Complement	A'
Intersection	$A \cap B$
Union	$A \cup B$

$P(A \cup B \cup C) =$
$P(A)+P(B)+P(C)-P(A\cap B)-P(B\cap C)-P(C\cap A)+P(A\cap B\cap C)$
$P(A|B) = \frac{P(A\cap B)}{P(B)}$
$P(A \cap B) = P(A)\cdot P(B|A) = P(B)\cdot P(A|B)$
$P(A) = P(A\cap B) + P(A \cap B') = P(B)\cdot P(A|B) + P(B')\cdot P(A|B')$
$$P(B_k|A) = \frac{P(B_k)\cdot P(A|B_k)}{\sum_{i=1}^{n} P(B_i)\cdot P(A|B_1)} = \frac{P(B_k)\cdot P(A|B_k)}{P(A)}$$

A and B independent $P(B|A) = P(B)$ $P(A|B) = P(A)$
$P(A\cap B) = P(A)\cdot P(B)$

${}_nP_r = \frac{n!}{(n-r)!}$ ${}_nC_r = \binom{n}{k} = \frac{{}_nP_r}{r!} = \frac{n!}{r!(n-r)!}$

	Order important	Order not important
replace	n^r	${}_{n+r-1}C_r$
no replace	${}_nP_r$	${}_nC_r$

$\binom{n}{k} = \binom{n}{n-k}$ $\binom{n}{0} = \binom{n}{n} = 1$ $\binom{n}{1} = \binom{n}{n-1} = n$

$E(x) = \mu_x = \sum x\cdot f(x) = \int x \cdot f(x)dx$
$Var(x) = \sigma_x^2 = \sum (x-\mu_x)^2 \cdot f(x)$
$= \int (x-\mu_x)^2 \cdot f(x)dx = \int x^2 \cdot f(x)dx - \mu_x^2$
$(x+y)^n = \sum_{n}^{k=0} \binom{n}{k} x^k y^{n-k}$ $\sum_{n}^{k=0} \binom{n}{k} p^k (1-p)^{n-k} = 1$
$\sigma_x^2 = E(x^2) - E(x)^2$ $E(aX+bY) = aE(X) + bE(Y)$
$Var(X) = E((X-\mu)^2)$ $Var(aX+b) = a^2 Var(X)$
$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$ $\sum_{n=1}^{\infty} r^n = \frac{r}{1-r}$ $(|r|<1)$ $\sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^{\lambda}$
When start from $n > 1$, always reduce to $n = 1$.

Moment Generation Function

$M_X(t) = E(e^{tx}) = \sum e^{tx} f(x) = \int e^{tx} f(x)$ $M_X'(0) = E(X)$
$M_X^{(k)}(0) = E(X^k)$ $Y = aX + b$ $M_Y(t) = e^{bt} M_X(at)$
$(\ln M_X(t))'|_{t=0} = E(X) = \mu_x$
$(\ln M_X(t))''|_{t=0} = E(X^2) - E^2(X) = \sigma_x^2$
$M_Y(t) = \sum_{k=o}^{\infty} \frac{t^k}{k!} M_Y^{(k)}(0) = \sum_{k=o}^{\infty} \frac{t^k}{k!} E(Y^k)$

Binomial Distribution

The number of trials, n, is fixed. The probability of success, p, is same. The trials are independent. X = number of successes.
$P(X=k) = \binom{n}{k}\cdot p^k \cdot (1-p)^{n-k} ={}_n C_k \cdot p^k \cdot (1-p)^{n-k}$
$E(x) = n\cdot p$ $Var(x) = n\cdot p \cdot (1-p)$

Geometric Distribution

X = the number of independent trials until the first success.
$P(X=x) = (1-p)^{x-1}\cdot p$ $E(x) = \frac{1}{p}$ $\sigma^2 = \frac{1-p}{p^2}$
$P(X > a) = \sum_{k=a+1}^{\infty} (1-p)^{k-1} p = \frac{(1-p)^a p}{1-(1-p)} = (1-p)^a$
$P(X > a+b | X > a) = \frac{P(X>a+b \cap X>a)}{P(X>a)} = (1-p)^b = P(X>b)$

