Chapter 1

Binomial Test

- tests of medians
- it tests if the observed median is the adequate estimate of the expected median because in theory, binomial distributions approach normal distribution, and in normal distribution, the median and mean should be the same.

Chapter 2

3 assumptions of two-sample t-test :

- the samples should be randomly selected from two infinite populations and independently
- the populations have normal distributions
- the variances of the two populations should be the same.

Permutation Tests

- comparing the means of 2 groups of distribution without the assumptions of the two-sample t-test.
- calculate all of permutations' mean difference and compute the probability.

Wilcoxon Rank-sum Test (No ties)

- similar process with permutation tests except for the use of ranks

Wilcoxon Rank-sum Test (with ties)

- stays the same except for adjusting ranks

Family?,

Mann-Whitney Test

- make every possible pair for each X_i and count the number of cases $X_i < Y_i$
- Use the table for Mann-Whitney Test of $n(X_i)$ and $n(Y_i)$.

Hodges Lehmann Estimate

assuming there is a shift parameter Δ, $F(X) = F(Y-\Delta)$

- We are making an inference about the Δ, the shift.
- arrange all mn pairwise differences of $X_i - Y_i$ in an ascending order.
- and using the Mann-Whitney table, find $k_a = l_{1-\alpha} + 1$, $k_b = U_{1-\alpha}$

To this point, we have discussed to distinguish between the effects of two treatments when the observations from one of the treatments tend to be larger than the other. However, in some situations the "variability" of the observations for the two treatments is important.

Siegel-Tukey Test:

- Arrange the data in an ascending order and assign rank 1 to the smallest observation, rank 2 to the largest observation, rank 3 to the next largest observation, rank 4 to the next smallest observation, and so on.
- Apply the Wilcoxon rank-sum test. The smaller ranks and smalle rank sum are associated with the treatment that has the larger variability.

* the treatment with the smaller rank sum has the larger variability.

** Ansari-Bradley Tests 는 중요한게 아닌가봄

Tests on Deviances

- tests the location parameter using mean values of 2 groups.
- obtain RMD of all possible permutation, and

$dev_{ix} = X_i - \mu_1$
$dev_{jy} = Y_j - \mu_2$

$$RMD = \frac{\sum_{i=1}^{m} |dev_{ix}| / m}{\sum_{j=1}^{n} |dev_{jy}| / n}$$

- If we use two-sided test,

$$RMD = \frac{max\left(\sum_{i=1}^{m} |dev_{ix}| / m,\ \sum_{j=1}^{n} |dev_{jy}| / n\right)}{min\left(\sum_{i=1}^{m} |dev_{ix}| / m,\ \sum_{j=1}^{n} |dev_{jy}| / n\right)}$$

Kolmogorov - Smirnov Test

- This test can be applied even when you don't have any clue or information about the population distributions of two data (Shape, scale, location)
- It is an omnibus test that is designed to identify any form of difference between 2 groups

$T_{KS} = \max_{z} \left| \hat{F}_1(z) - \hat{F}_2(z) \right|$, the maximum absolute value of the difference between the two sample

---

Permutation F-test

- designed specifically for detecting difference in mean for any pair of distributions among K different distributions.
- Obtain the F-statistics of all possible permutation of K treatments including the observed F-statistic, and compute $\frac{\#\text{ of obs } F_i(x) \geq F_{obs}}{\text{all permutation}}$

☆ F-distribution is always an upper-tail test.

☆☆ We may calculate the p-value of the observed F-statistic using the F-statistic table with K-1 and N-K degrees of freedom. However, it is better to use a permutation F-test when distributions are not normally distributed.

Kruskal-Wallis Statistic

- obtaining a statistic that is equivalent to the F-statistic applied to ranks with a permutation distribution that may be approximated by χ^2 distribution with $k-1$ degrees of freedom.

* We use χ^2 distribution because the p-value obtained from the permutation test will always be smaller than that of χ^2 statistic.

- assign ranks to all of the observations and compute the mean rank of each group, then

$$KW = \underbrace{\frac{12}{N(N+1)}}_{\text{scaling factor}} \underbrace{\sum_{i=1}^{k} n_i \left(\bar{R}_i - \frac{N+1}{2} \right)^2}_{\text{treatment sum of squares}}$$

** We need a different computation of Kruskal-Wallis statistic

$$\left.\begin{array}{l}\text{The data have potential outliers}\\ \text{The population distributions have heavy tails}\\ \text{The population distributions are significantly skewed}\end{array}\right\} \text{Kruskal-Wallis}$$

$$\left.\begin{array}{l}\text{The population distributions are normal}\\ \text{The population distributions are light tailed}\\ \text{The population distributions are symmetric}\end{array}\right\} \text{F-test}$$

Multiple Comparisons

Previous tests only verify if there are differences among treatments; they cannot verify which treatment/s differ.

Multiple comparisons using pairwise tests allow us to verify which treatment differs from the others, if any.

