Examen Statistiek en Gegevensverwerking
2de Bachelor Fysica en Sterrenkunde

Eerste zittijd 2010 - 2011

Theorie

1. Bepaal de bezettingsgraad van een wachtrij van lengte N, waarin gemiddeld λ_I evenementen per tijdseenheid binnenkomen en λ_0 evenementen per tijdseenheid verwerkt worden. Wat is de fractie van de inkomende evenementen die verloren gaat? Werk de sommaties of integralen die voorkomen expliciet uit.
(3 pt)

2. Bepaal het algoritme om met de inverse transformatiemethode toevalsgetallen te genereren volgens een Gauss-verdeling met gemiddelde μ en variantie σ^2? (Het is voldoende om 1 formule af te leiden)
(2 pt)

3. Bewijs dat de variantie van een (onbevooroordeelde) schatter niet willekeurig klein kan zijn door een ondergrens af te leiden voor de grootte van de variantie.
(3 pt)

4. Je hebt 2 types evenementen a en b die elk volgens een Poissonverdeling optreden. Bewijs dat de som van de 2 opnieuw een Poisson is. Hoe rijm je dit met het CLT dat stelt dat de som van een groot aantal evenementen, verdeeld volgens om het even welke distributie, steeds een Gauss-verdeling volgt?
(2 pt)

Oefeningen

1. Uit een steekproef met 944 mannen en 924 vrouwen blijkt dat het IQ van mannen en vrouwen licht verschillend verdeeld is. Beide distributies zijn in goede benadering Gaussisch met volgende parameters:

$$\mu_m = 103.08 \qquad \sigma_m = 14.54$$
$$\mu_f = 101.41 \qquad \sigma_f = 13.55$$

(a) Met welke confidentie kan je stellen dat het gemiddelde IQ van mannen hoger is dan dat van vrouwen?

(b) Is het verschil in de breedte van de distributies significant?

(c) Mensa is een vereniging van mensen die een IQ in de 98ste percentiel hebben. In de praktijk aanvaarden ze mensen met een IQ boven de 130 (gebaseerd op een gemiddeld IQ van 100, en een σ van 15). In de Mensa-afdeling van Toronto zijn er 150 mannen en 83 vrouwen. Is dit consistent met de gegevens die je hierboven kreeg? Werk uit!

(3 pt)

2. De waarschijnlijkheid om het virus influenza te hebben is 20%. Hoe groot is dan de kans dat in een zaal van 200 mensen 30 tot 50 mensen het virus hebben?

(2 pt)

3. In een test voor EPS (extra perceptory senses) wordt een voorwerp in een doos gestoken. Vervolgens wordt aan 100 vrijwilligers gevraagd om de doos met het voorwerp uit een reeks dozen, genummerd van 1-5, te halen. Het voorwerp zit in doos 3 en 31 vrijwilligers kiezen voor doos 3.

(a) Bereken de significantie van dit bewijs voor EPS.

(b) Er wordt een controletest gedaan waarbij er geen voorwerpen in de dozen zitten. Hierbij kiezen opnieuw 25 van de 100 vrijwilligers voor doos 3. Wat betekent dit voor de significantie van de test.

(3pt)

4. Een variant van een Gaussische verdeling

$$p(x) = \sqrt{\frac{2}{\pi}}\, e^{-x^2/2}$$

kan niet gegenereerd worden met behulp van de inverse transformatie (waarom niet?). We zullen hier dus de methode van 'weighted rejection' toepassen met enveloppe-functie

$$f(x) = k e^{-x}$$

(Opgelet: deze is niet genormeerd vanwege de enveloppeconstante k)

(a) Stel een ~~der~~ formule op om een getal volgens de verdeling te genereren.

(b) Voor welke k is deze methode het efficiëntst? Wat is dan het percentage verworpen getallen?

(2pt)

Opm: Dit is naar mijn geheugen, de middag na het examen, er kunnen dus, vooral tijdens de opgaven, wat andere bewoordingen gebruikt geweest zijn op het originele examen.