

# Tentamen Statistiek MBW/KW (deel 2, tweede kans)

Afdeling: Propedeuse MBW/KW 2020-2021

Examinator: Dr. J.B.M. Melissen

Datum: 19 november 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 drie opgaven (55, 25, 20 punten). Score = Puntentotaal/10.

## Opgave 1 (Totaal 55 punten)

Op 3 november 2021 kwam het RIVM met een rapportage over de effectiviteit van COVID-19 vaccins over de periode 11 juli en 31 oktober 2021. De gegevens waren gebaseerd op de volgende gegevens:

*Tabel 1 Aantal ziekenhuis- en IC-opnames met COVID-19 naar vaccinatiestatus per maand\* en percentage niet, deels en volledig gevaccineerden. Bij het berekenen van de percentages is de groep met onbekende vaccinatiestatus niet meegerekend in de noemer.*

<b>Ziekenhuisopnames<sup>#</sup></b>								
<b>Maand 2021</b>	<b>Totaal</b>	<b>Niet gevaccineerd</b>		<b>Deels gevaccineerd</b>		<b>Volledig gevaccineerd</b>		<b>Vaccinatiestatus onbekend</b>
		aantal	%	aantal	%	aantal	%	
Juli	1498	1054	73%	139	10%	244	17%	61
Augustus	1961	1381	74%	66	4%	409	22%	105
September	1386	940	71%	31	2%	352	27%	63
Oktober	2054	1064	55%	21	1%	841	44%	128
<b>Totaal</b>	<b>6899</b>	<b>4439</b>	<b>68%</b>	<b>257</b>	<b>4%</b>	<b>1846</b>	<b>28%</b>	<b>357</b>
<b>IC-opnames</b>								
<b>Maand 2021</b>	<b>Totaal</b>	<b>Niet gevaccineerd</b>		<b>Deels gevaccineerd</b>		<b>Volledig gevaccineerd</b>		<b>Vaccinatiestatus onbekend</b>
		aantal	%	aantal	%	aantal	%	
Juli	313	230	78%	30	10%	34	12%	19
Augustus	435	335	83%	10	2%	58	14%	32
September	290	219	82%	3	1%	45	17%	23
Oktober	387	239	69%	2	1%	105	30%	41
<b>Totaal</b>	<b>1425</b>	<b>1023</b>	<b>78%</b>	<b>45</b>	<b>3%</b>	<b>242</b>	<b>18%</b>	<b>115</b>

\*COVID-19 ziekenhuis- en IC-opnames geregistreerd in NICE van 11 juli tot en met 31 oktober 2021. Door de start van de delta variant periode is de maand juli niet compleet en door rapportagevertraging is de maand oktober nog niet compleet.

<sup>#</sup> Totaal aantal ziekenhuisopnames, inclusief de IC-opnames

In het rapport werden de volgende beweringen gedaan:

1. De COVID-19 vaccin-effectiviteit (VE) is in Nederland nog steeds hoog, namelijk 94% (95% BI 94-94) tegen ziekenhuisopname en 97% (95% BI 96-97) tegen intensive care (IC)-opname over de hele periode juli-oktober.
2. Op dit moment is 84 procent van iedereen van 18 jaar en ouder gevaccineerd.
3. De werkzaamheid van COVID-19 vaccinatie tegen ziekenhuisopname (94%) en IC-opname (97%) over de periode juli-oktober is hoog.
4. Dit betekent dat de kans voor volledig gevaccineerden om met COVID-19 in het ziekenhuis opgenomen te worden 17 keer kleiner is dan voor ongevaccineerden. De kans om op de IC opgenomen te worden is 33 keer kleiner.

De VE is de kans dat een gevaccineerde persoon in een bepaalde periode niet hoeft te worden opgenomen.

**1a. [10pt]** Bereken op grond van de gegevens uit Tabel 1 de vergelijking van de regressielijn voor de totale ziekenhuisopnamen door COVID per maand als functie van de maanden juli (1), augustus (2), september (3) en oktober (4) (gebruik deze nummers voor de maanden in je berekeningen).

Bereken ook de correlatiecoëfficiënt van deze lijn. Geef de coëfficiënten van de lijn en de correlatiecoëfficiënt in drie cijfers nauwkeurig.

**1b. [10pt]** Gebruik de lijn  $y = 109x + 1452$  (met  $x$  het maandnummer zoals eerder genoemd, en  $y$  de VE in procenten, om een 95% voorspellingsinterval te berekenen voor de VE in de maand november.

**1c. [10pt]** Toets de bewering

**H<sub>0</sub>:** De vergelijking die werd genoemd in 1c geeft een goede weerspiegeling van de totale ziekenhuisopnamen over de maanden juli tot en met oktober.

Doe hiervoor een  $\chi^2$  aanpassingstoets met een betrouwbaarheidsniveau van 97% en werk in de berekening met een kritiek gebied.

Formuleer een begrijpelijke conclusie van je berekening.

**1d. [10pt]** Onderzoek met behulp van een  $\chi^2$  homogeniteitstoets of er een afhankelijkheid bestaat tussen de volledig gevaccineerden met IC opname en de deels gevaccineerden met IC opname. Gebruik een betrouwbaarheidsniveau van 99% en doe de berekeningen met behulp van een  $p$ -waarde.

