# Resampling Magic

BERDC Special Topics Talk 18

DaCCoTA
DAKOTA COMMUNITY COLLABORATIVE ON TRANSLATIONAL ACTIVITY

**Dr. Mark Williamson**

**Biostatistics, Epidemiology, and Research Design Core**

# Introduction

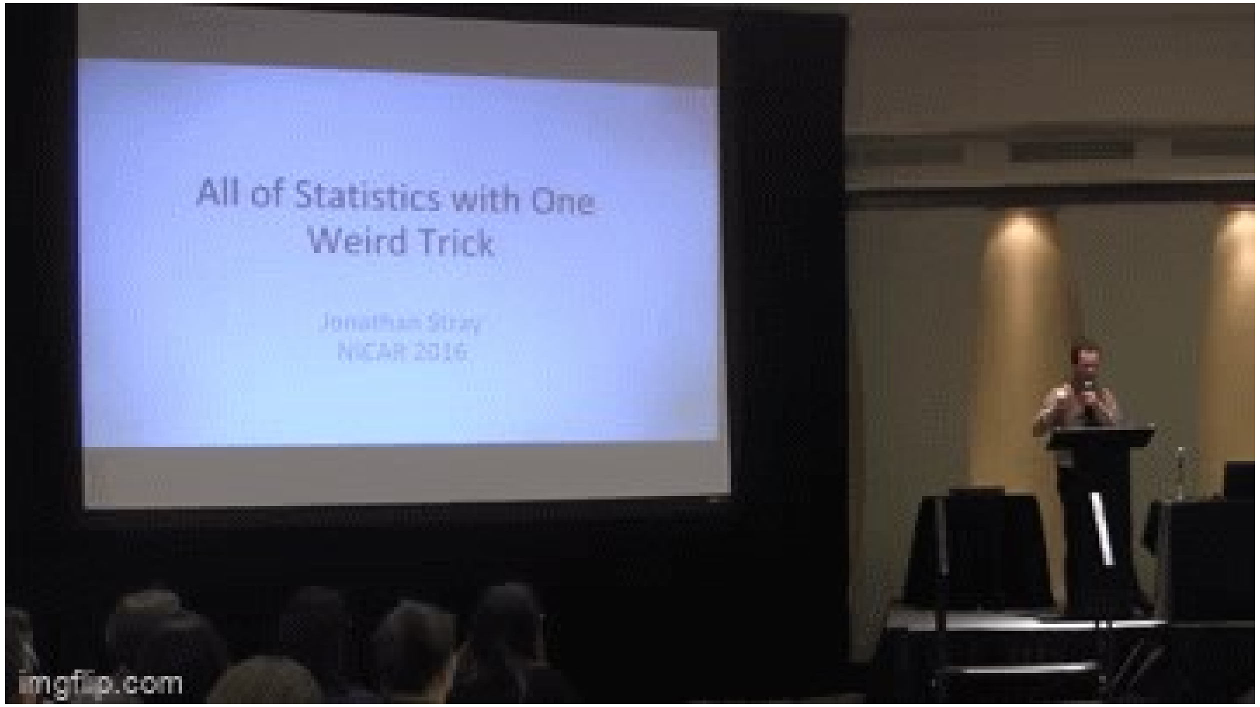
**Goal**: Spell out the benefits and methods of resampling

- What is resampling?
- What are the major types?
- What can resampling be used for?
- How does resampling work?
- Worked examples

# Interlude

**Solve Every Statistics Problem with One Weird Trick [1]**

- Resampling and Visualization
- Monte Carlo
- Confidence Intervals
- Significance Testing
- Error Detection
- And Beyond



DaCCoTA
DAKOTA COMMUNITY COLLABORATIVE ON TRANSLATIONAL ACTIVITY

# The What

UNIVERSITY OF NORTH DAKOTA

- **Resample Definition:**

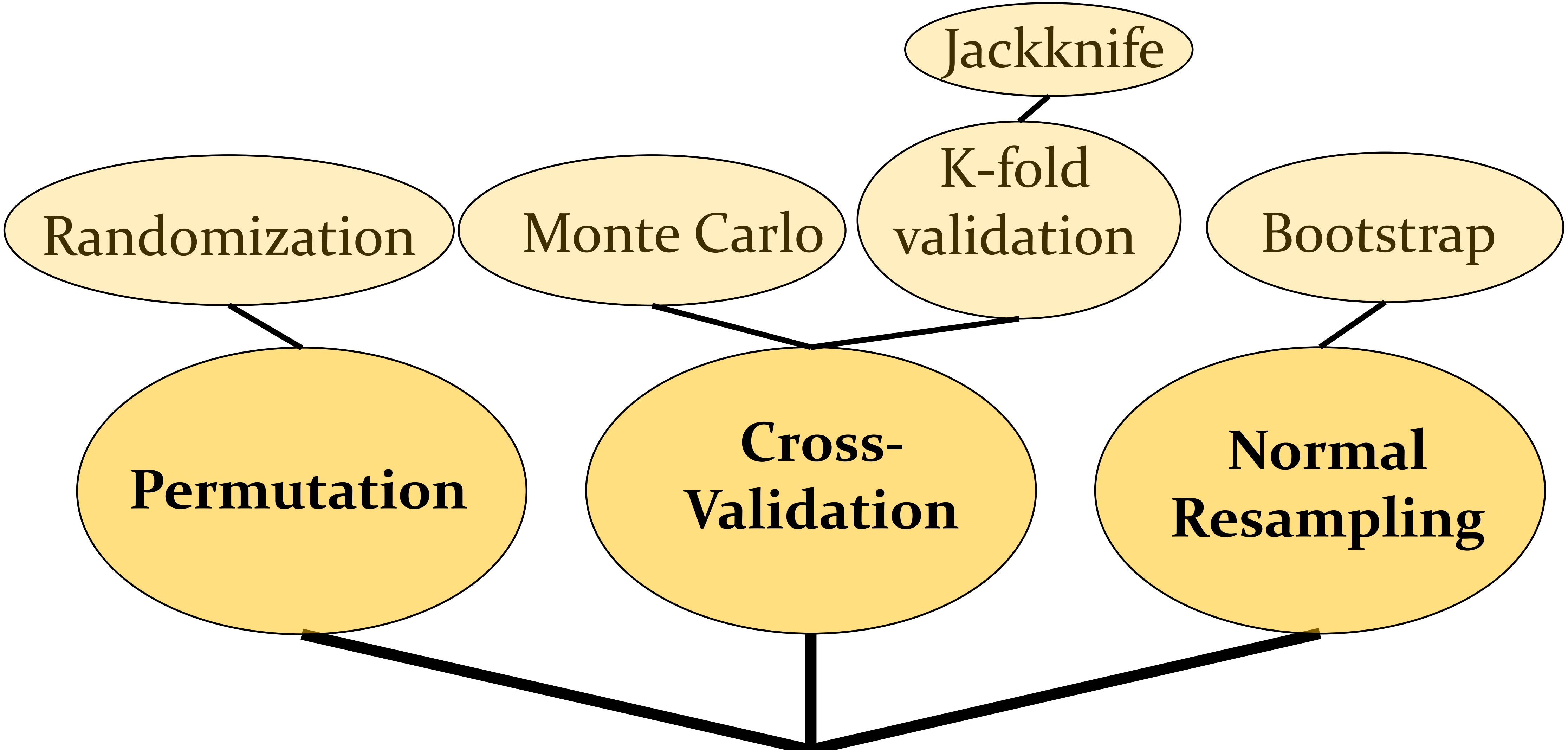
  *"Creation of new samples based on one observed sample."*

  -Wikipedia [2]

  *"Series of techniques used in statistics to gather more information about a sample. This can include retaking a sample or estimating its accuracy."*

  -Indeed [3]

# The Types [4-7]


Jackknife
Randomization
Monte Carlo
K-fold validation
Bootstrap
Permutation
Cross-Validation
Normal Resampling

DaCCoTA
DAKOTA COMMUNITY COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA

