Non-Parametric Hypothesis Testing

Prof. Richard B. Goldstein

test name	replaces	statistic	efficiency/comments
Sign test	single mean or two paired means	$z = \dfrac{r - 0.5}{\sqrt{\dfrac{0.25}{n}}}$ where $r = \dfrac{\#\text{of plus signs}}{n}$	0.63
Mann-Whitney test or Wilcoxin two-sample test	two independent means	n_1 = smaller sample n_2 = larger sample $\mu_R = \dfrac{n_1(n_1 + n_2 + 1)}{2}$ $\sigma_R = \sqrt{\dfrac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$ $z = \dfrac{R - \mu_R}{\sigma_R}$ R = sum of the ranks of the smaller sample	0.95 n_1, n_2 at least 8 Otherwise, use tables
Kruskal-Wallis test	$k \geq 2$ means	$H = \dfrac{12}{n(n+1)} \sum_{i=1}^{k} \dfrac{r_i^2}{n_i} - 3(n+1)$ is approx. χ^2 with $k - 1$ d.f. r_i = sum of ranks of i^{th} sample n_i = i^{th} sample size $n = n_1 + n_2 + \cdots + n_k$	All n_i at least 5

Runs test	randomness	V = number of runs of odds/evens or any two similar groups $\mu_V = \frac{2n_1n_2}{n_1 + n_2} + 1$ $\sigma_V = \sqrt{\frac{(2n_1n_2)(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 + 1)}}$ $z = \frac{V - \mu_V}{\sigma_V}$	
Spearman Rank Correlation	correlation	$r_S = 1 - \frac{6\sum_{i=1}^{n}(d_i)^2}{n(n^2 - 1)}$ where d_i = difference of rankings of n pairs of data	0.91 If H_0: no correlation, then for large n $z = \frac{r_s - 0}{\frac{1}{\sqrt{n-1}}} = r_s\sqrt{n-1}$
Kolmogorov-Smirnov test	goodness of fit	$D_n^+ = \max_{1 \le i \le n}\left\{\frac{i}{n} - F(x_{(i)}, \theta)\right\}$, $D_n^- = \max_{1 \le i \le n}\left\{F(x_{(i)}, \theta) - \frac{i-1}{n}\right\}$ $D_n = \max(D_n^+, D_n^-)$ $P\{\sqrt{n}D_n > t\} \to K(t) = 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2t^2} = 2\left(e^{-2t^2} - e^{-8t^2} + \cdots\right)$	