Statistiek voor Psychologen
Deel 1

GRATIS Proefexamen
+
Rekenmachine Uitwerking

StatHulp

Academiejaar 2025-2026

Voorwoord

Waarschijnlijk lees je dit voorwoord nu, na meermaals de theorie doorgenomen te hebben, het maken van talloze oefeningen die je inmiddels vanbuiten kent (en misschien zelfs bijna over droomt), en met de aanstaande examens in gedachten. De onzekerheid knaagt wellicht nog. Weet dat je niet alleen bent.

Duizenden psychologiestudenten vóór jou hebben zich in dezelfde situatie bevonden, worstelend met abstracte theorieën om de complexe statistische leerstof te doorgronden.

Wees niet langer onzeker over je capaciteiten. Het feit dat je actief op zoek bent gegaan naar extra studiemateriaal, toont al aan hoe toegewijd je bent aan dit vak. Deze bundel is precies ontworpen om je hierbij verder te helpen.

Deze gratis Proefexamenbundel biedt je **20 unieke extra oefeningen** die je verder uitdagen en je kennis grondig trainen. Elke oefening is voorzien van een zeer uitgebreide oplossing, **inclusief gedetailleerde uitwerkingen op je rekenmachine**. Dit zal je ongetwijfeld helpen om efficiënt te werken tijdens het echte examen.

Statistiek hoeft geen nachtmerrie te zijn!

Voor meer vragen, bundels en studiehulp kun je ons bereiken via:
stathulp@gmail.com of stathulp.be

Instructies voor het Examen

Deze examenbundel bestaat uit 15 meerkeuzevragen en 5 open vragen.

- Lees elke vraag zorgvuldig door.
- Voor meerkeuzevragen: selecteer de meest correcte optie. Er is slechts één correct antwoord per vraag.
- Voor open vragen: geef een volledig uitgewerkte oplossing en vermeld alle aannames die je maakt.
- Gebruik je rekenmachine waar nodig en noteer relevante stappen.
- Je hebt een totaal van **3 uur** de tijd om dit examen af te leggen.

Veel succes!

Meerkeuzevragen

1) De populatie IQ-scores van een specifieke bevolkingsgroep is normaal verdeeld met een gemiddelde $\mu = 110$ en een standaarddeviatie $\sigma = 12$. Een onderzoeker trekt willekeurig een steekproef van $n = 36$ personen uit deze populatie. Wat is de kans dat het steekproefgemiddelde van de IQ-scores tussen 108 en 113 ligt?

   A) 0.1977
   B) 0.6247
   C) 0.7734
   D) 0.8123

2) De kans dat een bepaalde psychologische test betrouwbaar is, is 0.85. Een psycholoog neemt de test af bij 5 onafhankelijke patiënten. Wat is de kans dat de test bij precies 4 van deze patiënten betrouwbaar blijkt te zijn?

   A) 0.3915
   B) 0.2768
   C) 0.0076
   D) 0.4063

3) Een continue toevalsvariabele X is uniform verdeeld over het interval $[5, 15]$. Wat is de kans dat X een waarde aanneemt die minstens 2 eenheden verwijderd is van het gemiddelde van de verdeling?

   A) 0.40
   B) 0.60
   C) 0.80
   D) 0.20

4) Gegeven zijn twee gebeurtenissen A en B. We weten dat $P(A) = 0.4$, $P(B) = 0.6$, en $P(A \cap B) = 0.2$. Wat is $P(A \cup B)$?

   A) 0.8
   B) 0.6
   C) 0.7
   D) 0.9

5) Een psycholoog bestudeert het aantal paniekaanvallen dat een patiënt per maand heeft. Dit aantal volgt een Poissonverdeling met een gemiddelde van $\lambda = 3$ paniekaanvallen per maand. Wat is de kans dat een patiënt in een willekeurige maand geen paniekaanvallen heeft?

   A) 0.0498
   B) 0.1991
   C) 0.0149
   D) 0.0033

6) Welke van de volgende beweringen over de Centrale Limiet Stelling (CLS) is CORRECT?

   A) De CLS stelt dat de populatieverdeling normaal verdeeld zal zijn als de steekproefomvang groot genoeg is.

   B) De CLS garandeert dat de steekproevenverdeling van het gemiddelde normaal verdeeld is, ongeacht de steekproefomvang.

   C) De CLS stelt dat de steekproevenverdeling van het gemiddelde een normale vorm zal aannemen als de steekproefomvang voldoende groot is, ongeacht de vorm van de populatieverdeling.

   D) De CLS is alleen van toepassing op discreet verdeelde populaties.

7) Gegeven een toevalsvariabele X met $E[X] = 20$ en $\sigma_X^2 = 5$. Definieer een nieuwe variabele $Y = 3X - 10$. Wat is de variantie van Y (σ_Y^2)?

   A) 15

   B) 35

   C) 45

   D) 75

8) De proportie studenten die online cursussen volgt, is 0.40. Een onderzoeker wil dit schatten door een steekproef te trekken. Als de onderzoeker de standaardfout van de steekproefproportie wil halveren, met welke factor moet de steekproefomvang dan worden vergroot?

   A) 2

   B) 4

   C) $\sqrt{}$

   D) 16

9) De tijd (in dagen) die nodig is om een patiënt te revalideren na een bepaalde operatie volgt een Exponentiële verdeling met een gemiddelde revalidatietijd van 14 dagen. Wat is de kans dat een willekeurige patiënt binnen 7 dagen gerevalideerd is?

   A) 0.3935

   B) 0.6065

   C) 0.2500

   D) 0.7500

10) Voor twee toevalsvariabelen X en Y geldt: $\sigma_X^2 = 10$, $\sigma_Y^2 = 8$, en $\sigma_{XY} = 4$. Wat is de variantie van de somvariabele $X + Y$ (σ_{X+Y}^2)?

    A) 18

    B) 22

    C) 26

    D) 30

11) Een onderzoeker verzamelt gegevens over de score op een geheugentest (X) en het aantal uren slaap per nacht (Y) voor een groep van 100 personen. De correlatiecoëfficiënt tussen X en Y is $r_{XY} = -0.30$. Welke van de volgende uitspraken is CORRECT op basis van deze correlatie?

    A) Hogere scores op de geheugentest gaan gepaard met meer uren slaap.

    B) Er is geen lineaire relatie tussen de geheugentestscore en het aantal uren slaap.

    C) Lagere scores op de geheugentest gaan gepaard met meer uren slaap.

    D) De correlatie impliceert dat minder slaap direct een lagere geheugenscore veroorzaakt.