Formuleer een begrijpelijke conclusie van je berekening.

**1e. [10pt]** In Bewering 1 wordt een 95% betrouwbaarheidsinterval (BI) voor de VE voor ziekenhuisopname genoemd van (94, 94). Dit is merkwaardig, want dit is een interval met lengte 0, dat zou dus betekenen dat de werkelijke waarde van VE **gelijk is** aan 94% met een betrouwbaarheid van 95%. Bereken dit interval (gebruik de binomiale verdeling) en geef het resultaat in twee decimalen nauwkeurig.

**1f. [5pt]** In bewering 4 wordt gesteld dat de kans om op de IC te worden opgenomen voor volledig gevaccineerden 33 keer kleiner is dan voor ongevaccineerden. Dit zou betekenen dat

$$\frac{P(\text{IC opname} \mid \text{niet gevaccineerd})}{P(\text{IC opname} \mid \text{volledig gevaccineerd})} = 33$$

Ofwel:

$$\frac{P(\text{IC opname én niet gevaccineerd})}{P(\text{IC opname én volledig gevaccineerd})} \cdot \frac{P(\text{volledig gevaccineerd})}{P(\text{niet gevaccineerd})} = 33$$

Ga met behulp van de gegevens uit Tabel 1 of dit juist is of niet. (Je hebt tijdens je berekening het totale aantal inwoners van Nederland nodig, maar dat valt uiteindelijk weer weg, dus het maakt niet uit wat je daar doet). Doe hiervoor drie verschillende berekeningen:

1. Bekijk alleen de volledig gevaccineerden en de niet gevaccineerden.
2. Neem de Deels gevaccineerden en de Vaccinatiestatus onbekend wel mee en neem aan dat die allemaal niet gevaccineerd zijn.
3. Neem de Deels gevaccineerden en de Vaccinatiestatus onbekend wel mee en neem aan dat die allemaal wel gevaccineerd zijn.

Maak eventueel gebruik van Bewering 2 en dat de verdeling over de groepen Volledig gevaccineerd, Niet gevaccineerd, Deels gevaccineerd en Vaccinatiestatus onbekend in de hele populatie gelijk is aan de verdeling zoals die bij de IC opnames is geconstateerd.

Trek een conclusie uit de drie getallen die je hebt berekend.

## Opgave 2 (Totaal 25 punten)

De bottelier op een mijnenveger van de Koninklijke Marine is verantwoordelijk voor de bevoorrading van proviand. De mijnenveger heeft een bemanning van 40 opvarenden en de laatste acht maanden is het maandelijks verbruik van aardappels als volgt geweest:

400, 340, 390, 370, 370, 400, 320, 410 kg aardappels.

**2a. [5pt]** Bereken van de bekende waarden het steekproefgemiddelde en de steekproefstandaarddeviatie.

**2b. [2pt]** In welke eenheid zou de steekproefvariantie moeten worden uitgedrukt?

Opties:

- Geen, gewoon een getal, dimensieloos
- In aardappels
- In kilo aardappels
- In kilo vierkante aardappels
- In vierkante kilo aardappels (kg<sup>2</sup>)

**2c. [8pt]** Bereken een 95% betrouwbaarheidsinterval voor het verwachte maandelijkse aardappelverbruik op grond van bovengenoemde steekproef, zonder daarbij gebruik te maken van de optie TESTS/Interval van de grafische rekenmachine. De aardappels worden geleverd in zakken van 10 kg, dus geef de grenzen van het interval in aantal zakken. Leg uit hoe en waarom je de grenzen hebt afgerond.

**2d. [5pt]** Toets:  $H_0$ : Elke maand zijn minstens 31 zakken van 10 kg nodig. Kies daarbij als onbetrouwbaarheid  $\alpha = 0,05$ .

Leg in simpele bewoordingen uit wat de uitslag van deze toets betekent voor het maandelijks aardappelverbruik.

**2e. [5pt]** Hoeveel 10 kg zakken moeten er op een missie van twee maanden worden meegenomen om met 99% zekerheid te garanderen dat er zonder tussentijdse bevoorrading voldoende aardappels zijn.

### Opgave 3 (Totaal 20 punten)

Het wekelijks aantal meldingen van inbreuk op persoonlijke integriteit bij Defensie is gedurende een jaar (52 weken) geregistreerd (zie tabel, de gegevens zijn volstrekt fictief. Het werkelijk aantal integriteitsschendingen van het type “Ongewenste omgangsvormen” was volgens het Jaarverslag Integriteit Defensie 2020: 82).

Meldingen per week	Frequentie
0	18
1	21
2	9
3	4
$\geq 4$	0

**3a. [5pt]** Hoeveel meldingen hebben in dit jaar plaatsgevonden?

**3b. [10pt]** Toets of het aantal meldingen per week is te beschouwen als een kansvariabele die een Poissonverdeling volgt met  $\mu = 1$ . Gebruik hierbij als betrouwbaarheid 97%.

**3c. [5pt]** Leg uit waarop het niet onlogisch is te verwachten dat hier de Poissonverdeling met  $\mu = 1$  een rol zou kunnen spelen.

===== XXXXXXXX =====