:

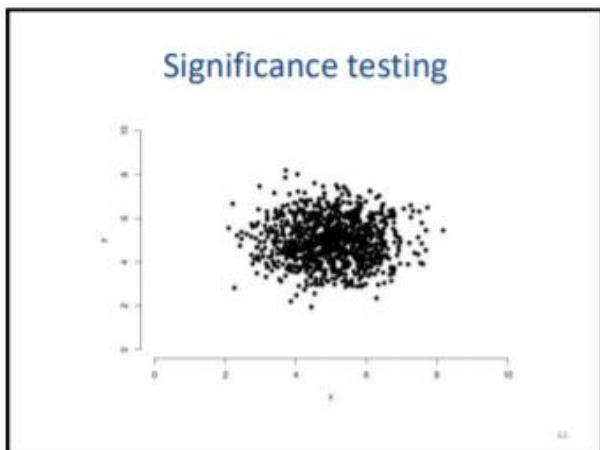
- We verwerpen de  $H_0$  hypothese
- We verwerpen de  $H_0$  hypothese niet.

Wanneer verwerp je de  $H_0$  hypothese?

- De P-waarde moet kleiner zijn dan Alpha (vaak 0.5) of gelijk aan.

Significantie testing (model)

Plaatje als de  $H_0$  hypothese waar zal zijn:



62

Geen relatie tussen de 2 variabelen. Lijn is dan horizontaal

$RC = 0$

- Voor het toetsen kun je dus kijken of de RC gelijk is aan 0
- 2<sup>e</sup> optie = de variantie verklaren. Kun je met de X variabele beter voorspellen, dan zonder de X variabele?
- Spreiding rond regressielijn is ong. zelfde als spreiding rond het gemiddelde van Y (verklaard 0,0 van de spreiding van Y variabele).

Ook gaan toetsen of  $R^2$  gelijk = aan 0

De 2 manieren van toetsen doen eigenlijk hetzelfde

Manier 1: Richtingscoëfficiënt

- Als we een hypothese toetsen dan toetsen we altijd over een populatie parameter.
- Het geen dat we in de steekproef vinden geven we aan met de letter B 1 (RC)

Voorbeeld: Is het gemiddelde in de populatie 0?

→ Is hetzelfde als een t-toets.

T-toets kun je gebruiken om te toetsen of de RC significant verschilt van 0.

(Wordt automatisch gedaan op SPSS)

(staat in tabel 3)

Tabel 3 → laatste kolom

- T-score & P-waarde
- T tussen -3 & 3, als de hypothese waar is.  
Lage/kleine P-waarde.  
En significantie niveau van .05 → 5%

Kun je de 0-hypothese verwerpen?

Ja dit kan want de P-waarde is enorm klein.

- Conclusie:  
RC → verschilt significant van 0, dus de voorspellingen (met PTSD als onafhankelijke variabelen) zijn significant beter dan de voorspellingen die we zullen doen zonder PTSD.

Manier 2: P 2 gekwadraterde waarde

- Standaard in SPSS
- Geen T-toets, maar een F-toets.

Hiervoor kijken in Tabel 4.

Kolom F

Als de 0-hypothese waar is dan is F ongeveer 1.

(0-hypothese niet waar is de F-waarde al snel heel groot)

→ P-waarde ook heel klein.

Dus de 0-hypothese kan verworpen worden.

- Conclusie:  
Dit model verklaard een significant deel van de spreiding in de productiviteitscores.

## TOE Hoorcollege 4

01-05-2020

$$\hat{Y} = 33,5 - 12x$$

Voor elke heel punt dat de PTSD score omhoog gaat, gaat de voorspelde productiviteit met 12 omhoog.

- Naar kolom 1 hebben we de vorige keer niet gekeken: standardized.

Standardized

= Gestandaardiseerde richtingscoëfficiënt/regressie coëfficiënt

→ Wordt ook wel Bèta genoemd.

De waarden van Bèta in een enkelvoudige regressiemode = gelijk aan de correlatie (maar wordt heel anders geïnterpreteerd).

→ Als we variabele standaardiseren halen we de eenheden weg.

Productiviteitsscores: 1-40

PTSD scores: 1-4

Dit zijn verschillende schalen

Eenheidsory maken

Je centreert een variabele door het gemiddelde er af te halen & de eenheden weg te halen door het te delen door de standaardafwijking.

De nieuwe gestandaardiseerde variabele

Hoeveel standaardafwijkingen valt een observatie boven/onder het gemiddelde.

→ Wanneer we met die variabele gaan werken, krijgen we als RC precies de correlatie.

