NOTE: Normalfordeling

KBJ, februar 2025 — *3u MA*

Definition: Kontinuert stokastisk variabel

En kontinuert stokastisk variabel X beskrives ved en tæthedsfunktion f, hvis det gælder at $f(x) \geq 0$ for $x \in \mathbb{R}$ samt

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Da siges den stokastiske variabel X at have middelværdien:

$$\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$$

Endvidere siges X at have variansen:

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\,dx$$

Dermed bliver spredningen $\sigma = \sqrt{\mathrm{Var}(X)}$

Definition: Standardnormalfordelingen

Tæthedsfunktionen ϕ for en *standardnormalfordeling* kan bestemmes som en løsning til differentialligningen:

$$y' = -x \cdot y$$

Løsningen skal være positiv, dvs. $\phi(x) \geq 0$, for at opfylde definitionen af en tæthedsfunktion.

Sætning 1: Standardnormalfordelingens tæthedsfunktion

Standardnormalfordelingens tæthedsfunktion er givet ved:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{1}{2}x^2}$$

Bevis:

Med separationsmetoden bestemmes løsningen til differentialligningen $\frac{dy}{dx} = -x \cdot y$ som:

$$\int \frac{1}{y} dy = \int -x\, dx$$

$$\ln|y| + c_1 = -\frac{1}{2}x^2 + c_2$$

Da vi har krævet $y > 0$ fås $\ln|y| = \ln(y)$ og dermed:

$$\ln(y) = -\frac{1}{2}x^2 + c_3$$

hvor $c_3 = c_2 - c_1$

$$\mathrm{e}^{\ln(y)} = \mathrm{e}^{-\frac{1}{2}x^2 + c_3}$$

$$y = \mathrm{e}^{c_3} \cdot \mathrm{e}^{-\frac{1}{2}x^2}$$

$$y = \mathrm{c} \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}$$

hvor $c = \mathrm{e}^{c_3} > 0$.

Konstanten c bestemmes ud fra definition af tæthedsfunktion, i det $\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}x^2}\, dx = \sqrt{2\pi}$:

$$\int_{-\infty}^{\infty} c \cdot \mathrm{e}^{-\frac{1}{2}x^2}\, dx = c \cdot \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}x^2}\, dx = c \cdot \sqrt{2\pi} = 1$$

$$c = \frac{1}{\sqrt{2\pi}}$$

Dermed er det bestemt at:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}$$

Da der for alle $x \in \mathbb{R}$ gælder $\frac{1}{\sqrt{2\pi}} > 0$ og $\mathrm{e}^x > 0$, er $\phi(x) \geq 0$ for alle $x \in \mathbb{R}$.

QED

Sætning 2: Standardnormalfordelingens middelværdi

En standardnormalfordelt stokastisk variabel har middelværdi 0.

Bevis:

$$\mu = E(X) = \int_{-\infty}^{\infty} x \cdot \phi(x)\, dx = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}\, dx = \frac{-1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} (-x) \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}\, dx$$

Det ses, at den sammensatte funktion $g\big(h(x)\big) = \mathrm{e}^{-\frac{1}{2}x^2}$ har indre $h(x) = -\frac{1}{2}x^2$ med $h'(x) = -x$ og ydre $g(x) = \mathrm{e}^x$, som har stamfunktion $G(x) = \mathrm{e}^x$.

Da integranden kan skrives på formen $g\big(h(x)\big) \cdot h'(x)$ kan integralet skrives $G\big(h(x)\big) + c$.

Da $G\big(h(x)\big) = \mathrm{e}^{-\frac{1}{2}x^2}$ fås altså:

$$\begin{aligned}
\int_{-\infty}^{\infty} (-x) \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}\, dx &= \lim_{a \to -\infty} \left(\lim_{b \to \infty} \int_a^b (-x) \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}\, dx \right) \\
&= \lim_{a \to -\infty} \left(\lim_{b \to \infty} \left[\mathrm{e}^{-\frac{1}{2}x^2} \right]_a^b \right) \\
&= \lim_{a \to -\infty} \left(\lim_{b \to \infty} \left(\mathrm{e}^{-\frac{1}{2}b^2} - \mathrm{e}^{-\frac{1}{2}a^2} \right) \right) \\
&= \lim_{a \to -\infty} \left(\lim_{b \to \infty} \mathrm{e}^{-\frac{1}{2}b^2} - \lim_{b \to \infty} \mathrm{e}^{-\frac{1}{2}a^2} \right) \\
&= \lim_{a \to -\infty} \left(0 - \mathrm{e}^{-\frac{1}{2}a^2} \right) \\
&= \lim_{a \to -\infty} -\mathrm{e}^{\frac{1}{2}a^2} \\
&= 0
\end{aligned}$$