Negative Binomial Distribution

X = the number of independent trials until the k success.
$P(X=x) = \binom{x-1}{k-1}\cdot p^k \cdot (1-p)^{x-k}$ $E(x) = \frac{k}{p}$
$V(x) = \frac{k\cdot(1-p)}{p^2}$

Hypergeometric Distribution

N=population size. S=number of successes. n=sample size.
X=number of successes in the sample without replacement.
$P(X=x) = \frac{\binom{S}{x}\cdot\binom{N-S}{n-x}}{\binom{N}{n}} = \frac{{}_SC_x \cdot {}_{N-S}C_{n-x}}{{}_NC_n}$

Multinomial Distribution

Fixed, n, trail. k possible outcomes, with probabilities
$P_1, P_2, P_3, \cdots, P_k.$ $\sum_{i=1}^{k} P_i = 1.$
Trails are independent. X are numbers of times of outcome.
$P(X_1 = x_1, \cdots, X_k = x_k) = \frac{n!}{x_1!x_2!\cdots x-k!} P_1^{x_1} P_2^{x_2} \cdots P_k^{x_k}$

Poisson Distribution

X = the number of occurrences of a particular event in an interval of time or space. $\lambda = n\cdot p$.
$P(X=x) = \frac{\lambda^x \cdot e^{-\lambda}}{x!}$ $E(X) = \lambda$ $\sigma^2 = \lambda$
Binomial probabilities can be approximated by Poisson probabilities.

Uniform Distribution

Uniform Distribution over an interval $[a,b]$, p.d.f $f(x) = \frac{1}{b-a}$.
$P(c \le x \le d) = \frac{d-c}{b-a}$ $E(x) = \frac{a+b}{2}$ $Var(x) = \frac{(b-a)^2}{12}$

Exponential Distribution

$f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & \text{for } x \ge 0 \\ 0 & \text{otherwise} \end{cases}$ $f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \ge 0 \\ 0 & \text{otherwise} \end{cases}$
$E(X) = \mu = \theta = \frac{1}{\lambda}$ $Var(X) = \sigma^2 = \theta^2 = \frac{1}{\lambda^2}$
$CDF = 1 - e^{-\lambda x}$ $M(t) = \frac{\lambda}{\lambda - t}$

Gamma Distribution

$f(x) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\theta}}$ $f(x) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$
$E(x) = \alpha\theta$ $E(x) = \frac{\alpha}{\lambda}$
$Var(x) = \alpha\theta^2$ $Var(x) = \frac{\alpha}{\lambda^2}$
$\mathbf{\Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u} du}$ $\Gamma(\frac{1}{2}) = \sqrt{\pi}$
$\Gamma(x) = (x-1)\Gamma(x-1)$ $\Gamma(n) = (n-1)!$

Multivariate Distributions

$P((x,y)\in A) = \sum\sum p(x,y) = \iint_A f(x,y)dxdy$
$p_X(x) = \sum_y p(x,y) = \int_{-\infty}^{\infty} f(x,y)dy$
$E(g(x,y)) = \sum\sum g(x,y)\cdot p(x,y) = \iint g(x,y)\cdot p(x,y)dxdy$
If X and Y are independent, $p(x,y) = p_X(x,y)\cdot p_Y(x,y)$
$\sigma_{XY} = Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)] = E(XY) - \mu_x\mu_y$
Covariance $Cov(X,X) = Var(X)$
$Cov(aX+b, Y) = aCov(X,Y)$
$Cov(aX+bY, cX+dY) =$
$acVar(X) + (ad+bc)Cov(X,Y) + bdVar(Y)$
$Var(aX+bY) = a^2Var(X) + 2abCov(X,Y) + b^2Var(Y)$

