MATERIALEN BIJ STATISTIEK 2 (199212) JANUARI 2010

Sheets hoorcollege 1 (over par. 7.1): Gevolgtrekkingen over μ als σ onbekend is

- De kansverdeling van $T = \dfrac{\overline{X} - \mu}{s/\sqrt{n}}$
- Toetsen van hypothesen over μ en betrouwbaarheidsintervallen voor μ.
- Het vergelijken van twee behandelingen met gekoppelde paren met t-toets
- Het vergelijken van twee behandelingen met gekoppelde paren met de tekentoets

---

De kansverdeling van $t = \dfrac{\overline{X} - \mu}{s/\sqrt{n}}$

- Veronderstelling: EAS uit populatieverdeling
- Als n groot is, dan heeft $T = \dfrac{\overline{X} - \mu}{s/\sqrt{n}}$ (vrijwel) dezelfde verdeling als

  $Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$: de standaardnormale verdeling.

- Kleine n en populatieverdeling normaal: $T = \dfrac{\overline{X} - \mu}{s/\sqrt{n}}$ heeft t-verdeling met n-1 'vrijheidsgraden'

Opgave 1

a) Bepaal het getal t^* zodanig dat rechts ervan een oppervlak 0.05 ligt onder de curve van t-verdeling met 9 vrijheidsgraden.

b) Bepaal het getal t^* zodanig dat links ervan een oppervlak 0.01 ligt onder de curve van t-verdeling met 19 vrijheidsgraden.

c) Bepaal het getal t^* zodanig dat geldt: $P\left(\left|\dfrac{\overline{X} - \mu}{s/\sqrt{n}}\right| \le t^*\right) = 0.95$ als n=30.

d) Bepaal het getal t^* zodanig dat geldt: $P\left(\left|\frac{\bar{X}-\mu}{s/\sqrt{n}}\right| \leq t^*\right) = 0.95$ als n=50.

Antwoord opgave 1

a) 1.833

b) -2.539

c) 2.045

d) ongeveer 2.009

Toetsen van hypothesen over μ en betrouwbaarheidsintervallen voor μ.

	σ bekend	σ onbekend
BI voor μ	$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$	$\bar{x} \pm t^* \frac{s}{\sqrt{n}}$
Toetsgrooth voor $H_0 : \mu = \mu_0$	$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$

Voor het bewijs van de formule voor het betrouwbaarheidsinterval: zie uitgewerkte opgave 1b!

Opgave 2

In EAS van 20 personen: $\bar{x} = 106.10$, s=16.00.

a) Stel 95% BI op voor μ.

b) We toetsen $H_0 : \mu = 100$ tegen $H_a : \mu \neq 100$. Bepaal de overschrijdingskans.

c) We toetsen $H_0 : \mu \leq 100$ tegen $H_a : \mu > 100$. Bepaal de overschrijdingskans.

d) We toetsen $H_0 : \mu \geq 100$ tegen $H_a : \mu < 100$. Bepaal de overschrijdingskans.

Antwoord opgave 2

a) Interval loopt van
106.10 - 2.093(16/$\sqrt{20}$) =
106.10 - 7.49 = 98.61
tot 106.10 + 7.49 = 113.59

b) Bij elk van de volgende onderdelen is de toetsingsgrootheid: $T = \frac{\bar{X} - 100}{s/\sqrt{20}}$. De gevonden waarde ervan: $t = \frac{106.10 - 100}{16/\sqrt{20}} = 1.705$.

Bij b): P = P($T \geq 1.705$) + P($T \leq -1.705$) ligt tussen 2*0.05=0.10 en 2*0.10=0.20.

c) P($T \geq 1.705$) ligt tussen 0.05 en 0.10.

d) P($T \leq 1.705$) ligt tussen 0.90 en 0.95.

Toetsen van hypothesen over μ en betrouwbaarheidsintervallen voor μ: SPSS

Opgave

In het bestand 'college1_PDI_1.sav' staan de scores op Psychomotor Development Index van 64 aselect gekozen te vroeg geboren baby's. Stel een 95% betrouwbaarheidsinterval op voor het populatiegemiddelde.

Antwoord:

(*Analyze->Compare Means->One-sample T-Test; vul in 'Test Value = 0' als je een interval voor mu wilt. Tweede manier om betrouwbaarheidsinterval te krijgen: Analyze->Descritive Statistics->Explore; zet PDI in 'Dependent List'; de betrouwbaarheidscoëfficiënt kun je aanpassen in 'Statistics'*)

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
PDI	64	95.1719	13.01715	1.62714

One-Sample Test

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
PDI	58.490	63	.000	95.17188	91.9203	98.4235

Toetsen van hypothesen over μ en betrouwbaarheidsintervallen voor μ: SPSS

In de populatie van baby's die op het juiste moment worden geboren is de gemiddelde PDI-score gelijk aan 100. Gebruik de data uit 'college1_PDI_1.sav' om te toetsen of het gemiddelde in de populatie van de te vroeg geboren baby's daarvan afwijkt. (Geef de overschrijdingskans)

Antwoord:
(*Analyze->Compare Means->One-sample T-Test; vul in 'Test Value = 100'*)

P=0.004

One-Sample Test

	Test Value = 100					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
PDI	-2.967	63	.004	-4.82812	-8.0797	-1.5765

Toetsen van hypothesen over μ en betrouwbaarheidsintervallen voor μ: SPSS

Stel dat tevoren werd verwacht dat in de populatie van te vroeg geboren baby's de gemiddelde PDI-score lager is dan 100. Dan is het relevant om $H_0 : \mu 1{=}0\ \ 0$ (of: $H_0 : \mu 1 \geq 0\ \ 0$) te toetsen tegen $H_a : \mu 1 < 0\ \ 0$.

a. Voer de toetsing uit met de data van 'college1_PDI_1.sav' . (Geef de overschrijdingskans.)

b. Idem voor de data uit 'college1_PDI_2.sav'.

c. Idem voor de data uit 'college1_PDI_3.sav'.

Toetsen van hypothesen over μ en betrouwbaarheidsintervallen voor μ: SPSS

Antwoorden:
a. *P=0.002*

One-Sample Test

	Test Value = 100					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
PDI	-2.967	63	.004	-4.82812	-8.0797	-1.5765

b. *P*=0.31

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
PDI	64	99.1719	13.01715	1.62714

One-Sample Test

	Test Value = 100					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
PDI	-.509	63	.613	-.82812	-4.0797	2.4235

c. *P*=0.99 (1-0.013/2)

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
PDI	64	104.1719	13.01715	1.62714

One-Sample Test

	Test Value = 100					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
PDI	2.564	63	.013	4.17188	.9203	7.4235

Het vergelijken van twee behandelingen met gekoppelde paren met t-toets

Opgave 3
10 aselect gekozen Middelbare scholieren beoordelen de aantrekkelijkheid van een oude en een nieuwe voorlichtingsfolder in aselecte volgorde:

leer-ling	1	2	3	4	5	6	7	8	9	10
oud	25	29	18	33	26	27	19	35	24	25
nieuw	30	30	22	33	30	31	26	34	29	26

a) Zijn de scores bij nieuwe folder hoger? Toets met $\alpha = 0.05$.
b) *Hoeveel* zijn de scores bij nieuwe folder gemiddeld hoger dan bij oude? Stel een 90% BI op om dit te kunne beantwoorden.

Antwoord opgave 3

a) Gebruik de verschilscores: 5, 1, 4, 0, 4, 4, 7, -1, 5, 1.
$H_0 : \mu = 0,\ H_a : \mu > 0$
$t = 3.67$, de eenzijdige P ligt tussen 0.0025 en 0.005.

b) Het interval loopt van 1.503 tot 4.497.

met SPSS
(*Data in 'hoorcollegePar7_1Opgave3.sav'; Kies: Analyze->Compare Means->Paired-Sample T-Test*)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	OUD	26.1000	10	5.3635	1.6961
	NIEUW	29.1000	10	3.5730	1.1299

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	OUD & NIEUW	10	.910	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	90% Confidence Interval of the Difference Lower	Upper			
Pair 1	OUD - NIEUW	-3.00	2.5820	.8165	-4.4967	-1.5033	-3.67	9	.005

<u>Het vergelijken van twee behandelingen met gekoppelde paren met de tekentoets</u>

<u>Opgave</u>
10 aselect gekozen Middelbare scholieren beoordelen de aantrekkelijkheid van een oude en een nieuwe voorlichtingsfolder in aselecte volgorde:

leer-ling	1	2	3	4	5	6	7	8	9	10
oud	25	29	18	33	26	27	19	35	24	25
nieuw	30	30	22	33	30	31	26	34	29	26

Zijn de scores bij nieuwe folder hoger? Toets met $\alpha = 0.05$. *Neem <u>niet</u> aan dat populatieverdeling van verschilscores normaal is.*

<u>Deel antwoord</u>

- Er is één paar met verschilscore 0; deze laten we weg. Voor de tekentoets geldt hier dus: n=9.

- Als de nulhypothese waar is, dan zijn de verschilscores symm. verdeeld rondom 0. De kans op een pos. verschil p is dan 1/2.

- Onder de nulhyp.: X= "aantal pos. verschilscores is" is
  B(n, 1/2)-verdeeld. Hier: n=9.

- We vinden X=8. De overschrijdingskans: $P = P(X \geq 8) = 0.0176 + 0.0020 = 0.0196$ (zie tabel binomiale verdeling)

Het vergelijken van twee behandelingen met gekoppelde paren met de tekentoets: SPSS

Kies: Analyze-> Nonparametric Tests
➔ 2 Related Samples . Vink als enige toets aan: 'Sign' (Sign-test = tekentoets). SPSS toetst (zoals vaker) tweezijdig!

Frequencies

		N
nieuw - oud	Negative Differences[a]	1
	Positive Differences[b]	8
	Ties[c]	1
	Total	10

a. nieuw < oud

b. nieuw > oud

c. nieuw = oud

Test Statistics[b]

	nieuw - oud
Exact Sig. (2-tailed)	.039[a]

a. Binomial distribution used.

b. Sign Test

Uitgewerkte opgaven week 1

Opgave 1

a. Onder welke veronderstellingen is $\overline{X} \pm t^* \frac{s}{\sqrt{n}}$ een 95% betrouwbaarheidsinterval voor het populatiegemiddelde μ? (Hierbij is t^* het getal waarboven een kans van 0.025 ligt onder de curve van de t-verdeling met n-1 vrijheidsgraden.) Anders geformuleerd: wat is het kansmodel waar we vanuit gaan bij het gebruik van deze formule?

b. Bewijs dat $\overline{X} \pm t^* \frac{s}{\sqrt{n}}$ onder de genoemde veronderstellingen een 95% betrouwbaarheidsinterval voor het populatiegemiddelde μ is. Gebruik daarbij de informatie uit het tweede 'hok' van bladzijde 308.

c. Leg in woorden op statistisch verantwoorde wijze uit wat het betekent dat $\overline{X} \pm t^* \frac{s}{\sqrt{n}}$ een 95% betrouwbaarheidsinterval voor het populatiegemiddelde μ is. Neem in de uitleg aan dat n=25.

Opgave 2

We nemen aan dat de scores op een neuroticismeschaal in de populatie van alle volwassen Nederlanders bij benadering normaal verdeeld zijn. In een EAS van 25 personen vinden we een gemiddelde score van 46.45. Verder is de standaardafwijking van de scores in de steekproef gelijk aan 9.78.

a. Stel een 99% betrouwbaarheidsinterval op voor het populatiegemiddelde.

b. Leg voor een iemand die niets van statistiek weet uit wat het berekende interval betekent. Begin je antwoord als volgt: *"Op basis van de steekproefgegevens kunnen we er redelijk zeker van zijn dat …"*

c. Is het erg als de scores in de populatie niet normaal verdeeld zouden zijn? Motiveer je antwoord.

Opgave 3a

We willen aantonen dat de gemiddelde score op een toets in de populatie van alle leerlingen kleiner is dan 50. We vinden in een EAS van 6 leerlingen vinden we de volgende scores: 42, 41, 36, 49, 51, 45.
Is dit een overtuigend bewijs voor de bewering dat het populatiegemiddelde kleiner is dan 50?
Voer een relevante statistische toetsing uit. Hanteer een significantieniveau van 0.05. Volg daarbij de volgende stappen (bij stap 3 is de keuze al gemaakt!):

1. Formuleer het kansmodel.
2. Formuleer de nulhypothese en de alternatieve hypothese in termen van de parameters van het kansmodel. Kies daarna het significantieniveau (onbetrouwbaarheidsdrempel) α .
3. Formuleer een geschikte toetsgrootheid in termen van de voorkomende stochastische variabelen.
4. Geef de kansverdeling van de toetsgrootheid onder (het randpunt van) H_0.
5. Bereken of geef de waarde van de toetsgrootheid.

6. Bepaal de overschrijdingskans.
7. Formuleer de conclusie over het al dan niet verwerpen van H_0.
8. Vermeld de conclusie ‘in gewone woorden’.

<u>Opgave 3b</u>
Waarom is het in de praktijk af te raden om bij de toetsing in het vorige onderdeel gebruik te maken van een steekproef van slechts 6 personen?

<u>Opgave 4</u>
We willen aantonen dat de gemiddelde score op een toets in de populatie van alle leerlingen kleiner is dan 50. We vinden in een EAS van 30 leerlingen een steekproefgemiddelde van 47. Verder is in deze steekproef de standaardafwijking van de scores gelijk aan 7.11.
Is dit een overtuigend bewijs voor de bewering dat het populatiegemiddelde kleiner is dan 50?
Voer de toetsing uit de ‘aanpak met een overschrijdingskans. Hanteer een significantieniveau van 0.05. Volg daarbij de volgende stappen:

1. Formuleer het kansmodel.
2. Formuleer de nulhypothese en de alternatieve hypothese in termen van de parameters van het kansmodel. Kies daarna het significantieniveau (onbetrouwbaarheidsdrempel) α .
3. Formuleer een geschikte toetsgrootheid in termen van de voorkomende stochastische variabelen.
4. Geef de kansverdeling van de toetsgrootheid onder (het randpunt van) H_0.
5. Bepaal de kritieke waarde(n) en geef het kritieke gebied.
6. Bereken of geef de waarde van de toetsgrootheid.
7. Formuleer de conclusie over het al dan niet verwerpen van H_0 bij de gegeven onbetrouwbaarheidsdrempel.
8. Vermeld de conclusie ‘in gewone woorden’.

<u>Opgave 5</u>
We willen aantonen dat de gemiddelde score in de populatie van alle leerlingen afwijkt van 50. We vinden in een EAS van 30 leerlingen een steekproefgemiddelde van 47. Verder is in deze steekproef de standaardafwijking van de scores gelijk aan 7.11.
Is dit een overtuigend bewijs voor de bewering dat het populatiegemiddelde afwijkt van 50?
Voer de toetsing uit de ‘aanpak met een verwerpingsgebied’ (zie opgave 4). Hanteer een significantieniveau van 0.05.

<u>Opgave 6</u>
We willen aantonen dat de gemiddelde score in de populatie van alle leerlingen afwijkt van 50. We vinden in een EAS van 30 leerlingen een steekproefgemiddelde van 47. Verder is in deze steekproef de standaardafwijking van de scores gelijk aan 7.11.
Is dit een overtuigend bewijs voor de bewering dat het populatiegemiddelde afwijkt van 50?

Voer de toetsing uit de ‘aanpak met een overschrijdingskans. Hanteer een significantieniveau van 0.05.

Opgave 7
We willen aantonen dat het populatiegemiddelde in de populatie kleiner is dan 50. We vinden in een EAS van 30 leerlingen een steekproefgemiddelde van 49 en een steekproefstandaardafwijling van 7.11.
Is dit een overtuigend bewijs voor de bewering dat het populatiegemiddelde kleiner is dan 50?
Voer de toetsing uit de ‘aanpak met een verwerpingsgebied’ (zoals beschreven onder ‘Inferentie als beslissing’). Hanteer een significantieniveau van 0.05. Volg daarbij de volgende stappen (bij stap 3 is de keuze al gemaakt!):

1. Formuleer het kansmodel.
2. Formuleer de nulhypothese en de alternatieve hypothese in termen van de parameters van het kansmodel. Kies daarna het significantieniveau (onbetrouwbaarheidsdrempel) α .
3. Formuleer een geschikte toetsgrootheid in termen van de voorkomende stochastische variabelen.
4. Geef de kansverdeling van de toetsgrootheid onder (het randpunt van) H_0.
5. Bereken of geef de waarde van de toetsgrootheid.
6. Bepaal de overschrijdingskans.
7. Formuleer de conclusie over het al dan niet verwerpen van H_0.
8. Vermeld de conclusie ‘in gewone woorden’.

Opgave 8
(Hypothesen worden geformuleerd voordat de gegevens verzameld worden waarmee de hypothesen getoetst worden. Het is natuurlijk mogelijk dat achteraf blijkt dat de gegevens helemaal niet in overeenstemming zijn met onze vermoedens (zoals weergegeven in de alternatieve hypothese). In deze opgave bekijken we zo’n situatie.)
We willen aantonen dat de gemiddelde score in de populatie van alle leerlingen groter is dan 50. We vinden in een EAS van 30 leerlingen een steekproefgemiddelde van 47. Verder is in deze steekproef de standaardafwijking van de scores gelijk aan 7.11.
Is dit een overtuigend bewijs voor de bewering dat het populatiegemiddelde kleiner is dan 50?
Voer de toetsing uit de ‘aanpak met een overschrijdingskans’. Hanteer een significantieniveau van 0.05.

Opgave 9
De Open Universiteit wil aantonen dat door invoering van vragenuren bij de schriftelijke cursus Statistiek II de scores op de toets stijgen. Er zijn in totaal 800 studenten die Statistiek I gedaan hebben en nu Statistiek II gaan volgen. Men vormt 400 paren van twee personen met (nagenoeg) dezelfde score op de toets van Statistiek I. Hieruit kiest men aselect 8 paren studenten. Uit elk paar wordt aselect één student gekozen die mee gaat doen aan de vragenuren; de andere student uit hetzelfde paar krijgt niet de mogelijkheid

om hieraan deel te nemen. De scores op Statistiek II (=aantal goed van 50 meerkeuzevragen) zijn als volgt **:**

Paar	1	2	3	4	5	6	7	8
Wel vragenuur	35	30	40	25	40	33	23	34
Geen vragenuur	20	25	34	28	35	27	18	45

a. Ga met een geschikte t-toets na of de scores met vragenuur inderdaad hoger uitvallen dan zonder vragenuur. Volg de stappen uit opgave 3.
b. Kan met de tekentoets geconcludeerd worden dat met het vragenuur de scores over het algemeen hoger uitvallen dan zonder het vragenuur? Geef de waarde van de toetsgrootheid en de overschrijdingskans. Formuleer de conclusie.

Antwoorden uitgewerkte opgaven week 1

Opgave 1

a. In ieder geval nemen we aan dat we een EAS uit de populatie trekken. Als we verder aannemen dat de scores in de populatie normaal verdeeld zijn, geeft de formule een exact 95% betrouwbaarheidsinterval (ook bij kleine n!). Als n niet te klein is, geeft de formule bij benadering een 95% betrouwbaarheidsinterval, ook als de scores in de populatie niet normaal verdeeld zijn, omdat de t-procedures bij wat grotere steekproeven vrij robuust zijn.

b.

$$0.95 = P\left(-t^* \le \frac{\overline{X} - \mu}{s/\sqrt{n}} \le t^*\right) = P\left(-t^* \frac{s}{\sqrt{n}} \le \overline{X} - \mu \le t^* \frac{s}{\sqrt{n}}\right) =$$

$$P\left(-\overline{X} - t^* \frac{s}{\sqrt{n}} \le -\mu \le -\overline{X} + t^* \frac{s}{\sqrt{n}}\right) = P\left(\overline{X} - t^* \frac{s}{\sqrt{n}} \le \mu \le \overline{X} + t^* \frac{s}{\sqrt{n}}\right)$$

c. Als we telkens opnieuw een EAS van 25 personen kiezen, dan geeft de formule $\overline{X} \pm t^* \frac{s}{\sqrt{25}}$ telkens een ander interval. Op de hele lange duur (bij heel veel steekproeven), zal ongeveer 95% van die intervallen het (onbekende) populatiegemiddelde μ bevatten.

Opgave 2

a. Het interval loopt van $46.45 - 2.797\frac{9.78}{\sqrt{25}} = 46.45 - 5.47 = 40.98$ tot 46.45+5.47=51.92

b. Op basis van de steekproefgegevens kunnen we er redelijk zeker van zijn dat de gemiddelde score in de populatie ergens tussen 40.98 en 51.92 ligt.

c. De echte kans dat het populatiegemiddelde gevangen wordt door het interval zal niet meer precies 0.99 zijn. Omdat de steekproefomvang redelijk groot is, zal de afwijking van 0.99 vermoedelijk niet al te groot zijn (‘robuustheid’).

Opgave 3a

1. Er is een EAS uit de populatie getrokken. De populatieverdeling van scores is een normale verdeling met onbekend gemiddelde μ en onbekende standaardafwijking σ.
2. $H_0 : \mu \ge 50$, $H_a : \mu < 50$; $\alpha = 0.05$.
3. $T = \dfrac{\overline{X} - 50}{s/\sqrt{6}}$
4. Als $\mu = 50$, dan is T t-verdeeld met df=n-1=5 vrijheidsgraden.
5. $t = \dfrac{44 - 50}{5.5136/\sqrt{6}} = -2.666$.

6. Alleen kleine waarden van de toetsgrootheid pleiten voor H_a. Dus: $P = P(T \leq -2.666)$. Volgens tabel D ligt deze overschrijdingskans in tussen 0.02 en 0.025 (want -2.666 ligt tussen – 2.757 en -2.571).
7. Omdat de overschrijdingskans P kleiner is dan α, wordt de nulhypothese verworpen.
8. Er is met de steekproefgegevens overtuigend aangetoond dat de gemiddelde score in de populatie kleiner is dan 50.

Opmerking:
Het toetsen van hypothesen vertoond overeenkomsten met een rechtzaak. De rechter gaat er vanuit dat de verdachte onschuldig is, totdat de gegevens er duidelijk op wijzen dat hij schuldig is. Net zo nemen wij voorlopig aan dat de nulhypothese waar is. Pas als de toetsgrootheid een waarde aanneemt die wel erg onwaarschijnlijk is als de nulhypothese waar is, verwerpen we die nulhypothese.
Gevolg van de aanpak van de rechter is dat de beslissing om de verdachte te veroordelen een ‘harde beslissing’ is. De rechter neemt deze beslissing namelijk pas als hij er veel bewijsmateriaal voor is. Daarentegen is de beslissing om de verdachte niet te veroordelen een ‘zachte beslissing’. Als de rechter een verdachte niet veroordeeld, betekent dat niet dat de rechter durft te beweren dat de verdachte onschuldig is.
Net zo is de beslissing om H_0 te verwerpen een harde beslissing.
Als H_0 niet verworpen wordt, is er sprake van een zachte beslissing.

Opgave 3b
Omdat het onderscheidend vermogen van de statistische toets dan erg laag zal zijn. (Dit moet je kunnen toelichten!)

Opgave 4
1. Er is een EAS uit de populatie getrokken. De populatieverdeling van scores is een normale verdeling met onbekend gemiddelde μ en onbekende standaardafwijking σ.
2. $H_0: \mu \geq 50$, $H_a: \mu < 50$; $\alpha = 0.05$.
3. $T = \dfrac{\overline{X} - 50}{s/\sqrt{30}}$
4. Als $\mu = 50$, dan is T t-verdeeld met df=n-1=29 vrijheidsgraden.
5. We verwerpen de nulhypothese als $t \leq -1.699$.
6. We vinden: $t = \dfrac{47 - 50}{7.11/\sqrt{30}} = -2.31$
7. We verwerpen H_0 en aanvaarden daarmee H_a. Reden: de gevonden waarde van de toetsgrootheid ligt in het verwerpingsgebied.
8. Er is statistisch aannemelijk gemaakt (aangetoond) dat de gemiddelde score in de populatie kleiner is dan 50.

Opgave 5

1. is een EAS uit de populatie getrokken. De populatieverdeling van scores is een normale verdeling met onbekend gemiddelde μ en onbekende standaardafwijking σ
2. $H_0: \mu = 50$, $H_a: \mu \neq 50$; $\alpha = 0.05$
3. $T = \dfrac{\bar{X} - 50}{s/\sqrt{30}}$
4. Als $\mu = 50$, dan is T t-verdeeld met df=n-1=5 vrijheidsgraden.
5. We verwerpen de nulhypothese als $t \leq -2.045$ en ook als $t \geq 2.045$.
6. We vinden: $t = \dfrac{47 - 50}{7.11/\sqrt{30}} = -2.31$
7. We verwerpen H_0 en aanvaarden daarmee H_a. Reden: de gevonden waarde van de toetsgrootheid ligt in het verwerpingsgebied.
8. Er is statistisch aannemelijk gemaakt (aangetoond) dat de gemiddelde score in de populatie afwijkt van 50.

Opgave 6

1. is een EAS uit de populatie getrokken. De populatieverdeling van scores is een normale verdeling met onbekend gemiddelde μ en onbekende standaardafwijking σ
2. $H_0: \mu = 50$, $H_a: \mu \neq 50$; $\alpha = 0.05$
3. $T = \dfrac{\bar{X} - 50}{s/\sqrt{30}}$
4. Als $\mu = 50$, dan is T t-verdeeld met df=n-1=29 vrijheidsgraden.
5. We vinden: $t = \dfrac{47 - 50}{7.11/\sqrt{30}} = -2.31$.
6. De overschrijdingskans is de kans dat de toetsgrootheid een waarde aanneemt die minstens evenveel voor H_a pleit als de gevonden waarde. Dus: $P = P(T \leq -2.31) + P(T \geq 2.31)$. Uit tabel van t-verdeling: $0.02 < P < 0.04$.
7. We verwerpen H_0 en aanvaarden daarmee H_a. Reden: Reden: de overschrijdingskans *P)* is kleiner dan (of gelijk aan) het significantieniveau α =0.05.
8. Er is statistisch aannemelijk gemaakt (aangetoond) dat de gemiddelde score in de populatie afwijkt van 50.

Opgave 7

1. Er is een EAS uit de populatie getrokken. De populatieverdeling van scores is een normale verdeling met onbekend gemiddelde μ en onbekende standaardafwijking σ.
2. $H_0: \mu \geq 50$, $H_a: \mu < 50$; $\alpha = 0.05$.
3. $T = \dfrac{\bar{X} - 50}{s/\sqrt{30}}$
4. Als $\mu = 50$, dan is T t-verdeeld met df=n-1=29 vrijheidsgraden.

5. We verwerpen de nulhypothese als $t \leq -1.699$.
6. We vinden: $t = \dfrac{49-50}{7.11/\sqrt{30}} = -0.77$.
7. We verwerpen H_0 niet Reden: de gevonden waarde van de toetsgrootheid ligt niet in het kritieke gebied.
8. Er is niet statistisch aannemelijk gemaakt (aangetoond) dat de gemiddelde score in de populatie kleiner is dan 50.

<u>Opmerking:</u>
Let op de voorzichtige formulering bij de punten 7 en 8 hierboven. Dit heeft te maken met de opmerking die je vindt na de uitwerking van opgave 3.

<u>Opgave 8</u>

1. Er is een EAS uit de populatie getrokken. De populatieverdeling van scores is een normale verdeling met onbekend gemiddelde μ en onbekende standaardafwijking σ.
2. $H_0: \mu \leq 50$, $H_a: \mu > 50$; $\alpha = 0.05$
3. $T = \dfrac{\overline{X} - 50}{s/\sqrt{30}}$
4. Als $\mu = 50$, dan is T t-verdeeld met df =29.
5. We vinden: $t = \dfrac{47-50}{7.11/\sqrt{30}} = -2.31$.
6. De overschrijdingskans is de kans dat de toetsgrootheid een waarde aanneemt die minstens evenveel voor H_a pleit als de gevonden waarde. Dus: $P = P(T \geq -2.31)$. Met tabel: 0.98<P<0.99.
7. We verwerpen H_0 niet Reden: de overschrijdingskans P is groter dan het significantieniveau α.
8. Er is geen statistisch bewijs voor de bewering dat het populatiegemiddelde groter is dan 50.

<u>Opgave 9a</u>

1. De verschilscores vormen een EAS steekproef uit een populatie van verschilscores. De populatieverdeling van verschilscores is een normale verdeling met onbekend gemiddelde μ en onbekende standaardafwijking σ.
2. $H_0: \mu = 0$ (of: $H_0: \mu \leq 0$); $H_a: \mu > 0$. Hierbij is μ het gemiddelde van de verschilscore X in de populatie. X='score met vragenuur- score zonder vragenuur'. $\alpha = 0.05$.
3. $T = \dfrac{\overline{X}}{s/\sqrt{8}}$
4. t-verdeling met df=7.
5. De verschilscores zijn: 15, 5, 6, -3, 5, 6, 5, -11. Het gemiddelde en de standaardafwijking van deze verschilscores: $\overline{X} = 3.50$, s=7.60. De toetsgrootheid heeft als waarde: $t = \dfrac{3.5}{7.60/\sqrt{8}} = 1.30$.

6. $0.10 < P < 0.15$.
7. Omdat de overschrijdingskans groter is dan het significantieniveau (=onbetrouwbaarheidsdrempel), wordt de nulhypothese niet verworpen.
8. Er is niet overtuigend aangetoond dat in de populatie de scores gemiddeld stijgen als het vragenuur wordt ingevoerd.

