

# Middelværdi og spredning

## Beregninger i Maple

Vi husker fra sidst, at definitionen på en binomialfordelt stokastisk variabel er følgende.

**Definition 1.1** (Binomialfordelt stokastisk variabel). En stokastisk variabel  $X$  siges at være *binomialfordelt* med antalsparameter  $n$  og sandsynlighedsparameter  $p$  hvis sandsynlighedsfunktionen for  $X$  er givet ved

$$f(r) = P(X = r) = K(n, r) \cdot p^r \cdot (1 - p)^{n-r}.$$

Vi skriver, at  $X \sim B(n, p)$ .

Vi kan bruge følgende funktion i Maple, når vi skal bruge sandsynlighedsfunktionen for den binomialfordelte stokastiske variabel.

`binpdf(n,p,x),`

hvor  $n$  er antalsparameteren,  $p$  er sandsynlighedsparameteren og  $x$  er udfaldet, du ønsker at bestemme sandsynligheden for.

**Eksempel 1.2.** Vi ønsker i Maple at bestemme sandsynligheden for at få 2 seksere i 5 forsøg. Vi skriver i Maple

```
restart
with(Gym):
binpdf(5, 1/6, 2)
```

og får outputtet 0.16.

## Fordelingsfunktion

Vi er ofte interesserede i sandsynligheden for, om vi har fået mindre end et bestemt antal succeser. Derfor introduceres *fordelingsfunktionen*.

**Definition 1.3** (Fordelingsfunktion). Fordelingsfunktionen for en stokastisk variabel  $X$  er en funktion  $F$ , der opfylder, at

$$F(x) = P(X \leq x).$$

**Eksempel 1.4.** Vi skal bestemme sandsynligheden for at slå mindre end eller lig 2 seksere i 5 forsøg med en terning. Vi skal derfor for den binomialfordelte stokastiske variabel  $X \sim B\left(5, \frac{1}{6}\right)$  bestemme

$$P(X \leq 2).$$

Dette gøres ved at bestemme

$$P(X = 0) + P(X = 1) + P(X = 2) \approx 0.40 + 0.40 + 0.16 = 0.96,$$

men kan også gøres i Maple ved at bruge funktionen `bincdf` (for binomial cumulative distribution function).

```
restart
with(Gym):
bincdf(5, 1/6, 2)
```

og vi får outputtet 0.96.

## Middelværdi og spredning

Har vi en stokastisk variabel  $X$ , så kunne vi godt tænke os at beskrive det mest sandsynlige udfald af den stokastiske variabel. Betragter vi 1000 slag med en terning, så vil det gennemsnitlige slag være summen af alle de slåede slag delt med 1000. Lad os se på et eksempel:

**Eksempel 1.5.** Vi slår 1000 gange med en terning og får følgende resultater

1	2	3	4	5	6
157	162	165	176	173	167

Vi vil så bestemme det gennemsnitlige slag som

$$\frac{1 \cdot 157 + 2 \cdot 162 + 3 \cdot 165 + 4 \cdot 176 + 5 \cdot 173 + 6 \cdot 167}{1000} = 3,445.$$

Vi har derfor eksperimentelt bestemt den forventede værdi, når vi slår med en terning.

Vi vil bruge mere eller mindre det samme princip, når vi definerer *middelværdien for en stokastisk variabel*.

**Definition 1.6** (Middelværdi). Lad  $X$  være en (diskret) stokastisk variabel med værdimængde  $\{x_1, x_2, \dots, x_n\}$  og fordelingsfunktion  $P$ . Så defineres middelværdien  $\mathbb{E}(X)$  for en stokastisk variabel som

$$\begin{aligned}\mathbb{E}(X) &= \mu = x_1 \cdot P(X = x_1) + x_2 \cdot P(X = x_2) + \dots + x_n \cdot P(X = x_n) \\ &= \sum_{i=1}^n x_i P(X = x_i).\end{aligned}$$

Middelværdien giver os et indblik i, hvad den stokastiske variabel mest sandsynligt antager af værdi, men den siger ikke noget om, hvor meget værdierne, variablen kan antage varierer omkring middelværdien  $\mu$ . Vi vil definere variansen som den ”gennemsnitlige variation fra middelværdien”. Derfor ville det give mening at definere variansen som

$$\text{Var} = \mathbb{E}(X - \mu),$$

men dette vil tage højde for, at  $X - \mu$  nogle gange er negativ og nogle gange er positiv. Vi vil bare gerne vide, hvor langt  $X$  ligger fra  $\mu$ , så derfor definerer vi variansen på følgende vis:

**Definition 1.7** (Varians og spredning). Lad  $X$  være en (diskret) stokastisk variabel med værdimængde  $\{x_1, x_2, \dots, x_n\}$  og fordelingsfunktion  $P$ . Så defineres *variansen* for  $X$  som

$$\begin{aligned}\text{Var}(X) &= \sigma^2 = E[(X - \mu)^2] \\ &= (x_1 - \mu)^2 \cdot P(X = x_1) + \dots + (x_n - \mu)^2 \cdot P(X = x_n) \\ &= \sum_{i=1}^n (x_i - \mu)^2 \cdot P(X = x_i).\end{aligned}$$

Tallet  $\sigma = \sqrt{\text{Var}(X)}$  kaldes for spredningen af  $X$ .

