Lecture 10: Two-sample problem II

Pratheepa Jeganathan

04/24/2019

Recall

- One sample sign test, Wilcoxon signed rank test, large-sample approximation, median, Hodges-Lehman estimator, distribution-free confidence interval.
- Jackknife for bias and standard error of an estimator.
- Bootstrap samples, bootstrap replicates.
- Bootstrap standard error of an estimator.
- Bootstrap percentile confidence interval.
- Hypothesis testing with the bootstrap (one-sample problem.)
- Assessing the error in bootstrap estimates.
- Example: inference on ratio of heart attack rates in the aspirin-intake group to the placebo group.
- The exhaustive bootstrap distribution
- Discrete data problems (one-sample, two-sample proportion tests, test of homogeneity, test of independence)
- Two-sample problems (location problem - equal variance, unequal variance, exact test or Monte Carlo, large-sample approximation, H-L estimator)

Two-sample location problem

Behrens-Fisher problem

- Let $X_1, \cdots X_m$ and $Y_1, \cdots Y_n$ be independent random samples from continuous distributions that are symmetric about the population medians θ_X and θ_Y, respectively.
- Behrens-Fisher problem: testing $\mathrm{H}_0 : \theta_X = \theta_Y$ without assuming equal variances.

Behrens-Fisher problem (Fligner-Policello)

- $P_i =$ [number of sample $\boldsymbol{Y}$ observations less than X_i] $, i = 1, \cdots, m$.
- $Q_i =$ [number of sample $\boldsymbol{X}$ observations less than Y_j] $, j = 1, \cdots, n$.
- Average $\boldsymbol{X}$ sample placement $\bar{P} = \dfrac{\sum_{i=1}^{m} P_i}{m}$.
- Average $\boldsymbol{Y}$ sample placement $\bar{Q} = \dfrac{\sum_{j=1}^{n} Q_j}{n}$.
- $V_1 = \sum_{i=1}^{m} \left(P_i - \bar{P}\right)^2$.
- $V_2 = \sum_{j=1}^{n} \left(Q_j - \bar{Q}\right)^2$.
- The standardized test-statistic $\hat{U} = \dfrac{\sum_{j=1}^{n} Q_j - \sum_{i=1}^{m} P_i}{2\sqrt{V_1 + V_2 + \bar{P}\bar{Q}}}$.
- $\hat{U}$ has a symmetric distribution.
- $\hat{U}$ resembles Welch's t statistic for the normal theory when variances are unequal.

Behrens-Fisher problem (Fligner-Policello)

- Let u_α is the upper α quantile of $\hat{U}$.
- u_α can be computed exactly or estimated using Monte Carlo simulation using the R command `cFligPoli`.
- To do exact test (either exact or Monte Carlo)

```
pFligPoli(x = X, y = Y,
  method = "Monte Carlo")
```

Behrens-Fisher problem (large-sample approximation)

- $\hat{U}$ has an asymptotic ($\min(m, n) \to \infty$) $\mathrm{N}(0, 1)$ distribution.

```
pFligPoli(x = X, y = Y,
  method = "Asymptotic")
```

Example (Behrens-Fihser problem)

- **HWC**: Example 4.5 Plasma Glucose in Geese.
- Examining whether plasma glucose in lead-poisoned geese are greater than plasma glucose in healthy (normal) Canadian geese.
- Measured plasma glucose in eight healthy and seven lead-poisoned geese.

```r
# Table4.7
healthy.geese = c(297,340,325, 227,
  277, 337, 250, 290)
lead.poisoned.geese = c(293, 291, 289,
  430, 510, 353, 318)
```

- Let X be plasma glucose in healthy geese and Y be plasma glucose in lead-poisoned geese.

Example (Behrens-Fihser problem)

- Let θ_X be the location parameter of plasma glucose in healthy geese and θ_Y be the location parameter of plasma glucose in lead-poisoned geese.
- Hypothesis: $\mathrm{H}_0 : \theta_X = \theta_Y$ versus $\mathrm{H}_A : \theta_X < \theta_Y$.
- Let's use P-value approach:
- $\binom{8+7}{7} = 6435 < 10,000$ can do exact test.

```r
library(NSM3)
pFligPoli(x = healthy.geese,
  y = lead.poisoned.geese,
  method = "Exact")
```

```
## Number of X values: 8 Number of Y values: 7
## Fligner-Policello U Statistic: 1.4676
## Exact upper-tail probability: 0.0808
## Exact two-sided p-value: 0.1616
##
```

Example (Behrens-Fihser problem)

- Because $\hat{U}$ is symmetric, exact lower-tail probability $=$ P-value $= .0808$.
- We do not have enough evidence to reject H_0 at 5% significance level.