Dermed fås altså:

$$\mu = E(X) = \int_{-\infty}^{\infty} x \cdot \phi(x)\, dx = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} (-x) \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}\, dx = \frac{1}{\sqrt{2\pi}} \cdot 0 = 0$$

QED

Sætning 3

En standardnormalfordelt stokastisk variabel X har spredningen 1.

Bevis:

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot \phi(x)\, dx = \int_{-\infty}^{\infty} x^2 \cdot \phi(x)\, dx = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} x^2 \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}\, dx$$

Om bestemt partiel integration gælder, i det G er en stamfunktion til g:

$$\int_a^b f(x) \cdot g(x)\, dx = [f(x) \cdot G(x)]_a^b - \int_a^b f'(x) \cdot G(x)\, dx$$

Vi vælger her $f(x) = x$ og $g(x) = x \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}$, således at $f(x) \cdot g(x) = x^2 \cdot \mathrm{e}^{-\frac{1}{2}x^2}$.

Heraf følger $f'(x) = 1$ og $G(x) = -\mathrm{e}^{-\frac{1}{2}x^2}$.

Da $\lim\limits_{x \to \infty} x \cdot \mathrm{e}^{-x} = 0$ (eksponentialfunktion aftager ”hurtigere" end lineær funktion vokser), fås:

$$\begin{aligned}
\int_a^b x^2 \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}\, dx &= \left[x \cdot \left(-\mathrm{e}^{-\frac{1}{2}x^2}\right)\right]_a^b - \int_a^b 1 \cdot \left(-\mathrm{e}^{-\frac{1}{2}x^2}\right) dx \\
&= -\left[x \cdot \mathrm{e}^{-\frac{1}{2}x^2}\right]_a^b + \int_a^b \mathrm{e}^{-\frac{1}{2}x^2}\, dx \\
&= \int_a^b \mathrm{e}^{-\frac{1}{2}x^2}\, dx = -\left[x \cdot \mathrm{e}^{-\frac{1}{2}x^2}\right]_a^b
\end{aligned}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} x^2 \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}\, dx &= \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2} \cdot x^2}\, dx - \lim_{a \to -\infty} \left(\lim_{b \to \infty} \left(b \cdot e^{-\frac{1}{2}b^2} - a \cdot e^{-\frac{1}{2}a^2} \right) \right) \\
&= \sqrt{2\pi} - \lim_{a \to -\infty} \left(\lim_{b \to \infty} b \cdot e^{-\frac{1}{2}b^2} - \lim_{b \to \infty} a \cdot e^{-\frac{1}{2}a^2} \right) \\
&= \sqrt{2\pi} - \lim_{a \to -\infty} \left(0 - a \cdot e^{-\frac{1}{2}a^2} \right) \\
&= \sqrt{2\pi} - 0 \\
&= \sqrt{2\pi}
\end{aligned}$$

Dermed fås

$$\mathrm{Var}(X) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} x^2 \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}\, dx = \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = 1$$

Og således kan vi slutte om spredningen:

$$\sigma(X) = \sqrt{\mathrm{Var}(X)} = \sqrt{1} = 1$$

QED.

Sætning 4: Monotoniforhold for tæthedsfunktionen

Standardnormalfordelingens tæthedsfunktion er jf. sætning 1 bestemt ved:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{1}{2}x^2}$$

Det gælder om $\phi(x)$ at den er voksende for $x \in]-\infty; 0]$ og aftagende for $x \in [0; \infty[$.