**Eksempel 1.8.** Lad os fortsætte eksemplet med terningekastet. Vi vil bestemme middelværdien  $\mathbb{E}(X)$  for den stokastiske variabel  $X$ , der beskriver et terningekast. Dette bestemmes ifølge definitionen som

$$\mathbb{E}(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = 3.5,$$

hvilket er tæt på det vi eksperimentelt havde bestemt gennemsnittet til at være. Variansen er så

$$\begin{aligned}\text{Var}(X) &= (1 - 3.5)^2 \cdot \frac{1}{6} + (2 - 3.5)^2 \cdot \frac{1}{6} + (3 - 3.5)^2 \cdot \frac{1}{6} \\ &\quad + (4 - 3.5)^2 \cdot \frac{1}{6} + (5 - 3.5)^2 \cdot \frac{1}{6} + (6 - 3.5)^2 \cdot \frac{1}{6} \approx 2.92\end{aligned}$$

Spredningen er derfor

$$\sigma \approx \sqrt{2,92} \approx 1.71.$$

Det er svært at give en god fortolkning af spredningen ud over at det er den gennemsnitlige afstand et slag vil have fra middelværdien.

Hvis vi kender sandsynlighedsparametren  $p$  og antalsparametren  $n$  for en binomialfordelt stokastisk variabel, så er det muligt at bestemme middelværdien og variansen.

**Sætning 1.9.** *Lad  $X$  være en binomialfordelt stokastisk variabel med antalsparameter  $n$  og sandsynlighedsparameter  $p$ . Så er middelværdien  $\mu$  og spredningen  $\sigma$  for  $X$  givet ved*

$$\mathbb{E}(X) = \mu = n \cdot p,$$

og

$$\sqrt{\text{Var}(X)} = \sigma = \sqrt{n \cdot p \cdot (1 - p)}$$

## Opgave 1

Et lægemiddel bruges til en bestemt behandling, og gives til 10 personer. Sandsynligheden for helbredelse er 20%.

- i) Hvad er sandsynligheden for, at to personer helbredes?
- ii) Hvad er sandsynligheden for, at alle personer helbredes?
- iii) Hvad er sandsynligheden for, at mindst 4 helbredes?

## Opgave 2

I en by har 15% stemt på partiet "liste Q". 100 Personer udvælges tilfældigt og spørges "har du stemt liste Q?"

- i) Hvad er sandsynligheden for, at netop 15 personer svarer ja?
- ii) Hvad er sandsynligheden for, at ingen svarer ja?
- iii) Bestem sandsynligheden for, at mere end 20 svarer ja.

## Opgave 3

Til kurset "Introduktion til matematisk analyse" på Husum Universitet dumper 50% af deltagerne. 20 procent af dem, der har dumpet vil lyve, når de bliver spurgt, om de har dumpet kurset. 15 personer udvælges tilfældigt og spørges, om de er dumpet.

- i) Hvad er sandsynligheden for, at 3 personer svarer, at de er dumpet?
- ii) Bestem sandsynligheden for, at mere end 5 personer svarer, at de er dumpet.
- iii) Bestem sandsynligheden for, at mellem 2 og 4 personer svarer, at de er dumpet.

## Opgave 4

- i) Slå plat og krone, og lad  $X = 1$ , hvis udfaldet er plat og  $X = 0$  ellers. Hvad er middelværdien og spredningen for  $X$ ?
- ii) Slå plat og krone to gange og lad  $X$  være antallet af plat. Hvad er middelværdien og spredningen af  $X$ ?
- iii) Kast en 12-sidet terning og lad  $X$  være antallet af øjne. Hvad er middelværdien og spredningen af  $X$ ?

## Opgave 5

Fordelingen for en stokastisk variabel  $X$  kan beskrives ved følgende tabel.

$x$	-5	2	3	4	7	11
$P(X = x)$	0.1	0.2	0.1	0.4	0.05	0.15

- i) Bestem middelværdien for  $X$ .
- ii) Bestem spredningen for  $X$ .

## Opgave 6

- i) Lad  $X$  være en Bernoulli-fordelt stokastisk variabel med sandsynlighedsparameter  $p$ . Hvad er middelværdien og variansen af  $X$ ?
- ii) Forklar med ord, hvorfor det giver god mening, at middelværdien for en binomialfordelt stokastisk variabel er  $n \cdot p$ .

## Opgave 7

- i) Et lægemiddel helbreder 15% af behandlede personer. Lad  $X$  beskrive antallet af helbrede personer, når vi prøver at helbrede 10000 personer. Hvad er middelværdien og variansen af  $X$ ?
- ii) Vi slår fem gange med en terning, og lader  $X$  beskrive antallet af gange, vi slår mere end 4. Hvad er middelværdien og variansen for  $X$ ?