MORE DISCRETE PROBABILITY DISTRIBUTIONS (*FOR IB*)

BERNOULLI: - This is a Binomial Distribution where only 1 trial is conducted. i.e. n = 1

For example if we play one game with the probability of winning being p, the distribution will be:

X:	0	1
Pr:	1-p	p

HYPERGEOMETRIC: - If we sample without replacement then we have a Hypergeometric Distribution.

For example if 3 balls are chosen with replacement from a bag containing 5 red and 7 black balls. The random variable X is the number of red balls.

Note that the number of ways of choosing any 3 from 12 is $\binom{12}{3}$ which is $\frac{12!}{8!3!}$ = 220 and this will be the denominator.

The number of ways of choosing m red balls and therefore (3 – m) black balls is the number of ways of choosing **m** reds from the 5 available red balls and **(3 – m)** black balls from the available 7 black balls. $\binom{5}{m} \times \binom{7}{3-m}$. This will be the numerator.

We will use the simplest definition of probability to construct the probability distribution. Pr(event) = $\frac{\text{number of favourable outcomes}}{\text{Total number of outcomes}}$

X:	**0**	**1**	**2**	**3**
Pr:	$\dfrac{\binom{5}{0} \times \binom{7}{3}}{\binom{12}{3}}$	$\dfrac{\binom{5}{1} \times \binom{7}{2}}{\binom{12}{3}}$	$\dfrac{\binom{5}{2} \times \binom{7}{1}}{\binom{12}{3}}$ 3	$\dfrac{\binom{5}{3} \times \binom{7}{0}}{\binom{12}{3}}$

X:	**0**	**1**	**2**	**3**
Pr:	$\frac{35}{220}$	$\frac{105}{220}$	$\frac{70}{220}$	$\frac{10}{220}$

GEOMETRIC: We use this distribution to model continual trials until success is achieved. For instance when throwing a dart and stopping when a bull's-eye is scored. We assume that the events are independent and the probability of success is constant. The random variable X is the number of trials needed until the first success is achieved. Let Pr(success) = p

Pr(x = 1) is the probability that success is achieved at the first attempt = p
Pr(x = 2) is the probability that the first success is achieved at the second attempt = failure then a success = (1-p)p

Pr(x = r) is the probability that the first success is achieved at the **r^th^** attempt = (r-1) failures then a success = $(1-p)^{(r-1)}p$

So the probability distribution can be written **$\Pr(X = x) = (1-p)^{(x-1)}p$ for x = 1, 2, 3, ………………….**

The mean, $E(X) = \sum_{1}^{\infty} x(1-p)^{(x-1)}p = p\sum_{1}^{\infty} x(1-p)^{(x-1)} = p(1 + 2q + 3q^2 + 4q^3 + \ldots\ldots\ldots\ldots)$ **where q = 1 - p**

The **series in the bracket** is neither an arithmetic nor a geometric series. It is an arithmetico-geometric series and is the differential of the series

$(q + q^2 + q^3 + q^4 + \ldots\ldots)$ This geometric series has a sum to infinity of $\frac{q}{1-q}$ **and its differential using the quotient rule is:**

$\frac{d}{dx}\left(\frac{q}{1-q}\right) = \frac{(1-q)-q(-1)}{(1-q)^2} = \frac{1}{(1-q)^2} = \frac{1}{p^2}$. **Therefore $E(X) = p\frac{1}{p^2} = \frac{1}{p}$. The variance of the geometric distribution is $\frac{q}{p^2}$.**

<u>NEGATIVE BINOMIAL:</u> - The geometric distribution describes the number of trials required to get a single success. E.g. The number of darts you have to throe before you get a bull's-eye. Pr(bulls-eye) = p

If we consider the number of trials required for 2 or more successes then we have a **negative binomial distribution.**

I need 2 successes: Pr(x = 5) is the probability that it takes 5 trials to get the 2 successes. That is one success in 4 trials followed by a success.
One success in 4 trials is a Binomial probability $\binom{4}{1}p^1q^3$ therefore the probability that it takes 5 trials to get the 2 successes = $\binom{4}{1}p^1q^3\,p$

I need 3 successes as in a 5 set tennis match: Pr(x = 5) is the probability that it takes 5 trials to get the 3 successes.
That is two success in 4 trials followed by a success.
Two success in 4 trials is a Binomial probability $\binom{4}{2}p^2q^2$ therefore the probability that it takes 5 trials to get the 3 successes = $\binom{4}{2}p^2q^2\,p$

And in general, the probability that it takes x trials to get r successes is $\binom{x-1}{r-1}p^rq^{(r-1)}$ The mean and variance can be seen in the table below.

Discrete distributions

Distribution	Notation	Probability mass function	Mean	Variance
Bernoulli	$X \sim \mathrm{B}(1, p)$	$p^x(1-p)^{1-x}$ for $x = 0, 1$	p	$p(1-p)$
Binomial	$X \sim \mathrm{B}(n, p)$	$\binom{n}{x} p^x (1-p)^{n-x}$ for $x = 0, 1, \ldots, n$	np	$np(1-p)$
Hypergeometric	$X \sim \mathrm{Hyp}(n, M, N)$	$\dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$ for $x = 0, 1, \ldots, n$	np where $p = \dfrac{M}{N}$	$np(1-p)\left(\dfrac{N-n}{N-1}\right)$ where $p = \dfrac{M}{N}$
Poisson	$X \sim \mathrm{P_o}(m)$	$\dfrac{m^x e^{-m}}{x!}$ for $x = 0, 1 \ldots$	m	m
Geometric	$X \sim \mathrm{Geo}(p)$	pq^{x-1} for $x = 1, 2, \ldots$	$\dfrac{1}{p}$	$\dfrac{q}{p^2}$
Negative binomial	$X \sim \mathrm{NB}(r, p)$	$\binom{x-1}{r-1} p^r q^{x-r}$ for $x = r, r+1, \ldots$	$\dfrac{r}{p}$	$\dfrac{rq}{p^2}$
Discrete uniform	$X \sim \mathrm{DU}(n)$	$\dfrac{1}{n}$ for $x = 1, \ldots, n$	$\dfrac{n+1}{2}$	$\dfrac{n^2-1}{12}$

©www.funkyfactsheets.co.uk