Package: pbANOVA (via r-universe)

October 20, 2026

Type Package

Title Parametric Bootstrap for ANOVA Models

Version 0.1.0

Maintainer Sarah Alver <salver@unm.edu>

Description Parametric bootstrap (PB) has been used for three-way ANOVA model with unequal group variances.

License GPL-2 | GPL-3

Imports MASS, Rmisc, plyr, DescTools, lmtest,dplyr

Encoding UTF-8

LazyData true

RoxygenNote 7.1.2

NeedsCompilation no

Author Sarah Alver [aut, cre], Guoyi Zhang [aut]

Depends R (>= 3.5.0)

Config/pak/sysreqs cmake make libuv1-dev libssl-dev libx11-dev zlib1g-dev

Repository https://sarahkatharine.r-universe.dev

Date/Publication 2022-05-12 18:10:02 UTC

RemoteUrl https://github.com/cran/pbANOVA

RemoteRef HEAD

RemoteSha 6fe38ce19affbef3b5ca149d334d3cb8a156642a

Contents

alg.A1 2
alg.ABC 3
alg.BC 3
alg.C 4
alg.C.AB 5
barleyh20 6
dunnett.PB 6

fedata 9
potato 10
Q.ABmc 11
Q.Amc 13
Q.Amc_oneway 17

Index **19**

`alg.A1` *PB test for main effect of one-way ANOVA*

Description

Using Parametric Bootstrap to simulate a distribution for the main effect of one-way ANOVA

Usage

```
alg.A1(ns, ybars, s2, a,L)
```

Arguments

`ns`	sample size for each group
`ybars`	sample mean for each group
`s2`	sample variance for each group
`a`	level of the factor
`L`	Number of simulated values for the distribution

Value

the simulated p-value

Examples

```
#see Q.Amc_oneway
```

alg.ABC *test three factor interaction*

Description

Using Parametric Bootstrap to simulate a distribution and find a p-value for the test

Usage

```
alg.ABC(ns, ybars, s2, a, b, c, L)
```

Arguments

ns	sample size for each group
ybars	sample mean for each group
s2	sample variance for each group
a	Number of levels for factor A
b	Number of levels for factor B
c	Number of levels for factor C
L	Number of simulated values for the distribution

Value

Q: p_value for the three factor interaction test

Examples

```
#See Q.ABmc
```

alg.BC *test two factor interaction*

Description

Using Parametric Bootstrap to simulate a distribution and find a p-value for the test.

Usage

```
alg.BC(ns, ybars, s2, a, b, c, L)
```

Arguments

ns	sample size for each group
ybars	sample mean for each group
s2	sample variance for each group
a	Number of levels for factor A
b	Number of levels for factor B
c	Number of levels for factor C
L	Number of simulated values for the distribution

Value

Q: p_value for the two factor interaction test

Examples

```
#See Q.ABmc
# note that the ns, ybars and s2 vectors need to be in the order reflecting subscripts

# 111, 112, 113...., 121, 122, 123, ... , ... abc. The summarySE function from the package

# Rmisc is handy for doing this.  The order the user specifies the "groupvars" argument will

# put the factors in order A, B, C.  This order will matter when testing different two-way

# interaction terms and different main effects.  See comments in the potato example.
```

---

alg.C *for three-way ANOVA that has no significant interaction terms to test main effects*

Description

Using Parametric Bootstrap to simulate a distribution and find a p-value for the test

Usage

```
alg.C(ns, ybars, s2, a, b, c, L)
```

Arguments

ns	sample size for each group
ybars	sample mean for each group
s2	sample variance for each group
a	Number of levels for factor A

b	Number of levels for factor B
c	Number of levels for factor C
L	Number of simulated values for the distribution

Value

a simulated p-value for testing a main effect

Examples

```
#See Q.Amc
```

`alg.C.AB` *tests factor C when only AB interaction is present.*

Description

Using Parametric Bootstrap to simulate a distribution and find a p-value for the test

Usage

```
alg.C.AB(ns, ybars, s2, a, b, c, L)
```

Arguments

ns	sample size for each group
ybars	sample mean for each group
s2	sample variance for each group
a	Number of levels for factor A
b	Number of levels for factor B
c	Number of levels for factor C
L	Number of simulated values for the distribution

Value

Q: p_value for the test for factor C main effect when only AB interaction is present.

Examples

```
#See Q.ABmc
```

barleyh20 *barleyh20 data*

Description

Water uptake in barley for 2 genotypes, 2 sites, 2 years for various periods of steeping time. 2 reps per treatment combination.