# The Uses

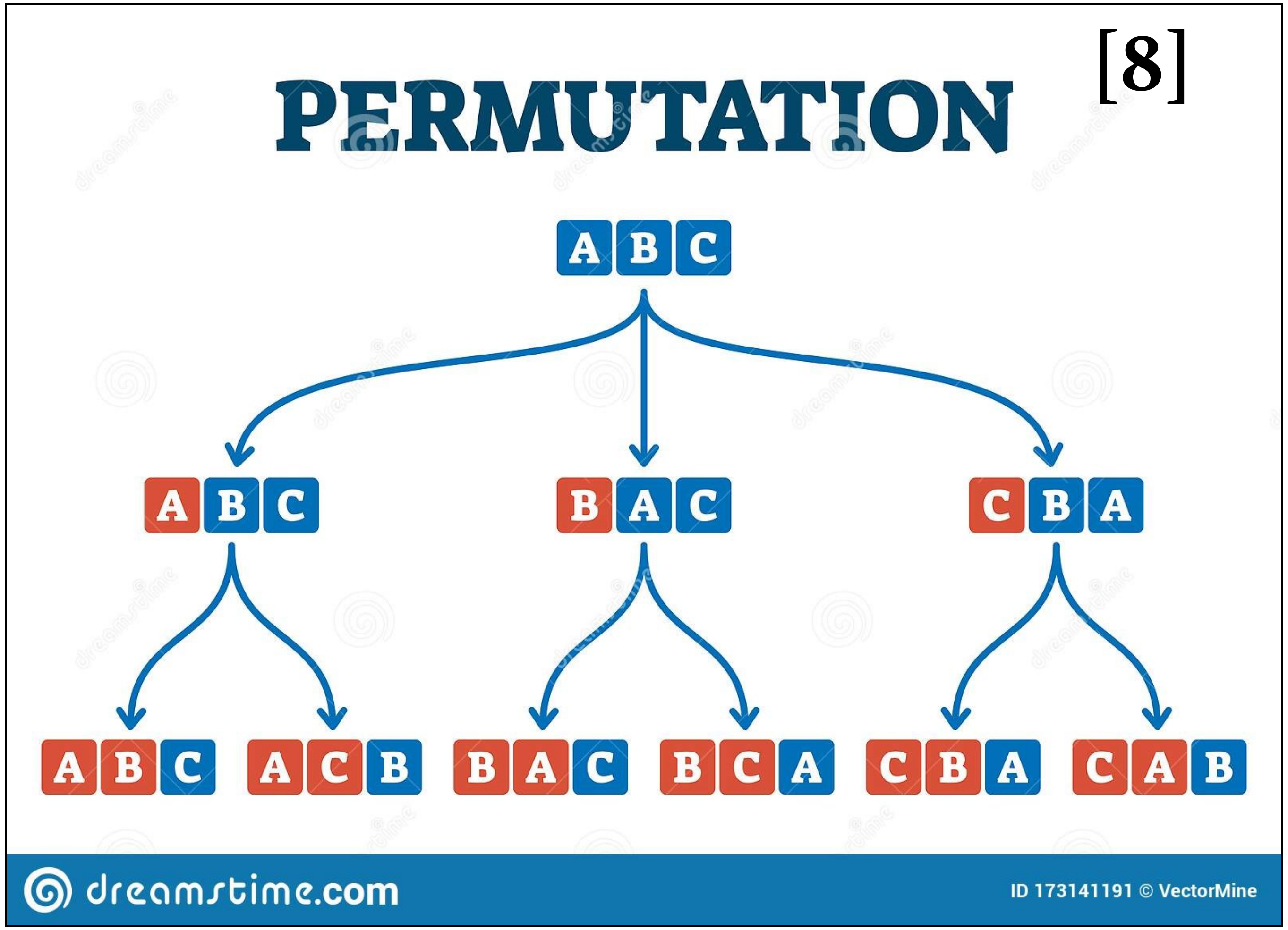
- Null Hypothesis significance testing (permutation)
- Accuracy, confidence intervals (bootstrap)
- Sampling distribution of test statistic (bootstrap)
- Validating a predictive model (jackknife, Monte Carlo, etc.)

DaCCoTA
DAKOTA COMMUNITY COLLABORATIVE ON TRANSLATIONAL ACTIVITY

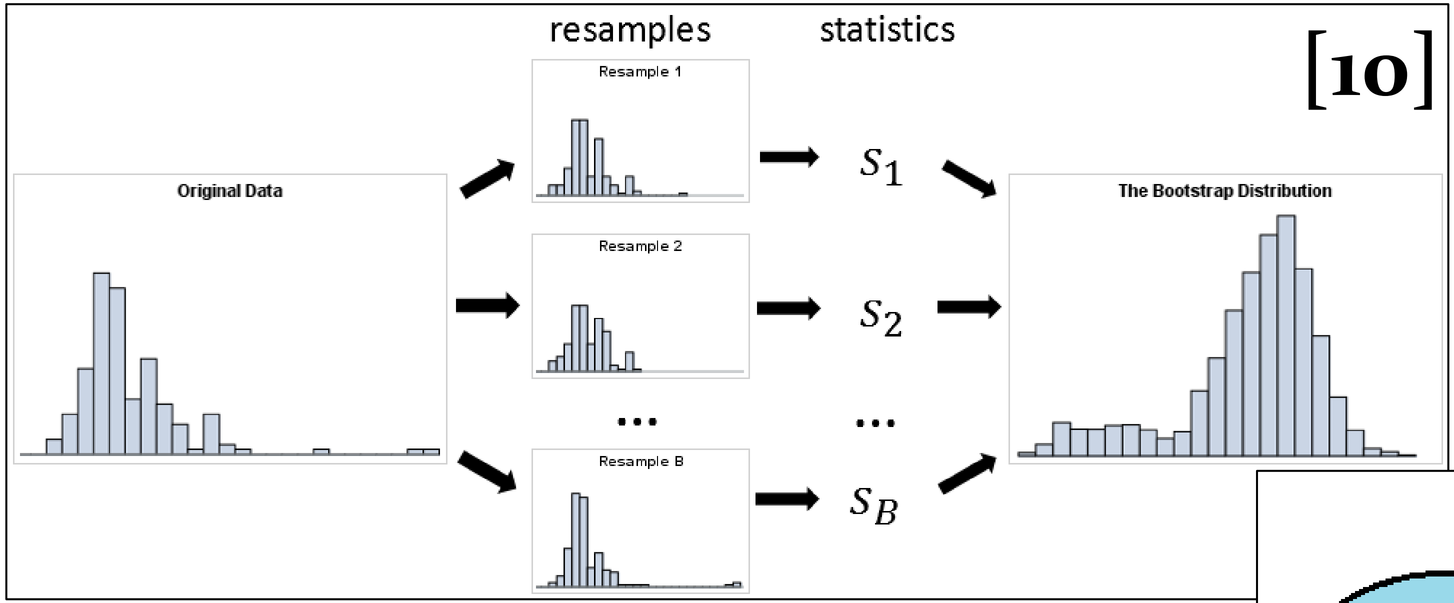
# The How

UND
UNIVERSITY OF NORTH DAKOTA

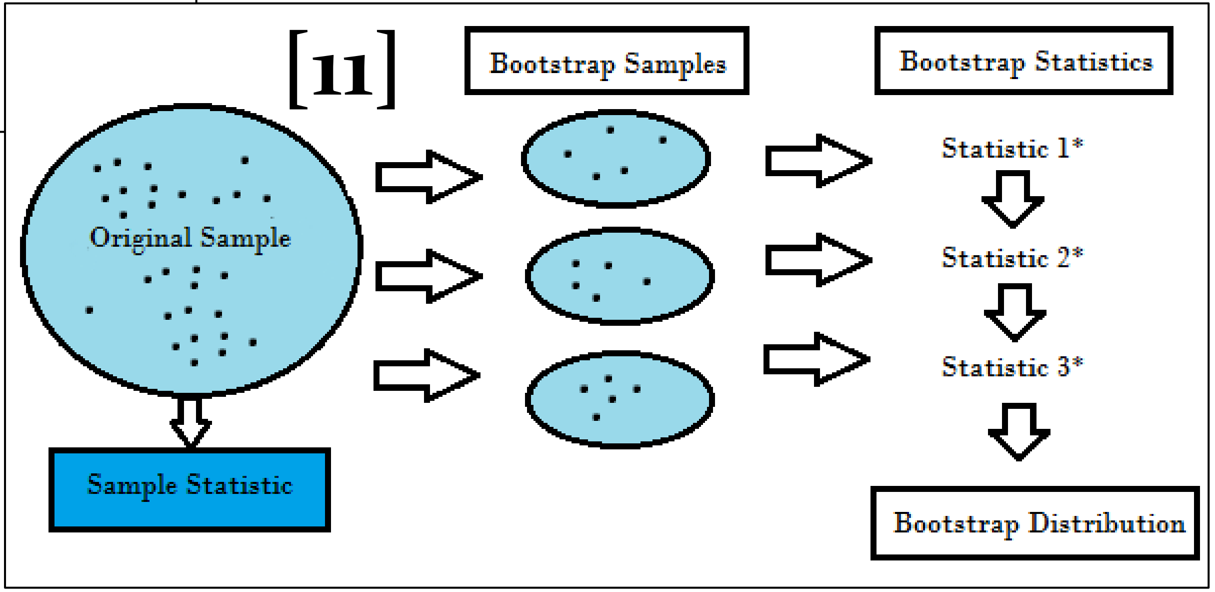
- Mostly done by computers
- R, SAS, Python, etc.
- Need dataset and objective