12) De lifetime van een psychologische batterij is normaal verdeeld met een gemiddelde van 800 uur en een standaarddeviatie van 50 uur. Een batterij wordt als 'defect' beschouwd als zijn lifetime minder dan 720 uur bedraagt. Wat is de Z-score die hoort bij een lifetime van 720 uur?

    A) -1.6

    B) -0.8

    C) 1.6

    D) 0.8

13) Welke van de volgende beweringen is ALTIJD waar met betrekking tot de machtsverzameling (2^S) van een verzameling S?

    A) $2^{A \cup B} = 2^A \cup 2^B$

    B) $2^{A \cap B} = 2^A \cap 2^B$

    C) $2^{A \setminus B} = 2^A \setminus 2^B$

    D) Geen van bovenstaande.

14) De gemiddelde reactietijd op een visuele stimulus is 350 ms met een standaarddeviatie van 60 ms. Volgens de ongelijkheid van Tchebychev, wat is de MINIMALE proportie reactietijden die binnen 2 standaarddeviaties van het gemiddelde vallen?

    A) 0.68

    B) 0.75

    C) 0.88

    D) 0.95

15) Een toevalsvariabele X heeft de volgende kansmassafunctie: $p_X(x) = k \cdot x$ voor $x = 1, 2, 3, 4$. Wat is de waarde van de constante k?

    A) 1/4

    B) 1/10

    C) 1/6

    D) 1/12

Open Vragen

1) Een onderzoeker wil de relatie tussen studiemotivatie (X) en examenprestaties (Y) analyseren aan de hand van een scatterplot met 5 punten. Teken een scatterplot die voldoet aan de volgende voorwaarden:

    - Alle X-waarden zijn groter dan of gelijk aan 0.
    - Het conditioneel gemiddelde van Y gegeven X, $E[Y|X]$, is constant voor alle X-waarden.
    - De verklaarde variantie van Y door X, s^2_{verkl}, is gelijk aan 0.
    - De totale variantie van Y, s^2_Y, is groter dan 0.

    Geef de coördinaten van de 5 punten en leg uit hoe je scatterplot aan elke voorwaarde voldoet.

2) De kans dat een individu een bepaalde infectie (I) oploopt is 0.05. Als een individu de infectie heeft, is de kans dat de test positief (T+) is 0.98. Als een individu de infectie niet heeft, is de kans dat de test positief (T+) is 0.01 (vals positief).

    a) Wat is de kans dat een willekeurig gekozen individu een positieve testuitslag heeft?
    b) Wat is de kans dat een individu de infectie heeft, GEGEVEN dat de test positief is? (Gebruik de Regel van Bayes).

3) Gegeven is de volgende tabel van gemiddelden voor de afhankelijk variabele Y, voorspeld door twee kwalitatieve predictoren X_1 en X_2:

	$X_2 = -1$	$X_2 = +1$
$X_1 = -1$	10	12
$X_1 = +1$	14	16

    a) Leg uit of er in dit scenario een hoofdeffect van X_1 is. Zo ja, bereken de grootte ervan.
    b) Leg uit of er in dit scenario een interactie-effect tussen X_1 en X_2 is. Zo ja, motiveer je antwoord met berekeningen.
    c) Pas één cel van de tabel aan zodat er **geen interactie-effect** meer is, maar de hoofdeffecten (indien aanwezig) behouden blijven. Leg uit welke cel je aanpast en waarom.

4) De gemiddelde reactietijd van mannelijke studenten op een cognitieve taak is 400 ms met een standaarddeviatie van 60 ms. Voor vrouwelijke studenten is de gemiddelde reactietijd 380 ms met een standaarddeviatie van 50 ms. De reactietijden zijn normaal verdeeld. Een onderzoeker trekt een willekeurige steekproef van 50 mannelijke studenten en 40 vrouwelijke studenten.

    a) Wat is de verwachte waarde van de steekproevenverdeling van het verschil in reactietijden $(\overline{X}_M - \overline{X}_V)$?
    b) Wat is de standaardfout van de steekproevenverdeling van het verschil in reactietijden?

c) Wat is de kans dat het gemiddelde van de mannelijke studenten in de steekproef minder dan 10 ms langer is dan die van de vrouwelijke studenten?

5) Gegeven een discrete toevalsvariabele X met de volgende kansmassafunctie:

x	1	2	3
$p_X(x)$	0.3	0.5	0.2

We trekken een steekproef van $n = 2$ onafhankelijke waarnemingen uit deze populatie, X_1 en X_2.

a) Construeer de kansmassafunctie van het steekproefgemiddelde $\overline{X}$.

b) Bereken de verwachte waarde van $\overline{X}$ ($E[\overline{X}]$).

c) Bereken de variantie van $\overline{X}$ ($\sigma^2_{\overline{X}}$).

Oplossing Meerkeuzevragen

1) De populatie IQ-scores van een specifieke bevolkingsgroep is normaal verdeeld met een gemiddelde $\mu = 110$ en een standaarddeviatie $\sigma = 12$. Een onderzoeker trekt willekeurig een steekproef van $n = 36$ personen uit deze populatie. Wat is de kans dat het steekproefgemiddelde van de IQ-scores tussen 108 en 113 ligt?
Antwoord: C Uitleg: Gegeven: $\mu = 110$, $\sigma = 12$, $n = 36$. De steekproevenverdeling van het gemiddelde $\overline{X}$ is normaal verdeeld met: $E[\overline{X}] = \mu = 110$ $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{36}} = \frac{12}{6} = 2$. We zoeken $P(108 < \overline{X} < 113)$. Transformeer naar Z-scores: Voor $\overline{X}_1 = 108$: $Z_1 = \frac{108-110}{=}\frac{-2}{=} - 1.00$. Voor $\overline{X}_2 = 113$: $Z_2 = \frac{113-110}{=}\frac{3}{=}1.50$. $P(-1.00 < Z < 1.50) = P(Z < 1.50) - P(Z \leq -1.00)$. Uit Z-tabel: $P(Z < 1.50) = 0.9332$. Uit Z-tabel: $P(Z \leq -1.00) = 0.1587$. $P(-1.00 < Z < 1.50) = 0.9332 - 0.1587 = \mathbf{0.7745}$. De dichtstbijzijnde optie is C (0.7734).

Rekenmachine (TI-30XB) Stappen

(a) Bereken de standaardfout: $12/\sqrt{36} = 2$.
(b) Bereken de Z-scores: $(108 - 110)/2 = -1.00$ en $(113 - 110)/2 = 1.50$.
(c) Zoek de cumulatieve kansen op in een Z-tabel voor -1.00 en 1.50.
(d) Trek de kansen van elkaar af: $0.9332 - 0.1587 = 0.7745$.

