TMA4245 Statistikk
Vår 2015

Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

Øving nummer Ekstraøving, blokk I
Løsningsskisse

Oppgave 1

X er kontinuerlig fordelt med sannsynlighetstetthet

$$f_X(x) = \begin{cases} 2x\exp(-x^2) & x \geq 0 \\ 0 & \text{ellers} \end{cases}$$

Vi bruker i oppgavene teorem 7.3.

a) Vi har

$$U = X - 2 = g(X); x \geq 0$$

slik at

$$X = U + 2 = h(U)$$

med

$$h'(u) = 1.$$

Funksjonen $g(x) = x - 2$ er strengt monoton og deriverbar for alle x. Vi har dermed

$$\begin{aligned} f_U(\,u) &= f_X(\,h(u))\,|h'(u)| \\ &= 2\,(u+2)\,\exp(-\,(u+2)^2)\cdot 1 \\ &= 2\,(u+2)\,\exp(-\,(u+2)^2); u \geq -2. \end{aligned}$$

b) Vi har her

$$V = -2X = g(X); x \geq 0$$

der

$$X = -\tfrac{1}{2}V = h(V)$$

med

$$h'(v) = -\tfrac{1}{2}.$$

Funksjonen $g(x) = -2x$ er strengt monoton og deriverbar for alle x. Dette gir

$$
\begin{aligned}
f_V(\,v) &= f_X(\,h(v))\,|h'(\,v)| \\
&= 2\,(-\tfrac{1}{2}\,v)\ \exp(-\,(-\tfrac{1}{2}\,v)^2)\cdot\tfrac{1}{2} \\
&= -\tfrac{1}{2}\,v\ \exp(-\,(\tfrac{1}{2}\,v)^2); v \leq 0.
\end{aligned}
$$

c) Vi har

$$W = X^2 = g(X); x \geq 0$$

som gir

$$X = \sqrt{W} = h(W)$$

med

$$h'(w) = \frac{1}{2\sqrt{w}}.$$

Funksjonen $g(x) = x^2$ er strengt monoton og deriverbar for alle $x \geq 0$.

$$
\begin{aligned}
f_W(\,w) &= f_X(\,h(\,w))\,|h'(\,w)| \\
&= 2\,\sqrt{w}\ \exp(\,-w)\tfrac{1}{2\sqrt{w}} \\
&= \exp(-w); w \geq 0.
\end{aligned}
$$

Oppgave 2

$$X \sim N(\mu, \sigma^2), \quad Y = \frac{X-\mu}{\sigma}$$

Den kumulative sannsynlighetsfordelingen til Y er

$$F_Y(y) = P(Y \leq y) = P\left(\frac{X-\mu}{\sigma} \leq y\right) = F(X \leq \mu + \sigma y) = F_X(\mu + \sigma y).$$

Derivasjon med hensyn på y gir sannsynlighetstettheten $f_Y(y)$,

$$
\begin{aligned}
f_Y(y) &= \frac{dF_Y(y)}{dy} = \frac{d}{dy}F_X(\mu + \sigma y) \\
&= F'_X(\mu + \sigma y)\cdot\frac{d}{dy}(\mu + \sigma y) \\
&= f_X(\mu + \sigma y)\cdot\sigma \\
&= \frac{\sigma}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}((\mu+\sigma y)-\mu)^2} \\
&= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}.
\end{aligned}
$$

Siden dette er tettheten til normalfordelingen med $\mu = 0$ og $\sigma = 1$, har vi at $Y \sim N(0,1)$.

