Landmålingens fejlteori
Sandsynlighedsregning
Lektion 1

Torben Tvedebrink - tvede@math.aau.dk
http://www.math.aau.dk/∼tvede/teaching/L4

Institut for Matematiske Fag
Aalborg Universitet

27. april 2009

Landmålingens fejlteori - lidt om kurset

Kursusholder
Torben Tvedebrink
Institut for Matematiske Fag
`tvede@math.aau.dk`

Litteratur (interne L4-websider)

- Mats Rudemo, *Statistik og sandynlighedslære med biologiske anvendelser, del 1. Grundbegreber.*
- Poul Winding & Jens Møller Pedersen, *Noter i Fejlteori.*

Spisesedler & Slides http://www.math.aau.dk/∼tvede/teaching/L4

Kursusform
30 min.:Repetition,
90 min.:Opgaveregning,
90 min.:Forelæsning.

Formål

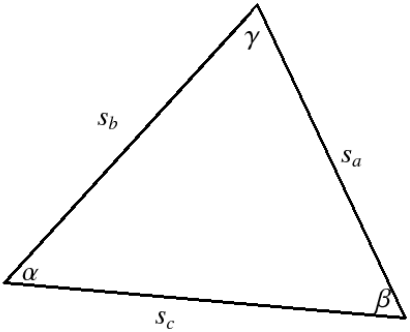
At give de studerende det nødvendige statistiske grundlag for:

- at kunne foretage kvalificerede valg af metode og instrumenter i relation til givne nøjagtighedskrav
- en vurdering af de udførte målingers kvalitet.

Konkrete problemer som behandles i L4-projektrapporterne:

- Kontrol af totalstationens og andre instrumenters præcision.
- Udjævning af måleusikkerhed på vinkelsum.
- Slutfejl på højdemålinger i forbindelse med nivellement.
- ...

Formål



Arealet af ovenstående trekant kan bestemmes vha: $A = s_b \times s_c \times \sin\alpha$.
Til at vurdere nøjagtigheden af A indgår nøjagtigheden af s_b, s_c og α.
Fejlforplantningsloven giver et approximativt udtryk for As nøjagtighed.

Terminologi

Gentag et forsøg n gange.

U	:	**udfaldsrum**, mængden af mulige udfald.
$u \in U$	:	et **udfald**.
A	:	en **hændelse**, delmængde af U, $A \subseteq U$.
		Hændelsen A indtræffer, hvis $u \in A$.
		Bemærk $\emptyset$ og U er hændelser.
n_A	:	antal gange A indtræffer.
$\frac{n_A}{n}$	:	**relativ hyppighed** for A.
$P(A)$	:	**sandsynlighed** for A.

Når n er stor, gælder der

$$P(A) \approx \frac{n_A}{n}. \tag{1}$$

De relative hyppigheder går mod $P(A)$, når n vokser.
Bemærk, de relative hyppigheder er beregnet fra data, dvs. de er **observeret værdier**.

Eksempel:

Kast med symmetrisk terning. Kast $n = 20$ gange, og observer $\#$ øjne.
Udfaldsrummet er

$$U = \{1, 2, 3, 4, 5, 6\}.$$

Observer flg. **udfald**

$$1, 3, 3, 6, 5 \quad 5, 4, 1, 2, 3 \quad 4, 5, 4, 3, 4 \quad 6, 6, 2, 4, 5.$$

Lad **hændelserne** A og B være

A : "$\#$ øjne er lige", $A = \{2, 4, 6\}$.

B : "$\#$ øjne er mindre end eller lig med 4", $B = \{1, 2, 3, 4\}$.

