1. Moment Generating Function
   $M_X(t) = E(e^{tx})$
   Properties:
   (a) $M_x(t) = (1-p+pe^t)^n$ for binomial(n,p)
   (b) $M_x^{(n)}(0) = E(X^n)$
   $\Rightarrow VAR[x] = E[x^2] - E^2[x] = M_x''(0) - [M_x'(0)]^2$
   (c) $M_x(t) = M_y(t) \Rightarrow XY$ has same distribution
   (d) $M_x(0) = 1$
   (e) $M_{ax+b}(t) = M_x(at)e^{bt}$
   (f) $M_{X+Y}(t) = M_X(t)M_Y(t)$ (X, Y independent)
2. Gama Function $\Gamma(n) = \int_0^\infty u^{n-1}e^{-u}du = (n-1)!$
3. Basic probability property
   (a) Conditional Probability and Bayles Theorem
   1. $P(A|B) = \frac{P(A\cup B)}{P(B)}$
   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. $E(X) = \sum xP_X(x)$ or $\int_{-\infty}^{\infty} xf_X(x)$
   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. $VAR(X) = E(X^2) - E^2(X) = \sum[(x-E(X)]P_X(x) = E[(x-\mu)^2] = \sigma^2$
   2. $VAR(c) = 0$
   3. $VAR(aX) = a^2VAR(x)$
   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 trails where p is the success probability.
   $P_X(x) = \binom{x-1}{r-1}p^r(1-p)^{x-r}, x = r, r+1, r+2\dots$
   $M_X(t) = (\frac{p}{1-e^t(1-p)})^r, E(X) = \frac{r}{p}, VAR(X) = \frac{r(1-p)}{p^2}$
   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
   $P(\mu - k\sigma \le x \le \mu + k\sigma) \ge 1 - \frac{1}{k^2}$
6. Continuous Distributions
   (a) Uniform(rectangle) Distribution
   $f_X(x) = \frac{1}{b-a}, (a \le x \le b); F_X(x) = \frac{x-a}{b-a}, a \le x \le b$
   $E(X) = \frac{a+b}{2}, VAR(X) = \frac{(b-a)^2}{12}$
   (b) Exponential Distribution
   $f_X(x) = \frac{1}{\theta}e^{-\frac{x}{\theta}}, (x \ge 0); F_X(x) = 1 - e^{-\frac{x}{\theta}}, x \ge 0$
   $M_X(t) = \frac{1}{1-\theta t}, E(X) = \theta, VAR(X) = \theta^2$
   Note as $X \sim exp(\theta)$, In terms of λ: $\lambda = \frac{1}{\theta}$
   (c) Gamma Distribution
   $f_X(x) = \frac{1}{\Gamma(n)\theta^n}x^{n-1}e^{-\frac{x}{\theta}}, x \ge 0$
   $M_X(t) = \frac{1}{(1-\theta t)^n}, E(X) = n\theta, VAR(X) = n\theta^2$
   In terms of α, β : $\alpha = n, \beta = \frac{1}{\theta}$
7. Central Limit Theorem
   Let $\{X_1, X_2, \dots, X_n\}$ be the independent random variable all of which have the same distribution and mean μ and standard deviation σ. If n is large $n \ge 30$, then $S = X_1 + X_2 + \cdots + X_n$ will be approximately normal with mean $n\mu$ and variance $n\sigma^2$. ($S \sim N(n\mu, n\sigma^2)$)
   OR $S' = \frac{X_1+X_2+\cdots+X_n}{n}$ will be approximately normal with mean μ, variance $\frac{\sigma^2}{n}$.($S \sim N(\mu, \frac{\sigma^2}{n})$)
8. Finding CDF for $Y = g(X)$
   $g(X)$ is strictly increasing or decreasing, $f_Y(y) = \frac{d}{dx}F_X(h(y)) \cdot |h'(y)|$