# The How [9]

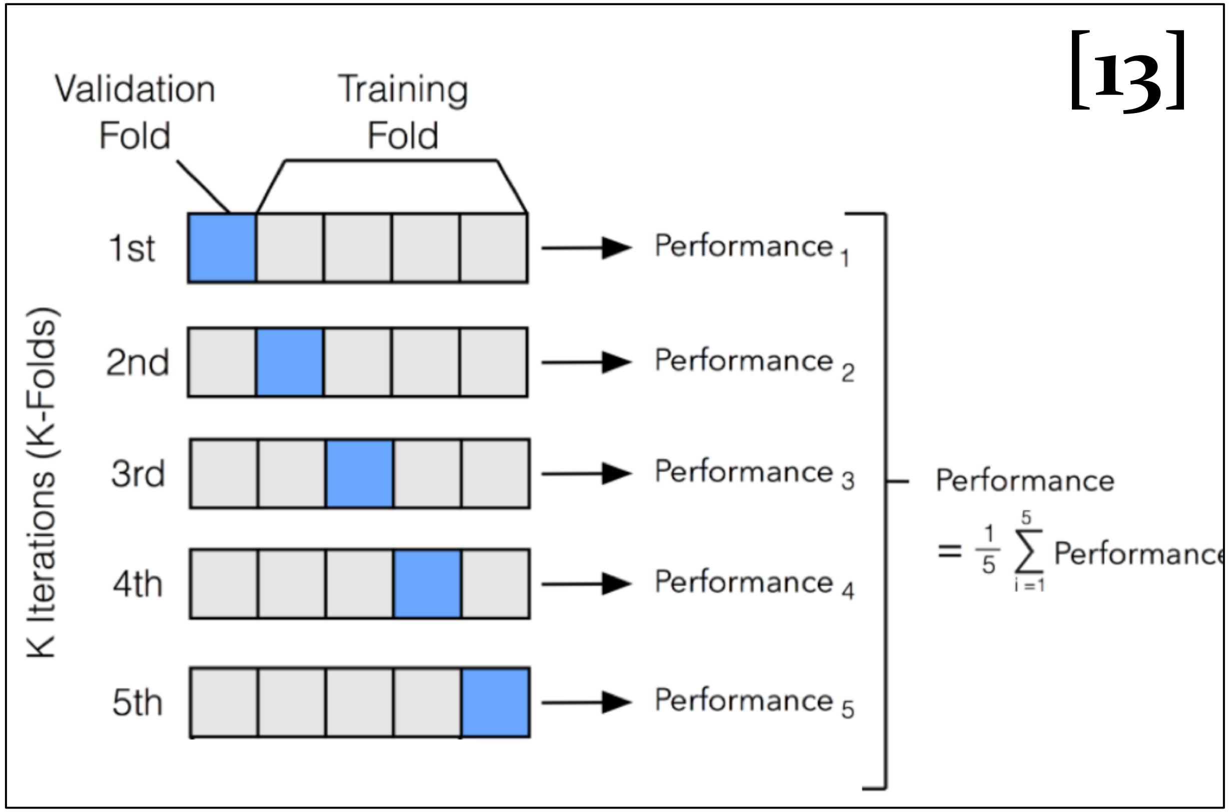


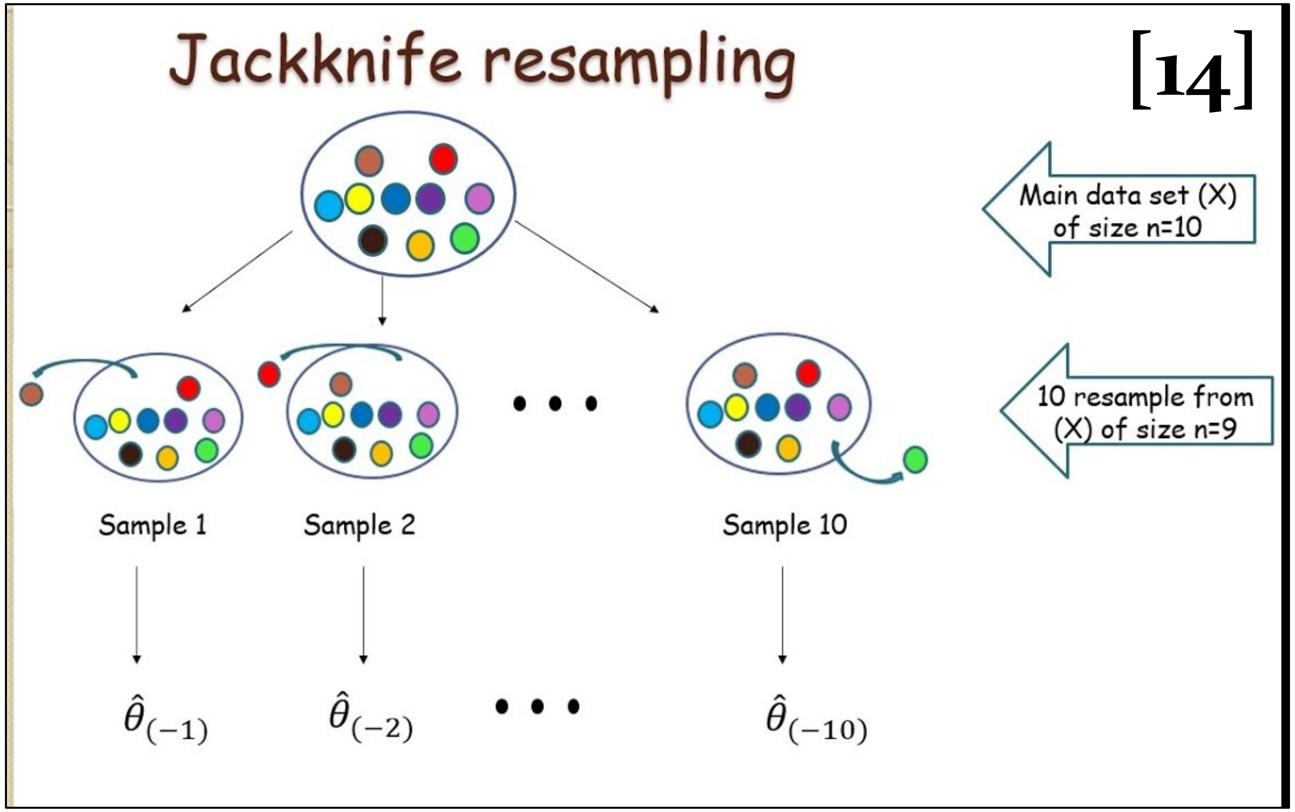
resamples
statistics
[10]
Resample 1
Original Data
Resample 2
The Bootstrap Distribution
Resample B
$s_1$
$s_2$
$s_B$