2) De kans dat een bepaalde psychologische test betrouwbaar is, is 0.85. Een psycholoog neemt de test af bij 5 onafhankelijke patiënten. Wat is de kans dat de test bij precies 4 van deze patiënten betrouwbaar blijkt te zijn?
Antwoord: A Uitleg: Dit is een Binomiale verdeling met $n = 5$ (aantal proeven) en $p = 0.85$ (kans op succes). We zoeken $P(X = 4)$. $P(X = k) = \binom{n}{k}p^k(1-p)^{n-k}$. $P(X = 4) = \binom{5}{4}(0.85)^4(1 - 0.85)^{5-4}$ $P(X = 4) = \binom{5}{4}(0.85)^4(0.15)^1$. $\binom{5}{4} = 5$. $(0.85)^4 \approx 0.5220$. $(0.15)^1 = 0.15$. $P(X = 4) = 5 \times 0.5220 \times 0.15 = \mathbf{0.3915}$.

Rekenmachine (TI-30XB) Stappen

(a) Bereken $\binom{5}{4}$ met **nCr**.
(b) Bereken $(0.85)^4$ met de y^x knop.
(c) Bereken $(0.15)^1 = 0.15$.
(d) Vermenigvuldig de resultaten: $5 \times 0.5220 \times 0.15 = 0.3915$.

3) Een continue toevalsvariabele X is uniform verdeeld over het interval $[5, 15]$. Wat is de kans dat X een waarde aanneemt die minstens 2 eenheden verwijderd is van het gemiddelde van de verdeling?
Antwoord: B Uitleg: Het interval is $[a, b] = [5, 15]$. Het gemiddelde van een uniforme verdeling is $\mu = \frac{a+b}{=}\frac{5+15}{=}10$. We zoeken de kans dat X minstens 2 eenheden verwijderd is van 10. Dit betekent $X \leq 10 - 2$ of $X \geq 10 + 2$. Dus $X \leq 8$ of $X \geq 12$. De dichtheidsfunctie is $\varphi_X(x) = \frac{1}{b-a} = \frac{1}{15-5} = \frac{1}{10}$ voor $5 \leq x \leq 15$. $P(X \leq 8) = (8 - 5) \times$

$\frac{1}{10} = 3 \times \frac{1}{10} = 0.3$. $P(X \geq 12) = (15 - 12) \times \frac{1}{10} = 3 \times \frac{1}{10} = 0.3$. De totale kans is $P(X \leq 8) + P(X \geq 12) = 0.3 + 0.3 = \mathbf{0.60}$.

Rekenmachine (TI-30XB) Stappen

(a) Bereken het gemiddelde: $(5 + 15)/2 = 10$.
(b) Bepaal de intervallen: $X \leq 8$ en $X \geq 12$.
(c) Bereken de hoogte van de dichtheidsfunctie: $1/(15 - 5) = 0.1$.
(d) Bereken de kansen voor elk interval: $(8 - 5) \times 0.1 = 0.3$ en $(15 - 12) \times 0.1 = 0.3$.
(e) Tel de kansen op: $0.3 + 0.3 = 0.60$.

4) Gegeven zijn twee gebeurtenissen A en B. We weten dat $P(A) = 0.4$, $P(B) = 0.6$, en $P(A \cap B) = 0.2$. Wat is $P(A \cup B)$?
Antwoord: A Uitleg: Gebruik de algemene optelregel voor kansen: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. $P(A \cup B) = 0.4 + 0.6 - 0.2 = 1.0 - 0.2 = \mathbf{0.8}$.

Rekenmachine (TI-30XB) Stappen

(a) Tel $P(A)$ en $P(B)$ op: $0.4 + 0.6 = 1.0$.
(b) Trek $P(A \cap B)$ af: $1.0 - 0.2 = 0.8$.

5) Een psycholoog bestudeert het aantal paniekaanvallen dat een patiënt per maand heeft. Dit aantal volgt een Poissonverdeling met een gemiddelde van $\lambda = 3$ paniekaanvallen per maand. Wat is de kans dat een patiënt in een willekeurige maand geen paniekaanvallen heeft?
Antwoord: A Uitleg: Dit is een Poissonverdeling met $\lambda = 3$. We zoeken $P(X = 0)$. De kansmassafunctie voor een Poissonverdeling is $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$. $P(X = 0) = \frac{3^0 e^{-3}}{0!} = \frac{1 \cdot e^{-3}}{1} = e^{-3}$. $e^{-3} \approx \mathbf{0.049787}$. De dichtstbijzijnde optie is A (0.0498).

Rekenmachine (TI-30XB) Stappen

(a) Bereken e^{-3} met de e^x knop (vaak **2nd** dan **LN**).

6) Welke van de volgende beweringen over de Centrale Limiet Stelling (CLS) is CORRECT?
Antwoord: C Uitleg: De Centrale Limiet Stelling stelt dat de steekproevenverdeling van het gemiddelde een normale vorm zal aannemen als de steekproefomvang (n) voldoende groot is, ongeacht de vorm van de populatieverdeling.

- A is onjuist: CLS gaat over de steekproevenverdeling van het gemiddelde, niet de populatieverdeling zelf.
- B is onjuist: De normaliteit wordt benaderd, en vereist een voldoende grote steekproefomvang (typisch $n \geq 30$).
- D is onjuist: CLS is van toepassing op zowel discreet als continu verdeelde populaties.

7) Gegeven een toevalsvariabele X met $E[X] = 20$ en $\sigma_X^2 = 5$. Definieer een nieuwe variabele $Y = 3X - 10$. Wat is de variantie van Y (σ_Y^2)?
Antwoord: C Uitleg: Gebruik de eigenschap van variantie van een lineaire transformatie: $\sigma_{aX+b}^2 = a^2\sigma_X^2$. Hier is $a = 3$ en $b = -10$. $\sigma_Y^2 = \sigma_{3X-10}^2 = (3)^2\sigma_X^2 = 9 \times 5 = \mathbf{45}$.

Rekenmachine (TI-30XB) Stappen

(a) Kwadrateer de constante a: $3^2 = 9$.
(b) Vermenigvuldig met de variantie van X: $9 \times 5 = 45$.

