

Sandsynlighedsteori og Statistik: Ugeseddel 8

Levi van Boekel

Uge 44, 2024

Opgave 1

1.1

Det gælder generelt, at man kan transformere en vilkårlig normalfordeling til en standardnormalfordeling ved $N \sim (0, 1)$:

$$Y = \frac{X - \mu}{\sigma}$$

Dvs. ved at fratrække middelværdien og dividere med standardafvigelsen. I vores tilfælde er variansen σ^2 og middelværdien μ , da har vi:

$$Y = \frac{1}{\sqrt{\sigma^2}} \cdot (X - \mu) \quad (1)$$

Som er en almindelig transformation, hvor vi kan bruge transformationssætningen. Da grænserne for en normalfordeling altid er fra minus uendelig til plus uendelig, og ovenstående er en lineær transformation, der ikke ændrer grænserne, har vi:

$$t^{-1}(x) = y \cdot \sigma + \mu \quad (2)$$

Og:

$$\frac{d}{dy} t^{-1}(x) = \sigma \quad (3)$$

Da standardnormalfordelingen er givet ved:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4)$$

Skal vi blot bestemme: $p(t^{-1}(y)) \cdot \left| \frac{d}{dy} t^{-1}(y) \right|$, som er:

$$\begin{aligned} q(y) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y\sigma + \mu - \mu)^2}{2\sigma^2}} \sigma \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{(y^2\sigma^2)}{2\sigma^2}} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \end{aligned}$$

Vi kan nu vise, at den forventede værdi er 0, da:

$$\mathbb{E}\left((X - \mu) \cdot \frac{1}{\sigma}\right) = \frac{1}{\sigma} \cdot [\mathbb{E}(X) - \mathbb{E}(\mu)] = 0 \quad (5)$$

og:

$$\text{Var}\left((X - \mu) \cdot \frac{1}{\sigma}\right) = \frac{1}{\sigma^2} \cdot \text{Var}(X) = 1 \quad (6)$$

Vi havde kunne vise nøjagtig samme svar ved at indse, at vi kan omskrive transformationen til:

$$Y = -\frac{\mu}{\sigma} + \frac{1}{\sigma}X \quad (7)$$

Dvs:

$$Y = a + bX \quad (8)$$

Hvor $a = -\frac{\mu}{\sigma}$ og $b = \frac{1}{\sigma}$. Vi ved nu, at, hvis X er normalfordelt, er Y normalfordelt ved:

$$N \sim \left(-\frac{\mu}{\sigma} + \frac{\mu}{\sigma}, \frac{1}{\sigma^2} \cdot \sigma^2\right) = N \sim (0, 1)$$

1.2

Den marginale fordeling af X har middelværdi μ_X og varians σ_x^2 . Vi har da igen:

$$Y = \frac{-\mu_x}{\sigma_x} + \frac{1}{\sigma_x}X \quad (9)$$

Hvorfor Y også her er standardnormalfordelt med $N \sim (0, 1)$.

1.3

Summen (differencen) mellem to normalfordelte stokastiske variable er normalfordelte, hvorfor vi blot skal vise middelværdi og varians. Middelværdien er:

$$\mathbb{E}(Y - \beta X) = \mathbb{E}(Y) - \beta \cdot \mathbb{E}(X) = 0 \quad \text{da den forventede værdi for begge er nul} \quad (10)$$

Variansen findes som:

$$\begin{aligned}
 \text{Var}(Y - \beta X) &= \text{Var}(Y) + \beta^2 \text{Var}(X) - 2\beta \text{Cov}(XY) \\
 &= \sigma_Y^2 + \beta^2 \sigma_X^2 - 2\beta \sigma_{XY} \\
 &= \sigma_Y^2 + \frac{\sigma_{XY}^2}{\sigma_X^2} \cdot \sigma_X^2 - 2 \cdot \frac{\sigma_{XY}}{\sigma_X^2} \cdot \sigma_{XY} \\
 &= \sigma_Y^2 + \frac{\sigma_{XY}^2}{\sigma_X^2} - 2 \frac{\sigma_{XY}^2}{\sigma_X^2} \\
 &= \sigma_Y^2 - \beta \cdot \sigma_{XY}
 \end{aligned}$$

