Stat 265 – Unit 3 – Video Set 5 – Class Examples

Summary:

1) Poisson(λ) Random Variable:

- Consider a continuous time period where successes occur *uniformly* during the time period, with an average of λ successes per time period.
- $Y =$ the number of "successes" in the time period.
- $p_Y(y) = P(Y=y) = \dfrac{\lambda^y}{y!}e^{-\lambda}, \; y = 0,1,2,\ldots, \; \lambda > 0.$
- $\mu_Y = \sigma_Y^2 = \lambda.$

e.g. $Y_1 \sim Poisson(\lambda_1 = 4)$, $Y_2 \sim Poisson(\lambda_2 = 2)$ } If ind.
$Y_1 + Y_2 \sim Poisson(\lambda = 4 + 2 = 6)$

2) Sums of independent Poisson Random Variables:

- If $Y_i \overset{ind}{\sim} Poisson(\lambda_i), \; i = 1,\ldots,n,$ then $U = \sum_{i=1}^{n} Y_i \sim Poisson\left(\lambda = \sum_{i=1}^{n} \lambda_i\right).$

3) Poisson Approximation to the Binomial: $\longrightarrow$ $\lim_{\substack{n \to \infty \\ np = \lambda}} \binom{n}{y} p^y (1-p)^{n-y} = \dfrac{\lambda^y e^{-\lambda}}{y!}$

- Suppose $Y \sim Binomial(n, p)$. Then, for "large" n and "small" p,
  $\Rightarrow Y \approx Poisson(\lambda = np).$

4) Poisson Counting Process:

- Suppose the number of successes in one time unit is a *Poisson* random variable with a mean of λ.
- Let Y_t be the number of successes in a time interval of length t. $(t > 0)$
- It can be shown that, if the number of arrivals in disjoint time intervals are independent and the success rate per time unit is fixed at λ, then $Y_t \sim Poisson\left(\mu_{Y_t} = \lambda t\right).$

e.g. If Y = # of successes per hour $\sim Poisson(\lambda = 3)$. Then,
$Y_{0.5}$ = # of successes in 30 minutes $\sim Poisson(1.5)$.

Example A: A food manufacturer uses an extruder (a machine that produces bite-sized cookies and snack food) that yields revenue for the firm at a rate of $200 per hour when in operation. However, the extruder breaks down according to a Poisson distribution at an average of two times per day it operates. If Y denotes the number of breakdowns per day, the daily revenue generated by the machine is $R = 1600 - 50Y^2$.

a) Find the expected revenue over a 5-day period. (Assume all days are iid.)

b) Find *P*(*R* > 1500**)**. (for a single day)

Y = # of breakdowns / day $\sim$ Poisson$(\mu_Y = \lambda = 2)$, $\sigma_Y^2 = \lambda = 2$.

$$p_Y(y) = \frac{2^y e^{-2}}{y!}, \quad y = 0, 1, 2, \ldots$$

a) R_i = revenue for day i $= 1600 - 50Y_i^2$, $i = 1, \ldots, 5$.

$$E[R] = E[1600 - 50Y^2] = 1600 - 50E[Y^2]$$

$$= 1600 - 50(\sigma_Y^2 + \mu_Y^2) = 1600 - 50(2 + 2^2) = 1300 \text{ / day.}$$

$$E[R_1 + \cdots + R_5] = 5(1300) = 6500.$$

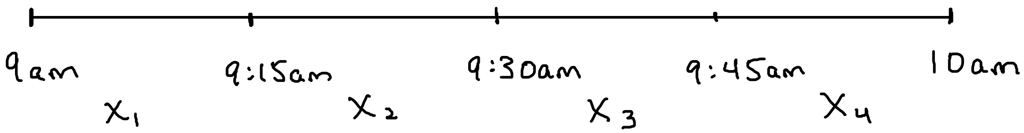
b) $P(R > 1500) = P(1600 - 50Y^2 > 1500) = P(Y^2 < 2)$

$= P(Y < \sqrt{2}) = p_Y(0) + p_Y(1)$

$= e^{-2} + 2e^{-2} = 3e^{-2} \approx 0.40601.$

Example B: Suppose the number of customers arriving at a car dealership in a 1-hour shift between 9:00am and 10:00am follows a Poisson distribution with mean 4. Suppose you know that exactly 1 arrived between 9:00am and 9:30am, exactly 3 arrived between 9:15am and 9:45am, and that nobody arrived in the last 15 minutes. Given this information, what is the expected number of customers that arrived in the 1-hour time period?

Y = # of arrivals / hour $\sim$ Poisson $(\lambda = 4)$.



