



Ugeseddel 3 - Sandsynlighedsteori og Statistisk

Levi van Boekel

Opgave 1

Vi lader X være den stokastiske variabel, der tæller antallet af score som en basketball får på 10 skud. Fordelingen for X er givet ved:

$$P(x = 0) = 0.5$$

$$P(x = 1) = 0.5$$

Hvor $x = 1$ betegner succes.

1.1

Hvis vi antager, at udfaldende er indbyrdes uafhængige, er de binomialfordelte og ssh. for at få 8 succeser kan derfor bestemmes som:

$$p(8) = \binom{10}{8} \left(\frac{1}{2}\right)^8 \left(\frac{1}{2}\right)^2 \approx 0.04$$

1.2

Den forventede værdi er givet ved bestemmes som:

$$\mathbb{E}(X) = np = \frac{1}{2} \cdot 10 = 5$$

Mens variansen bestemmes som:

$$\mathbb{V}(X) = np(1-p) = \frac{1}{2} \cdot 10 \cdot \frac{1}{2} = 2.5$$

Opgave 2

Vi lader X være en diskret stokastisk variabel med sandsynlighedsfunktion $f(x) = \frac{x}{8}$ for $x = \{1, 2, 5\}$.

2.1

Den forventede værdi findes ved det vægtede gennemsnit af udfaldende. Da:

$$\mathbb{E}(X) = \frac{1}{8} \cdot 1 + \frac{2}{8} \cdot 2 + \frac{5}{8} \cdot 5 = 3.75$$

2.2

Variansen bestemmes hernæst:

$$\mathbb{V}(X) = \left(\frac{1}{8} \cdot 1 + \frac{2}{8} \cdot 4 + \frac{5}{8} \cdot 25\right) - 3.75^2 \approx 2.69$$

2.3

Den forventede værdi af $\mathbb{E}(2X + 3)$ kan bestemmes som gjort nedenfor, idet middelværdien er en linear transformation:

$$3 + 2 \cdot \mathbb{E}(X) = 3 + 7.5 = 10.5$$

Opgave 3

3.1

Den forventede værdi er:

$$\mathbb{E}(X) = \frac{1}{10} \cdot 0 + \frac{3}{10} \cdot 1 + \frac{3}{10} \cdot 2 + \frac{2}{10} \cdot 3 + \frac{1}{10} \cdot 4 = 1.9$$

3.2

Variansen findes nedenfor:

$$\mathbb{V}(X) = \left(\frac{1}{10} \cdot 0 + \frac{3}{10} \cdot 1^2 + \frac{3}{10} \cdot 2^2 + \frac{2}{10} \cdot 3^2 + \frac{1}{10} \cdot 4^2 \right) - 66.5^2 = 1.29$$

3.3

Ejerens nettoprofit kan skrives som:

$$\pi = 35X - 40 \tag{1}$$

Den forventede værdi er altså:

$$\mathbb{E}(X)' = -40 + 35 \cdot 1.9 = 26.5$$

Variansen bliver:

$$\mathbb{V}(X)' = 35^2 \cdot \mathbb{V}(X) = 35^2 \cdot 1.29 \approx 1.580$$

Opgave 4

4.1

Punktsandsynligheden i en Poisson fordeling er givet ved:

$$P(X = k) = \frac{\mu^k e^{-\mu}}{k!} \tag{2}$$

I vores tilfælde er ssh for præcis 7 opkald altså:

$$P(X = 7) = \frac{10^7 e^{-7}}{7!} \approx 0.09$$

4.2

Vi kan finde ssh. for at der kommer højst 7 opkald som den komplementære ssh. Dvs:

$$P(X \leq 7) = 1 - \frac{10^{10} e^{-10}}{10!} - \frac{10^9 e^{-9}}{9!} - \frac{10^8 e^{-8}}{8!} \approx 0.22$$

4.3

Vi kan finde ssh. for at der kommer mellem 3 og 7 (inklusive) opkald ved at fratrække $P(X = 0)$, $P(X = 1)$, $P(X = 2)$ fra ovenstående:

$$P(3 \leq x \leq 7) = 0.63 - \frac{10^0 e^{-0}}{0!} - \frac{10^1 e^{-1}}{1!} - \frac{10^2 e^{-2}}{2!} \approx 0.217$$

Opgave 5

En stokastisk variabel X , har middelværdi 5 og varians 2. Vi skal nu finde: $\mathbb{E}(7 + 8X + X^2)$. Vi ved, at:

$$\mathbb{E}(X') = 7 + 8 \cdot 5 + \mathbb{E}(X^2)$$

Vi ved, at variansen af en stokastisk variabel bestemmes som $E(X^2) - E(X)^2$. Da vi ved, at $2 = E(X^2) - 25$, ved vi, at $E(X^2) = 27$, hvor middelværdien er:

$$\mathbb{E}(X') = 7 + 8 \cdot 5 + 27 = 74$$

Opgave 6

Igennem opgaven antager vi, at vi har en cykel til en værdi af 4000 kr. Sandsynligheden for at cyklen bliver stjålet i løbet af et år er 5 pct. Der findes en forsikring, der erstatter hele beløbet, hvis cyklen bliver stjålet.