1.4

Vi har:

$$\mathbb{E}(XY) - \beta \cdot \mathbb{E}(X^2)$$

Kovariansen er defineret som: $\text{Cov} = \mathbb{E}(XY) - \mathbb{E}(X) \cdot \mathbb{E}(Y)$ og variansen er givet ved: $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. Idet de enkelte forventede værdier er 0, har vi:

$$\begin{aligned}
 \sigma_{XY} - \beta \sigma_X^2 \\
 \sigma_{XY} - \frac{\sigma_{XY}}{\sigma_X^2} \cdot \sigma_X^2 \\
 \sigma_{XY} - \sigma_{XY} &= 0
 \end{aligned}$$

Idet kovariansen af $Y - \beta X$ og X er defineret til:

$$\text{Cov} = \mathbb{E}((Y - \beta X)X) - \mathbb{E}(X) \cdot \mathbb{E}(Y - \beta X) \quad (11)$$

Og $\mathbb{E}(X)$ er nul, har vi vist, at kovariansen er nul. Det gælder unikt for normalfordelte stokastiske variable, at en kovarians på 0 er ensbetydende med uafhængighed.

Opgave 2

2.1, 2.2 & 2.3

X er normalfordelt med $N \sim (1, 1)$, mens Y er normalfordelt med $N \sim (0, 1)$. Kovariansen er ρ .

2.4

Den betingede middelværdi er givet ved:

$$\begin{aligned}\mu_{Y|X} &= \mu_Y + \frac{\sigma_{XY}}{\sigma_X^2} \cdot (X - \mu_X) \\ &= 0 + \rho \cdot (X - 1) \\ &= \rho \cdot X - \rho\end{aligned}$$

2.5

Vi har:

$$\begin{aligned}\mu_{X|Y} &= \mu_X + \frac{\sigma_{XY}}{\sigma_Y^2} \cdot (Y - \mu_Y) \\ &= 1 + \rho \cdot (Y - 0) \\ &= 1 + \rho \cdot Y\end{aligned}$$

2.6

Den betingede varians er:

$$\begin{aligned}\sigma_{Y|X}^2 &= \sigma_Y^2 - \frac{\sigma_{XY}^2}{\sigma_X^2} \\ &= 1 - \rho^2\end{aligned}$$

2.7

Hvis kovariansen er 0, er de stokastiske variable uafhængige (fordi de er normalfordelte). Hvis $\rho > 0$ samvarierer de positivt.

Opgave 3

3.1

Vi ved, at $\alpha = \mu_Y - \omega\mu_X$ og $\beta = \frac{\sigma_{XY}}{\sigma_X^2}$. Beta angiver dermed en kovariansen mellem X og Y , normeret med variansen på X .

3.2

Vi får oplyst: $\mu_Y = 0$ og $\sigma_Y^2 = \sigma_Y^2$. Hvis det ikke skal være i modstrid med den oprindelige model, skal der gælde, at:

$$\mathbb{E}(\beta X + \epsilon) = \beta \mathbb{E}(X) + \mathbb{E}(\epsilon) = \beta \mathbb{E}(X) = 0$$

Da vi får oplyst, at middelværdien af X er 0, er der ikke modstrid her. Nu tjekker vi variansen af modellen:

$$\text{Var}(\beta X + \epsilon) = \beta^2 \sigma_X^2 + \sigma^2 \quad \text{pga. uafhængighed} \quad (12)$$

Det skal derfor strengt taget gælde, at $\sigma_Y^2 = \beta^2 \sigma_X^2 + \sigma^2$ for at der ikke er modstrid.

