Distributions

For Discrete Random Variable

Binomial / Bernoulli

Shape

- $p = q$ = symmetric
- $p < q$ +vely skewed
- $p > q$ -vely skewed

- conditions:
① Repeated experiments for fixed times
② Only 2 outcomes possible, Success and Failure
③ p, q are constant
④ Independant Trials

- $P(X = r) = {}^nc_r \, p^r q^{n-r}$

X → success

r → Desired Number of times success happens

n = no of Trials
r = no of success
p = prob of success
q = prob of failure

- parameter : n, p
- mean = np
- variance = npq

$b(x; n, p)$

Approximation of Binomial to Poisson

① if $\underset{\geq 20}{n \to \infty}$, $\underset{p < 0.05}{p \to 0}$

$\mu = np$

Hypergeometric

- Conditions:
① Trials are not independant
② p, q are not constant

N = objects in pop
n = objects in sample
k = No of success in pop
$N-k$ = No of failure in pop
x = No of success in sample
$n-k$ = No of failure in sample

- $f(x; N; n; k)$

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

- $p = \frac{k}{N}$; prob of success

- Mean = $n \times \frac{k}{N}$

- var = $n \cdot \frac{k}{N}\left(1 - \frac{k}{N}\right)\left(\frac{N-n}{N-1}\right)$

- Approximation of hypergeometric to binomial:

- when $\frac{n}{N} \leq 0.05$

→ p is required ⇒ $p = \frac{k}{N}$

→ $\mu = np = n \times \overset{p}{\frac{k}{N}}$

→ $\sigma^2 = npq = n \times \underbrace{\frac{k}{N}}_{p} \times \underbrace{\left(1 - \frac{k}{N}\right)}_{q}$

$\frac{N-n}{N-1}$ is negligible when n is small relative to N

Poisson

- Used for estimating the number of occurrences over a specified interval of time/space
  e.g: No of car arrivals in an hour

- Conditions:
① Prob of an occurrence is the same for any 2 intervals of equal length
② Occurrences / Non-occurrences are independent of Non/ocurrences in any other interval

- $P(X=x) = \dfrac{\mu^x e^{-\mu}}{x!}$; $\mu = \lambda t$, $x = 0, 1, 2 \ldots$

  $= \dfrac{(\lambda t)^x \cdot e^{-\lambda t}}{x!}$; λ = avg no. of occurrences per time, distance, area...
  t = specific time, distance, area of volume
  x = No of occurrences

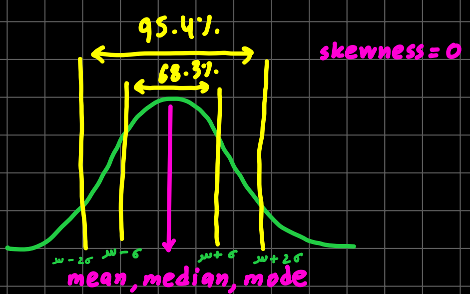
- parameter = $\mu = \lambda t$
- mean, var = μ

Continuous Random variables

Normal

- Prob Density Function:

- $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ · μ = mean
  σ = std



- parameters = μ, σ , $N(x; \mu, \sigma)$
- Range of x : $(-\infty, \infty)$
- $\mu \rightarrow$ shape parameter
- $\sigma \rightarrow$ location parameter
- Dispersion increase $\rightarrow$ curve flat, wider

X ―――――――――――――――― X

- When $\mu = 0, \sigma = 1$, random variable has standard normal probability distribution

- Standard normal Density Function

- $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$

- $z = \dfrac{x - \mu}{\sigma}$

$\rightarrow$ z is interpreted as the number of standard deviations that the normal random variable 'x' is from its mean 'μ'.