Resampling Techniques and their Application

-Class 10-

Frank Konietschke

Institut für Biometrie und Klinische Epidemiologie

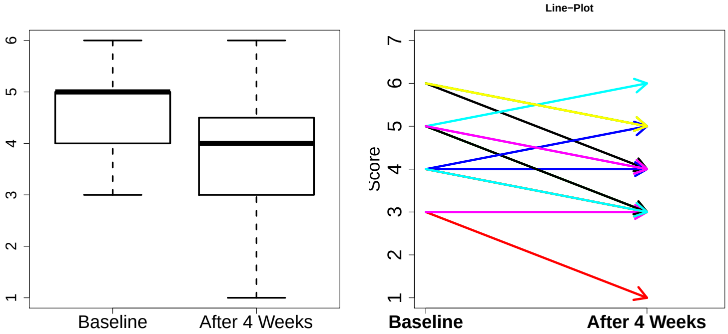
Charité - Universitätsmedizin Berlin, Berlin

frank.konietschke@charite.de

CHARITÉ
UNIVERSITÄTSMEDIZIN BERLIN

Example

- Panic disorder longitudinal trial
  - Specific physical exercise therapy
  - $n = 15$ patients
  - Response: CGI-score **before and after** 4 weeks of treatment
  - $0 =$ patient is healthy, . . ., $6 =$ patient is critically ill



Example: Discussion

- What is the research question?
- What is the data scale?
- Are means appropriate?

Statistical Model

- $\mathbf{X}_k = (X_k, Y_k)'$, $k = 1, \ldots, n$
- $X_k \sim F_1$, $Y_k \sim F_2$
- F_i is any distribution
    - How to define an effect? ("Difference between the samples")
    - Means are not defined
    - Other options?

Nonparametric Effects

- $\mathbf{X}_k = (X_k, Y_k)', k = 1, \ldots, n$
- $X_k \sim F_1$, $Y_k \sim F_2$
- Effect: Difference of the Medians
  - $\theta =$ Median(X) - Median(Y)
  - Crude Measure
  - What else could we measure?

```
x=c(6,3,5,4,5,3,4,5,5,4,
6,4,4,5,6)

y=c(4,1,3,4,6,3,3,4,3,3,
5,5,3,4,5)

xM<-median(x)
xY<-median(y)

theta<- xM-xY
```

Nonparametric Effects

- $\mathbf{X}_k = (X_k, Y_k)', k = 1, \ldots, n$
- $X_k \sim F_1,\ Y_k \sim F_2$
- Effect: %{X is smaller than Y}
  - $p =$ Proportion of $X < Y$
  - Issue: Equal observations
  - $\%(X < Y) + \%(X > Y) + \%(X = Y) = 100\%$
  - Reasonable:

$$\%(X < Y) + \frac{1}{2}\%(X = Y)$$

```
#X<Y +1/2(X=Y)

grid1=expand.grid(x,y)
prop1 = (grid1[,1]<grid1[,2]) +
1/2*(grid1[,1]==grid1[,2])
p=mean(prop1)
p

#Y<X+1/2(X=Y)
prop2 = (grid1[,2]<grid1[,1]) +
1/2*(grid1[,2]==grid1[,1])
q=mean(prop2)
q
p+q
```

Statistical Model

- $\mathbf{X}_k = (X_1, Y_k)', k = 1, \ldots, n$
- $X_k \sim F_1, Y_k \sim F_2$
- Effect and hypothesis
  - Relative marginal effect $p = P(X_1 < Y_2) + 0.5P(X_1 = Y_2)$
  - Interpretation
    - If $p < \frac{1}{2}$: X tends to be larger than Y
    - If $p = 1/2$: No tendency to smaller or larger values
  - Hypothesis:

$$H_0 : p = 1/2$$

Relative Effects: Other Examples

- Two samples: What is the chance, data in the first are smaller than in the second?
- What is the probability that men are taller than women?
- What is the probability, blood pressure under treatment is higher than under control?
- What is the probability, that scores after 4 weeks of treatment are smaller than at the beginning?
- **The relative effect thus measures, whether scores are smaller/larger between the groups**

Rel. Effect: Point Estimator

- **Ranks of the data**
  - Sort the data
  - Smallest number gets rank 1, largest gets rank $2n$
  - Mid-ranks are used to adjust for ties
  - In R: rank(x)

```
x <-c(5,3,2,1,4)
rank(x)

Two samples:
x=c(6,3,5,4,5,3,4,5,5,4,
6,4,4,5,6)

y=c(4,1,3,4,6,3,3,4,3,3,
5,5,3,4,5)

xy=c(x,y)
rank(xy)
```

