Institutionen för matematik och statistik
Introduktion till sannolikhetskalkyl
Lösningsförslag till räkneövning 5
16.2.2015

1. (2:35) Väntetiden X för en bankkund är en stokastisk variabel som är $Exp(1/10)$-fördelad med minuter som enhet. (i) Vilken är sannolikheten att kunden måste vänta över 15 minuter i banken? (ii) Vilken är sannolikheten att kunden måste vänta över 15 minuter, om kunden redan har väntat 10 minuter?

   Lösning: Om $X \sim Exp(\lambda)$ så är frekvensfunktionen för X

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \lambda e^{-\lambda x}, & x > 0 \end{cases}$$

   och fördelningsfunktionen

$$F(x) = P(X \leq x) = \begin{cases} 0, & x \leq 0 \\ \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}, & x > 0. \end{cases}$$

   (i) Eftersom X är väntetiden så söker vi sannolikheten $P(X > 15)$. Vi vet att X är $Exp(1/10)$-fördelad. så genom att beräkna komplementet får vi svaret:

$$P(X > 15) = 1 - P(X \leq 15) = 1 - F(15) = 1 - 1 + e^{-\frac{15}{10}} = e^{-\frac{3}{2}} = \frac{1}{\sqrt{e^3}} \approx 0,22.$$

   (ii) Vi söker $P(A|B)$ där A = "väntetid över 15 minuter" och B = "väntetid över 10 minuter". Vi noterar att $A \cap B = A$ vilket betyder att vi kan beräkna den betingade sannolikheten på följande vis:

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{P(X > 15)}{P(X > 10)} \\ &= \frac{1/\sqrt{e^3}}{1 - F(10)} = \frac{1/\sqrt{e^3}}{e^{-\frac{10}{10}}} = \frac{e}{\sqrt{e^3}} = \frac{1}{\sqrt{e}} \approx 0,60. \end{aligned}$$

   Notera att detta stämmer överens med egenskapen hos exponentiellt fördelade slumvariabler som säger att sådana fördelningar "saknar minne". Om $h, t \in \mathbb{R}$ så gäller

$$P(X > h + t | X > h) = P(X > t)$$

   då $X \sim Exp(\lambda)$. Detta betyder att frågan i uppgiften kan även besvaras genom att observera att

$$\begin{aligned} P(A|B) &= P(X > 15 | X > 10) \\ &= P(X > 10 + 5 | X > 10) = P(X > 5) = e^{-1/2} = \frac{1}{\sqrt{e}} \approx 0,60. \end{aligned}$$

□

2. Anta att $Y \sim N(\mu, \sigma^2)$ är en normalfördelad stokastisk variabel med parametrarna $\mu \in \mathbf{R}$ och $\sigma > 0$. Beräkna sannolikheterna

   (i) $P(\mu - \sigma < Y < \mu + \sigma)$,

   (ii) $P(\mu - 2\sigma < Y < \mu + 2\sigma)$.

   Lösning: Vi kan normalisera Y till en standard normalfördelad stokastisk variabel

$$X = \frac{Y - \mu}{\sigma} \sim N(0,1)$$

   och på det viset få fördelningsfunktionen till Y som

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

   där $\Phi(y)$ är alltså fördelningsfunktionen till den normaliserade variabeln X. Funktionen $\Phi(y)$ har den viktiga egenskapen av att $\Phi(-y) = 1 - \Phi(y)$ för alla $y \in \mathbb{R}$. Värden för denna fördelningsfunktion hittas i t.ex. MAOL:s tabeller.

   (i)

$$\begin{aligned} P(\mu - \sigma < Y < \mu + \sigma) &= P(Y < \mu + \sigma) - P(Y < \mu - \sigma) \\ &= \Phi\left(\frac{\mu + \sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - \sigma - \mu}{\sigma}\right) \\ &= \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 \approx 2 \cdot 0,8413 - 1 = 0,6826. \end{aligned}$$

   (ii)

$$\begin{aligned} P(\mu - 2\sigma < Y < \mu + 2\sigma) &= P(Y < \mu + 2\sigma) - P(Y < \mu - 2\sigma) \\ &= \Phi\left(\frac{\mu + 2\sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - 2\sigma - \mu}{\sigma}\right) \\ &= \Phi(2) - \Phi(-2) = 2\Phi(2) - 1 \approx 2 \cdot 0,9773 - 1 = 0,9546. \end{aligned}$$

   □

3. (2:41) En fabrik producerar kuvert. Anta att vikten av kuverten följer en $N(1.95, (0.05)^2)$-fördelning, där enheten är gram. Vilken är sannolikheten att

   (i) vikten av ett godtyckligt valt kuvert ligger mellan 1.8 g och 2.1 g,

   (ii) vikten av ett godtyckligt valt kuvert är större än 2 g?

