Statistiek voor MBW / KW
Uitwerkingen huiswerkopgaven

Dr. ir. D.A.M.P. Blom & Dr. J.B.M. Melissen

2025

Week 12: verschiltoetsen

Hoofdstuk 11

Opdracht 11.m1: Voor een tweetal normale verdelingen is gegeven dat hun standaarddeviatie respectievelijk $\sigma_1 = 4$ en $\sigma_2 = 6$ bedragen. Men wil toetsen $H_0 : \mu_1 = \mu_2$. Een steekproef van negen waarnemingen uit de eerste verdeling leverde een gemiddelde uitkomst op van 48, en een steekproef van zestien waarnemingen uit de tweede verdeling leverde een gemiddelde op van 42. De toetsingsgrootheid bij deze toets levert een z-waarde op van ...

(a) 6.

(b) $2,99$.

(c) $6,63$.

(d) $2,67$.

Uitwerking

We hebben te maken met twee normaal verdeelde kansvariabelen, namelijk $X_1 \sim N(\mu_1 =?; \sigma_1 = 4)$ en $X_2 \sim N(\mu_2 =?; \sigma_2 = 6)$. Aangezien we respectievelijk steekproeven hebben van groottes $n = 9$ en $m = 16$, geldt voor de theoretische steekproefgemiddeldes $\overline{X_1} \sim N(\mu = \mu_1; \sigma = \frac{\sigma_1}{\sqrt{n}})$ en $\overline{X_2} \sim N(\mu = \mu_2; \sigma = \frac{\sigma_2}{\sqrt{m}})$, oftewel

$$\overline{X_1} \sim N(\mu_1 =?; \sigma = \frac{4}{3})$$
$$\overline{X_2} \sim N(\mu_2 =?; \sigma = \frac{3}{2})$$

De verschilvariabele $V = \overline{X_1} - \overline{X_2}$ is dan normaal verdeeld met verwachtingswaarde $\mu_V = \mu_1 - \mu_2$ en standaardafwijking $\sigma_V = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} = \sqrt{\frac{16}{9} + \frac{9}{4}} \approx 2,0069$. Onder de nulhypothese geldt $\mu_1 = \mu_2$, oftewel de verschilvariabele heeft verwachtingswaarde $\mu_1 - \mu_2 = 0$. De geobserveerde waarde $v = \overline{x_1} - \overline{x_2} = 48 - 42 = 6$.
De z-waarde die bij deze uitkomst hoort is gelijk aan

$$z = \frac{v - \mu_V}{\sigma_V} = \frac{6 - 0}{2,0069} \approx 2,99$$

Het juiste antwoord is dus (b).

Opdracht 11.m3: Voor twee populaties mag worden verondersteld dat deze allebei een normale verdeling volgen en dezelfde onbekende variantie hebben. Men wil toetsen of deze populaties hetzelfde gemiddelde hebben. Daartoe neemt men uit beide populaties een steekproef van tien waarnemingen. Bij de toets is het aantal vrijheidsgraden van de toetsingsgrootheid gelijk aan ...

(a) 9.

(b) 18.

(c) 19.

(d) $-$ (er is bij deze toetsingsgrootheid geen sprake van vrijheidsgraden).

Uitwerking

We hebben te maken met twee normaal verdeelde kansvariabelen $X_1 \sim N(\mu_1 =?; \sigma_1 = ?)$ en $X_2 \sim N(\mu_2 =?; \sigma_2 =?)$ met dezelfde onbekende variantie, dat wil zeggen $\sigma = sigma_1 = \sigma_2$. In dat geval mogen de twee steekproefschattingen s_1 en s_2 voor de standaardafwijking (op basis van steekproeven van grootte $n = m = 10$ respectievelijk) worden gecombineerd tot één schatter (de zogenaamde "pooled variance").
De kansverdeling die de bijbehorende toetsingsgrootheid volgt is de t-verdeling met $df = n + m - 2 = 10 + 10 - 2 = 18$ vrijheidsgraden. Het juiste antwoord is dus (b).

Opdracht 11.m6: Voor een tweetal normale verdelingen willen we toetsen of ze dezelfde waarde van de variantie hebben. Voor de eerste variabele werd met een steekproef van twaalf waarnemingen de waarde $s^2 = 240$ gevonden. Bij de tweede variabele leverde de steekproef van tien waarnemingen de waarde $s^2 = 160$. De toetsingsgrootheid krijgt hier dan de waarde ...

(a) 80.

(b) $1,50$.

(c) 4.

(d) $1,25$.

Uitwerking

We hebben te maken met twee normaal verdeelde kansvariabelen $X_1 \sim N(\mu_1 =?; \sigma_1 = ?)$ en $X_2 \sim N(\mu_2 =?; \sigma_2 =?)$ waarvan we willen toetsen of ze dezelfde standaardafwijking hebben. In andere woorden, we willen toetsen met $H_0 : \sigma_1^2 = \sigma_2^2$ tegen $H_1 : \sigma_1^2 \neq \sigma_2^2$.
Dit kunnen we doen met behulp van een F-toets. We hebben voor beide kansvariabelen een schatting $s_1^2 = 240$ en $s_2^2 = 160$ bepaald van de variantie, op basis van steekproeven van grootte respectievelijk $n = 12$ en $m = 10$. De bijbehorende toetsingsgrootheid is $F = \frac{S_1^2}{S_2^2}$ volgt de $F(12, 10)$-verdeling. De geobserveerde toetsingsgrootheid is $f = \frac{s_1^2}{s_2^2} = \frac{240}{160} = 1,5$.
Het juiste antwoord is dus (b).