Point Estimator

- Combine the two samples in a joint sample
- $X_1, \dots, X_n, Y_1, \dots, Y_n$
- Assign Ranks: $R_{x1}, \dots, R_{xn}, R_{y1}, \dots, R_{yn}$
- Compute

$$\overline{R}_{x\cdot} = \frac{1}{n}\sum_{k=1}^{n} R_{xk} \quad \text{"Mean of the first"}$$

$$\overline{R}_{y\cdot} = \frac{1}{n}\sum_{k=1}^{n} R_{yk} \quad \text{"Mean of the second"}$$

- Estimator:

$$\widehat{p} = \frac{1}{2n}(\overline{R}_{y\cdot} - \overline{R}_{x\cdot}) + 1/2$$

```
n<-length(x)

xy<-c(x,y)
rxy <- rank(xy)

mRx <-mean(rxy[1:n])
mRy <-mean(rxy[(n+1):(2*n)])

phat<-1/(2*n)*(mRy-mRx)+1/2
```

Variance Estimator

- Need to estimate the variance of $\widehat{p}$
- Needed: Ranks and internal ranks per sample
- Ranks internal (only x): $R_{x1}^{(x)}, \ldots, R_{xn}^{(x)}$
- Ranks internal (only y): $R_{y1}^{(y)}, \ldots, R_{yn}^{(y)}$
- Compute

$$
\begin{aligned}
Z_{xk} &= \frac{1}{n} * (R_{xk} - R_{xk}^{(x)}) \\
Z_{yk} &= \frac{1}{n} * (R_{yk} - R_{yk}^{(y)}) \\
D_k &= Z_{xk} - Z_{yk}
\end{aligned}
$$

- Estimator:

$$\widehat{\sigma}^2 = \frac{1}{n-1} \sum_{k=1}^{n} (D_k - \overline{D}_{\cdot})^2$$

```
xy<-c(x,y)
rxy <- rank(xy)
rx <-rank(x)
ry <-rank(y)

Z1k <- 1/n*(rxy[1:n]-rx)
Z2k <- 1/n*(rxy[(n+1):(2*n)]-ry)
Dk <- Z1k-Z2k

sigmahat <-var(Dk)
```

Test Procedures: Munzel's Test

- Test Statistic

$$T = \sqrt{n}\frac{(\widehat{p} - 1/2)}{\widehat{\sigma}}$$

- p-values and critical values are computed from a t-distribution with $n-1$ degrees of freedom
- $(1-\alpha)100\%$ Confidence Intervals:

$$\widehat{p} \pm t_{1-\alpha/2}(n-1)\widehat{\sigma}/\sqrt{n}$$

- Is the method applicable when samples are small?

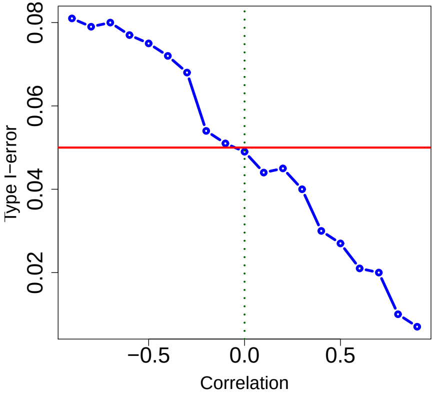
```
T=sqrt(n)*(phat-1/2)/sqrt(sigmahat)
pvalue=2*min(pt(T,n-1),1-pt(T,n-1))

crit <-qt(0.975,n-1)
SE <- sqrt(sigmahat)/sqrt(n)

Lower = phat -crit*SE
Upper = phat +crit*SE
```

Simulation: Munzel's test (1999)

- Type-1 error simulation ($\alpha = 5\%$, $n_{sim} = 10,000$)
- $\mathbf{X}_k = (X_{1k}, X_{2k}) \sim N(\mathbf{0}, \mathbf{V})$, $k = 1, \ldots, 10$
- $\mathbf{V}$: compound symmetric (variance = 1, covariance = ρ)



