4.1 Paired-Comparison Permutation Test

Steps for a paired-comparison permutation test

1. Compute D_i's for n pairs of data, and compute the $\bar{D}_{obs}$.
2. Permute 2^n possible assignments and compute $\bar{D}_i$ for each permutation.
3. Upper-tail p-value is

$$P_{upper} = \frac{\# \text{ of } \bar{D}_i\text{'s} \geq \bar{D}_{obs}}{2^n}$$

Hypotheses

$H_0 : F(x) = 1 - F(x)$

4.1.2 Randomly Selected Permutations

- Let U_i be n independent random variables such that $U_i = -1$ or 1 with the equal probability

Then a randomly selected mean of differences is, $\bar{D} = \frac{\sum_{i=1}^{n} U_i |D|}{n}$, $n \geq 1000$ or 5000 typically

$S_+ = \sum_{1}^{n} V_i |D_i|$, $V_i = 0$ or 1

4.1.3 Large-Sample Approximations

$E(\bar{D}) = 0$, $Var(\bar{D}) = \frac{1}{n^2} \sum_{i=1}^{n} |D_i|^2$, $Z = \bar{D} / \sqrt{Var(\bar{D})}$

$E(S_+) = \frac{1}{2} \sum_{i=1}^{n} |D_i|$, $Var(S_+) = \frac{1}{4} \sum_{i=1}^{n} |D_i|^2$, $Z = (S_+ - E(S_+)) / \sqrt{Var(S_+)}$

4.1.4 A Test for the Median of a symmetric population

4.2 Signed-Rank Test

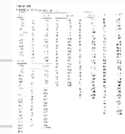
- a nonparametric test for paired-comparison experiments based on ranks.
- assume that the differences of the pairs have no ties and that none of the differences is 0. We rank the absolute values of the differences, and then we attach the signs of the differences to the ranks.

TABLE 4.2.1
Signed Ranks for Data in Table 4.1.1

	Student				
	1	*2*	*3*	*4*	*5*
Differences	240	−120	510	60	150
Ranks of Absolute Values	4	2	5	1	3
Signed Ranks	4	−2	5	1	3

The signed-rank statistic is the sum of positive signed ranks. SR_+

4.2.1 The Wilcoxon Signed-Rank Test without Ties in the data.

1. Obtain the signed ranks, and compute SR_{+obs}, the observed value of the sum of the positive signed ranks
2. Compute SR_+ for all 2^n possible assignments of plus and minus signs to the ranks of the absolute values of the differences.
3. The upper-tail p-value is the fraction of the SR_+'s that are greater than or equal to SR_{+obs}. Computing the lower-tail p-values are the opposite. The two-tail p-value is twice the one-tail p-value. Appendix A9 

4.2.2 Large-Sample Approximation

$$SR_+ = \sum_{i=1}^{n} i V_i, \quad V_i = 1 \text{ or } 0 \text{ (equally likely)}$$

$$E(SR_+) = \frac{1}{2}\sum_{i=1}^{n} i = \frac{n(n+1)}{4}, \quad Var(SR_+) = \sum_{i=1}^{n} \frac{i^2}{4} = \frac{n(n+1)(2n+1)}{24}$$

$$Z = \frac{SR_+ - E(SR_+)}{\sqrt{Var(SR_+)}}$$, and refer the result to the standard normal distribution

4.2.3 Adjustment for Ties

2 types of ties with paired data

- absolute values of differences are the same in value
- 2 different independent observations are the same in value

Ranking with zeros

- the absolute values of the differences, including zeros, are ranked.
- The average rank is assigned to differences that are tied.
- Plus and minus signs are attached to ranks that correspond to positive and negative differences
- 0 is fixed

Ranking Without Zeros

- only nonzero differences are ranked, with average ranks given to the nonzero tied values.

$$SR_+ = \sum_{i=1}^{m} V_i R_i \quad , \quad m = \text{\# of differences that are not zero}$$

$$E(SR_+) = \frac{1}{2}\sum_{i=1}^{m} R_i \quad , \quad Var(SR_+) = \frac{1}{4}\sum_{i=1}^{m} R_i^2$$

