Lecture 11: Permutation tests

Pratheepa Jeganathan

04/26/2019

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, dispersion problem, general distribution)

Permutation tests

Permutation tests

- The Mann-Whitney/Wilcoxon test is actually a special case of a permutation test.
  - Re-assignment of combined ranks with probability of each re-assignment is $\frac{1}{\binom{N}{n}}$ under H_0.
- Permutation test are computationally-intensive but exists before computers.
- R. A Fisher (1930's) introduced to support theoretical argument of Student's t-test.

Permutation test

- Two-sample problem
  - $\boldsymbol{X} = (X_1, \cdots, X_m)$ and $\boldsymbol{Y} = (X_1, \cdots, X_m)$ are drawn from $F(\cdot)$ and $G(\cdot)$, respectively.
- Test $\mathrm{H}_0 : F = G$.
  - If H_0 is true, any of the observations could have come equally well from either F or G.

Permutation test recipe (two-sample problem)

- Decide a test statistic T.
- Compute the test statistic for the given data t_0.
- H_0 assigns equal probabilities to all possible re-assignments of combined observations.
- P-value is defined to be the probability of observing at least that large a value when the null hypothesis is true.

$$\text{p-value} = P_{\mathrm{H}_0}\left(T^* \geq t_0\right),$$

  where t_0 is fixed and the random variable T^* has the null hypothesis distirbution, the distribution of T if H_0 is true.
- The smaller value of p-value, the stronger the evidence against H_0.

Example (Permutation test)

- The mouse data in **ET** page 11, table 2.1.
  - Sixteen mice were randomly assigned to a treatment group or a control group.
  - Measured their survival times, in days, following a test surgery.
  - Did the treatment prolong survival?

```
ET.table.2.1 = list(treatment =
    c(94, 197, 16, 38, 99, 141,23),
  control =
    c(52, 104, 146, 10, 51, 30, 40, 27, 46))
ET.table.2.1
```

```
## $treatment
## [1]  94 197  16  38  99 141  23
##
## $control
## [1]  52 104 146  10  51  30  40  27  46
```

Example (Permutation test)

- Let X be survival times of mice in treatment group and Y be survival times of mice in control group.
- Observed small amount of data

```r
lapply(ET.table.2.1, function(x){length(x)})
```

```
## $treatment
## [1] 7
##
## $control
## [1] 9
```

Example (Wilcoxon rank sum test)

- Use Wilcoxon test (dispersion is equal)

```
wilcox.test(x = ET.table.2.1$treatment,
  ET.table.2.1$control, alternative = "greater",
  paired = FALSE, exact = TRUE)
```

```
##
##  Wilcoxon rank sum test
##
## data:  ET.table.2.1$treatment and ET.table.2.1$control
## W = 36, p-value = 0.3403
## alternative hypothesis: true location shift is greater t
```

Example (Wilcoxon rank sum test)

- Use permutation to find the null distribution of Wilcoxon rank sum test statistic.
- The exact p-value is computed using this null distribution.

Example (Mean difference test statistic)

- Permutation test for the general hypothesis $F = G$.
  - Assume F and G only differ in location.
- $T = \bar{X} - \bar{Y}$. Test statistic is mean difference.
- t_0

```
t.0 = round(mean(ET.table.2.1$treatment) -
  mean(ET.table.2.1$control), digits = 2); t.0
```

```
## [1] 30.63
```

- This indicates treatment distribution F gives longer survival times than does control distribution G.
- Is it significant?

Example (Mean difference test statistic)

- Find the all possible re-assignments of combined observations under H_0.
  - Combine all the $N = m + n$ observations from both treatment and control.
  - Take a sample of size m without replacement to treatment group.
  - Assign remaining n observations to control.
  - There are $\binom{N}{n}$ possible re-assignments with $\frac{1}{\binom{N}{n}}$ probability.

Example (Mean difference test statistic)

- Order statistic representation

```
library(dplyr)
combined.sample = data.frame(group = c(rep("treatment",
  length(ET.table.2.1$treatment)),
  rep("control",
    length(ET.table.2.1$control))),
  value = c(ET.table.2.1$treatment,
    ET.table.2.1$control))

combined.sample = mutate(combined.sample,
  rank.combined = rank(value))
combined.sample =
  combined.sample[order(combined.sample$rank.combined,
    decreasing = FALSE), ]
```

Example (Mean difference test statistic)

```
combined.sample
```

```
##          group value rank.combined
## 11    control    10             1
## 3   treatment    16             2
## 7   treatment    23             3
## 15    control    27             4
## 13    control    30             5
## 4   treatment    38             6
## 14    control    40             7
## 16    control    46             8
## 12    control    51             9
## 8     control    52            10
## 1   treatment    94            11
## 5   treatment    99            12
## 9     control   104            13
## 6   treatment   141            14
```

Example (Mean difference test statistic)

- Find all possible re-assignments.

```
library(gtools)
all.possible.assignments  = combinations(n = dim(combined.s
  r = length(ET.table.2.1$treatment),
  v = combined.sample$rank.combined)
saveRDS(all.possible.assignments, "all.possible.assignments
```

