Portfolio III - Histograms & Distribution

Figure 1 shows histogram of trials until success, which I sampled from negative binomial distribution, with parameters $r = 5$ and $p = 0.5$ (where r is number of successes and p success probability). Negative binomial's probability mass function (PMF) is given by:

$$P(X = k) = \binom{k + r - 1}{r - 1} p^r (1 - p)^k, \quad k \geq 0$$

p is success probability, k is number of failures before achieving r successes.

In producing this plot I:

- Generated random samples from negative binomial distribution using Python's scipy's `scipy.stats.nbinom.rvs`[1] for above parameters.
- Made histogram by resampling sample means - used 50 samples, repeated 500 times, i.e. histogram shows sampling distribution of means.
- Estimated median of the sample means and 66% confidence intervals (lower = 17th percentile, upper = 83rd percentile) from distribution of resampled outcomes.

Science the histogram's bar heights represent sample means, each bar's height is a random variable. Figure 1 shows the sampling distribution of sample means and the variability due to resampling (through a 66% confidence interval).

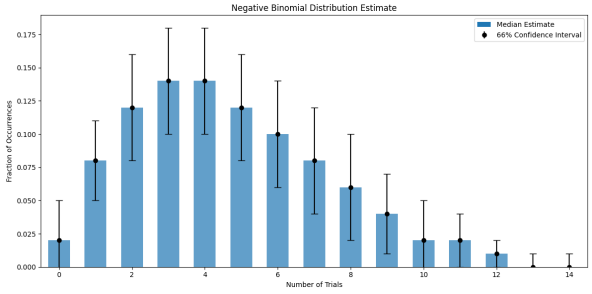

Figure 1: Histogram of trials until success for a negative binomial distribution. Error bars show 66% confidence intervals.

[1] `https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.nbinom.html`