

# 1 Negative Binomial Distribution (NBD)

A Bernoulli trial is an experiment that can result is either a 'success' or a 'failure', but not both.

Consider a sequence of Bernoulli trials with probability of success  $p$  and probability of failure  $q$  such that  $0 \leq p \leq 1$  and  $p + q = 1$ . If  $X$  is the number of failures before the  $r^{th}$  success,  $X$  is said to follow a NBD with parameters  $r$  and  $p$ , denoted by:

$$X \sim \text{NBin}(r, p)$$

## 1.1 Probability Mass Function Of NBD

$$P(X = n) = \binom{n+r-1}{r-1} p^r q^n, \text{ for } n \in \mathbb{N}, \text{ where } q = 1 - p$$

## 1.2 Expected Value and Variance of NBD

$$E(X) = \frac{r(1-p)}{p}$$

$$\text{Var}(X) = \frac{r(1-p)}{p^2}$$

## 1.3 Conditions for NBD

- Experiment must 2 mutually exclusive outcomes denoted as 'success' or 'failure'
- Probability of success must be constant for each trial
- Each trial must be independent
- The experiment must have a finite number of success(es)

## 1.4 Relationship between Binomial Distribution and NBD

Consider  $n$  independent Bernoulli trials with the same probability of success  $p$ . If  $Y$  is the number of successes, it is said to follow a binomial distribution with parameters  $n$  and  $p$ .

$$Y \sim \text{Bin}(n, p)$$

Upon comparison, both the BD and NBD are based upon independent Bernoulli trials. However, they differ in what they are counting. The BD counts the

number of successes in a fixed number of trials  $n$  while the NBD counts the number of failures until a fixed number of successes  $r$ .

## 1.5 Illustration of NBD

To illustrate, we will use data from Statistics Mauritius pertaining to grades of student in Economics A Level during the 2023 seating. Below is a summary of the data collected, along with mean and variance.

Grade	Point Range	$f_i$	$x_i$	$f_i \cdot x_i$	$(x_i - \mu)^2$	$f_i \cdot (x_i - \mu)^2$
A*	129-180	75	154.5	11587.5	5731.441922	429858.1441
A	113-129	261	120.5	31450.5	1739.414392	453987.1564
B	95-112	435	103.5	45022.5	610.4006273	265524.2729
C	83-95	419	89	37291	104.1682985	43646.51705
D	71-83	513	77	39501	3.2174056	1650.529073
E	60-71	490	65.5	32095	176.7227999	86594.17197
F	0-60	495	30	14850	2380.826409	1178509.072

$$\mu = \frac{\sum f_i \cdot x_i}{\sum f_i} = 78.79$$

$$\sigma = \frac{1}{\sum f_i} \cdot \sum f_i \cdot (x_i - \mu)^2 = 915.09$$

**Scenario:** Consider an event where A-Level economics students are gathered, the event organiser wants an  $r^{th}$  number of students who obtained A\* and starts approaching attendees about their grades. Let the random variable  $X$  represent the probability that the organiser gathers those students after  $n$  number of attempts. From this, the following NBD can be constructed:

$$X \sim \text{NBD}(3, 0.0279)$$

Taking a sample of 50, the rnbinom function in R outputs the following:

x	0	1	2	3	4	5	6	7	8	9	10	11
f	4	10	8	6	8	6	3	4	0	0	0	1

Based on the above, the mean and variance can be calculated:

$$\mu = \frac{\sum x \cdot f}{\sum f} = \frac{163}{50} = 3.26$$

$$\sigma^2 = \frac{(x_i - \mu)^2}{n - 1} = 5.46$$

### 1.5.1 Bar Chart

### 1.5.2 Issues with this application

## 2 Hypergeometric Distribution (HD)

Consider a population of  $N$  objects which are divided into 2 types: type A and type B. There are  $n$  objects of type A and  $N - n$  objects of type B. Suppose a random sample of size  $r$  is taken (without replacement) from the entire population of  $N$  objects. If  $X$  is the number of objects of type A in the sample, then  $X$  follows a HD with parameters  $n$ ,  $N - n$  and  $r$  denoted by:

$$X \sim \text{HGeom}(n, N - n, r)$$

### 2.1 Probability Mass Function of HD

$$p(k) = \frac{{}^nC_k \cdot {}^{(N-n)}C_{(r-k)}}{{}^NC_r}, \text{ for } \max\{0, r - (N - n)\} \leq k \leq \min\{r, n\}$$

### 2.2 Expected Value and Variance of HD

$$E(X) = \frac{nr}{N}$$
$$\text{Var}(X) = \frac{nr}{N} \cdot \frac{N - r}{N} \cdot \frac{N - n}{N - 1}$$

### 2.3 Conditions for HD

- Finite population
- Population can be separated into 2 types
- Sampling is done without replacement (dependent trials).