Föreläsning 3

Betingad Sannolikhet

Definition Låt E och F vara två händelser. Antag att $P(F) > 0$. Sannolikheten för E, betingat av F, betecknas med $P(E|F)$ och definieras som

$$P(E|F) = \frac{P(E \cap F)}{P(F)}\,.$$

Exempel Kast med en röd och en vit tärning.

$$A = \{\text{ögonsumma högst } 4\}$$

$$B_k = \{\text{vita tärningen visar } k \text{ ögon }\}$$

Observera: $P(B_k|A) = 0$ om $k \geq 4$ (för summan ≤ 4)

- \# gynnsamma utfall för $A \cap B_k$ är $4 - k$: $(v, r) = (1, 3), (1, 2), (1, 1)$ för $k = 1$, $(v, r) = (2, 2), (2, 1)$ för $k = 2$, $(v, r) = (3, 1)$ för $k = 3$,
- \# gynnsamma utfall för A är 6 $(v, r) = (1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (3, 1)$,
- \# möjliga utfall är 36

Detta ger för $k < 4$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{(4-k)/36}{6/36} \begin{cases} \frac{1}{2} & \text{om } k = 1 \\ \frac{1}{3} & \text{om } k = 2 \\ \frac{1}{6} & \text{om } k = 3 \end{cases}\,.$$

Egenskaper

Sats (a) (lagen om total sannolikhet). Om $F_1, \ldots, F_n$ är disjunkta händelser, har positiv sannolikhet och uppfyller hela utfallsrummet (dvs. $\bigcup_{i=1}^{n} F_i = S$), så gäller för varje händelse $E \subseteq F$

$$P(E) = \sum_{i=1}^{n} P(E|F_i) \cdot P(F_i)\,.$$

(b) (Bayes' formel) Under samma villkor som i (a), men $P(E) > 0$, så gäller för alla $j = 1, \ldots, n$.

$$\begin{aligned} P(F_j|E) &= \frac{P(E|F_j) \cdot P(F_j)}{P(E)} \quad \text{(kort version)} \\ &= \frac{P(E|F_j) \cdot P(F_j)}{\sum_{i=1}^{n} P(E|F_i) \cdot P(F_i)} \quad \text{(lång version)}\,. \end{aligned}$$

(c) (Multiplikationsregeln) Om $E_1 \ldots, E_n$ är händelser med $P(E_1 \cap \ldots \cap E_{n-1}) > 0$, så gäller

$$P(E_1 \cap \ldots \cap E_n) = P(E_1) \cdot P(E_2|E_1) \cdot P(E_3|E_1 \cap E_2) \cdot \ldots \cdot P(E_n|E_1 \cap \ldots \cap E_{n-1})\,.$$

Bevis (a)

$$\begin{aligned}
\sum_{i=1}^{n} P(E|F_i) \cdot P(F_i) &= \sum_{i=1}^{n} \frac{P(E \cap F_i)}{P(F_i)} \cdot P(F_i) \quad \text{(Definition betingad sannolikhet)} \\
&= \sum_{i=1}^{n} P(E \cap F_i) = P\left(\bigcup_{i=1}^{n} E \cap F_i\right) \quad (E \cap F_i \text{ disjunkta, axiom (iii)}) \\
&= P\left(E \cap \left(\bigcup_{i=1}^{n} F_i\right)\right) \quad \text{(distributiva lagen)} \\
&= P(E \cap S) = P(E) \quad (\text{förutsättning } \textstyle\bigcup_{i=1}^{n} F_i = S)\ .
\end{aligned}$$

(b)

$$\begin{aligned}
P(F_j|E) &= \frac{P(E \cap F_j)}{P(E)} \\
&= \frac{\left(P(E \cap F_j)/P(F_j)\right) \cdot P(F_j)}{P(E)} \\
&= \frac{P(E|F_j) \cdot P(F_j)}{P(E)} \quad \text{(kort version)}\,.
\end{aligned}$$

Den långa versionen följer från (a)-delen.