Alternative Expressions for $E(SR_+)$ and $Var(SR_+)$ with Ties

- when the data have ties and ranking is done with zeros

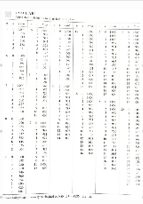
$$E(SR_+) = \frac{n(n+1) - t_0(t_0+1)}{4} \quad , \quad t_0 = \text{\# of differences that are } 0$$

$$Var(SR_+) = \frac{n(n+1)(2n+1) - t_0(t_0+1)(2t_0+1)}{24} - \frac{\sum^{g}(t_i^3 - t_i)}{48} \quad , \quad g = i^{th} \text{ group}$$

* if the ranking is done without zeros, we replace n for m, and drop t_0

4.3 Other Paired-Comparison Tests

4.3.1 Sign Test



- If the two treatments in the paired comparison have the same effect, then a difference has probability 0.5 of being positive, and SN_+ has a binomial distribution with $p = 0.5$
- If there is a difference between treatments, then SN_+ tends to be larger or smaller than one would expect of a binomial random variable with $p = 0.5$.

$$P(SN_+ \geq k) = \sum_{i=k}^{n} \binom{n}{i}(0.5)^n \quad , \quad \text{upper-tail } p\text{-value for } SN_+ = k$$

$$P(SN_+ \leq k) = 1 - P(SN_+ \geq k+1) \quad , \quad \text{and for further investigation,}$$

$$Z = \frac{SN_+ - 0.5n}{\sqrt{0.25n}} \quad , \quad \text{normal distribution}$$

* Use a continuity correction for small samples.

4.3.2 General Scoring Systems

4.3.3 Selecting Among Paired-Comparison Tests

- There are various assumptions about the probability distribution of the differences.
  - → For shifting by Δ, the signed-rank test performs better than the t-test for paired data when the distribution of the differences has heavy tails
  - → If the distribution has lighter tails, the t-test will do better.
- Sign tests generally do not perform well as signed-rank test. $(\frac{2}{3})$
  But sign tests may perform better in situations where there is a tendency for occasional large positive or negative differences relative to the other differences. (Cauchy distribution)

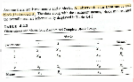
4.4 A Permutation Test for a Randomized Complete Block Design

Blocking is a technique that is used when the experimental units to which treatments are to be applied are not homogeneous, or when the conditions under which the experiment is to be conducted cannot be held constant throughout. For instance,

3 features of RCBD

1. Experimental units are divided into blocks in such a way that units or experimental conditions within blocks are homogeneous.
2. Blocks have the same number of experimental units as there are treatments.
3. The treatments are randomly assigned to experimental units within blocks.

4.4.1 F_1 statistic for Randomized Complete Block Designs



- Suppose the observations follow the model, $X_{ij} = \mu + t_i + b_j + \varepsilon_{ij}$

overall effect   treatment effect   block effects

with $k-1$ and $(k-1)(b-1)$ d.f.

$$F_1 = \frac{b\sum_{i=1}^{k}(\bar{X}_{i.}-\bar{X})^2/(k-1)}{\sum_{i=1}^{k}\sum_{j=1}^{b}(X_{ij}-\bar{X}_{i.}-\bar{X}_{.j}+\bar{X})/[(k-1)(b-1)]}$$

, where hypotheses are

$H_0: t_1 = t_2 = \cdots = t_k$

$H_a:$ Not all t_i's are the same

4.4.2 Permutation F-Test for Randomized Complete Block Design

Process of Permutation F-test ; (suppose $\varepsilon \not\sim N(0,1)$)

1. Compute F_{obs}.
2. Permute the observation within each of the block. $(K!)^b$ permutations
3. Compute F-statistic for each permutation
4. $$P_{upper} = \frac{\text{number of } F\text{'s} \geq F_{obs}}{(K!)^b}$$, or alternatively,

$$SST^* = \sum_{i=1}^{K} (\bar{X}_{i\cdot} - \bar{X})^2, \text{ or } \quad SSX^* = \sum_{i=1}^{K} (\bar{X}_{i\cdot})^2$$

