Nonparametric Hypothesis Tests

1 Nonparametric Tests

1.1 General idea: rank data, test for differences in ranks
1.2 Hypotheses in words, not parameters

2 Two-Sample Rank testing (Mann-Whitney Test)

2.1 Procedure
- 2.1.1 rank data (direction of ranking does not matter)
- 2.1.2 calculate Mann-Whitney statistic, U

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \qquad R_1 = \Sigma(\text{sample 1 ranks}) \qquad \text{Zar, eq. 8.45}$$

$$U' = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2 \qquad U' = n_1 n_2 - U$$

- 2.1.3 Compare w/ critical (in Zar, table B.11)
  2-tailed test: compare both U & U' w/ critical value $U_{\alpha(2),n1,n2}$
- 2.1.4 Check that $R_1 + R_2 = \frac{N(N+1)}{2}$, esp if have tied ranks
- 2.1.5 Dealing w/ tied ranks: if equal values, ranks tied, give each the mean tied rank
- 2.1.6 Power of test: $\geq$ 95% (=3/π) power of two-sample t-test

2.2 One-Tailed Mann-Whitney Test
2.3 Normal approximation for large samples (sample sizes larger than 20, 40)

3 Nonparametric Paired-Sample Test (Wilcoxon paired-sample test)

3.1 Procedure:
- 3.1.1 arrange data in pairs
- 3.1.2 rank pairs from low to high, according to | differences |
- 3.1.3 deal w/ tied measurements
- 3.1.4 discard any differences of zero; $d_i = 0$
- 3.1.5 assign + or – sign to each difference
- 3.1.6 sum + ranks & – ranks: T_+, T_-
- 3.1.7 $T_- = \frac{n(n+1)}{2} - T_+$ or $T_+ = \frac{n(n+1)}{2} - T_-$
- 3.1.8 compare w/ critical values of T-distribution, $T_{\alpha(2),n}$ (in Zar, table B.12)

3.2 Reject H_0 if either T_+, T_- *less than or equal to* critical value T_α

3.3 One-tailed tests:
Hypotheses (reverse signs for other tail):
H_0: measurements in pop. 1 $\leq$ measurements in pop. 2
H_A: measurements in pop. 1 $>$ measurements in pop. 2
reject H_0 if $T_- \leq T_{\alpha(1),n}$

3.4 Assumption: population of differences (d_i) symmetric about median

3.5 Power: if differences normally distributed,
power of Wilcoxon test ≈ 95% (=3/π) power of paired-sample t-test

3.6 Normal approximation to Wilcoxon Paired-Sample Test

$$Z = \frac{|T - \mu_T|}{\sigma_T} \quad ; \quad \text{use either } T_+ \text{ or } T_-$$

$$\mu_T = \frac{n(n+1)}{4} \qquad \sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$