Oppgave 3

a) Vi benytter den kumulative fordelingsfunksjonen i oppgaveteksten. Regner først ut sannsynligheten for generell verdi av β, for så å regne ut for $\beta = \pi/8$. Dette gir

$$P(Y > \pi/4) = 1 - P(Y \leq \pi/4) = 1 - \frac{1 - \exp\left\{-\frac{\pi}{4\beta}\right\}}{1 - \exp\left\{-\frac{\pi}{2\beta}\right\}} = \frac{\exp\left\{-\frac{\pi}{4\beta}\right\} - \exp\left\{-\frac{\pi}{2\beta}\right\}}{1 - \exp\left\{-\frac{\pi}{2\beta}\right\}} = \underline{\underline{0.1192}}$$

$$\begin{aligned} P(\pi/4 < Y < \pi/3) = P(Y < \pi/3) - P(Y < \pi/4) &= \frac{1 - \exp\left\{-\frac{\pi}{3\beta}\right\}}{1 - \exp\left\{-\frac{\pi}{2\beta}\right\}} - \frac{1 - \exp\left\{-\frac{\pi}{4\beta}\right\}}{1 - \exp\left\{-\frac{\pi}{2\beta}\right\}} \\ &= \frac{\exp\left\{-\frac{\pi}{4\beta}\right\} - \exp\left\{-\frac{\pi}{3\beta}\right\}}{1 - \exp\left\{-\frac{\pi}{2\beta}\right\}} = \underline{\underline{0.0671}} \end{aligned}$$

$$\begin{aligned} P(Y > \pi/4 | Y < \pi/3) = \frac{P(Y > \pi/4 \cap Y < \pi/3)}{P(Y < \pi/3)} &= \frac{0.0671}{\left(1 - \exp\left\{-\frac{\pi}{3\beta}\right\}\right) / \left(1 - \exp\left\{-\frac{\pi}{2\beta}\right\}\right)} \\ &= \underline{\underline{0.0708}}. \end{aligned}$$

b) Siden Y er en kontinuerlig, kan vi finne sannsynlighetstettheten ved å derivere den kumulative fordelingsfunksjonen i oppgaveteksten

$$\begin{aligned} f(y;\beta) = \frac{\mathrm{d}}{\mathrm{d}y} F(y;\beta) = \frac{1}{1 - \exp\left\{-\frac{\pi}{2\beta}\right\}} \left(0 - \left(-\frac{1}{\beta}\right) \exp\left\{-\frac{y}{\beta}\right\}\right) &= \\ &= \underline{\underline{\frac{1}{\beta - \beta \exp\left\{-\frac{\pi}{2\beta}\right\}} \exp\{-y/\beta\}}} \end{aligned}$$

Fra figuren i oppgaveteksten har vi at $\tan(Y) = X$, altså har vi en-til-en relasjon mellom vinkelen Y og avstanden X. Det betyr at vi kan benytte transformasjon av variable (kap 7.2 i læreboka) til å finne fordelingen til X. La $y = \arctan(x) = w(x)$, altså den omvendte funksjonen av funksjonen over. Vi har da at sannsynlighetsfordelingen til X, $g(x;\beta)$, er gitt ved

$$g(x;\beta) = f(w(x);\beta) \cdot |w'(x)|.$$

Opplysningen i oppgaven eller oppslag i Rottmann gir at $w'(x) = 1/(1 + x^2)$ som gir

$$\underline{\underline{g(x;\beta) = \frac{1}{\beta - \beta \exp\left\{-\frac{\pi}{2\beta}\right\}} \exp\{-\arctan(x)/\beta\} \cdot \frac{1}{1 + x^2}, \quad x > 0.}}$$

Oppgave 4

a) La X_T og X_F være antall scoringer til henholdsvis Tyskland og Frankrike i første del av straffesparkkonkurransen. Utfra opplysningene i oppgaveteksten kan vi da anta at X_T og X_F er binomisk fordelte,

$$X_T \sim \text{Bin}(5, p_T), \quad X_F \sim \text{Bin}(5, p_F).$$

Sannsynligheten for at stillingen blir 5-5 etter første del er

$$\begin{aligned} P(\text{5-5}) &= P(X_T = 5, X_F = 5) \\ &= P(X_T = 5)P(X_F = 5) \\ &= \binom{5}{5} p_T^5 (1-p_T)^0 \cdot \binom{5}{5} p_F^5 (1-p_F)^0 \\ &= p_T^5 p_F^5 \\ &= 0.8^5 \cdot 0.7^5 \\ &= 0.0551. \end{aligned}$$