6.2

Vi lader X være en stokastisk variabel, der angiver værdien af cyklen i tilfældet af, at der ikke er tegnet en forsikring. Vi lader $X = 0$ betegne tilfældet, hvor cyklen er stjålet. Det gælder, at:

$$P(X = 0) = 0.05$$

$$P(X = 1) = 0.95$$

Middelværdien af X bliver:

$$\mathbb{E}(X) = 0.05 \cdot 0 + 0.95 \cdot 4000 = 3800$$

6.3

Vi lader nu Y være en stokastisk variabel som angiver værdien af cyklen minus udgifter til forsikring (400 kr.). Fordelingen er uændret. Middelværdien bliver:

$$\mathbb{E}(Y) = 0.05 \cdot 3600 + 0.95 \cdot 3600 = 3600$$

Cyklen er altså 3600 værd uanset om den stjæles eller ej.

6.4

Fordelingen er stadigvæk uændret. Middelværdien for den stokastiske variabel Z er vist nedenfor:

$$\mathbb{E}(Z) = 0.05 \cdot 2850 + 0.95 \cdot 3850 = 3800$$

6.5

Middelværdien for er den samme i 6.2 og 6.4, men lavest i 6.3. Alligevel kan 6.3 være den mest fordelagtige, hvis forbrugeren er meget risikoavers og oplever udgiften til selvrisiko som værende meget *negativ*.

6.6

Vi lader nyttefunktionen være givet ved:

$$u(v) = 10v - 0.001v^2 \quad v \in (0, 4000) \quad (3)$$

Nyttefunktionen er en parabel, der peger nedad. FOC er: $10 = 0.002v$, dvs. $v_* = 5000$, da det ligger udenfor definitionsområdet, har vi dog, at $v_* = 4000$ i vores tilfælde. Den forventede nytte i de tre tilfælde bliver:

$$\mathbb{E}(U(X)) = 0.05 \cdot 0 + 0.95 \cdot (10 \cdot 4000 - 0.001 \cdot 4000^2) = 22.800$$

$$\mathbb{E}(U(Y)) = (10 \cdot 3600 - 0.001 \cdot 3600^2) = 23.040$$

$$\mathbb{E}(U(Z)) = 0.05(10 \cdot 2850 - 0.001 \cdot 2850^2) + 0.95(10 \cdot 3850 - 0.001 \cdot 3850^2) \approx 23.512$$

Personen vil under den givne nytte vælge forsikring Z , da den giver den højeste nytte.

6.7

Det generelle udtryk for den forventede værdi af den stokastiske variabel $u(W)$ er givet ved:

$$10 \cdot \mathbb{E}(W) - 0.001 \cdot \mathbb{E}(W^2)$$

Vi ved, at der gælder, at:

$$\mathbb{V}(X^2) = \mathbb{E}(U(X^2)) - \mathbb{E}(U(X))^2 \longleftrightarrow \mathbb{V}(X^2) + \mathbb{E}(U(X))^2 = \mathbb{E}(U(X^2))$$

Derfor har vi:

$$10 \cdot \mathbb{E}(W) - 0.001(\mathbb{V}(W^2) + \mathbb{E}(U(W))^2) \longleftrightarrow 10 \cdot \mathbb{E}(W) - 0.001\mathbb{V}(W^2) - 0.001\mathbb{E}(U(W))^2$$

Som giver det ønskede udtryk. Fortolkningen er, at den forventede værdi af nyten, afhænger positivt af den forventede værdi af cyklen (med aftagende grænseprodukt), men afhænger negativt af variansen, idet forbrugeren til en vis grad er risikoavers.

6.8

Variansen i de tre tilfælde er givet ved:

$$\mathbb{V}(X) = 0.95 \cdot 4000^2 - 3800^2 = 760.000$$

$$\mathbb{V}(Y) = (0.05 \cdot 3600^2 + 0.95 \cdot 3600^2) - 3600^2 = 0$$

$$\mathbb{V}(Z) = (0.05 \cdot 2850 + 0.95 \cdot 3850) - 3800^2 = 47.500$$

Resultatet stemmer overens med 6.6 og 6.7, idet muligheden med *moderat* varians foretrækkes af forbrugeren.