Conclusie:

De gestandaardiseerde regressiecoëfficiënt → hoeveel standaardafwijkingen gaat de afhankelijke variabele omhoog/naar beneden, als de onafhankelijke variabele met 1 standaarddeviatie groter wordt.

Voorbeeld:

Als de PTSD score met 1 standaardafwijking toe neemt.

Gaat de voorspelde productiviteitsscore met 0.5 naar beneden (kun je zien in tabel → standardized)

→ Zo interpreteer je de gestandaardiseerde regressiecoëfficiënt.

Waarvoor meet je dit?

1. Als je variabele hebt die op verschillende schalen zijn gemeten.  
(Soms extreem grote getallen & kun je die hiermee veel nauwkeuriger zien)  
(Dit wordt meestal ook gerapporteerd in artikelen)
2. Komt verderop.

Wanneer mogen we een lineaire regressie gebruiken?

- Om model te maken.
- Om voorspellingen mee te doen.

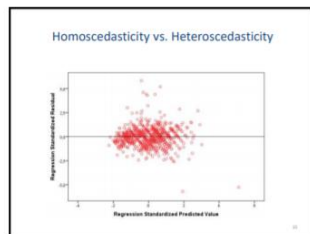
Voorwaarden waar de data aan moet voldoen voordat zo'n lineaire regressie model valide is:

1. Er moet een lineaire relatie zijn.  
(Niet lineair = geen enkelvoudige regressie)
2. Het meetniveau van beide variabele moet van interval/ratio meetniveau zijn.

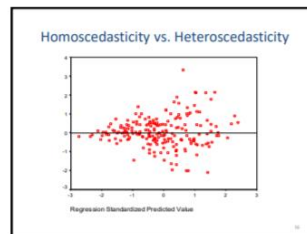
3. Residuen moeten normaal verdeel zijn.  
(Ten opzichte van die regressielijn)  
(Meeste waarden liggen dicht bij het gemiddelde)  
(De meeste observaties zijn redelijk klein/dicht bij de regressierechten)  
(Paar mogen wat verder weg liggen)
4. Homoscedasticity  
(Er vanuit gaan dat de residuen rond de regressierechten qua afstand gem. even groot zijn, zie afbeelding aantekening) '  
(= Gelijke verdeling van de residuen)
5. Geen uitschieters.

#### Controleren

1. Aan de hand van de spreidingsdiagram.
2. PTSD → schaalscore  
Productiviteit → schaalscore  
→ Schaalscore = interval meetniveau (score 2 = niet 2 keer zoveel als score 4)
3. Zijn ze normaal verdeeld? In SPSS kun je aanvinken dat je een histogram van de residuen wil.  
(Hierin kun je controleren of de klokvorm wordt gevolgd)
4. De spreiding is ongeveer even groot.
  - Grafiek maken van de residuen.
  - De grafiek met residuen op de Y-as & op de X-as kun je 2 dingen doen:
    1. X-variabele  
Ongeacht waar we langs de X variabele zitten, willen we dat de spreiding gelijk is.
    2. Voorspelde waarden  
De waarden van  $\hat{y}$ , is de spreiding rond de X-as & de lijn van 0 ongeveer gelijk?



15



16

Kleine standaardfout bij het rechter plaatje.

5. Normaal verdeling  
(waarden tussen 3 & -3)  
Valt het hierboven/onder dan heb je iets heel extreems te pakken.

#### Deel 2

Er is een soort weegschaal hoeveel variabele we willen toevoegen & de nauwkeurigheid ervan.

→ Je krijgt een vergelijking met meerdere onafhankelijke variabele.

$\hat{Y} = B_0 + B_1 X_1 + B_2 X_2 + \dots + B_K X_K \rightarrow K =$  hoeveel variabele we nog willen toevoegen.

→ Het wordt natuurlijk veel ingewikkelder zo, maar achter de schermen wordt nog steeds gezocht naar de vergelijking waardoor de som van de gekwadrateerde residuen het allerkleinst wordt.

→ Dus: de lijn die het dichtst bij alle punten in de puntenwolk ligt.

Hoe meer variabele je toevoegt, hoe meer spreiding er gaat komen in het model.

Hoe meer variabele in het model, hoe kleiner de standaardsschatingsfout. (Dus je kunt nauwkeuriger schatten/voorspellen)

Meer variabel toevoegen wil niet zeggen dat het beter is

- ➔ Alle onafhankelijke variabelen hoeven helemaal niet significant te zijn.
- ➔ Verschil met kleine model hoeft niet significant te zijn.

Meer spreiding → R<sup>2</sup> groter

Hoe hoger, hoe beter het model werkt.