Bevis:

Med kædereglen ses: $\left(\mathrm{e}^{-\frac{1}{2}x^2}\right)' = -\frac{1}{2} \cdot 2x \cdot \mathrm{e}^{-\frac{1}{2}x^2} = -x \cdot \mathrm{e}^{-\frac{1}{2}x^2}$

Dermed fås:

$$\phi'(x) = \frac{-1}{\sqrt{2\pi}} \cdot x \cdot \mathrm{e}^{-\frac{1}{2}x^2}$$

De stationære punkter for ϕ bestemmes:

$$\phi'(x) = 0$$
$$\frac{-1}{\sqrt{2\pi}} \cdot x \cdot \mathrm{e}^{-\frac{1}{2}x^2} = 0$$

Med nulreglen fås da følgende løsning, da $\mathrm{e}^{-\frac{1}{2}x^2} > 0$ for alle x:
$x = 0$

Da $-\frac{1}{\sqrt{2\pi}} < 0$ og $\mathrm{e}^{-\frac{1}{2}x^2} > 0$ for alle værdier af x, afgøres fortegnet for ϕ' alene af fortegnet for x.

For $x < 0$ fås således $\phi'(x) > 0$ og $x > 0$ fås $\phi'(x) < 0$

Således følger det at ϕ er voksende for $x \leq 0$ og aftagende for $x \geq 0$.

QED

Sætning 5: Krumningsforhold for tæthedsfunktion

Om $\phi(x)$ gælder, at den er konveks for $x \in]-\infty; -1]$ og $x \in [1; \infty[$ samt konkav for $x \in [-1; 1]$.

Bevis:

Med produktreglen ses:

$$\left(\mathrm{e}^{-\frac{1}{2} \cdot x^2}\right)'' = \left(-x \cdot \mathrm{e}^{-\frac{1}{2}x^2}\right)' = (-x)' \cdot \mathrm{e}^{-\frac{1}{2}x^2} + (-x) \cdot \left(\mathrm{e}^{-\frac{1}{2}x^2}\right)' = -1 \cdot \mathrm{e}^{-\frac{1}{2}x^2} - x \cdot (-x) \cdot \mathrm{e}^{-\frac{1}{2}x^2} =$$

$$x^2 \cdot \mathrm{e}^{-\frac{1}{2}x^2} - 1 \cdot \mathrm{e}^{-\frac{1}{2}x^2} = (x^2 - 1) \cdot \mathrm{e}^{-\frac{1}{2}x^2}$$

Dermed fås:

$$\phi''(x) = \frac{1}{\sqrt{2\pi}} \cdot (x^2 - 1) \cdot \mathrm{e}^{-\frac{1}{2}x^2}$$

For at finde mulige vendepunkter løses ligningen $\phi''(x) = 0$:

$$\frac{1}{\sqrt{2\pi}} \cdot (x^2 - 1) \cdot \mathrm{e}^{-\frac{1}{2}x^2} = 0$$

Med nulreglen fås, idet $\mathrm{e}^{-\frac{1}{2}x^2} > 0$ for alle x:

$$x^2 - 1 = 0$$
$$x^2 = 1$$
$$x = -1 \text{ eller } x = 1$$

Da $\frac{1}{\sqrt{2\pi}} > 0$ og $\mathrm{e}^{-\frac{1}{2}x^2} > 0$ for alle x, har ϕ'' samme fortegn som $(x^2 - 1)$.

For $x^2 > 1$ er $x^2 - 1 > 0$. Dette gælder når $x < -1$ eller $x > 1$.

For $x^2 < 1$ er $x^2 - 1 < 0$. Dette gælder når $-1 < x < 1$

Da $\phi''(x) > 0$ for $x < -1$ er ϕ konveks for $x \in]-\infty; 1]$.

Da $\phi''(x) < 0$ for $-1 < x < 1$ er ϕ konkav for $x \in [-1; 1]$.

Da $\phi''(x) > 0$ for $x > 1$, er ϕ konveks for $x \in [1; \infty[$.

QED

Definition: Fordelingsfunktion

Standardnormalfordelingens fordelingsfunktion er defineret ved:

$$\Phi(x) = \int_{-\infty}^{x} \phi(t)\, dt$$

Sætning 6: Monotoni- og krumningsforhold for Φ

Φ er voksende på hele ℝ.