<u>Opgave 9b</u>

Bij 6 van de acht paren valt het verschil uit in het voordeel van degene die wel het vragenuur volgt. Dus: X=6. De overschrijdingskans bepalen we met tabel C: $P = P(X \geq 6) = 0.1094 + 0.0312 + 0.0039 = 0.1445$. Zelfs bij een significantieniveau van 0.10 kan niet geconcludeerd worden dat met het vragenuur de scores over het algemeen hoger uitvallen dan zonder het vragenuur.

Op de volgende ongenummerde bladzijden staan de power-point sheets van sheets van college 2!!

Uitgewerkte opgaven week 2

De opgaven 1 tot en met 4 hebben betrekking op één onderzoek. Doel van het onderzoek was aan te tonen dat chauffeurs op een gestandaardiseerde rijtest meer fouten maken na inname van een bepaald medicijn dan naar inname van een placebo.

Er zaten in de experimentele groep (medicijn) vier aselect gekozen personen uit zekere populatie. De aantallen fouten die zij maakten zijn: 12, 15, 9 en 8. In de controlegroep (placebo) zitten drie aselect gekozen personen. Zij maakten 11, 8 en 5 fouten.

<u>Opgave 1</u>

a. Waarom is het in de praktijk af te raden om met het onderzoek uit te voeren met dergelijke kleine aantallen proefpersonen als je een statistische toetsing wilt uitvoeren?
b. Waarom is het in de praktijk af te raden om met het onderzoek uit te voeren met dergelijke kleine aantallen proefpersonen als je een betrouwbaarheidsinterval wilt opstellen voor $\mu_{medicijn} - \mu_{placebo}$?

<u>Opgave 2</u>

a. Stel een 95% betrouwbaarheidsinterval op voor $\mu_{medicijn} - \mu_{placebo}$. Neem aan dat in de populatie de standaardafwijking van het aantal fouten bij gebruik van het medicijn gelijk is aan de standaardafwijking bij het gebruik van het placebo. Formuleer het gehanteerde kansmodel.
b. Stel een 95% betrouwbaarheidsinterval op voor $\mu_{medicijn} - \mu_{placebo}$. Neem niet aan dat in de populatie de standaardafwijking van het aantal fouten bij gebruik van het medicijn gelijk is aan de standaardafwijking bij het gebruik van het placebo. Formuleer het gehanteerde kansmodel. (Gebruik de formule van bladzijde 345 om het aantal vrijheidsgraden te bepalen. Rond dit aantal vrijheidsgraden af op een geheel getal.)
c. (Dit onderdeel kun je ook maken voor of op het eerste SPSS-practicum.) Bereken de beide intervallen ook met SPSS.

<u>Opgave 3</u>

Voer een toetsing uit om na te gaan of chauffeurs in de rijtest gemiddeld meer fouten maken na inname van het medicijn dan naar inname van het placebo. Neem aan dat in de populatie de standaardafwijking van het aantal fouten bij gebruik van het medicijn gelijk is aan de standaardafwijking bij het gebruik van het placebo. Doorloop de volgende stappen:

1. Formuleer het kansmodel.
2. Formuleer de nulhypothese en de alternatieve hypothese in termen van de parameters van het kansmodel. Kies daarna het significantieniveau (onbetrouwbaarheidsdrempel) α .
3. Formuleer een geschikte toetsgrootheid in termen van de voorkomende stochastische variabelen.
4. Geef de kansverdeling van de toetsgrootheid onder (het randpunt van) H_0.

5. Bereken of geef de waarde van de toetsgrootheid.
6. Bepaal de overschrijdingskans.
7. Formuleer de conclusie over het al dan niet verwerpen van H_0.
8. Vermeld de conclusie 'in gewone woorden'.

Opgave 4

Voer een toetsing uit om na te gaan of chauffeurs in de rijtest gemiddeld meer fouten maken na inname van het medicijn dan naar inname van het placebo. Neem niet aan dat in de populatie de standaardafwijking van het aantal fouten bij gebruik van het medicijn gelijk is aan de standaardafwijking bij het gebruik van het placebo. Formuleer het gehanteerde kansmodel. (Gebruik de formule van bladzijde 345 om het aantal vrijheidsgraden te bepalen. Rond dit aantal vrijheidsgraden af op een geheel getal.) Doorloop de volgende stappen:

1. Formuleer het kansmodel.
2. Formuleer de nulhypothese en de alternatieve hypothese in termen van de parameters van het kansmodel. Kies daarna het significantieniveau (onbetrouwbaarheidsdrempel) α .
3. Formuleer een geschikte toetsgrootheid in termen van de voorkomende stochastische variabelen.
4. Geef de kansverdeling van de toetsgrootheid onder (het randpunt van) H_0.
5. Bepaal de kritieke waarde(n) en geef het kritieke gebied.
6. Bereken of geef de waarde van de toetsgrootheid.
7. Formuleer de conclusie over het al dan niet verwerpen van H_0 bij de gegeven onbetrouwbaarheidsdrempel.
8. Vermeld de conclusie 'in gewone woorden'.

De opgaven 5 en 6 hebben betrekking op een onderzoek waarbij 29 aselect gekozen volwassen Nederlanders intelligentietest 1 maken, terwijl test 2 gemaakt wordt door 35 aselect gekozen personen. In de onderstaande tabel staan de gemiddelde scores en de standaardafwijkingen van beide groepen.

Group Statistics

	test	N	Mean	Std. Deviation
score	1	29	97.4483	14.79041
	2	35	100.1429	15.21637

Opgave 5

a. Stel een 95% betrouwbaarheidsinterval op voor $\mu_1 - \mu_2$. Neem aan dat in de populatie de standaardafwijking van de scores op beide testen gelijk zijn. Formuleer het gehanteerde kansmodel.

b. Stel een 95% betrouwbaarheidsinterval op voor $\mu_1 - \mu_2$. Neem niet aan dat in de populatie de standaardafwijking van de scores op beide testen gelijk zijn. Formuleer het gehanteerde kansmodel. (Gebruik de formule van bladzijde 345 om het aantal vrijheidsgraden te bepalen. Rond dit aantal vrijheidsgraden af op een geheel getal.)

c. (Dit onderdeel kun je ook maken voor of op het eerste SPSS-practicum.) Bereken de beide intervallen ook met SPSS. De gegevens staan in ‘werkcollege2 IQ.sav’.

Opgave 6

a. Toets of de gemiddelde scores op de beide toetsen in de populatie van elkaar verschillen. Neem aan dat de beide standaarddeviaties gelijk zijn. Doorloop de stappen van opgave 3.
b. Toets of de gemiddelde scores op de beide toetsen in de populatie van elkaar verschillen. Neem *niet* aan dat de beide standaarddeviaties gelijk zijn. Doorloop de stappen van opgave 4.

Opgave 7

In het vak “Inleiding Empirische Cyclus” hebben we geleerd om een aantal designs in ‘X-Y-notatie’ weer te geven. Hieronder staan een aantal van deze designs weergegeven. We nemen aan dat de afhankelijke variabele telkens kwantitatief is en dat de nametingen verkregen zijn met hetzelfde of met een vergelijkbaar instrument als de voormetingen. Geef bij elk van de designs aan hoe je de gegevens kunt analyseren als je wilt nagaan of de behandeling leidt tot hogere scores. Geef bij elk design tevens aan of een ‘significant resultaat’ een voldoende voorwaarde is om te concluderen dat de behandeling effect heeft. (De ‘R’ van ‘randomiseren’ noteren wij, in tegenstelling tot Dooley, in elke rij.)

Design 1:

Y	X	Y

Design 2:

R	X	Y
R	-	Y

Design 3:

X	Y
-	Y

Design 4:

R	Y	X	Y
R	Y	-	Y

Design 5:

Y	X	Y
Y	-	Y

Opgave 8

Op basis van de scores van de personen in een aselecte steekproef van Nederlandse mannen tussen de 45 en 50 jaar, is een 98% betrouwbaarheidsinterval opgesteld voor het gemiddelde gewicht in deze populatie. Dit interval loopt van 84.17 kg tot 86.18 kilogram. Wat kun je zonder verder rekenwerk zeggen over de overschrijdingskans die we zullen vinden als we de aangegeven nulhypothese toetsen tegen de aangegeven alternatieve hypothesen? Motiveer telkens je antwoord.
(Opmerking:
De relaties tussen toetsen en betrouwbaarheidsintervallen die in het vorige hoofdstuk aan de orde zijn geweest, gelden ook nu!)

a. $H_0: \mu = 85$, $H_a: \mu \neq 85$.

b. $H_0: \mu = 86.50$, $H_a: \mu \neq 86.50$.

c. $H_0: \mu = 86.18$, $H_a: \mu \neq 86.18$.

<u>Opgave 9</u>

a. Onderzoeker A wil aantonen dat vrouwen gemiddeld hoger scoren dan mannen op een sociale intelligentietest. De test wordt gemaakt door een aselecte steekproef van vrouwen en een even grote steekproef van mannen. In de volgende tabel staan voor beide steekproeven de gemiddelden en de standaardafwijkingen van de testscores.

	gemiddelde	standaardafwijking
Vrouwen	98.89	11.85
Mannen	98.60	12.39

Merk op dat het verschil tussen de beide gemiddelden erg klein is ten opzichte van de spreiding binnen de beide groepen. Laat zien dat het toch mogelijk is dat bij deze uitkomsten zelfs bij een significantieniveau van 0.01 wordt geconcludeerd dat vrouwen gemiddelde hoger scoren dan mannen.
Ontbrekende gegevens van het onderzoek mag je zelf verzinnen. Motiveer je antwoord door de relevante toetsgrootheid uit te rekenen en de overschrijdingskans daarbij te bepalen.
b. Onderzoeker B wil aantonen dat vrouwen gemiddeld hoger scoren dan mannen op een sociale intelligentietest. De test wordt gemaakt door een aselecte steekproef van vrouwen en een even grote steekproef van mannen. In de volgende tabel staan voor beide steekproeven de gemiddelden en de standaardafwijkingen van de testscores.

	gemiddelde	standaardafwijking
Vrouwen	98.89	11.85
Mannen	85.60	12.39

Merk op dat het verschil tussen de beide gemiddelden groot is ten opzichte van de spreiding binnen de beide groepen. Laat zien dat het toch mogelijk is dat bij deze uitkomsten zelfs bij een significantieniveau van 0.10 niet kan worden geconcludeerd dat vrouwen gemiddelde hoger scoren dan mannen.

Ontbrekende gegevens van het onderzoek mag je zelf verzinnen. Motiveer je antwoord door de relevante toetsgrootheid uit te rekenen en de overschrijdingskans daarbij te bepalen.
c. Geef in tekeningen globaal aan hoe de 95% betrouwbaarheidsintervallen voor $\mu_{vrouw} - \mu_{man}$ er uitzullen zien in de onderdelen a en b. Ga daarbij weer uit van de keuzen die je eerder gemaakt hebt m.b.t. ontbrekende gegevens uit het onderzoek. Geef in de tekening het getal 0 aan en het midden van het betrouwbaarheidsinterval. Schets in beide gevallen hoe het interval ten opzichte van deze beide getallen ligt.

Opgave 10

In deze opgave moet je de relatie tussen tweezijdige t-toetsen op basis van onafhankelijke steekproeven in verband brengen met betrouwbaarheidsintervallen voor $\mu_1 - \mu_2$. Maak daartoe de volgende zin af:

$H_0 : \mu_1 = \mu_2$ wordt bij $\alpha = 0.01$ verworpen ten gunste van $H_a : \mu_1 \neq \mu_2$
dan en slechts dan als
het ….% betrouwbaarheidsinterval voor $\mu_1 - \mu_2$ …….

Antwoorden uitgewerkte opgaven week 2

Opgave 1

a. Het onderscheidingsvermogen is te laag. (Leg zelf verder uit wat het onderscheidende vermogen is.)
b. Het betrouwbaarheidsinterval zal heel erg ‘breed’ worden: je zegt dus heel weinig over de parameter $\mu_{medicijn} - \mu_{placebo}$.

Opgave 2

a. $\bar{x}_1 = \frac{12+15+9+8}{4} = 11$, $\bar{x}_2 = \frac{11+8+5}{3} = 8$, $s_1^2 = \frac{(12-11)^2 + \ldots + (8-11)^2}{4-1} = 10$,

$s_2^2 = \frac{(11-8)^2 + (8-8)^2 + (5-8)^2}{3-1} = 9$, $s_p^2 = \frac{(4-1)10 + (3-1)9}{(4-1)+(3-1)} = 9.60$, $s_p = \sqrt{9.60} = 3.098$

. Het interval loopt van $3 - 2.571 * 3.098\sqrt{1/4 + 1/3} = -3.08$ tot

$3 + 2.571 * 3.098\sqrt{1/4 + 1/3} = 9.08$

b. De formule van bladzijde 345 invullen geeft: $df = 4.60$ (ga dat zelf na!). We ronden af: $df = 5$. He interval loopt van $3 - 2.57\sqrt{10/4 + 9/3} = -3.03$ tot

$3 + 2.57\sqrt{10/4 + 9/3} = 9.03$.

c. Met SPSS vind je bij onderdeel a precies hetzelfde interval als wij berekend hebben. Bij onderdeel b vindt je met SPSS een iets ander interval dan wij: [-3.19, 9.19]. Dit komt doordat SPSS het aantal vrijheidsgraden niet afrondt.

Opgave 3

1. De beide steekproeven zijn onafhankelijk. De ene steekproef is een EAS uit een normale verdeling met verwachting μ_1 en standaardafwijking σ, de andere steekproef is een EAS uit een normale verdeling met verwachting μ_2 en standaardafwijking σ. (Dus: de standaardafwijkingen in de beide populatieverdelingen zijn gelijk.)
2. $H_0 : \mu_1 = \mu_2$ (of: $H_0 : \mu_1 \leq \mu_2$), $H_a : \mu_1 > \mu_2$. Hierbij is μ_1 het gemiddelde aantal fouten in de populatie als iedereen het medicijn zou innemen. Ik kies $\alpha = 0.05$.
3. De toetsgrootheid is: $T = \frac{\bar{X}_1 - \bar{X}_2}{s_p\sqrt{1/4 + 1/3}}$.
4. Onder $H_0 : \mu_1 = \mu_2$ heeft T de t-verdeling met df=5.
5. $t = \frac{11 - 8}{3.098\sqrt{1/4 + 1/3}} = 1.27$
6. $0.10 < P < 0.15$
7. Omdat de overschrijdingskans groter is dan de onbetrouwbaarheidsdrempel (het significantieniveau), wordt H_0 niet verworpen.
8. Er is niet aangetoond dat na inname van het medicijn in de populatie gemiddeld meer fouten worden gemaakt dan na inname van het placebo.

Opgave 4

1. De beide steekproeven zijn onafhankelijk. De ene steekproef is een EAS uit een normale verdeling met verwachting μ_1 en standaardafwijking σ_1, de andere steekproef is een EAS uit een normale verdeling met verwachting μ_2 en standaardafwijking σ_2.
2. $H_0: \mu_1 = \mu_2$ (of: $H_0: \mu_1 \leq \mu_2$), $H_a: \mu_1 > \mu_2$. Hierbij is μ_1 het gemiddelde aantal fouten in de populatie als iedereen het medicijn zou innemen. Ik kies $\alpha = 0.05$.
3. De toetsgrootheid is: $T = \dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{s_1^2}{4} + \dfrac{s_2^2}{3}}}$.
4. Onder $H_0: \mu_1 = \mu_2$ heeft T de t-verdeling. Het aantal vrijheidsgraden van die t-verdeling bepalen we als in opgave 2b. Dus: *df=5.*
5. De nulhypothese wordt verworpen als $t \geq 2.015$. Anders geformuleerd: het kritieke gebied is [2.015, ∞).
6. $t = \dfrac{11-8}{\sqrt{10/4+9/3}} = 1.28$
7. Omdat de gevonden waarde van de toetsgrootheid niet in het kritieke gebied ligt, wordt H_0 niet verworpen.
8. Er is niet aangetoond dat na inname van het medicijn in de populatie gemiddeld meer fouten worden gemaakt dan na inname van het placebo.

Opgave 5

a.

$s_p = \sqrt{\dfrac{28*14.79^2+34*15.22^2}{28+34}} = 15.03$. Het interval loopt van

$(97.45-100.14)-2.000*15.03\sqrt{\dfrac{1}{29}+\dfrac{1}{35}} = -2.69-7.55 = -10.24$ tot $-2.69+7.55 = 4.86$

.

b.

De formule van bladzijde 345 invullen geeft, na afronding: $df = 60$ (ga dat zelf na!). Het interval loopt van $(97.45-100.14)-2.000\sqrt{\dfrac{14.79^2}{29}+\dfrac{15.22^2}{35}} = -2.69-7.53 = -10.22$ tot $-2.69+7.53 = 4.84$.

c. Op afrondingen na vind je met SPSS dezelfde intervallen.

Opgave 6a

1. De beide steekproeven zijn onafhankelijk. De ene steekproef is een EAS uit een normale verdeling met verwachting μ_1 en standaardafwijking σ, de andere steekproef is een EAS uit een normale verdeling met verwachting μ_2 en standaardafwijking σ. (Dus: de standaardafwijkingen in de beide populatieverdelingen zijn gelijk.)

2. $H_0: \mu_1 = \mu_2 \quad H_a: \mu_1 \neq \mu_2$. Hierbij is μ_1 de gemiddelde in de populatie op test 1. Ik kies $\alpha = 0.05$.
3. De toetsgrootheid is: $T = \dfrac{\overline{X}_1 - \overline{X}_2}{s_p\sqrt{1/29 + 1/35}}$.
4. Onder $H_0: \mu_1 = \mu_2$ heeft T de t-verdeling met df=62. (We benaderen deze t-verdeling met die gebaseerd op df=60)
5. $t = \dfrac{97.45 - 100.14}{15.03\sqrt{1/29 + 1/35}} = -0.71$
6. $0.40 < P < 0.50$
7. Omdat de overschrijdingskans groter is dan de onbetrouwbaarheidsdrempel (het significantieniveau), wordt H_0 niet verworpen.
8. Er is niet aangetoond dat in de populatie de gemiddelde scores op beide toetsen verschillen.

<u>Opgave 6b</u>

1. De beide steekproeven zijn onafhankelijk. De ene steekproef is een EAS uit een normale verdeling met verwachting μ_1 en standaardafwijking σ_1, de andere steekproef is een EAS uit een normale verdeling met verwachting μ_2 en standaardafwijking σ_2.
2. $H_0: \mu_1 = \mu_2 \quad H_a: \mu_1 \neq \mu_2$. Hierbij is μ_1 de gemiddelde in de populatie op test 1. Ik kies $\alpha = 0.05$.
3. De toetsgrootheid is $T = \dfrac{\overline{X}_1 - \overline{X}_2}{s_p\sqrt{1/29 + 1/35}}$:.
4. Onder $H_0: \mu_1 = \mu_2$ heeft T de t-verdeling. Het aantal vrijheidsgraden van die t-verdeling bepalen we als in opgave 5b. Dus: *df=60.*
5. De nulhypothese wordt verworpen als $t \leq -2.000$ en ook als $t \geq 2.000$. Anders geformuleerd: het kritieke gebied is: $(-\infty, -2.000] \cup [2.000, \infty)$.
6. $t = \dfrac{97.45 - 100.14}{\sqrt{\dfrac{14.79^2}{29} + \dfrac{15.22^2}{35}}} = -0.71$
7. Omdat de gevonden waarde van de toetsgrootheid niet in het kritieke gebied ligt, wordt H_0 niet verworpen.
8. Er is niet aangetoond dat in de populatie de gemiddelde scores op beide toetsen verschillen.

<u>Opgave 7</u>

Met statistische toetsen kunnen we nagaan of er meer dan toevallige (‘significante’) verschillen bestaan tussen de scores in verschillende condities of tussen voor- en nametingen. Allen bij goed uitgevoerd experimenteel onderzoek (designs 2 en 4) kun je bij een ‘significant resultaat’ concluderen dat de behandeling effect heeft gehad.

Van de ons bekende statistische toetsen kunnen we gebruiken:
<u>Design 1</u> T-toets voor gekoppelde paren of tekentoets om na te gaan of de scores op de nameting systematisch hoger uitvallen dan die op de voormeting.
<u>Designs 2</u> T-toets voor onafhankelijke steekproeven om na te gaan of de scores in de experimentele groep systematisch hoger uitvallen dan die in de controlegroep.
<u>Design 3</u> T-toets voor onafhankelijke steekproeven om na te gaan of de scores in de groep die de behandeling krijgt systematisch hoger uitvallen dan die in de andere groep.
<u>Design 4</u> Er zijn hier twee mogelijke ‘afhankelijke variabelen’ waarop we beide groepen kunnen vergelijken: a. De scores op de natoets b. de verschilscores (= score op nameting – score op voormeting). De verschilscores zul je vooral gebruiken als er tevoren al grote verschillen bestaan tussen de personen. Met één van deze beide afhankelijke variabelen kun je gebruik maken van de t-toets voor onafhankelijke steekproeven om na te gaan of de scores in de experimentele groep systematisch hoger uitvallen dan die in de controlegroep.
<u>Design 5</u> Er zijn hier twee mogelijke ‘afhankelijke variabelen’ waarop we beide groepen kunnen vergelijken: a. De scores op de natoets b. de verschilscores (= score op nameting – score op voormeting). De verschilscores zul je vooral gebruiken als er tevoren al grote verschillen bestaan tussen de personen. Met één van deze beide afhankelijke variabelen kun je gebruik maken van de t-toets voor onafhankelijke steekproeven om na te gaan of de scores in de behandelde groep systematisch hoger uitvallen dan die in de andere groep.

<u>Opgave 8</u>
Zie voor deze opgave het hok op bladzijde 286 van het leerboek!

a. ‘het getal 85 ligt binnen het 95% betrouwbaarheidsinterval voor μ’ $\Rightarrow$ ‘$H_0: \mu = 85$ wordt niet verworpen ten gunste van $H_a: \mu \neq 85$ bij $\alpha = 0.05$’ $\Rightarrow$ ‘$P > 0.05$’
b. ‘het getal 86.5 ligt buiten het 95% betrouwbaarheidsinterval voor μ’ $\Rightarrow$ ‘$H_0: \mu = 86.5$ wordt wel verworpen ten gunste van $H_a: \mu \neq 86.5$ bij $\alpha = 0.05$’ $\Rightarrow$ ‘$P < 0.05$’ (we zitten niet net op de grens tussen wel en niet verwerpen!)
c. We zitten hier ‘net op de grens tussen wel en niet verwerpen’. $P = 0.05$.

<u>Opgave 9</u>

a. Bij hele grote steekproeven (neem bijvoorbeeld $n_1 = n_2 = 20000$), is zelfs dit kleine verschil statistisch ‘zeer significant’ (dan geldt: $t = 2.39$, $P < 0.01$; ga na!). Het 95% betrouwbaarheidsinterval voor $\mu_{vrouw} - \mu_{man}$ zal geheel rechts van nul liggen. Dit interval zal echter in zijn geheel wel ‘heel dicht bij 0’ liggen.
b. Bij hele kleine steekproeven is zelfs dit grote verschil niet statistisch ‘significant’ (ga na!). Het 95% betrouwbaarheidsinterval voor $\mu_{vrouw} - \mu_{man}$ zal heel breed zijn en het getal 0 bevatten.

<u>Opgave 10</u>
$H_0: \mu_1 = \mu_2$ (ofwel: $H_0: \mu_1 - \mu_2 = 0$) wordt bij $\alpha = 0.01$ verworpen ten gunste van $H_0: \mu_1 \neq \mu_2$ (ofwel: $H_0: \mu_1 - \mu_2 \neq 0$)
dan en slechts dan als het 99.% betrouwbaarheidsinterval voor $\mu_1 - \mu_2$ het getal 0 <u>**niet**</u> bevat.

Hoor- en werkcollege 3: HET ONDERSCHEIDINGSVERMOGEN VAN DE
SAMENGESTELDE T-TOETS VOOR ONAFHANKELIJKE STEEKPROEVEN

§1 Inleiding

We zijn in hoofdstuk 7 twee verschillende t-toetsen tegengekomen.
In paragraaf 7.1 werden t-toetsen gebruikt om hypothesen te toetsen over het populatiege - middelde μ van één normale verdeling; in paragraaf 7.2 ging het om het toetsen van hypothesen over het verschil $\mu_1 - \mu_2$ tussen de gemiddelden van twee normale verdelingen met gelijke varianties. Bij deze t-toetsen is telkens alleen op de fout van de eerste soort gelet: de toetsen zijn zo gemaakt dat de kans om H_0 ten onrechte te verwerpen beneden een bepaalde onbetrouwbaarheidsdrempel α blijft. We zijn bij de behandeling van de t-toetsen echter volledig voorbij gegaan aan de kans op een fout van de tweede soort. In de volgende paragrafen zullen we daar alsnog aandacht aan besteden. We beperken ons daarbij tot de samengestelde t-toets voor twee onafhankelijke steekproeven.

§ 2 Het onderscheidingsvermogen van de twee-steekproeven t-toets

We beschouwen hier de toetsingsproblemen die behandeld zijn in § 7.2 van het leerboek. We gaan ervan uit dat voldaan is aan de veronderstellingen die vermeld staan in de kaders op bladzijde 347 van het boek. We zullen het dus alleen hebben over de zogenaamde "samengestelde t-toets voor twee onafhankelijke steekproeven".
Stel we toetsten

$$H_0 : \mu_1 \leq \mu_2$$

tegen

$$H_a : \mu_1 > \mu_2$$

Als μ_1 en μ_2 gelijk zijn, weten we dat de toetsgrootheid

$$T = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{1/n_1 + 1/n_2}}$$

t-verdeeld is met d.f. = $n_1 + n_2 - 2$.
Als we H_0 verwerpen als
$T \geq t^*_{n_1+n_2-2;\alpha}$ (het getal waarboven kansmassa α ligt onder de curve van de t-verdeling met $df = n_1 + n_2 - 2$), dan hebben we dus een toets met onbetrouwbaarheidsdrempel α. Het is echter ook interessant om te weten hoe groot de kans is om H_0 te verwerpen als H_a waar is (dit is het onderscheidingsvermogen) d.w.z. bij bepaalde waarden van σ, μ_1 en μ_2 met $\mu_1 > \mu_2$. Om het onderscheidingsvermogen van de t-toets te bepalen bij gegeven n_1, n_2, α, σ, μ_1 en μ_2 hebben we het volgende resultaat nodig:

Als μ_1, μ_2 en σ de echte waarden zijn van de parameters, dan is de kansverdeling van T de niet-centrale t-verdeling met d.f. = $n_1 + n_2 - 2$ en niet-centraliteits-parameter

$$\delta = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\mu_1 - \mu_2\right) / \sigma .$$

De wiskundige formules van de niet-centrale t-verdelingen laten we natuurlijk achterwege. Op Teletop staan tabellen van een aantal niet-centrale t-verdelingen. Het praktisch gebruik ervan illustreren we aan de hand van enkele voorbeelden.

Voorbeeld 1

Stel $n_1 = n_2 = 10$ en $\alpha = 0.05$.
Als we $H_0 : \mu_1 \leq \mu_2$ toetsen tegen $H_1 : \mu_1 > \mu_2$, wat is dan het onderscheidingsvermogen van de t-toets als $\mu_1 = 23$, $\mu_2 = 20$ en $\sigma = 4$?
<u>Oplossing:</u>
(Uit de tabel van de (centrale) t-verdeling zien we dat we H_0 verwerpen als $T > 1.734$.)
Er geldt:

$$\delta = \sqrt{\frac{100}{20}}\,\frac{3}{4} = 1.68.$$

In de tabel van de niet-centrale t-verdeling zoeken we de rij voor d.f. = 18 en zien daarin de getallen 1.71 en 1.45. Hier ligt 1.68 tussen. We zien dat ß (de kans op een fout van de tweede soort) ligt tussen 0.50 en 0.60. Het onderscheidingsvermogen $\gamma = 1 - ß$ ligt dus tussen 0.40 en 0.50.

Voorbeeld 2

Stel $n_1 = n_2 = 10$ en $\alpha = 0.10$. We willen nu toetsen $H_0 : \mu_1 = \mu_2$ tegen het tweezijdige alternatief $H_1 : \mu_1 \neq \mu_2$. We willen opnieuw het onderscheidingsvermogen berekenen als $\mu_1 = 23$, $\mu_2 = 20$ en $\sigma = 4$. Dit is gelijk aan
$P_{\delta = 1.68}(\,|\,T\,|>1.734) = P_{\delta = 1.68}(T>1.734) + P_{\delta = 1.68}(T<-1.734)$.
De kans $P_{\delta = 1.68}(T<-1.734)$ zal echter verwaarloosbaar klein zijn. We kunnen daarom concluderen dat het onderscheidingsvermogen ongeveer gelijk zal zijn aan dat in voorbeeld 1.

Slordig geformuleerd:
'power bij tweezijdig toetsen en $\alpha = 0.10$'
$\approx$
'power bij eenzijdig toetsen en $\alpha = 0.05$'

NB: De antwoorden op de onderstaande opgaven zijn de vinden op de bladzijden 36 en volgende!

Opgave 1

Een onderwijskundige heeft twee verschillende programma's ontwikkeld om kinderen van de tweede klas van de MAVO elementaire alg ebra te leren. Om te kijken of de beide cursussen een verschillend effect hebben, worden aselect 50 leerlingen gekozen. De helft van hen volgt programma A, de andere helft volgt programma B. Na afloop wordt een al bestaande toets afgenomen. Bekend is dat de standaardafwijking in de populatie van alle leerlingen ongeveer gelijk is aan $\sigma = 10$.
De onderzoeker vindt het belangrijk dat de juiste conclusie getrokken wordt als in de populatie de gemiddelde score bij een van de programma's 4 punten hoger is dan bij het andere.

a. Formuleer H_0 en H_a.
b. Bereken voor de onderzoeker het onderscheidingsvermogen. Stel $\alpha = 0.05$.
c. Van welke veronderstellingen bent u bij b. uitgegaan?

