

# Oefentamen Statistiek KW/MBW (deel 2)

Afdeling: Propedeuse KW/MBW 2020-2021

Examinator: Dr. J.B.M. Melissen

Datum: 15 juli 2021, 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 (30, 20, 20, 30 punten). Score = Puntentotaal/10

## Opgave 1 (Totaal 30 punten)

Tijdens een landmachtoefening wordt een week lang dagelijks bijgehouden hoeveel liter drinkwater wordt verbruikt. Dit leidt tot de volgende waarden: 5215, 5879, 7021, 3983, 3974, 6003, 7972. Neem aan dat de hoeveelheden  $x$  (kansvariabele die het waterverbruik 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. [7pt]** Bereken een 96% betrouwbaarheidsinterval voor het verwachte dagelijkse waterverbruik  $\mu$  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 96% betrouwbaarheid gewaarborgd blijft.

**1c. [8pt]** Toets:  $H_0: \mu \leq 4630$  tegen  $H_1: \mu > 4630$ . Bepaal de toetsuitslag door het berekenen van een kritiek gebied op basis van de gegeven steekproef van zeven dagen waterverbruik. Kies in dit geval als onbetrouwbaarheid  $\alpha = 0,05$ .

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

**1d. [5pt]** Hoeveel liter water 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 water 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).**

Het aantal acute meldingen per week waarvoor de Explosieven Opruimingsdienst moest uitrukken is gedurende 40 weken geregistreerd (zie tabel).

Meldingen per week	Frequentie
0	10
1	16
2	9
3	5
$\geq 4$	2
Totaal	42

**2a. [4pt]** Leg uit waarom het aannemelijk is dat aantal meldingen per week beschreven kan worden door een Poissonverdeling. Welke waarde van  $\mu$  kies je daarbij?

**2b. [10pt]** Toets of het aantal meldingen per week is te beschouwen als een kansvariabele die een Poissonverdeling volgt. Doe deze toetsing door middel van uitrekenen van de  $p$ -waarde. Kies als betrouwbaarheid 98%.

**2c. [6pt]** Bereken een 95% betrouwbaarheidsinterval voor  $\mu$ .

### Opgave 3 (Totaal 20 punten).

Uit gegevens van het RIVM is een overzicht gemaakt van de vijftigplussers die tot en met mei 2021 zijn overleden aan corona (drie kolommen: Overleden man / vrouw / totaal), uitgesplitst naar geslacht en leeftijd.

Verder is uit gegevens van het CBS op basis van de overleden totalen per leeftijdscategorie (kolom Overleden totaal) geschat hoe deze totalen normaal gesproken (buiten corona) zouden zijn verdeeld over mannen en vrouwen (zie kolommen Overleden verwacht m / v).

Leeftijd	Overleden man	Overleden vrouw	Overleden totaal	Overleden verwacht m	Overleden verwacht v
50-59	246	124	370	217	153
60-69	820	463	1283	752	531
70-79	2774	1554	4326	2538	1788
80-89	4140	3389	7529	3259	4270
90-99	1335	2246	3581	1550	2031
Totaal	9315	7776	17089	8316	8773

**3a. [10pt]** Voer een  $\chi^2$  homogeniteitsanalyse uit op de kolommen “Overleden man” en “Overleden vrouw”. Bereken daarvoor de waarde van  $\chi^2$  en het kritieke gebied bij een betrouwbaarheid van 99%.

**3b. [10pt]** Voer een  $\chi^2$  aanpassingsanalyse uit op de kolommen “Overleden man” en “Overleden verwacht man”. Bereken daarvoor de waarde van  $\chi^2$  en de  $p$ -waarde en houd een betrouwbaarheid van 99% aan.

Leeftijd	Overleden man	Overleden verwacht m
50-59	246	217
60-69	820	752
70-79	2774	2538
80-89	4140	3259
90-99	1335	1550
Totaal	9315	8316

#### Opgave 4 (Totaal 30 punten)

In de tabel hieronder is van zes studenten het eindexamencijfer wiskunde en het eindcijfer Statistiek vermeld.

Student →	1	2	3	4	5	6
Eindexamencijfer wiskunde	8	6	7	6	9	8
Cijfer Statistiek	7,6	5,7	7,5	5,8	8,8	7,2

**4a [8pt]** Bereken handmatig de correlatiecoëfficiënt van Pearson. Bepaal of er sprake is van een lineaire correlatie tussen de twee cijfers.

**4b [12pt]** Bereken de regressielijn, maak een schets van de puntenwolk en en bepaal hiermee een voorspelling van het verwachte Statistiekcijfer dat een student haalt als hij op zijn eindexamen een 5 zou hebben gehaald.

**4c [10pt]** Bereken een 95% voorspellingsinterval voor het cijfer Statistiek van de student die een 5 op zijn eindexamen haalde voor wiskunde.

===== XXXXXXXX =====