Stokastiske variable

Ordet 'stokastisk' er avledet av det greske ordet for å sikte. Når man sikter på noe, vet man ikke helt hvor man treffer. En stokastisk variabel X er en funksjon som tilordner en tallverdi x til hvert utfall A

$$X(A) = x.$$

Eksempel. Vi triller to terninger. La

$$A_{ij} = i \text{ øyne på den ene, } j \text{ øyne på den andre}$$

Det er nå trettiseks utfall som alle har sannsynlighet $\frac{1}{36}$. Et eksempel på en stokastisk variabel er summen av øynene på terningen

$$X(A_{ij}) = i + j. \qquad \triangle$$

Det er vanlig å enten operere med et noe diffust skille mellom utfall og stokastisk variabel. To grunner er

- De kan de være til forveksling like.
- Vi kan bruke den stokastiske variabelen til å definere utfall.

Eksempel. Vi triller to terninger igjen. La A_{ij} og X være som i forrige eksempel. Vi kan sette opp en sannsynlighetsfordeling på utfallsrommet definert av verdiene til den stokastiske variabelen.

x	$P(X = x)$
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

Du bør sjekke at $2 \leq X \leq 12$ definerer et utfallsrom og at tabellen er en sannsynlighetsfordeling. $\triangle$

Sannsynligheter på formen $P(X < a)$ kalles kumulative sannsynligheter.

Eksempel. Vi triller to terninger igjen for tredje gang idag. De kumulative sannsynlighetene er

x	$P(X \leq x)$
2	1/36
3	3/36
4	6/36
5	10/36
6	15/36
7	21/36
8	26/36
9	30/36
10	33/36
11	35/36
12	1

$\triangle$

Forventningsverdi

Forventningsverdi er, som navnet tilsier, det man forventer skal skje i tilfeldige forsøk. Anta at vi har et tilfeldig forsøk med n utfall A_i, sannsynlighetsfunksjon P, og stokastisk variabel X. Da er

$$\mu = E(X) = \sum_{i=1}^{n} X(A_i)P(A_i).$$

Hvis vi bruker den stokastiske variabelens verdier $x_i = X(A_i)$ til å definere utfallene, blir formelen

$$\mu = E(X) = \sum_{i=1}^{n} x_i P(X = x_i).$$

Eksempel. Vi triller nok en gang to terninger. Forventet antall øyne er:

$$\begin{aligned} E(X) &= \sum_{x=2}^{12} xP(X = x) \\ &= 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + \cdots + 12 \cdot \frac{1}{36} \\ &= 7 \end{aligned} \qquad \triangle$$

Noen ganger er det nødvendig å definere forskjellige stokastiske variable på et og samme utfallsrom.

Eksempel. Dusjanlegg på campingplass: en person kan dusje om gangen, og det kan stå inntil 3 personer i kø utenfor dusjen. Vi definerer følgende utfall:

- A_0 = ingen folk i dusjen
- A_1 = en person i dusjen
- A_2 = en person i dusjen og en i kø
- A_3 = en person i dusjen og to i kø
- A_4 = en person i dusjen og tre i kø

med sannsynlighet

Utfall	A_0	A_1	A_2	A_3	A_4
$P(A_i)$	15/30	8/30	4/30	2/30	1/30

I dette eksemplet skal vi finne forventet antall personer i dusjanlegget og forventet antall personer i kø. Vi setter først opp en stokastisk funksjon som måler antall personer i dusjanlegget:

$$X(A_i) = i,$$

Forventet antall personer i anlegget er:

$$\begin{aligned} E(X) &= \sum_{i=0}^{4} X(A_i)P(A_i) \\ &= 0 \cdot \frac{15}{30} + 1 \cdot \frac{8}{30} + 2 \cdot \frac{4}{30} + 3 \cdot \frac{2}{30} + 4 \cdot \frac{1}{30} = \frac{26}{30} \end{aligned}$$

Vi definerer så en ny stokastisk variabel Y, som måler antall personer i kø, gitt ved $Y(A_0) = Y(A_1) = 0$,