Opgave 2

Stel dat de onderwijskundige uit opgave 1 wil aantonen dat de gemiddelde score in de populatie hoger uitvalt bij programma A dan bij programma B.

a. Formuleer H_0 en H_a.
b. We willen toetsen met een significantieniveau van 0.05. Geef de toetgrootheid en het kritieke gebied.
c. De onderzoeker vindt het belangrijk dat de juiste conclusie getrokken wordt als in de populatie de gemiddelde score bij programma A 4 punten hoger is dan bij B. Bereken de kans dat de toetsgrootheid in het kritieke gebied (zie onderdeel b) terechtkomt als dit verschil van 4 punten zou bestaan.
d. Van welke veronderstellingen bent u bij b. uitgegaan?

Soms is het moeilijk om te bepalen bij welke waarden van μ_1, μ_2 en σ het interessant is om het onderscheidingsvermogen te bepalen.
In het boek 'Statistical Power Analysis for the Behavioral Sciences' (Cohen, 1977) wordt daarom een voorstel gedaan om te komen tot afspraken over wat zal worden verstaan onder 'kleine', 'middelgrote' en 'grote' effecten.
Mede op basis van onderzoek naar gevonden verschillen tussen gemiddelden in sociaal wetenschappelijk onderzoek stelt Cohen voor van een 'klein' effect te spreken als

$$\left|\frac{\mu_2 - \mu_1}{\sigma}\right| = 0.2 .$$

Van 'middelgrote' en 'grote' effecten wordt door Cohen gesproken als respectievelijk

$$\left|\frac{\mu_2 - \mu_1}{\sigma}\right| = 0.5 \text{ en } \left|\frac{\mu_2 - \mu_1}{\sigma}\right| = 0.8 .$$

Hoewel deze definities uiteraard vrij arbitrair zijn, kan het toch handig zijn er mee te werken.
Bij gegeven α, n_1 en n_2 is het natuurlijk niet moeilijk om het onderscheidingsvermogen van bijvoorbeeld de tweezijdige t-toets te berekenen als

$$\left|\frac{\mu_2 - \mu_1}{\sigma}\right| = 3 .$$

De definities van Cohen suggereren echter dat dit weinig zinvol is: dergelijke spectaculaire 'effecten' komen we (in de sociale wetenschappen) toch vrijwel nooit tegen.

Opgave 3

a. Bepaal het onderscheidingsvermogen van de t-toets uit opgave 1 opgave bij 'kleine', 'middelgrote' en 'grote' effecten.
b. Bepaal het onderscheidingsvermogen van de t-toets uit opgave 2 opgave bij 'kleine', 'middelgrote' en 'grote' effecten.

§ 3 Het bepalen van de steekproefgrootte om een bepaald onderscheidingsvermogen te bereiken bij de twee steekproeven t-toets.

In de vorige paragraaf hebben we gekeken hoe bij gegeven steekproefgroottes n_1 en n_2 en onbetrouwbaarheidsdrempel α het onderscheidingsvermogen van t-toetsen bepaald kan worden in een gegeven 'punt' (μ_1, μ_2, σ). Het is echter verstandig om omgekeerd te werk te gaan en de steekproefgroottes n_1 en n_2 zodanig te kiezen dat het onderscheidingsvermogen een gewenst nivo bereikt bij een bepaalde waarde van de 'effectgrootte'

$$\Delta = \frac{\mu_1 - \mu_2}{\sigma} .$$

We beschouwen hier het eenzijdig toetsingsprobleem $H_0 : \mu_1 \leq \mu_2$, $H_a : \mu_1 > \mu_2$.
Verder zullen we de beide steekproefgroottes gelijk nemen: $n_1 = n_2 = n$. We kiezen verder $\alpha = 0.05$. Hoe groot moet n nu bijvoorbeeld gekozen worden om bij een 'groot effect' (d.w.z. als $\Delta = \frac{\mu_1 - \mu_2}{\sigma} = 0.8$) een onderscheidingsvermogen van, bijvoor beeld, 0.90 te krijgen?
We zullen dit probleem op twee manieren oplossen: één keer m.b.v. de tabel van de niet-centrale t-verdeling en één keer met een speciaal voor dit soort problemen ontworpen tabel.

Oplossing methode 1. (*Aan deze methode zullen we geen aandacht besteden)*
We gebruiken de tabel van de niet-centrale t-verdeling 'proberenderwijs'. We doorlopen daarbij een aantal stappen:

1. Omdat $n = n_1 = n_2$ onbekend is, weten we niet in welke rij van de tabel we moeten kijken: $f = n_1 + n_2 - 2 = 2n - 2$ is onbekend. We beginnen daarom maar in de onderste rij ($f = \infty$). We zien dat

$$\delta = \frac{\mu_1 - \mu_2}{\sigma}\sqrt{\frac{n_1 n_2}{n_1 + n_2}} = 0.8\sqrt{\frac{n}{2}}$$

dan gelijk zou moeten zijn aan 2.93.
We lossen op: $0.8\sqrt{\frac{n}{2}} = 2.93$.
Dit geeft een eerste gissing voor n: n = 27 (na afronding).

2. Bij $n \approx 27$ horen we echter niet in de onderste rij ('$f = 2n - 2 = \infty$') te kijken, maar in de rij $f = 60$.
Een volgende gissing krijgen we daarom door de vergelijking $0.8\sqrt{\frac{n}{2}} = 2.96$ op te lossen. Dit levert, na afronding, n = 28.

In dit geval kunnen we na twee stappen ophouden, want bij n = 28 horen we inderdaad (ongeveer) in de rij '$f = 60$' te kijken.

Oplossing methode 2.
We gebruiken nu de tabel 'Sample Size for Comparing the Means of Two Normal Variables'. Deze tabel treft u na deze paragraaf aan. We hebben

$\alpha = 0.05$ (one-sided),

$\beta = 0.10$,

$\Delta = 0.80$.

Uit de tabel lezen we in één keer af: $n = 28$.

Opgave 4
We toetsen

$H_0 : \mu_1 = \mu_2$

tegen

$H_a : \mu_1 \neq \mu_2$.

We kiezen $\alpha = 0.05$. Bij een 'middelgroot effect', d.w.z. als $\frac{|\mu_2 - \mu_1|}{\sigma} = 0.5$, willen we een onderscheidingsvermogen van minstens 0.80. Hoe groot moet $n = n_1 = n_2$ gekozen worden?

Tenslotte volgen hieronder een aantal gemengde opgaven.

Opgave 5
Huisartsen geven aan patiënten die aan een bepaalde ziekte lijden een voorlichtingsfolder mee. De folder wordt grondig herzien door het aanbrengen van kopjes, het vermelden van kernbegrippen in de marge etc. Men wil nagaan of patiënten die de herziene folder meekrijgen er systematisch meer van opsteken dan patiënten die de oude versie krijgen. Men gaat daartoe aan n aselekt gekozen patiënten de nieuwe versie aanbieden en aan evenveel patiënten geeft men de oude versie. Een week nadat de folders uitgedeeld zijn, worden alle patiënten geïnterviewd om te achterhalen welke kennis hij/zij van de ziekte heeft. Er worden 25 vragen gesteld, waarbij de score van de patient gelijk is aan het aantal goed beantwoorde vragen.

Verwacht wordt dat de gemiddelde score bij gebruik van de nieuwe folder in de populatie van alle patiënten ongeveer 0.3 standaarddeviatie hoger uit zal vallen dan bij gebruik van de oude versie. Er wordt gebruik gemaakt van de t-toets met $\alpha = 0.05$.

a. Neem aan dat $n = 30$. Hoe groot is dan de kans dat H_0 wordt verworpen, indien de boven geformuleerde verwachting juist is?

b. Als de boven geformuleerde verwachting juist is, willen we dat er 95 % kans is dat de t-toets tot de juiste conclusie leidt. Hoe groot moet n gekozen worden?

Opgave 6
We willen twee lesmethoden vergelijken. We wijzen aan el ke lesmethode n aselect gekozen leerlingen toe. Alle 2n personen maken na afloop dezelfde studietoets. Met de t-toets gaan we na of er verschil in effect van de twee lesmethoden is. We kieze n een onbetrouwbaarheidsdrempel van 5%.

a. Stel dat $n = 16$. Hoe groot is de kans dat H_0 verworpen wordt als de gemiddelde score in de populatie bij de ene lesmethode een halve standaarddeviatie hoger uitvalt dan bij de andere lesmethode?

b. We willen dat er 90% kans is dat H_0 wordt verworpen als de gemiddelde score in de populatie bij de ene lesmethode een halve standaarddeviatie hoger uitvalt dan bij de andere lesmethode. Hoe groot moeten we n kiezen?

Opgave 7
We willen aantonen dat meisjes systematisch beter presteren op een gestandaardiseerd biologie practicum dan jongens. Neem aan dat het practicum objectief gescoord wordt en dat de scores van zowel de meisjes als van de jongens (ongeveer) normaal verdeeld zijn en dat de standaarddeviaties van de scores in beide groepen gelijk zijn.
We kiezen aselekt n meisjes en n jongens die het practicum gaan doen. Op de scores van deze leerlingen gaan we een t-toets toepassen. We hanteren daarbij een onbetrouwbaarheidsdrempel van $\alpha = 0.05$. Als in de populatie van de meisjes de gemiddelde score 0.3 standaarddeviatie hoger is dan bij d e jongens, willen we dat er 90% kans is dat met de t-toets de juiste conclusie getrokken wordt. Hoe groot moeten we n kiezen?

Antwoorden opgaven uit studietekst 'Het onderscheidend vermogen van de samengestelde t-toets voor onafhankelijke steekproeven

Opgave 1

a. $H_0 : \mu_A = \mu_B$, $H_a : \mu_A \neq \mu_B$

b. $\delta = \sqrt{\frac{n_1 n_2}{n_1+n_2}} \frac{\mu_1 - \mu_2}{\sigma} = \sqrt{\frac{25*25}{25+25}} \frac{4}{10} = 1.41$. df=25+25-2=48. We kijken bij df=40. We kijken in Table C.13 bij 'one sided test with α =0.025'.

1.15<1.41<1.47 $\Rightarrow 0.70 < \beta < 0.80 \Rightarrow 0.20 < \gamma < 0.30$. Of: de berekende waarde 1.41 ligt heel dicht bij de tabelwaarde 1.47, dus: $\beta \approx 0.70 \Rightarrow \gamma \approx 0.30$.

c. Er is sprake van twee onafhankelijke steekproeven. De populatieverdelingen zijn normale verdelingen met gelijke standaardafwijkingen.

Opgave 2

a. $H_0 : \mu_A \leq \mu_B$, $H_a : \mu_A > \mu_B$

b. De toetsgrootheid: $T = \frac{\overline{X}_1 - \overline{X}_2}{s_p \sqrt{\frac{1}{25} + \frac{1}{25}}}$. Het kritieke gebied: $[1.676, \infty)$.

c. Er wordt gevraagd naar het onderscheidend vermogen!

$\delta = \sqrt{\frac{n_1 n_2}{n_1+n_2}} \frac{\mu_1 - \mu_2}{\sigma} = \sqrt{\frac{25*25}{25+25}} \frac{4}{10} = 1.41$. df=25+25-2=48. We kijken bij df=40. We kijken in Table C.13 bij 'one sided test with α =0.05'.

1.14<1.41<1.42 $\Rightarrow 0.60 < \beta < 0.70 \Rightarrow 0.30 < \gamma < 0.40$. Of: de berekende waarde1.41 ligt heel dicht bij de tabelwaarde 1.42, dus: $\beta \approx 0.60 \Rightarrow \gamma \approx 0.40$.

d. Er is sprake van twee onafhankelijke steekproeven. De populatieverdelingen zijn normale verdelingen met gelijke standaardafwijkingen.

Opgave 3

Ik geef alleen het antwoord voor een 'middelgroot effect', dus voor $\frac{\mu_A - \mu_B}{\sigma} = 0.5$. Voor 'kleine' en 'grote' effecten moet in de berekening van δ het getal 0.5 vervangen worden door respectievelijk 0.2 en 0.8.

a. $\delta = \sqrt{\frac{n_1 n_2}{n_1+n_2}} \frac{\mu_1 - \mu_2}{\sigma} = \sqrt{\frac{25*25}{25+25}} 0.5 = 1.77$. df=25+25-2=48. We kijken bij df=40. We kijken in Table C.13 bij 'one sided test with α =0.025'.

1.75<1.77<2.01 $\Rightarrow 0.50 < \beta < 0.60 \Rightarrow 0.40 < \gamma < 0.50$. Of: de berekende waarde 1.77 ligt heel dicht bij de tabelwaarde 1.75, dus: $\beta \approx 0.60 \Rightarrow \gamma \approx 0.40$.

b. $\delta = \sqrt{\frac{n_1 n_2}{n_1+n_2}} \frac{\mu_1 - \mu_2}{\sigma} = \sqrt{\frac{25*25}{25+25}} 0.5 = 1.77$. df=25+25-2=48. We kijken bij df=40. We kijken in Table C.13 bij 'one sided test with α =0.05'.

1.67<1.77<1.93 $\Rightarrow 0.40<\beta<0.50 \Rightarrow 0.50<\gamma<0.60$. Of: de berekende waarde 1.77 ligt vrij dicht bij de tabelwaarde 1.67, dus: $\beta \approx 0.50 \Rightarrow \gamma \approx 0.50$.

Opgave 4

$\alpha = 0.05$ (tweezijdig), $\Delta = 0.5$, $\gamma = 0.80$ $(\beta = 0.20) \Rightarrow n = 64$. In beide condities minstens 64 personen.

Opgave 5

a. $\delta = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}\frac{\mu_1 - \mu_2}{\sigma} = \sqrt{\frac{30*30}{30+30}}\,0.3 = 1.16$. *df*=30+30-2=58. We kijken bij *df*=60. We kijken in Table C.13 bij 'one sided test with α =0.05'.

1.13<1.16<1.41 $\Rightarrow 0.60<\beta<0.70 \Rightarrow 0.30<\gamma<0.40$. Of: de berekende waarde 1.16 ligt heel dicht bij de tabelwaarde 1.13, dus: $\beta \approx 0.70 \Rightarrow \gamma \approx 0.30$.

b. $\alpha = 0.05$ (eenzijdig), $\Delta = 0.3$, $\gamma = 0.95$ $(\beta = 0.05) \Rightarrow n = 242$. In beide condities minstens 242 personen.

Opgave 6

a. $\delta = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}\frac{\mu_1 - \mu_2}{\sigma} = \sqrt{\frac{16*16}{16+16}}\,0.5 = 1.41$. *df*=16+16-2=30. We kijken in Table C.13 bij 'one sided test with α =0.025'.

1.16<1.41<1.48 $\Rightarrow 0.70<\beta<0.80 \Rightarrow 0.20<\gamma<0.30$. Of: de berekende waarde1.41 ligt heel dicht bij de tabelwaarde 1.48, dus: $\beta \approx 0.70 \Rightarrow \gamma \approx 0.30$.

b. $\alpha = 0.05$ (tweezijdig), $\Delta = 0.5$, $\gamma = 0.90$ $(\beta = 0.10) \Rightarrow n = 86$. In beide condities minstens 86 personen.

Opgave 7

$\alpha = 0.05$ (eenzijdig), $\Delta = 0.3$, $\gamma = 0.90$ $(\beta = 0.10) \Rightarrow n = 191$. Kies dus minstens 191 meisjes en 191 jongens.

Op de volgende ongenummerde bladzijden staan de power-point sheets van sheets van college 3!!

<u>Sheets hoorcollege 4 (over paragrafen 15.1 en 15.2)</u>

<u>Opzet</u>	Toets gebaseerd op normaliteit	Niet-parametrische toets
Gekoppelde paren	t-toets voor gekoppelde paren (7.1)	Tekentoets (7.1)/ Rangtekentoets van Wilcoxon (15.2)
2 onafh. steekpr.	Twee-steekproeven t-toets (7.2)	Rangsomtoets van Wilcoxon (=Mann-Whitney; 15.1)

Rangsomtoets van Wicoxon

<u>Voorbeeld</u>
Scores op intelligentietest systematisch hoger na training?

Design:

R	X	Y
R	_	Y

Scores:

Met training	110 98 107 121 115
Zonder training	105 111 97 104 109 103

Toetsgrootheid: *W = 'som van rangnummers in één van beide (bijvoorbeeld de eerste) steekproef' [Bij gebruik van de tabellen van de exacte verdeling van de toetsgrootheid,* ***moet*** *de som van de rangnummers in de* ***kleinste steekproef*** *genomen worden!]*

<u>Berekening W in voorbeeld:</u>

- Orden alle scores, onderscheid score in beide groepen (bijv.:vette letters voor steekproef 1):

97 **98** 103 104 105 **107** 109 **110** 111 **115** **121**

- Vervang scores door rangnummers:

scores	97	**98**	103	104	105	**107**	109	**110**	111	**115**	**121**
Rangn.	1	**2**	3	4	5	**6**	7	**8**	9	**10**	**11**

- Bereken som van rangnummers in (bijv.) eerste steekproef:

W = 2+6+8+10+11=37.

<u>Om bijbehorende overschrijdingskans te vinden moeten we kansverdeling van W kennen bij n_1=5 en n_2=6.</u>

- Bij kleine steekproeven: exacte verdeling gemakkelijk te bepalen. Tabellen staan in syllabus (‘Table 9’; niet in ons boek!). SPSS gebruikt tabellen.

- De kansverdeling van W onder de nulhypothese is gemakkelijk af te leiden. Zie opgave 1 van ‘uitgewerkte opgaven’!

- Bij niet al te kleine steekproeven:

$$z = \frac{W - \mu_W}{\sigma_W} = \frac{W - n_1(n_1 + n_2 + 1)/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}}$$

is bij benadering standaard-normaal verdeeld.

<u>Overschrijdingskans in ons voorbeeld:</u>

- Tabellen geven: $P(eenzijdig) = P(W \geq 37) = 0.123$.
- Met normale benadering:

  $z = \frac{37 - 5 * 12/2}{\sqrt{5 * 6 * 12/12}} = 1.28$, *p(éénzijdig) = 0.1003.*

  - Conclusie: ....

<u>SPSS-uitvoer</u>

- Kies: *Analyze, Non-parametric Tests, 2 Independent Samples, Mann-Whitney U*

Ranks

	groep	N	Mean Rank	Sum of Ranks
score	1	5	7.40	37.00
	2	6	4.83	29.00
	Total	11		

Test Statistics[b]

	score
Mann-Whitney U	8.000
Wilcoxon W	29.000
Z	-1.278
Asymp. Sig. (2-tailed)	.201
Exact Sig. [2*(1-tailed Sig.)]	.247[a]

a. Not corrected for ties.

b. Grouping Variable: groep

- Hoe te handelen in geval van knopen?

- <u>vergelijking Rangsomtoets van Wilcoxon met t-toets voor onafhankelijke steekproeven</u>

<u>Nadeel t-toets (voordeel Wilcoxon)</u>
- Vereist normaliteit, voor al bij kleine steekproeven. Geen normaliteit: echte significantie-niveau kan verschillen van nominale niveau.
- Wilcoxon's toets is geschikt voor *ordinale gegevens.*
- (Kansverdeling toetsgrootheid moeilijk af te leiden.)

<u>Voordeel t-toets (nadeel Wilcoxon)</u>
- hoger onderscheidingsvermogen in geval van (ongeveer) normale verdelingen. (Verschil in power tussen beide toetsen valt overigens mee!!)

<u>Rangtekentoets van Wilcoxon</u>

<u>Voorbeeld</u>

Scores op intelligentietest systematisch hoger na training?

<u>Design:</u> Uit 8 eeneiige tweelingen kiezen we aselect één persoon die training volgt. Broertje/zusje volgt geen training.

Tweeling	1	2	3	4	5	6	7	8
Wel training	131	87	98	119	105	120	102	105
Geen training	125	89	91	122	104	110	102	100

<u>Verschilscores:</u>

6 –2 7 -3 1 10 0 5

Bij deze toets worden paren met verschilscore van 0 weggelaten (kan worden gerechtvaardigd!).

We doen dus alsof we de volgende data hebben:

Tweeling	1	2	3	4	5	6	7
Wel training	131	87	98	119	105	120	105
Geen training	125	89	91	122	104	110	100

Verschilscores:

6 –2 7 -3 1 10 5

Toetsgrootheid:

W^+= 'som rangnummers van positieve verschillen'

Berekening W^+ in voorbeeld

- Orden absolute waarden verschilscores (pos. Verschillen: vet) en vervang door rangnummers:

Abs. verschil	**1**	2	3	**5**	**6**	**7**	**10**
Rangnr.	**1**	2	3	**4**	**5**	**6**	**7**

- $W^+ = 1+4+5+6+7 = 23$

Om overschrijdingskans te bepalen moeten we kansverdeling van W^+ onder H_0 kennen

- Exacte verdeling staat in andere boeken. SPSS gebruikt exacte verdeling.
- Normale benadering:
- $z = \dfrac{W^+ - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}$ is bij benadering $N(0,1)$-verdeeld.

Berekening éénzijdige overschrijdingskans in vb. met normale benadering

$$z = \frac{23 - 7*8/4}{\sqrt{7*8*15/24}} = 1.52 \text{ , } p=0.0643$$

SPSS-uitvoer

- Kies: *Analyze, Non-parametric Tests, 2 Related Samples,Wilcoxon*

Ranks

		N	Mean Rank	Sum of Ranks
geen training - wel training	Negative Ranks	5[a]	4.60	23.00
	Positive Ranks	2[b]	2.50	5.00
	Ties	0[c]		
	Total	7		

a. geen training < wel training

b. geen training > wel training

c. wel training = geen training

Test Statistics[b]

	geen training - wel training
Z	-1.521[a]
Asymp. Sig. (2-tailed)	.128

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

Uitgewerkte opgaven week 4

Opgave 1

Leidt de kansverdeling onder de nulhypothese af van de rangsomtoetsgrootheid W (som van de rangnummers in groep 1) als er geen knopen voorkomen en als $n_1=2$ en $n_2=3$.

Opgave 2

In een onderzoek naar het proefleidereffect kreeg elk van 20 aselect gekozen eerstejaars studenten een rat toegewezen. De studenten moesten deze ratten leren met behulp van visuele tekens door doolhoven te lopen. De helft van de studenten kreeg te horen dat ze een zeer slimme rat hadden gekregen; aan de anderen werd juist verteld dat ze een domme rat hadden gekregen. In werkelijkheid waren de ratten aselect over de studenten verdeeld. Na de training werd van alle ratten de vaardigheid om de weg te vinden in doolhoven getest door een onafhankelijke onderzoeker, die niet wist welke ratten eerder als 'dom' en welke als 'slim' waren aangemerkt. Hieronder staan de s cores. Een hoge score geeft een hoge vaardigheid aan.

'slimme' ratten	43 50 48 45 64 37 49 63 51 52
'domme' ratten	47 41 61 44 46 36 38 58 56 55

Tevoren werd vermoed dat de scores systematisch hoger uit zouden vallen bij de 'slimme' ratten. Toets of dit vermoeden overtuigend ondersteund wordt door deze gegevens. Gebruik een niet-parametrische toets. Hanteer een significantieniveau dat zo dicht mogelijk bij 0.05 ligt. Gebruik in deze opgave de exacte kansverdeling van de toetsgrootheid W. Doorloop de volgende stappen.

9. Formuleer het kansmodel.
10. Formuleer de nulhypothese en de alternatieve hypothese in woorden. Kies daarna het significantieniveau (onbetrouwbaarheidsdrempel) α .
11. Formuleer een geschikte toetsgrootheid in termen van de voorkomende stochastische variabelen.
12. Geef de kansverdeling van de toetsgrootheid onder (het randpunt van) H_0.
13. Bepaal de kritieke waarde(n) en geef het kritieke gebied.
14. Bereken of geef de waarde van de toetsgrootheid.
15. Formuleer de conclusie over het al dan niet verwerpen van H_0 bij de gegeven onbetrouwbaarheidsdrempel.
16. Vermeld de conclusie 'in gewone woorden'.

Opgave 3

In een onderzoek naar het proefleidereffect kreeg elk van 20 aselect gekozen eerstejaars studenten een rat toegewezen. De studenten moesten deze ratten leren met behulp van visuele tekens door doolhoven te lopen. De helft van de studenten kreeg te horen dat ze een zeer slimme rat hadden gekregen; aan de anderen werd juist verteld dat ze een domme rat hadden gekregen. In werkelijkheid waren de ratten aselect over de studenten verdeeld. Na de training werd van alle ratten de vaardigheid om de weg te vinden in doolhoven getest door een onafhankelijke onderzoeker, die niet wist welke ratten eerder als ‘dom’ en welke als ‘slim’ waren aangemerkt. Hieronder staan de s cores. Een hoge score geeft een hoge vaardigheid aan.

‘slimme’ ratten	43 50 48 45 64 37 49 63 51 52
‘domme’ ratten	47 41 61 44 46 36 38 58 56 55

Tevoren werd vermoed dat de scores systematisch hoger uit zouden vallen bij de ‘slimme’ ratten. Toets of dit vermoeden overtuigend ondersteund wordt door deze gegevens. Gebruik een niet-parametrische toets. Hanteer een significantieniveau dat zo dicht mogelijk bij 0.05 ligt. Gebruik in deze opgave de exacte kansverdeling van de toetsgrootheid W. Doorloop de volgende stappen.

1. Formuleer het kansmodel.
2. Formuleer de nulhypothese en de alternatieve hypothese in termen van de parameters van het kansmodel. Kies daarna het significantieniveau (onbetrouwbaarheidsdrempel) α .
3. Formuleer een geschikte toetsgrootheid in termen van de voorkomende stochastische variabelen.
4. Geef de kansverdeling van de toetsgrootheid onder (het randpunt van) H_0.
5. Bereken of geef de waarde van de toetsgrootheid.
6. Bepaal de overschrijdingskans.
7. Formuleer de conclusie over het al dan niet verwerpen van H_0.
8. Vermeld de conclusie ‘in gewone woorden’.

<u>Opgave 4</u>

a. Maak opgave 2 nog eens, maar gebruik nu de gestandaardiseerde toetsgrootheid $Z = \dfrac{W - n_1(n_1+n_2+1)/2}{\sqrt{n_1 n_2 (n_1+n_2+1)/12}}$.

b. Maak opgave 3 nog eens, maar gebruik nu de gestandaardiseerde toetsgrootheid $Z = \dfrac{W - n_1(n_1+n_2+1)/2}{\sqrt{n_1 n_2 (n_1+n_2+1)/12}}$.

<u>Opgave 5</u>

Toets de relevante hypothese uit opgave 1 nog eens, maar nu met een geschikte t-toets. Neem aan ‘dat de beide populatiestandaardafwijkingen gelijk zijn’. Voer de toetsing uit de ‘aanpak met een overschrijdingskans. Hanteer een significantieniveau van 0.05. Volg daarbij de volgende stappen:

1. Formuleer het kansmodel.
2. Formuleer de nulhypothese en de alternatieve hypothese in termen van de parameters van het kansmodel. Kies daarna het significantieniveau (onbetrouwbaarheidsdrempel) α .
3. Formuleer een geschikte toetsgrootheid in termen van de voorkomende stochastische variabelen.
4. Geef de kansverdeling van de toetsgrootheid onder (het randpunt van) H_0.
5. Bepaal de kritieke waarde(n) en geef het kritieke gebied.

6. Bereken of geef de waarde van de toetsgrootheid.
7. Formuleer de conclusie over het al dan niet verwerpen van H_0 bij de gegeven onbetrouwbaarheidsdrempel.
8. Vermeld de conclusie ‘in gewone woorden’.

NB: Als we de slimme ratten met ‘groep 1’aanduiden, dan geldt:
$\bar{x}_1 = 50.20$, $\bar{x}_2 = 48.20$, $s_1 = 8.28$, $s_2 = 8.79$ *Laat zien hoe dit berekend wordt uit de gegeven getallen!*

Opgave 6
In deze opgave vergelijken we de voor- en nadelen van de rangsomtoets van Wilcoxon ten opzichte van de t-toets voor onafhankelijke steekproeven. Vul in:
a. Het voordeel van de toets van Wilcoxon ten opzichte van de t-toets is dat ….
b. Dit voordeel telt vooral zwaar bij ….. steekproeven.
c. Als de beide populatieverdelingen wel (nagenoeg) normale verdelingen zijn, dan heeft de toets van Wilcoxon het volgende nadeel ten opzichte van de t-toets: …..

Opgave 7
In een onderzoek nar het ‘proefleidereffect’ kreeg elk van 10 studenten twee ratten toegewezen. Doordat de ratten een chip droegen, waren ze van elkaar te onderscheiden. Aan elke student werd verteld dat één van de beide ratten slim was en de andere dom. In werkelijkheden waren de ratten aselect gekozen. De studenten moesten deze ratten allerlei taken leren uitvoeren in zo weinig mogelijk tijd. Na deze training werden door een onafhankelijke onderzoeker vastgesteld na hoeveel pogingen elke rat een nieuwe standaardtaak voor het eerst perfect uitvoerde. Tevoren werd vermoed dat de ‘slimme’ ratten systematisch minder pogingen nodig zouden hebben dan de ‘domme’. In de onderstaande tabel staan voor elk van de 20 ratten de aantallen benodigde pogingen.