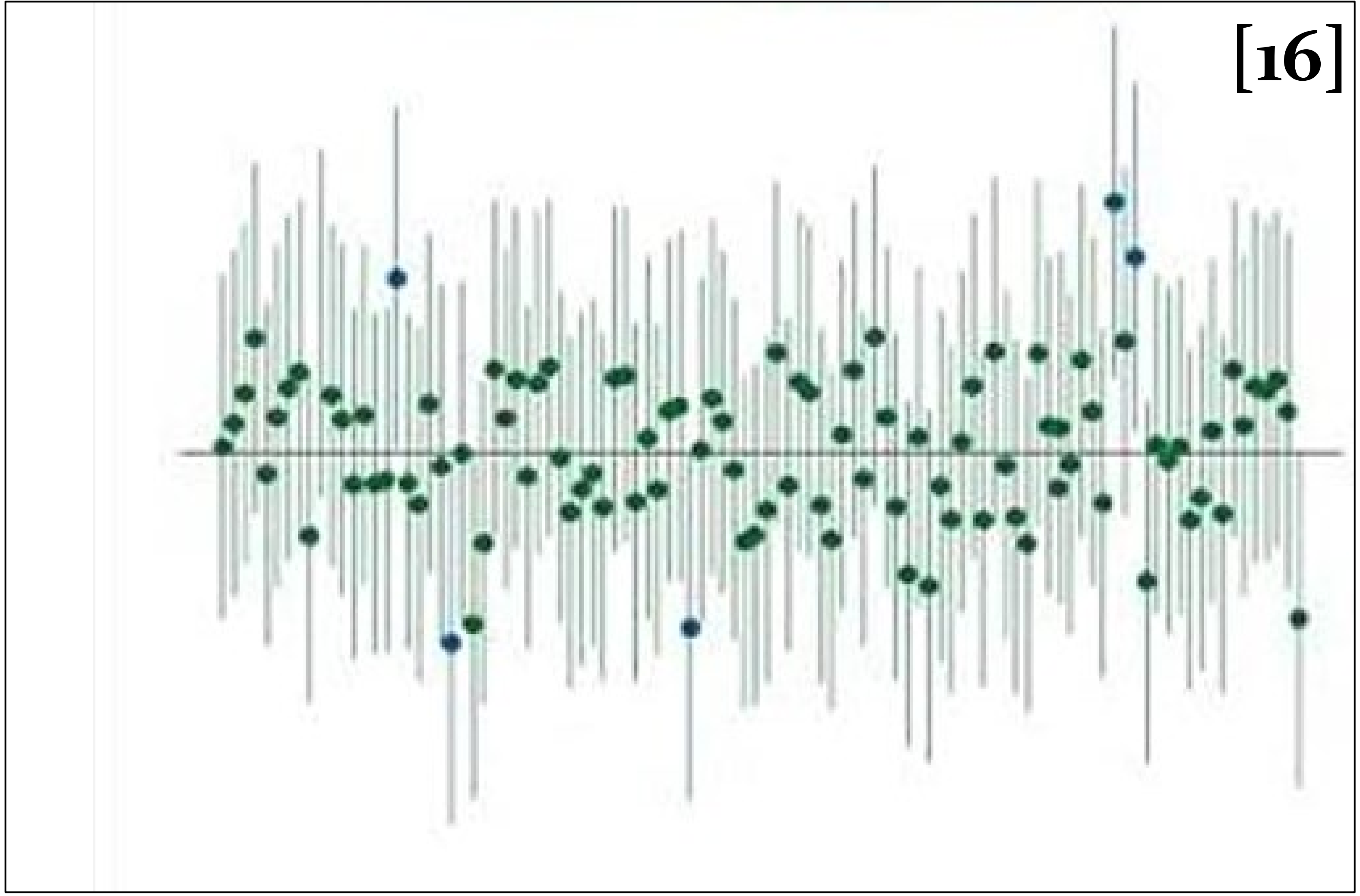
[11]
Bootstrap Samples
Bootstrap Statistics
Original Sample
Statistic 1*
Statistic 2*
Statistic 3*
Sample Statistic
Bootstrap Distribution

# The How [12]


[13]
Validation Fold
Training Fold
1st
2nd
3rd
4th
5th
K Iterations (K-Folds)
Performance 1
Performance 2
Performance 3
Performance 4
Performance 5
Performance
$= \frac{1}{5} \sum_{i=1}^{5}$ Performance


Jackknife resampling
[14]
Main data set (X) of size n=10
10 resample from (X) of size n=9
Sample 1
Sample 2
Sample 10
$\hat{\theta}_{(-1)}$
$\hat{\theta}_{(-2)}$
$\hat{\theta}_{(-10)}$

# The How [15]


[16]

# Examples

## Dataset

Magic the Gather Playing card data [17]

```
> head(magic)
```

| Name | Colors | Set Code | Mana Cost | Power | Toughness | Types |
|---|---|---|---|---|---|---|
| Ancestor's Chosen | W | 10 | 7 | 4 | 4 | Creature |
| Angel of Mercy | W | 10 | 5 | 3 | 3 | Creature |
| Angelic Blessing | W | 10 | 3 | NA | NA | Sorcery |
| Angelic Wall | W | 10 | 2 | 0 | 4 | Creature |
| Aven Cloudchaser | W | 10 | 4 | 2 | 2 | Creature |
| Benalish Knight | W | 10 | 3 | 2 | 2 | Creature |
| … | … | … | … | … | … | … |

```r
#Resampling Magic R-code:

#References:
  #   https://rcompanion.org/handbook/K_01.html
  #   https://www.statology.org/bootstrap-standard-error-in-r/
  #   https://stackoverflow.com/questions/58393608/bootstrapped-correlation-in-r
  #   https://www.statology.org/k-fold-cross-validation-in-r/
  #   #https://search.r-project.org/CRAN/refmans/pls/html/jack.test.html

#Get data and packages
library(lmPerm)
library(dplyr)
library(ggplot2)
library(boot)
library(caret)
library(pls)


  #magic cards:
setwd("C:/Users/Mark.Williamson.2/OneDrive - North Dakota University System/Desktop/Williamson Data/Magic")
magic<-read.csv("cards_condensed_10E_clean.csv")
head(magic)

boxplot(magic$convertedManaCost~magic$colors, col=c("grey", "green", "red", "blue", "white"))

boxplot(magic$convertedManaCost[magic$types=="Creature"]~magic$colors[magic$types=="Creature"],
        col=c("grey", "green", "red", "blue", "white"))

boxplot(magic$toughness[magic$types=="Creature"]~magic$colors[magic$types=="Creature"],
        col=c("grey", "green", "red", "blue", "white"))

boxplot(magic$power[magic$types=="Creature"]~magic$colors[magic$types=="Creature"],
        col=c("grey", "green", "red", "blue", "white"))

creatures <-magic[magic$types=="Creature",]
```

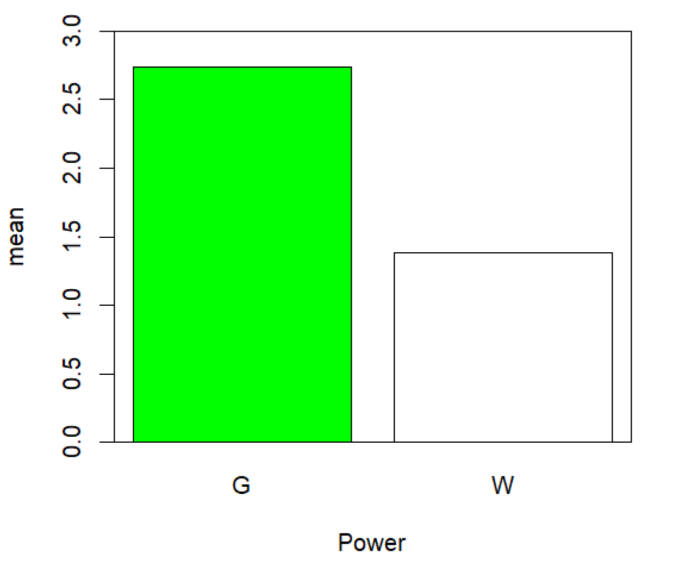
```
par(mfrow=c(2,3))
hist(creatures$convertedManaCost[creatures$colors=="R"], col="red", main="Mana Cost", xlab="Cost")
hist(creatures$convertedManaCost[creatures$colors=="U"], col="blue", main="", xlab="Cost")
hist(creatures$convertedManaCost[creatures$colors=="B"], col="grey", main="", xlab="Cost")
hist(creatures$convertedManaCost[creatures$colors=="W"], col="white", main="", xlab="Cost")
hist(creatures$convertedManaCost[creatures$colors=="G"], col="green", main="", xlab="Cost")
par(mfrow=c(1,1))

par(mfrow=c(2,3))
hist(creatures$power[creatures$colors=="R"], col="red", main="Power", xlab="Power")
hist(creatures$power[creatures$colors=="U"], col="blue", main="", xlab="Power")
hist(creatures$power[creatures$colors=="B"], col="grey", main="", xlab="Power")
hist(creatures$power[creatures$colors=="W"], col="white", main="", xlab="Power")
hist(creatures$power[creatures$colors=="G"], col="green", main="", xlab="Power")
par(mfrow=c(1,1))

par(mfrow=c(2,3))
hist(creatures$toughness[creatures$colors=="R"], col="red", main="Toughness", xlab="Toughness")
hist(creatures$toughness[creatures$colors=="U"], col="blue", main="", xlab="Toughness")
hist(creatures$toughness[creatures$colors=="B"], col="grey", main="", xlab="Toughness")
hist(creatures$toughness[creatures$colors=="W"], col="white", main="", xlab="Toughness")
hist(creatures$toughness[creatures$colors=="G"], col="green", main="", xlab="Toughness")
par(mfrow=c(1,1))

creatures %>%
  group_by(colors) %>%
  summarize(mean_Mana = mean(convertedManaCost),
            mean_power= mean(power),
            mean_toughness=mean(toughness))
```

