

Stat 394 Probability I

Lecture 6

Richard Li

July 9, 2017

Review: binomial distribution

Theorem

If $X \sim Bin(n, p)$, then as $n \rightarrow \infty$,

$$P(X = i) \rightarrow \frac{e^{-\lambda} \lambda^i}{i!}, \text{ for } i = 0, 1, 2, \dots$$

where $\lambda = np$.

Pf. $P(X = i) = \binom{n}{i} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i}$

$$= \frac{n(n-1)(n-2)\cdots(n-i+1)}{i!} \cdot \frac{\lambda^i}{n^i} \left(1 - \frac{\lambda}{n}\right)^n$$

$$= \frac{\lambda^i}{i!} \cdot \underbrace{\frac{n(n-1)\cdots(n-i+1)}{n^i}}_{A} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{B} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-i}}_{C}$$

$$\lim_{n \rightarrow \infty} A = \lim_{n \rightarrow \infty} \left(\frac{n}{n} \cdot \frac{n-1}{n} \cdot \cdots \cdot \frac{n-i+1}{n} \right) = \lim_{n \rightarrow \infty} \left(1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdots \cdot \frac{\left(1 - \frac{i-1}{n}\right)}{\left(1 - \frac{i-1}{n}\right)} \right) = 1$$

Theorem

$$\lim_{n \rightarrow \infty} B = \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$$

$$\lim_{n \rightarrow \infty} C : \lim_{n \rightarrow \infty} \frac{1}{\left(1 - \frac{\lambda}{n}\right)^i} = 1$$

Poisson distribution

Definition: We say $X \sim \text{Poisson}(\lambda)$ with $\lambda > 0$ if it has the following p.m.f

$$P(X = i) = \frac{e^{-\lambda} \lambda^i}{i!}, \quad \text{for } i = 0, 1, 2, \dots$$

Example: Suppose the number of typos on a single page of my lecture slides has a Poisson distribution with parameter $\lambda = 0.2$. Calculate the probability that there is at least one typo on this page.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - \frac{e^{-0.2} \cdot (0.2)^0}{0!} = 1 - e^{-0.2} \approx 0.181 \end{aligned}$$

Example

Suppose that the probability that a person is killed by lightning in a year is, independently, $1/(500 \text{ million})$. Assume that the US population is 300 million.

- Compute $P(3 \text{ or more people will be killed by lightning next year})$ exactly.

X : # of people killed by lightning.

$$X \sim \text{Bin}(300m, \frac{1}{500m})$$

$$P(Y \geq 3) = 1 - P(X=0) - P(X=1) - P(X=2) = 1 - (1-p)^n - np(1-p)^{n-1}$$

- Approximate the above probability.

$$X \sim \text{Po}\left(\frac{3}{5}\right) \quad (\text{Apprx}), \lambda = \frac{3}{5} \quad - \binom{n}{2} p^2 (1-p)^{n-2}$$

$$\begin{aligned} P(X \geq 3) &= 1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2}{2} e^{-\lambda} \\ &\approx 0.02311529 \end{aligned}$$

$Y \sim \text{Bin}(n, p)$

$E(Y) = np$

Approx by $X \sim \text{Poi}(\lambda)$
 $\lambda = np$

$X \sim \text{Poi}(\lambda)$

$E(X) = \sum_{i=0}^{\infty} i \cdot p(i) = \sum_{i=0}^{\infty} \frac{i \cdot e^{-\lambda} \lambda^i}{i!}$