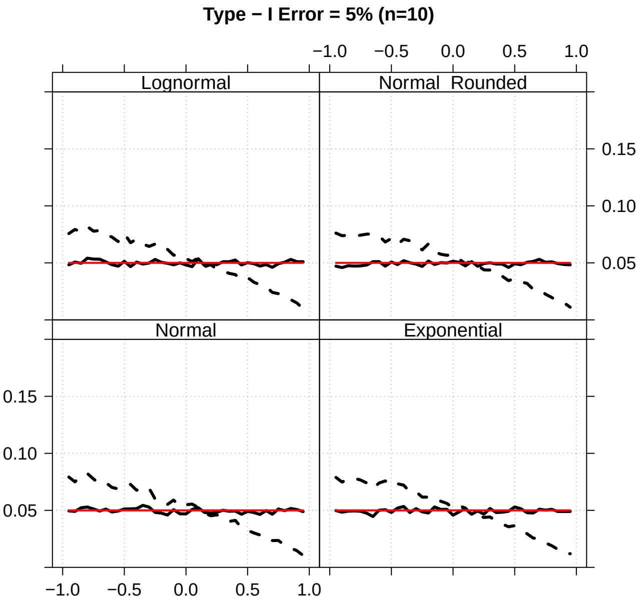
Improve Munzel's Test

- How?
  - Explore Resampling Methods
  - For example, Wild-Bootstrap Approach
  - Permute data per subject
- Procedure
  - **Permute randomly X_{1k} and X_{2k} within each pair $\mathbf{X}_k$**
  - Assign new ranks for permuted data
  - Compute $T^* = \sqrt{n}(\widehat{p}^* - 1/2)/\widehat{\sigma}*$
  - Repeat this step several times
  - CI: $\widehat{p} \pm z^*_{1-\alpha/2}\widehat{\sigma}/\sqrt{n}$
- Project: Explore other resampling methods
- Note: Permuting data overall is not applicable in this model

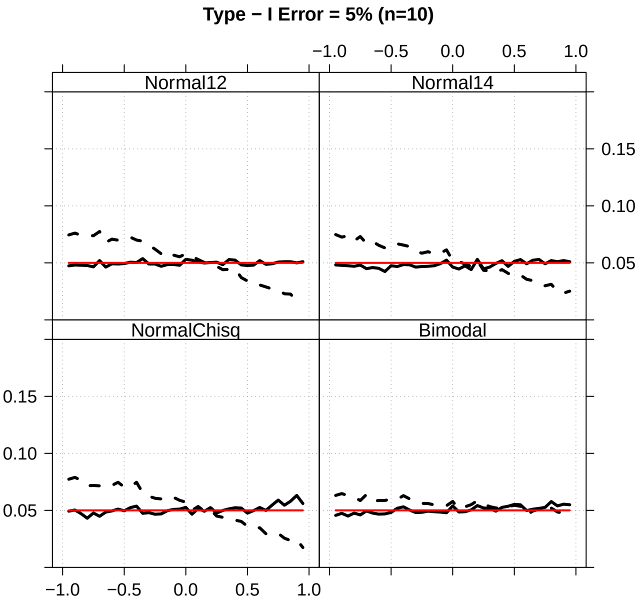
Studentized Permutation Test - Theory

- Theoretically shown
  - Procedure is asymptotically exact under non-exchangeable data
  - Exact under exchangeability
  - Confidence intervals are asymptotically exact
  - Both procedures have equal power (asymptotically)
  - Reference: Konietschke and Pauly (2012), A studentized permutation test for the Nonparametric Behrens-Fisher Problem in Paired Data

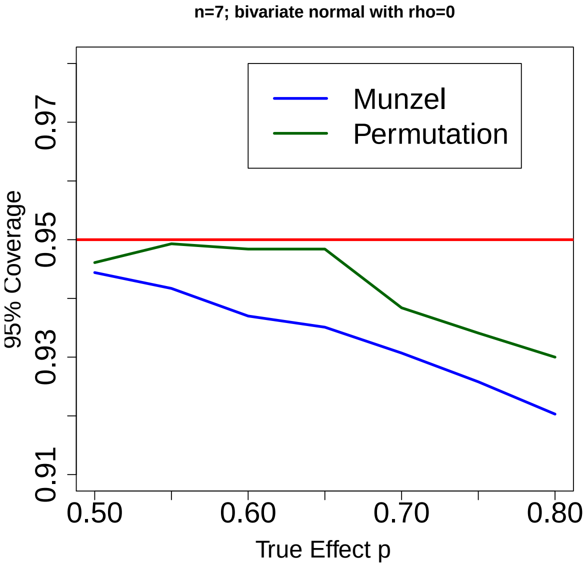
Simulations - Exchangeable data



Simulations - Non Exchangeable data



Simulations - 95%-Coverage Probability



Example: Evaluation

- Point estimate $\widehat{p} = .29$
  - Scores tend to smaller after 4 weeks than under baseline
- Hypothesis $H_0 : p = 1/2 \to p - value = .0006$
- 95%-CI: $[0.16; 0.43]$
- Munzel's test: similar results

Patient	PGI		Ranks	
k	Baseline	Week 4	Baseline	Week 4
1	6	4	28.5	14.0
2	3	1	5.5	1.0
3	5	3	22.5	5.5
4	4	4	14.0	14.0
5	5	6	22.5	28.5
6	3	3	5.5	5.5
7	4	3	14.0	5.5
8	5	4	22.5	14.0
9	5	3	22.5	5.5
10	4	3	14.0	5.5
11	6	5	28.5	22.5
12	4	5	14.0	22.5
13	4	3	14.0	5.5
14	5	4	22.5	14.0
15	6	5	28.5	22.5

