1. Expected values and standard errors of common point estimators

$\overline{Y}$	μ	$\frac{\sigma}{\sqrt{n}}$
$\hat{p} = Y/n$	p	$\sqrt{\frac{pq}{n}}$
$\overline{Y}_1 - \overline{Y}_2$	$\mu_1 - \mu_2$	$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
$\hat{p}_1 - \hat{p}_2$	$p_1 - p_2$	$\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$

2. Small-sample normal population confidence interval summary

parameter	confidence interval	Notes
μ	$\overline{Y} \pm t_{\alpha/2,\nu}$	$\nu =$ degrees of freedom $= n - 1$
$\mu_1 - \mu_2$	$\overline{Y}_1 - \overline{Y}_2 \pm t_{\alpha/2,\nu} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$	$\nu = n_1 + n_2 - 2$

note: $S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 + 1}$

3. Confidence interval for σ^2

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}} \right)$$

4. Fisher information, large sample distribution of maximum likelihood statistics

If $Y_1, \ldots, Y_n$ is a random sample from a distribution with density $f(y)$ depending on parameter θ, the Fisher information $I(\theta)$ is defined as:

$$I(\theta) = \left[nE\left(-\frac{\partial^2 \ln f(y)}{\partial \theta^2} \right) \right]^{-1}$$

If t is a differentiable function of the parameter θ then under the usual regularity conditions, for large sample sizes

$$Z = \frac{t(\hat{\theta}) - t(\theta)}{\sqrt{\left[\frac{\partial t(\theta)}{\partial \theta}\right]^2 I(\theta)}}$$

has an approximately standard normal distribution.

5. Hypothesis tests for means

type	H_0	H_a	test statistic	rejection region
large sample	$\theta = \theta_0$	$\theta \neq \theta_0$	$z = \frac{\hat{\theta}-\theta_0}{\sigma/\sqrt{n}}$	$\lvert z \rvert > z_{\alpha/2}$
small sample	$\theta = \theta_0$	$\theta \neq \theta_0$	$t = \frac{\hat{\theta}-\theta_0}{S/\sqrt{n}}$	$\lvert t \rvert > t_{\alpha/2,n-1}$

6. Large sample distribution of likelihood ratio test statistic

If λ is the test statistic for a likelihood ratio test of

$$H_0 : \theta \in \Omega_0 \quad \text{versus} \quad H_a : \theta \in \Omega$$

where the number of free parameters specified by the statement $\theta \in \Omega$ is one, then $-2 \ln \lambda$ is approximately distributed as a χ^2 random variable with one degree of freedom.

7. linear models

For the linear model

$$Y = X\beta + e$$

where the e_i are IID $N(0, \sigma_e^2)$, the least squares estimates $\hat{\beta}$ of the parameters β are given by:

$$\hat{\beta} \ = \ (X'X)^{-1}X'Y$$

where the columns of X are linearly independent. Furthermore, the vector of 'fitted values' is given by

$$X\hat{\beta} = X(X'X)^{-1}X'Y$$

and the vector of residuals is given by

$$e = Y - X\hat{\beta} = [I - X(X'X)^{-1}X']Y$$

8. MISCELLANEOUS FORMULAS

If Y is a vector of random variables, μ is the vector of expected values of the Y_i, that is, $E(Y) = \mu$, and V is the variance-covariance matrix of Y, then if t is a vector and A is a matrix,

$$E(t'Y) = t'\mu \quad V(t'y) = t'Vt \quad E(A'Y) = A'\mu \quad V(A'Y) = A'VA$$

Suppose Y has a multivariate normal distribution with mean vector 0 and variance-covariance matrix I (i.e, Y is a vector of IID $N(0,1)$ random variables). Then:

1) The quadratic form $Y'AY$ has a chi-square distribution with degrees of freedom equal to the rank of A if and only if A is idempotent, that is, $A^2 = A$.

2) The quadratic forms $Y'AY$ and $Y'BY$ are independently distributed if $AB = 0$

9. CRITICAL VALUES

If z is a $N(0,1)$ random variable, then

$$P(-1.96 < z < 1.96) = 0.95$$

If t has a t-distribution with n degrees of freedom, then:

If $n = 8$ $\quad P(-2.306 < t < 2.306) = 0.95$
If $n = 9$ $\quad P(-2.262 < t < 2.262) = 0.95$
If $n = 10$ $\quad P(-2.228 < t < 2.228) = 0.95$
If $n = 11$ $\quad P(-2.201 < t < 2.201) = 0.95$

If x has a χ^2 distribution with one degree of freedom, then

$$P(x > 0.00393) = .95$$