Sannsynligheten for at stillingen blir 3-3 etter første del er

$$\begin{aligned} P(\text{3-3}) &= P(X_T = 3, X_F = 3) \\ &= P(X_T = 3)P(X_F = 3) \\ &= \binom{5}{3} p_T^3 (1-p_T)^2 \cdot \binom{5}{3} p_F^3 (1-p_F)^2 \\ &= \binom{5}{3}^2 (p_T p_F)^3 [(1-p_T)(1-p_F)]^2 \\ &= 10^2 (0.8 \cdot 0.7)^3 (0.2 \cdot 0.3)^2 \\ &= 0.0632. \end{aligned}$$

Sannsynligheten for at lagene står likt etter første del, er lik sannsynligheten for at $X_T = X_F$,

$$p_{\text{D1}-\text{Lik}} = P(X_T = X_F) = \sum_{x=0}^{5} P(X_T = x)P(X_F = x) = 0.2728 \approx 0.27.$$

b) Skal vi ha en vinner etter første runde i del 2, må vi enten ha at Tyskland scorer og Frankrike ikke scorer, eller vice versa. Sannsynligheten for dette blir

$$\begin{aligned} p_{\text{V1}|\text{D2}} &= p_T(1-p_F) + (1-p_T)p_F \\ &= 0.8(1-0.7) + (1-0.8)0.7 = 0.38. \end{aligned}$$

Den betingede sannsynligheten for at det er Tyskland som vinner, gitt at resultatet er klart etter første runde i del 2, finnes ved å ta sannsynligheten for at Tyskland vinner første runde delt på sannsynligheten for at vi har en vinner etter første runde,

$$p_{\text{T}|\text{V1,D2}} = \frac{p_T(1-p_F)}{p_{\text{V1}|\text{D2}}} = \frac{0.8(1-0.7)}{0.38} = 0.6316.$$

Sannsynligheten for at Tyskland vinner del 2 av konkurransen er lik sannsynligheten for at Tyskland vinner runde 1, pluss sannsynligheten for at vi går videre til runde 2, og at Tyskland vinner runde 2, pluss sannsynligheten for at vi går videre til runde 3, og at Tyskland vinner runde 3, og så videre. Sannsynligheten for at det er Tyskland som vinner en runde, gitt at ett av lagene vinner, er $p_{\text{T}|\text{V}} = p_{\text{T}|\text{V1,D2}} = 0.6316$. Sannsynligheten for at ett av lagene vinner en gitt runde er $p_V = p_{\mathrm{V1|D2}} = 0.38$. Sannsynligheten for at Tyskland vinner konkurransen, gitt at vi er kommet til del 2, er dermed

$$\begin{aligned}
p_{\mathrm{T|D2}} &= p_{\mathrm{V1|D2}} p_{\mathrm{T|V1,D2}} \\
&\quad + (1 - p_{\mathrm{V1|D2}}) p_{\mathrm{V2|D2}} p_{\mathrm{T|V2,D2}} \\
&\quad + (1 - p_{\mathrm{V1|D2}})(1 - p_{\mathrm{V2|D2}}) p_{\mathrm{V3|D2}} p_{\mathrm{T|V3,D2}} \\
&\quad + \ldots \\
&= \sum_{r=1}^{\infty} (1 - p_V)^{r-1} p_V p_{\mathrm{T|V}} \\
&= p_V p_{\mathrm{T|V}} \sum_{r=1}^{\infty} (1 - p_V)^{r-1} \\
&= p_V p_{\mathrm{T|V}} \sum_{k=0}^{\infty} (1 - p_V)^{k} \\
&= p_V p_{\mathrm{T|V}} \frac{1}{1 - (1 - p_V)} \\
&= \frac{p_V p_{\mathrm{T|V}}}{p_V} = p_{\mathrm{T|V}} = 0.6316.
\end{aligned}$$

c) La fortsatt p_V være sannsynligheten for at kampen avgjøres i inneværende runde. Og la X være antall runder som spilles i del 2, inntil kampen er avgjort. Da er X geometrisk fordelt med parameter p_V.

