

## 1 Vraag 1

Zij  $X_1, \dots, X_n$  i.i.d verdeeld volgens de Paretdistributie met dichtheidsfunctie

$$f(x) = \begin{cases} \theta x_0^\theta x^{-\theta-1} & x \geq x_0 \\ 0 & x < x_0 \end{cases}$$

Hierbij is  $\theta$  een onbekende parameter en  $x_0$  een gegeven getal.

1. Gegeven de statistiek  $T_n = \frac{1}{n} \sum_{i=1}^n \log \frac{X_i}{x_0}$ . bewijs dat dit een onvertekende schatter is van  $\frac{1}{\theta}$ .
2. Bereken de variantie van  $T_n$ .
3. Stel dat  $n$  groot genoeg zodat je kan aannemen dat  $T_n$  normaal verdeeld is (CLS). Geef dan een 95% betrouwbaarheidsinterval voor  $\theta$ .

## 2 Vraag 2

Zij  $S^2$  de steekproefvariantie van een normaal verdeelde stochastische veranderlijke  $X$  met gemiddelde  $\mu$  en standaard deviatie  $\sigma^2$ . Gegeven het volgende 95% betrouwbaarheidsinterval voor de variantie:

$$\left[ \frac{(n-1)S^2}{b}, \frac{(n-1)S^2}{a} \right]$$

met  $a$  en  $b$  positieve getallen en  $a < b$ .

1. Wat is het verband tussen  $a$  en  $b$ ? (wat vooral belangrijk is, is dat je één of andere relatie kan vinden zodat je deelvraag (b) kan oplossen)
2. Toon aan dat de volgende vergelijking geldt als je het betrouwbaarheidsinterval zo klein mogelijk wil maken:

$$g(a)a^2 = g(b)b^2$$

Hierbij is  $g$  de dichtheidsfunctie van  $\chi_{n-1}^2$ .

## 3 Vraag 3

Een vraag over de vergelijking van proporties. Een verhaaltje over iets van de sterkte van een bepaald object bij twee verschillende soorten objecten (A en B). Bij soort A waren er 22 van de 40 die boven een sterkte van 34 ksi zaten en bij B 29 van de 42.

1. Geef een 95% betrouwbaarheidsinterval voor het verschil tussen de twee proporties.
2. Doe een hypothesetest om te kijken of de proporties significant verschillend zijn op significantieniveau 5%. Bereken ook de p-waarde

## 4 Vraag 3

Gegeven de data op 6 locaties van de zinkconcentratie op de bodem van een beekje en aan het oppervlak van dat beekje (dus voor elke locatie waren er twee datapunten: één van de bodem en één van het oppervlak). Je moest een hypothesetest om te kijken of de concentratie op de bodem significant groter was dan die aan het oppervlak met significantieniveau 0.05. Je moest ook de

p-waarde berekenen. De data zag er ongeveer als volgt uit (niet helemaal, maar toch goed genoeg denk ik).

| Locatie     | 1    | 2     | 3    | 4    | 5    | 6    |
|-------------|------|-------|------|------|------|------|
| Bodem       | 0.44 | 0.43  | 0.56 | 0.37 | 0.34 | 0.55 |
| Oppervlakte | 0.41 | 0.402 | 0.44 | 0.2  | 0.24 | 0.45 |