   (iii) Vilket är väntevärdet på antalet kuvert som väger över 2 g i en packning med 100 kuvert?

   Lösning: Eftersom $X \sim N(1,95; (0,05)^2)$ så kan vi igen normalisera variabeln till den standarda normalfördelningen genom

$$Y = \frac{X - 1,95}{0,05} \sim N(0,1).$$

   Nu är fördelningsfunktionen för X enkelt uttryckbar via fördelningsfunktionen för Y:

$$F_X(x) = P(X \leq x) = \Phi\left(\frac{x - 1,95}{0,05}\right)$$

(i)

$$\begin{aligned} P(1,8 < X < 2,1) &= P(X < 2,1) - P(X < 1,8) \\ &= \Phi\left(\frac{2,1-1,95}{0,05}\right) - \Phi\left(\frac{1,8-1,95}{0,05}\right) \\ &= \Phi(3) - \Phi(-3) = 2\Phi(3) - 1 \approx 2 \cdot 0,9987 - 1 = 0,9974. \end{aligned}$$

(ii)

$$P(X > 2) = 1 - P(X \leq 2) = 1 - \Phi\left(\frac{2-1,95}{0,05}\right) = 1 - \Phi(1) \approx 0,1587.$$

(iii) Situationen kan tolkas som att vi plockar upp 100 kuvert och kollar om brevet väger över eller under 2g. Det finns alltså två utfallsmöjligheter för varje enskilt kuvert vilket betyder att antalet kuvert av 100 kuvert som väger över 2g (betecknad som Z) följer en binomialfördelning:

$$Z \sim Bin(100, P(X > 2)) = Bin(100, 1 - \Phi(1)).$$

Väntevärdet för en binomialfördelning $Bin(n,p)$ är np från vilket vi kan erhålla

$$\mathcal{E}(X) = 100 \cdot (1 - \Phi(1)) \approx 15,87.$$

□

4. (2:42) Bestäm väntevärdet $\mathcal{E}(X)$, då den stokastiska variabeln X har en kontinuerlig fördelning med frekvensfunktionen

(i) $f(x) = 8/x^3, \quad x > 2,$

(ii) $f(x) = xe^{-x^2/2}, \quad x > 0.$

Lösning: En stokastisk variabel X med kontinuerlig fördelning och frekvensfunktion f har väntevärdet

$$\mathcal{E}(X) = \int_{-\infty}^{\infty} xf(x)dx$$

då integralen existerar och är ändlig.

(i) Eftersom frekvensfunktionen är definierad för $x > 2$ (och underförstått att den är 0 då $x \leq 2$) får vi väntevärdet genom att integrera från 2 till ∞

$$\begin{aligned} \mathcal{E}(X) &= \int_2^{\infty} x\frac{8}{x^3}dx = \int_2^{\infty} \frac{8}{x^2}dx \\ &= \lim_{a\to\infty} \int_2^a \frac{8}{x^2}dx = \lim_{a\to\infty}\left(-\frac{8}{x}\Big|_2^a\right) = \lim_{a\to\infty}\left(-\frac{8}{a} + 4\right) = 4. \end{aligned}$$

(ii) Frekvensfunktionen är definierad för $x > 0$ (och noll annars) vilket betyder att vi kan integrera från 0 till ∞. Tricket är att tillämpa partiell integration samt kunskapen om att

$$\int_{-\infty}^{\infty} e^{-x^2/2}dx = \sqrt{2\pi}$$

vilket beräknades på föreläsningarna. Därmed är

$$
\begin{aligned}
\mathcal{E}(X) &= \int_0^\infty x(xe^{-x^2/2})dx = \lim_{a\to\infty}\int_0^a \underbrace{x}_{u}(\underbrace{xe^{-x^2/2}}_{v'})dx \\
&= \lim_{a\to\infty}\left(x(-e^{-x^2/2})\Big|_0^a - \int_0^a (-e^{-x^2/2})dx \right) \\
&= \underbrace{\lim_{a\to\infty}\left(-ae^{-a^2/2}\right)}_{0} + \lim_{a\to\infty}\left(\int_0^a e^{-x^2/2}dx\right) \\
&= \int_0^\infty e^{-x^2/2}dx \overset{(*)}{=} \frac{1}{2}\int_{-\infty}^\infty e^{-x^2/2}dx = \frac{1}{2}\sqrt{2\pi} = \sqrt{\frac{\pi}{2}}.
\end{aligned}
$$