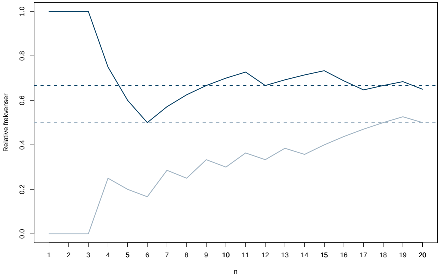
De **relative frekvenser** er for forskellige n:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$\frac{n_A}{n}$	$\frac{0}{1}$	$\frac{0}{2}$	$\frac{0}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{2}{7}$	$\frac{2}{8}$	$\frac{3}{9}$	$\frac{3}{10}$	$\frac{4}{11}$	$\frac{4}{12}$	$\frac{5}{13}$	$\frac{5}{14}$	$\frac{6}{15}$	$\frac{7}{16}$	$\frac{8}{17}$	$\frac{9}{18}$	$\frac{10}{19}$	$\frac{10}{20}$
$\frac{n_B}{n}$	$\frac{1}{1}$	$\frac{2}{2}$	$\frac{3}{3}$	$\frac{3}{4}$	$\frac{3}{5}$	$\frac{3}{6}$	$\frac{4}{7}$	$\frac{5}{8}$	$\frac{6}{9}$	$\frac{7}{10}$	$\frac{8}{11}$	$\frac{8}{12}$	$\frac{9}{13}$	$\frac{10}{14}$	$\frac{11}{15}$	$\frac{11}{16}$	$\frac{11}{17}$	$\frac{12}{18}$	$\frac{13}{19}$	$\frac{13}{20}$

Eksempel - fortsat

Dvs.

$$P(A) \approx \frac{10}{20} = 0.5 \text{ og } P(B) \approx \frac{13}{20} = 0.65.$$



Fra symmetri argumenter ses, at **sandsynlighederne** er

$$P(A) = 0.5 \text{ og } P(B) = \frac{4}{6} \approx 0.67.$$

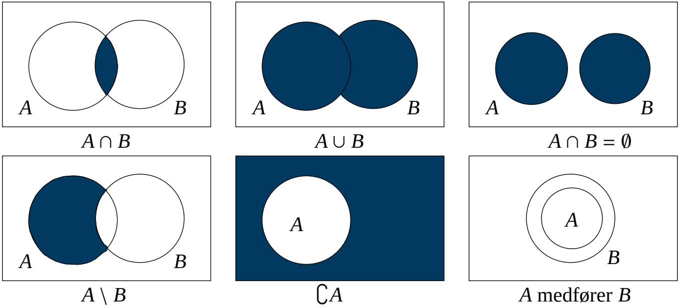
Dvs. $P(A)$ er stabiliseret, og $P(B)$ er næsten stabiliseret.

Hændelser

Lad A og B være to hændelser, da betegner

$A \cap B$: hændelsen **både** A og B indtræffer,
$A \cup B$: hændelsen **enten** A, **eller** B (incl. **både og**) indtræffer,
$A \setminus B$: hændelsen A, **men ikke** B, indtræffer,
$\complement A = U \setminus A$: hændelsen A indtræffer **ikke**, A **komplementær**.

Hvis $A \cap B = \emptyset$, så siges A og B at være **uforenlige**.
Hvis $A \subseteq B$, dvs. $A \cap B = A$, siges A at **medføre** B.



Hændelser - eksempel fortsat

Vi har

$$U = \{1,2,3,4,5,6\},\; A = \{2,4,6\} \text{ og } B = \{1,2,3,4\}.$$

Der gælder, at

$$\begin{aligned} A \cap B &= \{2,4\}, \\ A \cup B &= \{1,2,3,4,6\}, \\ A \setminus B &= \{6\}, \\ \complement A &= \{1,3,5\}. \end{aligned}$$

Lad C være hændelsen, at "# øjne er større end 4", dvs.

$$C = \{5,6\} \quad \text{og} \quad C \text{ og } B \text{ er } \textbf{uforenlige}.$$

Lad D være hændelsen, at "# øjne er mindre end 3", dvs.

$$D = \{1,2\} \quad \text{og} \quad D \textbf{ medfører } B.$$

Definition 1

Lad U være udfaldsrummet for et forsøg. En funktion P, som til hver hændelse $A \subseteq U$ tilordner et reelt tal $P(A)$, således at

