

# Tentamen Statistiek MBW/KW deel 2 29-07-2022

Afdeling: Propedeuse MBW/KW 2021-2022

Examinator: Dr. J.B.M. Melissen, Dr. R.J. Nijboer

Datum: vrijdag 29 juli 2022 09:00 – 11:00 (adelborsten 08:10 – 10.10), duur tentamen: 2 uur

- 1. Alle antwoorden moeten gemotiveerd worden!**
2. Rond eindantwoorden (kommagetallen) af op *vier* decimalen, tenzij anders vermeld.
3. Boeken, reader en aantekeningen mogen worden geraadpleegd.
4. De aanwezigheid van *communicatieapparatuur* is niet toegestaan.
5. Het gebruik van een (grafische) rekenmachine met statistische programmatuur en het raadplegen van de bijbehorende handleiding is toegestaan. Het *statistische* gebruik van deze rekenmachine is bij een aantal onderdelen ingeperkt. Let op de aanwijzingen!
- 6. De opgaven dienen na afloop van het tentamen ingeleverd te worden.**

Dit tentamen bestaat uit vier opgaven (25, 20, 30, 25 punten). Score = Puntentotaal/9

### Opgave 1 (Totaal 25 punten)

Tijdens een landmachtoefening wordt een week lang het dagelijks brandstofverbruik bijgehouden. Dit leidt tot de volgende waarden in liters: 4215, 4879, 6021, 2983, 2974, 5003, 6972. Neem aan dat de hoeveelheden  $x$  (kansvariabele die het brandstofverbruik per dag in liters voorstelt) normaal verdeeld zijn, elke dag met dezelfde verwachtingswaarde en standaarddeviatie.

**1a [5pt]** Bereken van de gemeten waarden het steekproefgemiddelde en de steekproefstandaarddeviatie.

**1b [5pt]** Bereken een 95% betrouwbaarheidsinterval voor het verwachte dagelijkse brandstofverbruik  $\mu$  van een dergelijke oefening, op grond van bovengenoemde steekproef, zonder daarbij gebruik te maken van de optie TESTS/Interval van de grafische rekenmachine. Rond de grenzen van dit interval af op gehele liters en wel zodanig dat de betrouwbaarheid gewaarborgd blijft.

**1c [5pt]** Toets:  $H_0: \mu \leq 3000$  tegen  $H_1: \mu > 3000$ . Bepaal de toetsuitslag door het berekenen van een kritiek gebied op basis van de gegeven steekproef van zeven dagen brandstofverbruik. Kies als onbetrouwbaarheid  $\alpha = 0,1$ .

Leg in simpele bewoordingen uit wat de uitslag van deze toets betekent voor het dagelijks waterverbruik.

**1d [5pt]** Hoeveel liter brandstof moet er **dagelijks** minimaal op voorraad zijn, wil met 98% zekerheid aan de dagelijkse behoefte kunnen worden voldaan? (Antwoord in gehele liters).

**1e [5pt]** Hoeveel liter brandstof moet er **op weekbasis** minimaal op voorraad zijn wil met 98% zekerheid aan de dagelijkse behoefte kunnen worden voldaan? (Antwoord in gehele liters).

Leg uit waarom deze hoeveelheid niet gelijk is aan zevenmaal de hoeveelheid die in 1d is berekend.

**Opgave 2 (Totaal 20 punten).** Bij een groot bedrijf zijn in een jaar (50 weken) 432 sollicitatiegesprekken gevoerd. In de onderstaande tabel is te zien hoeveel gesprekken er per week plaatsvonden.

Gesprekken per week	Frequentie
0 - 4	4
5 - 9	26
10 - 14	18
15 - 19	2

**2a [5pt]** Bereken/schat met behulp van de gegevens uit de tabel het aantal weken en het aantal gesprekken waarop de tabel is gebaseerd.

**2b [10pt]** Toets of het aantal gesprekken per week is te beschouwen als een kansvariabele die een Poissonverdeling volgt, door middel van uitrekenen van een  $p$ -waarde. Kies als betrouwbaarheid 95% en gebruik in je berekening de verwachte frequenties in één decimaal nauwkeurig.

**2c [5pt]** Voer de toets ook uit door berekening van het kritieke gebied.

**Opgave 3 (Totaal 30 punten).** Onlangs namen 85 eerstejaars MBW en KW cadetten en adelborsten deel aan een tentamen Statistiek. In de tabel hieronder is per krijgsmachtdeel te zien hoeveel studenten wel en hoeveel niet slaagden voor dit tentamen.

Krijgsmacht-Deel ↓	Niet geslaagd	Geslaagd	Totaal
KL	16	13	29
KLU	22	5	27
KM	8	12	20
KMAR	6	3	9
Totaal	52	33	85

**3a [10pt]** Zoek uit of je met behulp van een homogeniteitstoets de volgende hypothese met betrouwbaarheid van 97% kunt verwerpen:

$H_0$ : Het wel of niet slagen voor dit tentamen is onafhankelijk van het krijgsmachtdeel van de studenten.

Houd hierbij geen rekening met het feit dat het verwachte aantal geslaagden bij de KMAR eigenlijk te laag is om deze toets toe te passen.

**3b [2pt]** Welke groep studenten levert de grootste bijdrage aan de conclusie in 3a?

**3c [2pt]** Zoals al werd vermeld blijkt dat het verwachte aantal geslaagde marechaussees te laag is om de homogeniteitstoets uit te voeren. Leg uit waarom je dit probleem niet kunt oplossen door twee kolommen in de tabel samen te voegen.

**3d [3pt]** Het probleem is wel op te lossen door twee rijen van de tabel samen te voegen. Kun je de KMAR beter samenvoegen met de marine of met de landmacht? Geef een statistische verklaring voor je antwoord.

**3e [5pt]** Bereken met de gegevens uit de tabel een 90% betrouwbaarheidsinterval voor het slaagpercentage van de KLU cadetten.

**3f [5pt]** Bereken met de gegevens uit de tabel een 90% betrouwbaarheidsinterval voor het slaagpercentage van de adelborsten.

**3g [3pt]** Als het goed is zul je zien dat de twee intervallen uit 3e en 3f elkaar niet overlappen. Welke conclusie kun je uit dat gegeven trekken? Met welke betrouwbaarheid?

#### Opgave 4 (Totaal 25 punten)

Na het uitvoeren van een oefening buiten Defensierrein in Nederland komen bij Defensie altijd schadeclaims binnen van particulieren, boeren en gemeentes over (vermeende) beschadigingen aan infrastructuur, gewassen, en over aangetroffen vervuilingen. In de tabel hieronder zijn voor zes oefeningen de geclaimde kosten in kaart gebracht

	1	2	3	4	5	6
<b>Duur oefening (dagen)</b>	2	8	4	2	4	10
<b>Kosten (1000 euros)</b>	20	58	32	16	42	80

**4a [10pt]** Bereken met behulp van een tabel de correlatiecoëfficiënt van Pearson. Bepaal of er sprake is van een lineaire correlatie tussen de duur van de oefening en de veroorzaakte kosten. Leg uit hoe daarbij het teken en de grootte van de berekende coëfficiënt een rol spelen.

**4b [8pt]** Bereken de regressielijn door berekening van de coëfficiënten  $a$  en  $b$  met behulp van een tabel.

**4c [2pt]** Bereken met de regressielijn een statistisch verantwoorde voorspelling voor de te verwachten kosten voor een oefening die zes dagen gaat duren.

**4d [5pt]** Bereken een 95% voorspellingsinterval voor de waarde die in 4c is berekend.