

Landmålingens fejlteori

Repetition - Diskrete stokastiske variable

Lektion 3

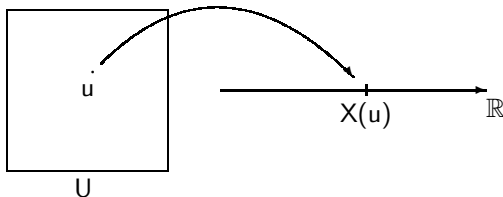
Torben Tvedebrink - tvede@math.aau.dk
<http://www.math.aau.dk/~tvede/teaching/L4>

Institut for Matematiske Fag
Aalborg Universitet

30. april 2009

Definition

En reel funktion defineret på et udfaldsrum (med sandsynlighedsfordeling) kaldes en **stokastisk variabel**.



Eksempler

Forsøg	Stokastisk variabel	Type
kast med terning	$\#$ øjne	diskret
kast med to terninger	$\sum$ øjne	diskret
familie i Danmark	$\#$ børn	diskret
1. gangsfødende	alder	diskret
fødsel	fødselsvægt	kontinuert
hustande i Danmark	indkomst	kontinuert
mænd i Danmark	højde	kontinuert

Diskret stokastisk variabel: antal værdier endelig eller tællelig.

Kontinuert stokastisk variabel: antager værdier i en delmængde af reelle tal.

Definition 1

Lad U være et udfaldsrum med sandsynlighedsfordeling P . Lad X være en funktion givet ved $X : U \mapsto V$, hvor V er endelig eller tællelig delmængde af $\mathbb{R}$. Funktionen X kaldes en **diskret stokastisk variabel**.

Definition

En funktion $f : V \mapsto \mathbb{R}$ givet ved

$$f(x) = P(X = x), \quad x \in V,$$

kaldes en **sandsynlighedsfunktion** for X .

Der gælder

$$f(x) \geq 0, \quad x \in V,$$

da $f(x)$ er en sandsynlighed, og

$$\sum_{x \in V} f(x) = 1,$$

$$\text{da } \sum_{x \in V} f(x) = \sum_{x \in V} P(X = x) = P(X \in V) = 1.$$

Definition 2

Middelværdien af en diskret stokastisk variabel X er

$$\mu = \mathbb{E}(X) = \sum_{x \in V} xf(x) < \infty,$$

- også kaldet **forventede værdi**.

Definition 2

Middelværdien af en diskret stokastisk variabel X er

$$\mu = \mathbb{E}(X) = \sum_{x \in V} xf(x) < \infty,$$

- også kaldet **forventede værdi**.

Definition 3

Variansen af en diskret stokastisk variabel X er

$$\sigma^2 = \mathbb{V}(X) = \mathbb{E}((X - \mu)^2) = \sum_{x \in V} (x - \mu)^2 f(x) < \infty,$$

hvor σ kaldes **spredning** eller **standard afvigelse** - mål for forventet variabilitet i X .

Varians formel: Der gælder $\sigma^2 = \mathbb{E}(X^2) - \mu^2$

Binomialfordeling

Forsøg med **to mulige udfald**: succes eller fiasko, hvor

$$P(\text{"succes"}) = p, \text{ hvor } 0 \leq p \leq 1.$$

Dvs.

$$P(\text{"fiasko"}) = 1 - p.$$

Per definition gælder $P(\text{"succes"}) + P(\text{"fiasko"}) = p + 1 - p = 1$.

Gentag forsøget n gange, og lad

X : antal succeser.

Sandsynlighed for x succeser

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x \in \{1, 2, \dots, n\}.$$

Binomialfordeling - fortsat

Sandsynlighedsfunktionen fremkommer ved at udregne sandsynligheden for x succeser ud af n mulige. Vi antager hvert forsøg er uafhængige af de øvrige.

De x succeser kan fx. forekomme ved $\overbrace{s, s, \dots, s}^x, \overbrace{f, f, \dots, f}^{n-x}$, men alle øvrige måder x kan vælges blandt n er gyldige.

$$\begin{aligned} P("x \text{ succeser}") &= \overbrace{p \times p \times \dots \times p}^x \times \overbrace{(1-p) \times (1-p) \times \dots \times (1-p)}^{n-x} \\ &\quad + \dots + \\ &\quad \overbrace{(1-p) \times (1-p) \times \dots \times (1-p)}^{n-x} \times \overbrace{p \times p \times \dots \times p}^x \\ &= \binom{n}{x} p^x (1-p)^{n-x} \end{aligned}$$

Idet der er $\binom{n}{x}$ led i summen/måder x kan vælges ud af n .

Definition 4

En stokastisk variabel X er **binomialfordelt** med parametre $n \in \mathbb{N}$ og p , hvor $0 \leq p \leq 1$, hvis sandsynlighedsfunktionen for X er

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, x \in \{0, 1, 2, \dots, n\}.$$

Notation: $X \sim b(n, p)$.

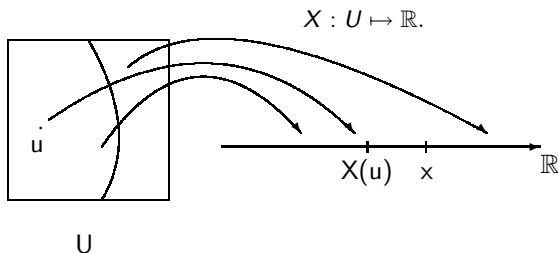
Sætning 1

Hvis $X \sim b(n, p)$, så er **middelværdi** og **varians** af X hhv.

$$\begin{aligned}\mathbb{E}(X) &= np, \\ \mathbb{V}(X) &= np(1-p).\end{aligned}$$

Fordelingsfunktion

Lad X være en stokastisk variabel (diskret eller kontinuert) givet ved



Lad x være et reelt tal, og betragt hændelsen

$$\{u \in U | X(u) \leq x\}.$$

Hvad er sandsynligheden for denne hændelse?

Definition 1

Fordelingsfunktion for en stokastisk variabel X er

$$F(x) = P(\{u \in U | X(u) \leq x\}) = P(X \leq x).$$

Bemærk $F : \mathbb{R} \mapsto [0, 1]$.

Fordelingen af en stokastisk variabel defineres ofte ud fra fordelingsfunktionen.