

# Resampling Methods Part (2)

Ragheed Al-Dulaimi, MD, MPH

December 17, 2014

# Resampling

- Population  $\rightarrow$  Sample  $\rightarrow$  Inference
- Resampling:
  - Resample from the original sample
  - No information on population characteristics or distribution
  - No assumption on distribution of the data
- Estimation (e.g. bias), hypothesis testing, model assessment

---

Type of Sample	Without Replacement	With Replacement
Full sample	Permutation Test	Bootstrap
Subsample	Cross-Validation, Jackknife	

---

## Cross-Validation

- **Procedure:** Data is randomly divided into subsets. Results validated across sub-samples.
- **Application:** Model Validation

## Jackknife

- **Procedure:** Samples consist of full data set, with one observation left out ( $n-1$ ).
- **Application:** Standard deviation, confidence interval, bias

## Bootstrap

- **Procedure:** Samples drawn at random with replacement.
- **Application:** Standard deviation, confidence interval, bias , hypothesis testing

## Permutation Test

- **Procedure:** Samples drawn at random without replacement.
- **Application:** Hypothesis testing

# Permutation Test

- The Permutation test is a technique that bases inference on “experiments” within the observed dataset. i.e. Permutation test create a null distribution by random permutation of the data
- In classical hypothesis testing, we start with assumptions about the underlying distribution and then derive the sampling distribution of the test statistic under  $H_0$ .
- In Permutation testing, the initial assumptions are not needed, and the sampling distribution of the test statistic under  $H_0$  is computed by using permutations of the data.

# Permutation Test Example (1):

- In an RCT, participants are randomly assigned to a treatment ( $T_x$ ) or control (C) group.
- Under  $H_0$ , the outcome does not depend on whether a participant takes the treatment ( $T_x$ ) or the control (C).
- Under  $H_1$ , the outcome tends to be different, say larger for participants in the  $T_x$  group.
- A test statistic  $T$  measures the difference in observed outcomes for the two groups.  $T$  may be the difference in the two group means (or medians), denoted as  $t$  for the observed data.

# Permutation Test Example (1)

- Under  $H_0$ , the individual labels of Tx and C are unimportant, since they have no impact on the outcome.
- Since they are unimportant, the label can be randomly shuffled among the participants without changing the joint null distribution of the data.
- Shuffling the data creates a “new” dataset. It has the same participants, but with the group labels changed so as to appear as there were different group assignments.

# Permutation Test Example (1)

- Consider data from two groups: Group 1: 20,24,30 Group 2: 10,14,18
- How to test if the means of the two groups are equal?
- T-test may not be appropriate to use, given the small sample size.
- permutation is equivalent to wilcoxon test



# Permutation Test Example (1)

- If the group means are truly equal, then shifting the group labels will not have a big impact the sum of the two groups (or mean with equal sample sizes).
- Some group sums will be larger than in the original data set and some will be smaller.
- There are 6 observations.
- There are  $6!=720$  permutations considering all 6 positions as unique. Of the 720 permutations, there are 20 unique combinations (equal sample sizes). Compute mean (or sum) for each of these.
- If all permutations were enumerated. Thus, this was an exact permutation test.
- When dataset is too large to enumerate all permutations, a large number of random permutations are selected.

## Permutation Test Example(2):

- It says, “Okay, these are the n result (20 score in the sleep data) we got, and the way they are divided up between the two groups is one possible permutation of them. There are, in fact. . .

## [1] 184756

. . . possible permutations (technically combinations) of these data into 2 groups of 10, and each has equal probability.

- Most of these permutations would give no or little difference between the group means, but a few of them would give large differences.
- How extreme is the obtained case?

# Principles of permutation test

- In other words, the logic is similar to t-test. If the obtained case is in the most extreme 5% of possible results, then we reject the null hypothesis.
- Permutation test does not make any assumption at all about parent distributions. [no assumption of normality].
- Furthermore, a permutation test is generally more powerful than a “traditional” nonparametric test.
- The disadvantage of a permutation test is the number of permutations that must be generated.

# Steps of permutation test:

- ① Calculate T and p value observed sample
- ② Permutation part
  - ① Sample without replacement n (20) observations, assign the first half (10) observations to the treatment group and the remaining observations to the control group
  - ② Compute the T value
  - ③ Store the T value in a vector/matrix
  - ④ Repeat steps(1- 3) [ R times]
- ③ Calculate empirical p value
  - How often  $T(\text{permutation}) > \text{observed. T value?}$

## Example from (sleep) data:

- Sleep data show the effect of two soporific drugs (increase in hours of sleep compared to control) on 10 patients.

##	extra	group	ID
##	Min. :-1.600	1:10	1 :2
##	1st Qu.: -0.025	2:10	2 :2
##	Median : 0.950		3 :2
##	Mean : 1.540		4 :2
##	3rd Qu.: 3.400		5 :2
##	Max. : 5.500		6 :2
##			(Other):8

## Example from (sleep) data: -cont

##	extra	group	ID
## 1	0.7	1	1
## 2	-1.6	1	2
## 3	-0.2	1	3
## 4	-1.2	1	4
## 5	-0.1	1	5
## 6	3.4	1	6

## 2 groups: 10 observation each

On hyoscine drug

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	-1.600	-0.175	0.350	0.750	1.700	3.700

Control

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	-0.100	0.875	1.750	2.330	4.150	5.500

Observed value of T test:

```
## $T_value
##          t
## -1.860813
##
## $P_Value
```

# Permutation test for sleep data

```
R = 999
scores = sleep$extra
t.values <- numeric(R)
  for (i in 1:R) {
    index = sample(1:20, size=10, replace=F)
    group1 = scores[index]
    group2 = scores[-index]
    studTest<- t.test(group1,group2)
    t.values[i] = studTest$statistic
  }
```



## P-value for permutation test:

-we compare those simulated t-values to the one we obtained above from the actual t-test. tells us that a little more than 9% of the sims gave us a result equal or more extreme than the obtained case.

```
permut.tvalues = abs(t.values)
# for a two-tailed test
mean(permut.tvalues<=1.8608)
```

```
## [1] 0.8988989
```

- The p.value from permutation test is pretty close to the p-value from the actual t-test may be due to nonnormality [t-test was a bit too generous here].

