2.1 Two-Sample Permutation Test

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.

2.1.1 Permutation Test (don't need the assumptions)

2-sample permutation test

- If there is N samples, m treatment 1 subjects, and n treatment 2 subjects, we need to first calculate all possible combinations $\binom{N}{m}$.

  ☆ because it needs full combinations, it provides the exact probability.

Permution Distribution

- the distribution of $\binom{N}{m}$ differences of means

Summary of steps used in a two-sample permutation Test

1. Calculate $\binom{N}{m} = \frac{N!}{m!(N-m)!}$, m = # of observations designated to new method

2. compute the difference, or any desirable computations, of all combinations.

3. Compute the number of cases where the difference of a combination is greater than or equal to that of the targeted combination.

$$\frac{\text{\# of qualified obs}}{\binom{N}{m}}$$

☆ $D = \frac{T_1}{m} - \frac{T_2}{n} = \frac{T_1}{m} - \frac{T - T_1}{n} = T_1\left(\frac{1}{m} + \frac{1}{n}\right) - \frac{T}{n}$, $T_1 = \sum_{i=1}^{m} X_i$, $T_2 = \sum_{i=1}^{n} X_i$

2.1.3 Hypotheses for the two-sample permutation test.

Two-sided

example $H_0 : F_1(x) = F_2(x)$ "two distributions are identical under the null hypothesis."

$H_a : F_1(x) \leq F_2(x)$ or $F_1(X) \geq F_2(X)$ or $F_1(X) \neq F_2(X)$

2.2 Permutation tests based on the median and trimmed means

Population with symmetric distribution and no outliers $\Rightarrow$ mean
Population with symmetric distribution but some outliers $\Rightarrow$ trimmed mean
population with asymmetric distribution $\Rightarrow$ median

2.3 Random Sampling the permutations

2.3.1 An Approximate p-value based on random sampling the permutations

1. Compute D_{obs}
2. Create a vector of $m+n$ observation
3. randomly assign m and n samples to designated treatment group. (for K times)
4. Compute D_{obs} for each shuffle.
5. Compute

$$D_{obs} \pm Z_{(1-\alpha/2)} \sqrt{\frac{P(1-P)}{K}} , \quad P = \frac{\text{\# of satisfying obs}}{K}$$

as $K \to \infty$, the approximation method $\to$ the exact probability

2.4 Wicoxon Rank-sum Test (No ties)

ties: having obs with same numbers

Appendix A3

Rank:

$R(X_i)$ = number of obs less than or equal to X_i

Obs with rank 1 is the smallest obs.

Obs with rank N is the largest obs.

Wilcoxon Rank-Sum Test

- Let W_i be the sum of the ranks of the observations from treatment i. The test is a two-sample permutation test based on W_i.

Process

1. Combine m+n obs and assign ranks to each.
2. Compute W_1 and W_2 according to the pre-assigned ranks.
3. Find all possible permutations of ranks (Combinations)
4. Compute W_i for each combination.
5. $$P_{upper\ tail} = \frac{\text{number of rank sums} \geq \text{observed rank sum } W}{\binom{m+n}{n}}$$

Rank tests can be used to difference in mean ranks.

$$\text{Difference of mean rank} = W_1\left(\frac{1}{m} + \frac{1}{n}\right) - \frac{N(N+1)}{2n}$$

2.5 Wilcoxon rank-sum test adjusted for ties

Process

- stays the same except adjusting ranks with same numbers to average rank.

If the number of ties is small, then approximate critical values may be obtained from the distribution of the rank-sum statistic without ties.

PDF 59

2.6 Mann-Whitney test and a confidence interval

Appendix A4

Mann-Whitney Statistics "U":

U = number of pairs (X_i, Y_j) for which $X_i < Y_j$

$H_0: F_X(X) = F_Y(Y)$ "the distributions of X and Y are the same"

TABLE 2.6.1
Hours Until Recharge of Batteries

Brand 1	3.6 3.9 4.0 4.3
Brand 2	3.8 4.1 4.5 4.8
Pairs (X_i, Y_j):	(3.6, 3.8) (3.6, 4.1) (3.6, 4.5) (3.6, 4.8)
$U = 12$	(3.9, 3.8) (3.9, 4.1) (3.9, 4.5) (3.9, 4.8)
	(4.0, 3.8) (4.0, 4.1) (4.0, 4.5) (4.0, 4.8)
	(4.3, 3.8) (4.3, 4.1) (4.3, 4.5) (4.3, 4.8)

Process: ← 참고용 $\{n(X_i)\}^2$ U

1. make pairs of (X_i, Y_j) and count the number of pairs $X_i < Y_j$.
2. Using the Mann-Whitney CI table, confirm whether U is within the CI.