8) De proportie studenten die online cursussen volgt, is 0.40. Een onderzoeker wil dit schatten door een steekproef te trekken. Als de onderzoeker de standaardfout van de steekproefproportie wil halveren, met welke factor moet de steekproefomvang dan worden vergroot?
Antwoord: B Uitleg: De standaardfout van de proportie is $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$. Als we de standaardfout willen halveren, dan willen we dat de nieuwe standaardfout $\sigma'_{\hat{p}} = \frac{1}{\sigma \hat{p}}$. $\sqrt{\frac{p(1-p)}{n'}} = \frac{1}{\sqrt{\frac{p(1-p)}{n}}}$. Kwadrateer beide zijden: $\frac{p(1-p)}{n'} = \frac{1}{4}\frac{p(1-p)}{n}$. Dit impliceert dat $\frac{1}{n'} = \frac{1}{4n}$, dus $n' = 4n$. De steekproefomvang moet met een factor **4** worden vergroot.

Rekenmachine (TI-30XB) Stappen

Dit is een conceptuele vraag over de relatie tussen standaardfout en steekproefomvang. Om de standaardfout te halveren, moet je de steekproefomvang kwadrateren, dus $2^2 = 4$.

9) De tijd (in dagen) die nodig is om een patiënt te revalideren na een bepaalde operatie volgt een Exponentiële verdeling met een gemiddelde revalidatietijd van 14 dagen. Wat is de kans dat een willekeurige patiënt binnen 7 dagen gerevalideerd is?
Antwoord: A Uitleg: Dit is een Exponentiële verdeling. Het gemiddelde is $E[X] = 1/\lambda = 14$ dagen, dus $\lambda = 1/14$. We zoeken $P(X \leq 7)$. De cumulatieve verdelingsfunctie is $F(x) = 1 - e^{-\lambda x}$. $P(X \leq 7) = 1 - e^{-(1/14)\cdot 7} = 1 - e^{-0.5} \approx 1 - 0.6065 = \mathbf{0.3935}$.

Rekenmachine (TI-30XB) Stappen

(a) Bereken λ: $1/14 \approx 0.0714$.
(b) Bereken $\lambda \times x$: $(1/14) \times 7 = 0.5$.
(c) Bereken $e^{-0.5}$ met de e^x knop.
(d) Bereken $1 - e^{-0.5}$: $1 - 0.6065 = 0.3935$.

10) Voor twee toevalsvariabelen X en Y geldt: $\sigma_X^2 = 10$, $\sigma_Y^2 = 8$, en $\sigma_{XY} = 4$. Wat is de variantie van de somvariabele $X + Y$ (σ_{X+Y}^2)?
Antwoord: C Uitleg: De variantie van een somvariabele is: $\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY}$. $\sigma_{X+Y}^2 = 10 + 8 + 2(4) = 18 + 8 = \mathbf{26}$.

Rekenmachine (TI-30XB) Stappen

(a) Tel de varianties van X en Y op: $10 + 8 = 18$.

(b) Bereken $2\times$ de covariantie: $2 \times 4 = 8$.

(c) Tel de resultaten op: $18 + 8 = 26$.

11) Een onderzoeker verzamelt gegevens over de score op een geheugentest (X) en het aantal uren slaap per nacht (Y) voor een groep van 100 personen. De correlatiecoëfficiënt tussen X en Y is $r_{XY} = -0.30$. Welke van de volgende uitspraken is CORRECT op basis van deze correlatie?
Antwoord: C Uitleg: Een negatieve correlatie ($r_{XY} = -0.30$) betekent dat er een negatieve lineaire samenhang is tussen de twee variabelen. Naarmate de ene variabele toeneemt, neemt de andere af.

- A is onjuist: Hogere scores op X gaan gepaard met *minder* uren slaap (negatieve correlatie).
- B is onjuist: Er is wel een lineaire relatie, zij het een zwakkere negatieve.
- C is juist: Lagere scores op de geheugentest (X) gaan gepaard met meer uren slaap (Y), en vice versa.
- D is onjuist: Correlatie impliceert geen causaliteit. Het toont alleen een samenhang aan.

12) De lifetime van een psychologische batterij is normaal verdeeld met een gemiddelde van 800 uur en een standaarddeviatie van 50 uur. Een batterij wordt als 'defect' beschouwd als zijn lifetime minder dan 720 uur bedraagt. Wat is de Z-score die hoort bij een lifetime van 720 uur?
Antwoord: A Uitleg: De formule voor een Z-score is $Z = \frac{X-\mu}{\sigma}$. Hier is $X = 720$, $\mu = 800$, $\sigma = 50$. $Z = \frac{720-800}{50} = \frac{-80}{50} = \mathbf{-1.6}$.

Rekenmachine (TI-30XB) Stappen

(a) Bereken het verschil: $720 - 800 = -80$.

(b) Deel door de standaarddeviatie: $-80/50 = -1.6$.

13) Welke van de volgende beweringen is ALTIJD waar met betrekking tot de machtsverzameling (2^S) van een verzameling S?
Antwoord: B Uitleg: De machtsverzameling 2^S is de verzameling van alle deelverzamelingen van S.

- A ($2^{A\cup B} = 2^A \cup 2^B$): Dit is ONJUIST. Deelverzamelingen van de unie zijn niet per se alleen deelverzamelingen van A of deelverzamelingen van B. Bijvoorbeeld, als $A = \{1\}$ en $B = \{2\}$, dan $A\cup B = \{1, 2\}$. $2^{A\cup B} = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$. $2^A = \{\emptyset, \{1\}\}$ en $2^B = \{\emptyset, \{2\}\}$. Dan $2^A \cup 2^B = \{\emptyset, \{1\}, \{2\}\}$, wat niet gelijk is aan $2^{A\cup B}$.

- B ($2^{A\cap B} = 2^A \cap 2^B$): Dit is CORRECT. Een verzameling is een deelverzameling van $A \cap B$ dan en slechts dan als het een deelverzameling van A is én een deelverzameling van B. Dus de machtsverzameling van de doorsnede is gelijk aan de doorsnede van de machtsverzamelingen.
- C ($2^{A\setminus B} = 2^A \setminus 2^B$): Dit is ONJUIST. Deelverzamelingen van $A \setminus B$ zijn niet noodzakelijkerwijs A-deelverzamelingen die geen B-deelverzamelingen zijn. Bijvoorbeeld, als $A = \{1,2\}$ en $B = \{2\}$, dan $A \setminus B = \{1\}$. $2^{A\setminus B} = \{\emptyset, \{1\}\}$. $2^A = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$ en $2^B = \{\emptyset, \{2\}\}$. Dan $2^A \setminus 2^B = \{\{1\}, \{1,2\}\}$, wat niet gelijk is aan $2^{A\setminus B}$.