$\Rightarrow X_1, \ldots, X_4 \overset{ind}{\sim}$ Poisson$(\lambda = 1)$

$p_X(x) = e^{-1}/x!\ , \ x = 0, 1, 2, \ldots$

Information: $X_1 + X_2 = 1$, $X_2 + X_3 = 3$, $X_4 = 0$

So

x_1	x_2	x_3	x_4	Total (y)
0	1	2	0	3
1	0	3	0	4

Y is 3 if $X_1 = 0$, $X_2 = 1$, $X_3 = 2$, $X_4 = 0$.

Y is 4 if $X_1 = 1$, $X_2 = 0$, $X_3 = 3$, $X_4 = 0$.

$$P(Y=3 \mid \text{Info.}) = \frac{P(Y=3 \cap \text{Info.})}{P(\text{Info.})} = \frac{p_{X_1}(0)\,p_{X_2}(1)\,p_{X_3}(2)\,p_{X_4}(0)}{p_{X_1}(0)\,p_{X_2}(1)\,p_{X_3}(2)\,p_{X_4}(0) + p_{X_1}(1)\,p_{X_2}(0)\,p_{X_3}(3)\,p_{X_4}(0)}$$

$$= \frac{(e^{-1})(e^{-1})(\frac{1}{2}e^{-1})(e^{-1})}{(e^{-1})(e^{-1})(\frac{1}{2}e^{-1})(e^{-1}) + (e^{-1})(e^{-1})(\frac{1}{6}e^{-1})(e^{-1})} = \frac{\frac{1}{2}e^{-4}}{\frac{1}{2}e^{-4} + \frac{1}{6}e^{-4}} = \frac{3}{4}.$$

$P(Y=4 \mid \text{Info}) = \frac{1}{4}$. $E[Y \mid \text{Info}] = 3(3/4) + 4(1/4) = 3.25$.

Example C: An insurance company offers yearly insurance policies to businesses to protect them from loss due to natural disasters (such as a flooding, hail damage etc.). After a natural disaster the business will make a "claim". The policy pays nothing for the first claim. Thereafter, the business is reimbursed a fixed amount of $25,000 for any claim for the rest of the year. The insurance company estimates the number of claims made in a year by their policyholders follows a Poisson distribution with a mean of 1.8. What is the minimum the insurance company should charge for the insurance policy to make an **expected** profit? $\hookrightarrow Y \sim \text{Poisson}(\mu_Y = \lambda = 1.8)$, $P_Y(y) = \dfrac{1.8^y e^{-1.8}}{y!}$, $y = 0, 1, 2, \ldots$

y	x = amount reimbursed
0	0
1	0
2	25000
3	50000

etc...

$$E[X] = \sum_{\text{all } x} x P(X = x)$$

$$= \sum_{y=1}^{\infty} 25000(y-1) P(Y = y)$$

$$= \sum_{y=1}^{\infty} 25000(y-1) \frac{\lambda^y e^{-\lambda}}{y!}$$

$$= \sum_{y=0}^{\infty} 25000(y-1) \frac{\lambda^y e^{-\lambda}}{y!} - 25000(0-1)e^{-\lambda}$$

$$= 25000\, E[Y-1] + 25000 e^{-\lambda}$$

$$= 25000(\lambda - 1 + e^{-\lambda}) = 25000(1.8 - 1 + e^{-1.8})$$

$$= \underbrace{24132.47}_{\text{The minimum they need to charge for the premium to make an expected profit.}}$$

Example D – The Matching Game (again): A room full of ***n*** students all throw their *OneCards* in a hat. Then, they each draw one *OneCard* at random from the hat and hold on to it. If a student selects their own card it is considered a match.

a) What is the probability distribution for the number of matches? What is the probability of at least one match?

$$P(Y \geq 1) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots + (-1)^{n-1}\frac{1}{n!}$$

$$\lim_{n \to \infty} P(Y \geq 1) = 1 - \frac{1}{e}.$$

b) What is the expected number of matches? $\Rightarrow 1$

c) What is the variance for the number of matches? $\Rightarrow 1$

Let the random variable Y denote the number of matches. Consider approximations for the questions above using the Poisson distribution.

Let $X_i = \begin{cases} 1, & \text{if } i\text{th student is a match} \\ 0, & \text{otherwise} \end{cases}, \quad i = 1, \ldots, n.$

Then $Y = X_1 + \cdots + X_n$.

<u>Approximation 1:</u> $Y \simeq \text{Bin}(n, p = 1/n)$.