$= e^{-\lambda} \sum_{i=0}^{\infty} \frac{i \cdot \lambda^i}{i!}$

$= e^{-\lambda} \sum_{i=1}^{\infty} \frac{i \cdot \lambda^i}{i!}$

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

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

$= e^{-\lambda} \cdot \lambda \cdot \sum_{j=0}^{\infty} \frac{\lambda^j}{j!}$

$= e^{-\lambda} \cdot \lambda \cdot e^{\lambda}$

$= \lambda$

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n$$

$$\sum_{i=0}^{\infty} \frac{x^i}{i!}$$

$$= \sum_{i=0}^{\infty} i^2 p(i)$$

$$E(X^2) = \sum_{i=0}^{\infty} \frac{i^2 \cdot e^{-\lambda} \lambda^i}{i!} \text{ Variance}$$

$$= 0 + \frac{1 \cdot e^{-\lambda} \cdot \lambda^1}{1!} + \sum_{i=2}^{\infty} \frac{i^2 e^{-\lambda} \lambda^i}{i!}$$

$$= \lambda e^{-\lambda} + e^{-\lambda} \sum_{i=2}^{\infty} \frac{i \lambda^i}{(i-1)!}$$

$$= \lambda e^{-\lambda} + e^{-\lambda} \sum_{i=2}^{\infty} \frac{(i-1+1) \lambda^i}{(i-1)!}$$

$$= \lambda e^{-\lambda} + e^{-\lambda} \left(\underbrace{\sum_{i=2}^{\infty} \frac{(i-1) \lambda^i}{(i-1)!}}_{\sum_{i=2}^{\infty} \frac{\lambda^i}{(i-2)!}} + \underbrace{\sum_{i=2}^{\infty} \frac{\lambda^i}{(i-1)!}}_{\sum_{i=2}^{\infty} \frac{\lambda^{i-2}}{(i-2)!}} \right)$$

$$\sum_{i=2}^{\infty} \frac{\lambda^i}{(i-2)!} = \lambda^2 \sum_{i=2}^{\infty} \frac{\lambda^{i-2}}{(i-2)!}$$

 $Y \sim \text{Bin}(n, p)$

$$\text{Var}(Y) = np(1-p)$$

$$\lambda = np$$

$$\text{Var}(Y) = \lambda(1 - \frac{\lambda}{n}) = \lambda - \frac{\lambda^2}{n}$$

$$\lim_{n \rightarrow \infty} \text{Var}(Y) = \lambda$$

Variance

$$= \lambda e^{-\lambda} + e^{-\lambda} \left(\lambda^2 \sum_{i=2}^{\infty} \frac{\lambda^{i-2}}{(i-2)!} + \lambda \sum_{i=2}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} \right)$$

$$= \lambda e^{-\lambda} + e^{-\lambda} \left(\lambda^2 \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} + \lambda \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \right)$$

$$= \lambda e^{-\lambda} + e^{-\lambda} \left(\lambda^2 \cdot \underbrace{e^\lambda}_{\lambda^0} + \lambda \left(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} - \frac{\lambda^0}{0!} \right) \right)$$

$$= \lambda e^{-\lambda} + e^{-\lambda} (\lambda^2 e^\lambda + \lambda(e^\lambda - 1))$$

$$= \lambda e^{-\lambda} + e^{-\lambda} (\lambda^2 e^\lambda + \lambda e^\lambda - \lambda)$$

$$= \cancel{\lambda e^{-\lambda}} + \lambda^2 + \lambda - \cancel{\lambda e^{-\lambda}}$$

$$\text{Var}(X) = E(X^2) - E(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

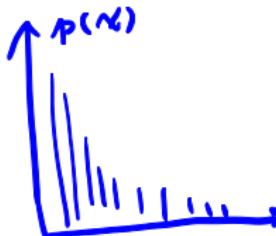
Additional OH

Fri	1-2	7
	2-3	8
	3-4	8

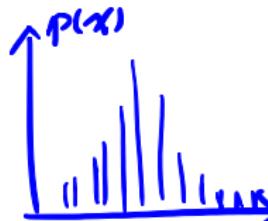
Poisson paradigm

This Friday

Consider n events, with p_i equal to the probability that event i occurs, $i = 1, \dots, n$. If all the p_i are “small” and the trials are either independent or at most “weakly dependent”, then the number of these events that occur approximately has a Poisson distribution with mean $\lambda = \sum_{i=1}^n p_i$.



$$p = 0.01$$

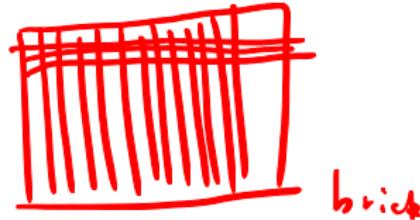


$$p = 0.5$$

- $p \approx 0.5$, moderate n
- $p \rightarrow 0$, large n .

