Discrete random variables and the binomial distribution

math612.1 612.1 Numbers, arithmetic and algebra

Gunnar Stefansson (editor) with contributions from very many students

September 3, 2018

Simple probabilities

Random variables

A random variable is a concept used to denote the outcome of an experiment before it is conducted.

Simple surveys with replacement

If we randomly draw individuals (with replacement) and ask a question with two possible answers (positive or negative), then the number of positive answers will come from a binomial distribution.

Note that the same methodology is commonly used as an approximation when sampling is done without replacement.

The binomial distribution

If we toss a biased coin n independent times, each with probability p of landing heads up, then the probability of obtaining x heads is

$$\binom{n}{x} p^x (1-p)^{n-x}$$

General discrete probability distributions

A general discrete probability distribution can be described by a list of all possible outcomes and associated probabilities.

The expected value or population mean

The expected value is the sum of the possible outcomes, weighted with the respective probabilities (discrete variable). Think of this in terms of an urn full of marbles, each labelled with number.

The population variance

The (population) variance, for a discrete distribution, is

$$\sigma^2 = E\left[(X-\mu)^2\right] = (x_1-\mu)^2 p_1 + (x_2-\mu)^2 p_2 + \ldots$$

where it is understood that the random variable X has this distribution and μ is the expected value.

In the case of the binomial distribution, it turns out that:
$\sigma^2 = np(1-p)$