Φ er konveks for $x < 0$ og konkav for $x > 0$.

Bevis:

Da $\Phi'(x) = \phi(x)$, og da $\phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{1}{2}x^2} > 0$ for alle x, så er Φ voksende for alle x.

Da $\Phi''(x) = \phi'(x)$, og da der jf. sætning 1, gælder at $\phi'(x) > 0$ for $x < 0$ og $\phi'(x) < 0$ for $x > 0$, så er Φ konveks for $x \leq 0$ og konkav for $x \geq 0$.

QED

Indledende bemærkninger om generel normalfordeling

Den normalfordelte stokastiske variabel X, med middelværdi μ og spredning σ gælder, at den har tæthedsfunktionen:

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \mathrm{e}^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2}$$

Endvidere har den fordelingsfunktionen

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$

Standardnormalfordelingens har $\mu = 0$ og $\sigma = 1$ hvilket netop giver:

$$\phi(x) = \frac{1}{\sqrt{2\pi} \cdot 1} \cdot \mathrm{e}^{-\frac{1}{2} \cdot \left(\frac{x-0}{1}\right)^2} = \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{1}{2} \cdot x^2}$$

Sætning 7:

Det gælder at $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$

Bevis

Med kædereglen fås:

$$\left(\Phi\left(\frac{x-\mu}{\sigma}\right)\right)' = \left(\frac{x-\mu}{\sigma}\right)' \cdot \Phi'\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right)$$
$$= \frac{1}{\sigma} \cdot \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2} = \frac{1}{\sqrt{2\pi} \cdot \sigma} \mathrm{e}^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2} = f(x)$$

Det er altså vist at $\Phi\left(\frac{x-\mu}{\sigma}\right)$ er en stamfunktion til f.

Da forskellen på to stamfunktioner til samme funktion højest er en konstant, må det således gælde at $F(x) - \Phi\left(\frac{x-\mu}{\sigma}\right) = c$. Derfor bestemmes c i det det vides at $F(\mu) = 0{,}5$ og $\Phi(0) = 0{,}5$:

$$F(\mu) - \Phi\left(\frac{\mu-\mu}{\sigma}\right) = F(\mu) - \Phi(0) = 0{,}5 - 0{,}5 = 0 = c$$

Da $c = 0$ gælder det altså at $\Phi\left(\frac{x-\mu}{\sigma}\right)$ og $F(x)$ er den samme stamfunktion til f.

Det gælder altså at $\Phi\left(\frac{x-\mu}{\sigma}\right) = F(x)$.

Sætning 8:

Hvis der for en normalfordelt stokastisk variabel $X \sim N(\mu, \sigma)$ gælder, at $P\left(X \leq x_p\right) = p$ og $P\left(X \leq x_q\right) = q$, dvs. at p- og q-fraktilen for X er hhv. x_q og x_p, da gælder at middelværdi og spredning er givet ved:

$$\mu = \frac{\Phi^{-1}(q) \cdot x_p - \Phi^{-1}(p) \cdot x_q}{\Phi^{-1}(q) - \Phi^{-1}(p)}$$

og

$$\sigma = \frac{x_q - x_p}{\Phi^{-1}(q) - \Phi^{-1}(p)}$$

hvor Φ^{-1} er fraktilfunktionen for standardnormalfordelingen.

Bevis:

Hvis vi kalder fordelingsfunktionen for X for F, så ved vi altså at

$F\left(x_p\right) = p$ og $F\left(x_q\right) = q$

Heraf følger med sætning 6 at:

$\Phi\left(\frac{x_p - \mu}{\sigma}\right) = p$ og $\Phi\left(\frac{x_p - \mu}{\sigma}\right) = p$

Ved at bruge standardnormalfordelingens fraktilfunktion Φ^{-1} på begge sider fås:

$$\Phi^{-1}\left(\Phi\left(\frac{x_p - \mu}{\sigma}\right)\right) = \Phi^{-1}(p) \text{ og } \Phi^{-1}\left(\Phi\left(\frac{x_p - \mu}{\sigma}\right)\right) = \Phi^{-1}(p)$$