14) De gemiddelde reactietijd op een visuele stimulus is 350 ms met een standaarddeviatie van 60 ms. Volgens de ongelijkheid van Tchebychev, wat is de MINIMALE proportie reactietijden die binnen 2 standaarddeviaties van het gemiddelde vallen?
Antwoord: B Uitleg: De ongelijkheid van Tchebychev stelt: $P(|X-\mu| < k\sigma) \geq 1-\frac{1}{k^2}$. Hier is $k = 2$ (binnen 2 standaarddeviaties). $P(|X-\mu| < 2\sigma) \geq 1-\frac{1}{2^2} = 1-\frac{1}{4} = \frac{3}{4} = \mathbf{0.75}$. De minimale proportie is 0.75.

Rekenmachine (TI-30XB) Stappen

(a) De waarde van k is 2.

(b) Bereken $1-(1/k^2)$: $1-(1/2^2) = 1-1/4 = 0.75$.

15) Een toevalsvariabele X heeft de volgende kansmassafunctie: $p_X(x) = k \cdot x$ voor $x = 1, 2, 3, 4$. Wat is de waarde van de constante k?
Antwoord: B Uitleg: Voor een geldige kansmassafunctie moet de som van alle kansen gelijk zijn aan 1: $\sum p_X(x) = 1$. $p_X(1)+p_X(2)+p_X(3)+p_X(4) = 1$. $(k\cdot 1)+(k\cdot 2)+(k\cdot 3)+(k\cdot 4) = 1$. $k(1+2+3+4) = 1$. $k(10) = 1$. $k = \mathbf{1/10}$.

Rekenmachine (TI-30XB) Stappen

(a) Tel de waarden van x op: $1+2+3+4 = 10$.

(b) Deel 1 door de som: $1/10$.

Oplossing Open Vragen

1) Een onderzoeker wil de relatie tussen studiemotivatie (X) en examenprestaties (Y) analyseren aan de hand van een scatterplot met 5 punten. Teken een scatterplot die voldoet aan de volgende voorwaarden:

   - Alle X-waarden zijn groter dan of gelijk aan 0.
   - Het conditioneel gemiddelde van Y gegeven X, $E[Y|X]$, is constant voor alle X-waarden.
   - De verklaarde variantie van Y door X, s^2_{verkl}, is gelijk aan 0.
   - De totale variantie van Y, s^2_Y, is groter dan 0.

   Geef de coördinaten van de 5 punten en leg uit hoe je scatterplot aan elke voorwaarde voldoet.

   Uitleg en Coördinaten: De voorwaarde dat $E[Y|X]$ constant is voor alle X-waarden betekent dat de gemiddelde Y-waarde niet verandert, ongeacht de X-waarde. Dit impliceert dat er geen lineaire relatie is, en ook geen andere systematische relatie waarbij de X-waarde de gemiddelde Y-waarde beïnvloedt. De voorwaarde dat $s^2_{verkl} = 0$ betekent dat de X-variabele geen enkele variantie in Y verklaart. Dit is consistent met het constante conditionele gemiddelde. De voorwaarde dat $s^2_Y > 0$ betekent dat er wel variatie is in Y, maar deze variatie wordt niet verklaard door X. Dit betekent dat voor een gegeven X-waarde, er nog steeds spreiding in Y-waarden moet zijn. Alle X-waarden moeten ≥ 0 zijn.

   Mogelijke coördinaten van 5 punten: (X, Y)

   - Punt 1: $(1, 5)$
   - Punt 2: $(1, 7)$
   - Punt 3: $(2, 5)$
   - Punt 4: $(2, 7)$
   - Punt 5: $(3, 6)$

   Controle van de voorwaarden:

   - **Alle X-waarden zijn groter dan of gelijk aan 0:** De X-waarden zijn 1, 2, 3, die allemaal ≥ 0 zijn.
   - **Het conditioneel gemiddelde van Y gegeven X, $E[Y|X]$, is constant voor alle X-waarden:**
     - Voor $X = 1$: $E[Y|X = 1] = (5 + 7)/2 = 6$.
     - Voor $X = 2$: $E[Y|X = 2] = (5 + 7)/2 = 6$.
     - For $X = 3$: $E[Y|X = 3] = 6$.

     Het conditioneel gemiddelde van Y is constant (6) voor alle waargenomen X-waarden.
   - **De verklaarde variantie van Y door X, s^2_{verkl}, is gelijk aan 0:** Omdat de conditionele gemiddelden van Y voor elke X-waarde hetzelfde zijn (allemaal 6), is er geen variatie in de voorspelde Y-waarden op basis van X. De regressielijn zou een horizontale lijn op $Y = 6$ zijn, wat betekent dat X geen deel van de variantie in Y verklaart. Dus $s^2_{verkl} = 0$.

- **De totale variantie van Y, s_Y^2, is groter dan 0:** De Y-waarden zijn 5, 7, 5, 7, 6. Deze waarden variëren (bijv. 5 en 7 zijn verschillend). Het algemene gemiddelde van Y is $\overline{Y} = (5 + 7 + 5 + 7 + 6)/5 = 30/5 = 6$. De variantie is $s_Y^2 = \frac{\sum(Y_i - \overline{Y})^2}{n-1} = \frac{(5-6)^2+(7-6)^2+(5-6)^2+(7-6)^2+(6-6)^2}{5-1} = \frac{(-1)^2+1^2+(-1)^2+1^2+0^2}{4} = \frac{1+1+1+1+0}{4} = 1$. Aangezien $s_Y^2 = 1 > 0$, voldoet deze voorwaarde.

Rekenmachine (TI-30XB) Stappen

(a) De punten moeten handmatig worden getekend in een scatterplot.

(b) Berekening $E[Y|X]$: Voor elke unieke X-waarde, tel de corresponderende Y-waarden op en deel door het aantal waarnemingen voor die X-waarde.

(c) Berekening s_Y^2: Voer de Y-waarden in de statistiekfunctie van de rekenmachine en bereken de steekproefvariantie (S_X^2 of s_x^2 symbool, maar pas op dat dit de steekproefvariantie is, niet de populatievariantie). Bijvoorbeeld, voor de TI-30XB:

  i. Druk op **DATA**.

  ii. Voer de Y-waarden in Lijst 1 in (bijv. 5, 7, 5, 7, 6).

  iii. Druk op **STAT**. Selecteer "1-Var Stats".

  iv. Zoek naar S_x (steekproefstandaarddeviatie) en kwadrateer deze om s_Y^2 te krijgen. ($S_x^2 = 1$).