Proef-leider	1	2	3	4	5	6	7	8	9	10
‘domme’ rat	46	49	47	40	38	58	33	54	67	24
‘slimme’ rat	40	31	55	37	40	54	38	37	47	31

Hanteer een significantieniveau van 0.05. Doorloop de volgende stappen.

1. Formuleer het kansmodel.
2. Formuleer de nulhypothese en de alternatieve hypothese in woorden. Kies daarna het significantieniveau (onbetrouwbaarheidsdrempel) α .
3. Formuleer een geschikte toetsgrootheid in termen van de voorkomende stochastische variabelen.
4. Geef de kansverdeling van de toetsgrootheid onder (het randpunt van) H_0.
5. Bereken of geef de waarde van de toetsgrootheid.
6. Bepaal de overschrijdingskans.
7. Formuleer de conclusie over het al dan niet verwerpen van H_0.
8. Vermeld de conclusie ‘in gewone woorden’.

Opgave 8

a. Analyseer de gegevens uit opgave 7 met een andere niet-parametrische toets uit deze cursus. Geef alleen de waarde van de toetsgrootheid en de overschrijdingskans.

b. Analyseer de gegevens uit opgave 7 ook nog eens met een t-toets. Geef alleen de waarde van de toetsgrootheid en de overschrijdingskans. *NB: Je kunt gebruiken:* $\bar{x} = 4.60, s = 10.56$. *Laat zien hoe dit uit de gegeven getallen berekend wordt!*

Opgave 9

We nemen een attitudeschaal af aan 10 aselect gekozen mannen en aan 10 aselect gekozen vrouwen. De (binnen elke groep geordende) scores staan in de onderstaande tabel.

Case Summaries[a]

				score
geslacht	vrouw	1		172.00
		2		186.00
		3		189.00
		4		191.00
		5		192.00
		6		194.00
		7		203.00
		8		208.00
		9		213.00
		10		229.00
		Total	N	10
	man	1		30.00
		2		90.00
		3		101.00
		4		116.00
		5		126.00
		6		127.00
		7		171.00
		8		190.00
		9		204.00
		10		267.00
		Total	N	10
	Total	N		20

a. Limited to first 100 cases.

Gebruik een geschikte niet-parametrische toets om te kijken of een van de beide geslachten systematisch hoger scoort dan het andere. Kies een significantieniveau van 0.05. Doorloop de stappen van opgave 3. Gebruik de niet-gestandaardiseerde toetsgrootheid.

Antwoorden uitgewerkte opgaven week 4

Opgave 1

Over de beide groepen worden de rangnummers 1, 2, 3, 4 en 5 verdeeld. Twee van deze rangnummers zullen bij groep 1 terechtkomen. Er zijn $\binom{5}{2}=10$ mogelijkheden wat betreft de deelverzameling van de rangnummers die bij groep 1 terechtkomen. Deze 10 mogelijkheden staan in de linkerkolom van de onderstaande tabel. Als de beide steekproeven uit dezelfde populatieverdeling komen, dan hebben elk van deze 10 mogelijkheden dezelfde kans 1/10.

Rang-nummers in groep 1	kans	Waarde W
1, 2	0.10	3
1, 3	0.10	4
1, 4	0.10	5
1, 5	0.10	6
2, 3	0.10	5
2, 4	0.10	6
2, 5	0.10	7
3, 4	0.10	7
3, 5	0.10	8
4, 5	0.10	9

De kansverdeling van W onder de nulhypothese wordt verkregen door de kans horend bij een bepaalde waarde van W op te tellen. Hieronder staat het resultaat.

Waarde van W	3	4	5	6	7	8	9
Kans	.1	.1	.2	.2	.2	.1	.1

Opgave 2

1. We hebben te maken met twee onafhankelijke aselecte steekproeven uit twee populatieverdelingen. De beide populatieverdelingen zijn gelijk of in één van beide populatieverdelingen zijn de scores systematisch hoger dan in de andere. (Het tweede deel van de laatste zin kan wiskundig gepreciseerd worden.)
2. H_0: de beide populatieverdelingen zijn gelijk.; Ha: in de populatieverdeling bij ‘slimme’ ratten zijn de scores systematisch hoger dan bij de ‘domme’ ratten. We kiezen $\alpha = 0.053$. (Uit Table 9 blijkt dat dit het significantieniveau is dat zo dicht moeglijk ligt bij 0.05.)
3. De toetsgrootheid is: W = ‘de som van de rangnummers bij de ‘slimme’ ratten’.
4. De verdeling van W is weergegeven in ‘Table 9’. Zie bij ‘smaller sample size’= ‘larger sample size’ =10
5. Verwerp als $W \geq 127$.
6. In de volgende tabel staan tussen haakjes de rangnummers.

'slimme' ratten	43(5) 50(12) 48(10) 45(7) 64(20) 37(2) 49(11) 63(19) 51(13) 52(14)
'domme' ratten	47(9) 41(4) 61(18) 44(6) 46(8) 36(1) 38(3) 58(17) 56(16) 55(15)

W=5+12+10+7+20+2+11+19+13+14=113.

7. Omdat de gevonden waarde van W (113) niet in het verwerpingsgebied (zie stap 5) ligt, wordt de nulhypothese niet verworpen. (Schrijf NIET: "De nulhypothese is waar/bewezen". Denk aan de vergelijking met de rechtspraak!)
8. Er is niet overtuigend aangetoond dat de scores bij de 'slimme' ratten systematisch hoger uitvallen dan bij de 'domme'. (Let ook nu weer op de voorzichtige formulering!)

<u>Opgave 3</u>

(Alleen de stappen die anders zijn dan in opgave 2!)

5. W=5+12+10+7+20+2+11+19+13+14=113.
6. $P = P(W \geq 113) > P(W \geq 120) = 0.109$. (De overschrijdingskans is groter dan 0.109)
7. Omdat de overschrijdingskans groter is dan α, wordt de nulhypothese niet verworpen.

<u>Opgave 4a</u>

1. We hebben te maken met twee onafhankelijke aselecte steekproeven uit twee populatieverdelingen. De beide populatieverdelingen zijn gelijk of in één van beide populatieverdelingen zijn de scores systematisch hoger dan in de andere. (Het tweede deel van de laatste zin kan wiskundig gepreciseerd worden.)
2. H_0: de beide populatieverdelingen zijn gelijk.; Ha: in de populatieverdeling bij 'slimme' ratten zijn de scores systematisch hoger dan bij de 'domme' ratten. $\alpha = 0.05$
3. De toetsgrootheid is: $Z = \dfrac{W - 10*21/2}{\sqrt{10*10*21/12}}$. Hierbij is W de som van de rangnummers bij de 'slimme' ratten.
4. Onder de nulhypothese is Z standaardnormaal verdeeld.
5. Verwerp als $z \geq 1.645$.
6. In de volgende tabel staan tussen haakjes de rangnummers.

'slimme' ratten	43(5) 50(12) 48(10) 45(7) 64(20) 37(2) 49(11) 63(19) 51(13) 52(14)
'domme' ratten	47(9) 41(4) 61(18) 44(6) 46(8) 36(1) 38(3) 58(17) 56(16) 55(15)

W=5+12+10+7+20+2+11+19+13+14=113; $z = \dfrac{113 - 105}{\sqrt{175}} = 0.60$

7. Omdat de gevonden waarde van z (0.60) niet in het verwerpingsgebied (zie stap 5) ligt, wordt de nulhypothese niet verworpen. (Schrijf NIET: “De nulhypothese is waar/bewezen”. Denk aan de vergelijking met de rechtspraak!)
8. Er is niet overtuigend aangetoond dat de scores bij de ‘slimme’ ratten systematisch hoger uitvallen dan bij de ‘domme’. (Let ook nu weer op de voorzichtige formulering!)

<u>Opgave 4b</u>
(alleen de stappen die anders zijn dan bij opgave 4a)

5. W=5+12+10+7+20+2+11+19+13+14=113; $z = \frac{113 - 105}{\sqrt{175}} = 0.60$
6. $P = P(Z \geq 0.60) = 0.2743$

<u>Opgave 5</u>
1. De beide steekproeven zijn onafhankelijk. De ene steekproef (‘slimme ratten’) is een EAS uit een normale verdeling met verwachting μ_1 en standaardafwijking σ , de andere steekproef is een EAS uit een normale verdeling met verwachting μ_2 en standaardafwijking σ . (Dus: de standaardafwijkingen in de beide populatieverdelingen zijn gelijk.)
2. $H_0 : \mu_1 \leq \mu_2$ $H_a : \mu_1 > \mu_2$. Ik kies $\alpha = 0.05$.
3. De toetsgrootheid is: $T = \frac{\overline{X}_1 - \overline{X}_2}{s_p \sqrt{1/10 + 1/10}}$.
4. Onder $H_0 : \mu_1 = \mu_2$ heeft T de t-verdeling met df=18.
5. $s_p = \sqrt{\frac{9 * 8.28^2 + 9 * 8.79^2}{18}} = 8.54$; $t = \frac{50.20 - 48.20}{8.54 \sqrt{1/10 + 1/10}} = 0.52$
6. $P > 0.25$
7. Omdat de overschrijdingskans groter is dan de onbetrouwbaarheidsdrempel (het significantieniveau), wordt H_0 niet verworpen.
8. Er is niet aangetoond dat in de populatie de scores gemiddeld hoger uitvallen als aan de proefleiders verteld wordt dat de ratten slim zijn dan wanneer hen gezegd wordt dat de ratten dom zijn.

<u>Opgave 6</u>
a. ...bij de toets van Wilcoxon niet hoeft te worden verondersteld dat de beide populatieverdelingen normale verdelingen zijn.
b. ...kleine steekproeven (bij iets grotere steekproeven is de t-toets heel robusst tegen afwijkingen van ‘normaliteit’: de echte kans op een fout van de eerste soort zal niet veel afwijken van de ‘nominale kans’.)
c. Het onderscheidend vermogen van de t-toets is dan groter. (Dit moet je kunnen toelichten!)
d.

Opgave 7

1. We hebben te maken met een EAS van n gekoppelde paren scores.
2. H_0: aantallen pogingen bij de 'domme' ratten hebben dezelfde populatieverdeling als bij 'slimme'; H_a: aantallen pogingen zijn bij 'domme' ratten systematisch hoger dan bij 'slimme'. Verder: $\alpha = 0.05$.
3. De (gestandaardiseerde) toetsgrootheid is: $Z = \dfrac{W^{+} - 10*11/4}{\sqrt{10*11*21/24}}$. Hierbij is W^{+} de som van de rangnummers bij de positieve verschilscores. (Verschilscores= aantal pogingen domme rat- aantal pogingen slimme rat.)
4. Onder de nulhypothese is Z standaardnormaal verdeeld.
5. De verschilscores (met tussen haakjes de rangnummers van hun absolute waarden) zijn: 6(5) 18(9) -8(7) 3(2) -2(1) 4(3) -5(4) 17(8) 20(10) -7(6).
   W^{+}=5+9+2+3+8+10=37. $z = \dfrac{37 - 27.5}{9.81} = 0.97$
6. $P = P(Z \geq 0.97) = 0.1660$.
7. Omdat de overschrijdingskans groter is dan het significantieniveau, wordt de nulhypothese niet verworpen. (Schrijf NIET: "De nulhypothese is waar/bewezen". Denk aan de vergelijking met de rechtspraak!)
8. Er is niet overtuigend aangetoond dat aantallen pogingen zijn bij 'domme' ratten systematisch hoger dan bij 'slimme' (Let ook nu weer op de voorzichtige formulering!)

Opgave 8

a. Pas de tekentoets toe! X = 'aantal proefleiders waarbij 'domme' rat meer pogingen nodig heeft dan 'slimme' rat'= 6, n=10 (geen knopen!). Met tabel C: P=0 .2051+0.1172+0.0439+0.0098+0.0010=0.3770.

b. T-toets voor gekoppelde paren: $t = \dfrac{4.60}{10.56/\sqrt{10}} = 1.38$, $P \approx 0.10$.

Opgave 9

a.

Ranks

		N	Mean Rank	Sum of Ranks
pogdom - pogslim	Negative Ranks	4[a]	4.50	18.00
	Positive Ranks	6[b]	6.17	37.00
	Ties	0[c]		
	Total	10		

a. pogdom < pogslim

b. pogdom > pogslim

c. pogdom = pogslim

Test Statistics[b]

	pogdom - pogslim
Z	-.968[a]
Asymp. Sig. (2-tailed)	.333

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

b. Met de tekentoets:

Frequencies

		N
pogdom - pogslim	Negative Differences[a]	4
	Positive Differences[b]	6
	Ties[c]	0
	Total	10

a. pogdom < pogslim

b. pogdom > pogslim

c. pogdom = pogslim

Test Statistics[b]

	pogdom - pogslim
Exact Sig. (2-tailed)	.754[a]

a. Binomial distribution used.

b. Sign Test

Met de t-toets voor gekoppelde paren:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	pogslim	41.0000	10	8.45905	2.67499
	pogdom	45.6000	10	12.51843	3.95868

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	pogslim & pogdom	10	.551	.099

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)
		Mean	Std. Deviation	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	pogslim - pogdom	-4.600	10.5641	-12.16	2.9571	-1.377	9	.202

<u>Sheets hoorcollege 5 (hoofdstuk 8) en hoorcollege 6 (hoofdstuk 9)</u>

- Inferentie over een fractie
- Vergelijken van 2 fracties
- Inferentie voor kruistabellen ('chi-kwadraat-toets voor onafhankelijkheid')

<u>Inferentie over een fractie</u>

- <u>Probleem</u> In populatie (omvang: N) heeft fractie p een bepaalde eigenschap. We kiezen EAS van omvang n uit de populatie. X is aantal 'successen' ('eenheden met eigenschap E') in steekproef. We willen X gebruiken 'om iets over p te zeggen'.

- <u>Model</u> EAS *met teruglegging*. X is dan *binomiaal* verdeeld met parameters n en p. (zie par. 5.1)

- Bij EAS *zonder teruglegging* beschrijft dit model de situatie goed als n/N ('steekproeffractie') erg klein is. (blz. 225)

- We weten al: $E(X)=\mu_X = np$ en $\sigma_X = \sqrt{np(1-p)}$ (zie blz. 229)

- <u>Zuivere schatter voor p:</u> $\hat{p} = \frac{X}{n}$ (immers: $E(\hat{p})=p$) (Gebruik blz. 197:

  $E(a+bX)=a+bE(X)$)

- <u>Variantie $\hat{p}$ is</u> : $p(1-p)/n$ (Gebruik blz. 200:

  $$\sigma^2(a+bX) = b^2\sigma^2(X)$$

- <u>Regel</u> (nodig om hypothesen over p te toetsen):

  $\dfrac{\hat{p}-p}{\sqrt{p(1-p)/n}}$ is bij benadering N(0, 1)-verdeeld (als n 'groot is').

(Dit komt op hetzelfde neer als hetgeen op blz. 232 staat:

$$\hat{p} \text{ is bij benadering } N\left(p, \sqrt{\frac{p(1-p)}{n}}\right) \text{ verdeeld}$$

- Gevolg: Als de nulhyp. $H_0 : p = p_0$ waar is, dan is de toetsgrootheid

$$\frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$ standaardnormaal verdeeld.

Opgave 1
In het eindexamen van 1995 maakte 45% van alle eindexamenkandidaten een bepaalde opgave uit het examen wiskunde A goed. We wilden aantonen dat in de populatie van alle kandidaten van 2007 dit percentage hoger ligt.
In een EAS van 200 leerlingen blijken 112 leerlingen de opgave correct op te lossen.

a. Formuleer H_0 en H_a.
b. Bereken de waarde die de relevante toetsgrootheid aanneemt. (Antw.: 3.13)
c. Bepaal de overschrijdingskans.
d. Gebruik de overschrijdingskans om te toetsen met $\alpha = 0.05$.
e. Leg uit wat het resultaat van de toetsing betekent.

Boek: blz. 369, oude druk: 476

Antwoord opgave 1
a. $H_0 : p = 0.45$, $H_a : p > 0.45$
b. 3.13
c. 0.0009
d. H_0 wordt verworpen
e. Er is voldoende bewijsmateriaal voor de stelling dat het percentage in de huidige populatie hoger is dan 45%.

- Regel (nodig om betr. interval voor p op te stellen; overslaan: par. 8.1.2):

$$\frac{\hat{p} - p}{\sqrt{\hat{p}(1 - \hat{p})/n}}$$ is bij benadering N(0, 1)-verdeeld (als n 'groot is').

- Gevolg:

BI voor p heeft grenzen

$$\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$$

<u>Opgave 2</u>

a. Stel voor de gegevens van opgave 1 een 95% BI op voor de *fractie* personen die het goede antwoord geeft.
b. Stel voor de gegevens van opgave 1 een 95% BI op voor het *percentage* personen dat het goede antwoord geeft.

<u>Antw. opgave 2</u>

<u>a.</u>

$\hat{p} = \frac{112}{200} = 0.56$; het interval loopt van

$0.56 - 1.96\sqrt{(0.56)(0.44)/200} = 0.56 - 0.07 = 0.49$ tot 0.56 + 0.07 = 0.63

c. Van 49% tot 63%.

Formule voor betrouwbaarheidsinterval voor p: blz. 364 boek.

Het begrip "foutmarge

- $\hat{p} - 1.96\sqrt{\hat{p}(1-\hat{p})/n}$ en $\hat{p} + 1.96\sqrt{\hat{p}(1-\hat{p})/n}$ *zijn de grenzen van 95% BI voor p*
- $1.96\sqrt{\hat{p}(1-\hat{p})/n}$ *is de halve lengte van het BI*
- *Er geldt ook:* $P\left(|\hat{p} - p| \leq 1.96\sqrt{p(1-p/n)}\right) = 0.95$

- *We noemen* $1.96\sqrt{\hat{p}(1-\hat{p})/n}$ *of* $1.96\sqrt{p(1-p)/n}$ *de 95% foutmarge (bij het schatten van p met* $\hat{p}$ *als schatter)*

- *Algemener: als* z^* *de bovenste (1-C)/2 kritieke waarde is voor N(0, 1), dan is* $z^*\sqrt{\hat{p}(1-\hat{p})/n}$ *of* $z^*\sqrt{p(1-p)/n}$ *de 100*C% foutmarge.*

<u>Het kiezen van de steekproefgrootte</u>

De drie volgende vragen komen (vrijwel) op hetzelfde neer:

1. *Hoe groot moet n zijn opdat de 100*C% (bijv.: 95%) foutmarge* $z^*\sqrt{\hat{p}(1-\hat{p})/n}$ *of* $z^*\sqrt{p(1-p)/n}$ *(ongeveer) gelijk is aan m?*

2. *Hoe groot moet n zijn opdat de steekproeffractie* $\hat{p}$ *met kans C niet verder dan m bij de populatie-*
   fractie p vandaan ligt?

(In formule: $P(|\hat{p}-p| \le m) = C$*)*

3. *Hoe groot moet n zijn opdat de halve lengte van het C*100% BI voor p gelijk is aan m (en de lengte dus gelijk aan 2m)?*

<u>Antwoord:</u>

$z^*\sqrt{\hat{p}(1-\hat{p})/n} \le m$ als

(*) [illegible]

[En dus: $z^* \sqrt{p(1-p)/n} \leq m$ als

(**) $n \geq \left(\frac{z^*}{m}\right)^2 p(1-p)$.]

Omdat p(1-p) maximaal is voor p=.5, is aan (*) en (**) zeker voldaan als we kiezen:

$$n = \left(\frac{z^*}{m}\right)^2 .5(1-.5) = \frac{1}{4}\left(\frac{z^*}{m}\right)^2 = \left(\frac{z^*}{2m}\right)^2$$

Boek: blz. 374;

<u>Opgave 3</u>

a. We willen er voor (minstens) 90% zeker van zijn dat $\hat{p}$ niet verder dan 0.02 bij p vandaan ligt. Hoe groot moet n zijn?
b. Het 90% BI voor p mag niet langer worden dan 0.02. Hoe groot moet n zijn?
c. Het 90% BI voor p mag niet langer worden dan 0.04. Hoe groot moet n zijn?

<u>Antw. opgave 3</u>

a. $n = (1.645/(2*0.02))^2 = 1691.3$, neem dus n = 1692.
b. $n = (1.645/(2*0.01))^2 = 6765.06$.
c. Zelfde antwoord als bij a.

<u>Vergelijken van twee fracties</u>

- Onderzoeksopzet: twee onafhankelijke EAS uit twee verschillende dichotome populaties. We willen op basis van steekproeven conclusies trekken over verschil tussen succeskansen/fracties p_1-p_2.

Popu latie	Pop. fractie	Steekpromvang	Aantal succes	Steek-proef-fractie
1	p_1	n_1	X_1	p_1-dak
2	p_2	n_2	X_2	p_2-dak

- Model: X_1 en X_2 zijn onafhankelijk; beide binomiaal verdeeld met parameters (n_1, p_1), resp. (n_2, p_2).

- Voorbeeld: experiment 1 van Loftus (Artikel op Teletop; zie archief)

Regel :

$$\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}} =$$

$$\frac{\text{'schatter'} - \text{'parameter'}}{\text{'standaardafw. van schatter'}}$$

is bij benadering N(0,1)-verdeeld.

Variant 1 (voor opstellen van betr.intervallen):

is bij benadering normaal verdeeld.

Boek: blz. 379 (oud: 485)

Variant 2 (voor toetsen van H_0: $p_1=p_2$):

Als $p_1=p_2$, dan is de volgende (toets)grootheid N(0,1)-verdeeld:

[illegible]

[illegible]

is bij benadering normaal verdeeld.

Hierbij: $\hat{p} = \dfrac{X_1 + X_2}{n_1 + n_2}$

Boek: blz. 385; (oud: 488)

Opgave 4

We kiezen een EAS van 100 studenten Psychologie en een EAS van 80 studenten Economie. Van de psychologen lezen er 35 de Volkskrant, van de economen lezen er 20 de Volkskrant.

a. Stel een 95% BI op voor het verschil in beide populatiefracties.

b. Toets met sign.niveau 0.05 of de beide populatiefracties verschillen.

c. Toets met sign.niveau 0.05 of het percentage Volkskrantlezers bij de psychologen hoger ligt dan bij de economen.

Deel van antw. 4:

a. 0.100 ± 0.133

b./c. z = 1.447

Kruistabellen

- Het vergelijken van de beide fracties 'Volkskrantlezers' (zie vorige opgave) kan ook met de chi-kwadraat-toets, *als we ten minste de tweezijdige alternatieve hyp. (* $H_a : p_1 \neq p_2$*) hanteren.*

- We berekenen de toetsgrootheid in enkele stappen:

Stap 1 De tabel met *geobserveerde frequenties* (en tussen haakjes de 'randtotalen') is:

	Studierichting		
Krant	Psychol.	Economie	Totaal
Geen Volkskr	65	60	(125)
Wel Volkskr	35	20	(55)
Totaal	(100)	(80)	(180)

Stap 2 We berekenen nu de *'verwachte frequenties'.* Bij de berekening daarvan nemen we even aan dat $H_0 : p_1 = p_2$ waar is. Verder beschouwen we de randtotalen als gegeven:

	Studierichting		
Krant	Psychol.	Econ.	Totaal
Geen Volkskr	…	…	(125)
Wel Volkskr	…	…	(55)
Totaal	(100)	(80)	(180)

De fractie Volkskrantlezers in de gehele groep is 55/180. Als de nulhypothese waar is, 'verwachten' we dat de fractie Volkskrantlezers bij de psychologen ook (ongeveer) gelijk zal zijn aan 55/180. Het 'verwachte aantal' is dan dus (55/180)*100 = $\frac{55*100}{180}$=30.56.

[Algemeen: *verwachtefrekw.= (rijtotaal*kolomtotaal)/totaal*]

De tabel met verwachte frekw.:

	Studierichting		
Krant	Psychol.	Econ.	Totaal
Geen Volkskr	69.44	55.56	(125)
Wel Volkskr	30.56	24.44	(55)
Totaal	(100)	(80)	(180)

- De 'chi-kwadraat-toetsgrootheid' drukt in één getal uit 'hoeveel de tabel met waargenomen frekw. afwijkt van die met verwachte frekw.':

$$X^2 = \sum \frac{(waargenomen\,frekw. - verwachte\,frekw.)^2}{verwachte\,frekw.}$$

$$\frac{(65-69.44)^2}{69.44} + \frac{(60-55.56)^2}{55.56} +$$

$$\frac{(35-30.56)^2}{30.56} + \frac{(20-24.44)^2}{24.44} = 2.09$$

[Deze toetsgrootheid. is kwadraat van]

- overschrijdingskans

Als H_0 waar is, is X^2 (ongeveer) chi-kwadraat verdeeld met df. = (r-1)(c-1)=1.
Met tabel: $0.10<p<0.15$.

- Relatie 'z-toets' voor vergelijken twee kansen en χ^2-toets in twee-bij-twee tabel. Zie uitgewerkte opgave 2.

    - De toetsgrootheid X^2 is gelijk aan het kwadraat van de toetsgrootheid z (zie blz 411)
    - Aan de toetsgrootheid z is de richting van het verschil tussen de fracties successen in de steekproeven te zien.
    - Aan de toetsgrootheid X^2 is de richting van het verschil tussen de fracties successen in de steekproeven niet te zien.
    - Gevolg: z is zowel geschikte toetsgrootheid bij tweezijdig toetsingsprobleem ($H_0: p_1 = p_2$ tegen

$H_0 : p_1 \neq p_2$; in dit geval wordt H_0 zowel verworpen voor hele grote als voor hele kleine waarden van z) als voor eenzijdig toetsprobleem (bijvoorbeeld: $H_0 : p_1 \leq p_2$ tegen $H_0 : p_1 > p_2$; in dit geval wordt H_0 alleen verworpen voor hele grote waarden van

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)}}$$)

- De toetsgrootheid X^2 is alleen geschikt bij tweezijdig toetsingsprobleem ($H_0 : p_1 = p_2$ tegen $H_0 : p_1 \neq p_2$; H_0 wordt alleen verworpen voor hele grote waarden van X^2)

- Bij een tabel met meer dan 2 kolommen of meer dan 2 rijen (bijvoorbeeld: vergelijken van drie succeskansen) komt alleen X^2 in aanmerking!!!

- <u>Chi-kwadraat-toetsen bij rxc-tabel (voorbeeld)</u>

- Waarschuwing vooraf: we gaan kijken naar de samenhang tussen twee categorische variabelen. Uit Inleiding Empirische Cyclus (M&T1) weten we al dat de samenhang tussen twee variabelen sterk kan veranderen als een derde variabele constant wordt gehouden.

- Zie ook paragrafen 9.1.5 en 9.1.6 en de opgaven 6 en 7 van werkcollege 2. *Soms moet je ingewikkelder technieken gebruiken, waarmee de samenhang tussen meer dan 2 variabelen tegelijk kunt bestuderen!*

- De steekproefgegevens (2 mogelijke steekproefopzetten: uit elke studierichting een EAS of één EAS; zie par. 9.3.5 (oud: 9.2))

	Studierichting			
Krant	Psych	Econ	Recht	Totaal
Volkskr	65	60	60	185
NRC	35	20	15	70
Telegraaf	10	19	35	64
Trouw	8	7	10	25
Anders	20	25	24	69
Geen	52	69	66	187
Totaal	190	200	210	600

<u>Eerst wat beschrijvende statistiek:</u>

De gezamenlijke steekproefverdeling van Krant en Studierichting (en de marginale steekproefverdelingen van beide variabelen):

krant * studie Crosstabulation

% of Total

		studie			Total
		Psy	Econ	Totaal	
krant	Volkskrant	10.8%	10.0%	10.0%	30.8%
	NRC	5.8%	3.3%	2.5%	11.7%
	Telegraaf	1.7%	3.2%	5.8%	10.7%
	Trouw	1.3%	1.2%	1.7%	4.2%
	Anders	3.3%	4.2%	4.0%	11.5%
	Geen	8.7%	11.5%	11.0%	31.2%
Total		31.7%	33.3%	35.0%	100.0%

<u>De voorwaardelijke steekproefverdelingen van Krant gegeven elk van de waarden van Studie</u>

krant * studie Crosstabulation

% within studie

		studie			Total
		Psy	Econ	Rechten	
krant	Volkskrant	34.2%	30.0%	28.6%	30.8%
	NRC	18.4%	10.0%	7.1%	11.7%
	Telegraaf	5.3%	9.5%	16.7%	10.7%
	Trouw	4.2%	3.5%	4.8%	4.2%
	Anders	10.5%	12.5%	11.4%	11.5%
	Geen	27.4%	34.5%	31.4%	31.2%
Total		100.0%	100.0%	100.0%	100.0%

–

- Vergelijking van de drie voorwaardelijke steekproefverdelingen suggereert dat er samenhang zou kunnen bestaan. (Bij Psy lijkt NRC relatief populair, bij Rechten de Telegraaf,..)

- Om te zien of deze verschillen op toeval zouden kunnen berusten, gaan we toetsen.

- H_0: geen samenhang tussen studierichting en 'krantkeuze'.

