

U krijgt de beschikking over een tabellenboekje. Het gebruik van een niet-programmeerbare rekenmachine is toegestaan. Succes!

1. Veronderstel dat de waarneming X de $\text{bin}(n, p)$ -verdeling bezit, waarin n bekend is en $0 \leq p \leq 1$ onbekend. We gaan p Bayesiaans schatten. Als klasse van a priori dichtheden op $[0, 1]$ nemen we de Bèta-dichtheid

$$\pi(y) = \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha, \beta)}$$

met bekende getallen α en β . Het getal $B(\alpha, \beta)$ (de Bèta-functie) is gekozen zodanig dat dit een kansdichtheid is (voor $\alpha, \beta > 0$).

- (a) Bepaal de a-posteriori dichtheid. Tot welke klasse bekende verdelingen behoort de a-posteriori verdeling? Met welke parameters?
- (b) Geef de Bayes-schatting voor p . Hierbij mag u gebruik maken van het feit dat de verwachting van een $\text{Bèta}(\alpha, \beta)$ verdeelde stochastische grootte gelijk is aan de breuk $\frac{\alpha}{\alpha+\beta}$.
2. Stel $X_1, \dots, X_n$ is een steekproef uit de verdeling met kansdichtheid

$$p_\theta(x) = 3\theta x^2 e^{-\theta x^3}, \quad \text{voor } x \geq 0$$

waar $\theta > 0$ een onbekende parameter is.

- (a) Laat zien dat de meest aannemelijke schatter voor θ gelijk is aan de breuk $\frac{n}{\sum_{i=1}^n X_i^3}$.
- (b) Bepaal de waargenomen informatie.
- (c) Geef een benaderend betrouwbaarheidsinterval voor θ op basis van de meest aannemelijke schatter met betrouwbaarheid $1 - \alpha$.

3. Stel dat $X_1, \dots, X_n$ onderling onafhankelijk en alle Poisson verdeeld zijn met onbekende parameter θ .
- Bepaal de momentenschatter T voor θ .
 - Is de schatter T^2 zuiver voor θ^2 ? Indien dit niet het geval is, bepaal dan een zuivere schatter voor θ^2 .
 - Geef een voldoende en volledige statistiek.
 - Bepaal een UMVZ-schatteer voor θ^2 .
4. In 1938 publiceerde Benford een wetenschappelijk artikel waarin hij beschrijft dat in veel datasets 30% van de getallen met een 1 beginnen (bijvoorbeeld de getallen: 12.3, 1231,8 en 1.1). Dit geldt ook voor financiële overzichten, bijvoorbeeld de boekhouding van bedrijven. Om te controleren of in de boekhouding van een groot bedrijf 30% van de getallen met een 1 begint, trekken we geheel willekeurig, met teruglegging, 1000 getallen uit de boekhouding. Definieer X als het aantal getallen in de steekproef dat met een 1 begint. Dan is X binomiaal verdeeld met parameters 1000 en p , waarbij p de kans is dat een willekeurig getal in de boekhouding met een 1 begint.
- Van de 1000 getallen in de steekproef beginnen 320 getallen met een 1. Toets de nulhypothese dat 30% van de getallen in de boekhouding met een 1 begint. Geef de standaard tweezijdige toets voor onbetrouwbaarheidsdrempel $\alpha = 0.05$. Welke veronderstellingen maak je? Geef in ieder geval: de nul- en de alternatieve hypothese in termen van p , de toetsingsgrootte en het kritieke gebied of de overschrijdingskans. Wat is je conclusie?
 - In zijn artikel schrijft Benford ook dat 18% van de getallen met een 2 begint. Om de nulhypothese dat in de boekhouding 30% van de getallen met een 1 en 18% van de getallen met een 2 begint te toetsen, gebruiken we de likelihood ratio toets. Definieer X_1, X_2 en X_3 als het aantal getallen in de steekproef dat begint met, respectievelijk een 1, een 2, en met iets anders. Nu heeft (X_1, X_2, X_3) een multinomiale verdeling met parameters $n = 1000$ en (p_1, p_2, p_3) . Dat betekent dat

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3) = \frac{1000!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

voor $x_1 + x_2 + x_3 = 1000$ en $p_1 + p_2 + p_3 = 1$. Bepaal de likelihood ratio statistiek, λ_n , voor het toetsen van de nulhypothese dat $p_1 = 0.30$ en $p_2 = 0.18$ tegen de alternatieve hypothese dat dit niet het geval is. (U mag gebruik maken van het feit dat de meest aannemelijke schatter voor p_1, p_2 en p_3 gelijk zijn aan $\hat{p}_1 = X_1/1000, \hat{p}_2 = X_2/1000$ en $\hat{p}_3 = X_3/1000$.)

Normering:

1(a):	5	2(a):	3	3(a):	3	4(a):	5
1(b):	2	2(b):	3	3(b):	3	4(b):	4
		2(c):	3	3(c):	4		
				3(d):	1		

Eindcijfer: (totaal + 4)/4