$Y(A_2) = 1$, $Y(A_2) = 3$ og $Y(A_4) = 3$. Forventningen til denne variabelen blir

$$\begin{aligned} E(Y) &= \sum_{i=0}^{4} Y(A_i)P(A_i) \\ &= 0 \cdot \frac{15}{30} + 0 \cdot \frac{8}{30} + 1 \cdot \frac{4}{30} + 2 \cdot \frac{2}{30} + 3 \cdot \frac{1}{30} \\ &= \frac{11}{30}. \end{aligned}$$

$\triangle$

Eksempel. Et annet vittig eksempel på forventning kalles St. Petersburgparadokset. Anta at du spiller et spill der man kaster mynt og kron. Man kaster til man får den første kron, og dersom man får den første kron i kast nr k, får man premien 2^k kr. Vi lar A_k være første kron i kast k og $X(A_k) = 2^k$ være premien i dette utfallet. Sannsynlighetstabellen blir

A_i	$X(A_i)$	$P(A_i)$
K	2	1/2
MK	4	1/4
MMK	8	1/8
MMMK	16	1/16
$\vdots$	$\vdots$	$\vdots$
M$\cdots$ M K	2^k	$1/2^k$

Vi beregner forventningen

$$E(X) = \sum_{k=1}^{\infty} 2^k \cdot \frac{1}{2^k} = \sum_{k=1}^{\infty} 1 = 1+1+1+1+\cdots = \infty.$$

Som du ser er det lurt å spille dette spillet uansett hva prisen er for å delta. $\triangle$

Varians og standardavvik

Anta igjen at vi har et tilfeldig forsøk med n utfall A_i, sannsynlighetsfunksjon P, og stokastisk variabel X. Variansen til en stokastisk funksjon er

$$Var(X) = \sum_{i=1}^{n} (X(A_i) - E(X))^2 P(A_i).$$

Vi kan nok en gang bruke den stokastiske variabelens verdier x_i som utfallsrom, og skrive

$$Var(X) = \sum_{i=1}^{n} (x_i - \mu)^2 P(X = x_i).$$

Standardavviket er kvadratroten av variansen.

$$\sigma = \sqrt{Var(X)}.$$

Standardavviket σ er mest relevant i praktisk bruk, for det har samme benevning som X og μ. Det er derimot lettere å skrive opp regneregler for variansen enn for standardavviket.

Eksempel. Vi ser på variabelen X for dusjanlegget. Variansen blir (husk at $X(A_i) = i$)

$$\begin{aligned} Var(X) &= \sum_{i=0}^{4} (i - \mu)^2 P(X = i) \\ &= (0 - \frac{26}{30})^2 \cdot \frac{15}{30} + (1 - \frac{26}{30})^2 \cdot \frac{8}{30} \\ &+ (2 - \frac{26}{30})^2 \cdot \frac{4}{30} + (3 - \frac{26}{30})^2 \cdot \frac{2}{30} \\ &+ (4 - \frac{26}{30})^2 \cdot \frac{1}{30} \\ &= 1.26 \end{aligned}$$

og standardavviket

$$\sigma = \sqrt{Var(X)} = 1.12.$$

$\triangle$

Eksempel. Forvnentet antall øyne når du kaster en terning er

$$E(X) = \frac{1}{6}(1+2+3+4+5+6) = 3.5.$$

Standardavviket er

$$\sigma = \sqrt{\frac{1}{6}\sum_{i=1}^{6}[i - \mu]^2} = 1.71.$$

$\triangle$

Binomisk fordeling

Anta at du har n innbyrdes uavhengige og helt identiske forsøk. Hvert forsøk har to utfall med sannsynligheter p og $1 - p$, kalt suksess og ikke suksess. La $X = r$, $0 \leq r \leq n$ være en stokastisk variabel som teller antall suksesser på de n forsøkene. Merk at verdiene til X definerer et gyldig utfallsrom. Sannsynligheten for r suksesser er

$$P(X = r) = \binom{n}{r} p^r (1-p)^{n-r}.$$

Eksempel. Av og til når du skal ta av vinterdekkene sitter de fast etter at boltene er løsnet. Da må du dælje løs på felgen med en slegge, og hvis du bryr deg om utseendet på felgen, er det i så fall bare å kaste den etter du har fått den av. Anta at sannsynligheten for at en felg må dæljes løs er 0.2, og at du bytter alle seks hjul på din Mercedes-Benz G63 AMG 6x6 Geländewagen. Sannsynligheten for at to felger blir ødelagt er

$$\binom{6}{2} 0.2^2 0.8^4 = 0.24576.$$

$\triangle$

Eksempel. Anta at du har tre transistorradioer, kalt A, B og C, alle fra 70-tallet. På et gitt tidspunkt virker hver av dem med uavhengig av hverandre med sannsynlighet p . Vi kan sette opp et utfallsrom ved å definere A, B og C som at radio A, B og C virker. Siden $P(A \cup B \cup C) = P(A)P(B)P(C)$ får vi