Observera att (*) gäller eftersom funktionen $f(x) = e^{-x^2/2}$ är symmetrisk runt origo, dvs. $f(x) = f(-x)$ för alla $x \in \mathbb{R}$.

$\square$

5. (2:45) Anta att X_1 och X_2 är oberoende stokastiska variabler samt att $X_i \sim Tas(0,1)$ då $i = 1, 2$. Bestäm väntevärdet av de stokastiska variablerna

$$
Y(\omega) = \max\{X_1(\omega), X_2(\omega)\}, \quad Z(\omega) = \min\{X_1(\omega), X_2(\omega)\}, \quad \text{där } \omega \in \Omega.
$$

Tips: börja med att söka fördelningsfunktionerna till Y och Z, där Exempel 2.6.11 från kompendiet hjälper.

Lösning: Eftersom X_1 och X_2 är båda likformigt fördelade på intervallet $(0,1)$ så är

$$
f_{X_1}(x) = f_{X_2}(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & x \leq 0 \text{ eller } x \geq 1 \end{cases}
$$

vilket betyder att

$$
F_{X_1}(x) = F_{X_2}(x) = \begin{cases} 1, & x \geq 1 \\ x, & 0 < x < 1 \\ 0, & x \leq 0. \end{cases}
$$

Vi börjar med att beräkna väntevärdet $\mathcal{E}(Y)$ vilket vi gör genom att först beräkna fördelningsfunktionen för Y. Ifall $Y(\omega) \leq x$ så följer det från definitionen att både $X_1(\omega) \leq x$ och $X_2(\omega) \leq x$ måste gälla. Därtill är X_1 och X_2 oberoende från vilket följande härledning av fördelningsfunktionen följer:

$$
\begin{aligned}
F_Y(x) &= P(Y \leq x) = P(\{X_1 \leq x\} \cap \{X_2 \leq x\}) \\
&= P(X_1 \leq x)P(X_2 \leq x) \\
&= F_{X_1}(x)F_{X_2}(x) = \begin{cases} 1, & x \geq 1 \\ x^2, & 0 < x < 1 \\ 0, & x \leq 0. \end{cases}
\end{aligned}
$$

För att hitta frekvensfunktionen noterar vi att

$$
\begin{aligned}
& F_Y(x) = \int_{-\infty}^{x} f(x)dx \\
\Leftrightarrow \quad & \frac{d}{dx}F_Y(x) = f_Y(x) - \underbrace{f_Y(0)}_{0} \\
\Leftrightarrow \quad & \frac{d}{dx}F_Y(x) = f_Y(x)
\end{aligned}
$$

Från detta får vi frekvensfunktionen:

$$
f_Y(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & x \geq 1 \text{ eller } x \leq 0 \end{cases}
$$

Genom att använda frekvensfunktionen kan vi enkelt beräkna väntevärdet. Det räcker att integrera från 0 till 1 eftersom funktionen är 0 utanför intervallet.

$$
\mathcal{E}(Y) = \int_{-\infty}^{\infty} x f_Y(x)dx = \int_0^1 x2xdx = \frac{2x^3}{3}\Big|_0^1 = \frac{2}{3}.
$$

Nu fortsätter vi med att beräkna väntevärdet $\mathcal{E}(Z)$ vilket vi liknande som tidigare gör genom att beräkna fördelningsfunktionen $F_Z(x)$. Ifall $Z(\omega) \leq x$ så följer det från definitionen att endera $X_1(\omega) \leq x$ eller $X_2(\omega) \leq x$ måste gälla. Genom att använda summaformeln för sannolikheten av unioner får vi att

$$
\begin{aligned}
F_Z(x) &= P(Z \leq x) = P(\{X_1 \leq x\} \cup \{X_2 \leq x\}) \\
&= P(X_1 \leq x) + P(X_2 \leq x) - P(\{X_1 \leq x\} \cap \{X_2 \leq x\}) \\
&= F_{X_1}(x) + F_{X_2}(x) - F_{X_1}(x)F_{X_2}(x) = \begin{cases} 2, & x \geq 1 \\ 2x - x^2, & 0 < x < 1 \\ 0, & x \leq 0. \end{cases}
\end{aligned}
$$