# Examples

## Permutation Test

Is average power different between Green and White creatures?


```
#Example 1: Permutation test

  #Question: Is average power different between Green and White creatures?
creatures_GW <-subset(creatures, colors =="G" | colors == "W")
head(creatures_GW)

glm1 <- glm(power ~ colors, data=creatures_GW, family="poisson")
summary(glm1)

lm1 <-lm(power ~ colors, data=creatures_GW)
summary(lm1)
anova(lm1)

barplot(c(2.74,1.38)~c("G","W"), xlab="Power", ylab="mean", col=c("Green", "White"), ylim=c(0,3))
box()

lmp1 <-lmp(power ~ colors, data=creatures_GW, perm="Prob", seqs=FALSE)
summary(lmp1)
anova(lmp1)
```




```
> summary(glm1)

Call:
glm(formula = power ~ colors, family = "poisson", data = creatures_GW)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-2.34247  -0.71120   0.08007   0.15248   2.57099

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.00927    0.09667  10.440   <2e-16 ***
colorsW     -0.37328    0.15509  -2.407   0.0161 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 73.171  on 74  degrees of freedom
Residual deviance: 67.247  on 73  degrees of freedom
AIC: 262.53

Number of Fisher Scoring iterations: 5
```

```
> summary(lm1)

Call:
lm(formula = power ~ colors, data = creatures_GW)

Residuals:
    Min      1Q  Median      3Q     Max
-2.7436 -0.8889  0.1111  0.2564  5.2564

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.7436     0.2479  11.066   <2e-16 ***
colorsW      -0.8547     0.3578  -2.388   0.0195 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.548 on 73 degrees of freedom
Multiple R-squared:  0.07248,   Adjusted R-squared:  0.05978
F-statistic: 5.705 on 1 and 73 DF,  p-value: 0.01951

> anova(lm1)
Analysis of Variance Table

Response: power
          Df  Sum Sq Mean Sq F value  Pr(>F)
colors     1  13.675 13.6752  5.7048 0.01951 *
Residuals 73 174.991  2.3971
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> summary(lmp1)

Call:
lmp(formula = power ~ colors, data = creatures_GW, perm = "Prob",
    seqs = FALSE)

Residuals:
    Min      1Q  Median      3Q     Max
-2.7436 -0.8889  0.1111  0.2564  5.2564

Coefficients:
        Estimate Iter Pr(Prob)
colors1   0.4274 2732   0.0355 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.548 on 73 degrees of freedom
Multiple R-Squared: 0.07248,    Adjusted R-squared: 0.05978
F-statistic: 5.705 on 1 and 73 DF,  p-value: 0.01951

> anova(lmp1)
Analysis of Variance Table

Response: power
          Df R Sum Sq R Mean Sq Iter Pr(Prob)
colors1    1   13.675   13.6752 2732  0.03551 *
Residuals 73  174.991    2.3971
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

DaCCoTA
DAKOTA COMMUNITY COLLABORATIVE ON TRANSLATIONAL ACTIVITY

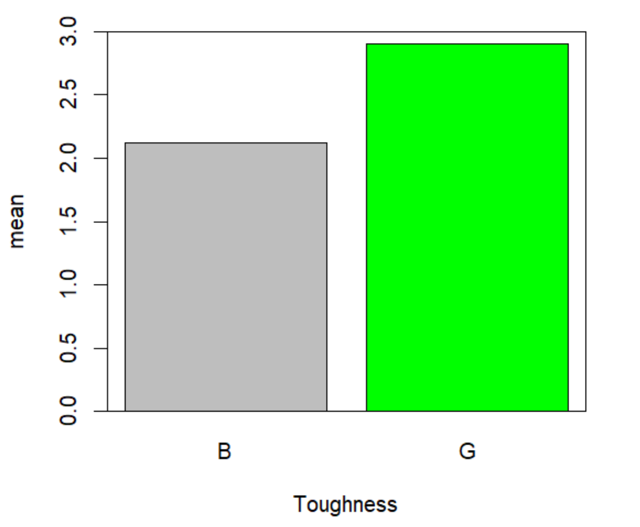
UND UNIVERSITY OF NORTH DAKOTA

# Examples

## Permutation Test

Is average toughness different between Green and Black creatures?


```
#Question: Is average toughness different between Black and Green creatures?
creatures_BG <-subset(creatures, colors =="B" | colors == "G")
head(creatures_BG)

glm2 <- glm(toughness ~ colors, data=creatures_BG, family="poisson")
summary(glm2)

lm2 <-lm(toughness ~ colors, data=creatures_BG)
summary(lm2)
anova(lm2)

barplot(c(2.12,2.90)~c("B","G"), xlab="Toughness", ylab="mean", col=c("Grey", "Green"), ylim=c(0,3))
box()

lmp2 <-lmp(toughness ~ colors, data=creatures_BG, perm="Prob", seqs=FALSE)
summary(lmp2)
anova(lmp2)
```




```
> summary(glm2)

Call:
glm(formula = toughness ~ colors, family = "poisson", data = creatures_BG)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.29121  -0.85933  -0.08404   0.56760   2.63768

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   0.7520     0.1195   6.292 3.14e-10 ***
colorsG       0.3118     0.1521   2.050   0.0403 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 67.728  on 71  degrees of freedom
Residual deviance: 63.438  on 70  degrees of freedom
AIC: 260.04

Number of Fisher Scoring iterations: 5
```

```
> summary(lm2)

Call:
lm(formula = toughness ~ colors, data = creatures_BG)

Residuals:
    Min      1Q  Median      3Q     Max
-1.8974 -1.1212 -0.1212  0.8788  5.1026

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.1212     0.2815   7.534 1.31e-10 ***
colorsG       0.7762     0.3825   2.029   0.0462 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.617 on 70 degrees of freedom
Multiple R-squared:  0.05555,	Adjusted R-squared:  0.04206
F-statistic: 4.117 on 1 and 70 DF,  p-value: 0.04625

> anova(lm2)
Analysis of Variance Table

Response: toughness
          Df Sum Sq Mean Sq F value  Pr(>F)
colors     1  10.77 10.7701  4.1174 0.04625 *
Residuals 70 183.10  2.6158
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

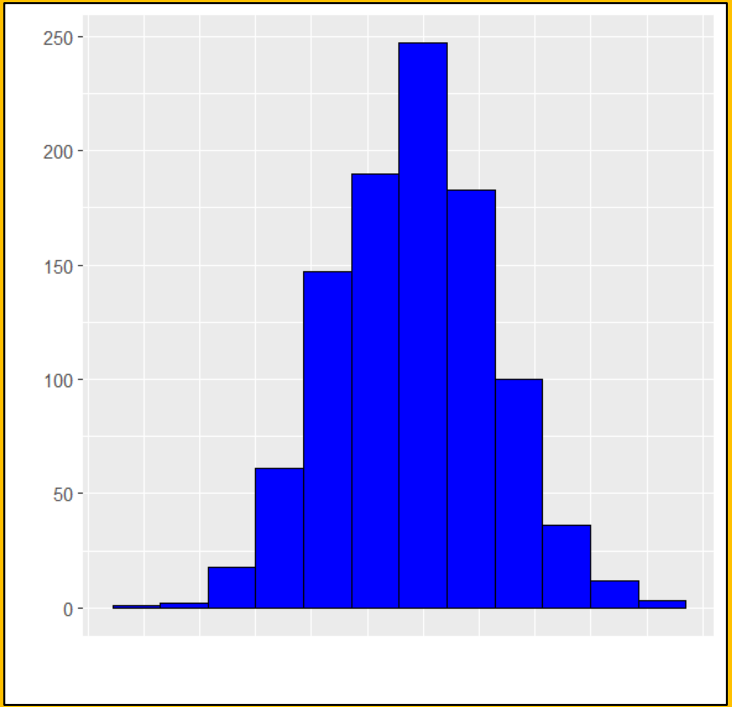
```
> summary(lmp2)

Call:
lmp(formula = toughness ~ colors, data = creatures_BG, perm = "Prob",
    seqs = FALSE)

Residuals:
    Min      1Q  Median      3Q     Max
-1.8974 -1.1212 -0.1212  0.8788  5.1026

Coefficients:
        Estimate Iter Pr(Prob)
colors1  -0.3881 3148   0.0308 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.617 on 70 degrees of freedom
Multiple R-Squared: 0.05555,	Adjusted R-squared: 0.04206
F-statistic: 4.117 on 1 and 70 DF,  p-value: 0.04625

> anova(lmp2)
Analysis of Variance Table

Response: toughness
          Df R Sum Sq R Mean Sq Iter Pr(Prob)
colors1    1    10.77    10.7701 3148  0.03081 *
Residuals 70   183.10     2.6158
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

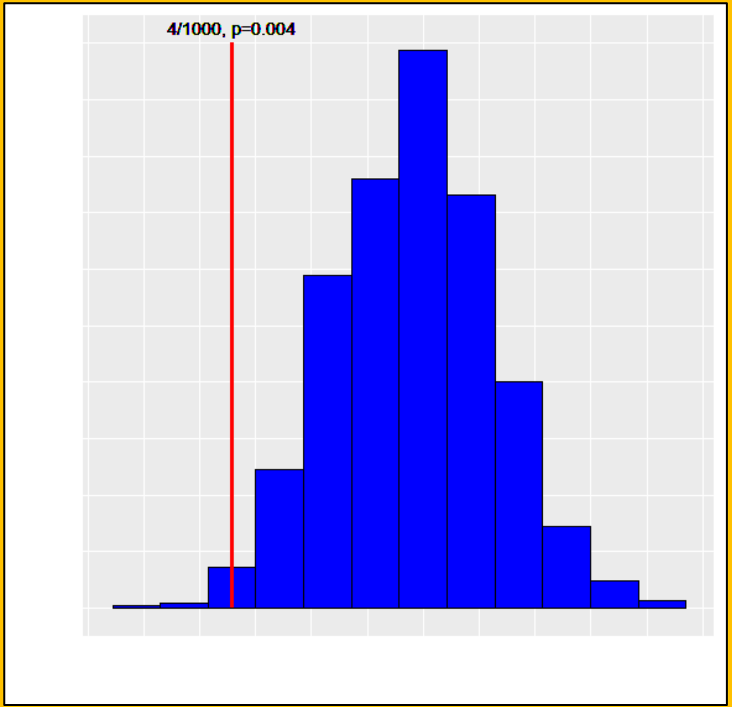
# Examples

## Permutation Test

Visualization