Utfra egenskapene til den geometriske fordelingen vet vi at X har forventingsverdi

$$\mathrm{E}(X) = \frac{1}{p_V} = \frac{1}{0.38} = 2.6316$$

og varians

$$\mathrm{Var}(X) = \frac{1 - p_V}{p_V^2} = \frac{1 - 0.38}{0.38^2} = 4.2936 = 2.0721^2.$$

d) La Y være totalt antall runder i hele straffesparkkonkuransen (del 1 og 2). Hvis det kåres en vinner i del 1, blir $Y = 5$. Hvis lagene står likt etter de fem rundene i del 1, vil $Y = 5 + X$, hvor X er antall runder som spilles i del 2 (som i forrige punkt).

Siden $P(X = x) = (1 - p_V)^{x-1} p_V$, vil punktsannsynligheten til $Y = 5 + X$ være

$$P(Y = y) = P(5 + X = y) = P(X = y - 5) = (1 - p_V)^{y-5-1} p_V = (1 - p_V)^{y-6} p_V$$

når det spilles runder i del 2, altså når $Y > 5$.

Generelt, for $y \geq 5$, er punktsannsynligheten til Y gitt ved

$$P(Y = y) = \begin{cases} 1 - p_{\mathrm{D1-Lik}}, & y = 5 \\ p_{\mathrm{D1-Lik}} (1 - p_V)^{y-6} p_V, & y = 6, 7, \ldots \end{cases}.$$

Det forventede antall runder finnes ved å dele opp utfallsrommet for Y i de to hendelsene

$\{Y \leq 5\}$ og $\{Y > 5\}$, for så å betinge på hver hendelse for seg,

$$\begin{aligned}
\mathrm{E}(Y) &= \mathrm{E}(Y|Y \leq 5)P(Y \leq 5) + \mathrm{E}(Y|Y > 5)P(Y > 5) \\
&= 5(1 - p_{\mathrm{D1-Lik}}) + (5 + \mathrm{E}(X))p_{\mathrm{D1-Lik}} \\
&= 5 \cdot (1 - 0.27) + \left(5 + \frac{1}{0.38}\right) \cdot 0.27 = 5.7105.
\end{aligned}$$

Variansen til Y finnes også ved å se på tilfellene $Y = 5$ og $Y > 5$ hver for seg. La μ_Y være forventingsverdien til Y. Da gir definisjonen av varians for diskrete tilfeldige variable at

$$\begin{aligned}
\mathrm{Var}(Y) &= \sum_{y=5}^{\infty} (y - \mathrm{E}(Y))^2 \, P(Y = y) \\
&= (5 - \mu_Y)^2(1 - p_{\mathrm{D1-Lik}}) + p_{\mathrm{D1-Lik}} \sum_{y=6}^{\infty} (y - \mu_Y)^2 (1 - p_V)^{y-6} p_V \\
&= (5 - \mu_Y)^2(1 - p_{\mathrm{D1-Lik}}) + p_{\mathrm{D1-Lik}} \sum_{y=0}^{\infty} (y + \underbrace{6 - \mu_Y}_{a})^2 (\underbrace{1 - p_V}_{q})^y p_V \\
&= (5 - \mu_Y)^2(1 - p_{\mathrm{D1-Lik}}) + p_{\mathrm{D1-Lik}} p_V \sum_{y=0}^{\infty} (y + a)^2 q^y \\
&= (5 - \mu_Y)^2(1 - p_{\mathrm{D1-Lik}}) + p_{\mathrm{D1-Lik}} p_V \frac{a^2(-q^2) + 2a^2q - a^2 + 2aq^2 - 2aq - q^2 - q}{(q-1)^3}
\end{aligned}$$

Overgangen fra nest siste til siste linje er gyldig når $|q| = |1 - p_V| < 1$, som er tilfredsstilt her, siden $0 < p_V < 1$. Ved å sette inn tallverdier for p_V, $p_{\mathrm{D1-Lik}}$ og μ_Y i dette uttrykket, finner vi at $\mathrm{Var}(Y) = 2.5242 = 1.5888^2$.

