

## Hertentamen I00099

*Vermeld op ieder blad je naam en studentnummer. Lees eerst de opgaven voor dat je aan de slag gaat. Geef uitleg over je oplossingen, antwoorden zonder heldere afleiding worden als niet gegeven beschouwd!*

*Het gebruik van een rekenmachine voor de uitwerking van numerieke resultaten (zo als  $\sqrt{\pi}$  of  $\log(2)$ ) is toegestaan, maar niet het gebruik van de statistische functies.*

*De uitgegeven tabellen mogen natuurlijk wel gebruikt worden.*

### Opgave 1. (25 punten)

In een aselechte steekproef van 55 studenten geven 23 studenten aan dat ze bekend met de normale verdeling zijn.

- (i) Bepaal betrouwbaarheidsintervallen op de levels 90%, 95% en 99% voor het relatieve aantal  $\bar{p}$  van studenten die de normale verdeling kennen.
- (ii) Hoe groot moet de steekproef zijn om op de drie betrouwbaarheidslevels uit deel (i) te kunnen concluderen dat minder dan de helft van de studenten de normale verdeling kent. Hierbij wordt verondersteld, dat ook bij grotere steekproeven de gevonden relatieve frequentie  $\hat{p} = \frac{23}{55}$  is.
- (iii) Beschrijf met je eigen woorden de samenhang tussen betrouwbaarheidsintervallen en het toetsen van hypothesen.

### Opgave 2. (30 punten)

300 mensen met een bepaalde ziekte worden in drie groepen van telkens 100 patiënten ingedeeld. De eerste groep wordt met het nieuwe serum  $A$  behandeld, de tweede groep met het veel gebruikte serum  $B$  en de derde groep met een placebo (suikerwater). Na een bepaalde tijd zijn 70 van de met serum  $A$  behandelde patiënten genezen, 60 van de met serum  $B$  behandelde patiënten, maar ook 50 van de patiënten uit de placebo-groep zijn gezond.

- (i) Toets op een onbetrouwbaarheidslevel van  $\alpha = 0.05$  de nulhypothese dat er geen verschil tussen de effecten van de sera  $A$  en  $B$  is. Hoe zit het bij een onbetrouwbaarheid van  $\alpha = 0.10$ ?
- (ii) Kan je op een onbetrouwbaarheidslevel van  $\alpha = 0.05$  de nulhypothese verwerpen, dat het serum  $A$  evenveel effect heeft als de placebo? Wat kan je (met behulp van de uitgegeven tabellen) over de  $P$ -waarde van de toets tussen serum  $A$  en de placebo zeggen?
- (iii) Toets ook of het serum  $B$  evenveel effect heeft als de placebo. Hanteer deze keer een onbetrouwbaarheidslevel van  $\alpha = 0.10$ .
- (iv) Hoeveel mensen moeten na behandeling met een serum minstens gezond worden, om op een onbetrouwbaarheidslevel van  $\alpha = 0.05$  de nulhypothese te kunnen verwerpen dat het serum evenveel effect heeft als de placebo?

**Opgave 3.** (25 punten)

Een drietal omroepverenigingen stelt een gezamenlijke enquête op, om in een steekproef de opinie van hun medewerkers te peilen omtrent een nauwere samenwerking in het programma-beleid. De uitslagen van de enquêtes zijn in de onderstaande tabel aangegeven:

omroepvereniging	voor	tegen	geen mening
A	105	61	87
B	118	60	130
C	88	58	62

- (i) Bereken de theoretisch verwachte resultaten onder de nulhypothese dat de uitslagen van de enquêtes onafhankelijk van de omroepvereniging zijn.
- (ii) Toets met een  $\chi^2$ -toets op een onbetrouwbaarheidslevel van 5% de nulhypothese dat de uitslagen onafhankelijk van de omroepvereniging zijn.
- (iii) Wat kan je (met behulp van de uitgegeven tabellen) over de  $P$ -waarde van de toets op onafhankelijkheid zeggen?

**Opgave 4.** (20 punten)

Zij  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  een steekproef voor twee kenmerken van een populatie en zij  $l(x)$  de regressielijn door de punten van deze steekproef.

- (a) Geef met je eigen woorden aan, hoe de regressielijn  $l(x)$  gedefinieerd is.
- (b) Geef commentaar over de volgende uitspraken over de regressielijn (d.w.z. leg uit waarom ze kloppen of niet kloppen):
  - (i) De correlatiecoëfficiënt geeft de stijging van de regressielijn aan.
  - (ii) Aan de hand van de regressielijn laten zich uitschieters identificeren.
  - (iii) De regressielijn loopt door het zwaartepunt  $(\bar{x}, \bar{y})$  van de punten in de steekproef.
  - (iv) Als het betrouwbaarheidsinterval voor de stijging van de regressielijn de waarde 0 bevat, zijn de twee kenmerken met hoge kans onafhankelijk.
  - (v) Als de punten van de steekproef dicht bij de regressielijn liggen is of het tweede kenmerk (de  $y$ -waarden) en gevolg van het eerste (de  $x$ -waarden), of het eerste kenmerk is een gevolg van het tweede.

Herinnering:

- De variantie voor de schatter van de verwachtingswaarde is  $\frac{\sigma^2}{n}$ .
- De variantie voor de schatter van de relatieve frequentie is  $\frac{p(1-p)}{n}$ .

Kritieke waarden  $z_\alpha$  voor de standaard-normale verdeling:

$\alpha$	0.10	0.05	0.025	0.02	0.01	0.005
$z_\alpha$	1.2816	1.6449	1.9600	2.0537	2.3263	2.5758

**Succes ermee!**