

# Den binomialfordelte stokastiske variabel

## Den binomialfordelte stokastiske variabel

En *binomialfordelt stokastisk variabel* er en stokastisk variabel, der beskriver sandsynligheden for et bestemt antal succeser i en række Bernoulli-eksperimenter. Dette kunne være sandsynligheden for at få krone to gange i syv forsøg eller at få to sekser i fem kast med en terning.

Vi definerer den binomialfordelte stokastiske variabel lidt anderledes og motiverer herefter definitionen. Først defineres dog, hvad der menes, når vi siger *sandsynlighedsfunktion* i forbindelse med stokastiske variable.

**Definition 1.1** (Sandsynlighedsfunktion). Lad  $X$  være en stokastisk variabel. Sandsynlighedsfunktionen for  $X$  defineres da til at være funktion  $f$ , der opfylder, at

$$f(x) = P(X = x).$$

**Definition 1.2** (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 vil nu motivere definitionen. Betragt situationen, hvor vi har udført  $n$  stokastiske eksperimenter, der hver har sandsynlighed  $p$  for succes. Desuden er de første  $r$  eksperimenter succeser og resten er fiaskoer. Sandsynligheden for at få  $r$  succeser i streg er (da eksperimenterne er uafhængige)

$$\underbrace{p \cdot p \cdots p}_{r \text{ gange}} = p^r$$

Sandsynligheden for at få  $n - r$  fiaskoer i streg er

$$\underbrace{(1 - p) \cdot (1 - p) \cdots (1 - p)}_{n-r \text{ gange}} = (1 - p)^{n-r}.$$

Skal vi først have succeserne *og* derefter fiaskoerne er sandsynligheden for dette produktet af de to tidligere sandsynligheder.

$$p^r \cdot (1 - p)^{n-r}$$

Vi er næsten i mål. Vi mangler blot følgende betragtning; vi er ligeglade med, hvornår vi får vores succeser. Vi skal derfor bestemme antallet af måder, vi kan placere vores succeser i vores  $n$  forsøg. Dette må tilsvare  $K(n, r)$ , da vi skal udtrække  $r$  pladser til vores succeser blandt de  $n$  forsøg. I alt må sandsynligheden for at få vores  $r$  succeser derfor være

$$K(n, r) \cdot p^r \cdot (1 - p)^{n-r},$$

da enhver placering af succeserne må have samme sandsynlighed.

Vi betragter et eksempel på denne argumentation.

**Eksempel 1.3.** Vi skal bestemme sandsynligheden for at slå 2 seksere i 5 forsøg. Sandsynligheden for at de to første slag er seksere er

$$\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

Sandsynligheden for, at de resterende slag ikke er seksere er

$$\frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \approx 0.58.$$

Sandsynligheden for kombinationen af succeser SSFFF er derfor

$$\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \approx \frac{1}{36} \cdot 0.58 \approx 0.016.$$

Sandsynligheden for andre kombinationer af succeser (eksempelvis SFSFF) må alle have denne sandsynlighed. Antallet af sådanne kombinationer er

$$K(5, 2) = \frac{5!}{(5-2)! \cdot 2!} = 10.$$

Sandsynligheden for at få netop 2 seksere i fem forsøg er derfor

$$\begin{aligned} P(X = 2) &= K(5, 2) \left(\frac{1}{6}\right)^2 \left(1 - \frac{1}{6}\right)^3 \\ &= K(5, 2) \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3 \\ &\approx 10 \cdot \frac{1}{36} \cdot 0.58 \\ &\approx 10 \cdot 0.016 \\ &= 0.16, \end{aligned}$$

så sandsynligheden for at få netop 2 seksere i 5 forsøg er altså omtrent 16%.

Det vil ofte være besværligt at bruge sandsynlighedsfunktionen for binomialfordelingen i hånden. I Maple hedder sandsynlighedsfunktionen for binomialfordelingen for `binpdf` (for binomial probability distribution function), og den bruges ved at skrive

$$\text{binpdf}(n, p, x),$$

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

**Eksempel 1.4.** 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 1.5** (Fordelingsfunktion). Fordelingsfunktionen for en stokastisk variabel  $X$  er en funktion  $F$ , der opfylder, at

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

**Eksempel 1.6.** 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  $X \sim B(5, \frac{1}{6})$  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.

## Opgave 1

- i) Hvad er sandsynligheden for at slå nøjagtigt fem seksere på seks slag med en terning?
- ii) Hvad er sandsynligheden for at få mindst 3 seksere?
- iii) Hvad er sandsynligheden for at få mindre en 4 nøjagtigt 2 gange?

## Opgave 2

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 3

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 4

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.