1. Obligatoriske Opgave

Anvendt Statistik - forårsemesteret 2001.

Følgende opgave er den første af to, som er obligatoriske for at kunne gå til eksamen i kurset *Anvendt Statistik*. Den vil blive udlevet onsdag d. 14. marts 2001 efter dagens forelæsning, og en skriftelig besvarelse skal være undertegnede i hænde senest onsdag d. 21. marts 2001. Det er tilladt at arbejde sammen om opgaven, dog skal der afleveres separate løsninger. Til eksamen vil samarbejde dog ikke være tilladt.

Der ikke blot tillades men anbefales at bruge computere og de programmer I har arbejdet med, og i den del af opgaven om Monte Carlo vil brugen af computere være nødvendig.

God arbejdslyst, Troels

Fordelinger:

- 1. Lille Per slår 10 gange med en almindelig terning og får 6 seksere. Hvad er sandsynligheden for en sådan hændelse eller noget mere ekstremt? Har lille Per snydt?
- 2. Lille Per går nu på kasino og spiller på ét nummer ad gangen ($p = 1/37$). Hvis han ikke snyder, hvad er så sandsynligheden for at han ud af 100 spil vil vinde mere end 3 gange?
- 3. Udregn middel og varians af følgende fordelinger:
  - $f(x) = \frac{1}{2}\sin(x), \quad x \in [0, \pi]$.
  - $f(x) = \frac{3}{4}(-x^3 + 2x^2), \quad x \in [0, 2]$.
  - $f(x) = \frac{1}{4}e^{-|x|} + \frac{1}{2}\frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{1}{2}(\frac{x-\mu}{\sqrt{2}\sigma})^2), \quad x \in [-\infty, \infty]$.

Fejlberegning:

- 1. Lad $x = y + 2z$. Givet at $\sigma_y = 0.2$, $\sigma_z = 0.3$ og $V_{yz} = 0.04$, hvad er fejlen på x?
- 2. Hvis i stedet $x = y^2/z$, hvad er så fejlen på x?
- 3. Begyndelsesaktiviteten, N_0, og levetiden, τ, af en radioaktiv kilde kendes med en nøjagtighed på 1 %. Ved estimering af aktiviteten $N = N_0\, e^{-t/\tau}$, vil usikkerheden i starten domineres af usikkerheden på N_0 og senere af usikkerheden på τ. For hvilken værdi af t/τ vil de to usikkerheder bidrage ligeligt til usikkerheden i N?
- 4. En studerende måler lydens hastighed tyve gange og konkluderer at spredningen på resultaterne er 12 m/s. Hvis fejlkilderne er mange og tilfældige, hvilken fordeling skulle målingerne så følge? Hvor mange målinger kræves der, for at usikkerheden når ned på 2 m/s?

Monte Carlo: Til denne del skal der bruges computere. Plots kan vedlægges opgavebesvarelsen.

- 1. Generer ud fra en flad fordeling i intervallet $[0,1]$ den normerede fordeling proportional med $-x^2+4x-3$ i intervallet $[1,3]$.
  - Bestem effektiviteten af din rutine.
- 2. Generer ud fra en flad fordeling i intervallet $[0,1]$ (Cauchy-) fordelingen $\frac{1}{\pi(1+x^2)}$ i intervallet $[-\infty,\infty]$.
  - Bestem middelværdi og spredningen af den genererede fordeling.
  - Konvergerer middelværdien mod nul?
  - Er spredningen unaturlig stor i forhold til, hvad den et gaussisk fit har?
- 3. Generer ud fra en flad fordeling i intervallet [0,1] den normerede fordeling proportional med $e^{-x}(1+\sin(x))$ i intervallet $[0,\infty]$.
  - Fit fordelingen med stigende antal bins. Konvergerer χ^2/N mod 1?

Statistiske tests:

- 1. For at kunne bruge χ^2-metoden, hvilket krav bør fejlene opfylde? Hvad er sandsynligheden for at opnå $\chi^2 = 1.35$ med 50 frihedsgrader?
- 2. I et eksperiment forventes baggrunden at være fordelt efter $f_{\text{baggrund}} = 0.25x + 2.5$. I et histogram med ti bins i intervallet $[0,10]$ fås 3. og 4. bin til at indeholde værdierne 4.79 og 5.83 i stedet for de forventede baggrundsværdier, henholdsvis 3.25 og 3.50.
  - Hvad er signifikansen af et sådant “signal”?
  - Er det rimeligt direkte at vælge de interessante bins, eller burde man benytte mindre subjektive metoder? Kan du kommer på nogle sådanne?
- 3. Nedenfor er angivet tyve målinger af samme størrelse, hvor usikkerheden på hver måling er bestemt til at være omkring 1.0.

9.43	10.04	15.19	10.27	10.26	8.98	8.88	7.71	9.74	10.34
9.65	11.90	9.11	10.99	8.94	11.42	9.81	10.58	9.09	10.33

  - Hvad er χ^2, og er det konsistent givet antallet af frihedsgrader?
  - Kunne én af målingerne mistænkes for at være forkert? Med hvilken sandsynlighed?
  - Hvis den mistænkte måling forkastes, er de resterende målinger da konsistente?

Bonus opgave:

- Hvor mange fysikstuderende findes der i Danmark?
  - Uden anden information, hvilken fejl ville du tilskrive dette tal?
  - Hvis en journalist påstod at antallet af fysikstuderende var faldet/steget med 3 %, hvad burde man så fortælle ham?

Those who are good at archery learnt from the bow and not from Yi the Archer. Those who know how to manage boats learnt from boats and not from Wo (the legendary mighty boatman). Those who can think learnt for themselves and not from the Sages.

[Kuan Yin Tze, 8th Century]