```
  #visualization:
t_coeff<-unname(lm1$coefficients[2]) #estimate of difference between Green and White creature power

set.seed(10)
coeff_list <-c()
for (i in seq(1:1000)) {
  perm_colors<-sample(creatures_GW$colors,replace=FALSE)
  perm_power <-creatures_GW$power
  perm_df <-data.frame(colors=perm_colors, power=perm_power)
  plm <-lm(power ~ colors, data=perm_df)
  perm_coeff <-unname(plm$coefficients[2])
  coeff_list <<-c(coeff_list, perm_coeff)
}
coeff_df <-data.frame(coeff=coeff_list)

plot1 <-ggplot(data=coeff_df, aes(x=coeff))+
  geom_histogram(, bins=12, fill="blue", col="black")
plot1

t_coeff #actual coefficient we got
1000-sum(coeff_list>t_coeff)

plot1+
  geom_segment(aes(x=t_coeff, y=0, xend=t_coeff, yend=250), colour="red", linewidth=1)+
  geom_text(aes(label=c("4/1000, p=0.004")), x=t_coeff, y=257, size=3)
#---------------------------------------------------------------------------------;
```



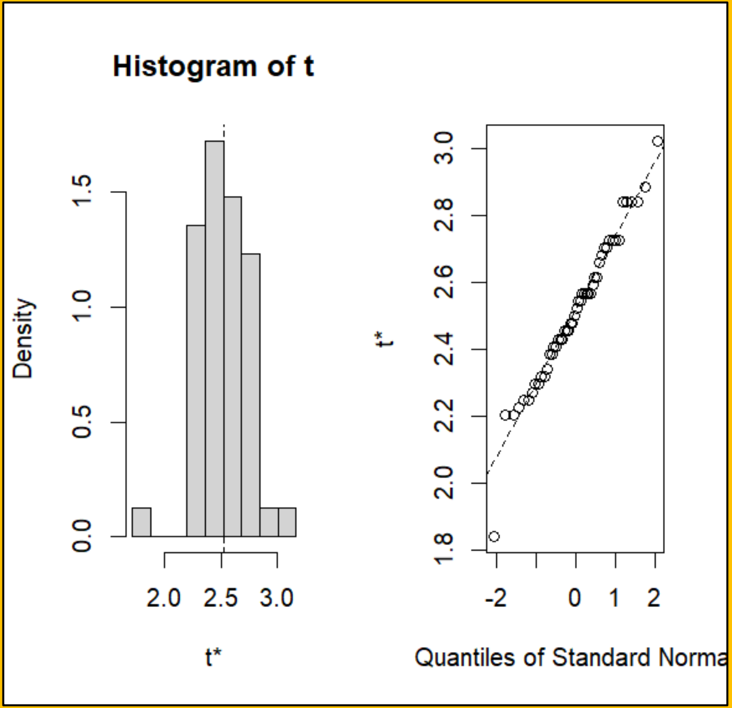
# Examples

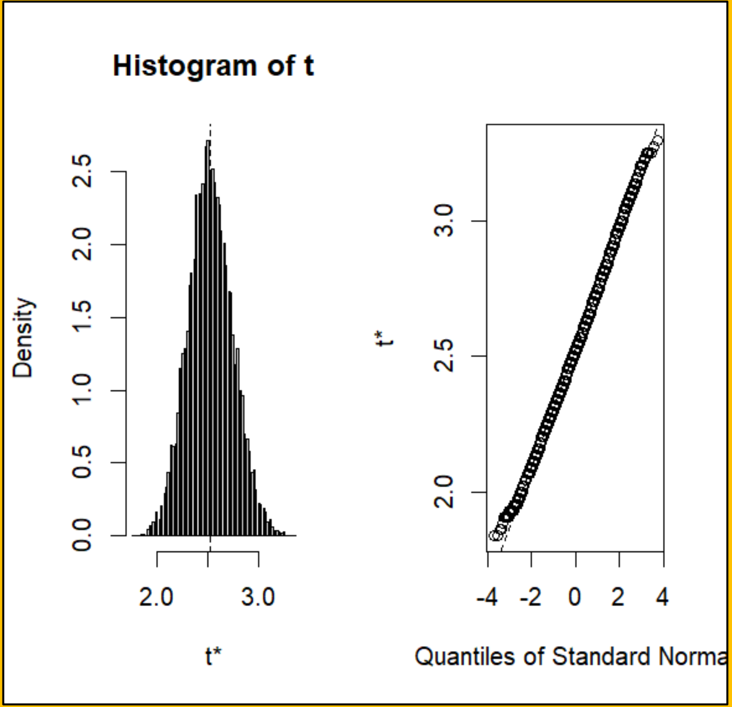
## Bootstrapping

What is the bootstrap confidence in mean cost for Instants?

```
#Example 2: Bootstrapping

  #Question: What is the bootstrap confidence in mean cost for Instants?
head(magic[magic$types=="Instant",])

Instant_Cost <-magic$convertedManaCost[magic$types=="Instant"]
mean(Instant_Cost)

    #Bootstrap standard error
set.seed(10)
meanFunc <- function(x,i){mean(x[i])}
reps1 <-boot(Instant_Cost, meanFunc, 50)
reps2 <-boot(Instant_Cost, meanFunc, 10000)

    #Plot Bootstrap distribution
plot(reps1)
plot(reps2)

    #Calculate confidence intervals
boot.ci(reps1, type="bca")
boot.ci(reps2, type="bca")
```

```
> boot.ci(reps1, type="bca")
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 50 bootstrap replicates

CALL :
boot.ci(boot.out = reps1, type = "bca")

Intervals :
Level       BCa
95%   ( 2.071,  3.021 )
Calculations and Intervals on Original Scale
Some BCa intervals may be unstable
> boot.ci(reps2, type="bca")
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = reps2, type = "bca")