Examples

- the number of winners in a lottery
- the number of people entering a store between 9:00 and 10:00 on weekdays
- the number of wrong numbers dialed each day in Seattle
- the number of raindrops on one brick on the red square at a given second



$P(X=0)$

$P(X=1) \rightarrow P(X \leq 1) = \underline{\underline{m}}$

$P(X=2) \rightarrow P(X \leq 2) = \underline{\underline{m}}$

$P(X=3) \rightarrow \underline{\underline{m}}$

Example

A company which sells flood insurance has three groups of clients.

- The first group, $N_1 = 10000$, is low-risk: each client has probability $p_1 = 0.01\%$ of a flood, independently of others.
- The second group, $N_2 = 1000$ clients, is medium-risk: $k = 5$
 $p_2 = 0.05\%$.
- The third group, $N_3 = 100$ clients, is high-risk, $p_3 = 0.5\%$.

For every flood, a company pays 100,000. How much should it charge its clients so that it does not go bankrupt with probability at least 95%? ~~The third group, $N_3 = 100$ clients is high risk, $p_3 = 0.5\%$~~

X : # of floods

$$X \sim \text{Poi}(\lambda), \quad \lambda = \sum \tilde{p}_i = N_1 p_1 + N_2 p_2 + N_3 p_3 = 2$$

$$\text{Find } k \quad P(X \leq k) \geq 95\% \rightarrow P(100000 \cdot X \leq 100000 \cdot k) \geq 95\%$$

Example: People v. Collins, 1968

Suppose there are $n = \underline{5 \text{ million}}$ couples in the LA area, and the probability that a randomly chosen couple fits the descriptions of the witness is $p = \underline{1/12 \text{ million}}$. One such couple is found. What is the probability that there exist other couple(s) who also fit the descriptions.

$X: \# \text{ of such couples}$

$\cancel{X \sim \text{Bin}(n, p)}$

$X \sim \text{Poi}(np)$

$P(X \geq 2 | X \geq 1) = \frac{P(X \geq 2 \& X \geq 1)}{P(X \geq 1)}$

$= \frac{1 - P(X=0) - P(X=1)}{1 - P(X=0)} = \frac{1 - e^{-\lambda} - \lambda e^{-\lambda}}{1 - e^{-\lambda}} \approx 0.19$

Continuous time interpretation of Poisson r.v.

For events happening in continuous time at a rate of λ , the number of event happening at any interval of length t follows a Poisson distribution with parameter λt if the following (informal) conditions are satisfied:

$$\rightarrow \sigma(h) = f(h) \text{ s.t. } \lim_{h \rightarrow 0} \frac{f(h)}{h} = \sigma$$

1. The probability of exactly one occurrence of the event in a given time interval h is $\lambda h + o(h)$.
 2. The probability of two or more occurrences of the event in a very small time interval is negligible.
 3. The numbers of occurrences of the event in disjoint time intervals are mutually independent.

三

Example

- Number of flying-bomb hits in the south of London during World War II
- Number of wars per year
- Number of earthquakes occurring during a fixed time span

Example

Suppose that the probability that a person is killed by lightning in a year is, independently, $1/(500 \text{ million})$. Assume that the US population is 300 million.

- Approximate $P(\text{two or more people are killed by lightning within the first 6 months of next year})$.

rate : $3/5$ per year

X : # killed within 6 months

$$\Rightarrow X \sim \text{Poi}(\lambda t) \rightarrow \text{Poi}\left(\frac{3}{5} \times \frac{1}{2}\right) \rightarrow \text{Poi}(0.3)$$

$$P(X \geq 2) = 1 - P(X=0) - P(X=1) = \dots$$

Combining multiple distributions together

Note: assume year-to-year independent.

