From binomial distribution to Poisson distribution

Binomial distribution $f(k,n,p) = P_r(k;n,p) = P_r(X=k)$

(n=1 → special case: Bernoulli distribution) $= C_n^k p^k (1-p)^{n-k}$

$\frac{1}{n}$ evenly divide 1 unit time into n pieces

each piece is $\underbrace{\frac{1}{n}}_{\text{a time bin's length}}$

Assumptions:

① An event occurs in a time bin independendy of evenes occuring in other time bins.

② At most one event can occur in one time bin

$$P_r\{X=k\} = C_n^k p^k (1-p)^{n-k}$$

$EX = np \triangleq \lambda$ average number of events that occur in a fixed interval (unit time)

$$\Downarrow$$

$$p = \frac{\lambda}{n}$$

$$P_r\{X=k\} = C_n^k \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

We need to let $n\to\infty$, then λ can be an average rate applicable at any time point.

$$P_r\{X=k\} = \lim_{n\to\infty} C_n^k \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$\Pr\{X=k\} = \lim_{n\to\infty} \underbrace{\frac{C_n^k}{n^k}} \lambda^k \cdot \underbrace{\left(1-\frac{\lambda}{n}\right)^n}_{e^{-\lambda}} \cdot \underbrace{\left(1-\frac{\lambda}{n}\right)^{-k}}_{1}$$

$$\lim_{n\to\infty} \frac{C_n^k}{n^k} = \frac{\overbrace{n(n-1)\cdots(n-k+1)}^{\approx 1}}{k!\cdot n^k} = \frac{1}{k!}$$

$$\Pr\{X=k\} = \frac{\lambda^k e^{-\lambda}}{k!}$$ Poisson distribution

We can define λ over any unit time, that is, the unit of λ can be per day, per year, or per second.

Actually, we can rewrite this equation as

$$\Pr\{k,\lambda\} = \frac{\lambda^k e^{-\lambda}}{k!},$$ then we have

$$\Pr\{k,\lambda t\} = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$ → This is the probability of k events occuring in time t, equivalent to resetting t as the unit time.

We can easily derive this equation by repeating the process from binomial distribution to Poisson distribution by setting the unit time as t instead of 1.

The time interval between two consecutive events in a Poisson distribution follows the exponential distribution.

Proof: $\Pr\{0 \text{ events in } (0,x]\} = \Pr\{0, \lambda x\} = e^{-\lambda x}$

$\Pr\{\text{at least 1 event in } (0,x]\} = 1 - e^{-\lambda x}$

Let Y be the time interval between first and second events.

$f(y) \triangleq$ the probability density function of Y. → Can be obtained by differentiating the cumulative distribution of Y.

$$f(y) = \frac{d\Pr\{Y>y\}}{dy} = \frac{d[1-\Pr\{Y\le y\}]}{dy} = \frac{d[1-\Pr\{\text{at least 1 event in } (0,x]\}]}{dy}$$

$$f(y) = \underline{\lambda e^{-\lambda y}}$$

↳ exponential distribution

Add up multiple Poisson distribution results in a new Poisson distribution with a mean equal to the sum of the means of individual distributions.

Let X and Y be two independent Poisson random variables with means λ and μ, respectively.

$Z = X + Y$ is also a Poisson random variable with mean $\lambda + \mu$.

$$P_r\{Z = k\} = \frac{(\lambda+\mu)^k e^{-(\lambda+\mu)}}{k!}$$

Adding up an infinite number of identical Poisson distributions results in a Guassian distribution.

This conclusion comes from the central limit theorem. (CLT)