Opgave 4

4.1

Vi har, at:

$$\mu_{Y|X} = X = \mu_Y + \frac{\sigma_{XY}}{\sigma_X^2} \cdot (X - \mu_X)$$

$$X = \mu_Y + \sigma_{XY} \cdot X$$

Derfor er det klart, at $\mu_Y = 0$ og $\sigma_{XY} = 1$. Vi mangler nu kun at finde variansen for Y , da:

$$\sigma_{Y|X}^2 = \sigma_Y^2 - \frac{\sigma_{XY}^2}{\sigma_X^2}$$

$$1 = \sigma_Y^2 - 1$$

$$2 = \sigma_Y^2$$

Hvorned det ønskede er fundet.

Opgave 5

5.1

Den betingede middelværdi er:

$$\begin{aligned} \mathbb{E}(C_t | P_t) &= \mathbb{E}(\hat{\alpha} P_t + \epsilon_t | P_t) \\ &= \hat{\alpha} \cdot P_t \quad \text{idet vi antager } \mathbb{E}(\epsilon_t | P_t) = 0 \\ &= \frac{\alpha}{1 - \alpha} P_t \end{aligned}$$

Parameteren måler altså hvordan forbrugsratioen reagerer på prisændringerne. Hvis parameteren er positiv, betyder det, at forbrugsratioen reagerer positivt på prisændringer og vice versa.

5.2

Fra LIE har vi, at:

$$\begin{aligned}\mathbb{E}(\hat{\alpha}_t \cdot P_t) &= \mathbb{E}(\mathbb{E}(\hat{\alpha}_t \cdot P_t \mid P_t)) \\ &= \mathbb{E}(P_t \cdot \mathbb{E}(\hat{\alpha}_t \mid P_t)) \\ &= 0 \quad \text{fordi vi antager uafhængighed mellem fejllid og } P_t\end{aligned}$$

Vi kan nu skrive fejllidet som:

$$\epsilon_t = C_t - \hat{\alpha}_t \cdot P_t \quad (13)$$

Vi kan nu finde OLS estimatoren ved at kræve, at produktet af prædiktorvariablen og residualerne er 0. Da:

$$\begin{aligned}\mathbb{E}((C_t - \hat{\alpha}_t \cdot P_t) \cdot P_t) &= 0 \\ \mathbb{E}(C_t P_t - \hat{\alpha}_t \cdot P_t^2) &= 0 \\ \mathbb{E}(C_t P_t) - \hat{\alpha}_t \cdot \mathbb{E}(P_t^2) &= 0 \\ \hat{\alpha}_{OLS_t} &= \frac{\mathbb{E}(C_t P_t)}{\mathbb{E}(P_t^2)} \\ &= \frac{\frac{1}{t} \sum_{i=1}^t C_t P_t}{\frac{1}{t} \sum_{i=1}^t P_t^2}\end{aligned}$$

5.3

OLS estimatet bliver da:

$$\tilde{\alpha}_{OLS} = \frac{\frac{1}{n} \sum_{i=1}^n c_t \cdot p_t}{\frac{1}{n} \sum_{i=1}^n p_t^2} \quad (14)$$

Opgave 6

6.1

Vi starter med at finde de marginale fordelinger, da:

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}(3Z_1) = 3 \cdot \mathbb{E}(Z_1) = 0 \\ \text{Var}(X) &= 3^2 \text{Var}(Z_1) = 9\end{aligned}$$

Og:

$$\begin{aligned}\mathbb{E}(Y) &= 2 \cdot \mathbb{E}(Z_1) + \mathbb{E}(Z_2) = 0 \\ \text{Var}(Y) &= 2^2 \text{Var}(Z_1) + \text{Var}(Z_2) = 4\end{aligned}$$