# Monte Carlo:

- What is a Monte Carlo study?
- How to conduct Monte Carlo simulation?
- What are some situations where Monte Carlo simulation is needed?

# What is a Monte Carlo study?

According to Webster's dictionary, Monte Carlo relates to or involves "the use of random sampling techniques and often the use of computer simulation to obtain approximate solutions to mathematical or physical problems especially in terms of a range of values each of which has a calculated probability of being the solution"

# What is a Monte Carlo study?

- When the “resampling” is done from a known theoretical distribution, the correct term is “Monte Carlo” simulation.
- Monte Carlo simulations are estimations, not analytic answers i.e. It uses random number generation, rather than analytic calculations
- Accuracy of estimate increases with larger  $n$
- It is a tool for combining distributions, and thereby propagating more than just summary statistics

# How to conduct Monte Carlo study?

- **Steps of Monte Carlo simulations:**

- 1 Sample randomly from the simple distributions in each step
- 2 Estimate the complex function for the sample
- 3 repeat this a large number of times to get the estimate

# Random number generator

Distribution	R DISTNAME	SAS DISTNAME
Beta	beta	BETA
Beta-binomial	betabin*	
binomial	binom	BINOMIAL
Cauchy	cauchy	CAUCHY
chi-square	chisq	CHISQUARE
exponential	exp	EXPONENTIAL
F	f	F
gamma	gamma	GAMMA
geometric	geom	GEOMETRIC
hypergeometric	hyper	HYPERGEOMETRIC
inverse Normal	inv.gaussian*	IGAUSS+
Laplace	laplace*	LAPLACE
logistic	logis	LOGISTIC
lognormal	lnorm	LOGNORMAL
negative binomial	nbinom	NEGBINOMIAL
normal	norm	NORMAL
Poisson	pois	POISSON
Student's t	t	T
Uniform	unif	UNIFORM
Weibull	weibull	WEIBULL

# Random number generator -R Function-

Distribution	Functions			
<a href="#">Beta</a>	pbeta	qbeta	dbeta	rbeta
<a href="#">Binomial</a>	pbinom	qbinom	dbinom	rbinom
<a href="#">Cauchy</a>	pcauchy	qcauchy	dcauchy	rcauchy
<a href="#">Chi-Square</a>	pchisq	qchisq	dchisq	rchisq
<a href="#">Exponential</a>	pexp	qexp	dexp	rexp
<a href="#">F</a>	pf	qf	df	rf
<a href="#">Gamma</a>	pgamma	qgamma	dgamma	rgamma
<a href="#">Geometric</a>	pgeom	qgeom	dgeom	rgeom
<a href="#">Hypergeometric</a>	phyper	qhyper	dhyper	rhyper
<a href="#">Logistic</a>	plogis	qlogis	dlogis	rlogis
<a href="#">Log Normal</a>	plnorm	qlnorm	dlnorm	rlnorm
<a href="#">Negative Binomial</a>	pnbinom	qnbinom	dnbinom	rnbinom
<a href="#">Normal</a>	pnorm	qnorm	dnorm	rnorm
<a href="#">Poisson</a>	ppois	qpois	dpois	rpois
<a href="#">Student t</a>	pt	qt	dt	rt
<a href="#">Studentized Range</a>	ptukey	qtukey	dtukey	rtukey
<a href="#">Uniform</a>	punif	qunif	dunif	runif
<a href="#">Weibull</a>	pweibull	qweibull	dweibull	rweibull
<a href="#">Wilcoxon Rank Sum Statistic</a>	pwilcox	qwilcox	dwilcox	rwilcox
<a href="#">Wilcoxon Signed Rank Statistic</a>	psignrank	qsignrank	dsignrank	rsignrank

## Example using Student's sleep data...

- Mean

```
##      1      2
## 0.75 2.33
```

- Standard deviation

```
##          1          2
## 1.789010 2.002249
```

- n for each group

```
##  1  2
## 10 10
```



# Power Calculation

```
power.tt<- power.t.test(n=10, delta=(2.33-.75),  
                        sd=1.9, sig.level=.05,  
                        type="two.sample", alternative="two.sided")
```

```
## $power  
## [1] 0.42
```

- we have a 42% chance of finding a two-tailed significant difference between 2 samples of 10 chosen from normal populations with a common SD of 1.9 (the pooled sd of the 2 samples) but with a true difference in means of 1.58.

## To do the same calculation by simulation,:

- Use the `rnorm( )` function to draw the samples,
- the `t.test( )` function to get a p-value,
- then we will simply look to see what % of p-values are less than  $\alpha=.05$  after running this procedure a large number of times.

# R script

```
R = 999
alpha = numeric(R)
  for (i in 1:R) {
group1 = rnorm(10, mean=.75, sd=1.9)
group2 = rnorm(10, mean=2.33, sd=1.9)
alpha[i] = t.test(group1,group2)$p.value
}
```

```
mean(alpha<.05)
```

```
## [1] 0.4174174
```

## Other Uses of Monte Carlo simulation: