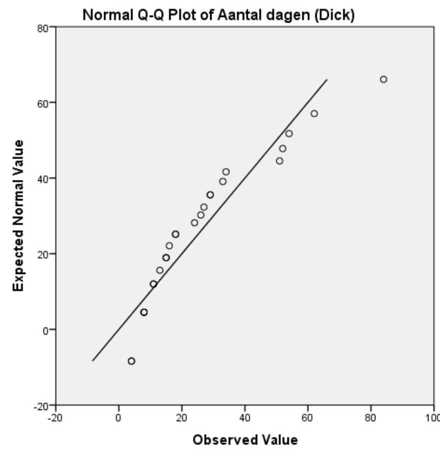
Voorbeeldtoets Statistiek voor INF en BIT (Module 6 -201400256) – tijdsduur 2.15 uur

Gebaseerd op HWO 1-4

Deze toets bestaat uit 5 opgaven, een formuleblad en de $N(0,1)$-, t-, χ^2- en Shapiro-Wilk-tabellen.
Een gewone rekenmachine is toegestaan, een programmeerbare (GR) niet.

1. Een garage houdt het aantal dagen bij dat een occasion te koop staat.
Van 25 occasions zijn in onderstaande tabel de aantallen verkoopdagen en de bijbehorende numerieke samenvatting gegeven. De meetgegevens zijn al gerangschikt van klein naar groot.



4	4	8	8	8
11	11	13	15	15
16	18	18	24	26
27	29	29	33	34
51	52	54	62	84

Numerieke samenvatting:	
Steekproefomvang	25
Steekproefgemiddelde	26.16
Steekproefstandaardafwijking	20.32
Steekproefvariantie	412.98
Steekproefscheefheidcoëfficiënt	1.30
Steekproefkurtosis	4.38

a. Bepaal het 10^{de} en het 80^{ste} percentiel van deze waarnemingen.

b. Ga na of er sprake is van uitschieters volgens de $\mathbf{1.5 \times \mathit{IKA} -}$ **regel**.

c. Ga met behulp van de numerieke samenvatting en het QQ-plot na of het redelijk is hier een normale verdeling voor de aantallen dagen te veronderstellen.

d. Omdat twijfel gerezen is ten aanzien van de normaliteitsveronderstelling wordt de toets van Shapiro-Wilk uitgevoerd: $\boldsymbol{W = 0.870}$. Bepaal het kritieke gebied bij deze toets en trek je conclusie ten aanzien van de normaliteitsveronderstelling met $\boldsymbol{\alpha = 10}\%$.

2. In een onderzoek naar de effectiviteit van een helpdesk werden onder meer de bedieningsduren van klanten, die de helpdesk een probleem voorlegden, onderzocht. Hieronder staan de gemeten bedieningsduren (in minuten) in een steekproef van 42 klanten, gerangschikt van klein naar groot.

0.20	0.62	0.63	1.02	1.08	1.23	1.23	1.24	1.38	1.45
1.80	1.85	1.86	1.91	1.93	1.99	2.10	2.11	2.16	2.21
2.24	2.26	2.29	2.37	2.41	2.42	2.49	2.57	2.81	2.94
3.10	3.34	3.66	3.69	3.81	3.98	4.52	4.67	4.95	5.22
5.76	6.44								

Het steekproefgemiddelde is $\boldsymbol{\overline{x}} = 2.570$ en de steekproefstandaardafwijking is $\boldsymbol{s} = 1.421$

a. Geef een 95%-betrouwbaarheidsinterval voor de verwachte bedieningsduur van klanten bij de helpdesk. Geef duidelijk aan op welke veronderstellingen dit interval gebaseerd is.

b. Iemand interpreteert het interval onder a. als volgt: “Als we de bedieningsduren van willekeurige klanten meten, zullen zo’n 95 van de 100 bedieningsduren in dit interval liggen”.
Is dit een correcte interpretatie? Waarom (niet)?

c. Schat de standaardafwijking van de bedieningsduren met een betrouwbaarheid van 95%.