Hvor det sidste følger af uafhængighed. Vi finder kovariansen ved:

$$\begin{aligned}\text{Cov}(Y, X) &= \text{Cov}(2Z_1 + Z_2, 3Z_1) \\ &= \text{Cov}(Y, X) = \text{Cov}(2Z_1, 3Z_1) + \text{Cov}(Z_2, 3Z_1) \\ &= \text{Cov}(2Z_1, 3Z_1) = 2 \cdot 3 \cdot \text{Var}(Z_1) = 6 \cdot 1 = 6\end{aligned}$$

Altså er fordelingen:

$$\begin{pmatrix} Y \\ X \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 & 6 \\ 6 & 9 \end{pmatrix} \right) \quad (15)$$

6.2

Den betingede middelværdi er:

$$\begin{aligned}\mathbb{E}(2Z_1 + Z_2 \mid Z_1) &= 2 \cdot \mathbb{E}(Z_1 \mid Z_1) + \mathbb{E}(Z_2 \mid Z_1) \\ &= 2Z_1 + \mathbb{E}(Z_2) = 2Z_1\end{aligned}$$

Hvor anden linje følger af uafhængighed.

6.3

Den betingede middelværdi er:

$$\mathbb{E}(3Z_1 \mid Z_2) = 3 \cdot \mathbb{E}(Z_1) = 0$$

6.4

Den betingede middelværdi er:

$$\begin{aligned}\mu_{Y|X} &= \mu_Y + \frac{\sigma_{XY}}{\sigma_X^2} \cdot (X - \mu_X) \\ &= \frac{9}{6} \cdot X \\ &= \frac{2}{3} \cdot X\end{aligned}$$

Opgave 7

7.1

Da den stokastiske variable er ligefordelt på intervallet er ssh. for at få de forskellige udfald lig hinanden. Den forventede værdi nu:

$$\sum_{i \in I} x_i \cdot P(X = x_i) = -1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = 0$$

Variansen er:

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\ &= \frac{2}{3} \end{aligned}$$

Sandsynligheden for den stokastiske variabel er større eller lig 0 er lig:

$$P(X \geq 0) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

7.2

Den betingede middelværdi kan bestemmes ved:

$$\begin{aligned} P(x_i | X \geq 0) &= \sum_{i=-1}^1 x_i \cdot P(x_i | x \geq 0) = \sum_{i=-1}^1 x_i \cdot \frac{P(x_i, x \geq 0)}{P(x \geq 0)} \\ &= \frac{\sum_{i=-1}^1 \mathbf{1}(x \geq 0) x_i P(X = x_i)}{\frac{2}{3}} = \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2} \end{aligned}$$

7.3

Vi har:

$$p(x) = \frac{1}{2} \mathbf{1}(X \in I) \tag{16}$$

Da er middelværdien:

$$\int_{-1}^1 \frac{1}{2} x dx = \left[\frac{1}{4} x^2 \right]_{-1}^1 = 0$$

For at bestemme variansen skal vi nu blot bestemme $\mathbb{E}(X^2)$:

$$\int_{-1}^1 \frac{1}{2} x^2 dx = \left[\frac{1}{6} x^3 \right]_{-1}^1 = \frac{1}{6} + \frac{1}{6} = \frac{2}{3}$$

Og:

$$\int_0^1 \frac{1}{2} dx = \left[\frac{1}{2}x \right]_0^1 = \frac{1}{2}$$

7.4

Den betingede middelværdi er:

$$\begin{aligned}\mathbb{E}(X \mid X \geq 0) &= \frac{\int_{-1}^1 \mathbf{1}(0 \leq x \leq 1) \frac{1}{2} x dx}{P(X \geq 0)} \\ &= \frac{\int_0^1 \frac{1}{2} x}{\frac{1}{2}} \\ &= \frac{\left[\frac{1}{4} x^2 \right]_0^1}{\frac{1}{2}} \\ &= \frac{1}{4} \cdot 2 = \frac{1}{2}\end{aligned}$$