Equivalence of Mann-Whitney and Wilcoxon Rank-Sum Statistics

Let,

$$\underbrace{R(Y_j)}_{\text{rank of } Y_j} = \underbrace{(\text{number of } Y\text{'s} \le Y_j)}_{\text{Wilcoxon Rank-Sum}} + \underbrace{(\text{number of } X_i \le Y_j)}_{\text{Mann-Whitney}}, \text{ then}$$

$$W_2 = \sum_{j=1}^{n} R(Y_j) = \underbrace{1+2+3+\cdots+n}_{=\frac{n(n+1)}{2}} + U$$

$$\therefore W_2 = \frac{n(n+1)}{2} + U$$

☆ a value of U or W_2 alone defines one another. Both tests provide a common conclusion.

☆☆ if the data contain ties, Mann-Whitney statistics may be computed by assigning $\frac{1}{2}$ to each pair in which a tie exists.

2.6.3 Confidence Interval for a Shift Parameter and the Hodges-Lehmann Estimate

Let,

$$F_1(X-\Delta) = F_1(Y) \quad \text{or} \quad F_1(X) = F_2(X-\Delta), \text{ then}$$

$$P(X_i \le x) = F_1(x) = F_2(x-\Delta) = P(Y_j \le x-\Delta) = P(Y_j + \Delta \le x)$$

Process

1. Form all pairwise differences of $X_i - Y_j$. mn pairwise differences
2. Arrange all pwd(k)'s in order.
3. Find k_a and k_b that satisfy

$$pwd(k_a) < \Delta \le pwd(k_b) \approx 1-\alpha \%$$

Appendix A4

* if we want 90% CI, apply $k_a = l_{0.05}+1$, $k_b = U_{0.05}$ in Mann-Whitney table

** for samples exceeding the Mann-Whitney table, normal approximation is used.

Hodges - Lehmann Estimate of Δ :

the median of all the pairwise differences of $X_i - Y_j$

* we may alternatively use pooled t-test, meaning

$$\bar{X}_1 - \bar{X}_2 \pm t_{1-\alpha} S_e \sqrt{\frac{1}{m}+\frac{1}{n}}, \quad m+n-2 \text{ df and}$$

$$S_e = \sqrt{\frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}}$$

2.7 Scoring Systems

Normal Scores : $S_i = E(Z_i)$ the expected value of Z_i is the i^{th} normal score

Appendix A5 Van der Waerden Scores : $V_i = \Phi^{-1}\left(\frac{i}{N+1}\right)$

inverse of CDF of standard normal distribution

the value of Z from standard normal distribution that has cumulative probability $\frac{i}{N+1}$

Exponential or Savage Scores : $\underbrace{\frac{1}{N}, \frac{1}{N}+\frac{1}{N-1}, \frac{1}{N}+\frac{1}{N-1}+\frac{1}{N-2}, \dots}_{\text{sum up to } 0}$

2.8 Tests for Equality of Scale Parameters and an Omnibus Test

Siegel-Tukey Test :

assume we have two distributions, $X_i = \mu + \sigma_1 \varepsilon_{iX}$

$Y_j = \mu + \sigma_2 \varepsilon_{iY}$, so

이쩔 때
Siegel-Tukey
Test 가
부적절한지
모르겠음
pdf 69 에 설명이
있긴한데
못알아듣겠음

$H_0: \sigma_1 = \sigma_2$

Process :

1. Arrange the observations in the combined dataset from smallest to largest.
2. 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.
3. Apply the Wilcoxon rank-sum test. The smaller ranks and smaller rank sum are associated with the treatment that has the larger variability.

Alternative for Siegel-Tukey Test

Ansari-Bradley Test :

Same process as Siegel-Tukey Test except for the CI table and, assign 1 to the largest and smallest observation, 2 to the next largest and smallest observation, ... so on.

※ but can no longer use the tables of Wicolxon rank-sum test.