Usage

```
data(barleyh20)
```

Format

This data frame contains the following columns:

genotype: 1 = Troubadour, 2 = mutant from Troubadour

site: 1 = Bell-lloc, Spain, 2 = Dundee, Scotland

year: 91 = 1991, 92 = 1992

replicate: Replicate

time: Steeping Time (Hours)

wt Weight of sample (grams)

References

J.L. Molina-Cano, T. Ramo, et al. (1995). "Effect of Grain Composition Water Uptake by Malting Barley: A Genetic and Environmental Study," Journal of the Institute of Brewing, Vol. 101, #2, pp. 79-83.

dunnett.PB *PB multiple comparisons of the levels of factor A aginst the control group*

Description

Using Parametric Bootstrap to simulate a distribution for the multiple comparisons of treatment groups against a control

Usage

```
dunnett.PB(L, ns, means, s2, alpha)
```

Arguments

L	Number of simulated values for the distribution
ns	sample size for each group
means	sample mean for each group
s2	sample variance for each group
alpha	significant level

Value

D.crit: The (1 - alpha) percentile of the simulated distribution

result: The differences, confidence intervals for the difference, and p-values for comparisons of each factor level vs. the control.

Examples

```
#This one gets a different result between the PB method and the traditional Dunnett's test.
#Constant variance assumption appears violated on residual plots.
#The breusch pagan test shows close to violating (p=0.0596) while levene's test (using #median)
#does not show a violation.  Data is mildly unbalanced.

#Traditional Dunnett's test says group 4-6 are different from "control", while the PB method
# only identifies group 5 and 6.
#Group 4 has a larger variance than the others. The pooled variance/MSE could be too small for
#this group and lead to arificially large test statistic.
#MSE is 0.0004274 and the sample variance of group 4 is 0.00169.

#The authors of the paper do not claim significance, they report the means and state
#that there is a delineation between 30 and 40 feet.  This seems true when looking at the
#means, but the measurements at 40 feet do have a larger variance than those at other depths.

#Practical interpretation:
#If your goal was to get the most iron rich water from as shallow depth as possible,
#knowing that the surface (control?) was not rich enough, and you decided to go 40 feet
# deep, you may still get water that didn't have enough iron content for your purpose.
#Suppose you wanted at least 0.1 content; based on the means you might use 40 feet, but
# from their data, the measurements were:
#0.098, 0.074, 0.154
#So 2/3 of these samples wouldn't contain enough iron for your purpose.


library(DescTools)  ##Dunnett's test
library(lmtest) #BP test for constant variance
library(dplyr) #data manipulation
library(MASS)

fedata$depth <- factor(fedata$depth)

femod <- lm(Y~depth, data=fedata)
plot(femod$fit, rstandard(femod),
    main="Fitted-Residual Plot, One-Way ANOVA Model", sub="Iron Data")
```

```
    #appears to violate equal variance assumption

bptest(femod) #close to violation

#what about normality?
qqnorm(rstandard(femod), main="Normal QQ-Plot, Standardized Residuals",
sub="One-Way ANOVA Model, Iron Data")
shapiro.test(rstandard(femod)) #does not violate

fe.sums <-fedata %>% group_by(depth) %>% summarise(means=mean(Y),
vars = var(Y), sd=sd(Y), ns=length(depth))
fe.sums
#summarySE(fedata, "Y", "depth") from Rmisc pkg does same thing as fe.sums above

pbd.fe <- dunnett.PB(L=5000, ns=fe.sums$ns, means=fe.sums$means,
                           s2=fe.sums$vars, alpha=0.05)

pbd.fe$result
pbd.fe$D.crit

##grp 5 and 6 sig diff from group 1
DunnettTest(Y~depth, data=fedata)

##Dunnett's test also says group 4 is different.
##group 4 has a larger variance than the others.
##pooled variance/MSE could be too small for this group
##and lead to arificially large test statistic.

anova(femod)
#MSE is 0.0004274
#sample variance of group 4 is 0.00169

#Use traditional Dunnett's test on transformed data for comparisons

#attempt log transformation
fedata$logY <-log(fedata$Y)
felogmod <- lm(logY~depth, data=fedata)
bptest(felogmod) #still violates
#LeveneTest(logY~depth, data=fedata)
plot(felogmod$fit, rstandard(felogmod))
shapiro.test(rstandard(felogmod)) #still for normality

#square root transformation?
fedata$srY <- sqrt(fedata$Y)
fesrmod <- lm(srY~depth, data=fedata)
bptest(fesrmod) #still violates, worse
plot(fesrmod$fit, rstandard(fesrmod))

boxcox(femod)
#lambda=-0.2

fedata$bcY <- with(fedata, (Y^(-0.2) - 1)/-0.2)
febcmod <- lm(bcY~depth, data=fedata)
```

```
#gives same F-statistic as lm(Y^(-0.2) ~ depth, data=fedata),


bptest(febcmod)  #still violates
plot(febcmod$fit, rstandard(febcmod))

##mg/L is a proportion, try arcsin
fedata$asY <- asin(sqrt(fedata$Y))
feasmod <-lm(asY~depth, data=fedata)
bptest(feasmod) #still close to violating

plot(felogmod$fit, rstandard(felogmod), main="Log Transform.",
    xlab="Fitted Values", ylab="Standardized Residuals")
plot(febcmod$fit, rstandard(febcmod), main="Box-Cox Transform.",
    xlab="Fitted Values", ylab="Standardized Residuals")
plot(feasmod$fit, rstandard(feasmod), main="Arcsin(Sq. Root) Transform.",
    xlab="Fitted Values", ylab="Standardized Residuals")

#P-values of BP test are similar for log and box-cox, plots look a little better
##log transform may be considered simpler, so try that

anova(felogmod)
DunnettTest(logY~depth, data=fedata)  #still identifies 40 feet and above

DunnettTest(bcY~depth, data=fedata) #still identifies 40 feet and above
shapiro.test(rstandard(febcmod)) # normality still ok
#W = 0.95535, p-value = 0.4556
```

