Lecture 14

Agenda

1. Negative Binomial Distribution continued
2. Poisson Distribution

Negative Binomial Distribution

We are running independent Bernoulli trials with success probability p, until we get r successes. If X denotes the number of failures before the r-th success, then $X \sim NegBin(r, p)$.

$$Range(X) = \{0, 1, 2, 3, \ldots\}$$

For $x \geq 0$,

$$P(X = x) = \binom{x+r-1}{x} p^r (1-p)^x$$

Mean and Variance

Let W_1 denote the number of failures before the 1st success
W_2 = number of failures between the 1st and 2nd success
$\vdots$
$\vdots$
W_r = number of failures between the $(r-1)$-th and r-th success
Since X is the total number of failures

$$X = W_1 + W_2 + \ldots + W_r$$

Note that

1. W_i is a $Geometric(p)$ random variable for $i = 1, 2, \ldots, r$.
2. W_1, W_2, $\ldots W_r$ arise out of independent experiments.

Hence

$$E(X) = E(W_1) + E(W_2) + \ldots + E(W_r) = \frac{r(1-p)}{p}$$

$$V(X) = V(W_1) + V(W_2) + \ldots + V(W_r) = \frac{r(1-p)}{p^2}$$

Example

A large lot of tires contains 5% defectives. 4 tires are to be chosen for a car.

- Find the probability that you find 2 defective tires before 4 good ones.
- Find the mean and variance of the number of defective tires you find before finding 4 good tires.

Let X = number of defective tires you find before you find 4 good tires.
$X \sim NegBin(4, 0.95)$
So for the first question,

$$P(X = 2) = \binom{4+2-1}{2}(0.95)^4(0.05)^2 = 0.02036$$

For the second part,

$$E(X) = \frac{4 \times 0.05}{0.95} = \frac{4}{19}$$
$$V(X) = \frac{4 \times 0.05}{0.95^2} = \frac{80}{361}$$

Poisson Distribution

Consider a random variable X, such that $X \sim Bin(n, p)$. If n is a very large number, and p is very small then calculating the probabilities $P(X = k)$, are a bit difficult.

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

For large n and small p, $\binom{n}{k}$ is very large and p^k is very small and calculating them and then multiplying them can create calculation errors. Furthermore if you have to calculate something like $P(X <= 100)$ our troubles multiply. Thus we have invented the Poisson Distribution.
Suppose for $n \geq 1$, $X_n \sim Bin(n, p_n)$ where the sequence p_n is defined as $p_n = \frac{\lambda}{n}$, for some $\lambda > 0$. Thus $p_n \to 0$ as $n \to \infty$.

Let's take a look at $P(X_n = k)$. It is 0 when $n < k$. For $n \geq k$,

$$\begin{aligned}
P(X_n = k) &= \binom{n}{k} p_n^k (1-p_n)^{(n-k)} \\
&= \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{(n-k)} \\
&= \frac{n \times (n-1) \times \ldots \times (n-k+1)}{k!} \times p_n^k \times (1-p_n)^{(n-k)} \\
&= \frac{n \times (n-1) \times \ldots \times (n-k+1)}{n^k} \times n^k \times \frac{1}{k!} \times p_n^k \times (1-p_n)^{(n-k)} \\
&= \frac{n \times (n-1) \times \ldots \times (n-k+1)}{n^k} \times (np_n)^k \times \frac{1}{k!} \times (1-p_n)^n \times \frac{1}{(1-p_n)^k}
\end{aligned}$$

As $n \to \infty$,

$$\begin{aligned}
\frac{n \times (n-1) \times \ldots \times (n-k+1)}{n^k} &\to 1 \\
(np_n)^k &= \lambda^k \\
\frac{1}{(1-p_n)^k} &\to 1
\end{aligned}$$

$$(1-p_n)^n = \left(1 - \frac{\lambda}{n}\right)^n \to e^{-\lambda}$$

Thus as $n \to \infty$, $P(X_n = k) \to e^{-\lambda} \frac{\lambda^k}{k!}$. So we define the poisson distribution.

Definition 1. *A random variable X is said to follow the Poisson Distribution with parameter $\lambda > 0$, [written as $X \sim Poisson(\lambda)$] if*

$$Range(X) = \{0, 1, 2, 3, \ldots\}$$

and for $x \geq 0$,

$$P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}$$

The earlier calculation is meant to be a intuitive explanation of how Poisson distribution arises. This does not mean that if $X \sim Poisson(\lambda)$, then X is actually $Bin(n, p_n)$. On the other hand if $X \sim Bin(n, p)$, for some large n and small p, then we may actually use the $Poisson(\lambda)$ approximation for $\lambda = np$.

Mean and Variance

$$X \sim Poisson(\lambda)$$

Then,

$$\begin{aligned}
E(X) &= \sum_{k=0}^{\infty} k \times e^{-\lambda}\frac{\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} k \times e^{-\lambda}\frac{\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} e^{-\lambda}\frac{\lambda^k}{(k-1)!} \\
&= \sum_{l=0}^{\infty} e^{-\lambda}\frac{\lambda^{(l+1)}}{l!} \\
&= \lambda \times \sum_{l=0}^{\infty} e^{-\lambda}\frac{\lambda^l}{l!} \\
&= \lambda \times 1 \\
&= \lambda
\end{aligned}$$

To calculate variance,

$$\begin{aligned}
E(X*(X-1)) &= \sum_{k=0}^{\infty} k \times (k-1) \times e^{-\lambda}\frac{\lambda^k}{k!} \\
&= \sum_{k=2}^{\infty} k \times (k-1) \times e^{-\lambda}\frac{\lambda^k}{k!} \\
&= \sum_{k=2}^{\infty} e^{-\lambda}\frac{\lambda^k}{(k-2)!} \\
&= \sum_{l=0}^{\infty} e^{-\lambda}\frac{\lambda^{(l+2)}}{l!} \\
&= \lambda^2 \times \sum_{l=0}^{\infty} e^{-\lambda}\frac{\lambda^l}{l!} \\
&= \lambda^2 \times 1 \\
&= \lambda^2
\end{aligned}$$

Thus,

$$
\begin{aligned}
&\Rightarrow \quad E(X^2) - E(X) = \lambda^2 \\
&\Rightarrow \quad E(X^2) = \lambda^2 + E(X) = \lambda^2 + \lambda \\
&\Rightarrow \quad V(X) = E(X^2) - [E(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda
\end{aligned}
$$

Thus for a $Poisson(\lambda)$ random variable,

$$E(X) = V(X) = \lambda$$

`Homework :: 3.90,3.96. We found the mean and variance of a Negative Binomial Distribution by breaking it into` r `independent parts, try to do it, using only algebra.`