Opgave 7

7.1

De marginale fordelinger af X_A og X_B er givet ved:

$$P(X_A = 0) = 0.11$$

$$P(X_A = 20) = 0.28$$

$$P(X_A = 40) = 0.61$$

og:

$$P(X_B = 0) = 0.08$$

$$P(X_B = 20) = 0.12$$

$$P(X_B = 40) = 0.8$$

7.2

Middelværdien og variansen af de de stokastiske variable findes nedenfor:

$$\mathbb{E}(X_A) = 0.28 \cdot 20 + 0.61 \cdot 40 = 30$$

$$\mathbb{V}(X_A) = (0.28 \cdot 20^2 + 0.61 \cdot 40^2) - 30^2 = 188$$

$$\mathbb{E}(X_B) = 0.12 \cdot 20 + 0.8 \cdot 40 = 34.4$$

$$\mathbb{V}(X_B) = (0.12 \cdot 20^2 + 0.8 \cdot 40^2) - 34.4^2 = 144.64$$

7.3

Vi starter med at bestemme kovariansen ved:

$$20 \cdot 20 \cdot 0.06 + 40 \cdot 20 \cdot 0.06 + 20 \cdot 40 \cdot 0.5 = 0.2 + 40 \cdot 40 \cdot 0.5 = 1032$$

Kovariansen er da:

$$\text{Cov}(X_A, X_B) = 1032 - 30 \cdot 34.4 = 0$$

Fordi kovariansen er 0, er korrelationskoefficienten også 0.

7.4

De betingede fordelinger af A 's arbejdstid givet, at B 's arbejdstid er hhv. 0, 20 og 40, er:

$$\begin{aligned} P(X_A = 0 \mid X_B = 0) &= \frac{0.01}{0.08} = 0.125 \\ P(X_A = 20 \mid X_B = 0) &= \frac{0.02}{0.08} = 0.25 \\ P(X_A = 40 \mid X_B = 0) &= \frac{0.02}{0.08} = 0.625 \\ P(X_A = 0 \mid X_B = 20) &= \frac{0}{0.12} = 0 \\ P(X_A = 20 \mid X_B = 20) &= \frac{0.06}{0.12} = 0.5 \\ P(X_A = 40 \mid X_B = 20) &= \frac{0.06}{0.12} = 0.5 \\ P(X_A = 0 \mid X_B = 40) &= \frac{0.1}{0.8} = 0.125 \\ P(X_A = 20 \mid X_B = 40) &= \frac{0.06}{0.8} = 0.25 \\ P(X_A = 40 \mid X_B = 40) &= \frac{0.06}{0.8} = 0.625 \end{aligned}$$

7.5

Nej, de stokastiske variable er ikke uafhængige. Betragt f.eks. ssh. for at $P(X_A = 40 \mid X_B = 40)$, her er $0.61 \cdot 0.8 \neq 0.5$

7.6

Vi betragter nu en ny stokastisk variabel: $X_c = \frac{X_a + X_b}{2}$, som beskriver den gennemsnitlige arbejdstid i husstanden. Middelværdien for den transformerede stokastiske variabel kan bestemmes som:

$$\mathbb{E}(X_c) = \frac{1}{2}\mathbb{X}_a + \frac{1}{2}\mathbb{X}_b = \frac{1}{2} \cdot 30 + \frac{1}{2} \cdot 34.4 = 32.2$$

Variansen bliver altså det følgende, idet de er ukorrelerede

$$\frac{1}{4} \cdot (\mathbb{V}(X_A) + \mathbb{V}(X_B)) = \frac{1}{4} \cdot (188 + 144.64) = 83.16$$

Opgave 8

Vi lader X_1, X_2, X_3 være identisk fordelte, uafhængige stokastiske variable med samme varians. Vi ønsker at finde nedenstående udtryk

$$\text{corr}(X_1 + X_2, X_2 + X_3) = \frac{\text{Cov}((X_1 + X_2), (X_2 + X_3))}{\sqrt{\text{Var}(X_1 + X_2) \text{Var}(X_2 + X_3)}} \quad (4)$$

Vi starter med at betragte:

$$\text{Cov}((X_1 + X_2), (X_2 + X_3)) = \text{Cov}(X_1, X_2) + \text{Cov}(X_1, X_3) + \text{Var}(X_2) + \text{Cov}(X_2, X_3) = \sigma^2$$

Idet alle har samme varians, σ^2 . Vi har altså:

$$\text{corr}(X_1 + X_2, X_2 + X_3) = \frac{\sigma^2}{\sqrt{4\sigma^4}} = \frac{1}{2} \quad (5)$$