II

Slumpvariabler (forts.)

Vi använder stora bokstäver X, Y, Z, T, osv. för slumpvariabler och små bokstäver x, y, z, t, k för data/resultat/utfall.

Varning! Boken fuskar med detta och använder alltid små bokstäver.

Nu när vi har introducerat slumpvariabler så kan vi smidigt skapa händelser utifrån slumpvariablerna.

Ex 1: Vi kastar en tärning två gånger och X är summan av de två kasten. Händelsen att vi slår 7 kan beskrivas som $\{X = 7\}$, och sannolikheten att slå 7 är då $P(X = 7)$. Händelsen att vi slår 7 eller mindre kan beskrivas som $\{X \leq 7\}$ och sannolikheten för detta $P(X \leq 7)$. Händelsen att slå 7 eller 10 kan beskrivas som $\{X = 7\} \cup \{X = 10\}$ och sannolikheten för detta

$$P(\{X = 7\} \cup \{X = 10\}) = P(X = 7) + P(X = 10) \quad \text{(disjunkt!)}$$

Diskreta slumpvariabler

En diskret slumpvariabel X bestäms genom att vi

1) Specificerar vilka värden X kan anta.

2) Beskriver vad de motsvarande sannolikheterna är.

I tärningsexemplet ovan (Ex. 1) så kan X anta värdena 2, 3, 4, ..., 12 och sannolikheterna är

$$P(X=k) = \begin{cases} \frac{1}{36} & \text{om } k=2 \\ \frac{2}{36} & \text{om } k=3 \\ \frac{3}{36} & \text{om } k=4 \\ \vdots \\ \frac{1}{36} & \text{om } k=12 \end{cases}$$

Övning I: Bekräfta att detta stämmer med hjälp av divisionsregeln. ~~för~~

Sannolikhetsfunktionen (S(f)) för en slumpvariabel X är en graf/tabell/formel som anger alla sannolikheter för alla värden som X kan anta.

$P(X=k)$ ovan är en sannolikhetsfunktion.

Ibland skrivs $p(x)$ istället för $P(X=x)$

$p(k)$ $P(X=k)$

$$\boxed{\begin{array}{l} \text{En sannolikhetsfunktion uppfyller alltid} \\ 1. \quad P(X=x) \geq 0 \quad \text{för alla } x \\ 2. \quad \sum P(X=x) = 1 \end{array}}$$

Summan $\sum P(X=x)$ är en summa över alla värden på x.

Ex1: Tärningsexemplet igen. Villkor 1) är uppenbarligen uppfyllt. Villkor 2)?

$$\sum P(X=k) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \dots + \frac{2}{36} + \frac{1}{36} = 1$$

Väntevärde

Ofta vill man veta det "förväntade värdet" ~~av~~ av en slumpvariabel

Definition: Väntevärdet av en diskret ~~slu~~ slumpvariabel X ~~...~~ betecknas av $E(X)$ och ges av

$$E(X) = \sum x P(X=x)$$

där summan går över alla möjliga värden på x.

Ex1: Låt X vara resultatet av ett tärningskast. Beräkna $E(X)$.

L: $$E(X) = \sum_{k=1}^{6} k P(X=k) = \sum_{k=1}^{6} k \cdot \frac{1}{6} = \frac{1}{6} \sum_{k=1}^{6} k =$$

$$= \frac{1}{6} \cdot \frac{6 \cdot 7}{2} = \frac{7}{2} = 3.5$$

OBS: → Beräkning av väntevärdet kräver kunskap om sannolikhetsfunktionen.

→ Väntevärdet kan vara ett värde som X aldrig kan anta, t.ex. 3.5 i exemplet ovan.

Det är också användbart att kunna beräkna väntevärdet av en funktion av en slumpvariabel.

Om f är en funktion och X är en ~~sla~~ diskret slumpvariabel så gäller att

$$\boxed{E\left(f(X)\right) = \sum f(x) P(X = x)}$$

Varians

Väntevärdet är ett ~~tal~~ värde som talar om hur ~~en~~ en slumpvariabel beter sig i "genomsnitt". Det är också ~~viktigt~~ viktigt att kunna analysera hur ~~mycket~~ stor variation en slumpvariabel ger upphov till.