Opdracht 11.1: Een instantie die toezicht houdt op de mate van luchtverontreiniging doet regelmatig metingen op diverse plaatsen in het land. De gemeten verontreiniging wordt uitgedrukt in een bepaalde index, die op alle werkdagen wordt bepaald. In 2018 werden enkele nieuwe richtlijnen voor de industrie afgekondigd met als doel het niveau van luchtverontreiniging te verlagen. Om het effect van deze maatregelen te bestuderen, werd een onderzoek gedaan. Hierbij werden 20 waarden (x_i) die bepaald zijn in februari 2017 vergeleken met 20 waarden (y_i) die zijn gemeten in februari 2018. Voor deze waarden werd berekend:

$$\overline{x} = 164,6 \text{ en } s_X = 17,2; \text{ en } \overline{y} = 143,2 \text{ en } s_Y = 15,9$$

Met een verschiltoets gaan we bepalen of er een significante verbetering is opgetreden.

Uitwerking

Opdracht 11.3: De variabele X is normaal verdeeld met onbekende μ en een variantie $\text{Var}(()X) = 100$. Ook de variabele Y is normaal verdeeld met onbekende verwachtingswaarde. Er geldt $\text{Var}(()Y) = 30$. Van de variabele X worden tien trekkingen gedaan. Die leveren als gemiddelde $\overline{x} = 50$. Er worden vijf trekkingen gedaan van de variabele Y die als gemiddelde opleveren $\overline{y} = 55$. Toets of beide variabelen dezelfde verwachtingswaarde kunnen hebben (kies $\alpha = 0,05$).

Uitwerking

Opdracht 11.5: Van werknemers in een bepaalde bedrijfstak wordt het verband tussen jaarsalaris en het bezit van een vakdiploma onderzocht. Groep 1 is een steekproef uit de populatie van werknemers met diploma, groep 2 is een steekproef van werknemers zonder diploma. De resultaten waren (jaarinkomens in duizenden euro's):

Groep 1	40	45	48	33	42	35	32	47	38	37	43
Groep 2	25	28	30	35	38	32	22				

Veronderstel dat beide inkomensverdelingen mogen worden beschouwd als normale verdelingen met dezelfde onbekende variantie. Toets of $\mu_1 = \mu_2$ (kies $\alpha = 0,05$).

Uitwerking

Opdracht 11.7: In een studie naar luchtvervuiling leverde een aselecte steekproef van acht luchtmonsters in de omgeving van een bepaalde fabriek gemiddeld $2,26$ microgram van een schadelijke stof per m^3, met een standaardafwijking van $0,56$. In een grote stad leverde een steekproef van tien monsters een resultaat van $1,54$ als gemiddelde en $0,42$ als standaardafwijking. Ook in dit geval ging het om de hoeveelheid van dezelfde stof per m^3. Onderzoek met $\alpha = 0,01$ of het gevonden verschil significant is. Hierbij mag men aannemen dat de steekproeven afkomstig waren uit twee normale verdelingen met dezelfde variantie.

Uitwerking

Opdracht 11.12: Een grootwinkelbedrijf verkoopt in alle filialen een bepaald afwasmiddel. In een week werd in zes filialen het volgende aantal flessen van dit afwasmiddel verkocht.

Filiaal	**A**	**B**	**C**	**D**	**E**	**F**
Aantal flessen	320	260	440	400	380	300

In de week nadat voorgaande gegevens werden verzameld, is een intensieve reclamecampagne gehouden voor het desbetreffende afwasmiddel. Volgend op deze reclamecampagne worden in tien filialen in één week de volgende aantallen flessen verkocht:

Filiaal	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**
Aantal flessen	400	3600	410	500	540	380	490	520	420	480

Toets of het gemiddelde aantal verkochte flessen voor en na de behandeling gelijk is ($\alpha = 0,05$). (Ga ervan uit dat de waargenomen aantallen mogen worden beschouwd als trekkingen uit normale verdelingen.)

Uitwerking

Opdracht 11.13: In zes filialen van een supermarktketen is het aantal verkochte flessen van een wasmiddel vlak voor en vlak na een reclamecampagne geregistreerd. De resultaten waren:

Filiaal	**A**	**B**	**C**	**D**	**E**	**F**
Aantal voor	320	260	440	400	380	310
Aantal na	420	380	520	490	490	400

Toets of er een opvallend verschil is in de omzet per filiaal voor en na de campagne ($\alpha = 0,05$). Voer de toets op twee manieren uit en geef aan van welke veronderstellingen men is uitgegaan bij de toetsen (vergelijk opgave 11.12).

Uitwerking

Opdracht 11.14: Een onderzoeker merkt naar aanleiding van de aanpak bij opgave 11.13 op dat het misschien beter is om per filiaal eerst de procentuele omzetstijging te berekenen. Voer op basis van deze uitkomsten nogmaals een t-toets uit. (Kies $\alpha = 0,05$.)

Uitwerking