Discrete Probability Distributions

Note Title 21/01/2009

(a) Random Variables

Suppose we have a die with 6 faces labelled with 3 '1's, 2 '2's and 1 '3'.

Let the RANDOM VARIABLE X be "the number obtained by rolling the die"

Then X can take a value x, where $x = 1$, 2 or 3.

[Note the distinction between

X - the variable described in words

x - the values the variable can take

- so we can write $p(X = x)$]

We can make a table showing the PROBABILITY DISTRIBUTION of X:

x	$p(X=x)$
1	$\frac{1}{2}$
2	$\frac{1}{3}$
3	$\frac{1}{6}$

Note that $\sum p(X=x) = 1$ - this must be true for any Probability Distribution.

Sometimes we can express the probability distribution using a PROBABILITY FUNCTION eg. we could write

$$p(X=x) = \begin{cases} \frac{4-x}{6} & \text{if } x = 1, 2, \text{ or } 3 \\ 0 & \text{otherwise.} \end{cases}$$

Sometimes we need to use the CUMULATIVE PROBABILITY FUNCTION which is written $F(x)$ (NB CAPITAL F!).

$$\boxed{F(x) = p(X \leq x)}$$

x	$F(x)$
1	$\frac{1}{2}$
2	$\frac{1}{2}+\frac{1}{3}=\frac{5}{6}$
3	1

Note that we can have something like

$$F(2.3) = \frac{5}{6}$$

Example! The random variable H has probability distribution defined by

$$p(H=h) = \begin{cases} kh & \text{if } h = 1, 2, 3 \text{ or } 4 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the value of k

(b) Find $F(2.9)$

(a)

h	$P(H=h)$	
1	k	$\frac{1}{10}$
2	$2k$	$\frac{2}{10}$
3	$3k$	$\frac{3}{10}$
4	$4k$	$\frac{4}{10}$

$$k + 2k + 3k + 4k = 1$$
$$10k = 1$$
$$k = \frac{1}{10}$$

(b) $F(2.9) = \frac{1}{10} + \frac{2}{10} = \frac{3}{10}$

p 150 Ex 8A Q 1, 2, 3, 5 by Tuesday.

Example 2 A r.v. X has cumulative probability function

$$F(x) = \begin{cases} \dfrac{x^2+k}{25} & (x = 1, 2, 3, 4) \\ 0 & (x < 1) \\ 1 & (x \geq 4) \end{cases}$$

$p(X \leq x) = 1$ if $x \geq 4$

(a) Find the value of k

(b) Find the probability distribution of X.

(a) $p(X \leq 4) = 1 \Rightarrow \dfrac{4^2+k}{25} = 1$

$\Rightarrow k = 9$

(b)

x	$F(x) = p(X \leq x)$	$p(X = x)$
1	$\frac{10}{25}$	$\frac{10}{25}$
2	$\frac{13}{25}$	$\frac{3}{25}$
3	$\frac{18}{25}$	$\frac{5}{25}$
4	1	$\frac{7}{25}$

Example 3 In a game, the player pays the bank 20p for a go, then throws a normal die and tosses a coin. She wins according to the following table:

	H	T
6	50p	35p
not a 6	30p	—

Let G be the player's gain in playing once.
Find the probability distribution of G.

g	$P(G=g)$	Working
30	$\frac{1}{12}$	$P(6 \text{ AND } H) = \frac{1}{6} \times \frac{1}{2}$
15	$\frac{1}{12}$	$P(6 \text{ AND } T) = \frac{1}{6} \times \frac{1}{2}$
10	$\frac{5}{12}$	$P(\text{not-}6 \text{ AND } H) = \frac{5}{6} \times \frac{1}{2}$
−20	$\frac{5}{12}$	$P(\text{not-}6 \text{ AND } T) = \frac{5}{6} \times \frac{1}{2}$

$\Sigma p = 1$

Ex 8A (cont) Q 6, 7, 8, 9, 10, 11, 12
start now – finish by next Tuesday.

(b) Expectation or Expected Mean $E(X)$

If we have real data in a frequency table

$\left(\begin{array}{c|c} x & f \\ \hline & \end{array} \right)$ we can find the mean $\bar{x}$ using the formula $\bar{x} = \dfrac{\sum fx}{\sum f}$.

A probability distribution is theoretical, so we can work out an EXPECTED mean of our random variable X, written $E(X)$

This is calculated in the same way except that we have probabilities instead of frequencies.

So $$E(X) = \frac{\sum px}{\sum p}$$

Since $\sum p$ is always 1, this can be written

$$\boxed{E(X) = \sum px}$$

Examples

① For the die with 3 '1's, 2 '2's and 1 '3' in the notes above, find the expected mean score per roll.

The probability distribution (p.d.) is

x	p
1	$\frac{1}{2}$
2	$\frac{1}{3}$
3	$\frac{1}{6}$

so $E(X) = \left(1 \times \frac{1}{2}\right) + \left(2 \times \frac{1}{3}\right) + \left(3 \times \frac{1}{6}\right)$

$= \underline{\underline{1 \frac{2}{3}}}$

(2) What is the expected gain per game on the die + coin game described above?

g	p
30	$\frac{1}{12}$
15	$\frac{1}{12}$
10	$\frac{5}{12}$
-20	$\frac{5}{12}$

$$E(G) = \left(30 \times \tfrac{1}{12}\right) + \left(15 \times \tfrac{1}{12}\right) + \left(10 \times \tfrac{5}{12}\right) + \left(-20 \times \tfrac{5}{12}\right)$$

$$= \frac{-5}{12}$$

The player can expect to lose $\frac{5}{12}$p per game on average.

p 157 Ex 8B Q 1(a), 8, 10

not variance.

