Hypothesis tests for two samples

Test	Null hypothesis	Test statistic	Distribution
F test on ratio of two variances	The ratio of the variances of the two populations is $\frac{\sigma_1^2}{\sigma_2^2}$.	$F = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \quad (s_1^2 > s_2^2)$	F_{n_1-1, n_2-1}
Kolmogorov-Smirnov 2-sample test	The two samples are drawn from the same underlying population.	$D^* = n_1 n_2 D$ where D = largest difference between cumulative probabilities based on the two samples and n_1, n_2 are the sample sizes.	As in statistical tables
Mann Whitney *U* test	The two samples are drawn from the same underlying population. [H$_1$: the two samples come from populations with different medians]	Rank all the data values from both samples together; for sample sizes $m, n : m \le n$, calculate $T_1 = \sum R - \frac{1}{2}m(m+1)$ and $T_2 = \sum S - \frac{1}{2}n(n+1)$ where $\sum R$ is the sum of ranks of sample size *m*, $\sum S$ is the sum of ranks of sample size *n*. The test statistic is the smaller of T_1, T_2. The test statistic can also be found by counting the number of values in sample 2 which exceed each of the values in sample 1, repeat for all the values in sample 2. There are online calculators such as: http://socr.stat.ucla.edu/Applets.dir/U_Test.html	As in statistical tables. Approx Normal for large samples with Mean $\frac{1}{2}mn$, Variance $\frac{1}{12}mn(m+n+1)$
Normal test for paired samples with known variance	The difference in the population means has value *k*.	$z = \frac{(\bar{x}_1 - \bar{x}_2) - k}{\sigma/\sqrt{n}} = \frac{\bar{d} - k}{\sigma/\sqrt{n}}$	N(0, 1)
Normal test for paired samples with unknown variance	The difference in the population means has value *k*.	$z = \frac{(\bar{x}_1 - \bar{x}_2) - k}{s/\sqrt{n}} = \frac{\bar{d} - k}{s/\sqrt{n}}$	N(0, 1) for large samples
Normal test for unpaired samples with common known variance	The difference in the means of the two populations is $\mu_1 - \mu_2$.	$z = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	N(0, 1)

Normal test for unpaired samples with common unknown variance	The difference in the means of the two populations is $\mu_1 - \mu_2$.	$z = \dfrac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ where $s = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	N(0, 1) for large samples
Normal test for unpaired samples with different known variances	The difference in the means of the two populations is $\mu_1 - \mu_2$.	$z = \dfrac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$	N(0, 1)
Normal test for unpaired samples with different unknown variances	The difference in the means of the two populations is $\mu_1 - \mu_2$.	$z = \dfrac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$	N(0, 1) for large samples
Sign test	Population of differences has median = 0.	r = number of values of $d_i > 0$ where $d_i = x_i - y_i$	B(n, ½) where n = number of values $\neq 0$
t test for paired samples	The difference in the population means has value *k*.	$t = \dfrac{(\bar{x}_1 - \bar{x}_2) - k}{s/\sqrt{n}} = \dfrac{\bar{d} - k}{s/\sqrt{n}}$	t_{n-1}
t test for unpaired samples with common unknown variance	The difference in the means of the two populations is $\mu_1 - \mu_2$.	$t = \dfrac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ where $s = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	$t_{n_1 + n_2 - 2}$
Wilcoxon paired sample test	Population of differences has median = *M*.	Test statistic: $T = \min[P, Q]$ P, Q are the sums of the ranks corresponding to positive and negative deviations $(d_i - M)$ where $d_i = x_i - y_i$	As in statistical tables
Wilcoxon rank sum 2-sample test	The two samples are drawn from the same underlying population. [H_1: the two samples come from populations with different medians]	Rank all the data values from both samples together; W = sum of ranks of sample size m. where sample sizes are $m, n : m \leq n$.	Statistical tables give critical values for the lower tail; critical value for the upper tail is $m(m+n+1) - W_C$ where W_C is the critical value from tables. Approx Normal for large samples with Mean $\frac{1}{2}mn + \frac{1}{2}m(m+1)$, Variance $\frac{1}{12}mn(m+n+1)$