Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

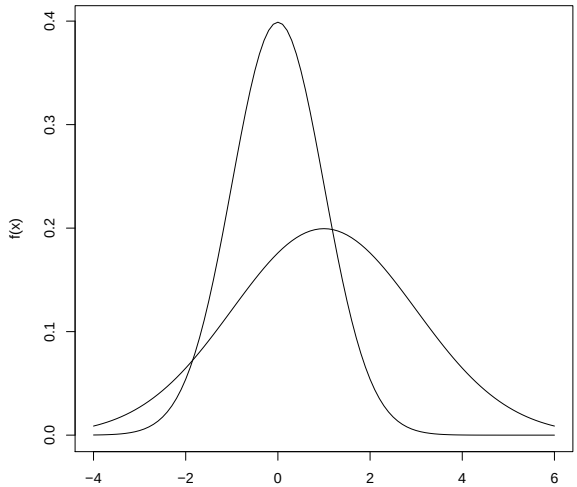
Brukerkurs i statistikk, ST0103

10. august, 2013

Løsningsforslag

Oppgave 1

a)



b) De to første sannsynlighetene blir

$$P(X \leq 1.2) = P(Z \leq \frac{1.2-0}{1}) = G(1.2) = 0.8849, \tag{1}$$

og

$$P(Y > 2) = 1 - P(Z \leq \frac{2-1}{2}) = 1 - G(0.5) = 0.3085. \tag{2}$$

Videre blir $E(X+Y) = 0+1 = 1$ og $\mathrm{Var}(X+Y) = \mathrm{Var}\,X + \mathrm{Var}\,Y = 1 + 2^2 = 5$ slik at

$$P(X+Y < 2) = P(Z \leq \frac{2-1}{\sqrt{5}}) = G(0.4472) = 0.673. \tag{3}$$

c) Kumulativ fordeling til $Z = e^X$ blir

$$F_Z = P(Z < z) = P(e^X < z) = P(X < \ln z) = F_X(\ln z), \tag{4}$$

slik at sannsynlighetstettheten blir

$$\begin{aligned} f_Z(z) &= \frac{d}{dz} F_X(\ln z) \\ &= f_X(\ln z) \frac{d}{dz} \ln z \\ &= \frac{1}{2\pi} e^{-\frac{1}{2}(\ln z)^2} \frac{1}{z} \\ &= \frac{1}{2\pi z} e^{-\frac{1}{2}(\ln z)^2}. \end{aligned} \tag{5}$$

Oppgave 2

a) To hendelser er disjunkte hvis de ikke kan inntreffe samtidig. Det er ikke tilstrekkelig at $P(A \cap B) = 0$ siden mulige hendelser kan ha 0 sannsynlighet i kontinuerlige modeller.

I oppgaven er det gitt at $P(A) = 0.5$, $P(B) = 0.2$ og $P(A \cup B) = 0.6$. I følge den generelle addisjonssetningen er

$$P(A \cup B) = P(A) + P(B) - P(A \cap B), \tag{6}$$

slik at

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.5 + 0.7 - 0.6 = 0.1. \tag{7}$$

Dermed er A og B ikke disjunkte.

b) To hendelser er uavhengige hvis $P(A \cap B) = P(A)P(B)$. For hendelsene i oppgaven er $P(A)P(B) = 0.5 \cdot 0.2 = 0.1 = P(A \cap B)$, altså er A og B uavhengige hendelser.

Oppgave 3

a) Det er rimelig å anta at X er binomisk fordelt. Dette forutsetter at hver spurv i utvalget er parasittert med samme sannsynlighet p og uavhengighet mellom delforsøkene (hvorvidt en spurv er parasittert skal ikke påvirke sannsynlighet for at en annen spurv er parasittert). Det siste er oppfyllt fordi vi trekker fra en populasjon som er mye større enn utvalgsstørrelsen.

b) Punktestimatet av p blir $\hat{p} = 45/50 = 0.9$ og tilnærmet 95% konfidensintervall for p blir

$$(\hat{p} - z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}, \hat{p} + z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}) = (0.817, 0.983) \tag{8}$$

c) Sannsynligheten for at hver spurv er parasittert kan uttrykkes som

$$p = P(Y \geq 1) = 1 - P(Y = 0) = 1 - e^{-\mu}\mu^0/0! = 1 - e^{-\mu}, \tag{9}$$

av parameteren μ i Poissonfordelingen.

d) Løser vi ligningen over m.h.p. μ får vi at

$$\mu = -\ln(1-p). \tag{10}$$

Et punktestimat (SME av μ) er da samme funksjon av $\hat{p}$,

$$\hat{\mu} = -\ln(1-\hat{p}) = -\ln 0.1 = 2.30. \tag{11}$$

Regel for transformasjon av konfidensintervall gir videre at

$$(-\ln(1-0.817), -\ln(1-0.983)) = (1.697, 4.084) \tag{12}$$

er et tilnærmet 95%-konfidensintervall for μ.

e) Fra egenskapene til Poissonfordelte variable vet vi at forventningen til antall parasitter observert på hver spurv $E(Y_i) = \mu$ slik at $\hat{\mu} = \sum Y_i/n = 121/50 = 2.42$ en et forventningsrett estimat av μ. Videre blir $\operatorname{Var}\hat{\mu} = (1/n^2)\sum \operatorname{Var} Y_i = \mu/n$ og normaltilnærming gir så at

$$P(-z_{\alpha/2} < \frac{\hat{\mu} - \mu}{\sqrt{\hat{\mu}/n}} < z_{\alpha/2}) = 1 - \alpha. \tag{13}$$

slik at

$$(\hat{\mu} - z_{\alpha/2}\sqrt{\mu/n}, \hat{\mu} + z_{\alpha/2}\sqrt{\mu/n}) = (2.10, 2.73). \tag{14}$$

Dette siste estimatet og konfidensintervallet er å foretrekke siden konfidensintervallet er kortere. Dette skyldes at det basert på data som inneholder mer informasjon (antall parasitter og ikke bare hvorvidt antall parasitter tilstedet på hvert individ er større eller lik 1).