3. Het nut van de marktwerking in de zorg wordt betwist door tal van politieke partijen en maatschappelijke organisaties. Een deel van hen is voorstander van een terugkeer naar het systeem van "ziekenfonds", nu in de vorm van een solidaire zorgverzekering voor alle bevolkingsgroepen. Een eerste indicatief onderzoek moet uitwijzen of een meerderheid voor afschaffen van de marktwerking is. Daartoe gaat een opiniepeiler 200 willekeurig gekozen Nederlanders na enige uitleg bij de vraagstelling de vraag voorleggen of hij/zij vóór het afschaffen van de markwerking is. $\boldsymbol{p}$ is de fractie van voorstanders van afschaffen onder alle Nederlanders. In de (aselecte) steekproef van 200 personen blijken er 111 voorstander van afschaffen te zijn.

a. Toont de steekproef aan dat de meerderheid van de Nederlanders vóór afschaffen is,?
Voer de toets uit in 8 stappen met $\boldsymbol{\alpha_0 = 0.05}$, door het kritiek gebied te bepalen.

b. Bepaal ook de overschrijdingskans van de toets in a en geef aan voor welke waarden van $\boldsymbol{\alpha_0}$ (tussen 1% en 10%) $\boldsymbol{H_0}$ wordt verworpen.

c. Bereken het onderscheidend vermogen van de toets in a. als het percentage voorstanders van afschaffen in werkelijkheid 60% is.

4. In een onderzoek onder UT-studenten is aantal zaken gemeten. Ten aanzien van het gewicht verwachtten de onderzoekers dat mannen gemiddeld zwaarder zijn dan vrouwen (zoals meestal uit dit soort onderzoeken blijkt). De meetgegevens waren het (o.m.):

Gewicht (in kg)	Aantal	Gemiddelde	Standaardafwijking
vrouw	21	61.3	7.0
man	60	73.9	10.8

a. Ga met een geschikte toets na of de verwachting van de onderzoekers juist is ("mannen zijn aantoonbaar zwaarder dan vrouwen"). Gebruik de 8 stappen van de toetsingsprocedure met $\boldsymbol{\alpha = 0.01}$.

b. Bij a. werd onder meer aangenomen dat de varianties gelijk zijn. Is dit een correcte aanname? Voer daartoe een geschikte toets uit: vermeld (alleen) 1. de hypothesen, 2. de toetsingsgrootheid en zijn waarde, 3. het kritieke gebied en 4. de conclusie die je daaruit trekt m.b.t. de toets onder a., met $\boldsymbol{\alpha = 5}$%.

c. Als de normaliteit van de twee gewichtspopulaties geen houdbare aanname blijkt, welke toets kunnen we dan al alternatief uitvoeren? Vermeld ook 1. de hypothesen, 2. De formule van de toetsingsgrootheid en 3. De benaderende verdeling die je in dit geval gebruikt om de overschrijdingskans te bepalen.

5. De gegevens in bovenstaand onderzoek zijn bij nader inzien afkomstig uit een tweetal enquêtes, één onder INF/BIT en één onder Create studenten (die nu samen module 6 doen). Nagegaan werd ook of zij het nieuwe TOM-onderwijs aantrekkelijk vinden. In de volgende tabel zie je de resultaten samengevat:

		Opinie over aantrekkelijkheid TOM		
		Mee eens/neutraal	Mee oneens	Totaal mee oneens
Studie	INF/BIT	15	12	3
	Create	6	19	28

We vatten deze cijfers op als resultaten van een steekproef uit een grotere populatie (van bijv. ook toekomstige studenten). Ga met een geschikte toets na of de twee groepen studenten verschillende opinies hebben over de aantrekkelijkheid van TOM. Gebruik de toetsingsprocedure met α = 1%.

Uitwerkingen:

Opgave 1

a. Het 10^de^ percentiel: 10% van 25 is 2.5, dus het 10^de^ percentiel is $\boldsymbol{x_{(3)} = 8}$

Het 80^ste^ percentiel: 80% van 25 is 20, dus het 80^ste^ percentiel is $\frac{x_{(20)}+x_{(21)}}{2} = \frac{34+51}{2} = \mathbf{42.5}$

b. 25% van 25 is 6.25, dus $\boldsymbol{Q_1 = x_{(7)} = 11}$ en $\boldsymbol{Q_3 = x_{(19)} = 33}$, dus IKA = 33 – 11 = 22.