(c)

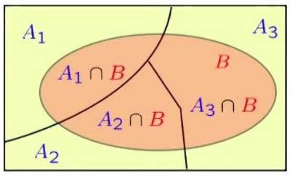
$$\begin{aligned}
&P(E_1) \cdot P(E_2|E_1) \cdot P(E_3|E_1 \cap E_2) \cdot \ldots \cdot P(E_n|E_1 \cap \ldots \cap E_{n-1}) \\
&\quad = P(E_1) \cdot \frac{P(E_1 \cap E_2)}{P(E_1)} \cdot \frac{P(E_1 \cap E_2 \cap E_3)}{P(E_1 \cap E_2)} \cdot \ldots \cdot \frac{P(E_1 \cap \ldots \cap E_{n-1} \cap E_n)}{P(E_1 \cap \ldots \cap E_{n-1})} \\
&\quad = P(E_1 \cap \ldots \cap E_n) \quad \text{(teleskopisk produkt)}\,.
\end{aligned}$$

Exempel I en fabrik tillverkas 25% av enheterna vid maskin 1, 35% vid maskin 2 och 40% vid maskin 3. Av produktionen är respektive 1%, 2% och 3% defekt. Man blandar enheterna och sänder dem till kunderna.

(a) Hur stor är sannolikheten att en slumpmässigt vald enhet är felaktig?

(b) Antag att en kund påträffar en felaktig enhet. Hur stor är sannolikheten att den har tillverkats vid maskin 1.

Svar: (a) Låt B beteckna händelsen {enhet är felaktig} och låt A_i beteckna händelsen {enhet har tillverkats vid maskin i}, $i = 1, 2, 3$.



Google Bild

Enligt lagen om "total sannolikhet" gäller det att

$$\begin{aligned} P(B) &= P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + P(B|A_3) \cdot P(A_3) \\ &= 0.01 \cdot 0.25 + 0.02 \cdot 0.35 + 0.03 \cdot 0.40 = 0.0215\,. \end{aligned}$$

(b) Bayes' formel medför att

$$P(A_1|B) = \frac{P(B|A_1) \cdot P(A_1)}{P(B)} = \frac{0.01 \cdot 0.25}{0.0215} = 0.116\,.$$

Oberoende händelser

Intuitivt är två händelser E och F oberoende om inträffandet av F inte ger någon information om huruvida E inträffar eller ej. I formler betyder detta:

$$P(E|F) = P(E) \quad \text{om} \quad P(F) > 0\,.$$

Ekvivalent med detta är

$$\frac{P(E \cap F)}{P(F)} = P(E) \quad \text{om} \quad P(F) > 0\,.$$

som också kan skrivas som

$$P(E \cap F) = P(E) \cdot P(F) \quad \text{om} \quad P(F) > 0\,.$$

Denna observation är anledning till

Definition Två händelser A och B kallas *oberoende* om

$$P(A \cap B) = P(A) \cdot P(B).$$

Om man har tre eller mer händelser blir definitionen mycket mer sofistikerad.

Definition Tre händelser A, B, C kallas *oberoende* om

$$\begin{aligned} P(A \cap B) &= P(A) \cdot P(B), \\ P(A \cap C) &= P(A) \cdot P(C), \\ P(B \cap C) &= P(B) \cdot P(C) \end{aligned}$$

och

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C).$$

Exempel Endast $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ räcker inte, vilket inses om man tar $A = B$ och $C = \emptyset$. I så fall är A, B, C inte oberoende.

Exempel Inte heller räcker parvis oberoende: Kast med röd och vit tärning.

- $A = \{$ vita tärningen visar jämnt antal ögon$\}$
- $B = \{$ röda tärningen visar jämnt antal ögon$\}$
- $C = \{$ jämnt ögonsumma$\}$

Händelserna A och B är oberoende av försöksskäl. Vidare gäller

$$P(A \cap B) = P(A) \cdot P(B) = \frac{1}{4}$$

och

$$P(A) \cdot P(C) = \frac{1}{4}\,. \tag{1}$$

Dessutom gäller $A \cap C = A \cap B$ som medför

$$P(A \cap C) = P(A \cap B) = \frac{1}{4}\,. \tag{2}$$

Alltså ger (1) och (2) att A och C är oberoende. På samma sätt får man att B och C är oberoende. Men $A \cap B \Longrightarrow C$, som betyder att A, B, C inte är oberoende.