4.4.3 Multiple Comparisons

4.5 Friedman's Test for a Randomized Complete Block Design

- involves ranking the observations within blocks and applying the permutation F-test for RCBD
- approximated by χ^2 distribution with $k-1$ degrees of freedom

1. assign ranks within blocks
2. Obtain mean rank for each treatment
3. $$FM = \underbrace{\frac{12b}{k(k+1)}}_{==C} \underbrace{\sum_{i=1}^{K} \left(\bar{R}_i - \frac{k+1}{2}\right)^2}_{==SSR}$$, and apply the statistic to χ^2 table with $k-1$ degrees of freedom.

4.5.2 Adjustment for Ties

- Since ranking is done within blocks, no adjustment for ties is needed when ties occur across blocks, but within each block, we assign the average of the ranks to tied observation.
- Let S^2_{Bj} denote the sample variance of the adjusted ranks with block j.

$$FM_{ties} = \frac{b^2}{\sum_{j=1}^{b} S^2_{Bj}} \sum_{i=1}^{K} \left(\bar{R}_i - \frac{k+1}{2}\right)^2, \text{ or alternatively, } FM_{ties} = \frac{FM}{1 - \sum_{j=1}^{b} \sum_{i=1}^{g_j} \frac{t_{ij}^3 - t_{ij}}{bk(k-1)}}$$, where

t_{ij} denote the number of tied observations in the i^{th} group within the j^{th} block, and g_j denote the number of groups of tied observations within the j^{th} block.

Use of Friedman's Test

Friedman's test is generally not as powerful as either the F-test or the permutation F-test for randomized complete block designs. If there are only two treatments, then Friedman's test is equivalent to a two-sided sign test, which generally has low power except when extreme observations tend to occur. When the observations satisfy the usual analysis of variance assumptions as discussed in Section 4.4.1, then the asymptotic relative efficiency of the test to the F-test is as low as 0.64 for the case of two treatments but reaches a limiting value of 0.955 as the number of treatments increases. Thus, Friedman's test is generally more effective for a larger number of treatments.

4.5.3 Cochran's Q and Kendall's W

Cochran's Q
- Used for experiments with binary outcomes.

both use Friedman's test

Kendall's W
- level of agreement (Concordance)

4.5.4 χ^2 Approximation for FM or FM_{ties}

$E(SSR) = \sum_{i=1}^{k} E\left(\bar{R}_i - \frac{k+1}{2}\right)^2 = \frac{k}{b^2}\sum_{j=1}^{b}\sigma_{Bj}^2$, σ_{Bj}^2 = population variance of the ranks in block j, $j = 1, \ldots, k$

Since $E[C(SSR)] = k-1$, $C = \dfrac{b^2(k-1)}{k\sum_{j=1}^{b}\sigma_{Bj}^2}$.

$\sigma_{Bj}^2 = k\sigma_{Bj}^2/(k-1) = S_{Bj}^2$

$= (k-1)(k+1)/12$ if no ties within blocks, then $C = \dfrac{12\,b}{k(k+1)}$

4.6 Ordered Alternatives for a Randomized Complete Block Design

4.6.1 Page's Test

- measure of the association between the presumed order of the treatments and the rank sum of the treatments
- Assume that the observations are ranked, Page's statistic is,

  $PG = \sum_{i=1}^{k} i \cdot R_i$

- same process with the permutation test.
- A large value of PG indicates that the treatment responses tend to increase as the index of the treatments increases

- Large-Sample Approximation

$E(PG) = \frac{bk(k+1)^2}{4}$, $Var(PG) = \frac{(k-1)k(k+1)}{12} \sum_{j=1}^{b} S_{Bj}^2$, if no ties within blocks,

$S_{Bj}^2 = k(k+1)/12$, so $Var(PG) = \frac{b(k-1)k^2(k+1)^2}{144}$, so $Z = \frac{PG - E(PG)}{\sqrt{Var(PG)}}$