$(\boldsymbol{Q_1 - 1.5 \times IKA}, \boldsymbol{Q_3 + 1.5 \times IKA}) = (\mathbf{-22, 66})$, dus 1 (potentiële) uitschieter: 84

c. 1. De numerieke waarden van de scheefheidcoëfficiënt 1.30 (> 0, dus scheefheid naar rechts) en de kurtosis 4.38 wijken af van de referentiewaarden 0 resp. 3 van de normale verdeling, (maar ook van de referentiewaarden van de exponentiële verdeling (2 resp. 9).

2. het normale Q-Q plot vertoont een duidelijk patroon (middenstuk boven de lijn $\boldsymbol{y = x}$ en de rest eronder)

<u>Conclusie</u>: al met al is vanwege de evidente scheefheid naar rechts de normale verdeling wellicht geen correct model.

d. uit de Shapiro-Wilk tabel met $\boldsymbol{n = 25}$ volgt:

- het kritieke gebied voor $\boldsymbol{\alpha = 10}$% is: $\boldsymbol{W \leq c = 0.931}$
- $\boldsymbol{W = 0.870} < \boldsymbol{c}$, dus $\boldsymbol{H_0}$ verwerpen: de verdeling van het aantal dagen is niet normaal met een onbetrouwbaarheid van 5%.

Opgave 2

a. We passen het normale model toe met onbekende μ en σ^2 (dus de "t-procedure"):

Model: de bedieningsduren $\boldsymbol{X_1, \ldots, X_{42}}$ zijn o.o. en alle N(μ, σ^2)-verdeeld

(Zie formuleblad:) het 95%-BI (μ) heeft grenzen $\boldsymbol{\bar{x} \pm c \cdot \frac{s}{\sqrt{n}}}$, met $\bar{x}$ = 2.57, $\boldsymbol{s}$ = 1.421, $\boldsymbol{n}$ = 42

en, uit de t_{41}-tabel: $P(T_{41} \geq \boldsymbol{c}) = ½\ \alpha = 0.025$, dus $\boldsymbol{c}$ = 2.02 (we nemen de t_{40}-tabel als "beste benadering").

Dus 95%-BI (μ) = (2.13, 3.01)

b. Deze interpretatie is onjuist (er liggen ook maar 12 van de 42 waarnemingen binnen dit interval, dus minder dan 30%). Het betrouwbaarheidsinterval heeft betrekking op de verwachte bedieningsduur (= het gemiddelde van alle mogelijke bedieningsduren) en niet op de waarde van één bedieningsduur.

c. 95%-betrouwbaarheidsinterval (σ) = $\left(\sqrt{\frac{(n-1)S^2}{c_2}}, \sqrt{\frac{(n-1)S^2}{c_1}}\right)$,

met $P(\chi^2_{n-1} \leq c_1) = \frac{1}{2}\alpha$ en $P(\chi^2_{n-1} \leq c_2) = 1 - \frac{1}{2}\alpha$ *(zie formuleblad!).*

Hierin is n = 42, $S^2 = 1.421^2$, $c_1 = 24.4$ en $c_2 = 59.3$ zodat $P(\chi^2_{41} \leq c_1) = 2.5\%$ en $P(\chi^2_{n-1} \geq c_2) = 2.5\%$.

Dus 95%-BI(σ) ≈ (1.18, 1.84)

Opgave 3

a. 1. $X =$ "aantal voorstanders in de steekproef met":
X is $B(200, p)$-verdeeld, met $p =$ "de onbekende fractievoorstanders inde populatie".

2. We toetsen $H_0: p = \frac{1}{2}$ tegen $H_1: p > \frac{1}{2}$ met $\alpha_0 = 5\%$

3. Toetsingsgrootheid X

4. Onder H_0 geldt: $X \sim B\left(200, \frac{1}{2}\right)$, dus bij benadering $N(100, 50)$

5. Waargenomen: $x = 111$

6. Verwerp H_0 als $X \geq c$.

$$P(X \geq c | H_0) \stackrel{\text{c.c.}}{=} P\left(X \geq c - \frac{1}{2} | H_0\right) = P(Z \geq \frac{c-0.5-100}{\sqrt{50}}) = 1 - \Phi\left(\frac{c-0.5-100}{\sqrt{50}}\right) \leq \alpha_0 = 0.05$$

Dus $\frac{c-0.5-100}{\sqrt{50}} \geq 1.645$, ofwel $c \geq 100.5 + 1.645 \cdot \sqrt{50} \approx 112.13$. Dus $c = 113$.