Stokastiska variabler

Då man mäter samma sak flera gånger, upptäcker man att mätvärdena varierar. Resultatet av en mätning, där slumpen inverkar, kallas *slumpvariabel* eller *stokastisk variabel* (s.v.). Matematiskt definierar man begreppet som en avbildning.

Definition En *slumpvariabel* eller *stokastisk variabel* X är en funktion

$$X : S \longrightarrow \mathbb{R}\,.$$

Exempel

- $X = pH - värdet$ i ett vattenprov (kontinuerliga resultat)
- $Y = \#emissioner$ från en radioaktiv kropp inom en sekond (diskreta resultat)

Definition (a) En stokastisk variabel sägs vara *diskret* om den bara kan anta ett ändligt eller uppräkneligt antal olika värden.
(b) För en diskret stokastisk variabel X definieras *sannolikhetsfunktionen* genom

$$p_X(k) = P(X = k)$$

för alla möjliga värden k på X.

Exempel Gör två kast med ett mynt. Låt

$$X = \text{antal kast som ger krona.}$$

X kan då anta värdena 0,1,2.

$$p_X(0) = P(X = 0) = P(\{kl, kl\}) = \frac{1}{4}$$
$$p_X(1) = P(X = 1) = P(\{kl, kr\}) + P(\{kr, kl\}) = \frac{1}{2}$$
$$p_X(2) = P(X = 2) = P(\{kr, kr\}) = \frac{1}{4}.$$

Exempel Sannolikhetsfunktionen av en viss stokastisk variabel ges av

$$p_X(k) = P(X = k) = c \cdot \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

där $\lambda > 0$. Hitta $c > 0$ sådant att $\sum_{k=0}^{\infty} p_X(k) = 1$.

Svar:

$$1 = \sum_{k=0}^{\infty} p_X(k) = \sum_{k=0}^{\infty} c \cdot \frac{\lambda^k}{k!} = c \cdot e^{\lambda} \quad \text{Taylor serie.}$$

Detta medför $c = e^{-\lambda}$.

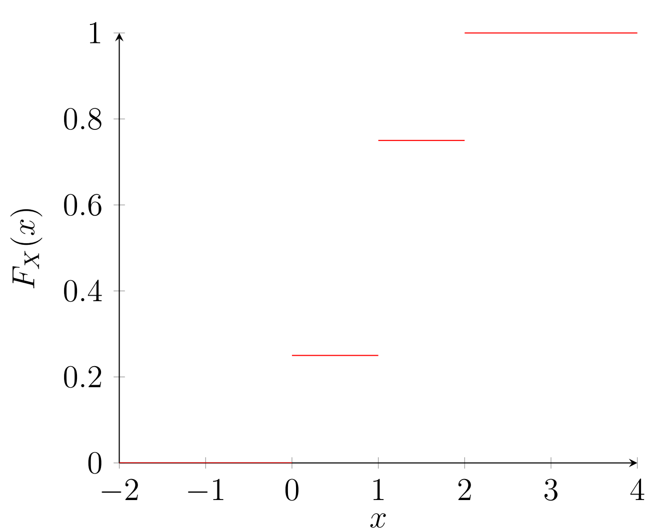
Definition

$$F_X(x) = P(X \leq x) \quad x \in \mathbb{R},$$

kallas *fördelningsfunktion* för X.

Exempel Två kast med ett mynt (se ovan).

$$F_X(x) = \begin{cases} 0 & \text{om } -\infty < x < 0 \\ \frac{1}{4} & \text{om } \quad 0 \leq x < 1 \\ \frac{3}{4} & \text{om } \quad 1 \leq x < 2 \\ 1 & \text{om } \quad 2 \leq x < \infty \end{cases}.$$



Jörg-Uwe Löbus, `jorg-uwe.lobus@liu.se`