1. $0 \leq P(A) \leq 1$,
2. $P(U) = 1$,
3. $P(A \cup B) = P(A) + P(B)$, hvis $A \cap B = \emptyset$,

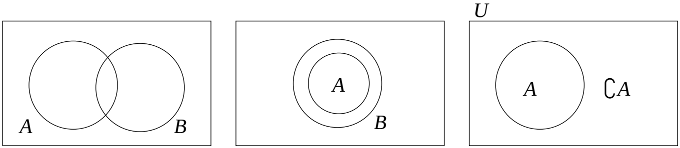
kaldes en **sandsynlighedsfordeling**.

$P(A)$ kaldes **sandsynligheden** for hændelsen A.

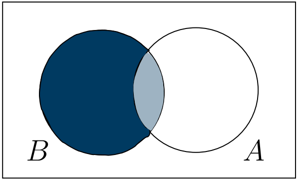
Sætning 1

Lad A og B være to hændelser. Da gælder:

4. $P(B \setminus A) = P(B) - P(A \cap B)$
5. $P(A) \leq P(B)$, hvis $A \subseteq B$
6. $P(\complement A) = 1 - P(A)$
7. $P(\emptyset) = 0$
8. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$



Bevis for Sætning 1*



Vi beviser 4. således:
Idet $B = (A \cap B) \cup (B \setminus A)$ og $(A \cap B) \cap (B \setminus A) = \emptyset$
Dermed har vi fra 3. at:
$P(B) = P(A \cap B) + P(B \setminus A) \quad \Leftrightarrow \quad P(B \setminus A) = P(B) - P(A \cap B)$

Tilsvarende kan 8. vises ved:
Da $A \cup B = A \cup (B \setminus A)$ og $A \cap (B \setminus A) = \emptyset$.
Igen anvendes 3. til at: $P(A \cup B) = P(A) + P(B \setminus A)$.
Anvendes nu 4. får vi: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Sætning 2

Sætning 2 generalisering af Definition 1 pkt. 3
Hvis $A_1, A_2, \ldots, A_k$ er **parvis uforenlige**, så er

$$P(\cup_{i=1}^{k} A_i) = P(A_1) + \cdots + P(A_k) = \sum_{i=1}^{k} P(A_i).$$

Specielt hvis $A_i = \{u_i\}$, og $A = \{u_1, u_2, \ldots, u_k\}$, så er

$$P(A) = P(u_1) + \cdots + P(u_k) = \sum_{i=1}^{k} P(u_i).$$

Ligefordelingen - Definition

Lad $U = \{u_1, u_2, \ldots, u_n\}$ bestå af n lige sandsynlige udfald

$$P(u_i) = \frac{1}{n} \quad \text{for} \quad i = 1, 2, \ldots, n.$$

Sandsynlighedsfordelingen P kaldes så en **ligefordeling**.
Lad $A \subseteq U$ være en hændelse, og lad $n(A)$ betegne antallet af elementer i A. Da er

$$P(A) = \sum_{i=1}^{n(A)} P(u_i) = \sum_{i=1}^{n(A)} \frac{1}{n} = \frac{n(A)}{n} = \frac{n(A)}{n(U)}.$$

Kende antal elementer i en mængde.
$\Downarrow$
Kombinatorik (baseret på **multiplikationsprincippet**).

Ligefordelingen - eksempel

Yatzy - kast med 5 symmetriske terninger. Antal mulige udfald

$$n(U) = 6 \cdot 6 \cdot 6 \cdot 6 \cdot 6 = 6^5.$$

Alle udfald er lige sandsynlige, dvs. **ligefordelingen**

$$P = \frac{1}{6^5}.$$

Hændelsen

$$A\text{: “Yatzy” (fem ens).}$$

Der gælder

$$n(A) = 6 \cdot 1 \cdot 1 \cdot 1 \cdot 1 = 6.$$

Sandsynligheden for “Yatzy” er

$$P(\text{“Yatzy”}) = \frac{n(A)}{n(U)} = \frac{6}{6^5} = \frac{1}{6^4} = \frac{1}{1296} = 7.7 \cdot 10^{-4}.$$

Bemærk, antal elementer i U og A er fundet vha. **multiplikationsprincippet**.