2) De kans dat een individu een bepaalde infectie (I) oploopt is 0.05. Als een individu de infectie heeft, is de kans dat de test positief (T+) is 0.98. Als een individu de infectie niet heeft, is de kans dat de test positief (T+) is 0.01 (vals positief).

  a) Wat is de kans dat een willekeurig gekozen individu een positieve testuitslag heeft?

  b) Wat is de kans dat een individu de infectie heeft, GEGEVEN dat de test positief is? (Gebruik de Regel van Bayes).

Uitleg: Gegeven: $P(I) = 0.05$ (Kans op infectie) $P(T + |I) = 0.98$ (Kans op positieve test gegeven infectie - sensitiviteit) $P(T + |I^c) = 0.01$ (Kans op positieve test gegeven geen infectie - vals positief percentage)

Hieruit kunnen we afleiden: $P(I^c) = 1 - P(I) = 1 - 0.05 = 0.95$ (Kans op geen infectie)

a) **Wat is de kans dat een willekeurig gekozen individu een positieve testuitslag heeft?** Gebruik de Wet van de Totale Kans: $P(T+) = P(T + |I)P(I) + P(T + |I^c)P(I^c)$ $P(T+) = (0.98 \times 0.05) + (0.01 \times 0.95)$ $P(T+) = 0.0490 + 0.0095$ $P(T+) =$ **0.0585**

Rekenmachine (TI-30XB) Stappen

  i. Bereken $0.98 \times 0.05 = 0.049$.

  ii. Bereken $0.01 \times (1 - 0.05) = 0.01 \times 0.95 = 0.0095$.

  iii. Tel de resultaten op: $0.049 + 0.0095 = 0.0585$.

b) **Wat is de kans dat een individu de infectie heeft, GEGEVEN dat de test positief is? (Gebruik de Regel van Bayes).** We zoeken $P(I|T+)$. Gebruik de

Regel van Bayes: $P(I|T+) = \frac{P(T+|I)P(I)}{P(T+)}$ We hebben $P(T+)$ al berekend in deel a): $P(T+) = 0.0585$. $P(I|T+) = \frac{0.98\times0.05}{0.0585}$ $P(I|T+) = \frac{0.0490}{0.0585} \approx \mathbf{0.8376}$

Rekenmachine (TI-30XB) Stappen

i. Bereken de teller: $0.98 \times 0.05 = 0.049$.
ii. Deel de teller door de noemer (resultaat van 2a): $0.049/0.0585 \approx 0.8376$.

3) Gegeven is de volgende tabel van gemiddelden voor de afhankelijk variabele Y, voorspeld door twee kwalitatieve predictoren X_1 en X_2:

	$X_2 = -1$	$X_2 = +1$
$X_1 = -1$	10	12
$X_1 = +1$	14	16

a) **Leg uit of er in dit scenario een hoofdeffect van X_1 is. Zo ja, bereken de grootte ervan.** Een hoofdeffect van X_1 betekent dat de gemiddelde Y-waarde verandert wanneer X_1 verandert, ongeacht de waarde van X_2. Bereken de rijgemiddelden (marginale gemiddelden voor X_1):

- $E[Y|X_1 = -1] = (10 + 12)/2 = 11$
- $E[Y|X_1 = +1] = (14 + 16)/2 = 15$

Aangezien $11 \neq 15$, is er een hoofdeffect van X_1. De grootte van het hoofdeffect van X_1 is het verschil tussen deze marginale gemiddelden: $15 - 11 = \mathbf{4}$.

Rekenmachine (TI-30XB) Stappen

i. Bereken de gemiddelden voor elke rij: $(10 + 12)/2 = 11$ en $(14 + 16)/2 = 15$.
ii. Bereken het verschil tussen deze gemiddelden: $15 - 11 = 4$.

b) **Leg uit of er in dit scenario een interactie-effect tussen X_1 en X_2 is. Zo ja, motiveer je antwoord met berekeningen.** Een interactie-effect betekent dat het effect van de ene variabele afhangt van de waarde van de andere variabele. Dit is te zien als de verschillen in Y over de niveaus van één variabele niet constant zijn over de niveaus van de andere variabele. Bereken de 'eenvoudige' effecten (verschillen) van X_1 voor elk niveau van X_2:

- Effect van X_1 wanneer $X_2 = -1$: $14 - 10 = 4$.
- Effect van X_1 wanneer $X_2 = +1$: $16 - 12 = 4$.

Aangezien de effecten (verschillen) van X_1 op Y constant zijn over de niveaus van X_2 $(4 = 4)$, is er **geen interactie-effect**.

Alternatieve methode (kruisverschillen): Als er geen interactie is, moeten de kruisverschillen nul zijn. Bijvoorbeeld, $(Y_{1,1} - Y_{1,2}) - (Y_{2,1} - Y_{2,2}) = 0$. $(10 - 14) - (12 - 16) = (-4) - (-4) = 0$. Aangezien de kruisverschillen nul zijn, is er **geen interactie-effect**.

Rekenmachine (TI-30XB) Stappen

i. Bereken de verschillen binnen kolommen (of rijen): $14 - 10 = 4$ $16 - 12 = 4$
ii. Als deze verschillen gelijk zijn, is er geen interactie.

c) **Pas één cel van de tabel aan zodat er geen interactie-effect meer is, maar de hoofdeffecten (indien aanwezig) behouden blijven. Leg uit welke cel je aanpast en waarom.** In dit specifieke geval is er al geen interactie-effect. De vraag lijkt te suggereren dat er wel een interactie-effect zou zijn en dat we dit zouden moeten verwijderen. Als we ervan uitgaan dat dit een algemene vraag is over het verwijderen van een interactie-effect door één cel aan te passen, dan is het doel om de 'parallelle lijnen' (of gelijke verschillen) te herstellen.

Laten we voor het voorbeeld een hypothetisch scenario nemen waarin er wél een interactie-effect was, en dan één cel aanpassen. Hypothetische starttabel (met interactie):

	$X_2 = -1$	$X_2 = +1$
$X_1 = -1$	10	12
$X_1 = +1$	14	**18**

Hier is het effect van X_1 wanneer $X_2 = -1$: $14 - 10 = 4$. Het effect van X_1 wanneer $X_2 = +1$: $18 - 12 = 6$. Omdat $4 \neq 6$, is er een interactie.

Om het interactie-effect te elimineren en de hoofdeffecten (of ten minste de gemiddelde verschillen) te behouden, moet het effect van X_1 constant zijn over de niveaus van X_2. We willen dat het verschil $16 - 12$ hetzelfde is als $14 - 10$, dus 4. Dus, de cel van 18 moet 16 zijn. Dan is $16 - 12 = 4$. Aanpassing: de cel Y als $X_1 = +1$ en $X_2 = +1$ van 18 naar 16 veranderen.

