Lecture 15

Agenda

1. Poisson Distribution Examples
2. Hypergeometric Distribution

Poisson Distribution Examples

Example 1

The manager of a industrial plant is planning to buy a machine of either type A or type B. For each day's operation the number of repairs X, that the machine A needs is a poisson random variable with mean 0.96. The daily cost of operating A is $C_A = 160 + 40 * X^2$. For machine B, let Y be the random variable indicating the number of daily repairs, which has mean 1.12, and the daily cost of operating B is $C_B = 128 + 40 * Y^2$. Assume that the repairs take negligible time and each night the machine are cleaned so that they operate like new machine at the start of each day. Which machine minimizes the expected daily cost ?

Solution. *The expected cost for machine A is,*

$$\begin{aligned} E(C_A) &= E(160 + 40 * X^2) \\ &= 160 + 40 * E(X^2) \\ &= 160 + 40 * (V(X) + [E(X)]^2) \\ &= 160 + 40 * (0.96 + 0.96^2) \\ &= 235.264 \end{aligned}$$

The expected cost for machine B is,

$$\begin{aligned} E(C_B) &= E(128 + 40 * Y^2) \\ &= 128 + 40 * E(Y^2) \\ &= 128 + 40 * (V(Y) + [E(Y)]^2) \\ &= 128 + 40 * (1.12 + 1.12^2) \\ &= 222.976 \end{aligned}$$

Example 2

The number of calls coming per minute into a hotels reservation center is Poisson random variable with mean 3.

(a) Find the probability that no calls come in a given 1 minute period.

(b) Assume that the number of calls arriving in two different minutes are independent. Find the probability that atleast two calls will arrive in a given two minute period.

Solution. (a) *Let X denote the number of calls coming in that given 1 minute period.*

$$X \sim Poisson(3)$$

$$P(X=0) = \frac{e^{-3}3^0}{0!} = e^{-3}$$

(b) *Let X_1 and X_2 be the number of calls coming in the first and second minutes rspectively.*
We want $P(X_1 + X_2 \geq 2)$.

$$\begin{aligned}
P(X_1 + X_2 \geq 2) &= 1 - P(X_1 + X_2 < 2) \\
&= 1 - [P(X_1 = 0, X_2 = 0) + P(X_1 = 1, X_2 = 0) + P(X_1 = 0, X_2 = 1)] \\
&= 1 - [P(X_1 = 0) * P(X_2 = 0) + P(X_1 = 1) * P(X_2 = 0) + \\
&\quad P(X_1 = 0) * P(X_2 = 1)] \\
&= 1 - e^{-3}e^{-3} - e^{-3}e^{-3}\frac{3^1}{1!} - e^{-3}\frac{3^1}{1!}e^{-3} \\
&= 1 - e^{-6} - 3e^{-6} - 3e^{-6} \\
&= 1 - 7e^{-6}
\end{aligned}$$

Hypergeometric Distribution

Example 3

Suppose 30% of the UF students are graduates. If I take a random sample of 10 students find the probability that the number of graduates is atmost 4.

Let X be the number of graduates among the 10 students. Then

$$X \sim Bin(10, 0.3)$$

Using our usual formulas,

$$P(X \leq 4) = 0.8497$$

Example 4

`Suppose I have 50 UF students in my class and 30% of them are graduates. If I take a random sample of 10 students from them find the probability that the number of graduates is atmost 4.`

Let X be the number of graduate students among the 10 students. In this case we cannot say that $X \sim Bin(10, 0.3)$, because for that to happen we need 10 **independent** Bernoulli trials. But here we have 50 students among whom 15 are graduates and 35 are undergraduates. So, for the first trial

$$P(Graduate) = \frac{15}{15+35} = 0.3$$

and that went as expected.
But now suppose in the first trial I got a graduate student. So at the beginning of the second trial we have 14 graduates and 35 undergrads. So now,

$$P(Graduate) = \frac{14}{14+35} = \frac{14}{49} \neq 0.3$$

That should not happened if they were independent Bernoulli trials with probability of success 0.3.
Thus we need **Hypergeometric Distribution**.

Definition 1. *A population of size N contains two types of items, k items of Type I and $N-k$ items of Type II. We randomly sample n items from N, without repitition. Let X denote the number of items of Type I in our sample.*
Then X follows Hypergeometric Distribution with parameters (N, k, n)

$$Range(X) = \{x | \max\{0, n-N+k\} \leq x \leq \min\{n, k\}\}$$

$$P(X = x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

For Example 4 we have, $N = 50$, $k = 15$ and $n = 10$. So now

$$Range(X) = \{x | \max\{0, 10-50+15\} \leq x \leq \min\{10, 15\}\} = \{x | 0 \leq x \leq 15\}$$

$$P(X \leq 4) = \sum_{x=0}^{4} \frac{\binom{15}{x}\binom{35}{10-x}}{\binom{50}{10}} = 0.8750$$

Since we have established the pmf we will head for mean and variance. They can be calculated the hard way i.e. doing it algebrically. But we won't do that. Let's learn a few things before doing those calculations.

Independence of discrete random variables

For any two events, A and B, we say A and B are independent, if

$$P(A \cap B) = P(A) \times P(B)$$

For the past few lectures we have been using the concept of independence of random variables without really defining it.

Definition 2. *Let $X_1, X_2, \ldots X_k$ be k discrete random variables. We say that $\{X_1, X_2, \ldots X_k\}$ are independent if for any*

$$x_1 \in Range(X_1),\ x_2 \in Range(X_2),\ \ldots,\ x_k \in Range(X_k)$$

$$P(X_1 = x_1, X_2 = x_2, \ldots X_k = x_k) = P(X_1 = x_1) \times P(X_2 = x_2) \times \ldots \times P(X_k = x_k)$$

We note that for $i = 1, 2, \ldots, k$; $\{X_i = x_i\}$ is an event. The above definition says that those k events are independent.
Please note one other thing; when we defined a discrete random variable X, we said it has a range like, $Range(X) = \{x_1, x_2, x_3, \ldots\}$. But here $x_1, x_2, x_3, \ldots$ don't belong to range of one variable X; but x_1 denotes some element of $Range(X_1)$, x_2 denotes some element of $Range(X_2)$,and so on.Please don't get confused over this.
Next we state another theorem, we will prove it next class.

Theorem 1. *Let $\{X_1, X_2, \ldots X_k\}$ be k independent random variables. Let $f_1(), f_2(), \ldots, f_k()$ be k functions.Then*

$$E(f_1(X_1) \times f_2(X_2) \times \ldots \times f_k(X_k)) = E(f_1(X_1)) \times E(f_2(X_2)) \times \ldots \times E(f_k(X_k))$$