7. $x = 111$ ligt niet in het kritieke gebied (< 113), dus H_0 niet verwerpen.
8. Met een onbetrouwbaarheidsdrempel van 10% is **niet** aangetoond dat meer dan de helft voor het afschaffen van de marktwerking in de zorg is.

b. Als $H_0: p = ½$, is X bij benadering N(100, 50). Dus (met continuïteitscorrectie):

$$P(X \geq 111|H_0) \overset{\text{c.c.}}{=} P(X \geq 110.5|H_0) = P\left(\frac{X-100}{\sqrt{50}} \geq \frac{110.5-100}{\sqrt{50}}\right) \approx 1-\Phi(1.48) \approx 6.9\%$$

De P-waarde = 6.9% $\leq \alpha_0$, als $\alpha_0 \geq 6.9\%$. Dus H_0 wordt alléén verworpen $\alpha_0 \geq 6.9\%$.

c. $\beta(0.6) = P(X \geq 113|p = 0.6) = P\left(Z \geq \frac{112.5-200\cdot 0.6}{\sqrt{200\cdot\ 0.6\cdot 0.4}}\right) = P(Z \geq -1.08) = \Phi(1.08) \approx 86.0\%$.

Opgave 4

a. 1. Modelaannames ("statistische veronderstellingen"):
het gaat om twee onafhankelijke, aselecte steekproeven van gewichten, uit de $N(\mu_1, \sigma^2)$-verdeling voor $n_1 = 21$ vrouwen en de $N(\mu_2, \sigma^2)$-verdeling voor $n_2 = 60$ mannen (gelijke σ's!)
Formeler: de opbrengsten $X_1, \ldots, X_{21}, Y_1, \ldots, Y_{60}$ zijn o. o., $X_i \sim N(\mu_1, \sigma^2)$ en $Y_j \sim N(\mu_2, \sigma^2)$

2. We toetsen $H_0: \mu_1 = \mu_2$ tegen $H_1: \mu_1 < \mu_2$ met α = 1%
3. Toetsingsgrootheid $T = \frac{\overline{X}_1-\overline{X}_2}{\sqrt{S^2\left(\frac{1}{21}+\frac{1}{60}\right)}}$ met $S^2 = \frac{20S_1^2+59S_2^2}{21+60-2}$
4. T is onder H_0 t-verdeeld met $df = n_1 + n_2 - 2 = 18$
5. Waargenomen: $s^2 = \frac{20\times 7.0^2 + 59\times 10.8^2}{79} \approx 99.52$ ($s \approx 9.98$), dus $t = \frac{61,3-73.9}{\sqrt{99.52\left(\frac{1}{21}+\frac{1}{60}\right)}} = -4.98$
6. De toets is tweezijdig: verwerp H_0 als $T \leq -c$ met $c = 2.374$ uit de $t_{79} \approx t_{80}$tabel
7. $t = -4.98$ ligt in het kritieke gebied, dus H_0 verwerpen.
8. De gewichten van de vrouwen zijn gemiddeld aantoonbaar lager dan die van mannen bij een onbetrouwbaarheid van 1%.

6./7. Met overschrijdingskans bij de waargenomen $t = -4.98$:
$P(T_{79} \leq -4.98) \approx P(T_{80} \geq 4.98) < 0.0005$, dus ook kleiner dan 1% = α, dus H_0verwerpen,

b. De F-toets op de gevraagde punten:

