

1. Moment Generating Function

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

Properties:

- (a) $M_X^{(n)}(0) = E(X^n)$
- (b) $M_X(t) = M_Y(t) \Rightarrow XY$ has same distribution
- (c) $M_X(0) = 1$
- (d) $M_{ax+b}(t) = M_X(at)e^{bt}$
- (e) $M_{X+Y}(t) = M_X(t)M_Y(t)$ (X, Y independent)

2. Gamma Function

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

$$\Gamma(n) = (n-1)\Gamma(n-1)$$

$$\int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

3. Basic probability property

(a) Conditional Probability and Bayes Theorem

$$1. P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$2. P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')P(A')}$$

(b) Expectation

$$1. E(X) = \sum x P_X(x) \text{ or } \int_{-\infty}^{\infty} x f_X(x)$$

$$2. E(c) = c$$

$$3. E(aX) = aE(X)$$

$$4. E(X+Y) = E(X) + E(Y)$$

$$5. E(X) = M'_X(0)$$

(c) Variance and Standard Deviation

$$1. VAR(X) = E(X^2) - E^2(X) = \sum [(x - E(X))]^2 P_X(x) = E[(x - \mu)^2] = \sigma^2$$

$$2. VAR(c) = 0$$

$$3. VAR(aX) = a^2 VAR(X)$$

$$4. VAR(X \pm Y) = VAR(X) + VAR(Y) \pm 2COV(X, Y), \\ COV(X, Y) = E(XY) - E(X)E(Y)$$

(d) z-score

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

4. Discrete Distributions

(a) Binomial Distribution

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

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

$$M_X(t) = (1-p + pe^t)^n, E(X) = np, VAR(X) = np(1-p)$$

(b) Hyper Geometric Distribution

Note as $X \sim H(N, n, r)$, N:total size, n:total pick, r:size of special subgroup

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

$$E(X) = n \frac{r}{N}, VAR(X) = n \frac{r}{N} (1 - \frac{r}{N}) (\frac{N-n}{N-1})$$

(c) Poisson Process

$$P_X(x) = \frac{e^{-\lambda} \lambda^x}{x!}, E(X) = \lambda, VAR(X) = \lambda$$

λ : average arrival in given time or space

When $\lambda = np < 10$, poisson approximates Binomial.

(d) Geometric Distribution

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

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

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

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

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

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

(e) Negative Binomial Distribution

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

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

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

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

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

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

5. Chebychev's Theorem

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

6. Continuous Distributions

(a) Uniform(rectangle) Distribution

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

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

(b) Exponential Distribution

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

$$M_X(t) = \frac{1}{1-\theta t}, E(X) = \theta, VAR(X) = \theta^2$$

Note as $X \sim \exp(\theta)$, In terms of λ : $\lambda = \frac{1}{\theta}$

(c) Gamma Distribution

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

$$M_X(t) = \frac{1}{(1-\theta t)^n}, E(X) = n\theta, VAR(X) = n\theta^2$$

In terms of α, β : $\alpha = n, \beta = \theta$

(d) Normal Distribution

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, E(X) = \mu, VAR(X) = \sigma^2$$

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

(e) Log-normal Distribution:

$$Y = e^X, X \sim N(\mu, \sigma^2), f_Y(y) = \frac{1}{\sigma y \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\ln y - \mu}{\sigma})^2}, y \geq 0$$

$$E(Y) = e^{\mu + \frac{\sigma^2}{2}}, VAR(Y) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$$

(f) Pareto Distribution:

$$f_X(x) = \frac{k}{\beta} (\frac{\beta}{x})^{k+1}, F_X(x) = 1 - (\frac{\beta}{x})^k, k > 2, x \geq \beta > 0$$

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

In terms of α, β : $\alpha = k, \beta = \beta$

(g) Weibull Distribution:

$$f_X(x) = k\lambda x^{k-1} e^{-\lambda k x^k}, F_X(x) = 1 - e^{-\lambda k x^k}$$

$$E(X) = \frac{\Gamma(1+\frac{1}{k})}{\lambda^{\frac{1}{k}}}, VAR(X) = \frac{1}{\lambda^{\frac{2}{k}}} [\Gamma(1+\frac{2}{k}) - \Gamma^2(1+\frac{1}{k})]$$

in terms of α, β : $k = \alpha, \lambda = \beta$

Note: $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

(h) Beta Distribution

$$f_X(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}, x \in (0, 1), a > 0, b > 0$$

$$E(X) = \frac{a}{a+b}, VAR(X) = \frac{ab}{(a+b)^2(a+b+1)}$$

(i) Chi-Square Distribution

$$f_X(x) = \frac{1}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}, x > 0$$