- Verwachte frekw. weer volgens *verwachtefrekw.= (rijtotaal*kolomtotaal)/totaal* :

	Studierichting			
Krant	Psych	Econ	Recht	Totaal
Volkskr	58.6	61.7	64.8	185
NRC	22.2	23.3	24.5	70
Telegraaf	20.3	21.3	22.4	64
Trouw	7.9	8.3	8.8	25
Anders	21.9	23.0	24.2	69
Geen	59.2	62.3	65.5	187
Totaal	190	200	210	600

- <u>de waarde van de toetsgrootheid:</u>

$$X^2 = \sum \frac{(waargenomen frekw. - verwachte frekw.)^2}{verwachte frekw.}$$

$$\frac{(65-58.6)^2}{58.6} + \ldots + \frac{(66-65.6)^2}{65.6} = 27.6$$

- overschrijdingskans:

df= (r-1)(c-1)=(6-1)(3-1)=10;
met tabel: 0.001<p<0.0025

Voorbeeld met SPSS

- Invoer data:

Voer voor elk van de 18 cellen 3 variabelen in: 'studie', krantkeuze' en 'aantal':

Case Summaries[a]

	STUDIE	krantkeuze	AANTAL
1	1.00	Volkskr	65.00
2	1.00	NRC	35.00
3	1.00	Telegraaf	10.00
4	1.00	Trouw	8.00
5	1.00	andere krant	20.00
6	1.00	geen krant	52.00
7	2.00	Volkskr	60.00
8	2.00	NRC	20.00
9	2.00	Telegraaf	19.00
10	2.00	Trouw	7.00
11	2.00	andere krant	25.00
12	2.00	geen krant	69.00
13	3.00	Volkskr	60.00
14	3.00	NRC	15.00
15	3.00	Telegraaf	35.00
16	3.00	Trouw	10.00
17	3.00	andere krant	24.00
18	3.00	geen krant	66.00
Total N	600	600	600

Geef vervolgens de opdracht "Weight by aantal" ('Weight' vindt je onder 'Data')

- Geef vervolgens de opdracht

```
WEIGHT
  BY aantal .
CROSSTABS
  /TABLES=krant  BY studie
  /FORMAT= AVALUE TABLES
  /STATISTIC=CHISQ
  /CELLS= COUNT EXPECTED COLUMN .
```

- Uitvoer:

krantkeuze * STUDIE Crosstabulation

			STUDIE			
			1.00	2.00	3.00	Total
krantkeuze	Volkskr	Count	65	60	60	185
		Expected Count	58.6	61.7	64.8	185.0
		% within STUDIE	34.2%	30.0%	28.6%	30.8%
	NRC	Count	35	20	15	70
		Expected Count	22.2	23.3	24.5	70.0
		% within STUDIE	18.4%	10.0%	7.1%	11.7%
	Telegraaf	Count	10	19	35	64
		Expected Count	20.3	21.3	22.4	64.0
		% within STUDIE	5.3%	9.5%	16.7%	10.7%
	Trouw	Count	8	7	10	25
		Expected Count	7.9	8.3	8.8	25.0
		% within STUDIE	4.2%	3.5%	4.8%	4.2%
	andere krant	Count	20	25	24	69
		Expected Count	21.9	23.0	24.2	69.0
		% within STUDIE	10.5%	12.5%	11.4%	11.5%
	geen krant	Count	52	69	66	187
		Expected Count	59.2	62.3	65.5	187.0
		% within STUDIE	27.4%	34.5%	31.4%	31.2%
Total		Count	190	200	210	600
		Expected Count	190.0	200.0	210.0	600.0
		% within STUDIE	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	27.551[a]	10	.002
Likelihood Ratio	27.220	10	.002
Linear-by-Linear Association	2.967	1	.085
N of Valid Cases	600		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.92.

Uitgewerkte opgaven week 5

Opgave 1

Regel: Als een EAS steekproef wordt getrokken uit een grote populatie met onbekende succesfractie p, dan is de stochastische variabele $\frac{X - np}{\sqrt{n\hat{p}(1-\hat{p})}} = \frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})/n}}$ bij benadering standaardnormaal verdeeld. (De regel geldt ook voor kleine populaties als er een EAS *met teruglegging* wordt getrokken.)

Gebruik deze regel om te laten zien dat $\hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/n}$ een 95% betrouwbaarheidsinterval voor de populatiefractie p definieert.

Opgave 2

We zijn geïnteresseerd in de fractie p van woorden uit een woordenboek Engels-Nederlands waarvan Fransje een goede vertaling weet te geven. We kiezen aselect 50 woorden; hiervan weet Fransje er 30 goed te vertalen.

a. Stel een 90% betrouwbaarheidsinterval op voor de fractie p.
b. Leg voor een niet statistisch geschoold iemand uit wat het berekende interval betekent. Begin je antwoord met als volgt: "Op basis van de steekproefgegevens kunnen we er redelijk op vertrouwen dat …"
c. Leg nu precies uit wat die "90% betrouwbaarheid" betekent: 'Als er telkens opnieuw een aselecte steekproef van 50 woorden zou worden gekozen (en de populatiefractie *p* blijft constant!), dan ….

Opgave 3

In een parapsychologisch experiment zit een proefpersoon in de een kamer. De proefpersoon krijgt 50 keer achter elkaar een kaart aangeboden. Bij elke herhaling wordt de kaart aselect uit een stapel van vijf gekozen. Op elk van de vijf kaarten staat een andere geometrische figuur. De proefpersoon moet zich telkens concentreren op de gekozen figuur. In een andere kamer zit een persoon van wie men denkt dat deze paranormaal begaafd is. Deze persoon moet telkens raden op welk van de vijf figuren de proefpersoon zich concentreert. Het blijkt dat het aantallen correcte gissingen bij dit experiment gelijk is aan 15. Toets met een significantieniveau van 5% of de 'paranormaal begaafde persoon' vaker goed heeft geraden dan op basis van toeval mag worden verwacht. Doorloop de volgende stappen:

1. Formuleer het kansmodel.
2. Formuleer de nulhypothese en de alternatieve hypothese in termen van de parameters van het kansmodel. Kies daarna het significantieniveau (onbetrouwbaarheidsdrempel) α.
3. Formuleer een geschikte toetsgrootheid in termen van de voorkomende stochastische variabelen. Gebruik de gestandaardiseerde versie Z.
4. Geef de kansverdeling van de toetsgrootheid onder (het randpunt van) H_0.
5. Bepaal de kritieke waarde(n) en geef het kritieke gebied.
6. Bereken of geef de waarde van de toetsgrootheid.

7. Formuleer de conclusie over het al dan niet verwerpen van H_0 bij de gegeven onbetrouwbaarheidsdrempel.
8. Vermeld de conclusie ‘in gewone woorden’.

Opdracht 4

Een multinational wil weten hoe groot het percentage werknemers is die een opleiding volgen. Daartoe wil men een EAS van n personen uit de populatie van alle werknemers kiezen.

a. Men wil er voor 90% zeker van zijn dat de fractie mensen in de steekproef die een opleiding volgen niet meer dan 0.10 bij de populatiefractie vandaan ligt. Hoe groot moet n gekozen worden?
b. We willen dat de lengte van het 90% betrouwbaarheidsinterval niet groter is dan 0.10. Hoe groot moet n gekozen worden?

Het onderscheidingsvermogen van de toetsen voor een populatiefractie: een voorbeeld ter voorbereiding op opgave 5.

Kijk eerst nog eens op bladzijde 293 en volgende van het leerboek als je niet meer weet wat het begrip onderscheidingsvermogen inhoudt.
Het bepalen van het onderscheidingsvermogen van de toetsen voor een populatiefractie gaat hier in grote lijnen net als in hoofdstuk 6. We weten dat bij voldoende grote n de toevalsvariabele $\dfrac{\hat{p}-p}{\sqrt{p(1-p)/n}}$ bij benadering standaardnormaal verdeeld is. Op pagina 369 is die regel gebruikt om toetsen voor hypothesen over p te construeren. We kunnen deze regel ook gebruiken om het onderscheidingsvermogen van deze toetsen te bepalen. Stel dat we $H_0: p \leq 0.40$ willen toetsen tegen $H_a: p > 0.40$. Neem verder aan dat n=100 en $\alpha = 0.10$. We bepalen hieronder het onderscheidingsvermogen $\gamma(0.50)$ van onze toets als p gelijk aan 0.50 zou zijn. We doen dit in een aantal stappen:

1. Als p=0.40, dan is $Z = \dfrac{\hat{p}-0.4}{\sqrt{0.4(1-0.4)/100}}$ standaardnormaal verdeeld. z is onze toetsingsgrootheid. We verwerpen de nulhypothese als $Z \geq 1.282$, ofwel als $\hat{p} \geq 0.4 + 1.282\sqrt{0.4(0.6)/100}$.

2. Als p=0.50,dan is $Z = \dfrac{\hat{p} - 0.5}{\sqrt{0.5(1-0.5)/100}}$ standaardnormaal verdeeld. Hiermee

kunt we het onderscheidingsvermogen bepalen:

$$\gamma(0.5) = P(H_0 \text{ wordt verworpen}, p = 0.5) =$$

$$P(\hat{p} \geq 0.4 + 1.282\sqrt{0.4(0.6)/100}; p = 0.5) =$$

$$P(\frac{\hat{p} - 0.5}{\sqrt{0.5(0.5)/100}} \geq \frac{0.4 + 1.282\sqrt{0.4(0.6)/100} - 0.5}{\sqrt{0.5(0.5)/100}}; p = 0.5) =$$

$$P(Z \geq \frac{0.4 + 1.282\sqrt{0.4(0.6)/100} - 0.5}{\sqrt{0.5(0.5)/100}}) = P(Z \geq -0.74) = 1 - 0.23 = 0.77$$

Als je het bovenstaande voorbeeld begrepen hebt, kun je de volgende opdracht maken.

<u>Opgave 5</u>

We willen aantonen dat minder dan 50% van de inwoners van een bepaalde stad voorstander is van samenvoeging met een naburige stad/gemeente. We kiezen aselect 100 inwoners en toetsen op basis van het aantal voorstanders in de steekproef een relevante hypothese met significantieniveau α .

a. Formuleer H_0 en H_a.

b. Welke (gestandaardiseerde) toetsingsgrootheid wordt gebruikt?

c. Voor welke waarden van de (gestandaardiseerde) toetsingsgrootheid wordt de nulhypothese verworpen als α =0.05 en n=100? Voor welke waarden van $\hat{p}$ wordt bij dit significantieniveau en deze steekproefgrootte de nulhypothese verworpen?

d. Bepaal het onderscheidingsvermogen van de toets met α =0.05, indien 30% van de inwoners voorstander zou zijn en als n=100.

<u>Opgave 6</u>

In het artikel ‘Animal companions and one-year survival of patients after discharge from a coronary car unit’ (gepubliceerd in *Public Health Reports,* 96 (1980), blz. 307-312) doen Erika Friedman et al verslag van een onderzoek waarin ze de relatie onderzoeken tussen het wel of niet overleven van het eerste jaar na het ontslag uit het ziekenhuis en het al dan niet hebben van een huisdier. In de onderstaande tabel staan de gegevens van de 92 onderzochte patiënten.

Toestand patiënt	Huisdier?	
	Nee	Ja
In leven	28	50
Overleden	11	3

a. Toets een relevante hypothese. Gebruik dat de onderzoekers verwachtten dat de overlevingskans in groep van personen met huisdier hoger zou zijn dan in die zonder huisdier. Kies voor de aanpak met een overschrijdingskans. Volg daarbij het bekende 'acht stappen plan' (zie uitgewerkte opgaven van eerdere weken).

b. Wat kun je zeggen over de invloed van het hebben van een huisdier op de overlevingskans?

Opgave 7

Een rijinstructeur wil aantonen dat de kans om voor het theorie-examen te slagen groter is als de nieuwe lesmethode B gevolgd wordt, dan wanneer de oude lesmethode A gevolgd wordt. Gebruik in deze opgave SPSS.

a. Bedenk zelf steekproefgegevens, waarbij 70% van de kandidaten slaagt bij lesmethode B en 50% bij lesmethode A. Zorg er echter voor zelfs bij een significantieniveau van 0.10 niet geconcludeerd mag worden dat de slaagkans bij lesmethode B hoger is dan bij A.
b. Bedenk zelf steekproefgegevens, waarbij 54% van de kandidaten slaagt bij lesmethode B en 50% bij de kandidaten van lesmethode A. Zorg er echter voor zelfs bij een significantieniveau van 0.01 wel geconcludeerd mag worden dat de slaagkans bij lesmethode B hoger is dan bij A.
c. Stel voor de gegevens die verzonnen zijn bij onderdeel a een 95% betrouwbaarheidsinterval op voor $p_B - p_A$. *(Terzijde: Kun je tevoren een beetje raden hoe het interval eruit zal zien (Wat is het middelpunt? Ligt 0 in het interval?)*
d. Stel voor de gegevens die verzonnen zijn bij onderdeel b een 95% betrouwbaarheidsinterval op voor $p_B - p_A$. *(Terzijde: Kun je tevoren een beetje raden hoe het interval eruit zal zien (Wat is het middelpunt? Ligt 0 in het interval?)*
e. Waarom is de informatie die het betrouwbaarheidsinterval geeft voor de gebruiker vaak nuttiger dan de uitkomst van de toetsing?

Antwoorden uitgewerkte opgaven week 5

Antwoord opgave 1

$$0.95 = P\left(-1.96 \le \frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} \le 1.96\right) =$$

$$P\left(-1.96\sqrt{\hat{p}(1-\hat{p})/n} \le \hat{p}-p \le 1.96\sqrt{\hat{p}(1-\hat{p})/n}\right) =$$

$$P\left(-\hat{p}-1.96\sqrt{\hat{p}(1-\hat{p})/n} \le -p \le -\hat{p}+1.96\sqrt{\hat{p}(1-\hat{p})/n}\right) =$$

$$P\left(\hat{p}-1.96\sqrt{\hat{p}(1-\hat{p})/n} \le p \le \hat{p}+1.96\sqrt{\hat{p}(1-\hat{p})/n}\right)$$

Antwoord opgave 2

a. $\hat{p} = \frac{30}{50} = 0.6$; het interval loopt van

$0.60 - 1.645\sqrt{0.6(1-0.6)/50} = 0.60 - 0.11 = 0.49$

tot 0.60 + 0.11 = 0.71

b. '….dat de fractie van alle ingangen uit het woordenboek die Fransje correct vertaalt ergens tussen 0.49 en 0.71 ligt.'

c. 'Als er telkens opnieuw een aselecte steekproef van 50 woorden zou worden gekozen (en de populatiefractie *p* blijft constant!), dan zal $\hat{p}$, en dus het interval $\hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/n}$, telkens variëren. Op de lange duur zal ongeveer 95% van deze intervallen de populatiefractie *p* bevatten.

Antwoord opgave 3

1. Er is sprake van n onafhankelijke herhalingen telkens met onbekende succeskans *p*. (Of: Er is sprake van een EAS uit een grote populatie herhalingen met onbekende succesfractie *p*.)
2. $H_0: p = 0.20$, $H_a: p > 0.20$; $\alpha = 0.05$.
3. $Z = \frac{\hat{p} - 0.2}{\sqrt{0.2(1-0.2)/50}}$
4. Onder H_0 is Z standaardnormaal verdeeld.
5. Verwerp H_0 als $z \ge 1.645$. (KG = [1.645, ∞))
6. $\hat{p} = \frac{15}{50} = 0.30$, $Z = \frac{0.3 - 0.2}{\sqrt{0.2(1-0.2)/50}} = 1.77$.
7. De nulhypothese wordt verworpen, omdat de gevonden waarde van de toetsgrootheid in het kritieke gebied ligt.

8. Op basis van dit experiment concluderen we dat de ‘paranormaal begaafde persoon’ vaker goed raadt dan op basis van toeval mag worden verwacht.

<u>Antwoord opgave 4</u>

In de formule $n = \left(\frac{z^*}{2m}\right)^2$ is *m* gelijk aan ‘de maximaal toegestane afstand tussen schatter $\hat{p}$ en parameter *p*’ (zie a.) of aan de ‘halve lengte van het betrouwbaarheidsinterval’ (zie b.). Iets soortgelijks gold voor *m* in de formule op bladzijde 266, alleen ging het daar om de paramater μ en zijn schatter $\overline{X}$.

a. $n = \left(\frac{1.645}{2(0.10)}\right)^2 = 67.7$ Kies dus maar n=68.

b. $n = \left(\frac{1.645}{2(0.05)}\right)^2 = 270.6$

<u>Antwoord opgave 5</u>

a. $H_0: p \geq 0.50 \quad H_a: p < 0.50$

b. $z = \frac{\hat{p} - 0.5}{\sqrt{0.5(0.5)/100}}$

c. Verwerp H_0 als $z \leq -1.645$, ofwel als

$\hat{p} \leq 0.50 - 1.645\sqrt{0.5(0.5)/100} = 0.42$

d.

$$P(\hat{p} \leq 0.42;\ p = 0.30) =$$

$$P\left(\frac{\hat{p} - 0.30}{\sqrt{0.3(0.7)/100}} \leq \frac{0.42 - 0.30}{\sqrt{0.3(0.7)/100}};\ p = 0.30\right)$$

$$= P(Z \leq 2.62) = 0.9956$$

<u>Antwoord opgave 6a</u>

1. (Zie bladzijde 382) Er wordt een EAS van omvang n_1 getrokken uit een grote populatie met fractie successen p_1 en een EAS van omvang n_2 getrokken uit een andere grote populatie met fractie successen p_1. **De beide steekproeven zijn onafhankelijk.** *NB: Als de populaties niet erg groot zijn, moet aangenomen worden dat de EAS getrokken worden met teruglegging.*

2. Definieer: p_1 = fractie overlevenden in populatie van patiënten zonder huisdier (of: kans op overleven met huisdier), p_2= kans op overleven met huisdier. $H_o: p_1 \geq p_2$,

$H_a: p_1 < p_2$. Ik kies $\alpha = 0.05$. $Z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{39}+\frac{1}{53})}}$

3. *NB Kleine waarden van deze toetsgrootheid pleiten voor de H_a.*

(Als $Z = \dfrac{\hat{p}_2 - \hat{p}_1}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{39}+\frac{1}{53})}}$ *gebruikt wordt als toetsgrootheid, dan pleiten*

natuurlijk grote waarden voor Ha.)

4. Als de beide slaagkansen gelijk zijn, dan is de toetsgrootheid standaardnormaal verdeeld.

5. $\hat{p}_1 = \frac{28}{39}, \hat{p}_2 = \frac{50}{53}, \hat{p} = \frac{28+50}{39+53}$. Invullen in de formule in stap 3 geeft: z=-2.92.

6. Bepaal de overschrijdingskans. $P = P(Z \leq -2.92) = 0.0018$.

7. Omdat de overschrijdingskans kleiner is dan het significantieniveau, wordt de nulhypothese verworpen.

8. Er is overtuigend bewijs dat de overlevingskans in de populatie met huisdier groter is dan in de populatie zonder huisdier,

<u>Antwoord opgave 6b</u>

Er is geen sprake van een experiment. Er is weliswaar een statistisch verband aangetoond tussen het hebben van een huisdier en het ‘overleven’, maar dat hoeft nog niet te betekenen dat het hebben van een huisdier <u>invloed</u> heeft op het ‘overleven’. Misschien is de overlevingskans het laagst voor de patiënten die er voor de operatie het slechts aan toe waren. Verder zouden deze mensen de huisdieren noodgedwongen al de deur uit hebben kunnen doen!

<u>Antwoord opgave 7</u>

Bij de onderdelen a en b toetsen we bij verschillende data $H_0: p_A \geq p_B$ tegen

$H_a: p_A < p_B$. Ik neem als toetsgrootheid $Z = \dfrac{\hat{p}_B - \hat{p}_A}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_A}+\frac{1}{n_B})}}$. Grote waarden van deze toetsgrootheid pleiten voor de alternatieve hypothese. (Neem je als toetsgrootheid

$Z = \dfrac{\hat{p}_A - \hat{p}_B}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_A}+\frac{1}{n_B})}}$, dan pleiten kleine waarden voor de alternatieve hypothese!)

Bij onderdeel a nemen we n_1 en n_2 natuurlijk klein kiezen (‘bij kleine steekproeven kunnen zelf grote verschillen in steekproeffracties nog op toeval berusten’).
Bij onderdeel b moeten we n_1 en n_2 natuurlijk heel groot kiezen (‘bij grote steekproeven kunnen zij zelfs kleine verschillen in steekproeffracties statistisch significant’).

a. Ik kies steekproefgroottes van 10. De aantallen geslaagden moeten dan zijn:

$X_A=5$, $X_B=7$. $z=\dfrac{\frac{7}{10}-\frac{5}{10}}{\sqrt{\frac{12}{20}\left(1-\frac{12}{20}\right)\left(\frac{1}{10}+\frac{1}{10}\right)}}=0.91$, $P=P(Z\geq 0.91)=0.1814$. De overschrijdingskans is inderdaad groter dan 0.10!

b. Ik kies steekproefgroottes van 10000. De aantallen geslaagden moeten dan zijn:

$X_A=5000$, $X_B=5400$. $z=\dfrac{\frac{5400}{10000}-\frac{5000}{10000}}{\sqrt{\frac{10400}{20000}\left(1-\frac{10400}{20000}\right)\left(\frac{1}{10000}+\frac{1}{10000}\right)}}=5.66$,

$P=P(Z\geq 5.66)<0.0002$. De overschrijdingskans is inderdaad kleiner dan 0.01!

c. *(Het middelpunt van het betrouwbaarheidsinterval is natuurlijk 0.20. Verder zal het getal 0 wel in het interval liggen.)* Het interval loopt van

$(0.70-0.50)-1.96\sqrt{\frac{0.7*0.3}{10}+\frac{0.5*0.5}{10}}=0.20-0.42=-0.22$ tot $0.20+0.42=0.62$.

d. *(Het middelpunt van het betrouwbaarheidsinterval is natuurlijk 0.04. Verder zal het getal 0 wel in het interval liggen.)* Het interval loopt van

$(0.54-0.50)-1.96\sqrt{\frac{0.54*0.46}{10000}+\frac{0.5*0.5}{10000}}=0.04-0.01=0.03$ tot

00.04+0.01=0.05.

e. *Bij de onderdelen a en c:* aan de toetsing zie je alleen maar dat je niet mag concluderen dat de slaagkans bij B groter is dan bij A. Aan het interval zie je (omdat het zo enorm breed is) dat je eigenlijk nog bijna niks weet over het verschil tussen de slaagkansen. *Bij de onderdelen b en d:* De toetsing leert ons alleen dat de slaagkans bij lesmethode B groter is dan bij A. Het betrouwbaarheidsinterval geeft deze informatie ook, want het ligt rechts van het getal 0. Daarnaast geeft het betrouwbaarheidsinterval ook aan *hoe groot* het verschil tussen de beide slaagkansen ongeveer is.

Uitgewerkte opgaven week 6

Sommige van deze opgaven horen wordt je geacht pas te maken voor of tijdens het volgende SPSS-practicum. Dat wordt dan telkens expliciet aangegeven!

Opgave 1

Deze opgave is gebaseerd op het volgende artikel: Lewis, C. and Neville, J. (1995). Images of Rosie: a content analysis of women in american magazine advertising, 1940-1946. *Journalism & Mass Communication Quaterly,* 72, 216-227.
Met dit onderzoek wilden ze onder andere nagaan of het beeld dat in advertenties van vrouwen wordt opgeroepen in de jaren 1940 (vlak voor de deelname van de VS aan WOII), 1943 (in WOII) en 1946 (na WOII), verschillend was. Uit elk van de drie jaren werden aselect een aantal advertenties uit tijdschriften gekozen waarin de vrouw werd neergezet als 'kostwinster', 'huisvrouw/moeder', 'geen beroepsrol' dan wel als 'militair/vrijwilligster'.
De resultaten zijn samengevat in de onderstaande tabel.

rol * jaar Crosstabulation

Count

		jaar			
		1940	1943	1946	Total
rol	kostwinner	17	49	34	100
	huisvrouw/moeder	119	77	156	352
	geen beroepsrol	196	97	290	583
	militair/vrijwilligster	0	33	0	33
	Total	332	256	480	1068

a. Voer een toetsing uit die helpt een antwoord te formuleren op de onderzoeksvraag. Hanteer een significantieniveau van 0.05. Voer de toetsing uit net de aanpak met een kritiek gebied. Volg de bekende stappen.
b. In onderdeel a is vastgesteld dat er samenhang bestaat tussen het jaar waarin de advertentie is verschenen en de rol die de vrouw erin speelt. Bereken nu rij- of kolompercentages (maak zelfs een verstandige keus!) en geef een tentatieve omschrijving van de *aard* van de genoemde samenhang.
c. **(Voor het practicum!)** Voer de toetsing uit met SPSS. Voer zelf de gegevens in. Maak daarbij gebruik van 'Data-> Weght Cases'. Maak 'pie charts' op basis waarvan de aard van de samenhang zichtbaar wordt gemaakt.

Opgave 2

Voer de toetsingen in de onderdelen van deze opgave daar waar dat mogelijk is zowel uit met de χ^2-toets uit hoofdstuk 9 als met de 'z-toets' uit hoofdstuk 8. Voer bij elk van de volgende onderdelen één of twee toetsingen uit met een significantieniveau van 0.05. Kies telkens de aanpak met overschrijdingskansen. Volg de bekende stappen.

a. We vermoeden dat een bepaalde afwijking bij vegetariërs procentueel minder voorkomt dan bij mensen die vlees eten. We kiezen aselect 500 vegetariërs; bij 7% van hen komt de afwijking voor. In een aselecte steekproef van 1200 vleeseters komt bij 9% de afwijking voor.
b. We willen weten of het percentage mensen dat een bepaalde afwijking heeft bij de vegetariërs afwijkt van het percentage bij mensen die vlees eten. We kiezen aselect 500 vegetariërs; bij 7% van hen komt de afwijking voor. In een aselecte steekproef van 1200 vleeseters komt bij 9% de afwijking voor.
c. We willen weten of het percentage mensen dat een bepaalde afwijking heeft in drie populaties gelijk zijn of niet. De eerste populatie bestaat uit de mensen die geregeld zowel vlees als vis eten. De twee populatie bestaat uit de mensen die wel geregeld vis eten, maar die nooit vlees eten. De derde populatie bestaat uit de mensen die vlees noch vis eten. Uit deze populaties worden aselect respectievelijk 1200, 500 en 200 mensen gekozen. In deze drie steekproeven zijn de percentages mensen met de bedoelde afwijking achtereenvolgens gelijk aan 9%, 7% en 5%.

Opgave 3

We beschouwen in deze opgave een onderzoek waarbij een aselecte steekproef wordt gekozen uit de populatie van alle studenten aan een zeer grote universiteit. We stellen van elke student vast aan welke faculteit hij studeert (Natuurwetenschappen, Sociale Wetenschappen en Economie) en op welk van de drie parijen hij gestemd heeft bij de laatste verkiezingen.

a.Vul in de onderstaande tabel aantallen in. Doe dit op een zodanige manier dat de χ^2-toets voor onafhankelijkheid een overschrijdingskans van 1.000 geeft. Verifieer je antwoord.

	Natuurweten-schappen	Sociale wetenschappen	Economie	
Partij A	20	….	….	….
Partij B	….	….	75	....
Partij C	….	….	….	….
	100	200	150	450

b.Vul in de onderstaande tabel aantallen in. Doe dit op een zodanige manier dat de χ^2-toets voor onafhankelijkheid een overschrijdingskans van 0.000 (dus: kleiner dan 0.0005) geeft. Verifieer je antwoord.

	Natuurweten-schappen	Sociale wetenschappen	Economie	Totaal
Partij A	20	….	….	….
Partij B	….	….	75	....
Partij C	….	….	….	….
	100	200	150	450

Opgave 4

Een rijinstructeur wil aantonen dat de kans om voor het theorie-examen te slagen groter is als de nieuwe lesmethode B gevolgd wordt, dan wanneer de oude lesmethode A gevolgd wordt.
a. Bedenk zelf steekproefgegevens, waarbij 70% van de kandidaten slaagt bij lesmethode B en 50% bij lesmethode A. Zorg er echter voor zelfs bij een significantieniveau van 0.10 niet geconcludeerd mag worden dat de slaagkans bij lesmethode B hoger is dan bij A.
b. Bedenk zelf steekproefgegevens, waarbij 54% van de kandidaten slaagt bij lesmethode B en 50% bij de kandidaten van lesmethode A. Zorg er echter voor zelfs bij een significantieniveau van 0.01 wel geconcludeerd mag worden dat de slaagkans bij lesmethode B hoger is dan bij A.
c. Wat kun je (vrijwel) zonder rekenwerk zeggen over het 95% betrouwbaarheidsinterval voor $p_B - p_A$ voor de gegevens uit onderdeel a? (Geef het middelpunt. Zal het getal 0 erin liggen? Schets het interval.) Bereken het interval.
d. Wat kun je (vrijwel) zonder rekenwerk zeggen over het 95% betrouwbaarheidsinterval voor $p_B - p_A$ voor de gegevens uit onderdeel b? (Geef het middelpunt. Zal het getal 0 erin liggen? Schets het interval.) Bereken het interval.

Opgave 5

We willen met de χ^2-toets nagaan of er samenhang bestaat tussen het geslacht en de partij (A, B, C of D) waarop men in een bepaalde populatie stemt. We kiezen daartoe een aselekte steekproef van n personen en stellen op basis daarvan een kruistabel ('contingency table') op. Het kan worden aangetoond dat het onderscheidend vermogen van de χ^2-toets gelijk is aan 0.19 als $\alpha = 0.05$, n = 50 en als de kansen op elk van de cellen zijn als weergegeven in de volgende tabel:

	PARTIJ A	PARTIJ B	PARTIJ C	PARTIJ D	TOTAAL
MAN	0.12	0.20	0.08	0.10	0.50
VROUW	0.20	0.12	0.08	0.10	0.50

Bij de onderdelen a t/m f hieronder wijzigen we telkens één of meer van bovenstaande getallen (de gewijzigde getallen zijn onderstreept).
U moet telkens aangeven of:

I. dat de kans dat de nulhypothese wordt verworpen kleiner is dan 0.19, òf
II. dat de kans dat de nulhypothese wordt verworpen gelijk is aan 0.19, òf
III. dat de kans dat de nulhypothese wordt verworpen groter is dan 0.19.