Intervals :
Level       BCa
95%   ( 2.114,  2.977 )
Calculations and Intervals on Original Scale
```



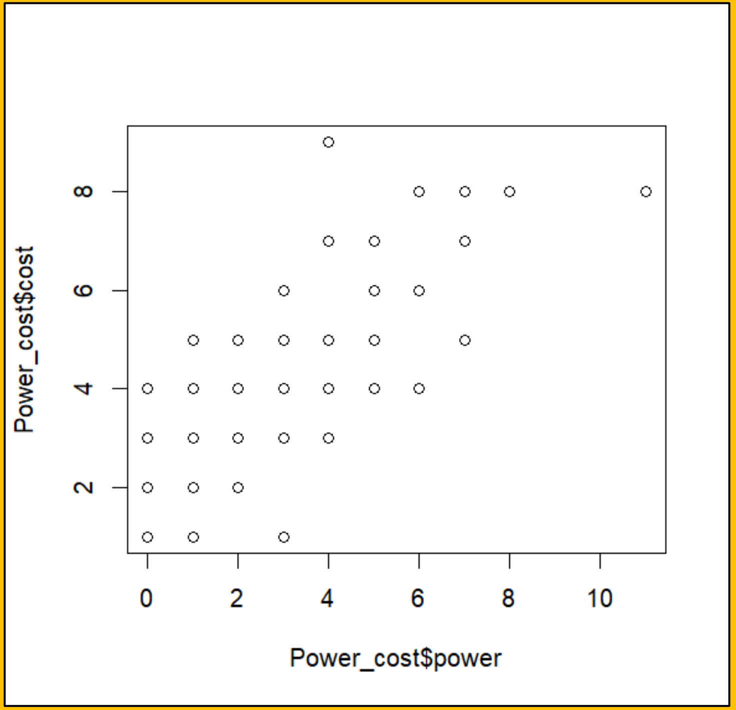



# Examples

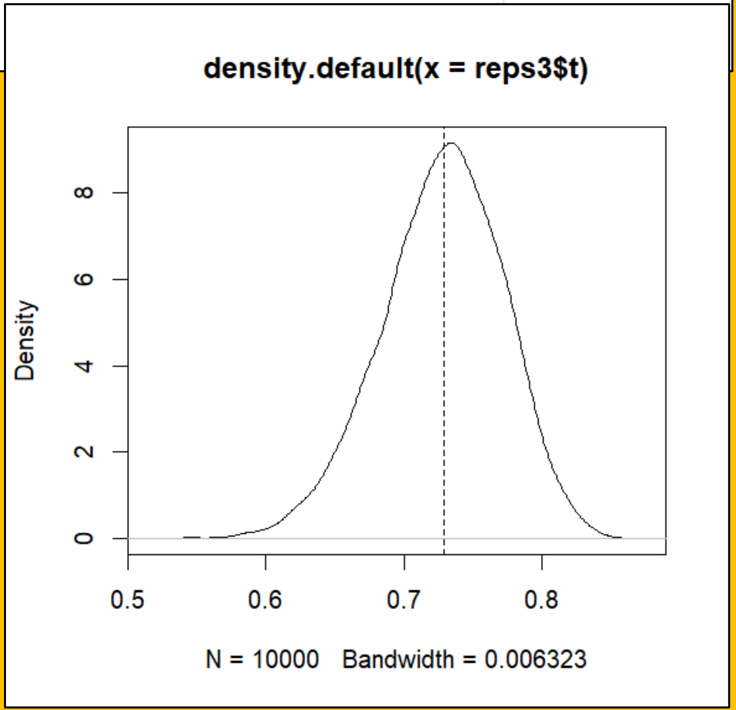
## Bootstrapping

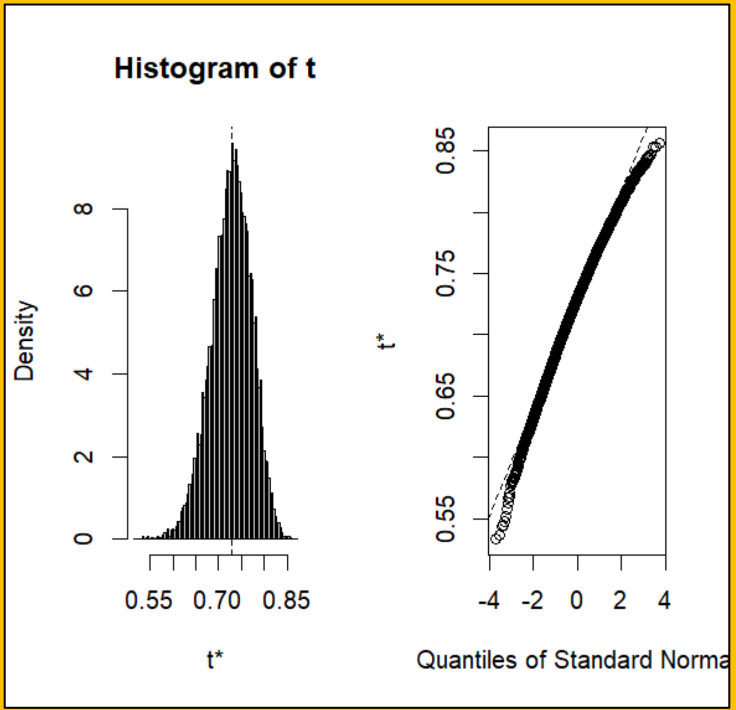
How confident can we be in a correlation between Mana cost and Creature power?

```r
#Question: How confident can we be in a correlation between Mana cost and Creature power?
Power_cost <-data.frame(power=creatures$power, cost=creatures$convertedManaCost)
head(Power_cost)

cor1<-cor(Power_cost$power, Power_cost$cost, method="spearman")
cor1

plot(Power_cost$power, Power_cost$cost)

corFunc <- function(data,i){cor(data[i,"power"], data[i,"cost"], method="spearman")}

reps3 <-boot(Power_cost, corFunc, 10000)
reps3

plot(reps3)

boot.ci(reps3, type="bca")

#density plot
plot(density(reps3$t))
abline(v=cor1, lty="dashed")
```




```
> boot.ci(reps3, type="bca")
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = reps3, type = "bca")

Intervals :
Level       BCa
95%   ( 0.6252,  0.8024 )
Calculations and Intervals on Original Scale
```



# Examples

## K-fold validation

How accurate is using power and toughness to predict Mana cost for creatures, using K-fold validation?

```
#Example 3: K-fold validation/ Jackknife

  #Question: How accurate is using power and toughness to predict Mana cost for creatures, using K-fold?
lm3 <-lm(convertedManaCost ~ power * toughness, data=creatures)
summary(lm3)

ctrl <- trainControl(method="cv", number =5)
modelA <-train(convertedManaCost ~ power, data=creatures, method="lm", trControl=ctrl)
modelB <-train(convertedManaCost ~ toughness, data=creatures, method="lm", trControl=ctrl)
modelC <-train(convertedManaCost ~ power + toughness, data=creatures, method="lm", trControl=ctrl)
modelD <-train(convertedManaCost ~ power * toughness, data=creatures, method="lm", trControl=ctrl)

    #RSME=root mean squared error: lower the better;
    #Rsquared=correlation: the higher the better;
    #MAE=mean absolute error: lower the better;

modelA$results; modelB$results; modelC$results; modelD$results
  #Model 4 had lowest RSME, highest Rsquared, and lowest MAE
modelD$finalModel
```

```
> summary(lm3)

Call:
lm(formula = convertedManaCost ~ power * toughness, data = creatures)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.1036 -0.5313 -0.0468  0.4688  3.4379 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept)      0.56212    0.21331   2.635   0.0092 ** 
power            0.72678    0.09178   7.918 3.12e-13 ***
toughness        0.63933    0.08303   7.700 1.11e-12 ***
power:toughness -0.06187    0.01446  -4.279 3.14e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9487 on 168 degrees of freedom
Multiple R-squared:  0.6911,	Adjusted R-squared:  0.6856 
F-statistic: 125.3 on 3 and 168 DF,  p-value: < 2.2e-16
```

```
> modelA$results; modelB$results; modelC$results; modelD$results
  intercept     RMSE  Rsquared       MAE    RMSESD RsquaredSD      MAESD
1      TRUE 1.100968 0.6108597 0.8388541 0.1124086 0.05188265 0.09605917
  intercept     RMSE  Rsquared       MAE    RMSESD RsquaredSD      MAESD
1      TRUE 1.112877 0.5833777 0.8348641 0.0970927 0.04482937 0.06114345
  intercept     RMSE  Rsquared       MAE     RMSESD RsquaredSD      MAESD
1      TRUE 1.005161 0.6442761 0.7370279 0.07047771 0.09868055 0.04877122
  intercept      RMSE  Rsquared       MAE     RMSESD RsquaredSD      MAESD
1      TRUE 0.9545428 0.6822408 0.6921889 0.09801078 0.08017481 0.09421169
>   #Model 4 had lowest RSME, highest Rsquared, and lowest MAE
> modelD$finalModel

Call:
lm(formula = .outcome ~ ., data = dat)

Coefficients:
      (Intercept)              power          toughness  `power:toughness`  
          0.56212            0.72678            0.63933           -0.06187  
```

# Examples

## K-fold validation

How accurate is using power and toughness to predict Mana cost for creatures, using jackknife validation?