Alternativt kan vi notere oss at andremomentet til $X \sim \mathrm{Geom}(p_V)$ er

$$\begin{aligned}
\mathrm{E}(X^2) &= \mathrm{Var}(X) + \mathrm{E}(X)^2 \\
&= \frac{1 - p_V}{p_V^2} + \left(\frac{1}{p_V}\right)^2 \\
&= \frac{2 - p_V}{p_V^2}.
\end{aligned}$$

Dette kan vi bruke til å finne et uttrykk for $\mathrm{E}(Y^2)$,

$$\begin{aligned}
\mathrm{E}(Y^2) &= \mathrm{E}(Y^2|Y \leq 5)P(Y \leq 5) + \mathrm{E}(Y^2|Y > 5)P(Y > 5) \\
&= 5^2 \cdot (1 - p_{\mathrm{D1-Lik}}) + \mathrm{E}\left((5 + X)^2\right) \cdot p_{\mathrm{D1-Lik}} \\
&= 5^2 \cdot (1 - p_{\mathrm{D1-Lik}}) + \left(5^2 + 2 \cdot 5 \cdot E(X) + E(X^2)\right) \cdot p_{\mathrm{D1-Lik}} \\
&= 5^2 \cdot (1 - p_{\mathrm{D1-Lik}}) + \left(5^2 + \frac{10}{p_V} + \frac{2 - p_V}{p_V}\right) \cdot p_{\mathrm{D1-Lik}} \\
&= 25 \cdot (1 - 0.27) + \left(25 + \frac{10}{0.38} + \frac{2 - 0.38}{0.38^2}\right) \cdot 0.27 \\
&= 35.1343.
\end{aligned}$$

Variansen til Y finner vi så ved å ta

$$\begin{aligned}\mathrm{Var}(Y) &= \mathrm{E}(Y^2) - \mathrm{E}(Y)^2 \\ &= \mathrm{E}(Y^2) - \mu_Y^2 \\ &= 35.1343 - 5.7105^2 \\ &= 2.5242.\end{aligned}$$

Oppgave 5

$$\begin{aligned}f(t|\lambda) &= \lambda e^{-\lambda t} \quad , t \geq 0 \\ f(\lambda) &= \theta e^{-\lambda\theta} \quad , \lambda \geq 0 \\ &\quad \text{Simultanfordelingen er da gitt ved} \\ f(t,\lambda) &= \lambda\theta e^{-\lambda(t+\theta)} \\ &\quad \text{Marginalfordelingen kan da finnes ved å integrere vekk } \lambda \\ f(t) &= \int_0^\infty f(t,\lambda)d\lambda = \int_0^\infty \lambda\theta e^{-\lambda(t+\theta)}d\lambda \\ &= \theta\left(\left[\lambda\left(-\frac{1}{t+\theta}\right)e^{-\lambda(t+\theta)}\right]_0^\infty + \int_0^\infty \frac{1}{t+\theta}e^{-\lambda(t+\theta)}d\lambda\right) \\ &= \theta\left(0 - 0 + \left[\frac{1}{(t+\theta)^2}e^{-\lambda(t+\theta)}\right]_0^\infty\right) \\ &= \theta\left(0 + \frac{1}{(t+\theta)^2}\right) = \frac{\theta}{(t+\theta)^2}\end{aligned}$$

Det vil altså si at fordelingen er gitt ved

$$\underline{\underline{f(t) = \frac{\theta}{(t+\theta)^2} \qquad ; t \geq 0}}$$