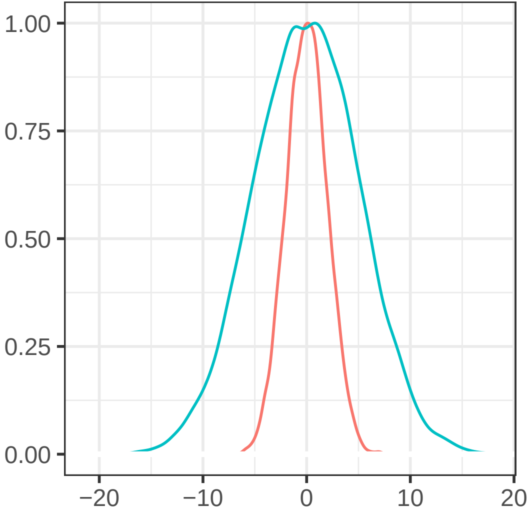
Other two-sample problems

Dispersion test (medians equal - ANSARI–BRADLEY)

- Data: $X_1, \cdots X_m \sim F(.)$ and $Y_1, \cdots Y_n \sim G(.)$.
- Assumption
  - X and Y are mutually independent.
  - $F(.)$ and $G(.)$ are continuous.
- Let θ_X, θ_Y be the population medians for the X and Y distributions.
- Let η_X, η_Y be the scale parameters associated with X and Y distributions.

Dispersion test (medians equal - ANSARI–BRADLEY)

- Example for the probability distributions with the same general form and equal medians but different scale parameters.



Dispersion test (medians equal - ANSARI–BRADLEY)

- Parameter of interest: $\gamma = \frac{\eta x}{\eta x}$.
- $\mathrm{H}_0 : \gamma^2 = 1$ versus $\mathrm{H}_A : \gamma^2 \neq 1$.
- Other alternative: $\mathrm{H}_A : \gamma^2 > 1$, $\mathrm{H}_A : \gamma^2 < 1$.
- Ansari–Bradley two-sample scale statistic C.
    - Assign the score 1 to both the smallest and largest observations in the combined sample.
    - Assign the score 2 to the second smallest and second largest, and continue in the manner.
    - Let R_j denote the score assigned in this manner to $Y_j, j = 1, \cdots, n$.
    - $C = \sum_{i=1}^{n} R_j$.

Dispersion test (medians equal - ANSARI–BRADLEY)

```
pAnsBrad(x = X, y = Y,
  method = "Exact")
```

- If there are ties among the X and/or Y observations, assign each of the observations in a tied group the average of the integer scores that are associated with the tied group.
  - text will not be exact.

Dispersion test (medians equal - large-sample approximation)

- $\mathbb{E}_0\left(C\right) = \dfrac{n(N+2)}{4}$ - expected value of C under H_0 and medians are equal.
- $\mathbb{V}_0\left(C\right) = \dfrac{mn(N+2)(N-2)}{48(N-1)}$ - variance of C under H_0 and medians are equal.
- $Z = \dfrac{C - \mathbb{E}_0\left(C\right)}{\sqrt{\mathbb{V}_0\left(C\right)}} \sim \mathrm{N}\left(0, 1\right)$.

```r
pAnsBrad(x = X, y = Y,
  method = "Asymptotic")
```

Example (Dispersion test)

- **HWC** Example 5.1 (Serum Iron Determination)
- From the point of view of procedural technique, the Jung–Parekh method competes favorably with the Ramsay method for serum iron determination.
- Test whether loss of accuracy when the Jung–Parekh procedure is used instead of the Ramsay procedure.
  - The alternative of interest in this example is greater dispersion or variation for the Jung–Parekh method of serum iron determination than for the method of Ramsay.
  - $\mathrm{H}_A : \gamma^2 > 1$.

```
serum = list(ramsay = c(111, 107, 100, 99, 102,
  106, 109, 108, 104, 99, 101, 96, 97, 102, 107,
  113, 116, 113, 110, 98),
jung.parekh = c(107, 108, 106, 98, 105, 103,
  110, 105, 104, 100, 96, 108, 103, 104, 114,
  114,113, 108, 106, 99))
```

```r
pAnsBrad(serum$ramsay, serum$jung.parekh,
  method = "Asymptotic")
```

```
## Ties are present, so p-values are based on conditional n
## Number of X values: 20 Number of Y values: 20
## Ansari-Bradley C Statistic: 234.5
## Asymptotic upper-tail probability: 0.9093
## Asymptotic two-sided p-value: 0.1815
##
```

- P-value for the upper-tail test is .9093.
- There is absolutely no evidence in the sample data to indicate any loss of accuracy with the Jung–Parekh method.

