

1. Moment Generating Function

$$M_X(t) = E(e^{tx})$$

$$\text{continuous: } M_X(t) = \int_{-\infty}^{\infty} e^{tx} P_X(x) dx$$

$$\text{discrete: } M_X(t) = \sum e^{tx} P_X(x)$$

$$\frac{dM_X(t)}{dt} = E\left(\frac{d}{dt} e^{tx}\right) = E(xe^{tx}) \Rightarrow M_X^{(n)}(t) = E(x^n e^{tx})$$

Properties:

- $M_X(t) = (1 - p + pe^t)^n$ for binomial(n,p)
- $M_X(t) = E\left(\sum_{n=0}^{\infty} \frac{x^n t^n}{n!}\right)$ according to Taylor Series
 $M_X'(t) = E(0 + x + x^2 t + \frac{x^3 t^2}{2!} + \dots) = E\left(\sum_{n=0}^{\infty} \frac{x^{n+1} t^n}{n!}\right)$
 $\Rightarrow M_X^{(k)}(t) = E\left(\sum_{n=0}^{\infty} \frac{x^{n+k} t^n}{n!}\right)$
- $M_X^{(n)}(0) = E(X^n)$
 $\Rightarrow \text{VAR}[x] = E[x^2] - E^2[x] = M_X''(0) - [M_X'(0)]^2$
- $M_X(t) = M_Y(t) \Rightarrow XY$ has same distribution
- $M_X(0) = 1$
- $M_{ax+b}(t) = M_X(at)e^{bt}$ Proof:
 $M_{ax+b}(t) = E(e^{t(ax+b)}) = E(e^{axt} e^{bt}) = E(e^{axt}) e^{bt} = M_X(at)e^{bt}$
- $M_{X+Y}(bt) = M_X(t)M_Y(t)$ (X, Y independent)

2. Gama Function $\Gamma(n) = \int_0^{\infty} u^{n-1} e^{-u} du = (n-1)!$

Proof:

$$\begin{aligned} 1. \int_0^{\infty} e^{-x} dx &= -e^{-x} \Big|_0^{\infty} = 1 \\ 2. \Gamma(n) &= - \int_0^{\infty} x^{n-1} d e^{-x} \\ &= -(x^{n-1} e^{-x} \Big|_0^{\infty} - \int_0^{\infty} e^{-x} (n-1) x^{n-2} dx) \\ &= 0 + \int_0^{\infty} e^{-x} (n-1) x^{n-2} dx \\ &= (n-1) \int_0^{\infty} e^{-x} x^{n-2} dx \\ &= (n-1) \Gamma(n-1) \end{aligned}$$

Use Induction to prove the formula

3. Basic probability property

- Basic Properties
 - $P(E) = \frac{n(E)}{n(S)}, \in [0, 1]$
 - $P(\phi) = 0$
 - $A \subseteq B, P(A) \leq P(B)$
 - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Conditional Probability and Bayles Theorem
 - $P(A|B) = \frac{P(A \cap B)}{P(B)}$
 - $P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')P(A')}$
- PDF and CDF

PDF: probability distribution function $\Rightarrow P_X(x)$
 CDF: cummulative distribution function $\Rightarrow F_X(x) = \phi_X(x) = N_X(x) = P(X < x)$
- Expectation
 - $E(X) = \sum x P_X(x)$ or $\int_{-\infty}^{\infty} x P_X(x)$
 - $E(c) = c$
 - $E(aX) = aE(X)$
 - $E(X + Y) = E(X) + E(Y)$
 - $E(X) = M_X'(0)$
- Variance and Standard Deviation
 - $\text{VAR}(X) = E(X^2) - E^2(X) = \sum [(x - E(X))]^2 P_X(x) = E[(x - \mu)^2] = \sigma^2$
 - $\text{VAR}(c) = 0$
 - $\text{VAR}(aX) = a^2 \text{VAR}(x)$

$$\begin{aligned} 4. \text{VAR}(X \pm Y) &= \text{VAR}(X) + \text{VAR}(Y) \pm 2\text{COV}(x, y), \\ \text{COV}(x, y) &= E(XY) - E(X)E(Y) \end{aligned}$$

(f) z-score

$Z = \frac{x - \mu}{\sigma}$, measures the distance of x from expected value in standard units.

4. Discrete Distributions

(a) Binomial Distribution

DEF: n time Bernoulli trials combined. probability of success and fail is (p, 1-p). Probability of success remains the same through the trails. X is the r.v of success times.

Note as $X \sim B(n, p)$

$$P_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$M_X(t) = (1 - p + pe^t)^n$$

$$E(X) = np, \text{VAR}(X) = np(1-p)$$

(b) Hyper Geometric Distribution

DEF: A sample of size n taken from a finite population of size N. The population has a subgroup of size $r \geq n$ that is of interest. x is the number of members of the subgroup taken.

