Testing the real standard dev of a sample:

$$W = \frac{(n-1)S^2}{\sigma_0^2}, W \sim \chi^2 \; df = n-1$$

Assume random/independent sample, normally distributed pop.

Testing if two population variances are the same:

$$F = \frac{S_1^2}{S_2^2}, \; df = n_1 - 1, n_2 - 1$$

Assume two samples are randomly and independently selected from their two populations, which are both normally distributed

One-Way Anova

$$SST = \sum_{k=1}^{K} n_k (Y_k - \hat{Y})^2$$

$$MST = \frac{SST}{K-1}$$

$$SSE = \sum_{k=1}^{K} \sum_{i=1}^{n_k} (Y_{ij} - Y_k)^2$$

$$s_p^2 = MSE = \frac{SSE}{n-K}$$

$$F = \frac{MST}{MSE}$$

$$TSS = \Sigma_{k=1}^{K} \Sigma_{i=1}^{n_k} (Y_{ij} - Y)^2 = (n-1)s^2$$

Assume random/independent selection, and each group is normally distributed

Confidence interval for difference between 2 groups if K = 2

$$y_1 - y_2 \pm t_{\alpha/2, n_1+n_2-2} \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}, \; df = n_1 + n_2 + 2$$

$$y_1 - y_2 \pm t_{\alpha/2, n-K} \sqrt{\frac{MSE}{n_i} + \frac{MSE}{n_j}} \; df = n - K$$

Bonferroni correction

$$\alpha = \frac{\alpha_F}{K(K-1)/2}$$

Two-Way Anova with block design

$$TSS = \sum_{i=1}^{K} \sum_{j=1}^{B} (Y_{ij} - Y)^2$$

$$SST = B \sum_{i=1}^{K} (Y_{i-} - Y)^2$$

$$SSB = K \sum_{j=1}^{B} (Y_{-j} - Y)^2$$

$$SSE = TSS - SSB - SST = \sum_{i_1}^{K} \sum_{j=1}^{B} (Y_{ij} - Y_{i-} - Y_{-j} + Y)^2$$

Assume errors are normally distributed, blocks are as homogenous as possible. Treatments are randomly assigned to the units in each block, treatment and block effects are all constants.

Two-Way Anova with Factors

Overall test

$$SSE = \sum_{i=1}^{K} \sum_{j=1}^{J} \sum_{r=1}^{R} (Y_{ijr} - Y_{ij-})^2$$

$$SST = R \sum_{i=1}^{K} \sum_{j=1}^{J} (Y_{ij-} - Y)^2$$

$$MST = \tfrac{SST}{JK-1} \; MSE = \tfrac{SSE}{n-KJ}$$

Decomposed tests

$$SS(A) = RJ \sum_{i=1}^{K} (\bar{y}_{i--} - \bar{y}_{---})^2$$

$$SS(B) = RK \sum_{j=1}^{J} (\bar{y}_{-j-} - \bar{y}_{---})^2$$

$$MS(B) = \tfrac{SS(B)}{J-1} \; MS(A) = \tfrac{SS(A)}{K-1}$$

$$SS(AB) = R \sum_{j=1}^{J} \sum_{i=1}^{K} (\bar{y}_{ij-} - \bar{y}_{i--} - \bar{y}_{-j-} + \bar{y}_{---})^2$$

$$MS(AB) = \frac{SS(AB)}{KJ - J - K + 1}$$

$$F_{AB} = \frac{MS(AB)}{MSE}, df = KJ - J - K + 1, n - JK$$

if not rejected, then:

$$F_A = \frac{MS(A)}{MSE}, df = K - 1, n - JK$$

Assume: samples are random/independent, each group normally distributed. same # of experimental units randomly assigned to each $R \times J$ possible factor combinations. Errors are normally distributed with the same var as population.

Linear Regression

$$\sigma_{\hat{\beta}_1} = \frac{\sigma}{\sqrt{S_{XX}}}$$

$$s^2 = \frac{\sigma_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-2} = \frac{SSE}{n-2}$$

$$SSE = SS_{YY} - \hat{\beta}_1 SS_{XY}$$

$$T = \frac{\hat{\beta}_1 - 0}{\sigma / \sqrt{SS_{XX}}}$$

$$\hat{\beta}_1 \pm t_{n-2, \alpha/2} \frac{s}{\sqrt{SS_{XX}}}$$

$$r = \frac{SS_{XY}}{\sqrt{SS_{XX} SS_{YY}}}$$

$$SS_{XY} = \sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})$$

$$SS_{XX} = \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$SS_{YY} = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$$r = \hat{\beta}_1 \frac{s_x}{s_y}$$