Dispersion problem (medians are unequal - MILLER(JACKKNIFE))

- Read **HWC 5.2**
- Ties: no adjustments are necessary - the jackknife procedures are well defined when ties within or between the X's and Y's occur.
- Compute Miller Jackknife Q statistic.

```
Q = MillerJack(x = X, y = Y)
```

- One-sided P-value (when n, m small.)

```
1 - pt(Q)
```

- One-sided P-value (when n, m large.)

```
1 - pnorm(Q)
```

General distribution test (KOLMOGOROV SMIRNOV)

- $X_1, \cdots X_m \sim F(.)$ and $Y_1, \cdots Y_n \sim G(.)$.
- $\mathrm{H}_0 : F(t) = G(.) \quad \forall t$ versus
  $\mathrm{H}_A : F(t) \neq G(.)$ for at least one t.
- Define empirical distribution functions for the X and Y samples.
  - $F_m(t) = \dfrac{\#\{X's \leq t\}}{m}$.
  - $G_n(t) = \dfrac{\#\{Y's \leq t\}}{n}$.
- Define d is the greatest common divisor of m and n.

```r
library(FRACTION)
gcd(4,3)
```

```
## [1] 1
```

General distribution test (KOLMOGOROV SMIRNOV)

- Kolmogorov–Smirnov general alternative (two-sided) statistic J
  - $J = \frac{mn}{d} \max_{i=1,\cdots,N} \{|F_m(Z_{(i)}) - G_n(Z_{(i)})|\}$, where $Z_{(1)}, \cdots, Z_{(N)}$ are ordered values for the combined sample.
- Reject H_0 if $J \geq j_\alpha$, where j_α is the upper α percentile of J under H_0.
  - Due to discreteness, j_α is not defined for all α values.
- Ties: no adjustments are necessary because empirical distributions $F_m(t)$ and $G_n(t)$ are well defined.

```
pKolSmirn(x= X, y = Y,
  method = "Exact")
```

General distribution test (large-sample approximation)

- Smirnov (1939)
- The large-sample $(\min(m,n) \to \infty)$ approximation is based on the asymptotic distribution of J.
- $J^* = \frac{d}{\sqrt{mnN}} J \sim Q(\cdot)$.
- For the alternative $\mathrm{H}_A : F(t) \neq G(.) \quad \text{for at least one} \quad t$, reject H_0 if $J^* \geq q_\alpha$.

```
pKolSmirn(x = X, y = Y,
  method = "Asymptotic")
```

Example (General distribution test)

- **HWC** Example 5.4 (Effect of Feedback on Salivation Rate.)
- Interest: The effect of enabling a subject to hear himself salivate while trying to increase or decrease his salivary rate.
- Experiment: Two groups of subjects were told to attempt to increase their salivary rates upon observing a light to the left and decrease their salivary rates upon observing a light to the right.
- Data: collected amount of saliva on feedback and no-feedback groups.

```
Table5.7 = list(feedback = c(-0.15, 8.6, 5, 3.71,
  4.29, 7.74, 2.48, 3.25, -1.15, 8.38),
  no.feedback = c(2.55, 12.07, 0.46, 0.35, 2.69,
    -0.94, 1.73, 0.73, -0.35, -0.37))
```

Example (General distribution test)

- Exact test

```r
pKolSmirn(x = Table5.7$feedback,
  y = Table5.7$no.feedback, method = "Exact")
```

```
## Number of X values: 10 Number of Y values: 10
## Kolmogorov-Smirnov J Statistic: 6
## Exact upper-tail probability: 0.0524
```

We conclude that there is some marginal evidence in the samples that feedback might have an effect on salivation rate.

Example (General distribution test)

- The large-sample test

```r
pKolSmirn(x = Table5.7$feedback,
  y = Table5.7$no.feedback, method = "Asymptotic")
```

```
## Number of X values: 10 Number of Y values: 10
## Kolmogorov-Smirnov J* Statistic: 1.3416
## Asymptotic upper-tail probability: 0.0546
```

Using large-sample test, we reach the same conclusion that there is some marginal evidence in the samples that feedback might have an effect on salivation rate.

Summary

- Testing procedure
  - Two-sample location problem (variance equal) (Wilcoxon rank sum test).
  - Two-sample location problem (variance unequal) (Fligner-Policello).
  - Two-sample dispersion problem (median equal) (ANSARI–BRADLEY).
  - Two-sample dispersion problem (median unequal) (Jackknife-Miller).
  - General distribution test (KOLMOGOROV SMIRNOV).
- In practice
  - Use boxplot to decide equal median or not.
  - Test for equal distribution.
  - Test for dispersion.
  - Choose appropriate location test.
  - If sample size is larger, choose large sample approximation.

References for this lecture

HWC Chapter 4.4

HWC Chapter 5