Tentamen 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.*

Opgave 1. (10 punten)
Zij $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ een steekproef voor twee kenmerken van een populatie en zij $l(x)$ de regressielijn door de punten van deze steekproef.

(i) Geef met je eigen woorden aan, hoe de regressielijn $l(x)$ gedefinieerd is.

(ii) Geef commentaar over de volgende uitspraken over de regressielijn (d.w.z. leg uit waarom ze kloppen of niet kloppen):

(a) De correlatiecoëfficiënt geeft de stijging van de regressielijn aan.

(b) Aan de hand van de regressielijn laten zich uitschieters identificeren.

(c) De regressielijn loopt door het zwaartepunt $(\overline{x}, \overline{y})$ van de punten in de steekproef.

(d) 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.

Opgave 2. (25 punten)
Van 1000 aselect gekozen personen is nagegaan of ze kleurenblind zijn. Van de 480 mannen bleken dit er 38 te zijn, bij de vrouwen was het aantal 23.

(i) Toets op onbetrouwbaarheidslevel $\alpha = 10\%$ de nulhypothese dat kleurenblindheid onafhankelijk is van het geslacht.

(ii) We gaan verder van 38 kleurenblinde mannen in dezelfde groep uit. Is het minimale aantal vrouwen dat kleurenblind moet zijn, opdat de nulhypothese op level $\alpha = 10\%$ net niet verworpen kan worden groter of kleiner dan 30?

De relevante χ^2-waarde is $\chi^2_{1,0.90} = 2.71$.

Opgave 3. (20 punten)
Bij een online-database heeft een steekproef van 400 aanvragen een gemiddelde verwerkingstijd van $\bar{t} = 9$ seconden opgeleverd. Uit langdurige ervaring is bekend dat de standaardafwijking voor de verwerkingstijd $\sigma = 3$ seconden bedraagt.

(i) Bepaal een betrouwbaarheidsinterval op level 95% voor de gemiddelde verwerkingstijd.

(ii) Hoe groot moet de steekproef minstens zijn om op level 95% een betrouwbaarheidsinterval met een lengte van hoogstens 0.5 seconden te hebben?

z.o.z.

Opgave 4. (10 punten)
Om de kwaliteit van een levering te testen, wordt op twee (aselecte, onafhankelijke) steekproeven uit de levering een toets met onbetrouwbaarheid α toegepast. De nulhypothese van de toets luidt natuurlijk, dat de producten in orde zijn.

De levering wordt gereclameerd als bij (minstens) een van de toetsen de nulhypothese verworpen wordt. Hoe groot moet α gekozen worden, opdat slechts 5% van de leveringen ten onrechte gereclameerd worden?

Opgave 5. (35 punten)
De tabel hieronder bevat de leeftijden van steekproeven van drie soorten computerchips. We willen de nulhypothese toetsen dat alle drie soorten chips dezelfde gemiddelde leeftijd μ hebben. Hiervoor veronderstellen we dat de leeftijden van de chips normaal verdeeld zijn met de gemeenschappelijke variantie σ^2.

soort chip	leeftijd				
1	407	411	409		
2	404	406	408	405	402
3	410	408	406	408	

(i) Bepaal de gemiddelde leeftijden $\overline{x_i}$ van de drie soorten chips. Bereken ook het gemiddelde $\overline{x}$ *en gros* over alle drie soorten chips samen.

(ii) Bepaal de kwadratische afwijking $v_b = \sum_{i,j}(x_{ij} - \overline{x_i})^2$ *binnen* de steekproeven en de kwadratische afwijking $v_t = \sum_i n_i(\overline{x_i} - \overline{x})^2$ *tussen* de steekproeven.

(Controle: $v_b = 36$, $v_t = 36$.)

(iii) Bepaal een schatting voor de variantie σ^2 van de leeftijden op basis van de kwadratische afwijkingen *binnen* de steekproeven.

(iv) Bepaal onder de aanname van de nulhypothese een schatting voor de variantie σ^2 op basis van de kwadratische afwijkingen *tussen* de steekproeven.

(v) Bepaal de F-waarde van de steekproeven en toets de nulhypothese met een F-toets op de onbetrouwbaarheidslevels $\alpha = 5\%$ en $\alpha = 1\%$.

Wat zijn de aantallen van vrijheidsgraden van de benodigde F-verdeling?

De relevante waarden van de F-verdeling zijn $f_{0.05} = 4.26$ en $f_{0.01} = 8.02$.

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_α voor de standaard-normale verdeling:

α	0.10	0.05	0.025	0.02	0.01	0.005
z_α	1.2816	1.6449	1.9600	2.0537	2.3263	2.5758

Succes ermee!