Note as $X \sim H(N, n, r)$

$$P_X(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}$$

$$E(X) = n \frac{r}{N}$$

$$\text{VAR}(X) = n \frac{r}{N} \left(1 - \frac{r}{N}\right) \left(\frac{N-n}{N-1}\right)$$

When $x \rightarrow \infty$, $H(N, n, r) \rightarrow B(n, \frac{r}{N})$. H samples without replacement while B samples with replacement.

(c) Poisson Process

DEF: model the number of random occurrence of some phenomenon in specific unit of space or time.

$$P_X(x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

λ : average arrival in given time or space

$$E(X) = \lambda, \text{VAR}(X) = \lambda$$

Poisson simulates Binomial when n is large. usually when $\lambda = np < 10$, we use poisson as an approximation of Binomial.

(d) Geometric Distribution

DEF: number of trials to get the first success in a sequence of Bernoulli trials where p is the success probability.

i. X is the r.v of number of total trials (x includes the first success)

$$P_X(x) = (1-p)^{x-1} p, x = 1, 2, 3 \dots$$

$$E(X) = 1/p, \text{VAR}(X) = \frac{1-p}{p^2}$$

ii. X is the r.v of number of failed trials (x excludes the first success)

$$P_X(x) = (1-p)^x p, x = 0, 1, 2, \dots$$

$$M_X(t) = \frac{p}{1 - e^t(1-p)}$$

$$E(X) = \frac{1-p}{p}, \text{VAR}(X) = \frac{1-p}{p^2}$$

(e) Negative Binomial Distribution

DEF: X is the r.v of number of trials need to observe the r^{th} success in a sequence of Bernoulli trails where p is the success probability.

$$P_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, x = r, r+1, r+2 \dots$$

$$M_X(t) = \left(\frac{p}{1 - e^t(1-p)}\right)^r$$

$$E(X) = \frac{r}{p}, \text{VAR}(X) = \frac{r(1-p)}{p^2}$$

Alternatively, X is the r.v of failures before the r^{th} success:

$$P_X(x) = \binom{x+r-1}{r-1} p^r (1-p)^x, x = 0, 1, 2, \dots$$

$$M_X(t) = \left(\frac{1-p}{1-pe^t} \right)^r$$

$$E(X) = \frac{r(1-p)}{p}, \text{VAR}(X) = \frac{r(1-p)}{p^2}$$

5. Chebychev's Theorem

$$P(\mu - k\sigma \leq x \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

6. Continuous Random Variable

(a) Basic Properties

$$1. \text{pdf} : f_X(x) \geq 0$$

$$2. \text{cdf} : F_X(x) = \int_{-\infty}^{\infty} f_X(x) dx$$

$$3. \int_{-\infty}^{\infty} f_X(x) dx = 1$$

$$4. f_X(x) = \frac{d}{dx} F_X(x)$$

(b) Mean, Medium and Variance

$$\text{mean: } \mu = E(x) = \int_{-\infty}^{\infty} x f_X(x) dx$$

$$\text{medium m: solve function } \int_{-\infty}^m f_X(x) dx = \frac{1}{2}$$

Variance:

$$\text{VAR}(X) = E(X^2) - E^2(X) = \int_{-\infty}^{\infty} [x - E(X)] f_X(x) dx$$

7. Continuous Distributions

(a) Uniform(rectangle) Distribution

$$f_X(x) = \frac{1}{b-a}, a \leq x \leq b$$

$$F_X(x) = \frac{x-a}{b-a}, a \leq x \leq b$$

$$E(X) = \frac{a+b}{2}, \text{VAR}(X) = \frac{(b-a)^2}{12}$$

(b) Exponential Distribution

$$f_X(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, x \geq 0$$

$$F_X(x) = 1 - e^{-\frac{x}{\theta}}, x \geq 0$$

$$M_X(t) = \frac{1}{1-\theta t}$$

$$E(X) = \theta, \text{VAR}(X) = \theta^2$$

Note as $X \sim \exp(\theta)$

(c) Gamma Distribution

$$f_X(x) = \frac{1}{\Gamma(n)\theta^n} x^{n-1} e^{-\frac{x}{\theta}}, x \geq 0$$

$$M_X(t) = \frac{1}{(1-\theta t)^n}$$

$$E(X) = n\theta, \text{VAR}(X) = n\theta^2$$

Note as $X \sim \Gamma(n, \theta)$

Exponential dist. is a special case of Gamma dist where $n = 1$. Gamma Distribution can be viewed as a sum of Exponential dists.

$$X \sim \Gamma(n, \theta) \Leftrightarrow X = \sum_{i=1}^n X_i, X_i \sim \exp(\theta)$$