Correlation coefficient
$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = E(\frac{X-\mu_X}{\sigma_X}, \frac{Y-\mu_Y}{\sigma_Y})$
If X and Y are independent, $Cov(X,Y) = \sigma_{XY} = \rho_{XY} = 0$
If $U = a_0 + a_1X_1 + a_2X_2 + \cdots + a_nX_n$
$E(U) = a_0 + a_1E(X_1) + a_2E(X_2) + \cdots + a_nE(X_n)$
$Var(U) = \sum_{i=1}^{n} a_i^2 Var(x_i) + 2\sum_{0<i<j}\sum a_ia_jCov(X_i,X_j)$
If $X_1, X_2, \cdots, X_n$ are independent,
$M_U(t) = e^{a_0t}M_{X_1}(a_1t)M_{X_2}(a_2t)\cdots M_{X_n}(a_nt)$

Central Limit Theorem

Population mean μ, standard deviation σ
$E(X_1+X_2+X_3+\cdots+X_n) = n\cdot\mu$
$Var(X_1+X_2+X_3+\cdots+X_n) = n\cdot\sigma^2$
$SD(X_1+X_2+X_3+\cdots+X_n) = \sqrt{n}\cdot\sigma$
Sample Mean $\bar{X} = \frac{X_1+X_2+X_3+\cdots+X_n}{n}$
$E(\bar{X}) = \mu$ $Var(\bar{X}) = \frac{\sigma^2}{n}$ $SD(\bar{X}) = \frac{\sigma}{\sqrt{n}}$ $M_{\bar{X}}(t) = (M_X\frac{t}{n})^n$
$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$

Normal Distribution

$Z = \frac{X-\mu}{\sigma}$ $X = \mu + \sigma Z$ p.d.f: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$
$E(x) = \mu$ $Var(x) = \sigma$ $M(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$
If n is large or population is normal distributed, $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$

Point Estimation of $f(x;\lambda)$

Likelihood Estimator of λ, $\hat{\lambda}$
$\mathcal{L}(\lambda) = \prod_{i=1}^{n} f(x_i;\lambda)$ $\frac{d(\ln \mathcal{L}(\hat{\lambda}))}{d\lambda} = \frac{d\sum f(x_i;\lambda)}{d\lambda} = 0$
Method of moments estimate of λ, $\tilde{\lambda}$
$E(X) = \bar{X}$, solve $\tilde{\lambda}$ in term of $\bar{X}$.
If $E(\hat{\theta}) = \theta$, $\hat{\theta}$ is unbaised for θ. $Bais(\hat{\theta}) = E(\hat{\theta}) - \theta$
Mean Squared Error of $\hat{\theta}$:
$MSE(\hat{\theta}) = E[(\hat{\theta}-\theta)^2] = (bais(\hat{\theta}))^2 + Var(\hat{\theta})$

Confidence Interval

If n is large or population is normal distributed, $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$
Sample variance $s^2 = \frac{\sum(x_i-\bar{x})^2}{n-1}$. $n-1$ freedom
If n is small and population is not normal distributed, $T = \frac{\bar{X}-\mu}{s/\sqrt{n}}$.

Mean $\bar{X} \pm z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}$ $n = [\frac{Z_{\alpha/2}\cdot\sigma}{\varepsilon}]^2$ (n rounds up)
$\bar{X} \pm t_{\alpha/2}\cdot\frac{s}{\sqrt{n}}$ $D.F. = n-1$

Population Variance, σ^2 $(\frac{(n-1)s^2}{\chi^2_{\alpha/2}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}})$

Standard Deviation, σ $(\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}}, \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}})$ $D.F. = n-1$

Sample Proportion $\hat{p} = \frac{x}{n}$ $E(\hat{p}) = p$ $SD(\hat{P}) = \sqrt{\frac{p(1-p)}{n}}$
$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $n = (\frac{z_{\alpha/2}}{\varepsilon})^2 p*(1-p*)$
Conservative Approach $p* = 0.5$. Choose $p* = 0.5$, or the closest to 0.5.