Definition: Variansen av en diskret slumpvariabel X betecknas med $\mathrm{Var}(X)$ och ~~beräknas~~ ges av

$$\mathrm{Var}(X) = E\left[(X - E(X))^2\right]$$

Tal: Låt X och Y vara två diskreta slumpvariabler.
X kan anta värdena $-1, 0, 1$ och har sannolikhetsfunktion $P(X = k) = \frac{1}{3}$ för alla k. $(-1, 0, 1)$
Y kan anta värdena ~~…~~ $-100, 0, 100$ och har sannolikhetsfunktion $P(Y = k) = \frac{1}{3}$ för ~~…~~ alla k. $(-100, 0, 100)$
Beräkna $E(X)$, $E(Y)$, $\mathrm{Var}(X)$ och $\mathrm{Var}(Y)$

$$E(X) = \sum k\,P(X=k) = -1\cdot\frac{1}{3} + 0\cdot\frac{1}{3} + 1\cdot\frac{1}{3} = -\frac{1}{3} + \frac{1}{3} = 0$$

$$E(Y) = \sum k\,P(Y=k) = -100\,\frac{1}{3} + 0\cdot\frac{1}{3} + 100\cdot\frac{1}{3} = \frac{-100}{3} + \frac{100}{3} = 0$$

$$Var(X) = E\Big[\big(X - \underbrace{E(X)}_{=0}\big)^2\Big] = E(X^2) =$$

$$= \sum k^2 P(X=k) = (-1)^2\cdot\frac{1}{3} + 0^2\cdot\frac{1}{3} + 1^2\cdot\frac{1}{3} = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

$$Var(Y) = E\Big[\big(Y - \underbrace{E(Y)}_{=0}\big)^2\Big] = E(Y^2) =$$

$$= \sum k^2 P(Y=k) = (-100)^2\cdot\frac{1}{3} + 0^2\cdot\frac{1}{3} + 100^2\cdot\frac{1}{3} =$$

$$= \frac{10\,000}{3} + \frac{10\,000}{3} = \frac{20\,000}{3}$$

Trots att det förväntade värdet för X är samma som för Y så är variansen <u>mycket</u> större för Y, dvs spridningen är större hos Y än hos X.

Ofta används notationen $\boxed{\mu = E(X)}$ och $\boxed{\sigma^2 = Var(X)}$

"my" "sigma kvadrat"

<u>Definition:</u> <u>Standardavvikelsen</u> av en slumpvariabel X betecknas med σ och ges av $\sigma = \sqrt{\sigma^2} = \sqrt{Var(X)}$.

~~Detecknas~~

Standardavvikelsen är precis som variansen ett mått på spridningen hos en slumpvariabel. Standardavvikelsen är lättare att tolka. ~~Om~~ Tex om X mäts i meter så kommer σ^2 ~~att~~ ha enheten kvadratmeter, medan σ kommer ha enheten meter.

OBS: Varken variansen eller standardavvikelsen kan vara negativa, men det kan väntevärdet.

Fördelningen av en slumpvariabel bestäms av dess sannolikhetsfunktion. Vi kommer prata om tre speciella (diskreta) fördelningar,

1) likformig fördelning
2) binomialfördelning
3) poissonfördelning.

Likformig fördelning

"likformig" = "lika mycket överallt"

Definition: En slumpvariabel X är likformigt fördelad ~~om~~ på mängden A om

$$P(X = x) = \frac{1}{\#\text{element i } A} \quad \text{för varje } x \in A$$

Exempel: Vi kastar en tärning en gång och X är resultatet.

$$P(X = k) = \frac{1}{6} \quad \text{för } k \in \{1, 2, 3, 4, 5, 6\}$$

alltså är X likformigt fördelad på mängden $\{1, 2, 3, 4, 5, 6\}$

Ex2 Vi kastar två tärningar och X = resultatet.

$$P(X = k) = \begin{cases} \frac{1}{36} & \text{om } k = 2 \\ \frac{2}{36} & \text{om } k = 3 \\ \frac{3}{36} & \text{om } k = 4 \\ \vdots & \end{cases}$$

X är ej likformigt fördelad på $\{2, 3, 4, \ldots, 12\}$

Ex 3: Samma upplägg som i exempel 1 där X var resultatet av ett tärningskast. Låt $Y = X^2$. Värdena som Y kan anta är $\{1^2, 2^2, 3^2, 4^2, 5^2, 6^2\} = \{1, 4, 9, 16, 25, 36\}$

Vi har att $P(Y = k) = \frac{1}{6}$ för alla $k \in \{1, 4, 9, 16, 25, 36\}$ så Y är likformigt fördelad på $\{1, 4, 9, 16, 25, 36\}$

En räkneregel för variansen

Ofta kan det vara besvärligt att beräkna variansen direkt från definitionen. Följande räkneregel kan då vara användbar.

$$\boxed{Var(X) = E(X^2) - (E(X))^2}$$