Example (Mean difference test statistic)

```
all.possible.assignments  = readRDS("all.possible.assignmen

compute.T.star = function(combined.sample,
  all.possible.assignments){

  T.star = apply(all.possible.assignments, 1,
    function(x){
    x.star  = combined.sample$value[x]
    y.star  = combined.sample$value[-x]
    mean(x.star) - mean(y.star)
  })

  return(T.star)
}
T.star = compute.T.star(combined.sample = combined.sample,
  all.possible.assignments = all.possible.assignments)
```

Example (Mean difference test statistic)

- Exact p-value

```
p.value.exact = mean(T.star >= t.0)
round(p.value.exact, digits = 3)
```

```
## [1] 0.141
```

Example (Permutation test in practice)

- Choose B possible re-assignments, each being randomly selected from the set of all $\binom{N}{n}$ possible re-assignments [B usually be at least 1000].

```
B = 1000
compute.T.star.Monte.Carlo = function(x,
  combined.sample){
  re.assignment.index = sample(combined.sample$rank.combine
    size = length(combined.sample$rank.combined),
    replace = FALSE)

  x.star = combined.sample$value[re.assignment.index[1:leng
  y.star  = combined.sample$value[re.assignment.index[(leng
  T.star = mean(x.star) - mean(y.star)
  return(T.star)
}
```

```r
T.star = lapply(seq_along(1:B),
  FUN = compute.T.star.Monte.Carlo,
  combined.sample = combined.sample)
T.star = unlist(T.star)
```

Example (Permutation test in practice)

- P-value using Monte Carlo method

```
P.value.MC = mean(T.star >= t.0)
round(P.value.MC, digits = 3)
```

```
## [1] 0.148
```

Example (Permutation test based on Student's t test statistic)

- $T = \dfrac{\bar{X} - \bar{Y}}{\sigma\sqrt{\dfrac{1}{m} + \dfrac{1}{n}}}$. - We don't know σ and an estimate for σ is $\bar{\sigma}$, the standard deviation of combined sample.

```
sigma.bar = sd(c(ET.table.2.1$treatment,
  ET.table.2.1$control)); round(sigma.bar, digit = 2)
```

```
## [1] 54.7
```

```
student.t.0 = (mean(ET.table.2.1$treatment) - mean(ET.table
round(student.t.0, digits = 2)
```

```
## [1] 1.11
```

Example (Permutation test based on Student's t test statistic)

```
B = 1000
compute.T.star.Monte.Carlo = function(x,
  combined.sample){
  re.assignment.index = sample(combined.sample$rank.combine
    size = length(combined.sample$rank.combined),
    replace = FALSE)

  x.star = combined.sample$value[re.assignment.index[1:leng
  y.star  = combined.sample$value[re.assignment.index[(leng

  sigma.bar.star = sd(c(x.star, y.star))
  student.t.star = (mean(x.star) - mean(y.star))/(sigma.bar

  T.star = student.t.star
  return(T.star)
}
```

Example (Permutation test based on Student's t test statistic)

```r
T.star = lapply(seq_along(1:B),
  FUN = compute.T.star.Monte.Carlo,
  combined.sample = combined.sample)
T.star = unlist(T.star)
```

Example (Permutation test based on Student's t test statistic)

```
p.value.student.t.test.MC = mean(T.star >= student.t.0)
round(p.value.student.t.test.MC, digits = 3)
```

```
## [1] 0.141
```

Example (Compare the permutation test results using different test statistics)

- Rank-based permutation test: exact p-value with $T = W$ is .340. (Rank-based permutation test)
- Exact p-value with $T = \bar{X} - \bar{Y}$ is .141.
- P-value with $T = \bar{X} - \bar{Y}$ and using Monte Carlo method is

```
round(P.value.MC, digits = 3)
```

```
## [1] 0.148
```

- P-value with $T = \dfrac{\bar{X} - \bar{Y}}{\sigma\sqrt{\dfrac{1}{m} + \dfrac{1}{n}}}$ and using Monte Carlo method is

```
round(p.value.student.t.test.MC, digits = 3)
```

```
## [1] 0.141
```

Permutation and bootstrap hypothesis tests

- Permutation tests
  - Define F_0 the null distirbution of test statistics using the possible re-assginments of observations.
  - Use for more general test of $F(t) = G(t)$.
  - Important aspect of permutation test is its accuracy.
    - If $\mathrm{H}_0 : F = G$ is true, $P_{\mathrm{H}_0}\{\text{p-value}_{\text{perm}} < \alpha\} = \alpha$.
- Bootstrap testing.
  - Uses plug-in estimator for F_0. Denote the combined sample $\boldsymbol{z}$ and let its empirical distirbution be $\hat{F}_0$, putting probability $\dfrac{1}{(n+m)}$ on each member of $\boldsymbol{z}$.
  - Can be used for many statistical problems when there is nothing to permute.
  - Bootstrap p-value is approximate.

Exchangeability

- A sufficient condition for permutation test is exchangeable of observations.
  - Consider random sample $X_1, \cdots X_n$.
  - If their joint distribution are equal under permuations Π $P_{X_1,\cdots X_n}(x_1, \cdots, x_n) = P_{X_{\Pi(1)},\cdots X_{\Pi(n)}}\left(x_{\Pi(1)}, \cdots, x_{\Pi(n)}\right)$, then $X_1, \cdots X_n$ are exchangable.
- This is a weaker assumption than indepdendence of observations.

References for this lecture

ET Chapter 15

Li:H1997: Holmes (1997). Lecture Notes on Computer Intensive Methods in Statistics.