Examen Kansrekening en Wiskundige Statistiek

S. Vansteelandt

Academiejaar 2009-2010

Alle vragen hebben een korte oplossing die weinig rekenwerk vergt, en kunnen onafhankelijk van elkaar worden opgelost. Vragen met een (*) zijn bonusvragen van een hogere moeilijkheidsgraad, en kunt u het best pas starten eenmaal de andere vragen goed werden opgelost.

1. Veronderstel dat $(X_1, Y_1), ..., (X_n, Y_n)$ onderling onafhankelijke, identiek verdeelde stochastische vectoren zijn met $Y_i \in \{0, 1\}$ en

$$P(Y_i = 1|X_i = x) = \frac{1}{1 + \exp(\theta x)}$$

met θ ongekend. De marginale dichtheidsfunctie van X_i is niet gekend, maar onafhankelijk van θ. Bepaal een vergelijking waar de maximum waarschijnlijkheidsschatter voor θ het nulpunt van is.

2. Stel dat de toevalsveranderlijken $X_1, ..., X_n$ i.i.d. uniform verdeeld zijn over $[-\theta, \theta]$ voor een zekere ongekende parameter $\theta > 0$. Om θ te schatten, beschouwen we de volgende schatters $S_n = (X_{(n)} - X_{(1)})/2$ met $X_{(n)}$ en $X_{(1)}$ het maximum en minimum van de observaties, respectievelijk, en $T_n = \sqrt{(3/n) \sum_{i=1}^n X_i^2}$.

   (a) (*) Toon aan dat de reeks $n(X_{(1)} + \theta, \theta - X_{(n)})$ in distributie convergeert naar een toevalsvector $2\theta(Z_1, Z_2)$ voor een paar onafhankelijke exponentieel verdeelde toevalsveranderlijken Z_1 en Z_2. Hint: stel eerst de gezamenlijke verdeling

   $$P\left(n(X_{(1)} + \theta) > x, n(\theta - X_{(n)}) > y\right)$$

   op.

   (b) Bepaal op basis van voorgaande resultaat de limietverdeling van $n(S_n - \theta)$.

   (c) Leid de dichtheidsfunctie van voorgaande limietverdeling. Indien u de vorige vraag niet kon beantwoorden, ga er dan hier en in de verdere vragen vanuit dat de limietverdeling de verdeling is van de som van 2 onafhankelijke exponentieel verdeelde metingen met gemiddelde 1.

   (d) Onderzoek of S_n een consistente schatter voor θ is.

   (e) Bereken de bias (vertekening) van S_n als schatter voor θ.

   (f) Toon aan hoe een asymptotisch betrouwbaarheidsinterval voor θ kan geconstrueerd worden op basis van S_n.

   (g) Bepaal de limietverdeling van $\sqrt{n}(T_n - \theta)$.

   (h) Welk van beide schatters S_n of T_n valt te verkiezen?

   (i) Bepaal de asymptotische relatieve efficiëntie van beide schatters.

3. Stel dat de toevalsveranderlijken $X_1, ..., X_n, Y_1, ..., Y_n$ onderling onafhankelijk zijn met elke X_i Poisson verdeeld met gemiddelde μ en elke Y_i Poisson verdeeld met gemiddelde ν. Om de nulhypothese te toetsen dat $\mu = \nu$ beschouwen we de teststatistiek

$$T_n = 2n \left(\sqrt{\bar{X}_n} - \sqrt{\bar{Y}_n}\right)^2.$$

(a) Bepaal de limietverdeling van T_n.

(b) Bepaal op basis van voorgaande resultaat een kritische waarde (drempelwaarde) van de test zodat het asymptotische significantieniveau α bedraagt. Indien u vorige vraag niet kon oplossen, ga er dan vanuit dat de limietverdeling een F-verdeling is met telkens 1 vrijheidsgraad.