1. Toets $H_0: \sigma_1^2 = \sigma_2^2$ tegen $H_1: \sigma_1^2 \neq \sigma_2^2$met $\alpha = 5\%$
2. Toetsingsgrootheid $F = \frac{S_1^2}{S_2^2} = \frac{7.0^2}{10.8^2} \approx 0.42$
3. Het is een tweezijdige toets: verwerp H_0 als $F \leq c_1$ of $F \geq c_2$.
$P(F_{59}^{20} \geq c_2) = \frac{\alpha}{2} = 0.05$, dus (volgens de F_{60}^{20}-tabel) $c_2 = 1.94$
$P(F_{59}^{20} \leq c_1) = P\left(F_{20}^{59} \geq \frac{1}{c_1}\right) = \frac{\alpha}{2} = 0.05$, dus $\frac{1}{c_1} = 2.22$, ofwel $c_1 \approx 0.45$
4. De waarde $F = 0.41$ ligt niet in het kritieke gebied (< 0.45), dus H_0 verwerpen.
We mogen dus niet gelijke varianties veronderstellen, bij een onbetrouwbaarheid van 5%

c. Wilcoxon's rangsomtoets: we toetsen $H_0: F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ tegen $H_1: F(x) \neq \Phi\left(\frac{x-\mu}{\sigma}\right)$
met $W = \sum_{i=1}^{21} R(X_i)$, die onder H_0 bij benadering normaal verdeeld is met:
$E(W) = \frac{1}{2}n_1(N+1) = \frac{1}{2}\cdot 21\cdot 82 = 861$ en $var(W) = \frac{1}{12}n_1n_2(N+1) = 8610$

Opgave 5

Er is hier sprake van twee (o.o.) aselecte steekproeven, dus een toets op homogeniteit van de meningsverdelingen van de twee populaties INF-BIT en Create.

De berekening van $\hat{E}_0 N_{ij} = \frac{\text{kolomsom} \times \text{rijsom}}{n}$ in onderstaande tabel levert $E_{ij} \geq 5$ op voor alle (i, j)

		Opinie over aantrekkelijkheid TOM			
		Mee eens/neutraal	Mee oneens	Totaal mee oneens	Totaal
Studie	INF/BIT	N_{11} =15, E_{11} =7.6	N_{12} =12, E_{12} = 11.2	N_{13} =3, E_{13} = 11.2	30
	Create	N_{21} = 6, E_{21} = 13.4	N_{22} =19, E_{22} =19.8	N_{23} =28, E_{23} = 19.8	53
	Totaal	21	31	31	$83 = n$

1. De aantallen N_{11}, N_{12}, N_{13} in de meningsklassen voor de INF-BIT studenten is multinomiaal verdeeld met $n_1 = 100$ en kansen p_{11}, p_{12} en p_{13}. En N_{21}, N_{22} en N_{23} analoog voor de Create studenten: multinomiaal verdeeld met $n_2 = 100$ en kansen p_{21}, p_{22} en p_{23}
2. We toetsen $H_0: p_{11} = p_{21}, p_{12} = \mathrm{p}_{22}$ en $\mathrm{p}_{13} = \mathrm{p}_{23}$ (gelijke meningsverdelingen)
   tegen $H_1: p_{1j} \neq p_{2j}$ voor minstens één waarde van j met $\alpha = 0.01$
3. Toetsingsgrootheid is $\chi^2 = \sum\sum \frac{(N_{ij} - \hat{E}_0 N_{ij})^2}{\hat{E}_0 N_{ij}}$ met schattingen $\hat{E}_0 N_{ij} = \frac{\text{kolomsom} \times \text{rijsom}}{n}$
4. Onder H_0 heeft χ^2 heeft een Chi kwadraat verdeling, aantal vrijheidsgraden $df = (r-1)(c-1) = 2$
5. We berekenen eerst de verwachte aantallen bij onafhankelijkheid: zie tabel hierboven: $\hat{E}_0 N_{ij} = E_{ij}$
   Waargenomen: $\chi^2 = \frac{(15-7.6)^2}{7.6} + \frac{(6-13.4)^2}{13.4} + \frac{(12-11.2)^2}{11.2} + \frac{(19-19.8)^2}{19.8} + \frac{(3-11.2)^2}{11.2} + \frac{(28-19.8)^2}{19.8} = 20.78$
6. We verwerpen H_0 als $\chi^2 \geq c$. In de χ^2-tabel met $df = 2$ vinden we $c \approx 9.21$
7. De uitkomst 20.78 ligt in het kritiek gebied (> 9.21), dus H_0 verwerpen.
8. Bij significantieniveau 1% is een verband tussen de mening over TOM en de studierichting aangetoond.