Three Rank-Based Procedures for Controlling Experiment-Wise Error Rate Assuming No Ties in the Data

1. Bonferroni Adjustment
2. Fisher's Protected Least Significant Difference (LSD)
3. Tukey's Honest Significant Difference (HSD)

Ordered Alternatives

- if treatments are not equal, it may be possible to anticipate the direction in which the treatments differ.
- We are interested in alternative hypotheses in which observations from treatment 1 tend to be smaller than observations from treatment 2, and so on.

$H_a: F_1(X) \geq F_2(X) \geq F_3(X) \geq \cdots \geq F_k(X)$

- If we have shift alternatives, like $F_i(x) = F(x - \mu_i)$, then

$H_a: \mu_1 \leq \mu_2 \leq \mu_3 \leq \cdots \leq \mu_k$

- A general form of a test statistic for testing the hypothesis is the sum of the pairwise statistics

$T = \sum_{i<j} T_{ij}$

Jonckeere - Terpstra Test

- compute T_{JT}, which is the Mann-Whitney statistic, for all possible permutations
- Count the number of statistics that are greater than or equal to the observed T_{JT}

Paired Comparison

Paired - Comparison Permutation Test

- Computing $\bar{D}_i$'s, the mean difference between a pair, for 2^n possible permutations and count, Upper-tail p-value

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

Hypotheses

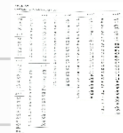
$H_0: F(x) = 1 - F(x)$

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.

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

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 

※ Adjustment for ties and ranking with 0's are not implied on this.

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 $p = 0.5$.

F_1 statistic for Randomized Complete Block Designs

- Suppose the observations follow the model, $X_{ij} = \underbrace{\mu}_{\text{overall effect}} + \underbrace{t_i}_{\text{treatment effect}} + \underbrace{b_j}_{\text{block effects}} + \varepsilon_{ij}$

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

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,

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

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

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

---

The Permutation Test

1. Compute the slope of the least squares line $\hat{\beta}_{1,obs}$ from the original data.
2. Permute the Y's among the X's in the $n!$ possible ways.
3. For each permutation, compute the slope $\hat{\beta}_1$ of the least squares line.
4. For an upper-tail test, the p-value is,

$$P_{upper\text{-}tail} = \frac{\text{number of } \hat{\beta}_1\text{'s} \geq \hat{\beta}_{1,obs}}{n!}$$, we may obtain lower-tail and two-tail p-values

* steps for the test of correlation are the same except that r is used

* Use sampling for large n.

Spearman Rank Correlation

- Suppose $(X_i, Y_i) = 1, \dots, n$, and let $R(X_i)$ and $R(Y_i)$
- The statistical significance of the Spearman rank correlation is determined by applying the permutation test for correlation with ranked pairs

Kendall's Tau :

- This idea leads us to define a measure of association between two variables based on counts of concordant and discordant pairs.

Permutation χ^2 Test

- there are $\binom{n}{r}$ possible permutations. n = # of units, r = # of rows
- compute the χ^2 statistic for each permutation.
- count the number of χ^2 statistic that is greater than or equal to the observed χ^2 statistic

Multiple Comparisons in Contingency Tables

- If the null hypothesis is rejected, then it may be of interest to determine where differences among proportions exist. These comparisons need conditional probabilities.
- For each pair of conditional row probabilities $P_{j|i}$ and $P_{j|i'}$ that we are interested in comparing, we compute the Z statistic for two proportions,

$$Z = \frac{\hat{P}_{j|i} - \hat{P}_{j|i'}}{\sqrt{\bar{P}(1-\bar{P})\left(\frac{1}{n_i}+\frac{1}{n_{i'}}\right)}}, \text{ where } \hat{P}_{j|i} = \frac{n_{ij}}{n_i}, \hat{P}_{j|i'} = \frac{n_{i'j}}{n_i}, \bar{P} = \frac{n_i\hat{P}_{j|i} + n_{i'}\hat{P}_{j|i'}}{n_i + n_{i'}}$$

and suppose there are k number of comparisons of interest, then the Z statistics upon which we base our multiple comparison procedure is, $Q^* = \max_i |Z_i|$

Fisher's Exact Test for a 2X2 Contingency Table

Fisher's Exact Test :

- special case of the permutation test applied to 2x2 contingency table.
- Compute the hypergeometric probabilities of X's possible assignments

Singly Ordered Tables

- run wilcoxon rank-sum test or Kruskal-Wallis test.

- The large difference between the p-value for the Kruskal-Wallis statistic and the χ^2 statistic is due to the type of alternatives that the two statistics are designed to detect.
- χ^2 statistic is designed to detect "any" association, whereas the Kruskal-Wallis statistic is designed to detect an ordering effect.

Mantel - Haenszel Test

- the responses are independent from a stratum to another.
- MH is a measure of how much the total number of observations in row 1 and column 1 deviates from what one would expect under the assumption of independence.

McNemar's Test

- Used for binary responses
- We are willing to know if a treatment affects the binary responses of 2 groups.
- χ^2 approximation using T_1, T_2, T_3, T_4.