Als de standaardsschatingsfout steeds kleiner wordt (SE → gem. fout die we maken/standaardafwijking van de residuen die zal steeds kleiner worden)

- Of het model echt beter wordt, kun je toetsen met de F-toets
- Je gaat dan toetsen of het model significant is.
- & of de onafhankelijke variabele die we toevoegen, kun je toetsen met de T-toets.

We hebben weer 4 tabellen

Tabel 1: Model Summary

- Onderaan: welke onafhankelijke variabele zijn er.
- VB. PTSD, Hoeveel ze werken, Jobsatisfaction.

☞

Onderste → enkelvoudige regressie → R square 0.285

Bovenste → meervoudige regressie → R square 0.497

= dus bijna 50% toegenomen.

R<sup>2</sup> = dus groter geworden, maar R (de fout die je maakt bij voorspellingen) is kleiner geworden.

Tabel 2: Coëfficiënts

### Coefficients

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized	t	Sig.
		B	Std. Error	Coefficients		
1	(Constant)	35,184	2,054		17,131	,000
	PTSD_scale	-7,910	1,860	-,352	-4,265	,000
	Job_Satisfact	-,178	,103	-,133	-1,734	,089
	Back_at_Work	-,121	,020	-,594	-6,184	,000

<sup>a</sup> Dependent Variable: Productivity\_at\_Work

$$\widehat{productivity} = 33.18 - 7.91PTSD - 0.18Satis - 0.12Work$$

Wanneer je met multipele regressie werkt → hebben verschillende variabele vaak verschillende eenheden.

→ Standardized Coëfficiënts (Beta)

Kun je zien welke de meeste invloed heeft op de regressie?

De grootste waarde van Beta, heeft de meeste invloed op de voorspellingen.

→ Back at work in dit geval dus.

Significantie

1. Toets voor de verklaarde variantie.

Is het deel van de spreiding dat verklaard wordt door het model, wel een significant deel van de spreiding?

→ Toets voor R<sup>2</sup>

2. We hebben een t-toets voor de RC

Deze gaan maar over 1 variabele tegelijk. (Bij enkelvoudige regressie toetsen t-toets & f-toets een beetje hetzelfde).

Bij multipele regressie is dit dus anders.

Je kunt wel variabele blijven toevoegen (R<sup>2</sup> stijgt), maar is dit wel significant?

Nauwkeurig, maar wel een simpel model = het beste.

Tabel 3: Anova

**Output**

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1809,138	3	603,046	29,924	,000 <sup>b</sup>
	Residual	1833,893	91	20,153		
	Total	3643,031	94			

<sup>a</sup>. Dependent Variable: Productivity\_at\_Work  
<sup>b</sup>. Predictors: (Constant), Back\_at\_Work, Job\_Satisfact, PTSD\_scale

- Using  $\alpha = .05$ , do we reject the null hypothesis?
- What is the conclusion?

30

0-hypothese kan verworpen worden want P = kleiner dan .05

Conclusie:

Het model (met 3 onafhankelijke variabele) verklaarde een significant deel van de spreiding.

→ In de productiviteitsscores.

T-toets

Voor elke onafhankelijke variabele die in dat model zit = er een T-toets.

Bij elke kun je dus een t-toets uitvoeren, waarmee je kan zien of die bepaalde regressiecoëfficiënt gelijk is aan 0. (In de aanwezigheid van andere variabele)



De T-toets kan dus maar uitgevoerd worden voor een variabele tegelijk.

→ Als er 1 variabele niet significant is, zou je die uit het model kunnen halen.

De voorspellingen die je maakt met die variabele in het model & zonder de variabele in het model.

→ Zit dan geen significant verschil tussen.

Tabel 4: Coëfficiënten a

**Output**

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	35,184	2,054		17,131	,000
	PTSD_score	-7,910	1,800	-,352	-4,385	,000
	Job_Sat/act	-,178	,103	-,133	-1,734	,086
	Back_at_Work	-,121	,520	-,504	-6,184	,000

<sup>a</sup>. Dependent Variable: Productivity\_at\_Work

- Using  $\alpha = .05$ , which predictor is *not* significant in this model?
- What is the conclusion?

Je kunt in de laatste kolom zien dat Job Sat. niet significant is.

Dat wil niet zeggen dat er geen relatie is, maar die relatie in aanwezigheid met andere variabelen = niet significant.

Conclusie:

Voorspellingen voor productiviteit gemaakt met Job Sat. & Voorspellingen gemaakt met het model zonder Job. Sat. = de verschillen zijn niet significant

Meer variabelen → hoe nauwkeuriger (  $R^2$  stijgt & Standaardfout daalt)

Maar er is dus geen significant verschil tussen die 2 modellen.

