Important Distributions

Binomial Distribution

The binomial distribution, denoted by $Bin(n, p)$, describes the distribution of successes that have probabililty p, in n trials.

Properties of the Binomial Distribution

- Probability Mass Function:

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

- Range:

$$\{0, 1, \ldots, n\}$$

- Mean:

$$\mu = np$$

- Variance:

$$\sigma^2 = np(1-p)$$

The binomial distribution can be thought of as the sum of n iid Bernouilli variables with probability p. Therefore, the probability that there are k successes in a specific order is $p^k(1-p)^{n-k}$ by independence. We then multiply this by $\binom{n}{k}$, to get the total probability of $P(X = k)$.

Example 1. What is the probability of getting four or more heads in six tosses of a fair coin?

$$P(X \geq 4) = P(4) + P(5) + P(6)$$

Since $p = 1 - p$ in the case of a fair coin, and the fact that combinations are symmetrical (remember that $\binom{n}{k} = \binom{n}{n-k}$), instead of directly calculating the sum of three probabilities, what we can do is

$$\begin{aligned} P(X \geq 4) &= \frac{1 - P(X = 3)}{2} \\ &= \frac{1 - \binom{6}{3}/2^6}{2} \\ &= \frac{11}{32} \end{aligned}$$

Example 2. Let's say we consider a trial to be six die rolls. What is the probability that among 5 trials, at least 4 of the families have 4 or more heads?
We know among a single trial, the probability that we get 4 or more heads is 11/32. Now, we can consider each trial to have probability $p = 11/32$, and therefore our distribution becomes $X \sim Bin(5, 11/32)$. Since $p \neq 1-p$, the probabilities are not symmetrical, so we cannot use the same trick as the previous example.

$$\begin{aligned}
P(X \geq 4) &= P(X=4) + P(X=5) \\
&= \binom{5}{3}\left(\frac{11}{32}\right)^4\left(\frac{21}{32}\right) + \binom{5}{5}\left(\frac{11}{32}\right)^5 \\
&\approx 0.051
\end{aligned}$$

Example 3. Consider if we have two independent random variables such that $X \sim Bin(n, p)$ and $Y \sim Bin(m, p)$. What is the distribution of $X + Y$?
Intuitively, $X + Y \sim Bin(n + m, p)$ because $X + Y$ is just simulating coin flips with probability p in $m + n$ trials because we are given that X, Y are independent. To prove this, we can calculate

$$\begin{aligned}
P(X+Y=k) &= \sum_{j=0}^{k} P(X=j, Y=k-j) \\
&= \sum_{j=0}^{k} P(X=j)P(Y=k-j) \quad \text{By independence} \\
&= \sum_{j=0}^{k} \binom{n}{j} p^j (1-p)^{n-j} \binom{m}{k-j} p^{k-j} (1-p)^{k-j} \\
&= \sum_{j=0}^{k} p^k (1-p)^{n+m-k} \binom{n}{j}\binom{m}{k-j} \\
&= p^k (1-p)^{n+m-k} \binom{n+m}{k}
\end{aligned}$$

The last line comes from Vandermonde's Identity

$$\sum_{j=0}^{k} \binom{n}{j}\binom{m}{k-j} = \binom{n+m}{k}$$

The proof of this is left to the reader (try a combinatorial proof).

Geometric Distribution

The geometric distribution, denoted by $Geom(p)$, describes the number of trials until a success, which occurs with probability p.

Properties of the Geometric Distribution

- Probability Mass Function:
$$P(X = k) = (1 - p)^{k-1}p$$
- CDF:
$$P(X \leq k) = 1 - (1 - p)^k$$
- Range:
$$\{1, 2, \ldots\}$$
- Mean:
$$\mu = \frac{1}{p}$$
- Variance:
$$\sigma^2 = \frac{1 - p}{p^2}$$
- Memoryless:
$$P(X > s + t | X > t) = P(X > s)$$

Example 4. Let $X \sim Geom(p)$. Find $E(X|X > n)$.
By the memoryless property, this is equal to $n + E(X) = n + \frac{1}{p}$. To see why this is true

$$\begin{aligned}
E(X|X > n) &= \sum_{x=0}^{\infty} xP(X = x|X > n) \\
&= \sum_{x=0}^{\infty} xP(X = x - n) \quad \text{By the memoryless property} \\
&= \sum_{u=0}^{\infty} (u + n)P(X = u) \\
&= \sum_{u=0}^{\infty} uP(X = u) + \sum_{u=0}^{\infty} nP(X = u) \\
&= E(X) + n
\end{aligned}$$

Example 5. Let X be the number of failures until the r-th success, which occurs with probability p. Find $E(X)$.
This can be thought of as r successive geometric distributions. We expect that the number of trials until the first success is $\frac{1}{p}$, therefore we expect the number of trials until the r-the success to be $\frac{r}{p}$. Thus, the number of failures until the r-th success has expectation $\frac{r}{p} - r = \frac{r(1-p)}{p}$. This is called the *negative binomial* distribution with parameters r, p.

Poisson Distribution

The Poisson distribution, denoted by $Poiss(\lambda)$, describes the number of events occuring within a fixed interval of time/space, where these events occur independently and with a known average rate λ.

Properties of the Poisson Distribution

- Probability Mass Function:
$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$
- Range:
$$\{0, 1, \dots\}$$
- Mean:
$$\mu = \lambda$$
- Variance:
$$\sigma^2 = \lambda$$
- Sums:
Let $X_1, X_2, \dots, X_n$ be indpendent Poisson with parameters $\lambda_1, \lambda_2, \dots, \lambda_n$. Their sum is a Poisson random variable with parameter $\sum_{i=1}^{n} \lambda_i$.

Example 6. Suppose that we receive telephone calls as a Poisson process with $\lambda = 0.5$ calls/min. What is the average number of calls in a 5 min interval? What is the probability of no calls in a 5 min interval?
The number of calls in a 5 min interval is distributed as $Poiss(\lambda = 2.5)$. The average number of calls in this interval is 2.5. The probability of zero calls is $\frac{\lambda^0 e^{-2.5}}{0!} = e^{-2.5}$.