- $\mu_Y = np = 1$
- $\sigma_Y^2 = np(1-p) = \frac{n-1}{n} \to 1$ as $n \to \infty$.
- $P(Y \geq 1) = 1 - p_Y(0) = 1 - \left(1 - \frac{1}{n}\right)^n$.

$$\lim_{n \to \infty} 1 - \left(1 - \frac{1}{n}\right)^n = 1 - \frac{1}{e} \approx 0.63212.$$

Approximation 2:

$Y \approx \text{Poisson}(\lambda = np = 1)$. (for "large" n)

Then $p_Y(y) \approx \dfrac{e^{-1}}{y!}$, $y = 0, 1, 2, \ldots$

doesn't even depend on n!

- $\mu_Y = 1$
- $\sigma_Y^2 = 1$
- $P(Y \geq 1) = 1 - p_Y(0) = 1 - e^{-1}$

Example ~~D~~ E - Birthday Coincidence: What is the minimum number of people you would need so that the probability that at least two people share the same birthday (a "pair") is at least 0.5? What are your assumptions?

A_k – at least two individuals among k individuals share a birthday (day and year)

Assuming all birthdays are equally likely (and ignoring leap year),

$$P(A_k) = 1 - P(\bar{A}_k) = 1 - \frac{P_k^{365}}{365^k}.$$

You will need a minimum of $k = 23$ so that $P(A_k) > 0.5$.

New Questions:

a) How many "pairs" would you expect from *k* individuals?

b) What is the minimum number of people you would need so that the probability that at least **three** people share the same birthday (a "triple") is at least 0.5?

c) How many "triples" would you expect from *k* individuals?

a) Consider all pairs in the group of k people.

$\Rightarrow n = \binom{k}{2}$ (e.g. If $k = 23$ we have $n = \binom{23}{2} = 253$ pairs.)

Let $p = P(\text{of two people sharing a birthday}) = \frac{1}{365}$.

Let $X_i = \begin{cases} 1, & \text{if pair } i \text{ is a match} \\ 0, & \text{otherwise} \end{cases}$, $i = 1, \ldots, n = \binom{k}{2}$.

Let Y = total # of matches $= X_1 + X_2 + \cdots + X_n$.

$\mu_Y = E[X_1 + \cdots + X_n] = n\left(\frac{1}{365}\right) = \binom{k}{2}\left(\frac{1}{365}\right)$.

Probability Approximations for Y:

$X_i \sim \text{Bernoulli}(p = 1/365)$.

$Y = X_1 + \cdots + X_n$

1) $Y \approx \text{Bin}\left(n = \binom{k}{2},\ p = \frac{1}{365}\right)$

2) $Y \approx \text{Poisson}\left(\mu_Y = \lambda = np = \binom{k}{2}\left(\frac{1}{365}\right)\right)$.

e.g. $k = 23$, $\quad p_Y(y) \approx \dfrac{(253/365)^y\, e^{-253/365}}{y!}$, $y = 0, 1, \ldots$

$P(Y \geq 1) = 1 - p_Y(0) = 1 - e^{-253/365} \approx 0.500002$.

b – c) Consider the number of "triple" matches.

n = # of triples $= \binom{k}{3}$

p = P(any 3 people share same birthday) $= \left(\frac{1}{365}\right)^2$.

Y = total # of triple matches

1) $\approx \text{Bin}\left(n = \binom{k}{3},\ p = \left(\frac{1}{365}\right)^2\right)$

2) $\approx \text{Poisson}\left(\lambda = np = \binom{k}{3}\left(\frac{1}{365}\right)^2\right)$.

$E[Y] = \binom{k}{3}\left(\frac{1}{365}\right)^2$.

Predictions: $k = 106$ (# of respondents to the eclass poll)

1) Pairs: $n = \binom{k}{2} = 5565$

$\lambda = E[\text{pairs}] = \binom{k}{2}\left(\frac{1}{365}\right) = 5565\left(\frac{1}{365}\right) = 15.25$

Actual Observed: 15

$P(Y \geq 1) \approx 1 - e^{-15.25} = 0.9999998.$

$\hookrightarrow Y \approx \text{Poisson}(\lambda = 15.25)$

2) Triples: $n = \binom{k}{3} = 192920$

$\lambda = E[\text{triples}] = \binom{k}{3}\left(\frac{1}{365}\right)^2 = 1.44808$

Actual Observed: 1

$P(Y \geq 1) \approx 1 - e^{-1.44808} = 0.765.$

$\hookrightarrow Y \approx \text{Poisson}(\lambda = 1.44808)$

3) Quadruples: $n = \binom{k}{4} = 4967690$

$\lambda = E[\text{quadruples}] = \binom{k}{4}\left(\frac{1}{365}\right)^3 = 0.10215$

Actual Observed: 0.

$P(Y \geq 1) \approx 1 - e^{-\lambda} = 1 - e^{-4967690/365^3} = 1 - e^{-0.10215}$

$Y \approx \text{Poisson}(\lambda = 0.10215)$

$= 0.097.$