Geef uw antwoord door achter elk van de letters a t/m f een van de romeinse cijfers I, II of III te plaatsen. Geef telkens een motivering voor de gemaakte keuze.

a.

PARTIJ	PARTIJ	PARTIJ	PARTIJ	TOTAAL

	A	B	C	D	
MAN	0.12	0.20	0.05	0.13	0.50
VROUW	0.20	0.12	0.13	0.05	0.50

$\alpha = 0.10$ en $n = 100$

b.	PARTIJ A	PARTIJ B	PARTIJ C	PARTIJ D	TOTAAL
MAN	0.12	0.20	0.08	0.10	0.50
VROUW	0.12	0.20	0.08	0.10	0.50

$\alpha = 0.05$ en $n = 50$

c.	PARTIJ A	PARTIJ B	PARTIJ C	PARTIJ D	TOTAAL
MAN	0.12	0.20	0.05	0.13	0.50
VROUW	0.20	0.12	0.13	0.05	0.50

$\alpha = 0.05$ en $n = 50$

d.	PARTIJ A	PARTIJ B	PARTIJ C	PARTIJ D	TOTAAL
MAN	0.12	0.20	0.08	0.10	0.50
VROUW	0.20	0.12	0.08	0.10	0.50

$\alpha = 0.05$ en $n = 30$

e.	PARTIJ A	PARTIJ B	PARTIJ C	PARTIJ D	TOTAAL
MAN	0.12	0.20	0.08	0.10	0.50
VROUW	0.20	0.12	0.08	0.10	0.50

$\alpha = 0.01$ en $n = 50$

f.	PARTIJ	PARTIJ	PARTIJ	PARTIJ	TOTAAL

	A	B	C	D	
MAN	0.20	0.12	0.08	0.10	0.50
VROUW	0.12	0.20	0.08	0.10	0.50

$\alpha = 0.05$ en n = 50

g.	PARTIJ A	PARTIJ B	PARTIJ C	PARTIJ D	TOTAAL
MAN	0.16	0.04	0.12	0.08	0.40
VROUW	0.24	0.06	0.18	0.12	0.60

α = 0.10 en n = 5000

Opgave 6 (De uitvoer met SPSS produceren is werk voor het practicum)

Een rijinstructeur wil aantonen dat de kans om voor het theorie-examen te slagen groter is als de nieuwe lesmethode B gevolgd wordt, dan wanneer de oude lesmethode A gevolgd wordt. Er worden aan lesmethode A 40 en aan lesmethode B 50 aselect gekozen leerlingen toegewezen. Bij lesmethode A slager er 21 personen; bij lesmethode B slagen er 32. Is dit een overtuigend bewijs dat de slaagkans bij de nieuwe lesmethode inderdaad hoger is dan bij de oude? Toets met een significantieniveau van 0.05.

a. Formuleer de beide hypothesen.
b. Welke toetsgrootheid gebruik je en voor welke waarden ervan verwerp je de nulhypothese?
c. Zet de data in een SPSS-bestand. Gebruik daarbij ‘weight cases’. Zorg dat je daarna de volgende uitvoer krijgt:

resultaat * lesmethode Crosstabulation

			lesmethode		
			A	B	Total
resultaat	geslaagd	Count	21	32	53
		% within lesmethode	52.5%	64.0%	58.9%
	gezakt	Count	19	18	37
		% within lesmethode	47.5%	36.0%	41.1%
Total		Count	40	50	90
		% within lesmethode	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.214[b]	1	.271		
Continuity Correction[a]	.785	1	.376		
Likelihood Ratio	1.213	1	.271		
Fisher's Exact Test				.290	.188
Linear-by-Linear Association	1.200	1	.273		
N of Valid Cases	90				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 16.44.

Bepaal nu de waarde van de toetsgrootheid z door informatie uit beide bovenstaande tabellen te combineren. Gebruik blz. 411 (oud: blz. 507-508).

d. Wordt de nulhypothese wel of niet verworpen? Motiveer je antwoord.
e. Formuleer de conclusie in gewone woorden.
f. Bereken de waarde van de toetsgrootheid ook nog eens zonder de SPSS-uitvoer te gebruiken.

Opgave 7 (Voor het practicum!)
We willen weten of de houding ten opzichte van euthanasie van Italianen afwijkt van de houding van Nederlanders. We leggen daarom de volgende stelling voor aan 200 aselect gekozen Nederlanders en aan 200 aselect gekozen Italianen: ‘Aan ongeneselijk zieke patiënten moet de mogelijkheid tot euthanasie geboden worden.’ Hierop kan geantwoord worden op een ‘vijfpuntschaal’ van 1 (‘helemaal oneens’) tot en met 5 (‘helemaal eens’). Op de gegevens die dit onderzoek oplevert kunnen onder andere de volgende toetsen worden toegepast:

I. De tweezijdige rangsomtoets van Wilcoxon.
II. De eenzijdige ransomtoets voor Wilcoxon waarbij verworpen wordt als de som van de rangnummers van de Nederlanders erg hoog is.
III. De chi-kwadraat-toets voor onafhankelijkheid.

Bij de keuze uit een van deze toetsen, is het verstandig te letten op de afwijkingen van de nulhypothese die je vermoedt. Je kiest een toets die een hoog onderscheidend vermogen heeft ten opzichte van de vermoede afwijkingen ten opzichte van de nulhypothese.

Hieronder staan drie mogelijke ‘vermoede afwijkingen van de nulhypothese’.

a. De scores van de Nederlanders vallen in het algemeen hoger uit dan die van de Italianen (Nederlanders staan positiever ten opzicht van euthanasie dan Italianen.)
b. In één van beide landen vallen de scores in het algemeen hoger uit dan in het andere land.
c. De beide populatieverdelingen verschillen op één of andere manier.

A. Geef bij elk van deze drie mogelijke afwijkingen van de nulhypothese aan welk van de drie genoemde toetsen je zou kiezen. Motiveer je antwoord.
B. Bedenk zelf gegevens waarbij de chi-kwadraat-toets een overschrijdingskans groter dan 0.10 oplevert, terwijl de tweezijdige toets van Wilcoxon een overschrijdingskans kleiner dan 0.05 geeft.
C. Bedenk zelf gegevens waarbij de chi-kwadraat-toets een overschrijdingskans kleiner dan 0.01 oplevert, terwijl de tweezijdige toets van Wilcoxon een overschrijdingskans van 1 geeft.

Antwoorden uitgewerkte opgaven week 6

Antwoord opgave 1a

1. (NB: Uit de opgave blijkt dat er sprake is van het model dat beschreven wordt op pagina 421, niet van dat op pagina422) De drie populaties zijn de drie verzamelingen van alle advertenties uit de genoemde jaren waarin de vrouw wordt weergegeven in één van de genoemde rollen. Model: uit elk van de drie populaties is een aselecte steekproef getrokken. De drie steekproeven zijn onafhankelijk.
2. H_0: ‘in de drie populaties (jaren) zijn de procentuele verdelingen over de vier rollen gelijk’; H_a: ‘de populatieverdelingen zijn niet alle drie gelijk’
3. $X^2 = \sum \frac{(O-E)^2}{E}$, E = verwachte aantal, O = geobserveerde aantal
4. Onder H_0 is de toetsgrootheid chi-kwadraat verdeeld met df=6.
5. Verwerp de nulhypothese als $X^2 = \sum \frac{(O-E)^2}{E} \geq 12.59$
6. In onderstaande tabel staan onder de geobserveerde aantallen de verwachte aantallen (moet je zelf kunnen uitrekenen!)

rol * jaar Crosstabulation

			jaar			
			1940	1943	1946	Total
rol	kostwinner	Count	17	49	34	100
		Expected Count	31.1	24.0	44.9	100.0
	huisvrouw/moeder	Count	119	77	156	352
		Expected Count	109.4	84.4	158.2	352.0
	geen beroepsrol	Count	196	97	290	583
		Expected Count	181.2	139.7	262.0	583.0
	militair/vrijwilligster	Count	0	33	0	33
		Expected Count	10.3	7.9	14.8	33.0
	Total	Count	332	256	480	1068
		Expected Count	332.0	256.0	480.0	1068.0

$X^2 = \frac{(17-31.1)^2}{31.1} + \dots + \frac{(0-14.8)^2}{14.8} = 158.6$ (NB: In SPSS-uitvoer misschien weergegeven als ‘1.586 E2’)

7. Omdat de gevonden waarde van de toetsgrootheid in het kritieke gebied ligt, wordt de nulhypothese verworpen.
8. Op basis van deze steekproefgegevens kunnen we concluderen dat de procentuele verdeling over de verschillende rollen (in de populaties van advertenties!) in de drie jaren niet gelijk zijn.

Antwoord opgave 1b

rol * jaar Crosstabulation

% within jaar

		jaar			
		1940	1943	1946	Total
rol	kostwinner	5.1%	19.1%	7.1%	9.4%
	huisvrouw/moeder	35.8%	30.1%	32.5%	33.0%
	geen beroepsrol	59.0%	37.9%	60.4%	54.6%
	militair/vrijwilligster		12.9%		3.1%
	Total	100.0%	100.0%	100.0%	100.0%

Alleen in het oorlogsjaar wordt de vrouw neergezet als ‘militair/vrijwilligster’. Ook komen in 1943 meer advertenties voor waarin de vrouw wordt weergegeven als ‘kostwinner’. In de beide ander jaren vallen de advertenties vooral meer in de categorie ‘geen beroepsrol’.

<u>Antwoord opgave 1c</u>

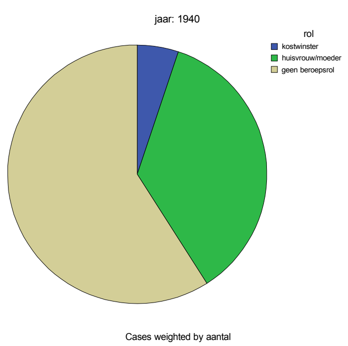
Chi-Square Tests

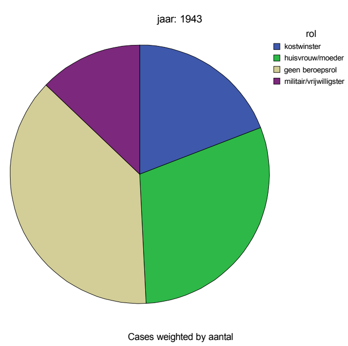
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	158.878[a]	6	.000
Likelihood Ratio	145.012	6	.000
Linear-by-Linear Association	.000	1	.984
N of Valid Cases	1071		

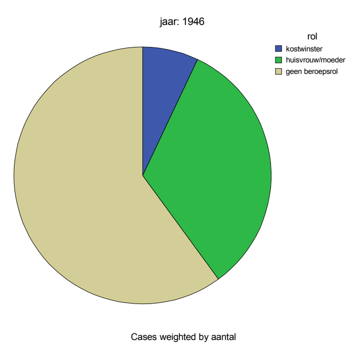
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.89.

rol * jaar Crosstabulation

			jaar			Total
			1940	1943	1946	
rol	kostwinster	Count	17	49	34	100
		Expected Count	31.0	23.9	45.1	100.0
		% within jaar	5.1%	19.1%	7.0%	9.3%
	huisvrouw/moeder	Count	119	77	159	355
		Expected Count	110.0	84.9	160.1	355.0
		% within jaar	35.8%	30.1%	32.9%	33.1%
	geen beroepsrol	Count	196	97	290	583
		Expected Count	180.7	139.4	262.9	583.0
		% within jaar	59.0%	37.9%	60.0%	54.4%
	militair/vrijwilligster	Count	0	33	0	33
		Expected Count	10.2	7.9	14.9	33.0
		% within jaar	.0%	12.9%	.0%	3.1%
Total		Count	332	256	483	1071
		Expected Count	332.0	256.0	483.0	1071.0
		% within jaar	100.0%	100.0%	100.0%	100.0%







<u>Antwoord opgave 2</u>

In onderdeel a is de chi-kwadraat-toets niet geschikt, omdat er de alternatieve hypothese eenzijdig is. In onderdeel c kan de ‘z-toets’ uit paragraaf 8.2 niet gebruikt worden omdat het niet het vergelijken van *twee* succeskansen betreft.

<u>Onderdeel a</u>

1. Zie bladzijde 382) Er wordt een EAS van omvang n_1=500 getrokken uit een grote populatie van vegetariërs met fractie afwijkingen p_1 en een EAS van omvang n_2=1200 getrokken uit een grote populatie van vleeseters met fractie afwijkingen p_2. **De beide steekproeven zijn onafhankelijk.** *NB: Als de populaties niet erg groot zijn, moet aangenomen worden dat de EAS getrokken worden met teruglegging.*

2. $H_o : p_1 \geq p_2$, $H_a : p_1 < p_2$. Ik kies α =0.05.

3. $Z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{500}+\frac{1}{1200}\right)}}$. *NB Kleine waarden van deze toetsgrootheid pleiten voor de H_a.*

(Als $Z = \dfrac{\hat{p}_2 - \hat{p}_1}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{500}+\frac{1}{1200}\right)}}$ *gebruikt wordt als toetsgrootheid, dan pleiten* natuurlijk grote waarden voor Ha.)

4. Als de beide slaagkansen gelijk zijn, dan is de toetsgrootheid standaardnormaal verdeeld.
5. $\hat{p}_1 = \frac{35}{500}$, $\hat{p}_2 = \frac{108}{1200}$, $\hat{p} = \frac{35+108}{500+1200}$. Invullen in de formule in stap 3 geeft: z=-1.35.
6. Bepaal de overschrijdingskans. $P = P(Z \le -1.36) = 0.0885$.
7. Omdat de overschrijdingskans groter is dan het significantieniveau, wordt de nulhypothese niet verworpen.
8. Er is niet aangetoond dat de fractie afwijkingen bij de vegetariërs kleiner is dan bij de vleeseters.

Onderdeel b met 'z-toets'

1. (Zie bladzijde 382) Er wordt een EAS van omvang n_1=500 getrokken uit een grote populatie van vegetariërs met fractie afwijkingen p_1 en een EAS van omvang n_2=1200 getrokken uit een grote populatie van vleeseters met fractie afwijkingen p_2. **De beide steekproeven zijn onafhankelijk.** *NB: Als de populaties niet erg groot zijn, moet aangenomen worden dat de EAS getrokken worden met teruglegging.*
2. $H_o: p_1 = p_2$, $H_a: p_1 \neq p_2$. Ik kies $\alpha = 0.05$.
3. $Z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{500}+\frac{1}{1200}\right)}}$. *NB Hele kleine en hele grote waarden van deze toetsgrootheid pleiten nu voor* H_a.
4. Als de beide slaagkansen gelijk zijn, dan is de toetsgrootheid standaardnormaal verdeeld.
5. $\hat{p}_1 = \frac{35}{500}$, $\hat{p}_2 = \frac{108}{1200}$, $\hat{p} = \frac{35+108}{500+1200}$. Invullen in de formule in stap 3 geeft: z=-1.36.
6. $P = P(Z \le -1.36) + P(Z \ge 1.36) = 2 * 0.0885 = 0.1770$.
7. Omdat de overschrijdingskans groter is dan het significantieniveau, wordt de nulhypothese niet verworpen.
8. Er is niet aangetoond dat de fractie afwijkingen bij de vegetariërs afwijkt van de fractie bij de vleeseters.

Onderdeel b met 'chi-kwadraat-toets'

1. (Zie bladzijde 382 en 421) Er wordt een EAS van omvang n_1=500 getrokken uit een grote populatie van vegetariërs met fractie afwijkingen p_1 en een EAS van omvang n_2=1200 getrokken uit een grote populatie van vleeseters met fractie

afwijkingen p_2. **De beide steekproeven zijn onafhankelijk.** *NB: Als de populaties niet erg groot zijn, moet aangenomen worden dat de EAS getrokken worden met teruglegging.*

2. $H_o: p_1 = p_2$, $H_a: p_1 \neq p_2$. Ik kies $\alpha = 0.05$.
3. $X^2 = \sum \frac{(O-E)^2}{E}$, E = verwachte aantal, O = geobserveerde aantal. *NB Hele grote waarden van deze toetsgrootheid pleiten voor H_a.*

4. Als de beide slaagkansen gelijk zijn, dan is de toetsgrootheid chi-kwadraat verdeeld met df=1.

5. $X^2 = 1.85$ (Ik laat het rekenwerk deze keer niet zien.)

6. Met tabel F: $0.15 < P < 0.20$.
7. Omdat de overschrijdingskans groter is dan het significantieniveau, wordt de nulhypothese niet verworpen.
8. Er is niet aangetoond dat de fractie afwijkingen bij de vegetariërs afwijkt van de fractie bij de vleeseters.

Onderdeel c

1. Model: uit elk van de drie populaties is een aselecte steekproef getrokken. De drie steekproeven zijn onafhankelijk.
2. H_0: 'in de drie populaties zijn de fracties personen met de afwijking gelijk'; H_a: 'de populatiefracties zijn niet alle drie gelijk'.
3. $X^2 = \sum \frac{(O-E)^2}{E}$, E = verwachte aantal, O = geobserveerde aantal
4. Onder H_0 is de toetsgrootheid chi-kwadraat verdeeld met df=2.
5. In onderstaande tabel staan onder de geobserveerde aantallen de verwachte aantallen (moet je zelf kunnen uitrekenen!)

groep * afwijking Crosstabulation

			afwijking		
			ja	nee	Total
groep	vlees en vis	Count	108	1092	1200
		Expected Count	96.6	1103.4	1200.0
	alleen vis	Count	35	465	500
		Expected Count	40.3	459.7	500.0
	vlees noch vis	Count	10	190	200
		Expected Count	16.1	183.9	200.0
	Total	Count	153	1747	1900
		Expected Count	153.0	1747.0	1900.0

$$X^2 = \frac{(108-96.6)^2}{96.6} + \frac{(190-183.9)^2}{183.9} = 4.72$$

6. Met tabel F: *0.05<P<0.10*
7. De nulhypothese wordt niet verworpen, want de overschrijdingskans is groter dan het significantieniveau van 0.05.
8. Er is niet aangetoond dat de drie populatiefracties verschillen.

Antwoord opgave 3

a. ‘overschrijdingskans 1’ ⇔ ’$X^2=0$’ ⇔ ’geobserveerde aantal in elke cel gelijk aan verwachte aantal’ ⇔ ’in alle kolommen dezelfde kolompercentages’ ⇔ ’in alle rijen dezelfde rijpercentages’. Er is maar één oplossing:

	Natuurweten-schappen	Sociale wetenschappen	Economie	Totaal
Partij A	20	40	30	90
Partij B	50	100	75	225
Partij C	30	60	45	135
	100	200	150	450

b. Er zijn heel veel oplossingen. Zorg ervoor dat de kolompercentages in de verschillende kolommen flink afwijken!

	Natuurweten-schappen	Sociale wetenschappen	Economie	
Partij A	20	60	50	130
Partij B	10	60	75	145
Partij C	70	80	25	175
	100	200	150	450

Antwoord opgave 4

a. *Bij hele kleine steekproeven kunnen zelfs hele grote verschillen in fracties geslaagden nog op toeval berusten! Ik neem $n_1=n_2=10$ en dus $X_1=7$ en $X_2=5$.*

$z = \frac{0.70-0.50}{\sqrt{0.6(1-0.6)(\frac{1}{10}+\frac{1}{10})}}$ =0.92. De eenzijdige overschrijdingskans is 0.1788.

b. *Bij hele grote steekproeven kunnen zelfs hele kleine verschillen in fracties geslaagden niet meer op toeval berusten! Ik neem $n_1=n_2=10000$ en dus $X_1=5400$ en $X_2=5000$.* $z = \frac{0.54-0.50}{\sqrt{0.52(1-0.52)(\frac{1}{10000}+\frac{1}{10000})}}$ =5.66. De eenzijdige overschrijdingskans is kleiner dan 0.0002.

c. *Het getal 0 zal wel in het interval liggen. Uit onderdeel a weten we immers dat er geen enkel bewijs is tegen de nulhypothese dat de beide slaagkansen gelijk zijn! Het interval heeft 0.20 als middelpunt en zal vrij lang zijn.* Het interval loopt van $(0.70-0.50) - 1.96\sqrt{\frac{0.50*0.50}{10}+\frac{0.70*0.30}{10}} = 0.20-0.42 = -0.22$ tot

0.20+0.42=0.62. *We weten dus nog heel weinig over het verschil tussen beide slaagkansen!*

d. *Het getal 0 zal niet in het interval liggen. Uit onderdeel b weten we immers dat er een sterk bewijs is tegen de nulhypothese dat de beide slaagkansen gelijk zijn! Het interval heeft 0.04 als middelpunt en zal heel kort zijn.* Het interval loopt van $(0.54-0.50)-1.96\sqrt{\frac{0.54*0.46}{10000}+\frac{0.50*0.50}{10000}}=0.04-0.01=0.03$ tot 0.04+0.01=0.05. *De toetsing uit b leert ons alleen dat de slaagkans bij methode B groter is dan bij methode B. Het betrouwbaarheidsinterval leert ons hoe groot dat verschil ongeveer is.*

<u>Antwoord opgave 5</u>

b. Hier bestaat er helemaal geen samenhang tussen ‘geslacht’ en ‘partijkeuze’. De nulhypothese is waar. De kans dat H_0 verworpen wordt is dus gelijk aan het significantieniveau 0.05.

c. De samenhang tussen beide variabelen is hier sterker dan in de uitgangstabel. Omdat een sterke samenhang gemakkelijker is te ontdekken dan een zwakke, is het onderscheidingsvermogen groter dan 0.19.

d. Met een kleinere steekproef kun je minder goed beslissen. Omdat de kans op een fout van de eerste soort gelijk blijft (aan 0.05), zal de kans op een fout van de tweede soort toenemen. Dus neemt het onderscheidingsvermogen (= 1- kans op fout van tweede soort) af. De kans op verwerpen van H_0 is dus kleiner dan 0.19.

e. Als de kans op een fout van de eerste soort kleiner wordt (en alle andere factoren blijven gelijk), zal de kans op een fout van de tweede soort toenemen. Dus neemt het onderscheidingsvermogen (= 1- kans op fout van tweede soort!) af. De kans op verwerpen van H_0 is dus kleiner dan 0.19.

f. Je kunt de verandering in de tabel opvatten als het verhangen van de labels. De partij waarop de mannen het meeste stemmen kreeg eerst het label ‘B’ en nu het label ‘A’. Bij de partij waarop de vrouwen het meeste stemmen gebeurt het omgekeerde. Intuïtief begrijpen we dat het onderscheidingsvermogen 0.19 blijft.

g. In deze tabel geldt: ‘De kans op elke cel is het product van de kans op de betreffende rij en de kans op de betreffende kolom’. ‘Geslacht’ en ‘partijkeuze’ zijn dus onafhankelijk. De kans dat de nulhypothese toch verworpen wordt is dus gelijk aan het significantieniveau 0.10.

Tenslotte: bij onderdeel a is het onderscheidingsvermogen groter dan 0.19. Er treden 3 veranderingen op die elk afzonderlijk al tot een verhoging van het onderscheidingsvermogen leiden!

<u>Antwoord opgave 6</u>

a. $H_0: p_B \le p_A \quad H_a: p_B > p_A$

b. $z=\dfrac{\hat{p}_b-\hat{p}_a}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{50}+\frac{1}{40})}}$ Verwerp de nulhypothese als $z \ge 1.645$.

c. $\chi^2 = 1.214$. Dus: $z = \sqrt{1.214} = 1.10$ of $z = -\sqrt{1.214} = -1.10$ (zie bladzijde 411 (oud: 507-508)). In de eerste tabel zien we dat $\hat{p}_b = 0.64$ en $\hat{p}_a = 0.525$. Dus is z positief. Conclusie: z=1.10 (Zorg dat je dit ook met de hand uit kunt rekenen!).
d. De nulhypothese wordt niet verworpen.
e. Op basis van deze steekproefgegevens kan niet geconcludeerd worden dat de slaagkans bij lesmethode B hoger is dan bij lesmethode A.

Antwoord opgave 7A
Als er keuze is uit een aantal toetsen (met dezelfde kans op een fout van de eerste soort), dan is het verstandig een toets te kiezen die een zo hoog mogelijk onderscheidingsvermogen zal hebben bij de ‘vermoede afwijkingen van de nulhypothese’. De chi-kwadraat-toets kijkt of er cellen zijn waarin de geobserveerde aantallen erg afwijken van de verwachte. Deze toets zal dus wel geschikt zijn als je enige ‘samenhang’ (van welke ‘vorm’ dan ook!) verwacht tussen ‘nationaliteit’ en ‘score’. Kortom: toets III is geschikt voor ‘vermoede afwijking c’.
Toets II is natuurlijk geschikt voor ‘vermoede afwijking a’ en toets I voor ‘vermoede afwijking b’.

Antwoord opgave 7B
Zie het bestand ‘opgave7B.sav’.

nationaliteit * score Crosstabulation

Count

		score					
		1	2	3	4	5	Total
nationaliteit	Nederlanders	40	40	40	40	40	200
	Italianen	55	40	40	40	25	200
	Total	95	80	80	80	65	400

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5.830[a]	4	.212
Likelihood Ratio	5.871	4	.209
Linear-by-Linear Association	4.540	1	.033
N of Valid Cases	400		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 32.50.

Ranks

	nationaliteit	N	Mean Rank	Sum of Ranks
score	Nederlanders	200	212.50	42500.00
	Italianen	200	188.50	37700.00
	Total	400		

Test Statistics[a]

	score
Mann-Whitney U	17600.000
Wilcoxon W	37700.000
Z	-2.121
Asymp. Sig. (2-tailed)	.034

a. Grouping Variable: nationaliteit

Antwoord opgave 7C
Zie het bestand ‘opgave7C.sav’.

nationaliteit * score Crosstabulation

Count

		score					
		1	2	3	4	5	Total
nationaliteit	Nederlanders	5	20	150	20	5	200
	Italianen	75	20	10	20	75	200
	Total	80	40	160	40	80	400

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2.450E2	4	.000
Likelihood Ratio	293.988	4	.000
Linear-by-Linear Association	.000	1	1.000
N of Valid Cases	400		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 20.00.

Ranks

	nationaliteit	N	Mean Rank	Sum of Ranks
score	Nederlanders	200	200.50	40100.00
	Italianen	200	200.50	40100.00
	Total	400		

Test Statistics[a]

	score
Mann-Whitney U	20000.000
Wilcoxon W	40100.000
Z	.000
Asymp. Sig. (2-tailed)	1.000

a. Grouping Variable: nationaliteit

Overzicht regels en technieken uit Statistiek 2

I . Hoe hoog zijn gemiddeld de scores op deze kwantitatieve variabele (bijv.: intelligentietest) in deze populatie?

- Veronderstelling: normale populatieverdeling en/of grote steekproef
- σ bekend (chapter 6): 'z-toets' + power, betrouwbaarheidsinterval voor μ
- σ onbekend (par. 7.1): 't-toets', betrouwbaarheidsinterval voor μ

II. Vergelijken van de (gemiddelde) hoogte van de scores (op kwantitatieve variabele) in twee groepen/populaties/condities

	Normale verdeling(en) verondersteld in populatie(s)	Geen normaliteit nodig; 'niet-parametrische procedures' (intervallen niet behandeld)
Onafhankelijke steekproeven	t-toetsen voor onafhankelijke steekproeven + betr. intervallen (par. 7.2)	Rangsomtoets van Wilcoxon (SPSS: Mann-Whitney) par. 15.1
Gekoppelde paren	t-toets voor gekoppelde paren + betr. interval (par. 7.1)	▪ Rangtekentoets (15.2) ▪ Tekentoets (7.1)

Bij samengestelde t-toets voor onafhankelijke steekproeven: onderscheidingsvermogen en steekproefgrootte. Tabellen!

V. 'Inferenties' over één of meer populatiefracties/ succeskansen

- Eén fractie (toetsen en betrouwbaarheidsintervallen): par. 8.1
- Steekproefgrootte bij het schatten van een fractie: $n = \left(\frac{z^*}{2m}\right)^2$ (blz. 374)
- Hierbij:
  - Bij betrouwbaarheidsintervallen: m = 'halve lengte van interval'
  - Bij puntschatting: m= 'maximale afstand tussen schatter $\hat{p}$ en parameter *p*'
  - Preciezer: $P(|\hat{p} - p| \le m) = 0.95$ als $n = \left(\frac{1.96}{2m}\right)^2$
- Vergelijken van twee fracties (toetsen en betrouwbaarheidsintervallen): par. 8.2

- 'Bestaat er (in de populatie) enige samenhang tussen deze twee categorische variabelen?'

- Chi-kwadraat-toets voor onafhankelijkheid (chapter 9)
- Hieronder ook: vergelijken van drie of meer fracties (en bij tweezijdige toetsen vergelijken van twee fracties (zie blz 411-412).

Regels en formules voor tentamen van Statistiek 2 en Statistiek 3

Hoofdstuk 7

Regel 1

Veronderstel dat een EAS van omvang n is getrokken uit een populatie en dat de populatieverdeling van de scores een normale verdeling is met (onbekend) populatiegemiddelde μ en onbekende populatiestandaardafwijking σ . Dan heeft de toevalsvariabele

$T = \dfrac{\overline{X} - \mu}{s/\sqrt{n}}$ de t-verdeling met *df*=n-1.