Utfall	Sannsynlighet
$A \cup B \cup C$	p^3
$\overline{A} \cup B \cup C$	$p^2(1-p)$
$A \cup \overline{B} \cup C$	$p^2(1-p)$
$A \cup B \cup \overline{C}$	$p^2(1-p)$
$\overline{A} \cup \overline{B} \cup C$	$p(1-p)^2$
$\overline{A} \cup B \cup \overline{C}$	$p(1-p)^2$
$A \cup \overline{B} \cup \overline{C}$	$p(1-p)^2$
$\overline{A} \cup \overline{B} \cup \overline{C}$	$(1-p)^3$

La nå X være antall radioer som virker. Hvis du ønsker sannsynligheten for at $X = 2$, ser du at det er de tre utfallene $\overline{A} \cup B \cup C$, $A \cup \overline{B} \cup C$ og $A \cup B \cup \overline{C}$ som må telles med. Siden alle tre utfall har sannsynlighet $p^2(1-p)$, blir

$$P(X = 2) = \binom{3}{1} p^2(1-p) \qquad \triangle$$

Eksempel. Anta at en fyr som skal ta jegerprøven, ikke har lest på stoffet, og blir nødt til å gjette på alle spørsmålene. Vi antar at det er 50 spørsmål, og at det er fire svaralternativer på hver. Hvis vill gjetting er fyrens metode for å bestå prøven, er det 25% sannsynlighet for å treffe på hvert enkelt spørsmål. Dette kan er et binomisk forsøk med $n = 50$, og $p = 0.25$. Sannsynligheten for 10 rette svar er

$$P(A_{10}) = \binom{50}{10} 0.25^{10} \cdot 0.75^{40} = 0.098.$$

Sannsynligheten for mindre enn 3 rette er

$$\begin{aligned} P(A_0 \cup A_1 \cup A_2) &= P(A_0) + P(A_1) + P(A_2) \\ &= \sum_{r=0}^{2} \binom{50}{r} 0.25^r 0.75^{50-r} = 0.000087. \end{aligned}$$

Så noen rette svar får han nok. $\triangle$

Poissonfordelingen

Anta at du har en stokastisk variabel X som kan ta alle ikkenegative heltallsverdier. Dersom sannsynlighetsfordelingen er gitt ved

$$P(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}$$

sier vi at X er poissonfordelt. Poissonfordelingen brukes gjerne i situasjoner der et visst antall forekomster av noe kan inntreffe, og dette antallet ikke har noen øvre begrensning. Et eksempel er antall sms du mottar pr time, antall punkteringer i et fylke på en dag, eller antall konsumerte fuglefrø på en typisk kveld på Omega Verksted.

Eksempel. Vi antar at antall mål Brann scorer på en typisk kamp er poissonfordelt med forventet antall mål $\lambda = \frac{1}{100}$. Vi kan beregne sannsynligheten for at de scorer to mål:

$$P(X = 2) = e^{-\frac{1}{100}} \frac{\frac{1}{100}^2}{2!} \approx 0.$$

Sannsynligheten for at de i det hele tatt scorer mål, er

$$1 - P(X = 0) = 1 - e^{-\frac{1}{100}} \frac{\frac{1}{100}^0}{0!} = 1 - e^{-\frac{1}{100}} \approx 0. \ \triangle$$

Oppgaver

1: Vis at den binomiske fordelingen tilfredsstiller aksiomene for sannsynlighetsfordelingen, og at den har forventning np.

2: Vis at den poissonfordelingen tilfredsstiller aksiomene for en sannsynlighetsfordeling, og at den har forventning λ.