

Tentamen Statistiek KW/MBW (deel 2, finale kans)

Afdeling: Propedeuse KW/MBW 2020-2021

Examinator: Dr. J.B.M. Melissen

Datum: 14 februari 2022, 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, 25, 30 punten). Score = Puntentotaal/10

Opgave 1 (Totaal 25 punten)

Uit het *Annual report Surveillance of influenza and other respiratory infections in the Netherlands: winter 2019/2020* van het RIVM en onderzoek van het Nivel (Nederlands instituut voor onderzoek van de gezondheidszorg) blijkt dat in het seizoen 2019/2020 de vaccinatiegraad voor de hele Nederlandse bevolking 20,8% bedroeg en dat gevaccineerden 48% minder kans op griep hadden. Verder kregen naar schatting 400.000 mensen in dit seizoen de griep en de oversterfte als gevolg van de griep bedroeg naar schatting 600 personen. De bevolking van Nederland bedroeg 17,4 miljoen mensen.

1a. [5pt] Geef voor het seizoen 2019/2020 een schatting van het aantal griepgevallen en het aantal doden als gevolg van griep bij Defensie (68.000 personen). Ga ervan uit dat voor defensiepersoneel dezelfde percentages gelden als voor de Nederlandse bevolking.

1b. [10pt] Bereken voor het seizoen 2019/2020 een 95% schattingsinterval voor het aantal griepgevallen bij Defensie.

1c. [5pt] Bereken hoeveel doden er in Nederland in het seizoen 2019/2020 naar schatting zijn voorkomen door de griepvaccinatie (Hint: Noem de kans dat iemand griep krijgt p en bereken hiermee voor de gevaccineerden en voor de ongevaccineerden hoeveel griepgevallen er zijn).

1d. [5pt] Bij Defensie overleden zes medewerkers als gevolg van de griep (fictief aantal). Kan hieruit worden geconcludeerd dat de kans op overlijden als gevolg van griep groter is dan het landelijk gemiddelde? Welke nauwkeurigheid zou hiervoor moeten worden gehanteerd?

Opgave 2 (Totaal 20 punten). Het aantal acute meldingen per week waarvoor de Explosieven Opruimingsdienst moest uitrukken is gedurende 42 weken geregistreerd (zie onderstaande tabel).

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

2a. [2pt] Hoe vaak is de EOD in deze 42 weken uitgerukt?

2b. [3pt] Leg uit waarom het aannemelijk is dat aantal meldingen per week beschreven kan worden door een Poissonverdeling. Welke waarde van μ kies je daarbij?

2c. [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 een kritiek gebied. Kies als betrouwbaarheid 95%.

2d. [5pt] Bereken een 95% betrouwbaarheidsinterval voor μ .

Opgave 3 (Totaal 25 punten). Uit gegevens van het RIVM is een overzicht gemaakt van de vijftigplussers die tot juli 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) een schatting gemaakt over 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 χ^2 homogeniteitsanalyse uit op de kolommen “Overleden man” en “Overleden vrouw”. Bereken daarvoor de waarde van χ^2 en de p -waarde.

3b. [10pt] Voer een χ^2 aanpassingsanalyse uit op de kolommen “Overleden man” en “Overleden verwacht man”. Bereken daarvoor de waarde van χ^2 en de p -waarde.

3c. [5pt] Doe aan de hand van je berekeningen zo sterk mogelijke, maar statistisch verantwoorde uitspraken voor conclusies die je kunt trekken, **zowel 3a, als voor 3b.**

Bijvoorbeeld (fictief):

- Het aantal vrouwen en mannen dat overlijdt aan corona is niet verschillend bij een significantieniveau van $\alpha = \dots$
- Er kan met een significantieniveau van $\alpha = \dots$ geconcludeerd worden dat het overlijden door corona afhankelijk is van leeftijd.

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	7,3	5,7	7,5	5,8	8,8	7,2

4a [10pt] Bereken handmatig de correlatiecoëfficiënt van Pearson. Bepaal of er sprake is van een lineaire correlatie tussen de twee cijfers.

4b [10pt] Bereken de regressielijn en bepaal hiermee een statistisch verantwoorde 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 Statistiekcijfer van iemand die een 10 als eindcijfer voor wiskunde haalde.

===== XXXXXXXX =====