# F5, Kontinuerliga stokastiska variabler

Christian Tallberg

Statistiska institutionen

Stockholms universitet

## Kontinuerlig stokastisk variabel

- Vikten på ett slumpmässigt valt nyfött barn
- Livslängden på en slumpmässigt vald glödlampa

Sannolikheten för en kontinuerlig stokastisk variabel kan illustreras med en kurva (täthetsfunktion).

Följande villkor gäller för täthetsfunktionen

1. $f(x) \geq 0$ för alla $x$
2. $\Pr(a \leq X \leq b) =$ Ytan under $f(x)$ mellan $a$ och $b$. (Beräknas som $\int_a^b f(x)\,dx$)
3. Totala ytan under $f(x)$ skall vara lika med 1 (dvs $\int_{-\infty}^{\infty} f(x)\,dx = 1$)
4. $p(x) = \Pr(X = x) = 0$
5. Fördelningsfunktionen $F(x) = \Pr(X \leq x) =$Ytan under $f(x)$ mellan det lägsta värdet och punkten $x$.

OBS! För kontinuerliga variabler gäller att:

$$\begin{aligned}
\Pr(a \leq X \leq b) &= \Pr(a < X < b) = \Pr(a < X \leq b) \\
&= \Pr(a \leq X < b) = F(b) - F(a) \\
&= \int_a^b f(x)\,dx
\end{aligned}$$

$$\Pr(X \leq a) = \Pr(X < a) = F(a) = \int_{-\infty}^{a} f(x)\,dx$$

Väntevärde och varians för en kontinuerlig stokastisk variabel beräknas som

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f(x)\,dx \\
V(X) &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx
\end{aligned}$$

## Normalfördelningen

$X$ är en normalfördelad variabel om täthetsfunktionen ges av: $f(x) = \frac{1}{\sigma sqrt(2\pi)} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ där $-\infty < x < \infty$, och där $\mu$ och $\sigma$ är parametrar sådana att $-\infty < \mu < \infty$ och $0 < \sigma < \infty$.
Kortfattat skriver vi att $X$ är $N(\mu;\sigma)$.

De två parametrarna $\mu$ och $\sigma$ (eller $\sigma^2$) bestämmer normalfördelningens utseende.

Egenskaper:

- $E(X) = \mu$
- $V(X) = \sigma^2$
- Normalfördelningen är symmetrisk kring sitt väntevärde $\mu$.

Kuriosa:

För en normalfördelning gäller att:

- Ca 68 % av fördelningen finns inom $\mu \pm \sigma$
- Ca 95 % av fördelningen finns inom $\mu \pm 2\sigma$
- Ca 99.7 % av fördelningen finns inom $\mu \pm 3\sigma$, dvs nästan hela fördelningen finns inom gränserna $\mu \pm 3\sigma$.

Exempel:

Variabeln $X$ är $N(0;1)$. Beräkna följande sannolikheter.

1.
$$\begin{aligned} \Pr(X \le 1) &= \int_{-\infty}^{1} f(x)\,dx = F(1) = \Phi(1) \\ &= 0.8413 \text{ enl tabell 3a} \end{aligned}$$

2.
$$\begin{aligned} \Pr(0 \le X \le 1) &= \Pr(X \le 1) - \Pr(X \le 0) \\ &= F(1) - F(0) = \Phi(1) - \Phi(0) \\ &= 0.8413 - 0.5 = 0.3413 \end{aligned}$$

3.
$$\begin{aligned} \Pr(X > 0.73) &= 1 - \Pr(X \le 0.73) \\ &= 1 - \Phi(0.73) \\ &= 1 - 0.7673 = 0.2327 \end{aligned}$$

4.
$$\begin{aligned} \Pr(X > -1.59) &= \Pr(X < 1.59) = \Phi(1.59) \\ &= 0.9441 \end{aligned}$$

5.
$$\begin{aligned} \Pr(-1.59 \le X \le 0.73) &= \Pr(X \le 0.73) \\ &\quad - \Pr(X \le -1.59) \\ &= \Phi(0.73) - [1 - \Phi(1.59)] \\ &= 0.7673 - [1 - 0.9441] \\ &= 0.7114 \end{aligned}$$

Endast $N(0;1)$, den normerade (standardiserade) normalfördelningen finns i tabell.

Om variabeln $X$ är $N(\mu;\sigma)$, så är den standardiserade variablen

$$Z = \frac{X-\mu}{\sigma} \sim N(0;1).$$

Bevis:

$$\begin{aligned} E(Z) &= E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}E(X-\mu) \\ &= \frac{1}{\sigma}[E(X)-\mu] = \frac{1}{\sigma}(\mu-\mu) = 0 \\ V(Z) &= V\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}V(X-\mu) \\ &= \frac{1}{\sigma^2}[V(X)+V(-\mu)] \\ &= \frac{1}{\sigma^2}[V(X)+0] = \frac{V(X)}{\sigma^2} = \frac{\sigma^2}{\sigma^2} = 1 \end{aligned}$$

Följande sannolikhet beräknas alltså som

$$\begin{aligned} \Pr(X \le a) &= \Pr\left(\frac{X-\mu}{\sigma} \le \frac{a-\mu}{\sigma}\right) \\ &= \Pr\left(Z \le \frac{a-\mu}{\sigma}\right) = \Phi\left(\frac{a-\mu}{\sigma}\right) \end{aligned}$$

Exempel:

Variabeln $X$ är $N(170;10)$. Beräkna följande sannolikheter.

1.

$$\begin{aligned} \Pr(X \le 190) &= \Pr\left(\frac{X-170}{10} \le \frac{190-170}{10}\right) \\ &= \Pr(Z \le 2) = \Phi(2) = 0.97725 \end{aligned}$$

2.

$$\begin{aligned} \Pr(X \le 160) &= \Pr\left(\frac{X-170}{10} \le \frac{160-170}{10}\right) \\ &= \Pr(Z \le -1) = \Phi(-1) \\ &= 1-\Phi(1) = 1-0.8413 \\ &= 0.1587 \end{aligned}$$

3.

$$\begin{aligned} \Pr(160 \le X \le 190) &= \Pr(X \le 190) \\ &\quad - \Pr(X \le 160) \\ &= 0.97725 - 0.1587 \\ &= 0.81855 \end{aligned}$$

4.

$$\begin{aligned} \Pr(X \ge 190) &= 1-\Pr(X \le 190) \\ &= 1-0.97725 = 0.0228 \end{aligned}$$

5. Bestäm talet $a$ så att $\Pr(X \ge a) = 0.05$. Enligt tabell 3a (eller 3b) är $\Phi(1.64) \approx 0.95$ vilket medför att $\Pr(Z \ge 1.64) = 0.05$.

$$\begin{aligned} \Pr(X \ge a) &= \Pr\left(\frac{X-\mu}{\sigma} \ge \frac{a-\mu}{\sigma}\right) \\ &= \Pr\left(Z \ge \frac{a-\mu}{\sigma}\right) = 0.05 \\ &\Rightarrow \frac{a-\mu}{\sigma} = 1.64 \\ \frac{a-170}{10} &= 1.64 \\ a &= 1.64 \cdot 10 + 170 = 186.4 \end{aligned}$$