BMI 713 / GEN 212

Lecture 4: Nonparametric Methods

- Wilcoxon Signed Rank Test
- Wilcoxon Rank Sum Test
- Parametric vs nonparametric tests

September 30, 2010

Nonparametric Test

- Assumptions in the *t*-test
  - normal distribution of the underlying population
  - or large sample size
- If the assumptions are not satisfied, the *t*-statistic does not follow a *t*-distribution closely.

- **Parametric tests** – based on the assumption that the shape of the population is known
- **Nonparametric tests** – fewer restriction on the underlying distribution of data; do not rely on the Central Limit Theorem
  - **Uses ranks rather than actual values**

Nonparametric Test

- Sometimes transforming the data (e.g., log) can make the data look more normal.

- A nonparametric test is **robust** to non-normality. This means the results are valid even when the distributions are highly non-normal.
  - Less sensitive to measurement error
  - Less sensitive to outliers
  - But loss of information

- Nonparametric tests is on the population median, not population mean
- Median is robust to outliers

Nonparametric Test

- So if there is a method that makes fewer assumptions, why not use that all the time?
- If the data really do follow normal distribution, there is loss of power!
- That means if H_0 is false, a larger sample size would be needed to provide sufficient evidence to reject it

- Steps are the same as before:
  - Specify the null and alternative hypotheses
  - Select a random sample of observations
  - Calculate a test statistic
  - Based on the value of the test statistic, either reject or do not reject H_0

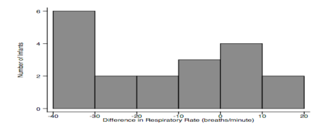
Wilcoxon Signed-Rank Test

- A nonparametric version of the paired t-test
- Like the paired t-test, this test focuses on the difference in values for each pair of observations
- Example: Respiratory rates in premature infants
  - A study was conducted to examine the effects of the transition from fetal to postnatal circulation in premature infants
  - Respiratory rate was measured for each infant at two different times: <15 days old and >25 days old
  - We wish to compare the respiratory rates measured at two different times in the infants' development

Wilcoxon Signed-Rank Test

- Let Δ be the true median difference in measurements
- H_0: $\Delta = 0$ vs H_A: $\Delta \neq 0$
- We first calculate the difference for each pair

rr1	rr2	diff	rr1	rr2	diff
48	47	-1	68	45	-23
38	40	2	67	31	-36
43	46	3	80	42	-38
48	42	-6	88	48	-40
35	42	7	84	45	-39
52	44	-8	75	38	-37
48	36	-12	67	44	-23
27	40	13	45	45	0
26	40	14	70	35	-35
62	46	-16			



Wilcoxon Signed-Rank Test

- We rank the absolute values of the differences from smallest to largest
- Pairs for which the difference is 0 are not ranked
  - These pairs provide no information about which group has higher or lower values
  - Exclude these pairs and adjust the sample size accordingly
- Tied observations are assigned an average rank

diff	rank	diff	rank
0	-	-16	10
-1	1	-23	11.5
2	2	-23	11.5
3	3	-35	13
-6	4	-36	14
7	5	-37	15
-8	6	-38	16
-12	7	-39	17
13	8	-40	18
14	9		

Sign Test: Simply compare the number of positive signs to the expected number under H_0

Wilcoxon Signed-Rank Test

- Each rank is assigned a plus or a minus sign, depending on the sign of the difference
- Under the null hypothesis, the sum of the positive ranks should be comparable in magnitude to the sum of the negative ranks
- The statistic: the sum of the positive ranks, denoted by T
- We reject H_0 if T is either too big or too small
- If H_0 is true and sample size is large,

$$\boxed{Z_T = \frac{T - \mu_T}{\sigma_T}}$$

$$\mu_T = \frac{n(n+1)}{4}$$

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

follows a N(0,1) distribution

Example

- The formula are slightly different when there are ties

- Back to the study on respiratory rates for premature infants (n=18)
- The sum of positive ranks:
  T = 2+3+5+8+9
- Under H_0, $\mu_T = n(n+1)/4 = 85.5$
- $Var(T) = n(n+1)(2n+1)/24 - C = 527.125$
- $Z_T = (27-85.5)/\mathrm{sqrt}(527.125) = -2.55$

diff	rank	diff	rank
0	-	-16	10
-1	1	-23	11.5
2	2	-23	11.5
3	3	-35	13
-6	4	-36	14
7	5	-37	15
-8	6	-38	16
-12	7	-39	17
13	8	-40	18
14	9		

- Since Z_T<-1.96, we reject H_0 and conclude that the median difference is not equal to 0.

Wilcoxon Rank Sum Test

- Analogous to the two-sample **unpaired** *t*-test
- Also called Mann-Whitney Test

- The null hypothesis being tested is that the median of the first population is equal to the median of the second population

- Like the signed-rank test, the rank sum test is performed on ranks rather than actual measurements

Example

- We test the supplement's effect on SBP by giving supplement to one sample and placebo to an independent sample.
- Data:
  - Sample 1 (supplement): 118, 148, 110, 120, 140
  - Sample 2 (placebo): 135, 116, 120, 145
- 1) combine samples into one group and order them
  - Ordered values: 110, 116, 118, 120, 120, 135, 140, 145, 148
- 2) Assign ranks from lowest to highest; assign average rank to tied values
  - Ranks: 1, 2, 3, 4.5, 4.5, 6, 7, 8, 9
- 3) Compute statistic
  - W= 2 + 4.5 + 6 + 8 = 20.5

Wilcoxon Rank Sum Test

- H0: median in group 1 = median in group 2
- HA: median in group 1 ≠ median in group 2
- If H0 is true and sample size is large enough,

$$\boxed{Z_W = \frac{W - \mu_W}{\sigma_W}}$$

$$\mu_W = \frac{n_S(n_S + n_L + 1)}{2}$$

$$\sigma_W = \sqrt{\frac{n_S n_L(n_S + n_L + 1)}{12}}$$

follows a N(0,1) distribution

More Groups

- You may wish to comparing **three or more** populations without assuming that the underlying data are not normally distributed
- The **Kruskal-Wallis Test**

- Paired t-test → Wilcoxon Signed Rank Test
- Unpaired t-test → Wilcoxon Rank Sum Test
- ANOVA → Kruskal-Wallis Test

R programming

```
wilcox.test(x, y = NULL,
  alternative = c("two.sided", "less",
  "greater"), mu = 0, paired = FALSE,
  exact = NULL, correct = TRUE,
  conf.int = FALSE, conf.level = 0.95, ...)
```