- Approximate P(in exactly 3 of next 10 years exactly 3 people are killed by lightning).

$X: \# \text{ killed within a year} \rightarrow X \sim \text{Poi}\left(\frac{3}{5}\right)$

$\Rightarrow P(X=3) = \frac{\lambda^3 e^{-\lambda}}{3!}$ where $\lambda = \frac{3}{5}$

$Y: \# \text{ years where 3 kills happen.} \hookrightarrow Y \sim \text{Bin}(10, P(X=3)) \Rightarrow P(Y=3) = \binom{10}{3} p^3 (1-p)^{10-3}$

- Compute the expected number of years, among the next 10, in which 2 or more people are killed by lightning.

$Y: \# \text{ such years in the ten}$

$$Y \sim \text{Bin}(10, \underbrace{P(X \geq 2)}_{= 1 - e^{-\lambda} - \lambda e^{-\lambda}}) \Rightarrow E(Y) = 10 P(X \geq 2) = \dots$$

More example

The number of students coming to the Suzzallo and Allen library in a given hour follow a Poisson distribution with parameter λ . Each student choose to stay at Suzzallo library with probability $1/4$, or stay at Allen library with probability $3/4$. What is the distribution of the number of students in Allen library in a given hour?

X : # student visiting S & A in an hour : $X \sim \text{Poi}(\lambda)$

Y : # student stay at A : $\hat{Y} \sim \text{Bin}(x, 3/4)$

$$P(Y=k) = \sum_{x=k}^{\infty} P(X=x) \left[\binom{x}{k} \left(\frac{3}{4}\right)^k \left(\frac{1}{4}\right)^{x-k} \right] \quad \text{given } X=x$$

$$= \sum_{x=k}^{\infty} \frac{\cancel{x!} e^{-\lambda}}{\cancel{x!}} \frac{\cancel{x!}}{k!(x-k)!} \left(\frac{3}{4}\right)^k \left(\frac{1}{4}\right)^{x-k}$$

More example

$$\begin{aligned}
 &= \frac{e^{-\lambda}}{k!} \left(\frac{3}{4}\right)^k \sum_{x=k}^{\infty} \frac{\lambda^x \left(\frac{1}{4}\right)^{x-k}}{(x-k)!} \\
 &= \frac{e^{-\lambda}}{k!} \left(\frac{3}{4}\right)^k \cdot \sum_{x=k}^{\infty} \frac{\lambda^{x-k} \cdot \left(\frac{1}{4}\right)^{x-k}}{(x-k)!} \cdot \lambda^k \\
 &= -1 - \sum_{x=k}^{\infty} \frac{\left(\frac{\lambda}{4}\right)^{x-k}}{(x-k)!} \cdot \lambda^k \quad Y \sim \text{Poi}\left(\frac{3}{4}\lambda\right) \\
 &= -1 - \sum_{x=0}^{\infty} \underbrace{\frac{\left(\frac{\lambda}{4}\right)^x e^{-\frac{\lambda}{4}}}{x!}}_{\lambda^k} \cdot \lambda^k \cdot e^{\frac{\lambda}{4}} \\
 &= \frac{e^{-\lambda}}{k!} \left(\frac{3}{4}\right)^k \cdot \underbrace{1}_{\lambda^k} \cdot \lambda^k \cdot e^{\frac{\lambda}{4}} = \frac{e^{-\frac{3\lambda}{4}} \cdot \left(\frac{3\lambda}{4}\right)^k}{k!}
 \end{aligned}$$

Final thoughts about Poisson distribution

Chap 9.1 Poisson process

- This use of Poisson random variable in continuous time is often times referred to as “Poisson process”.
- It goes much deeper beyond what we have looked at in this course.
 - e.g., how do we characterize the interarrival times between two events? (*You will learn in 395, I think*)
- It has a lot of cool applications in finance, physics, environmental science, etc.
- The example we just talked about is called “thinning/splitting of Poisson process”.