Stat 394 Probability I

Lecture 6

Richard Li

July 9, 2017

Review: binomial distribution

Theorem

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

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

where $\lambda = np$.

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

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

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

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

Theorem

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

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

Poisson distribution

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

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

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 leest one typo on this page.

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

$$= 1 - \frac{e^{-0.2} \cdot (0.2)^0}{0!} = 1 - e^{-0.2} \approx 0.181$$

Example

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

- Compute P(3 or more people will be killed by lightning next year) exactly.

X: # of people killed by lightning.

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

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

$= 0.02311530$

- Approximate the above probability.

$X \sim Poi\left(\frac{3}{5}\right)$ (Approx), $\lambda = \frac{3}{5}$

$P(X \geq 3) = 1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2}{2} e^{-\lambda}$

≈ 0.02311529

Expectation

Variance

Variance

Poisson paradigm

Consider n events, with p_i equal to the probability that event i occurs, $i = 1, ..., n$. 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 $\lambda = \sum_{i=1}^{n} p_i$.

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

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: $p_2 = 0.05\%$.
- 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 is does not go bankrupt with probability at least 95%?The third group, $N_3 = 100$ clients, is high-risk, $p_3 = 0.5\%$

Example: People v. Collins, 1968

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

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:

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.

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(two or more people are killed by lightning within the first 6 months of next year).

Combining multiple distributions together

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

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

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?

More example