`fedata` *fedata data*

Description

Iron content of water at various water depths.

Usage

```
data(fedata)
```

Format

This data frame contains the following columns:

depth: depth of the water

Y: Iron content of water

References

paper: https://www.jstor.org/stable/1351176

`potato` *potato data*

Description

This dataset potato is from an experiment on how plants adapt to cold climates. The investigators decided to study this problem after observing that plants that have been conditioned to cold previously appear to suffer less damage from the cold. Two species of potato were studied (species 1 and 2). Each plant was exposed to one of two acclimatization regimes (1= plant was kept in cold room; 0= plant was kept at room temperature) for several days. Later, plants were subjected to one of two cold temperatures (-4 degrees C is coded as 1; -8 degrees C is coded as 2). Two responses were measured: damage score for photosynthesis (photo), and damage score for ion leakage (leak).

Usage

```
data(potato)
```

Format

This data frame contains the following columns:

variety: Two species of potato were studied (species 1 and 2)

regime: Each plant was exposed to one of two acclimatization regimes (1= plant was kept in cold room; 0= plant was kept at room temperature) for several days.

temp: plants were subjected to one of two cold temperatures (-4 degrees C is coded as 1; -8 degrees C is coded as 2)

photo: damage score for photosynthesis

leak: damage score for ion leakage

Details

Use ion leakage to be the response variable. Some of the 80 plants originally assigned to the treatment combinations were lost during the experiment. Analyze the data from the plants that made it through, and assess the effects of the three experimental factors species, regime, and temperature on the response leakage.

References

Alver and Zhang (2022), Parametric Bootstrap Procedures for Three-Factor ANOVA and Multiple Comparison Procedures with Unequal Group Variances

Alver and Zhang (2022), Multiple Comparisons of Treatment vs Control Under Unequal Variances Using Parametric Bootstrap

Q.ABmc *multiple comparisons of the levels of AB interactions*

Description

Using Parametric Bootstrap to simulate a distribution and find a p-value for the test

Usage

```
Q.ABmc(L=5000, ns, means, s2, alpha=0.05, a, b, c)
```

Arguments

L	Number of simulated values for the distribution
ns	sample size for each group
means	sample mean for each group
s2	sample variance for each group
alpha	significant level
a	Number of levels for factor A
b	Number of levels for factor B
c	Number of levels for factor C

Value

Q.crit: The (1- alpha) percentile of the simulated distribution.

res.df: A dataframe containing the differences between each pair of factor levels, standard errors, confidence interval for the differences, test statistic for each pair, p-value, and indicator of whether the difference was statistically significant for each pair.