På samma sätt som tidigare får vi nu frekvensfunktionen och därmed också väntevärdet på följande vis:

$$
\begin{aligned}
& f_Z(x) = \frac{d}{dx}F_Z(x) = \begin{cases} 2 - 2x, & 0 < x < 1 \\ 0, & x \leq 0 \text{ eller } x \geq 1 \end{cases} \\
\Rightarrow \quad & \mathcal{E}(Z) = \int_0^1 x(2-2x)dx = \left(x^2 - \frac{2x^3}{3}\right)\Big|_0^1 = 1 - \frac{2}{3} = \frac{1}{3}.
\end{aligned}
$$

□

6. (2:58) Anta att $X \sim Exp(\lambda)$. Bestäm fördelnings- och frekvensfunktionen till den stokastiska variabeln

$$
Y = \sqrt{X},
$$

samt dess väntevärde.

Lösning: Eftersom X är $Exp(\lambda)$-fördelad vet vi (se uppgift 1)

$$
F_X(x) = 1 - e^{-\lambda x}, \text{ då } x \geq 0
$$

Därmed är

$$F_Y(y) = P(Y \le y) = P(\sqrt{X} \le y) = P(X \le y^2) = F_X(y^2) = 1 - e^{-\lambda y^2}$$

då $y > 0$. På samma vis som i uppgift 5. hittar vi frekvensfunktionen genom att derivera fördelningsfunktionen:

$$f_Y(y) = \frac{d}{dy}F_Y(y) = 2\lambda y e^{-\lambda y^2}$$

Nu finns det endast väntevärdet kvar vilket vi kan beräkna genom att använda partiell integrering:

$$\begin{aligned}
\mathcal{E}(Y) &= \int_{-\infty}^{\infty} yf(y)dy = \lim_{a\to\infty}\int_0^a \underbrace{y}_{u} \cdot \underbrace{2\lambda y e^{-\lambda e^2}}_{v'} dy \\
&= \lim_{a\to\infty}\left(-ye^{-\lambda y^2}\Big|_0^a - \int_0^a (-e^{-\lambda y^2})dy \right) \\
&= \underbrace{\lim_{a\to\infty}\left(-ye^{-\lambda y^2}\Big|_0^a \right)}_{0} + \lim_{a\to\infty}\left(\int_0^a e^{-\lambda y^2} dy \right) \\
&= \int_0^{\infty} e^{-\lambda y^2} dy = \frac{1}{2}\int_{-\infty}^{\infty} e^{-\lambda y^2} dy
\end{aligned}$$

Låt oss substituera $\sigma = 1/\sqrt{2\lambda} \Leftrightarrow \lambda = 1/(2\sigma^2)$. Vårt mål är att skriva om integralen som integralen av en frekvensfunktion för någon normalfördelad slumpvariabel. Eftersom integralen över hela reella axeln av en frekvensfunktion är 1 (från definitionen) så kan vi på det viset få bort den krångliga integralen. Nu är

$$\begin{aligned}
\frac{1}{2}\int_{-\infty}^{\infty} e^{-\lambda y^2} dy &= \frac{1}{2}\int_{-\infty}^{\infty} e^{-y^2/(2\sigma^2)} dy = \frac{\sigma\sqrt{2\pi}}{2} \cdot \underbrace{\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-((y-0)/\sqrt{2}\sigma)^2} dy}_{\substack{\text{integralen över frekvensfunktionen} \\ \text{av en } N(0,\sigma^2)\text{-fördelad variabel}}} \\
&= \frac{\sigma\sqrt{2\pi}}{2} \cdot 1 = \frac{\sqrt{2\pi}}{2\sqrt{2\lambda}} = \frac{1}{2}\sqrt{\frac{\pi}{\lambda}}.
\end{aligned}$$

Observera att genom att substituera $\lambda = 1/2$ får vi resultatet

$$\int_{-\infty}^{\infty} e^{-y^2/2} dy = \sqrt{2\pi}$$

vilket vi använde i uppgift 4. $\square$