Regel 2

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

Regel 3

Onder de veronderstellingen van ‘Regel 1’:

- Betrouwbaarheidsinterval voor μ : $\bar{x} \pm t^* s/\sqrt{n}$
- Toetsgrootheid voor het toetsen van $H_0 : \mu = 50$: $T = \dfrac{\overline{X} - 50}{s/\sqrt{n}}$

Regel 4

(*t*-procedures voor twee onafhankelijke steekproeven; gelijke populatie-variaties)

Veronderstel dat een EAS van omvang n_1 is getrokken uit een normale populatie met onbekende verwachting μ_1 en onbekende standaardafwijking σ en dat een EAS van omvang n_2 is getrokken uit een andere normale populatie met onbekende verwachting μ_2 en ook met standaardafwijking σ .Neem verder aan dat beide steekproeven onafhankelijk gekozen zijn.

- Het betrouwbaarheidsinterval voor $\mu_1 - \mu_2$ heeft dan de vorm

  $(\bar{x}_1 - \bar{x}_2) \pm t^* s_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$ Hier komt het getal *t** uit de $t(n_1 + n_2 - 2)$-verdeling.
- Om de hypothese $H_0 : \mu_1 = \mu_2$ te toetsen, berekent men de twee-steekproeven toetsgrootheid $t = \dfrac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ en bepaalt overschrijdingskansen of kritieke waarden met de $t(n_1 + n_2 - 2)$-verdeling.
- $s_p = \sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$

.

Regel 5
t-procedures voor twee onafhankelijke steekproeven; mogelijk ongelijke populatie-varianties)

Veronderstel dat een EAS van omvang n_1 is getrokken uit een normale populatie met onbekende verwachting μ_1 en onbekende standaardafwijking σ_1 en dat een EAS van omvang n_2 is getrokken uit een andere normale populatie met onbekende verwachting μ_2 en onbekende standaardafwijking σ_2 .Neem verder aan dat beide steekproeven onafhankelijk gekozen zijn.

- Het betrouwbaarheidsinterval voor $\mu_1 - \mu_2$ heeft dan de vorm
  $(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ Hier komt het getal t^* uit de $t(df)$-verdeling. Het aantal vrijheidsgraden df benaderen we met de formule
  $$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$
  (zie bladzijde 345 boek; we gebruiken niet de vrijheidsgraden uit het hok van bladzijde 339).
- Om de hypothese $H_0 : \mu_1 = \mu_2$ te toetsen, berekent men de twee-steekproevengrootheid
  $$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

en bepaalt overschrijdingskansen of kritieke waarden met de $t(df)$-verdeling.

Onderscheidend vermogen van de samengestelde t-toets voor onafhankelijke steekproeven

- $\delta = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \frac{\mu_1 - \mu_2}{\sigma}$
- $\Delta = \frac{\text{grootste } \mu_i - \text{kleinste } \mu_i}{\sigma}$

Nuttige formule voor het bepalen van de steekproefgrootte bij het schatten van μ :

$$n = \left(\frac{z^* \sigma}{m}\right)^2$$

De toetsen van Wilcoxon

- De exacte kansverdeling van de rangsomtoetsgrootheid W wordt, voor kleine steekproeven, weergegeven in ‘Table 9’. Bij gebruik van de tabel moet de som van de rangnummers worden bepaald in de groep met het geringste aantal waarnemingen (als de aantallen verschillen!).

- Bij grotere steekproeven (en bij kleine steekproeven alleen als dat in een opgave gevraagd wordt!) wordt gebruikt dat de gestandaardiseerde versie Z van de rangsomtoetsgrootheid onder de nulhypothese bij benadering standaardnormaal verdeeld is. Hierbij is: $Z = \frac{W - n_1(n_1 + n_2 + 1)/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}}$, n_1 = ‘aantal waarnemingen in groep waarin W is bepaald’.

- De gestandaardiseerde versie Z van de toetsgrootheid van de rangtekentoets is onder de nulhypothese bij benadering standaardnormaal verdeeld. $Z = \frac{W^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$.

Regels en formules bij hoofdstuk 9 (fracties)

(Zorg ervoor dat je zelf de voorwaarden kent waaronder deze regels gelden!)

- Onder de nulhypothese $H_0 : p = p_0$, is de toetsgrootheid $Z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$

  standaardnormaal verdeeld.
- Betrouwbaarheidsinterval voor p: $\hat{p} \pm z^* \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$
- Formule voor het bepalen van steekproefgrootte bij schatten van p: $n = \left(\dfrac{z^*}{2m}\right)^2$
- Onder de nulhypothese $H_0 : p_1 = p_2$, is de toetsgrootheid

  $Z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$ standaardnormaal verdeeld. Hierbij: $\hat{p} = \dfrac{X_1 + X_2}{n_1 + n_2}$
- Betrouwbaarheidinterval voor verschil tussen twee fracties:

  $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

De chi-kwadraat-toets voor onafhankelijkheid

- $X^2 = \sum \dfrac{(waargenomenaantal - verwachtaantal)^2}{verwachtaantal}$
- $verwachtaantal = \dfrac{rijtotaal * kolomtotaal}{n}$
- df= (r-1)(c-1)

Hoofdstuk 10 (Enkelvoudige regressieanalyse)

1. Kleinste kwadratenschatters van β_1 en β_0: $b_1 = r\dfrac{s_y}{s_x} = \dfrac{\sum xy - \frac{1}{n}(\sum x)(\sum y)}{\sum x^2 - \frac{1}{n}(\sum x)^2}$ en

   $b_0 = \bar{y} - b_1 \bar{x}$.
2. residuen: $e_i = y_i - \hat{y}_i$.
3. Schatter van σ^2: $s^2 = \dfrac{\sum_{i=1}^{n} e_i^2}{n-2}$, $s = \sqrt{s^2}$

4. Betrouwbaarheidsinterval voor β_1: $b_1 \pm t^* SE_{b_1}$, met $SE(b_1) = \frac{s}{\sqrt{\sum (x_i - \bar{x})^2}}$, t^* komt uit de t-verdeling met n-2 vrijheidsgraden.

5. Betrouwbaarheidsinterval 'voor verwachte reactie':
$(b_0 + b_1 x^*) \pm t^* s \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum (x_i - \bar{x})^2}}$, t^* komt uit de t-verdeling met n-2 vrijheidsgraden.

6. Voorspellingsinterval voor toekomstige waarneming:
$(b_0 + b_1 x^*) \pm t^* s \sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum (x_i - \bar{x})^2}}$, t^* komt uit de t-verdeling met n-2 vrijheidsgraden.

7. Toetsen van de nulhypothese $H_0 : \beta_1 = 0$: de toetsgrootheid $t = \frac{b_1}{SE(b_1)}$ is onder de nulhypothese t-verdeeld met df=n-2. $SE(b_1) = \frac{s}{\sqrt{\sum (x_i - \bar{x})^2}}$. ($t = \frac{b_1}{s / \sqrt{\sum (x_i - \bar{x})^2}}$)

8. Toetsen van de nulhypothese $H_0 : \beta_1 = 0$ (alleen bij tweezijdige alternatieve hypothese): de toetsgrootheid F uit de onderstaande 'variantie-analyse tabel' is onder de nulhypothese F-verdeeld met $df_{teller} = 1$ en $df_{noemer} = n - 2$.

Bron	DF	Kwadratensom (SS=sum of sqaures)	Gemiddelde Kwadratensom (MS= mean square)	*F*
Model	*DFM*=1	*SSM*= $\Sigma(\hat{y}_i - \bar{y})^2$	$MSM = \frac{SSM}{DFM}$	*MSM/MSE*
Error	*DFE*=n-2	*SSE*= $\Sigma(y_i - \hat{y}_i)^2$	$MSE = \frac{SSE}{DFE}$	
Totaal	n-1	$\Sigma(y_i - \bar{y})^2$		

De toetsgrootheid $F = \frac{MSM}{MSE}$ uit deze tabel is het kwadraat van *t*.

9. Formules om de steekproefcorrelatie uit te rekenen (naast de formule op bladzijde 78!):

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{[\sum_{i=1}^{n}(x_i - \bar{x})^2][\sum_{i=1}^{n}(y_i - \bar{y})^2]}} = \frac{\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)/n}{\sqrt{[\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2/n][\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2/n]}}.$$

10. Er geldt: $r^2 = \frac{SSM}{SST}$ (zie regel 8).

11. Toetsen van $H_0 : \rho = 0$: de toetsgrootheid $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ is onder de nulhypothese *t*-verdeeld met *df*=n-2.

Hoofdstuk 12 Één-factor variantie-analyse)

1. De variantie-analyse tabel:

Bron	df	Kwadratensom	Gemiddelde Kwadratensom	Toetsgrootheid F
Groepen	I-1	SSG= $\sum_{i=1}^{I} n_i(\bar{x}_i - \bar{x})^2$	MSG=SSG/(I-1)	$F = \frac{MSG}{MSE}$
Error	N-I	SSE= $\sum_{i=1}^{I}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2 =$ $= \sum_{i=1}^{I}(n_i - 1)s_i^2$	MSE=SSE/(N-I)	
Totaal	N-1	SST= $\sum_{i=1}^{I}\sum_{j=1}^{n_i}(x_{ij} - \bar{x})^2$		

De toetsgrootheid voor het toetsen van $H_0: \mu_1 = \mu_2 = \mu_3$ is dus:

$$F = \frac{\sum_{i=1}^{I} n_i(\bar{x}_i - \bar{x})^2 \,/\, (I-1)}{\sum_{i=1}^{I}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2 \,/\, (N-I)} = \frac{\sum_{i=1}^{I} n_i(\bar{x}_i - \bar{x})^2 \,/\, (I-1)}{\sum_{i=1}^{I}(n_i - 1)s_i^2 \,/\, (N-I)}$$

. Onder de nulhypothese is deze toetsgrootheid F-verdeeld met $df_{teller} = I - 1$ en $df_{noemer} = N - I$.

2. (Methode van Bonferroni voor opstellen van simultane betrouwbaarheidsintervallen voor verschillen tussen verwachtingen. We beperken ons hier tot het geval van I=3 treatments) Als we betrouwbaarheidsintervallen voor $\mu_2 - \mu_1$, $\mu_3 - \mu_1$ en $\mu_3 - \mu_2$ op willen stellen met een gezamenlijke betrouwbaarheidscoëfficiënt van minstens (bijvoorbeeld) 0.94 (=1-0.06) , kan dat door voor elk van de drie verschillen apart een betrouwbaarheidsinterval op te stellen met betrouwbaarheidscoëfficiënt 1-0.06/3= 0.98. Het interval voor $\mu_2 - \mu_1$ wordt dan bijvoorbeeld berekent volgens de formule $(\bar{x}_2 - \bar{x}_1) \pm t^{**} s_p \sqrt{\frac{1}{n_2} + \frac{1}{n_1}}$.

   Hierbij geldt:

   a. t^{**} is het getal waarboven een kansmassa 0.01 (=(1-0.98)/2) ligt in de t-verdeling met df=N-I.

   b. $s_p = \sqrt{MSE}$ (wordt dus berekend op basis van de scores in alle drie de groepen!)

Hoofdstuk 13: Twee-factor variantie-analyse (formules geldig als in elke cel evenveel waarnemingen zitten!)

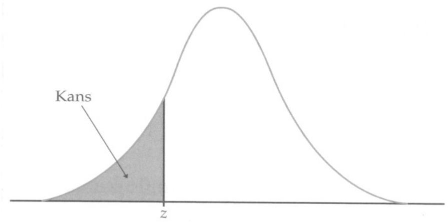
1. De ‘kolomfactor A’ kent I niveaus. De rijfactor B kent J niveaus. In elke cel r waarnemingen. $SST = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{r}(x_{ijk} - \bar{x}_{...})^2$; $SSE = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{r}(x_{ijk} - \bar{x}_{ij.})^2$;

$$SSA = Jr\sum_{i=1}^{I}(\bar{x}_{i..} - \bar{x}_{...})^2 \; ; \; SSB = Ir\sum_{j=1}^{J}(\bar{x}_{.j.} - \bar{x}_{...})^2 \; ;$$

$SSAB = \sum_{j=1}^{J} r(\bar{x}_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{x}_{...})^2$; N=IJr= ‘totaal aantal waarnemingen’

Bron	df	SS	MS	F-ratio's	P^*
A	DFA=I-1	SSA	MSA= SSA/DFA	MSA/MSE	…
B	DFB=J-1	SSB	MSB=SSB/DFB	MSB/MSE	…
AB	DFAB=(I-1)(J-1)	SSAB	MSAB=SSAB/DFAB	MSAB/MSE	…
Error	DFE=N-IJ	SSE	MSE=SSE/DFE		
Totaal	N-1	SST			

Tabellen

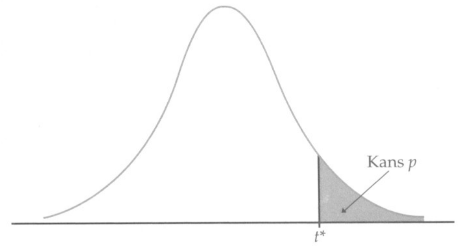


Tabelelement bij z is de kans $P(Z \leq z)$ voor standaardnormale Z.

Tabel A **Standaardnormale kansen**

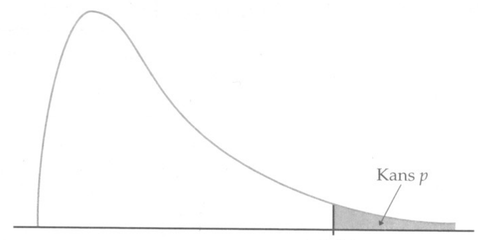
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Tabelelement voor p en C is het punt t^* met kans p op een waarde groter dan t^* en kans C op een waarde tussen $-t^*$ en t^*



Tabel D Kritieke waarden voor de t-verdeling

	Staartkans p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Betrouwbaarheidsniveau											



Tabelelement bij p is de kritiek waarde $(\chi^2)^*$ met kans p erboven.

Tabel F Kritieke waarden voor de χ^2-verdeling

	Staartkans p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2

Table C.13 Noncentral t Distribution†
Values of the noncentrality parameter for a one-sided test with $\alpha = .050$

f	β = type 2 error										
	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90
1	16.47	12.53	10.51	8.19	6.63	5.38	4.31	3.35	2.46	1.60	.64
2	6.88	5.52	4.81	3.98	3.40	2.92	2.49	2.07	1.63	1.15	.50
3	5.47	4.46	3.93	3.30	2.85	2.48	2.13	1.79	1.43	1.02	.46
4	4.95	4.07	3.60	3.04	2.64	2.30	1.99	1.67	1.34	.96	.43
5	4.70	3.87	3.43	2.90	2.53	2.21	1.91	1.61	1.29	.92	.42
6	4.55	3.75	3.33	2.82	2.46	2.15	1.86	1.57	1.26	.90	.41
7	4.45	3.67	3.26	2.77	2.41	2.11	1.82	1.54	1.24	.89	.40
8	4.38	3.62	3.21	2.73	2.38	2.08	1.80	1.52	1.22	.88	.40
9	4.32	3.58	3.18	2.70	2.35	2.06	1.78	1.51	1.21	.87	.39
10	4.28	3.54	3.15	2.67	2.33	2.04	1.77	1.49	1.20	.86	.39
11	4.25	3.52	3.13	2.65	2.31	2.02	1.75	1.48	1.19	.86	.39
12	4.22	3.50	3.11	2.64	2.30	2.01	1.74	1.47	1.19	.85	.38
13	4.20	3.48	3.09	2.63	2.29	2.00	1.74	1.47	1.18	.85	.38
14	4.18	3.46	3.08	2.62	2.28	2.00	1.73	1.46	1.18	.84	.38
15	4.17	3.45	3.07	2.61	2.27	1.99	1.72	1.46	1.17	.84	.38
16	4.16	3.44	3.06	2.60	2.27	1.98	1.72	1.45	1.17	.84	.38
17	4.14	3.43	3.05	2.59	2.26	1.98	1.71	1.45	1.17	.84	.38
18	4.13	3.42	3.04	2.59	2.26	1.97	1.71	1.45	1.16	.83	.38
19	4.12	3.41	3.04	2.58	2.25	1.97	1.71	1.44	1.16	.83	.38
20	4.12	3.41	3.03	2.58	2.25	1.97	1.70	1.44	1.16	.83	.38
21	4.11	3.40	3.03	2.57	2.24	1.96	1.70	1.44	1.16	.83	.38
22	4.10	3.40	3.02	2.57	2.24	1.96	1.70	1.44	1.16	.83	.37
23	4.10	3.39	3.02	2.56	2.24	1.96	1.69	1.43	1.15	.83	.37
24	4.09	3.39	3.01	2.56	2.23	1.95	1.69	1.43	1.15	.83	.37
25	4.09	3.38	3.01	2.56	2.23	1.95	1.69	1.43	1.15	.83	.37
26	4.08	3.38	3.01	2.55	2.23	1.95	1.69	1.43	1.15	.82	.37
27	4.08	3.38	3.00	2.55	2.23	1.95	1.69	1.43	1.15	.82	.37
28	4.07	3.37	3.00	2.55	2.22	1.95	1.69	1.43	1.15	.82	.37
29	4.07	3.37	3.00	2.55	2.22	1.94	1.68	1.42	1.15	.82	.37
30	4.07	3.37	3.00	2.54	2.22	1.94	1.68	1.42	1.15	.82	.37
40	4.04	3.35	2.98	2.53	2.21	1.93	1.67	1.42	1.14	.82	.37
60	4.02	3.33	2.96	2.52	2.19	1.92	1.66	1.41	1.13	.81	.37
100	4.00	3.31	2.95	2.50	2.18	1.91	1.66	1.40	1.13	.81	.37
∞	3.97	3.29	2.93	2.49	2.17	1.90	1.64	1.39	1.12	.80	.36

$$\Pr[\text{noncentral } t > t_{1-\alpha} \mid \delta = (\mu_1 - \mu_0)(\sqrt{n}/\sigma)] = 1 - \beta.$$

† This table is reproduced from: The power of Student's t test. *Journal of the American Statistical Association*, 1965, 60, 320–333, with the permission of the author, D. B. Owen, and the editors.

Table C.13 (*Continued*)
Values of the noncentrality parameter for a one-sided test with $\alpha = .025$

f	β = type 2 error										
	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90
1	32.83	24.98	20.96	16.33	13.21	10.73	8.60	6.68	4.91	3.22	1.58
2	9.67	7.77	6.80	5.65	4.86	4.21	3.63	3.07	2.50	1.88	1.09
3	6.88	5.65	5.01	4.26	3.72	3.28	2.87	2.47	2.05	1.57	.94
4	5.94	4.93	4.40	3.76	3.31	2.93	2.58	2.23	1.86	1.44	.87
5	5.49	4.57	4.09	3.51	3.10	2.75	2.43	2.11	1.76	1.37	.82
6	5.22	4.37	3.91	3.37	2.98	2.64	2.34	2.03	1.70	1.32	.80
7	5.06	4.23	3.80	3.27	2.89	2.57	2.27	1.98	1.66	1.29	.78
8	4.94	4.14	3.71	3.20	2.83	2.52	2.23	1.94	1.63	1.27	.77
9	4.85	4.07	3.65	3.15	2.79	2.48	2.20	1.91	1.60	1.25	.76
10	4.78	4.01	3.60	3.11	2.75	2.45	2.17	1.89	1.59	1.23	.75
11	4.73	3.97	3.57	3.08	2.73	2.43	2.15	1.87	1.57	1.22	.74
12	4.69	3.93	3.54	3.05	2.70	2.41	2.13	1.85	1.56	1.21	.74
13	4.65	3.91	3.51	3.03	2.69	2.39	2.12	1.84	1.55	1.21	.73
14	4.62	3.88	3.49	3.01	2.67	2.38	2.11	1.83	1.54	1.20	.73
15	4.60	3.86	3.47	3.00	2.66	2.37	2.09	1.82	1.53	1.19	.72
16	4.58	3.84	3.46	2.98	2.65	2.36	2.09	1.81	1.53	1.19	.72
17	4.56	3.83	3.44	2.97	2.64	2.35	2.08	1.81	1.52	1.18	.72
18	4.54	3.82	3.43	2.96	2.63	2.34	2.07	1.80	1.52	1.18	.72
19	4.52	3.80	3.42	2.95	2.61	2.33	2.06	1.80	1.51	1.17	.71
20	4.51	3.79	3.41	2.95	2.61	2.33	2.06	1.79	1.51	1.17	.71
21	4.50	3.78	3.40	2.93	2.60	2.32	2.05	1.79	1.50	1.17	.71
22	4.49	3.77	3.39	2.93	2.60	2.32	2.05	1.78	1.50	1.17	.71
23	4.48	3.77	3.39	2.93	2.59	2.31	2.05	1.78	1.50	1.17	.71
24	4.47	3.76	3.38	2.92	2.59	2.31	2.04	1.78	1.50	1.16	.71
25	4.46	3.75	3.37	2.92	2.58	2.30	2.04	1.77	1.49	1.16	.71
26	4.46	3.75	3.37	2.92	2.58	2.30	2.04	1.77	1.49	1.16	.70
27	4.45	3.74	3.36	2.91	2.58	2.30	2.03	1.77	1.49	1.16	.70
28	4.44	3.73	3.36	2.90	2.57	2.29	2.03	1.77	1.49	1.16	.70
29	4.44	3.73	3.35	2.90	2.57	2.29	2.03	1.77	1.48	1.16	.70
30	4.43	3.73	3.35	2.90	2.57	2.29	2.02	1.76	1.48	1.16	.70
40	4.39	3.69	3.32	2.87	2.55	2.27	2.01	1.75	1.47	1.15	.69
60	4.36	3.66	3.29	2.85	2.53	2.25	1.99	1.73	1.46	1.14	.69
100	4.33	3.64	3.27	2.83	2.51	2.23	1.98	1.73	1.45	1.12	.68
∞	4.29	3.60	3.24	2.80	2.48	2.21	1.96	1.71	1.44	1.12	.68

Pr [noncentral $t > t_{1-\alpha} \mid \delta = (\mu_1 - \mu_0)(\sqrt{n}/\sigma)] = 1 - \beta$.

Table C.13 (*Continued*)
Values of the noncentrality parameter for a one-sided test with $\alpha = .010$

f	β = type 2 error										
	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90
1	82.00	62.40	52.37	40.80	33.00	26.79	21.47	16.69	12.27	8.07	4.00
2	15.22	12.26	10.74	8.96	7.73	6.73	5.83	4.98	4.12	3.20	2.08
3	9.34	7.71	6.86	5.87	5.17	4.59	4.07	3.56	3.03	2.44	1.66
4	7.52	6.28	5.64	4.88	4.34	3.88	3.47	3.06	2.63	2.14	1.48
5	6.68	5.62	5.07	4.40	3.93	3.54	3.17	2.81	2.42	1.98	1.38
6	6.21	5.25	4.74	4.13	3.70	3.33	2.99	2.66	2.30	1.88	1.32
7	5.91	5.01	4.53	3.96	3.55	3.20	2.88	2.56	2.22	1.82	1.27
8	5.71	4.85	4.39	3.84	3.44	3.11	2.80	2.49	2.16	1.77	1.24
9	5.56	4.72	4.28	3.75	3.37	3.04	2.74	2.43	2.11	1.74	1.22
10	5.45	4.63	4.20	3.68	3.31	2.99	2.69	2.39	2.08	1.71	1.20
11	5.36	4.56	4.14	3.63	3.26	2.94	2.65	2.36	2.05	1.69	1.18
12	5.29	4.50	4.09	3.58	3.22	2.91	2.62	2.33	2.03	1.67	1.17
13	5.23	4.46	4.04	3.55	3.19	2.88	2.60	2.31	2.01	1.65	1.16
14	5.18	4.42	4.01	3.51	3.16	2.86	2.57	2.29	1.99	1.64	1.15
15	5.14	4.38	3.98	3.49	3.14	2.84	2.56	2.28	1.98	1.63	1.14
16	5.11	4.35	3.95	3.47	3.12	2.82	2.54	2.26	1.97	1.62	1.14
17	5.08	4.33	3.93	3.45	3.10	2.80	2.53	2.25	1.96	1.61	1.13
18	5.05	4.31	3.91	3.43	3.09	2.79	2.52	2.24	1.95	1.60	1.13
19	5.03	4.29	3.89	3.42	3.07	2.78	2.50	2.23	1.94	1.60	1.12
20	5.01	4.27	3.88	3.40	3.06	2.77	2.50	2.22	1.93	1.59	1.12
21	4.99	4.25	3.86	3.39	3.05	2.76	2.49	2.22	1.92	1.59	1.11
22	4.97	4.24	3.85	3.38	3.04	2.75	2.48	2.21	1.92	1.58	1.11
23	4.96	4.23	3.84	3.37	3.03	2.74	2.47	2.20	1.91	1.58	1.11
24	4.94	4.22	3.83	3.36	3.02	2.73	2.47	2.20	1.91	1.57	1.11
25	4.93	4.20	3.82	3.35	3.02	2.73	2.46	2.19	1.90	1.57	1.10
26	4.92	4.19	3.81	3.34	3.01	2.72	2.45	2.19	1.90	1.57	1.10
27	4.91	4.19	3.80	3.34	3.00	2.72	2.45	2.18	1.90	1.56	1.10
28	4.90	4.18	3.79	3.33	3.00	2.71	2.44	2.18	1.89	1.56	1.10
29	4.89	4.17	3.79	3.32	2.99	2.71	2.44	2.17	1.89	1.56	1.10
30	4.88	4.16	3.78	3.32	2.99	2.70	2.44	2.17	1.89	1.55	1.09
40	4.82	4.11	3.74	3.28	2.95	2.67	2.41	2.15	1.86	1.54	1.08
60	4.76	4.06	3.69	3.24	2.92	2.64	2.38	2.12	1.84	1.52	1.07
100	4.72	4.03	3.66	3.21	2.89	2.62	2.36	2.10	1.83	1.51	1.06
∞	4.65	3.97	3.61	3.17	2.85	2.58	2.33	2.07	1.80	1.48	1.04

$\Pr[\text{noncentral } t > t_{1-\alpha} \mid \delta = (\mu_1 - \mu_0)(\sqrt{n}/\sigma)] = 1 - \beta.$ 30

Table 4 SAMPLE SIZE FOR COMPARING THE MEANS OF TWO NORMAL VARIABLES (Table: pp. 9-10)

The two-sample t-test for two normal populations with common variance and possibly unequal means is to be used. Let $\Delta = |\mu_1 - \mu_2|/\sigma$ be the true absolute difference between the means in common standard deviation units. For given α (the significance level), Δ and β (the probability of a type II error), the table entry gives the minimum sample size needed for each group assuming that the two groups are to be sampled an equal number of times.