ybarij: estimated group mean for level i, j

var.YAB: estimated variance for each group mean

Q.test: largest test statistic from all pairs

Examples

```
#'# We need elapsed time less than 5 seconds to publish
# the package with this example. Hence, there are # before
#a couple of PB algrithms. Remove # before you run the code.

#Note that when running the example, the user should get similar p-values to the ones commented
# in the example, but they will not be identical.
attach(potato)
regime<-factor(regime)
variety<-factor(variety)
temp<-factor(temp)
```

```
#there are two levels for each factor, so a=b=c=2
library(Rmisc)
summarySE(potato, measurevar="leak", groupvars=c("variety","regime","temp"))


#need to extract group sizes (ns), group var's (s2), means (ybars) for function
pot.ns <- summarySE(potato, measurevar="leak", groupvars=c("variety","regime","temp"))$N
pot.means <- summarySE(potato, measurevar="leak", groupvars=c("variety","regime","temp"))$leak
pot.s2 <- summarySE(potato, measurevar="leak", groupvars=c("variety","regime","temp"))$sd^2
#alg.ABC(ns=pot.ns, ybars=pot.means,s2=pot.s2, a=2, b=2, c=2, L=5000)
#0.1626, pvalue, so we do not reject, so we should drop the three way term.




#alg.BC(ns=pot.ns, ybars=pot.means,s2=pot.s2, a=2, b=2, c=2, L=5000)
#0.202, not significant, so the regime:temp interaction is not significant
#to check the other two-way interactions we need to reorder the data so that
#the 'BC' term is either regime:variety or temp:variety

pot.ns.TRV <- summarySE(potato, measurevar="leak", groupvars=c("temp","regime","variety"))$N
pot.means.TRV <- summarySE(potato, measurevar="leak", groupvars=c("temp","regime","variety"))$leak
pot.s2.TRV <- summarySE(potato, measurevar="leak", groupvars=c("temp", "regime", "variety"))$sd^2
#alg.BC(ns=pot.ns.TRV, ybars=pot.means.TRV,s2=pot.s2.TRV, a=2, b=2, c=2, L=5000)
#p=0, reject H_0.  the regime:variety interaction is significant


pot.ns.RTV <- summarySE(potato, measurevar="leak", groupvars=c("regime", "temp","variety"))$N
pot.means.RTV <- summarySE(potato, measurevar="leak", groupvars=c("regime","temp","variety"))$leak
pot.s2.RTV <- summarySE(potato, measurevar="leak", groupvars=c("regime", "temp", "variety"))$sd^2
#alg.BC(ns=pot.ns.RTV, ybars=pot.means.RTV,s2=pot.s2.RTV, a=2, b=2, c=2, L=5000)
#p=0.0.3652, do not reject, the temp:variety interaction is not significant


##next we can see if we are able to drop the main effect 'temp',
##not involved with the regime:variety int.
##temp is factor 'A' in the TRV model above.

##algorithm 4 tests factor C when only AB interaction is present.
## so we need the order that makes 'temp' factor C
#the way we originally ordered it, to test the ABC interaction

#alg.C.AB(ns=pot.ns, ybars=pot.means,s2=pot.s2, a=2, b=2, c=2, L=5000)
#p-value is 0.002, so we cannot drop the temp term.  the final model is

#y = variety + regime + temp + variety:regime



Q.ABmc(ns=pot.ns, means=pot.means,s2=pot.s2, a=2, b=2, c=2, L=5000)
```

```
#     95%
#2.832115

#$res.df
#     groups differences std.errs      ci.lo      ci.hi test.stat      p sig
# 1 1 - 1 2   -1.803638 1.821063  -6.961098   3.353822 0.9904314 0.7676   -
# 1 1 - 2 1  -22.501936 2.895282 -30.701708 -14.302164 7.7719316 0.0000   *
# 1 1 - 2 2   -1.248118 1.615681  -5.823913   3.327677 0.7725027 0.8696   -
# 1 2 - 2 1  -20.698298 2.781589 -28.576079 -12.820517 7.4411765 0.0000   *
# 1 2 - 2 2    0.555520 1.401787  -3.414501   4.525541 0.3962943 0.9766   -
# 2 1 - 2 2   21.253818 2.651678  13.743962  28.763674 8.0152341 0.0000   *

#$ybarij
#         [,1]     [,2]
#[1,]  4.91472 6.718358
#[2,] 27.41666 6.162838

#$var.YAB
#        [,1]      [,2]
#[1,] 1.980845 1.3354254
#[2,] 6.401815 0.6295805

#$Q.test
#[1] 8.015234
```