Og da $\Phi^{-1}\left(\Phi(x)\right) = x$ fås:

$$\frac{x_p - \mu}{\sigma} = \Phi^{-1}(p) \text{ og } \frac{x_q - \mu}{\sigma} = \Phi^{-1}(q)$$

I begge ligninger isoleres nu σ:

$$x_p - \mu = \Phi^{-1}(p) \cdot \sigma \quad \text{og} \quad x_q - \mu = \Phi^{-1}(q) \cdot \sigma$$

$$\sigma = \frac{x_p - \mu}{\Phi^{-1}(p)} \quad \text{og} \quad \sigma = \frac{x_q - \mu}{\Phi^{-1}(q)}$$

og de to udtryk sættes lig hinanden og μ isoleres:

$$\frac{x_p - \mu}{\Phi^{-1}(p)} = \frac{x_q - \mu}{\Phi^{-1}(q)}$$

$$\frac{x_p}{\Phi^{-1}(p)} - \frac{\mu}{\Phi^{-1}(p)} = \frac{x_q}{\Phi^{-1}(q)} - \frac{\mu}{\Phi^{-1}(q)}$$

$$\frac{x_p}{\Phi^{-1}(p)} - \frac{x_q}{\Phi^{-1}(q)} = \frac{\mu}{\Phi^{-1}(p)} - \frac{\mu}{\Phi^{-1}(q)}$$

$$\frac{x_p}{\Phi^{-1}(p)} - \frac{x_q}{\Phi^{-1}(q)} = \mu\left(\frac{1}{\Phi^{-1}(p)} - \frac{1}{\Phi^{-1}(q)}\right)$$

$$\mu = \frac{\frac{x_p}{\Phi^{-1}(p)} - \frac{x_q}{\Phi^{-1}(q)}}{\frac{1}{\Phi^{-1}(p)} - \frac{1}{\Phi^{-1}(q)}}$$

Brøken på højreside forlænges nu med $\Phi^{-1}(p) \cdot \Phi^{-1}(q)$ og således opnås:

$$\mu = \frac{\Phi^{-1}(q) \cdot x_p - \Phi^{-1}(p) \cdot x_q}{\Phi^{-1}(q) - \Phi^{-1}(p)}$$

Herefter sættes den fundne μ-værdi ind i et af udtrykkene for σ:

$$\sigma = \frac{x_p - \frac{\Phi^{-1}(q) \cdot x_p - \Phi^{-1}(p) \cdot x_q}{\Phi^{-1}(q) - \Phi^{-1}(p)}}{\Phi^{-1}(p)}$$

$$\sigma = \frac{\frac{x_p \cdot \left(\Phi^{-1}(q) - \Phi^{-1}(p)\right)}{\Phi^{-1}(q) - \Phi^{-1}(p)} - \frac{\Phi^{-1}(q) \cdot x_p - \Phi^{-1}(p) \cdot x_q}{\Phi^{-1}(q) - \Phi^{-1}(p)}}{\Phi^{-1}(p)}$$

$$\sigma = \frac{\frac{x_p \cdot \Phi^{-1}(q) - x_p\Phi^{-1}(p) - \Phi^{-1}(q) \cdot x_p + \Phi^{-1}(p) \cdot x_q}{\Phi^{-1}(q) - \Phi^{-1}(p)}}{\Phi^{-1}(p)}$$

$$\sigma = \frac{\frac{\Phi^{-1}(p) \cdot x_q - \Phi^{-1}(p) \cdot x_p}{\Phi^{-1}(q) - \Phi^{-1}(p)}}{\Phi^{-1}(p)}$$

$$\sigma = \frac{x_q - x_p}{\Phi^{-1}(q) - \Phi^{-1}(p)}$$

Hermed er det altså vist at: $\mu = \frac{\Phi^{-1}(q) \cdot x_p - \Phi^{-1}(p) \cdot x_q}{\Phi^{-1}(q) - \Phi^{-1}(p)}$ og $\sigma = \frac{x_q - x_p}{\Phi^{-1}(q) - \Phi^{-1}(p)}$