ENGSCI 213: MATHEMATICAL MODELLING 2SE

Oliver Maclaren
oliver.maclaren@auckland.ac.nz

CURRENT TOPIC

2. *Discrete probability models* [3 lectures]

   Discrete random variables and discrete probability distributions. Probability mass functions and cumulative distribution functions. Binomial and Poisson distributions.

LECTURE 5

- Expectation and variance as reduced summaries of probability distributions/random variables
- The cumulative distribution as an alternative representation of the (whole) probability distribution

RECAP: BERNOULLI TRIALS AND THE BINOMIAL DISTRIBUTION

A *Bernoulli random variable* Y is a random variable which only takes 2 values, 0 and 1 say.

The probability (mass) function is

$$P(Y = y) = f_Y(y) = \begin{cases} p, & \text{if } y = 1 \\ 1 - p, & \text{if } y = 0 \end{cases}$$

RECAP: BERNOULLI TRIALS AND THE BINOMIAL DISTRIBUTION

A random experiment is called a *set of Bernoulli trials* if it consists of several trials such that

- Each trial has only *2 possible outcomes* (e.g. 1 or 0, often called "Success" or "Failure")
- The *probability of "Success"* p is the *same* for all trials
- The trials are *independent* i.e. "Success in trial i" doesn't affect the chances of "Success" in any other trial.

RECAP: BERNOULLI TRIALS AND THE BINOMIAL DISTRIBUTION

A *Binomial* random variable X *summarises* the results of a *set of Bernoulli trials* as just a count of the number of success.

$X \sim Bin(n, p)$ if X is the number of successes out of n independent Bernouilli trials

RECAP: BERNOULLI TRIALS AND THE BINOMIAL DISTRIBUTION

The probability (mass) function for a *Binomial* random variable is given by

$$P(X = x) = f_X(x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

SUMMARISING A PROBABILITY DISTRIBUTION/RANDOM VARIABLE

The *'full' information* about a probability distribution for a random variable X is contained in a tabulation or plot of the probability function $P(X = x)$ for *each value x*.

We could give a list of *'interesting features'* of the distribution - e.g. what values of X are 'most likely', how 'variable' X's values are etc.

SUMMARISING A PROBABILITY DISTRIBUTION/RANDOM VARIABLE

A full list of features would be enough to 'fully specify' the distribution. *In many cases, however, we want to only give 'reduced' summary of the full information* contained in the distribution e.g...

SUMMARISING A PROBABILITY DISTRIBUTION/RANDOM VARIABLE

- a measure of the *'centre'* of the distribution (for example the 'mean' or 'mode' - called *'location'* measures)
- a measure of the *'spread'* of a distribution (e.g. the 'standard deviation' or 'variance' - called *'scale'* measures')

We will look at how to define some of these as *types of averages* ('moments')

EXPECTATION

The expectation or 'mean' of a discrete random variable X is written $E(X)$ and defined by

$$E(X) := \sum_{x \in S_X} x P(X = x)$$

$E(X)$ is also often denoted by μ_X and called the 'mean'. *It is not random* and not a function of a particular sample - it is a *property of the whole/'true'/'population' distribution*. It represents what value of X you expect to get 'on average'.

EXPECTATION OF A CONDITIONAL RANDOM VARIABLE (CONDITIONAL EXPECTATION)

We can define a *conditional random variable* symbolised as $X|Y = y$ and called 'X given Y' (i.e. 'Y' is taken as 'fixed' at a given value y') which has a probability distribution

$$\boxed{P((X|Y = y) = x) = P(X = x \mid Y = y)}$$

This is just a name for a *new distribution over X values*, which takes into account 'background' information given by $Y = y$, i.e. $P(X = x \mid y = y) \neq P(X = x)$ in general.

EXPECTATION OF A CONDITIONAL RANDOM VARIABLE (CONDITIONAL EXPECTATION)

The *conditional expectation* of $X|Y = y$ is then simply determined by substituting into our definition

$$\boxed{E((X|Y = y)) := \sum_{x \in S_X} xP(X = x \mid y = y)}$$

(Again - note that $P(X = x \mid y = y)$ is just a name for a new distribution over X values)

VARIANCE AND STANDARD DEVIATION

A measure of the *'variabiilty'* of X values can be also defined as a type of expectation

The *variance* of a discrete random variable X is defined by

$$\boxed{\begin{aligned} Var(X) &:= E[(x - E[x])^2] \\ &= \sum_{x \in S_X} (x - E(X))^2 P(X = x) \end{aligned}}$$

(think: 'mean square difference'; note: we can also define 'higher' moments/averages)

VARIANCE AND STANDARD DEVIATION

The standard deviation is given by

$$SD(X) = \sqrt{Var(X)}$$

PROPERTIES OF MEAN

For constants a, b, c

$$E(aX + bY + c) = aE(X) + bE(Y) + c$$

(linearity)

If X and Y are independent random variables then $E(XY) = E(X)E(Y)$ (not true in general!)

PROPERTIES OF VARIANCE

$$Var(X) = E(X^2) - E(X)^2$$
$$Var(aX + b) = a^2 Var(X)$$

If X and Y are independent random variables then $Var(X + Y) = Var(X) + Var(Y)$ (not true in general!)

EXAMPLES

Some example calculations.

CUMULATIVE DISTRIBUTION FUNCTION

Rather than giving a reduced summary, we can also represent our probability (mass) distribution in a *different but equivalent (no loss of info)* manner as a sort of *'running total'*

This is called the *cumulative distribution function* and is defined as...

CUMULATIVE DISTRIBUTION FUNCTION

The *cumulative distribution* is the *'sum of probability up to x'*.
It is often denoted by a captial F (c.f. the mass function) $F_X(x)$ and given by

$$\boxed{\begin{aligned} F_X(x) &:= P(X \leq x) \\ &= \sum_{x' \in S_X | x' \leq x} P(X = x') \end{aligned}}$$

R EXAMPLES

See lecture supplement for plots etc!