(3) In a board game, the player throws the die in example 1 above. Before rolling she chooses how she will enhance her score, by either:

- doubling the number on the die
- adding two to the number
- squaring the number

Which method, on average, will give the highest score?

In symbols, if X is the score on the die, we need to find:

$$E(2X), \quad E(X+2) \quad \text{and} \quad E(X^2)$$

$2x$	p
2	$\frac{1}{2}$
4	$\frac{1}{3}$
6	$\frac{1}{6}$

$$E(2X) = \sum p\,(2x) = 1 + \tfrac{4}{3} + 1 = \underline{\underline{3\tfrac{1}{3}}}$$

$x+2$	p
3	$\frac{1}{2}$
4	$\frac{1}{3}$
5	$\frac{1}{6}$

$$E(X+2) = \sum p\,(x+2) = \tfrac{3}{2} + \tfrac{4}{3} + \tfrac{5}{6} = \underline{\underline{3\tfrac{2}{3}}}$$

x^2	p
1	$\frac{1}{2}$
4	$\frac{1}{3}$
9	$\frac{1}{6}$

$$E(X^2) = \sum p(x^2) = \tfrac{1}{2} + \tfrac{4}{3} + \tfrac{9}{6} = \underline{\underline{3\tfrac{1}{3}}}$$

Adding two to the roll gives the best result on average.

Note that for any function of X, we can calculate $E(f(X))$ using

$$\boxed{E(f(X)) = \sum p\,f(x)}$$

Note also that

$$E(2X) = 3\tfrac{1}{3} = 2\,E(X)$$

$$E(X+2) = 3\tfrac{2}{3} = E(X) + 2$$

BUT

$E(X^2) = 3\tfrac{1}{3}$ is <u>not</u> equal to $[E(X)]^2 = (1\tfrac{2}{3})^2 = 2\tfrac{7}{9}$

So for a LINEAR function of X

$$\boxed{E(aX+b) = a\,E(X) + b}$$

Example

Given that $E(X) = 4$, find if possible

(a) $E(3X)$ (b) $E(2X+5)$ (c) $E(3-X)$

(d) $E(X^2)$ (e) $E(X^3+X)$.

(a) $E(3X) = 3\,E(X) = 12$

(b) $E(2X+5) = 2\,E(X) + 5 = 13$

(c) $E(3-X) = 3 - E(X) = -1$

(d) $E(X^2)$ }

(e) $E(X^3+X)$ } cannot be found unless we know the whole distribution of X.

(c) Variance $Var(X)$

The variance of a frequency distribution is the standard deviation squared, ie,

$$\text{variance} = \frac{\sum fx^2}{\sum f} - \bar{x}^2$$

It is defined in the same way for a probability distribution :-

$$Var(X) = \frac{\sum px^2}{\sum p} - [E(X)]^2$$

(since $\sum p = 1$)

$$\boxed{\begin{aligned} Var(X) &= \sum px^2 - [E(X)]^2 \\ \text{or} \quad Var(X) &= E(X^2) - [E(X)]^2 \end{aligned}}$$

For a frequency distribution we use standard deviation far more than variance.

For a probability distribution we use variance more often.

<u>Example</u> Find $Var(X)$ where X is the number shown when a die with 3 '1's, 2 '2's and 1 '3' is tossed.

x	p
1	$\frac{1}{2}$
2	$1/3$
3	$1/6$

$E(X) = 1\,2/3$ (see above)

$E(X^2) = \left(1^2 \times \frac{1}{2}\right) + \left(2^2 \times \frac{1}{3}\right) + \left(3^2 \times \frac{1}{6}\right)$

$= 3\,1/3$

$Var(X) = E(X^2) - [E(X)]^2$

$= 3\,1/3 - (1\,2/3)^2$

$= 5/9$

(so s d of $X = \sqrt{5/9}$)

Rules for linear functions:

$$\boxed{\begin{aligned} Var(X+b) &= Var(X) \\ Var(aX) &= a^2\,Var(X) \\ \text{combining these} \quad Var(aX+b) &= a^2\,Var(X) \end{aligned}}$$

(same as for s.d)

(because variance $= (s.d)^2$)

Example A r.v. P has $E(P) = 8$, $Var(P) = 5$

(a) If $Q = 3P + 5$, find $E(Q)$ and $Var(Q)$

$$E(Q) = E(3P+5)$$
$$= 3E(P) + 5$$
$$= 3 \times 8 + 5 = \underline{\underline{29}}$$

$$Var(Q) = Var(3P+5)$$
$$= 3^2 \, Var(P)$$
$$= 9 \times 5 = \underline{\underline{45}}$$

(b) Find $E(P^2)$

[We CANNOT say $E(P^2) = 64$ since $E(P) = 8$
since this is not linear]

We know $Var(P) = 5$

so
$$E(P^2) - [E(P)]^2 = 5$$
$$E(P^2) - 64 = 5$$
$$\underline{\underline{E(P^2) = 69}}$$

(c) The r.v $R = mP + c$

If $E(R) = 6$ and $Var(R) = 20$, find m and c

$$Var(R) = Var(mP+c)$$
$$= m^2 \, Var(P)$$
$$20 = m^2 \times 5$$
$$\underline{\underline{m = 2}}$$

$$E(R) = E(2P + c)$$
$$= 2E(P) + c$$
$$6 = 2 \times 8 + c$$
$$\underline{\underline{c = -10}}$$

p 157 Ex 8B Q 1(a), 4, 6, 7, 13

variance (did mean already)

see Example 3 - but difference always +ve.

p 160 Ex 8C Q 1, 2, 3, 6