```
#Question: what about jackknife?
jlm <-pcr(convertedManaCost ~ power * toughness, data=creatures, validation="LOO", jackknife=TRUE)

jack.test(jlm)
```

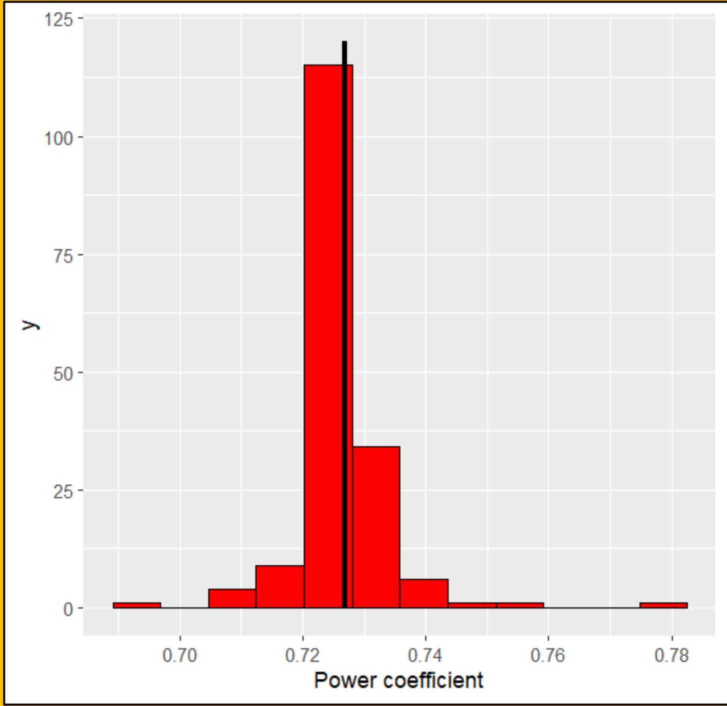
```
> jack.test(jlm)
Response convertedManaCost (3 comps):
                 Estimate Std. Error  Df t value  Pr(>|t|)    
power            0.726778   0.096417 171  7.5379 2.668e-12 ***
toughness        0.639332   0.099895 171  6.4000 1.440e-09 ***
power:toughness -0.061874   0.014454 171 -4.2808 3.097e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

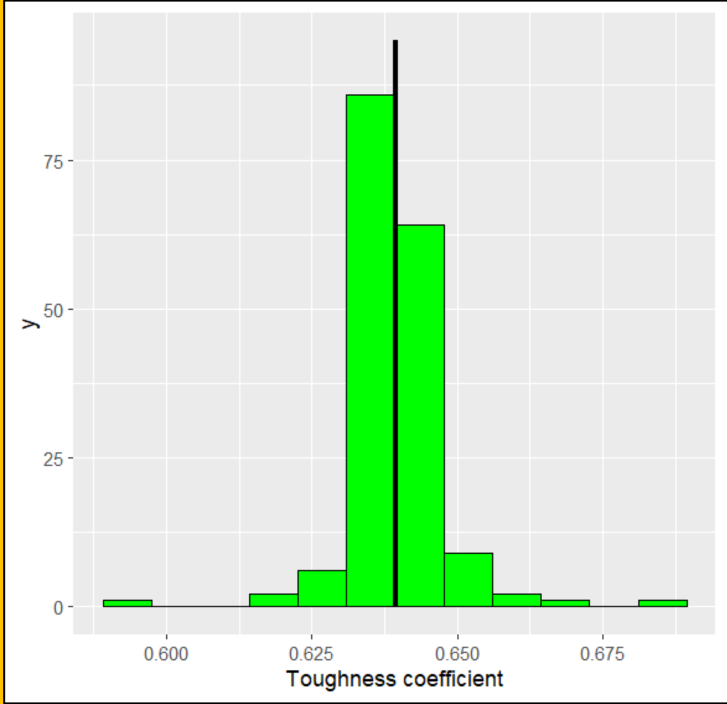
# Examples

## K-fold validation

Visualization (jackknife)

```
#visualization example:
lm3$coefficients

power_c <-unname(lm3$coefficients[2])
tough_c <-unname(lm3$coefficients[3])

lc<-length(creatures$convertedManaCost)
power_coeff_list <-c()
tough_coeff_list <-c()
for (i in seq(1:lc)){
  new_df <-creatures[-c(i), ]
  lm_loo <-lm(convertedManaCost ~ power * toughness, data=new_df)
  power_coeff <-unname(lm_loo$coefficients[2])
  tough_coeff <-unname(lm_loo$coefficients[3])
  power_coeff_list <<-c(power_coeff_list, power_coeff)
  tough_coeff_list <<-c(tough_coeff_list, tough_coeff)
}
power_jack_df <-data.frame(coeff=power_coeff_list)
tough_jack_df <-data.frame(coeff=tough_coeff_list)


plot2 <-ggplot(data=power_jack_df, aes(x=coeff))+
  geom_histogram(bins=12, fill="red", col="black")+
  geom_segment(aes(x=power_c, y=0, xend=power_c, yend=120), colour="black", linewidth=1.2)+
  labs(x="Power coefficient")
plot2

plot3 <-ggplot(data=tough_jack_df, aes(x=coeff))+
  geom_histogram(bins=12, fill="green", col="black")+
  geom_segment(aes(x=tough_c, y=0, xend=tough_c, yend=95), colour="black", linewidth=1.2)+
labs(x="Toughness coefficient")
plot3
```



DaCCoTA
DAKOTA COMMUNITY COLLABORATIVE ON TRANSLATIONAL ACTIVITY

# References

UND
UNIVERSITY OF NORTH DAKOTA


| Presentation | |
|---|---|
| [1] | https://www.youtube.com/watch?v=BhY-un6JURA |
| [2] | https://en.wikipedia.org/wiki/Resampling_(statistics) |
| [3] | https://www.indeed.com/career-advice/career-development/what-is-resampling |
| [4] | http://wise.cgu.edu/wp-content/uploads/2015/04/Introduction-to-Resampling-Techniques-110901.pdf |
| [5] | https://www.datasciencecentral.com/resampling-methods-comparison/ |
| [6] | https://strata.uga.edu/8370/lecturenotes/resampling.html#:~:text=There%20are%20four%20main%20types,intervals%20on%20a%20parameter%20estimate |
| [7] | https://machinelearningmastery.com/k-fold-cross-validation/ |
| [8] | https://thumbs.dreamstime.com/z/permutation-system-example-vector-illustration-diagram-permutation-system-example-vector-illustration-diagram-arranging-parts-173141191.jpg |
| [9] | https://towardsdatascience.com/an-introduction-to-the-bootstrap-method-58bcb51b4d60 |

# References


| Presentation | |
|---|---|
| **[10]** | https://blogs.sas.com/content/iml/files/2018/12/bootstrapSummary.png |
| **[11]** | http://www.statisticshowto.com/wp-content/uploads/2016/10/bootstrap-sample.png |
| **[12]** | https://machinelearningmastery.com/k-fold-cross-validation/ |
| **[13]** | https://d1zx6djv3kb1v7.cloudfront.net/wp-content/media/2020/04/K-Fold-Cross-Validation-2-i2tutorials.png |
| **[14]** | https://i.ytimg.com/vi/eKc9GHliSr4/maxresdefault.jpg |
| **[15]** | https://en.wikipedia.org/wiki/Monte_Carlo_method |
| **[16]** | https://i.gr-assets.com/images/S/compressed.photo.goodreads.com/books/1369156944i/17605831._UY630_SR1200,630_.jpg |
| **[17]** | https://mtgjson.com/getting-started/ |

DaCCoTA
DAKOTA COMMUNITY COLLABORATIVE ON TRANSLATIONAL ACTIVITY

# References

UND
UNIVERSITY OF NORTH DAKOTA

## R code


- https://rcompanion.org/handbook/K_01.html
- https://www.statology.org/bootstrap-standard-error-in-r/
- https://machinelearningmastery.com/k-fold-cross-validation/
- https://www.statology.org/k-fold-cross-validation-in-r/
- https://search.r-project.org/CRAN/refmans/pls/html/jack.test.html

DaCCoTA
DAKOTA COMMUNITY COLLABORATIVE ON TRANSLATIONAL ACTIVITY

# Acknowledgements

UNIVERSITY OF NORTH DAKOTA

The DaCCoTA is supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number U54GM128729.

For the labs that use the Biostatistics, Epidemiology, and Research Design Core in any way, including this Module, please acknowledge us for publications: "***Research reported in this publication was supported by DaCCoTA (the National Institute of General Medical Sciences of the National Institutes of Health under Award Number U54GM128729)***"

DaCCoTA
DAKOTA COMMUNITY COLLABORATIVE ON TRANSLATIONAL ACTIVITY