2.8.2 Tests on deviances

Suppose X and Y have different location parameters,

$$X_i = \mu_1 + \sigma_1 \varepsilon_{ix}$$

$$Y_j = \mu_2 + \sigma_2 \varepsilon_{jx}$$

, from here we may use F-statistics under ε's normality assumption but if not, it may leads to Type I Error (False Negative). Alternatively,

$$F = S_1^2 / S_2^2$$

$$dev_{ix} = X_i - \mu_1$$

$$dev_{jy} = Y_j - \mu_2$$

(Ratio of absolute Mean Deviances) $RMD = \dfrac{\sum_{i=1}^{m} |dev_{ix}| / m}{\sum_{j=1}^{n} |dev_{jy}| / n}$

Process of Permutation Test on Deviances (Known Location Parameters)

1. Obtain dev_{ix} and dev_{jy} and compute RMD from the original data.
2. Permute the deviances among two treatments and obtain RMD for each permutation
3. For small values of m and n, obtain all permutations; otherwise, samples.
4. p-value is the fraction of the permutations that $RMD_i \geq RMD_{obs}$ or $RMD_i \leq RMD_{obs}$

upper tailed lower tailed

Modifications for a 2-sided test:

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

Process:

- stays the same except the calculations of RMD's

PDF 73

2.8.3 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)

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

Process

- allign the data in an increasing order and specify where each obs belongs to.
- Count one by one and check where the obs belongs.
- calculate the absolute difference for each step.
- find the maximum absolute difference

TABLE 2.8.2
Computation of K-S Statistic for Data in Table 2.8.1

Combined Data	15.36	15.88	15.91	15.94	15.98	16.01	16.05	16.10	16.43	16.55
Treatment	(1)	(2)	(2)	(1)	(2)	(1)	(2)	(2)	(1)	(1)
cdf Treatment 1	1/5	1/5	1/5	2/5	2/5	3/5	3/5	3/5	4/5	1
cdf Treatment 2	0	1/5	2/5	2/5	3/5	3/5	4/5	1	1	1
Absolute Difference	1/5	0	1/5	0	1/5	0	1/5	2/5	1/5	0

To determine exact p-value, we need to find the Kolmogorov - Smirnov statistics for all possible permutation, and then determine the proportion of those values that are greater than or equal to the observed value.

2.9 Selecting Among Two-Sample Tests

2.9.1 The t-Test

- Under conditions that we know the population distribution is normal with equal variances, t-test has the correct probability of a Type I error and has the greatest power among unbiased tests.

※ modest violations of assumptions may not be critical to the results, though.

2.9.2 The Wilcoxon Rank-Sum Test versus the t-Test

- Wilcoxon rank-sum tests have advantage over t-test when either you don't know the distributions or dealing with outliers, or heavy-tailed or skewed
- unusually large or small observations compared to the rest of the data can have adverse effects on the power of the t-test.
- t-test has greater power than the Wilcoxon test for small samples and light-tailed distributions.

2.9.3 Relative Efficiency

$$RE(1,2) = \frac{N_2}{N_1}$$ "the relative efficiency of test 1 to test 2"

Asymptotic Relative Efficiency (ARE)

- measure of large-sample efficiency of one test relative to the other.

TABLE 2.9.2
Asymptotic Relative Efficiency of Wilcoxon Rank-Sum Test to *t*-Test

Distribution	*Efficiency*
Uniform	1.0
Normal	0.955
Laplace	1.5
Exponential	3.0
Cauchy	∞

2.9.4 Power of Permutation Tests

2.10 Large-Sample Approximation

2.10.1 Sampling Formulas

$$E(T_1) = m\mu, \quad Var(T_1) = \frac{mn\sigma^2}{N-1}, \quad \mu = \frac{\sum^N A_i}{N}$$

$$\sigma^2 = \frac{\sum_{i=1}^N (A_i - \mu)^2}{N} = \frac{\sum_{i=1}^N A_i^2}{N} - \mu^2, \quad Z = \frac{T_1 - E(T_1)}{\sqrt{Var(T_1)}}$$

2.10.2 Application to the Wilcoxon Rank-Sum Test

Formulas for Wilcoxon Rank-Sum Test

$\sum_{i=1}^{N} i = \frac{N(N+1)}{2}$, $\sum_{i=1}^{N} i^2 = \frac{N(N+1)(N+2)}{6}$

$$\sigma^2 = \frac{1}{N}\left(\sum_{i=1}^{N} i^2\right) - \left(\frac{1}{N}\sum_{i=1}^{N} i\right)^2$$
$$= \frac{(N-1)(N+1)}{12}$$

Expected Value and Variance of rank-sum statistic (from treatment with m obs)

$$E(W) = \frac{m(N+1)}{2}$$

$$Var(W) = \frac{mn(N+1)}{12}$$