Hypothesis Test

	H_0 ture	H_0 false
Accept H_0		Type II Error
Reject H_0	Type I Error	

Null	Alternative		Reject Condition
$H_0: p \geq p_0$	$H_1: p < p_0$	Left tailed	$Z < -z_\alpha$
$H_0: p \leq p_0$	$H_1: p > p_0$	Right tailed	$Z > z_\alpha$
$H_0: p = p_0$	$H_1: p \neq p_0$	Two tailed	$Z < -z_{\alpha/2}$ or $Z > z_{\alpha/2}$

P-value (observed level of significance). If P-value$>\alpha$, do not reject H_0. If P-value$<\alpha$, reject H_0.

Population Proportion, p $\quad Z = \dfrac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$

Population Mean, μ $\quad Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \quad t = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$

Population Variance, σ^2 $\quad \chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \quad n-1$ freedom

Combination of two population

Two large populations with success **proportions** p_1 and p_2. Sample proportions are $\hat{p}_1 = \frac{x_1}{n_1}$ and $\hat{p}_2 = \frac{x_2}{n_2}$.

$$SD(p_1 - p_2) = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

Confidence intervals of $(p_1 - p_2)$:

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Hypothesis Test:

$$H_0: P_1 = P_2 \qquad \begin{array}{l} H_1: P_1 < P_2 \\ H_1: P_1 > P_2 \\ H_1: P_1 \neq P_2 \end{array}$$

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p}(1/n_1 + 1/n_2))}} \qquad \hat{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

Population **means** are μ_1 and μ_2. Sample means are $\bar{X}_1$ and $\bar{X}_2$. Sample Std. Dev are s_1 and s_2. $SD(\bar{X}_1 - \bar{X}_2) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Confidence intervals of $(\mu_1 - \mu_2)$: $(\bar{X}_1 - \bar{X}_2) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$,

or $(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$, if σ_1 and σ_2 are unknown.

$$\text{D.F.} = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{1}{n_1-1}(\frac{s_1^2}{n_1})^2 + \frac{1}{n_2-1}(\frac{s_2^2}{n_2})^2}$$

If we assume $\sigma_1 = \sigma_2 = \sigma$, confidence interval is

$(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2} s_{pooled}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

$s_{pooled}^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}$. $D.F = n_1 + n_2 - 2$.

Hypothesis Test:

$$H_0: \mu_1 = \mu_2 \qquad \begin{array}{l} H_1: \mu_1 < \mu_2 \\ H_1: \mu_1 > \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{array}$$

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - \delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}, \text{ or } T = \frac{(\bar{X}_1 - \bar{X}_2) - \delta_0}{s_{pooled}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ if } \sigma_1 = \sigma_2 = \sigma.$$

Matched Pair Comparison

Pair			Difference
1	X_1	Y_1	$D_1 = X_1 - Y_1$
2	X_2	Y_2	$D_2 = X_2 - Y_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	X_n	Y_n	$D_n = X_n - Y_n$

Assume D_i has mean δ and Std. Dev σ_D. Confidence interval for δ is $\bar{D} \pm t_{\alpha/2}\frac{s_D}{\sqrt{n}}$. The degree of freedom is $n-1$.

$H_0: \delta = \delta_0$, test statistic $T = \frac{\bar{D} - \delta_0}{s_D/\sqrt{n}}$.

χ^2 test for goodness of fit

A random sample of size n is classified into k categories or cells. Let $Y_1, Y_2, Y_3, \cdots, Y_k$ denote the respective cell frequencies. $\sum_{i=1}^{k} Y_i = n$ Denote the cell probabilities by $p_1, p_2, p_3, \cdots, p_k$.

$$H_0: p_1 = p_{10}, p_2 = p_{20}, \cdots, p_k = p_{k0}. \qquad \sum_{i=1}^{k} p_{i0} = 1.$$

$$Q_{k-1} = \sum_{i=1}^{k} \frac{(Y_i - np_{i0})^2}{np_{i0}}$$

Reject H_0 if $Q_{k-1} \geq \chi_\alpha^2$, d.f.$=k-1$

Critical Normal Distribution Table

α	0.999	0.99	0.98	0.95	0.9	0.8
$Z_{\alpha/2}$	3.291	2.576	2.326	1.96	1.645	1.282

Small samples look up *t-distribution*!!