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

Special case of Gamma Dist. of $\alpha = \frac{k}{2}, \beta = 2$

$$\chi^2 = \sum_{i=1}^k Z_i^2 \text{ (sum of } k \text{ squared standard normal)}$$

7. Central Limit Theorem

Let $\{X_1, X_2, \dots, X_n\}$ be the indep. r.v. with same distribution and mean μ and std. deviation σ . If n is large $n \geq 30$,

$$S = X_1 + X_2 + \dots + X_n \Rightarrow S \sim N(n\mu, n\sigma^2)$$

$$\text{OR } S' = \frac{X_1 + X_2 + \dots + X_n}{n} \Rightarrow S' \sim N(\mu, \frac{\sigma^2}{n})$$

8. Finding PDF for $Y = g(X)$

If $g(X)$ is strictly increasing or decreasing

$$f_Y(y) = f_X(h(y)) \cdot |h'(y)|, h(y) = g^{-1}(x)$$

9. Multivariable Distributions

(a) Marginal Distribution

$$p_X(x) = \sum_y p(x, y), p_Y(y) = \sum_x p(x, y)$$

$$f_X(x) = \int_{-\infty}^{\infty} p(x, y) dy, f_Y(y) = \int_{-\infty}^{\infty} p(x, y) dx$$

(b) Conditional Distribution

$$P(X = x|Y = y) = \frac{p(x, y)}{p_Y(y)}, P(Y = y|X = x) = \frac{p(x, y)}{p_X(x)}$$

$$f(x|Y = y) = f(x|y) = \frac{f(x, y)}{f_Y(y)}$$

$$f(y|X = x) = f(y|x) = \frac{f(x, y)}{f_X(x)}$$

(c) Conditional Expected Value

$$E(y|X = x) = \sum_y y \cdot p(y|x), E(x|Y = y) = \sum_x x \cdot p(x|y)$$

$$E(y|X = x) = \int_{-\infty}^{\infty} y \cdot f(y|x) dy$$

$$E(x|Y = y) = \int_{-\infty}^{\infty} x \cdot f(x|y) dx$$

(d) Independence

$$p(x, y) = p_X(x) \cdot p_Y(y)$$

$$f(x|y) = f_X(x), f(y|x) = f_Y(y)$$

(e) Covariance

$$\text{COV}(x, y) = E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - E(X)E(Y)$$

Properties:

$$\text{i. } \text{COV}(X, Y) = \text{COV}(Y, X)$$

$$\text{ii. } \text{COV}(X, X) = \text{VAR}(X)$$

$$\text{iii. } \text{COV}(X, K) = 0, K \text{ is constant}$$

$$\text{iv. } \text{COV}(aX, bY) = ab \cdot \text{COV}(X, Y)$$

$$\text{v. } \text{COV}(X, Y + Z) = \text{COV}(X, Y) + \text{COV}(X, Z)$$

$$\text{vi. If } X, Y \text{ are independent, } \text{COV}(X, Y) = 0$$

(f) Correlation Coefficient ρ

$$\rho = \rho_{X,Y} = \frac{\text{COV}(X, Y)}{\sigma_X \cdot \sigma_Y}, \in [-1, 1]$$

$|\rho_{X,Y}|$ close to 1 : more related

$\rho > 0$ positively related, $\rho < 0$ negatively related

(g) Moment Generating Function of Joint Distribution

$$M_{X,Y}(t_1, t_2) = E(e^{t_1 X + t_2 Y})$$

$$E(X) = \frac{d}{dt_1} M_{X,Y}(t_1, t_2)|_{t_1=t_2=0}, E(Y) = \frac{d}{dt_2} M_{X,Y}(t_1, t_2)|_{t_1=t_2=0}$$

$$E(X^m Y^n) = \frac{\partial^{m+n}}{\partial t_1^m \partial t_2^n} M_{X,Y}(t_1, t_2)|_{t_1=t_2=0}$$

(h) Multinomial Distribution

Counting Partitions:

$$\binom{n}{n_1, n_2, \dots, n_k} = \frac{n!}{n_1! n_2! \dots n_k!}$$

Suppose that a random experiment has K mutually exclusive outcomes E_1, E_2, \dots, E_k , with $P(E_i) = p_i, i \in [1, k]$. Suppose you repeat this experiment in n independent trials. Then:

$$P(X_1 = E_1, \dots, X_k = E_k) = \binom{n}{n_1, n_2, \dots, n_k} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$$