Aangepaste Cel (terug naar oorspronkelijke tabel): De cel Y als $X_1 = +1$ en $X_2 = +1$ wordt **16**. **Waarom:** Door deze cel aan te passen naar 16, wordt het effect van X_1 bij $X_2 = +1$ ook $16 - 12 = 4$, wat consistent is met het effect bij $X_2 = -1$ ($14 - 10 = 4$). Dit elimineert het interactie-effect. De rijgemiddelden en kolomgemiddelden (hoofdeffecten) kunnen dan opnieuw berekend worden om te controleren of ze ook in lijn zijn met de wens om ze te behouden (indien de vraag daar expliciet om vroeg en het mogelijk was).

In dit specifieke geval van de oefening waren de hoofdeffecten al constant, dus was er geen interactie-effect. De aanpassing is dan om te laten zien hoe je het zou doen *als* er een interactie was en deze wilde elimineren.

Rekenmachine (TI-30XB) Stappen

i. Bereken de 'eenvoudige' effecten (verschillen) voor elke rij/kolom.
ii. Als deze verschillen niet gelijk zijn, is er interactie.
iii. Om interactie te elimineren, pas één cel aan zodat de verschillen weer consistent worden. Bijvoorbeeld, als $Y_{1,1} - Y_{1,2} = \text{verschil}_1$ en $Y_{2,1} - Y_{2,2} = \text{verschil}_2$, en $\text{verschil}_1 \neq \text{verschil}_2$, dan moet je één van de Y waarden aanpassen zodat $\text{verschil}_1 = \text{verschil}_2$. Bijvoorbeeld $Y'_{2,2} = Y_{2,1} - \text{verschil}_1$.

4) De gemiddelde reactietijd van mannelijke studenten op een cognitieve taak is 400 ms met een standaarddeviatie van 60 ms. Voor vrouwelijke studenten is de gemiddelde reactietijd 380 ms met een standaarddeviatie van 50 ms. De reactietijden zijn normaal verdeeld. Een onderzoeker trekt een willekeurige steekproef van 50 mannelijke studenten en 40 vrouwelijke studenten.

a) **Wat is de verwachte waarde van de steekproevenverdeling van het verschil**

in reactietijden ($\overline{X}_M - \overline{X}_V$)? Gegeven: Mannen (M): $\mu_M = 400$, $\sigma_M = 60$, $n_M = 50$. Vrouwen (V): $\mu_V = 380$, $\sigma_V = 50$, $n_V = 40$. De verwachte waarde van het verschil in steekproefgemiddelden is het verschil in populatiegemiddelden: $E[\overline{X}_M - \overline{X}_V] = \mu_M - \mu_V = 400 - 380 = \mathbf{20}$ ms.

Rekenmachine (TI-30XB) Stappen

i. Trek het gemiddelde van vrouwen af van het gemiddelde van mannen: $400 - 380 = 20$.

b) **Wat is de standaardfout van de steekproevenverdeling van het verschil in reactietijden?** De standaardfout van het verschil tussen twee onafhankelijke steekproefgemiddelden is: $\sigma_{\overline{X}_M - \overline{X}_V} = \sqrt{\frac{\sigma_M^2}{n_M} + \frac{\sigma_V^2}{n_V}}$ $\sigma_{\overline{X}_M - \overline{X}_V} = \sqrt{\frac{60^2}{50} + \frac{50^2}{40}} = \sqrt{\frac{3600}{50} + \frac{2500}{40}}$ $= \sqrt{72 + 62.5} = \sqrt{134.5} \approx \mathbf{11.597}$ ms.

Rekenmachine (TI-30XB) Stappen

i. Bereken σ_M^2/n_M: $60^2/50 = 72$.
ii. Bereken σ_V^2/n_V: $50^2/40 = 62.5$.
iii. Tel de resultaten op: $72 + 62.5 = 134.5$.
iv. Neem de wortel van de som: $\sqrt{134.5} \approx 11.597$.

c) **Wat is de kans dat het gemiddelde van de mannelijke studenten in de steekproef minder dan 10 ms langer is dan die van de vrouwelijke studenten?** We zoeken $P(\overline{X}_M - \overline{X}_V < 10)$. De steekproevenverdeling van het verschil ($\overline{X}_M - \overline{X}_V$) is normaal verdeeld met gemiddelde $E[\overline{X}_M - \overline{X}_V] = 20$ en standaardfout $\sigma_{\overline{X}_M - \overline{X}_V} = 11.597$. Transformeer de waarde 10 naar een Z-score: $Z = \frac{10 - E[\overline{X}_M - \overline{X}_V]}{\sigma_{\overline{X}_M - \overline{X}_V}} = \frac{10-20}{11.597} = \frac{-10}{11.597} \approx -0.862$. We zoeken $P(Z < -0.862)$. Uit de Z-tabel: $P(Z < -0.86) \approx 0.1949$. $P(\overline{X}_M - \overline{X}_V < 10) \approx \mathbf{0.1949}$.

Rekenmachine (TI-30XB) Stappen

i. Bereken de teller van de Z-score: $10 - 20 = -10$.
ii. Deel door de standaardfout (uit b): $-10/11.597 \approx -0.862$.
iii. Zoek de cumulatieve kans op in een Z-tabel voor -0.862.

5) Gegeven een discrete toevalsvariabele X met de volgende kansmassafunctie:

x	1	2	3
$p_X(x)$	0.3	0.5	0.2

We trekken een steekproef van $n = 2$ onafhankelijke waarnemingen uit deze populatie, X_1 en X_2.

a) **Construeer de kansmassafunctie van het steekproefgemiddelde $\overline{X}$.** Het steekproefgemiddelde is $\overline{X} = (X_1 + X_2)/2$. Mogelijke sommen $X_1 + X_2$:

- $1 + 1 = 2 \Rightarrow \overline{X} = 1$

- $1 + 2 = 3 \Rightarrow \overline{X} = 1.5$
- $1 + 3 = 4 \Rightarrow \overline{X} = 2$
- $2 + 1 = 3 \Rightarrow \overline{X} = 1.5$
- $2 + 2 = 4 \Rightarrow \overline{X} = 2$
- $2 + 3 = 5 \Rightarrow \overline{X} = 2.5$
- $3 + 1 = 4 \Rightarrow \overline{X} = 2$
- $3 + 2 = 5 \Rightarrow \overline{X} = 2.5$
- $3 + 3 = 6 \Rightarrow \overline{X} = 3$

Kansen voor $\overline{X}$ (product van kansen omdat X_1, X_2 onafhankelijk zijn):