Implementation

- Number of possible permutations: 2^n
- So, for small n, all possible permutations can be performed
- How?
- Need a matrix that contains all possible placement changes within each pair
- Notation

$$\mathbf{X}_k = (X_k, X_{k+n})', k = 1, \ldots, n$$

- $n = 3$

$$\mathbf{X}_k = \begin{pmatrix} X_1 & X_4 \\ X_2 & X_5 \\ X_3 & X_6 \end{pmatrix} \Rightarrow \mathbf{X}_k = (X_1, X_2, X_3, X_4, X_5, X_6)'$$

Implementation- II

- Now, generate a matrix...
  - that has 1 and $(n+1)$ how often in the first row and in which order?
  - that has 2 and $(n+2)$ how often in the second row and in which order?
  - that has 3 and $(n+3)$ how often in the second row and in which order?
  - ....

Implementation- III

n=4

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	1	1	1	1	1	1	5	5	5	5	5	5	5	5
2	2	2	2	6	6	6	6	2	2	2	2	6	6	6	6
3	3	7	7	3	3	7	7	3	3	7	7	3	3	7	7
4	8	4	8	4	8	4	8	4	8	4	8	4	8	4	8
5	5	5	5	5	5	5	5	1	1	1	1	1	1	1	1
6	6	6	6	2	2	2	2	6	6	6	6	2	2	2	2
7	7	3	3	7	7	3	3	7	7	3	3	7	7	3	3
8	4	8	4	8	4	8	4	8	4	8	4	8	4	8	4

```
permuall<-function(nsim,rho,n){
#-------Daten einlesen------#
n1<-n+1
n2<-2*n
x<-matrix(0,ncol=nsim,nrow=n2)
for (h in 1:nsim){
x11<-rnorm(n)
x22<-rho*x11+sqrt(1-rho^2)*(rnorm(n))
x[,h]<-c(x11,x22)}

#-------------Brunner-Munzel test-----#
tcrit<-qt(0.975,n-1)
x1<-x[1:n,]
x2<-x[n1:n2,]
rx<-apply(x,2,rank)
rx1<-rx[1:n,]
rx2<-rx[n1:n2,]
rix1<-apply(x1,2,rank)
rix2<-apply(x2,2,rank)
BM1<-1/n*(rx1-rix1)
BM2<-1/n*(rx2-rix2)
BM3<-BM1-BM2
BM4<-1/(2*n)*(rx1 - rx2)
pd<-colMeans(BM2)
m<-colMeans(BM3)
v<-(colSums(BM3^2)-n*m^2)/(n-1)
v0<-(v==0)
v[v0]<-1/n
T<-sqrt(n)*(pd-1/2)/sqrt(v)
```

```
#-------Studentized Permutation Test---#
nperm<-2^n
if(nperm <10000){
p<-0
for (i in 1:n){
a<-rep(c(rep(c(i,i+n),nperm/(2^i)),rep(c(i+n,i),nperm/(2^i))),2^(i-1))
p<-rbind(p,a)
}
p<-p[2:(n+1),]
P<-matrix(p,ncol=nperm)}
```

```
if (nperm >=10000){
nperm=10000
P<-matrix(0,nrow=(2*n),ncol=nperm)
for (h in 1:nperm){
P[,h]<-c(t(apply(cbind(1:n,(n+1):(2*n)),1,sample)))}}
```

```
#------Beginn der Simulationsschleife----#
BM=PERM=c()
for (s in 1:nsim){
xs<-x[,s]
rs<-rx[,s]
#---------Permutationstest----------#
xperm<-matrix(xs[P],nrow=n2,ncol=nperm)
rxperm<-matrix(rs[P],nrow=n2,ncol=nperm)
xperm1<-xperm[1:n,]
xperm2<-xperm[n1:n2,]
rperm1<-rxperm[1:n,]
rperm2<-rxperm[n1:n2,]
riperm1<-apply(xperm1,2,rank)
riperm2<-apply(xperm2,2,rank)
BMperm2<-1/n*(rperm2-riperm2)
BMperm3<-1/n*(rperm1-riperm1)-BMperm2
pdperm<-colMeans(BMperm2)
mperm3<-colMeans(BMperm3)
vperm3<-(colSums(BMperm3^2)-n*mperm3^2)/(n-1)
vperm30<-(vperm3==0)
vperm3[vperm30]<-1/n
Tperm<-sqrt(n)*(pdperm-1/2)/sqrt(vperm3)
p1perm<-mean(Tperm<=T[s]); p2perm<-mean(Tperm>=T[s])
pperm<-2*min(p1perm,p2perm)
PERM[s]<-(pperm<0.05)}
ergebnis<-data.frame(nsim=nsim,
 nperm=nperm,n=n,rho=rho,BM= mean(abs(T)>tcrit),
 PERM=mean(PERM))
ergebnis}
permuall(100,0,7)
```