Multiplikationsprincippet

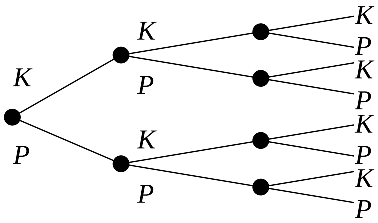
Sætning 3

Udfør k uafhængige operationer i en bestemt rækkefølge, hvor operation i kan udføres på n_i måder, $i = 1, 2, \ldots, k$.
Total antal måder de k operationer kan udføres på er

$$n_1 \cdot n_2 \cdot \cdots \cdot n_k.$$

Tælletræ:

Kast 3 gange med en mønt og observer plat P eller krone K.



ialt $2 \cdot 2 \cdot 2 = 8$ måder.

Sætning 5

Antal mulige måder/kombinationer at udtage k elementer fra en mængde med n elementer:

	uden tilbagelægning	med tilbagelægning
ordnet	$(n)_k = \frac{n!}{(n-k)!}$	n^k
uordnet	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$	-

Kombinatorik

Lad $U = \{a, b, c, d\}$ være udfaldsrummet. Da kan 2 elementer udvælges på følgende måde under forskellige udvælgelseskriterier:

	Uden tilbagelægning	Med tilbagelægning
Med hensyn til rækkefølge	(ab) (ac) (ad) (ba) (bc) (bd) (ca) (cb) (cd) (da) (db) (dc)	(aa) (ab) (ac) (ad) (ba) (bb) (bc) (bd) (ca) (cb) (cc) (cd) (ad) (db) (dc) (dd)
Uden hensyn til rækkefølge	(ab) (ac) (ad) (bc) (bc) (cd)	

Kombinatorik

Lad $U = \{a, b, c, d\}$ være udfaldsrummet. Da kan 2 elementer udvælges på følgende måde under forskellige udvælgelseskriterier:

	Uden tilbagelægning	Med tilbagelægning
Med hensyn til rækkefølge	(ab) (ac) (ad) (ba) (bc) (bd) (ca) (cb) (cd) (da) (db) (dc) $(n)_k = (4)_2 = \frac{4!}{2!} = 12$	(aa) (ab) (ac) (ad) (ba) (bb) (bc) (bd) (ca) (cb) (cc) (cd) (ad) (db) (dc) (dd) $n^k = 4^2 = 16$
Uden hensyn til rækkefølge	(ab) (ac) (ad) (bc) (bc) (cd) $\binom{n}{k} = \binom{4}{2} = \frac{4!}{2!(4-2)!} = 6$	

Eksempel

Kvalitetskontrol: Betragt et vareparti med 20 enheder, hvoraf 8 defekte.
Udtag 6 enheder **uordnet og uden tilbagelægning**.
Hændelsen

$$A : \text{ingen defekte enheder i stikprøven.}$$

Antal mulige stikprøver:

$$n(U) = \binom{20}{6}.$$

Antal mulige stikprøver uden defekte enheder:

$$n(A) = \binom{12}{6}.$$

Sandsynligheden for A

$$P(A) = \frac{\binom{12}{6}}{\binom{20}{6}} = \frac{12! \cdot 6! \cdot 14!}{6! \cdot 6! \cdot 20!} = \frac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7}{20 \cdot 19 \cdot 18 \cdot 17 \cdot 16 \cdot 15} = \frac{77}{3230} \approx 0.024.$$

Eksempel - fortsat

Hændelsen

$$B : \text{netop 2 defekte enheder i stikprøven.}$$

Antal mulige stikprøver med netop 2 defekte enheder:

$$n(B) = \binom{12}{4}\binom{8}{2}.$$

Sandsynligheden for B

$$P(B) = \frac{\binom{12}{4}\binom{8}{2}}{\binom{20}{6}} = \frac{12! \cdot 8! \cdot 6! \cdot 14!}{4! \cdot 8! \cdot 2! \cdot 6! \cdot 20!} \approx 0.3576.$$

Definition 2

Lad A og B være to hændelser, hvor $P(B) > 0$.
Den betingede sandsynlighed for A givet B er

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Bemærk

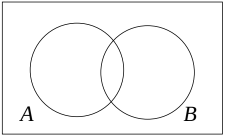
Betinget sandsynlighed er en sandsynlighed defineret på et nyt og mindre udfaldsrum. Dvs. almindelige regneregler gælder:

$$\begin{aligned} &P(A|B) + P(\complement A|B) = 1 \\ &P(A \cup B|C) = P(A|C) + P(B|C) - P(A \cap B|C) \\ &\vdots \end{aligned}$$

Bayes formel

Beskriver forholdet mellem $P(A|B)$ og $P(B|A)$

$$
\begin{aligned}
P(A \cap B) &= P(A|B)P(B) = P(B|A)P(A) \\
&\Downarrow \\
P(A|B) &= \frac{P(B|A)P(A)}{P(B)}
\end{aligned}
$$



Omskrives

$$
\begin{aligned}
P(A|B) &= \frac{P(B|A)P(A)}{P(B \cap A) + P(B \cap \complement A)} \\
&= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\complement A)P(\complement A)}
\end{aligned}
$$

Eksempel

Køns-fordeling i gymnasium

	kvinder	mænd
mat	45	55
sproglig	75	20

Betingede sandsynligheder

$$P(\text{mat}|\text{mand}) = \frac{P(\text{mat\&mand})}{P(\text{mand})} = \frac{55/195}{(55+20)/195} = 73.33\%$$

$$P(\text{mat}|\text{kvinde}) = \frac{P(\text{mat\&kvinde})}{P(\text{kvinde})} = \frac{45/195}{(45+75)/195} = 37.5\%$$

Eksempel

Lungesygdomme og rygere:
Iflg. "The American Lung Association" lider 7% af befolkningen af en lungesygdom, og 90% af disse er rygere. For folk uden lungesydomme er 25.3% rygere.
Lad hændelserne A og B være givet ved

$$\begin{array}{lcl} A & : & \text{personen har en lungesygdom.} \\ B & : & \text{personen er ryger.} \end{array}$$

Vi har

$$\begin{aligned} P(A) &= 0.07, \\ P(B|A) &= 0.9, \\ P(B|\complement A) &= 0.253. \end{aligned}$$

Eksempel - fortsat

Interessant at finde
$P(A|B) = P(\text{personen har en lungesygdom}|\text{personen er ryger})$.
Vi bruger **Bayes formel**

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\complement A)P(\complement A)} \\ &= \frac{0.900 \cdot 0.070}{0.900 \cdot 0.070 + 0.253 \cdot (1 - 0.070)} \\ &= 0.211. \end{aligned}$$

Bemærk

$$P(\complement A|B) = 1 - P(A|B) = 0.789.$$

Dvs. $P(\text{personen } \textbf{IKKE} \text{ har en lungesygdom}|\text{personen er ryger}) = 78.9\%$.

Uafhængige hændelser

Definition 3

To hændelser A og B siges at være **uafhængige**, hvis

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

Sætning 6

Lad A og B være to hændelser, hvor $P(A) > 0$ og $P(B) > 0$.
Flg. udsagn ækvivalente.

- A og B er uafhængige,
- $P(A|B) = P(A)$,
- $P(B|A) = P(B)$.

Eksempel - fortsat

Lungesygdomme og rygere:
Vi fandt

$$\begin{aligned} P(A) &= 0.07, \\ P(A|B) &= 0.211. \end{aligned}$$

Dvs. rygning og lungesygdomme er **afhængige**.