- $P(\overline{X} = 1) = P(X_1 = 1, X_2 = 1) = p_X(1) \cdot p_X(1) = 0.3 \cdot 0.3 = \mathbf{0.09}$
- $P(\overline{X} = 1.5) = P(X_1 = 1, X_2 = 2) + P(X_1 = 2, X_2 = 1) = (0.3 \cdot 0.5) + (0.5 \cdot 0.3) = 0.15 + 0.15 = \mathbf{0.30}$
- $P(\overline{X} = 2) = P(X_1 = 1, X_2 = 3) + P(X_1 = 2, X_2 = 2) + P(X_1 = 3, X_2 = 1) = (0.3 \cdot 0.2) + (0.5 \cdot 0.5) + (0.2 \cdot 0.3) = 0.06 + 0.25 + 0.06 = \mathbf{0.37}$
- $P(\overline{X} = 2.5) = P(X_1 = 2, X_2 = 3) + P(X_1 = 3, X_2 = 2) = (0.5 \cdot 0.2) + (0.2 \cdot 0.5) = 0.10 + 0.10 = \mathbf{0.20}$
- $P(\overline{X} = 3) = P(X_1 = 3, X_2 = 3) = 0.2 \cdot 0.2 = \mathbf{0.04}$

Kansmassafunctie van $\overline{X}$:

$\overline{x}$	1	1.5	2	2.5	3
$p_{\overline{X}}(\overline{x})$	0.09	0.30	0.37	0.20	0.04

Controle: $0.09 + 0.30 + 0.37 + 0.20 + 0.04 = 1.00$.

Rekenmachine (TI-30XB) Stappen

i. Maak een tabel van alle mogelijke paren (X_1, X_2) en hun gemiddelde $(X_1 + X_2)/2$.
ii. Bereken de kans voor elk paar door de kansen van X_1 en X_2 te vermenigvuldigen.
iii. Groepeer de kansen voor dezelfde $\overline{X}$-waarden.
iv. Tel de kansen op om te controleren of de som 1 is.

b) **Bereken de verwachte waarde van $\overline{X}$ ($E[\overline{X}]$).** Eerste methode (via de kansmassafunctie van $\overline{X}$): $E[\overline{X}] = (1 \cdot 0.09) + (1.5 \cdot 0.30) + (2 \cdot 0.37) + (2.5 \cdot 0.20) + (3 \cdot 0.04)$ $= 0.09 + 0.45 + 0.74 + 0.50 + 0.12 = \mathbf{1.90}$.

Tweede methode (via populatiegemiddelde): $E[X] = (1 \cdot 0.3) + (2 \cdot 0.5) + (3 \cdot 0.2) = 0.3 + 1.0 + 0.6 = 1.9$. Aangezien $E[\overline{X}] = E[X]$, is $E[\overline{X}] = \mathbf{1.90}$.

Rekenmachine (TI-30XB) Stappen

i. (Methode 1) Vermenigvuldig elke $\overline{x}$ met de corresponderende kans $p_{\overline{X}}(\overline{x})$ en tel op.
ii. (Methode 2) Bereken het populatiegemiddelde $E[X]$: $(1 \times 0.3) + (2 \times 0.5) + (3 \times 0.2) = 1.9$. De verwachte waarde van het steekproefgemiddelde is gelijk aan het populatiegemiddelde.

c) **Bereken de variantie van $\overline{X}$ ($\sigma^2_{\overline{X}}$).** Eerste methode (via $E[\overline{X}^2]$): $E[\overline{X}^2] = (1^2 \cdot 0.09) + (1.5^2 \cdot 0.30) + (2^2 \cdot 0.37) + (2.5^2 \cdot 0.20) + (3^2 \cdot 0.04) = (1 \cdot 0.09) + (2.25 \cdot 0.30) + (4 \cdot 0.37) + (6.25 \cdot 0.20) + (9 \cdot 0.04) = 0.09 + 0.675 + 1.48 + 1.25 + 0.36 = \mathbf{3.855}$. $\sigma^2_{\overline{X}} = E[\overline{X}^2] - (E[\overline{X}])^2 = 3.855 - (1.9)^2 = 3.855 - 3.61 = \mathbf{0.245}$.
Tweede methode (via populatievariantie): Eerst $\sigma^2_X = E[X^2] - (E[X])^2$. $E[X^2] = (1^2 \cdot 0.3) + (2^2 \cdot 0.5) + (3^2 \cdot 0.2) = 0.3 + 2.0 + 1.8 = 4.1$. $\sigma^2_X = 4.1 - (1.9)^2 = 4.1 - 3.61 = 0.49$. Voor onafhankelijke waarnemingen geldt $\sigma^2_{\overline{X}} = \frac{\sigma^2_X}{n}$. $\sigma^2_{\overline{X}} = \frac{0.49}{=}\mathbf{0.245}$.

Rekenmachine (TI-30XB) Stappen

i. (Methode 1) Bereken $E[\overline{X}^2]$ door $\overline{x}^2$ te vermenigvuldigen met $p_{\overline{X}}(\overline{x})$ en op te tellen. Trek vervolgens $(E[\overline{X}])^2$ af.

ii. (Methode 2) Bereken de populatievariantie σ^2_X:
   A. Voer de x-waarden in Lijst 1 en de $p_X(x)$ waarden in Lijst 2 van de TI-30XB.
   B. Ga naar STAT, selecteer "1-Var Statsën zorg dat Freq: L2 is.
   C. Zoek σ_x (populatiestandaarddeviatie) en kwadrateer deze om σ^2_X te krijgen. ($\sigma_x \approx 0.7$, $\sigma^2_x \approx 0.49$).

iii. Deel σ^2_X door n: $0.49/2 = 0.245$.

Nawoord

Met veel toewijding en zorg heb ik deze bundel samengesteld, speciaal voor jou, de psychologiestudent. Ik hoop oprecht dat deze oefeningen je niet alleen hebben geholpen maar ook dat het je zelfvertrouwen heeft gegeven om het zelf voor het echt te kunnen doen.

Het pad van een psychologiestudent door de statistiek is vaak een uitdagend parcours, en het is mijn missie bij StatHulp om dat pad een stukje toegankelijker te maken. Deze bundel is een eerste stap in die richting.

Ik sta klaar om je te ondersteunen gedurende je hele studietraject. Mocht je opmerkingen, vragen, of suggesties hebben over deze bundel, aarzel dan niet om contact met ons op te nemen. Je feedback is van onschatbare waarde.

Op naar de volgende practicumsessies en examens!

Met vriendelijke groet,

Peter Westgate

StatHulp Team
stathulp.be
stathulp@gmail.com