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$r^2 = 1 - \frac{SSE}{SS_{YY}}$$

$$r \pm t_{\alpha/2, n-2} \sqrt{(1 - r^2)/(n-2)}$$

In small samples:

$$Z = \ln \frac{1+r}{1-r}$$

$$Z \pm z_{\alpha/2} / \sqrt{(n-3)} = (c_L, c_U)$$

$$[\frac{exp(2*c_L-1)}{exp(2*c_L+1)}, \frac{exp(2*c_U-1)}{exp(2*c_U+1)}]$$

$$E(\hat{y}(x_0)) = \beta_0 + \beta_1 x_0$$

$$s_{\hat{y}(x_0)} = s\sqrt{\frac{1}{n} + \frac{(x_0-\bar{x})^2}{SS_{XX}}}$$

$$S_{\tilde{y}(x_0)} = s\sqrt{1 + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{SS_{XX}}}$$

Assume x as constant(error around the measurement of x is negligible), errors are independent random variables with the same variance as the population(Y), mean of zero. Y is a random variable with the same variance as errors. Assume that $\hat{\beta}_1$ and $\hat{\beta}_0$ are normally distributed. Multiple Regression Assume: $E(\epsilon_i) = 0$ for all i, $Var(\epsilon_i) = \sigma^2$, normally and independently distributed errors.

$$s^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-(K+1)}$$

$$t = \frac{\hat{\beta}_j - \beta_j^*}{s\sqrt{c_{jj}}}$$

c_{jj} is the variance of $\hat{\beta}_j$.
Assume x as constant(error around the measurement of x is negligible), errors are independent random variables with the same variance as the population(Y), mean of zero. Y is a random variable with the same variance as errors.
Measuring the fit of the model

$$SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

$$SS_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2$$

$$R^2 = 1 - \frac{SSE}{SS_{yy}}$$

$$R_a^2 = 1 - [\frac{n-1}{n-(K+1)}](\frac{SSE}{SS_{yy}})$$

$$F = \frac{(SS_{YY} - SSE)/k}{SSE/[n-(K+1)]} = \frac{R^2/K}{(1-R^2)/(n-(K+1))}$$

Comparing Nested Models

$$F = \frac{(SSE_{M_0} - SSE_{M_1})/(k-g)}{SSE_{M1}/(n-(k+1))}, \ df = k-g, n-k-1$$

$$e_i^{std} = \frac{e_i}{s}$$

Categorical Data

$$\chi^2 = \sum_{j=1}^{k} \frac{(n_j - np_j^{(0)})^2}{np_j^{(0)}}$$

$$\chi^2 = \sum_{j=1}^{k} \frac{(observed - expected)^2}{expected}$$

The degrees of freedom is the difference between k - 1 and the number of unspecified probabilities in H_0.

Test for independence $E(n_{jk}) = n\hat{p}_{j-}\hat{p}_{-k}$ $df = rc - 1 - (r+c-2) = (r-1)(c-1)$
Non-parametric statistics

$$D_i = X_i - \eta_0$$

use binomial test on D_i, $p = 0.5$

$$z = \frac{X - np}{\sqrt{npq}}$$

For matched pairs, $D_i = X_i - Y_i$, then do test on D_i
Wilcoxon Paired Rank Sum
T^+ is rank sum of positive D_i

$$E(T^+) = \frac{n(n+1)}{4}$$

$$Var(T^+) = \frac{n(n+1)(2n+1)}{24}$$

Use Z statistic
Wilcoxon Independent Rank Sum

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - W$$

Where W is the rank sum of first sample

$$Z = \frac{U - (n_1 n_2/2)}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}}$$

Assume independent and identically distributed data.

Source	df	SS	MS	F	p-value
Treatments	$K-1$	SST	$MST = \frac{SST}{K-1}$	$\frac{MST}{MSE}$	Pr(F > F)
Error	$n-K$	SSE	$MSE = \frac{SST}{K-1}$		
Total	$n-1$	TSS			

Source	df	SS	MS	F
Treatments	$K-1$	SST	$MST = \frac{SST}{K-1}$	$\frac{MST}{MSE}$
Blocks	$B-1$	SSB	$MSB = \frac{SSB}{B-1}$	$\frac{MSB}{MSE}$
Error	$n-K-B+1$	SSE	$MSE = \frac{SSE}{n-K-B+1}$	
Total	$n-1$	TSS		

Source	df	SS	MS	F
A	$K-1$	$SS(A)$	$MS(A) = \frac{SS(A)}{K-1}$	$\frac{MS(A)}{MSE}$
B	$J-1$	$SS(B)$	$MS(B) = \frac{SS(B)}{J-1}$	$\frac{MS(B)}{MSE}$
$A \times B$	$KJ-K-J+1$	$SS(AB)$	$MS(AB) = \frac{SS(AB)}{KJ-K-J+1}$	$\frac{MS(AB)}{MSE}$
Error	$n-KJ$	SSE	$MSE = \frac{SSE}{n-KJ}$	
Total	$n-1$	TSS		