Table 4 SAMPLE SIZE FOR COMPARING THE MEANS OF TWO NORMAL VARIABLES (Description: p. 6)

α = 0.05 (one-sided), 0.10 (two-sided)

β, probability of type II error

Δ	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.05
0.3	16	29	44	61	81	106	139	191	242
0.4	9	17	25	35	46	60	78	108	136
0.5	6	11	17	23	30	39	51	70	88
0.6	5	8	12	16	21	27	36	49	61
0.7	4	6	9	12	16	20	26	36	45
0.8	3	5	7	10	12	16	21	28	35
0.9	3	4	6	8	10	13	16	22	28
1.0	3	4	5	7	8	11	14	18	23
1.2	2	3	4	5	6	8	10	13	16
1.4	2	3	3	4	5	6	8	10	12
1.6	2	3	3	4	4	5	6	8	10
1.8	2	2	3	3	4	4	5	7	8
2.0	2	2	3	3	3	4	4	6	7
3.0	2	2	2	2	2	3	3	3	4
4.0	2	2	2	2	2	2	2	3	3

α = 0.025 (one-sided), 0.05 (two-sided)

β, probability of type II error

Δ	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.05
0.3	29	47	66	87	110	139	176	235	290
0.4	17	27	38	49	63	79	100	133	164
0.5	12	18	25	32	41	51	64	86	105
0.6	8	13	18	23	29	36	45	60	74
0.7	7	10	13	17	21	27	34	44	55
0.8	6	8	11	14	17	21	26	34	42
0.9	5	7	9	11	14	17	21	27	34
1.0	4	6	7	9	11	14	17	23	27
1.2	3	5	6	7	8	10	12	16	20
1.4	3	4	5	6	7	8	10	12	15
1.6	3	3	4	5	5	6	8	10	12
1.8	3	3	4	4	5	5	6	8	10
2.0	2	3	3	4	4	5	6	7	8
3.0	2	2	3	3	3	3	4	4	5
4.0	2	2	2	2	3	3	3	3	4

Table 4 SAMPLE SIZE FOR COMPARING THE MEANS OF TWO NORMAL VARIABLES (Description: p. 6)

α = 0.01 (one-sided), 0.02 (two-sided)

β, probability of type II error

Δ	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10
0.3	26	51	74	97	122	150	182	225	291
0.4	16	29	42	56	70	85	103	127	165
0.5	11	20	28	36	45	55	67	82	106
0.6	8	14	20	26	32	39	47	58	74
0.7	6	11	15	19	24	29	35	43	55
0.8	5	9	12	15	19	23	27	33	43
0.9	5	7	10	13	15	18	22	27	34
1.0	4	6	8	11	13	15	18	22	28
1.2	4	5	6	8	9	11	13	16	20
1.4	3	4	5	6	8	9	10	12	15
1.6	3	4	5	5	6	7	8	10	12
1.8	3	3	4	5	5	6	7	8	10
2.0	3	3	4	4	5	5	6	7	9
3.0	2	3	3	3	3	4	4	4	5
4.0	2	2	3	3	3	3	3	3	4

α = 0.005 (one-sided), 0.01 (two-sided)

β, probability of type II error

Δ	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10
0.3	39	69	96	122	150	180	216	262	333
0.4	23	40	55	70	85	102	122	148	188
0.5	16	26	36	45	55	66	79	96	121
0.6	12	19	26	32	39	47	56	67	85
0.7	9	14	19	24	29	35	41	50	63
0.8	7	12	15	19	23	27	32	39	49
0.9	6	10	13	16	19	22	26	31	39
1.0	6	8	11	13	15	18	21	26	32
1.2	5	6	8	10	11	13	16	18	23
1.4	4	5	7	8	9	10	12	14	17
1.6	4	5	6	7	7	9	10	11	14
1.8	3	4	5	6	6	7	8	10	11
2.0	3	4	4	5	6	6	7	8	10
3.0	3	3	3	4	4	4	5	5	6
4.0	2	3	3	3	3	3	4	4	4

TABLE 9

Selected Tail Probabilities for the Null Distribution of Wilcoxon's Rank-Sum Statistic

$P = P[W_s \geqslant x] = P[W_s \leqslant x^*]$

SMALLER SAMPLE SIZE = 2

LARGER SAMPLE SIZE

3			4			5			6		
x	P	x^*	x	P	x^*	x	P	x^*	x	P	x^*
8	.200	4	10	.133	4	11	.190	5	13	.143	5
9	.100	3	11	.067	3	12	.095	4	14	.071	4
10	0	2	12	0	2	13	.048	3	15	.036	3
						14	0	2	16	0	2

7			8			9			10		
x	P	x^*	x	P	x^*	x	P	x^*	x	P	x^*
15	.111	5	16	.133	6	18	.109	6	19	.136	7
16	.056	4	17	.089	5	19	.073	5	20	.091	6
17	.028	3	18	.044	4	20	.036	4	21	.061	5
18	0	2	19	.022	3	21	.018	3	22	.030	4
			20	0	2	22	0	2	23	.015	3

SMALLER SAMPLE SIZE = 3

LARGER SAMPLE SIZE

3			4			5			6		
x	P	x^*	x	P	x^*	x	P	x^*	x	P	x^*
13	.200	8	16	.114	8	18	.125	9	20	.131	10
14	.100	7	17	.057	7	19	.071	8	21	.083	9
15	.050	6	18	.029	6	20	.036	7	22	.048	8
16	0	5	19	0	5	21	.018	6	23	.024	7
						22	0	5	24	.012	6
									25	0	5

TABLE 9 (*Continued*)

7			8			9			10		
x	P	x^*	x	P	x^*	x	P	x^*	x	P	x^*
22	.133	11	24	.139	12	27	.105	12	29	.108	13
23	.092	10	25	.097	11	28	.073	11	30	.080	12
24	.058	9	26	.067	10	29	.050	10	31	.056	11
25	.033	8	27	.042	9	30	.032	9	32	.038	10
26	.017	7	28	.024	8	31	.018	8	33	.024	9
27	.008	6	29	.012	7	32	.009	7	34	.014	8
28	0	5	30	.006	6				35	.007	7
			31	0	5						

SMALLER SAMPLE SIZE = 4

LARGER SAMPLE SIZE

4			5			6			7		
x	P	x^*	x	P	x^*	x	P	x^*	x	P	x^*
22	.171	14	25	.143	15	28	.129	16	31	.115	17
23	.100	13	26	.095	14	29	.086	15	32	.082	16
24	.057	12	27	.056	13	30	.057	14	33	.055	15
25	.029	11	28	.032	12	31	.033	13	34	.036	14
26	.014	10	29	.016	11	32	.019	12	35	.021	13
27	0	9	30	.008	10	33	.010	11	36	.012	12
			31	0	9				37	.006	11

8			9			10		
x	P	x^*	x	P	x^*	x	P	x^*
34	.107	18	36	.130	20	39	.120	21
35	.077	17	37	.099	19	40	.094	20
36	.055	16	38	.074	18	41	.071	19
37	.036	15	39	.053	17	42	.053	18
38	.024	14	40	.038	16	43	.038	17
39	.014	13	41	.025	15	44	.027	16
40	.008	12	42	.017	14	45	.018	15
			43	.010	13	46	.012	14
						47	.007	13

TABLE 9 (*Continued*)

SMALLER SAMPLE SIZE = 5

LARGER SAMPLE SIZE

5			6			7			8		
x	P	x^*	x	P	x^*	x	P	x^*	x	P	x^*
34	.111	21	37	.123	23	41	.101	24	44	.111	26
35	.075	20	38	.089	22	42	.074	23	45	.085	25
36	.048	19	39	.063	21	43	.053	22	46	.064	24
37	.028	18	40	.041	20	44	.037	21	47	.047	23
38	.016	17	41	.026	19	45	.024	20	48	.033	22
39	.008	16	42	.015	18	46	.015	19	49	.023	21
			43	.009	17	47	.009	18	50	.015	20
									51	.009	19

9			10		
x	P	x^*	x	P	x^*
47	.120	28	51	.103	29
48	.095	27	52	.082	28
19	.073	26	53	.065	27
50	.056	25	54	.050	26
51	.041	24	55	.038	25
52	.030	23	56	.028	24
53	.021	22	57	.020	23
54	.014	21	58	.014	22
55	.009	20	59	.010	21

SMALLER SAMPLE SIZE = 6

LARGER SAMPLE SIZE

6			7			8			9		
x	P	x^*	x	P	x^*	x	P	x^*	x	P	x^*
47	.120	31	51	.117	33	55	.114	35	59	.112	37
48	.090	30	52	.090	32	56	.091	34	60	.091	36
49	.066	29	53	.069	31	57	.071	33	61	.072	35
50	.047	28	54	.051	30	58	.054	32	62	.057	34
51	.032	27	55	.037	29	59	.041	31	63	.044	33
52	.021	26	56	.026	28	60	.030	30	64	.033	32
53	.013	25	57	.017	27	61	.021	29	65	.025	31
54	.008	24	58	.011	26	62	.015	28	66	.018	30
			59	.007	25	63	.010	27	67	.013	29
									68	.009	28

TABLE 9 (*Continued*)

SMALLER SAMPLE SIZE = 6

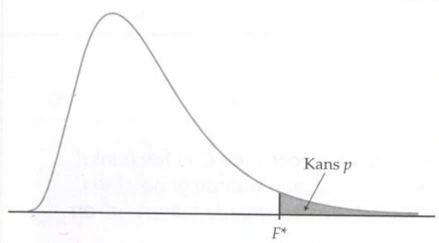
LARGER SAMPLE SIZE

10		
x	P	x^*
63	.110	39
64	.090	38
65	.074	37
66	.059	36
67	.047	35
68	.036	34
69	.028	33
70	.021	32
71	.016	31
72	.011	30
73	.008	29

SMALLER SAMPLE SIZE = 7

LARGER SAMPLE SIZE

7			8			9			10		
x	P	x^*	x	P	x^*	x	P	x^*	x	P	x^*
63	.104	42	67	.116	45	72	.105	47	76	.115	50
64	.082	41	68	.095	44	73	.087	46	77	.097	49
65	.064	40	69	.076	43	74	.071	45	78	.081	48
66	.049	39	70	.060	42	75	.057	44	79	.067	47
67	.036	38	71	.047	41	76	.045	43	80	.054	46
68	.027	37	72	.036	40	77	.036	42	81	.044	45
69	.019	36	73	.027	39	78	.027	41	82	.035	44
70	.013	35	74	.020	38	79	.021	40	83	.028	43
71	.009	34	75	.014	37	80	.016	39	84	.022	42
			76	.010	36	81	.011	38	85	.017	41
						82	.008	37	86	.012	40
									87	.009	39



Tabelelement bij p is de kritiek waarde F^* met kans p erboven.

Tabel E Kritieke waarden voor de F-verdeling

		Aantal vrijheidsgraden in de teller								
Aantal vrijheidsgraden in de noemer	p	1	2	3	4	5	6	7	8	9
1	.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
	.050	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
	.025	647.79	799.50	864.16	899.58	921.85	937.11	948.22	956.66	963.28
	.010	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4	5981.1	6022.5
	.001	405284	500000	540379	562500	576405	585937	592873	598144	602284
2	.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
	.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39
	.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	.001	998.50	999.00	999.17	999.25	999.30	999.33	999.36	999.37	999.39
3	.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
	.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
	.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
	.001	167.03	148.50	141.11	137.10	134.58	132.85	131.58	130.62	129.86
4	.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90
	.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	.001	74.14	61.25	56.18	53.44	51.71	50.53	49.66	49.00	48.47
5	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
	.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68
	.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	.001	47.18	37.12	33.20	31.09	29.75	28.83	28.16	27.65	27.24
6	.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
	.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
	.025	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52
	.010	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
	.001	35.51	27.00	23.70	21.92	20.80	20.03	19.46	19.03	18.69
7	.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
	.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	.025	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82
	.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	.001	29.25	21.69	18.77	17.20	16.21	15.52	15.02	14.63	14.33

Tabel E Kritieke waarden voor de F-verdeling *(Vervolg)*

Aantal vrijheidsgraden in de noemer	p	Aantal vrijheidsgraden in de teller: 1	2	3	4	5	6	7	8	9
8	.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	.025	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36
	.010	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	.001	25.41	18.49	15.83	14.39	13.48	12.86	12.40	12.05	11.77
9	.100	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	.050	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	.025	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03
	.010	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	.001	22.86	16.39	13.90	12.56	11.71	11.13	10.70	10.37	10.11
10	.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78
	.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	.001	21.04	14.91	12.55	11.28	10.48	9.93	9.52	9.20	8.96
11	.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	.025	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59
	.010	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	.001	19.69	13.81	11.56	10.35	9.58	9.05	8.66	8.35	8.12
12	.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	.025	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44
	.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	.001	18.64	12.97	10.80	9.63	8.89	8.38	8.00	7.71	7.48
13	.100	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
	.050	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
	.025	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31
	.010	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	.001	17.82	12.31	10.21	9.07	8.35	7.86	7.49	7.21	6.98
14	.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
	.025	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21
	.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	.001	17.14	11.78	9.73	8.62	7.92	7.44	7.08	6.80	6.58
15	.100	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	.050	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	.025	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12
	.010	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	.001	16.59	11.34	9.34	8.25	7.57	7.09	6.74	6.47	6.26
16	.100	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	.050	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	.025	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05
	.010	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	.001	16.12	10.97	9.01	7.94	7.27	6.80	6.46	6.19	5.98
17	.100	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	.050	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
	.025	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98
	.010	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68
	.001	15.72	10.66	8.73	7.68	7.02	6.56	6.22	5.96	5.75

Tabel E Kritieke waarden voor de *F*-verdeling *(Vervolg)*

Aantal vrijheidsgraden in de teller										
10	12	15	20	25	30	40	50	60	120	1000
2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.35	2.34	2.32	2.30
3.35	3.28	3.22	3.15	3.11	3.08	3.04	3.02	3.01	2.97	2.93
4.30	4.20	4.10	4.00	3.94	3.89	3.84	3.81	3.78	3.73	3.68
5.81	5.67	5.52	5.36	5.26	5.20	5.12	5.07	5.03	4.95	4.87
11.54	11.19	10.84	10.48	10.26	10.11	9.92	9.80	9.73	9.53	9.36
2.42	2.38	2.34	2.30	2.27	2.25	2.23	2.22	2.21	2.18	2.16
3.14	3.07	3.01	2.94	2.89	2.86	2.83	2.80	2.79	2.75	2.71
3.96	3.87	3.77	3.67	3.60	3.56	3.51	3.47	3.45	3.39	3.34
5.26	5.11	4.96	4.81	4.71	4.65	4.57	4.52	4.48	4.40	4.32
9.89	9.57	9.24	8.90	8.69	8.55	8.37	8.26	8.19	8.00	7.84
2.32	2.28	2.24	2.20	2.17	2.16	2.13	2.12	2.11	2.08	2.06
2.98	2.91	2.85	2.77	2.73	2.70	2.66	2.64	2.62	2.58	2.54
3.72	3.62	3.52	3.42	3.35	3.31	3.26	3.22	3.20	3.14	3.09
4.85	4.71	4.56	4.41	4.31	4.25	4.17	4.12	4.08	4.00	3.92
8.75	8.45	8.13	7.80	7.60	7.47	7.30	7.19	7.12	6.94	6.78
2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.04	2.03	2.00	1.98
2.85	2.79	2.72	2.65	2.60	2.57	2.53	2.51	2.49	2.45	2.41
3.53	3.43	3.33	3.23	3.16	3.12	3.06	3.03	3.00	2.94	2.89
4.54	4.40	4.25	4.10	4.01	3.94	3.86	3.81	3.78	3.69	3.61
7.92	7.63	7.32	7.01	6.81	6.68	6.52	6.42	6.35	6.18	6.02
2.19	2.15	2.10	2.06	2.03	2.01	1.99	1.97	1.96	1.93	1.91
2.75	2.69	2.62	2.54	2.50	2.47	2.43	2.40	2.38	2.34	2.30
3.37	3.28	3.18	3.07	3.01	2.96	2.91	2.87	2.85	2.79	2.73
4.30	4.16	4.01	3.86	3.76	3.70	3.62	3.57	3.54	3.45	3.37
7.29	7.00	6.71	6.40	6.22	6.09	5.93	5.83	5.76	5.59	5.44
2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.92	1.90	1.88	1.85
2.67	2.60	2.53	2.46	2.41	2.38	2.34	2.31	2.30	2.25	2.21
3.25	3.15	3.05	2.95	2.88	2.84	2.78	2.74	2.72	2.66	2.60
4.10	3.96	3.82	3.66	3.57	3.51	3.43	3.38	3.34	3.25	3.18
6.80	6.52	6.23	5.93	5.75	5.63	5.47	5.37	5.30	5.14	4.99
2.10	2.05	2.01	1.96	1.93	1.91	1.89	1.87	1.86	1.83	1.80
2.60	2.53	2.46	2.39	2.34	2.31	2.27	2.24	2.22	2.18	2.14
3.15	3.05	2.95	2.84	2.78	2.73	2.67	2.64	2.61	2.55	2.50
3.94	3.80	3.66	3.51	3.41	3.35	3.27	3.22	3.18	3.09	3.02
6.40	6.13	5.85	5.56	5.38	5.25	5.10	5.00	4.94	4.77	4.62
2.06	2.02	1.97	1.92	1.89	1.87	1.85	1.83	1.82	1.79	1.76
2.54	2.48	2.40	2.33	2.28	2.25	2.20	2.18	2.16	2.11	2.07
3.06	2.96	2.86	2.76	2.69	2.64	2.59	2.55	2.52	2.46	2.40
3.80	3.67	3.52	3.37	3.28	3.21	3.13	3.08	3.05	2.96	2.88
6.08	5.81	5.54	5.25	5.07	4.95	4.80	4.70	4.64	4.47	4.33
2.03	1.99	1.94	1.89	1.86	1.84	1.81	1.79	1.78	1.75	1.72
2.49	2.42	2.35	2.28	2.23	2.19	2.15	2.12	2.11	2.06	2.02
2.99	2.89	2.79	2.68	2.61	2.57	2.51	2.47	2.45	2.38	2.32
3.69	3.55	3.41	3.26	3.16	3.10	3.02	2.97	2.93	2.84	2.76
5.81	5.55	5.27	4.99	4.82	4.70	4.54	4.45	4.39	4.23	4.08
2.00	1.96	1.91	1.86	1.83	1.81	1.78	1.76	1.75	1.72	1.69
2.45	2.38	2.31	2.23	2.18	2.15	2.10	2.08	2.06	2.01	1.97
2.92	2.82	2.72	2.62	2.55	2.50	2.44	2.41	2.38	2.32	2.26
3.59	3.46	3.31	3.16	3.07	3.00	2.92	2.87	2.83	2.75	2.66
5.58	5.32	5.05	4.78	4.60	4.48	4.33	4.24	4.18	4.02	3.87

Tabel E **Kritieke waarden voor de F-verdeling *(Vervolg)***

Aantal vrijheidsgraden in de noemer	p	Aantal vrijheidsgraden in de teller: 1	2	3	4	5	6	7	8	9
18	.100	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00
	.050	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
	.025	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93
	.010	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	.001	15.38	10.39	8.49	7.46	6.81	6.35	6.02	5.76	5.56
19	.100	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98
	.050	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
	.025	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88
	.010	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	.001	15.08	10.16	8.28	7.27	6.62	6.18	5.85	5.59	5.39
20	.100	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	.050	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	.025	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84
	.010	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	.001	14.82	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24
21	.100	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95
	.050	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	.025	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80
	.010	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	.001	14.59	9.77	7.94	6.95	6.32	5.88	5.56	5.31	5.11
22	.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	.025	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76
	.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	.001	14.38	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.99
23	.100	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	.050	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	.025	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73
	.010	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	.001	14.20	9.47	7.67	6.70	6.08	5.65	5.33	5.09	4.89
24	.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	.025	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70
	.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	.001	14.03	9.34	7.55	6.59	5.98	5.55	5.23	4.99	4.80
25	.100	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	.050	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	.025	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68
	.010	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	.001	13.88	9.22	7.45	6.49	5.89	5.46	5.15	4.91	4.71
26	.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	.025	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65
	.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	.001	13.74	9.12	7.36	6.41	5.80	5.38	5.07	4.83	4.64
27	.100	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
	.050	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	.025	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63
	.010	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	.001	13.61	9.02	7.27	6.33	5.73	5.31	5.00	4.76	4.57

Tabel E Kritieke waarden voor de *F*-verdeling *(Vervolg)*

Aantal vrijheidsgraden in de teller										
10	12	15	20	25	30	40	50	60	120	1000
1.98	1.93	1.89	1.84	1.80	1.78	1.75	1.74	1.72	1.69	1.66
2.41	2.34	2.27	2.19	2.14	2.11	2.06	2.04	2.02	1.97	1.92
2.87	2.77	2.67	2.56	2.49	2.44	2.38	2.35	2.32	2.26	2.20
3.51	3.37	3.23	3.08	2.98	2.92	2.84	2.78	2.75	2.66	2.58
5.39	5.13	4.87	4.59	4.42	4.30	4.15	4.06	4.00	3.84	3.69
1.96	1.91	1.86	1.81	1.78	1.76	1.73	1.71	1.70	1.67	1.64
2.38	2.31	2.23	2.16	2.11	2.07	2.03	2.00	1.98	1.93	1.88
2.82	2.72	2.62	2.51	2.44	2.39	2.33	2.30	2.27	2.20	2.14
3.43	3.30	3.15	3.00	2.91	2.84	2.76	2.71	2.67	2.58	2.50
5.22	4.97	4.70	4.43	4.26	4.14	3.99	3.90	3.84	3.68	3.53
1.94	1.89	1.84	1.79	1.76	1.74	1.71	1.69	1.68	1.64	1.61
2.35	2.28	2.20	2.12	2.07	2.04	1.99	1.97	1.95	1.90	1.85
2.77	2.68	2.57	2.46	2.40	2.35	2.29	2.25	2.22	2.16	2.09
3.37	3.23	3.09	2.94	2.84	2.78	2.69	2.64	2.61	2.52	2.43
5.08	4.82	4.56	4.29	4.12	4.00	3.86	3.77	3.70	3.54	3.40
1.92	1.87	1.83	1.78	1.74	1.72	1.69	1.67	1.66	1.62	1.59
2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.94	1.92	1.87	1.82
2.73	2.64	2.53	2.42	2.36	2.31	2.25	2.21	2.18	2.11	2.05
3.31	3.17	3.03	2.88	2.79	2.72	2.64	2.58	2.55	2.46	2.37
4.95	4.70	4.44	4.17	4.00	3.88	3.74	3.64	3.58	3.42	3.28
1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.65	1.64	1.60	1.57
2.30	2.23	2.15	2.07	2.02	1.98	1.94	1.91	1.89	1.84	1.79
2.70	2.60	2.50	2.39	2.32	2.27	2.21	2.17	2.14	2.08	2.01
3.26	3.12	2.98	2.83	2.73	2.67	2.58	2.53	2.50	2.40	2.32
4.83	4.58	4.33	4.06	3.89	3.78	3.63	3.54	3.48	3.32	3.17
1.89	1.84	1.80	1.74	1.71	1.69	1.66	1.64	1.62	1.59	1.55
2.27	2.20	2.13	2.05	2.00	1.96	1.91	1.88	1.86	1.81	1.76
2.67	2.57	2.47	2.36	2.29	2.24	2.18	2.14	2.11	2.04	1.98
3.21	3.07	2.93	2.78	2.69	2.62	2.54	2.48	2.45	2.35	2.27
4.73	4.48	4.23	3.96	3.79	3.68	3.53	3.44	3.38	3.22	3.08
1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.62	1.61	1.57	1.54
2.25	2.18	2.11	2.03	1.97	1.94	1.89	1.86	1.84	1.79	1.74
2.64	2.54	2.44	2.33	2.26	2.21	2.15	2.11	2.08	2.01	1.94
3.17	3.03	2.89	2.74	2.64	2.58	2.49	2.44	2.40	2.31	2.22
4.64	4.39	4.14	3.87	3.71	3.59	3.45	3.36	3.29	3.14	2.99
1.87	1.82	1.77	1.72	1.68	1.66	1.63	1.61	1.59	1.56	1.52
2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.84	1.82	1.77	1.72
2.61	2.51	2.41	2.30	2.23	2.18	2.12	2.08	2.05	1.98	1.91
3.13	2.99	2.85	2.70	2.60	2.54	2.45	2.40	2.36	2.27	2.18
4.56	4.31	4.06	3.79	3.63	3.52	3.37	3.28	3.22	3.06	2.91
1.86	1.81	1.76	1.71	1.67	1.65	1.61	1.59	1.58	1.54	1.51
2.22	2.15	2.07	1.99	1.94	1.90	1.85	1.82	1.80	1.75	1.70
2.59	2.49	2.39	2.28	2.21	2.16	2.09	2.05	2.03	1.95	1.89
3.09	2.96	2.81	2.66	2.57	2.50	2.42	2.36	2.33	2.23	2.14
4.48	4.24	3.99	3.72	3.56	3.44	3.30	3.21	3.15	2.99	2.84
1.85	1.80	1.75	1.70	1.66	1.64	1.60	1.58	1.57	1.53	1.50
2.20	2.13	2.06	1.97	1.92	1.88	1.84	1.81	1.79	1.73	1.68
2.57	2.47	2.36	2.25	2.18	2.13	2.07	2.03	2.00	1.93	1.86
3.06	2.93	2.78	2.63	2.54	2.47	2.38	2.33	2.29	2.20	2.11
4.41	4.17	3.92	3.66	3.49	3.38	3.23	3.14	3.08	2.92	2.78

Tabel E **Kritieke waarden voor de *F*-verdeling *(Vervolg)***

Aantal vrijheidsgraden in de noemer	p	Aantal vrijheidsgraden in de teller: 1	2	3	4	5	6	7	8	9
28	.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	.025	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61
	.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	.001	13.50	8.93	7.19	6.25	5.66	5.24	4.93	4.69	4.50
29	.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	.025	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59
	.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	.001	13.39	8.85	7.12	6.19	5.59	5.18	4.87	4.64	4.45
30	.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	.025	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
	.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	.001	13.29	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39
40	.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
	.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	.025	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45
	.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	.001	12.61	8.25	6.59	5.70	5.13	4.73	4.44	4.21	4.02
50	.100	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.76
	.050	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07
	.025	5.34	3.97	3.39	3.05	2.83	2.67	2.55	2.46	2.38
	.010	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78
	.001	12.22	7.96	6.34	5.46	4.90	4.51	4.22	4.00	3.82
60	.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	.025	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33
	.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	.001	11.97	7.77	6.17	5.31	4.76	4.37	4.09	3.86	3.69
100	.100	2.76	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.69
	.050	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97
	.025	5.18	3.83	3.25	2.92	2.70	2.54	2.42	2.32	2.24
	.010	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59
	.001	11.50	7.41	5.86	5.02	4.48	4.11	3.83	3.61	3.44
200	.100	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66
	.050	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93
	.025	5.10	3.76	3.18	2.85	2.63	2.47	2.35	2.26	2.18
	.010	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50
	.001	11.15	7.15	5.63	4.81	4.29	3.92	3.65	3.43	3.26
1000	.100	2.71	2.31	2.09	1.95	1.85	1.78	1.72	1.68	1.64
	.050	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.89
	.025	5.04	3.70	3.13	2.80	2.58	2.42	2.30	2.20	2.13
	.010	6.66	4.63	3.80	3.34	3.04	2.82	2.66	2.53	2.43
	.001	10.89	6.96	5.46	4.65	4.14	3.78	3.51	3.30	3.13

Tabel E Kritieke waarden voor de *F*-verdeling *(Vervolg)*

Aantal vrijheidsgraden in de teller										
10	12	15	20	25	30	40	50	60	120	1000
1.84	1.79	1.74	1.69	1.65	1.63	1.59	1.57	1.56	1.52	1.48
2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.79	1.77	1.71	1.66
2.55	2.45	2.34	2.23	2.16	2.11	2.05	2.01	1.98	1.91	1.84
3.03	2.90	2.75	2.60	2.51	2.44	2.35	2.30	2.26	2.17	2.08
4.35	4.11	3.86	3.60	3.43	3.32	3.18	3.09	3.02	2.86	2.72
1.83	1.78	1.73	1.68	1.64	1.62	1.58	1.56	1.55	1.51	1.47
2.18	2.10	2.03	1.94	1.89	1.85	1.81	1.77	1.75	1.70	1.65
2.53	2.43	2.32	2.21	2.14	2.09	2.03	1.99	1.96	1.89	1.82
3.00	2.87	2.73	2.57	2.48	2.41	2.33	2.27	2.23	2.14	2.05
4.29	4.05	3.80	3.54	3.38	3.27	3.12	3.03	2.97	2.81	2.66
1.82	1.77	1.72	1.67	1.63	1.61	1.57	1.55	1.54	1.50	1.46
2.16	2.09	2.01	1.93	1.88	1.84	1.79	1.76	1.74	1.68	1.63
2.51	2.41	2.31	2.20	2.12	2.07	2.01	1.97	1.94	1.87	1.80
2.98	2.84	2.70	2.55	2.45	2.39	2.30	2.25	2.21	2.11	2.02
4.24	4.00	3.75	3.49	3.33	3.22	3.07	2.98	2.92	2.76	2.61
1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.48	1.47	1.42	1.38
2.08	2.00	1.92	1.84	1.78	1.74	1.69	1.66	1.64	1.58	1.52
2.39	2.29	2.18	2.07	1.99	1.94	1.88	1.83	1.80	1.72	1.65
2.80	2.66	2.52	2.37	2.27	2.20	2.11	2.06	2.02	1.92	1.82
3.87	3.64	3.40	3.14	2.98	2.87	2.73	2.64	2.57	2.41	2.25
1.73	1.68	1.63	1.57	1.53	1.50	1.46	1.44	1.42	1.38	1.33
2.03	1.95	1.87	1.78	1.73	1.69	1.63	1.60	1.58	1.51	1.45
2.32	2.22	2.11	1.99	1.92	1.87	1.80	1.75	1.72	1.64	1.56
2.70	2.56	2.42	2.27	2.17	2.10	2.01	1.95	1.91	1.80	1.70
3.67	3.44	3.20	2.95	2.79	2.68	2.53	2.44	2.38	2.21	2.05
1.71	1.66	1.60	1.54	1.50	1.48	1.44	1.41	1.40	1.35	1.30
1.99	1.92	1.84	1.75	1.69	1.65	1.59	1.56	1.53	1.47	1.40
2.27	2.17	2.06	1.94	1.87	1.82	1.74	1.70	1.67	1.58	1.49
2.63	2.50	2.35	2.20	2.10	2.03	1.94	1.88	1.84	1.73	1.62
3.54	3.32	3.08	2.83	2.67	2.55	2.41	2.32	2.25	2.08	1.92
1.66	1.61	1.56	1.49	1.45	1.42	1.38	1.35	1.34	1.28	1.22
1.93	1.85	1.77	1.68	1.62	1.57	1.52	1.48	1.45	1.38	1.30
2.18	2.08	1.97	1.85	1.77	1.71	1.64	1.59	1.56	1.46	1.36
2.50	2.37	2.22	2.07	1.97	1.89	1.80	1.74	1.69	1.57	1.45
3.30	3.07	2.84	2.59	2.43	2.32	2.17	2.08	2.01	1.83	1.64
1.63	1.58	1.52	1.46	1.41	1.38	1.34	1.31	1.29	1.23	1.16
1.88	1.80	1.72	1.62	1.56	1.52	1.46	1.41	1.39	1.30	1.21
2.11	2.01	1.90	1.78	1.70	1.64	1.56	1.51	1.47	1.37	1.25
2.41	2.27	2.13	1.97	1.87	1.79	1.69	1.63	1.58	1.45	1.30
3.12	2.90	2.67	2.42	2.26	2.15	2.00	1.90	1.83	1.64	1.43
1.61	1.55	1.49	1.43	1.38	1.35	1.30	1.27	1.25	1.18	1.08
1.84	1.76	1.68	1.58	1.52	1.47	1.41	1.36	1.33	1.24	1.11
2.06	1.96	1.85	1.72	1.64	1.58	1.50	1.45	1.41	1.29	1.13
2.34	2.20	2.06	1.90	1.79	1.72	1.61	1.54	1.50	1.35	1.16
2.99	2.77	2.54	2.30	2.14	2.02	1.87	1.77	1.69	1.49	1.22