Queen Mary University of London

MAS 108 Probability I

Notes 7 Autumn 2005

More about random variables

A random variable X is *symmetric* about a point m if $p(m+x) = p(m-x)$ for all x. The means that the line graph can be reflected in the vertical line at $x = m$ and it still looks the same. I won't prove the following proposition, but it does serve as a check on arithmetic.

Proposition 7 If X is symmetric about m then $E(X) = m$.

Functions of random variables

Let X be a random variable on a probability space $\mathcal{S}$ and let $g\colon \mathbb{R} \to \mathbb{R}$ be a real function. Then $g(X)$ is also a random variable, and it has its own probability mass function and expectation, and so on. In what follows, we assume that g is a 'nice' enough function for all the summations to be well behaved.

Proposition 8 $E(g(X)) = \sum_x g(x) p_X(x)$.

Proof

$$\begin{aligned}
E(g(X)) &= \sum_y y P(g(X) = y) \\
&= \sum_y y \left(\sum_{x \text{ such that } g(x) = y} P(X = x) \right) \\
&= \sum_x g(x) P(X = x) \\
&= \sum_x g(x) p_X(x). \quad \blacksquare
\end{aligned}$$

By taking $g(x) = x^2$ we see that X^2 is just the random variable whose values are the squares of the values of X. Thus

$$E(X^2) = \sum_x x^2 p(x).$$

In fact, $\mathrm{Var}(X) = E\left((X-\mu)^2\right) = E(X^2) - (E(X))^2$.

Example (Three coin tosses: part 2) I toss a fair coin three times. The random variable X gives the number of heads recorded. The possible values of X are 0, 1, 2,3, and its pmf is

a	0	1	2	3
$P(X=a)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

The event $X = 1$, for example, is $\{HTT, THT, TTH\}$, and has probability $3/8$.

What are the expected value and variance of X?

$$E(X) = 0 \times (1/8) + 1 \times (3/8) + 2 \times (3/8) + 3 \times (1/8) = 3/2,$$

$$\mathrm{Var}(X) = 0^2 \times (1/8) + 1^2 \times (3/8) + 2^2 \times (3/8) + 3^2 \times (1/8) - (3/2)^2 = 3/4.$$

Example (Child: part 4)

$$E(X^2) = 0^2 \times \frac{1}{2} + 1^2 \times \frac{1}{2} = \frac{1}{2},$$

so

$$\mathrm{Var}(X) = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}.$$

Example (One die: part 4)

$$E(X^2) = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) = \frac{91}{6},$$

so

$$\mathrm{Var}(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}.$$

Example (Two dice: part 3)

$$\begin{aligned}
\mathrm{Var}(Z) &= E((Z-7)^2) \\
&= \frac{1}{36}(5^2 \times 1 + 4^2 \times 2 + 3^2 \times 3 + 2^2 \times 4 + 1^2 \times 5 + \cdots + 5^2 \times 1) \\
&= \frac{35}{6}.
\end{aligned}$$

Theorem 4 (Properties of expectation) Let X be a random variable, let a and c be constants, and let g and h be real functions. Then

(i) $E(X+c) = E(X) + c$;

(ii) $E(aX) = aE(X)$;

(iii) $E(c) = c$;

(iv) $E(g(X)+h(X)) = E(g(X)) + E(h(X))$.

Proof (i)

$$\begin{aligned} E(X+c) &= \sum_x (x+c)p(x) \\ &= \sum_x [xp(x)+cp(x)] \\ &= \sum_x xp(x) + \sum_x cp(x) \\ &= E(X) + c\sum_x p(x) \\ &= E(X) + c. \end{aligned}$$

(ii) $E(aX) = \sum_x (ax)p(x) = a\sum_x xp(x) = aE(X)$.

(iii) $E(c) = \sum_x cp(x) = c\sum_x p(x) = x$.

(iv)

$$\begin{aligned} E((g(X)+h(X)) &= \sum_x (g(x)+h(x))\,p(x) \\ &= \sum_x [g(x)p(x)+h(x)p(x)] \\ &= \sum_x g(x)p(x) + \sum_x h(x)p(x) \\ &= E(g(X)) + E(h(X)). \quad \blacksquare \end{aligned}$$

Theorem 5 (Properties of variance) Let X be a random variable, and let a and c be constants. Then

(i) $\mathrm{Var}(X) \geq 0$;

(ii) $\mathrm{Var}(X+c) = \mathrm{Var}(X)$;

(iii) $\mathrm{Var}(ax) = a^2\,\mathrm{Var}(X)$.

Proof Put $\mu = E(X)$.

(i) $\mathrm{Var}(X) = \sum_x (x-\mu)^2 p(x)$. For each value of x, both $(x-\mu)^2 \geq 0$ and $p(x) \geq 0$, so $(x-\mu)^2 p(x) \geq 0$. A sum of non-negative terms is itself non-negative.

(ii) Theorem 5 shows that $E(X+c) = \mu + c$, so

$$\mathrm{Var}(X+c) = E[((X+c)-(\mu+c))^2] = E[(X-\mu)^2] = \mathrm{Var}(X).$$

(iii) Theorem 5 shows that $E(aX) = a\mu$, so

$$\mathrm{Var}(aX) = E[(aX-a\mu)^2] = E[(a(X-\mu))^2] = E[a^2(X-\mu)^2] = a^2 E[(X-\mu)^2],$$

using Theorem 5 again, which is $a^2\,\mathrm{Var}(X)$. ■

Note in particular that

(a) adding a constant c to the values of X adds c to $E(X)$ and doesn't change $\mathrm{Var}(X)$;

(b) multiplying the values of X by a constant a multiplies $E(X)$ by a and multiplies $\mathrm{Var}(X)$ by a^2.

Part (a) makes sense with the interpretation that $E(X)$ is a weighted average while $\mathrm{Var}(X)$ measures the spread; adding a constant shouldn't change the spread. For part (b), note that some people use the square root of the variance (which is called the *standard deviation*) as a measure of spread; then doubling the values doubles the standard deviation. But the square roots makes the formulae more complicated so I will stick to variance.