`Q.Amc` *PB multiple comparisons of the levels of factor A (output like TukeyHSD)*

Description

Using Parametric Bootstrap to simulate a distribution for the multiple comparisons and calculate a test stat

Usage

```
Q.Amc(L=5000, ns, means, s2, alpha=0.05, a, b, c)
```

Arguments

`L`	Number of simulated values for the distribution
`ns`	sample size for each group
`means`	sample mean for each group
`s2`	sample variance for each group
`alpha`	significant level
`a`	Number of levels for factor A
`b`	Number of levels for factor B
`c`	Number of levels for factor C

Value

Q.crit: The (1-alpha) percentile of the simulated distribution.

Q.test: largest test statistic from all pairs

res.df: A dataframe containing the differences between each pair of factor levels, standard errors, confidence interval for the differences, test statistic for each pair, p-value, and indicator of whether the difference was statistically significant for each pair.

Examples

```
# We need elapsed time less than 5 seconds to publish
# the package with this example. Hence, there are # before
#a couple of PB algrithms. Remove # before you run the code.

#function to make everything but the response a factor

make.factor <- function(dataset, fact.cols){
for( i in fact.cols){
 dataset[,i] <- factor(dataset[,i])
}
return(dataset)
}

barley_ex <- make.factor(barleyh20, 1:5)
##this dataset has 4 factors, ignore year

library(Rmisc)

library(MASS)

#library(lmtest)

summarySE(barley_ex, "wt", c("genotype", "site", "time"))
#ignore year, note that the data are balanced

summary(barley_ex$wt)
mod1 <- lm(wt~genotype*site*time, data=barley_ex)
anova(mod1)
plot(mod1$fit, mod1$resid)
qqnorm(mod1$resid)
shapiro.test(mod1$resid)

boxcox(mod1, lambda=seq(-4, -2, by=0.1)) #lambda approx -3.5
mod2 <- lm(wt^(-3.5)~genotype*site*time, data=barley_ex)
plot(mod2$fit, mod2$resid) #worse?

#go with untransformed data? drop 3way term
mod3 <- lm(wt~genotype + site + time + genotype:site + genotype:time + site:time, data=barley_ex)
anova(mod3) #site:time ns

anova(lm(wt~genotype + site + time + genotype:site + genotype:time, data=barley_ex))
anova(lm(wt~genotype + site + time + genotype:site, data=barley_ex))
```

```
anova(lm(wt~genotype + site + time, data=barley_ex))
anova(lm(wt~site + time, data=barley_ex))
anova(lm(wt~time, data=barley_ex))
TukeyHSD(aov(wt  ~ time, data=barley_ex))  #all sig except 35-30 and 20-15 (0.0569)


###use PB methods
summarySE(barley_ex, "wt", c("genotype", "site", "time"))
#note that the data are balanced

#need to extract group sizes (ns), group var's (s2), means (ybars) for function
barley.ns <- summarySE(barley_ex, "wt", c("genotype", "site", "time"))$N
barley.means <- summarySE(barley_ex, "wt", c("genotype", "site", "time"))$wt
barley.s2 <- summarySE(barley_ex, "wt", c("genotype", "site", "time"))$sd^2

#alg.ABC(ns=barley.ns, ybars=barley.means,s2=barley.s2, a=2, b=2, c=7, L=5000)
#p=0.9996, can drop 3way term

#can we drop the site:time int term?
#alg.BC(ns=barley.ns, ybars=barley.means,s2=barley.s2, a=2, b=2, c=7, L=5000)
#p=0.9998, drop

#reorder data to make the different two-way terms
barleyTSG.ns <- summarySE(barley_ex, "wt", c("time", "site", "genotype"))$N
barleyTSG.means <- summarySE(barley_ex, "wt", c("time", "site", "genotype"))$wt
barleyTSG.s2 <- summarySE(barley_ex, "wt", c("time", "site", "genotype"))$sd^2

#alg.BC(ns=barleyTSG.ns, ybars=barleyTSG.means, s2=barleyTSG.s2, a=7, b=2, c=2, L=5000)
#p=0.9988, drop site:genotype

#reorder to SGT, can we drop genotype:time?
barleySGT.ns <- summarySE(barley_ex, "wt", c("site", "genotype", "time"))$N
barleySGT.means <- summarySE(barley_ex, "wt", c("site", "genotype", "time"))$wt
barleySGT.s2 <- summarySE(barley_ex, "wt", c("site", "genotype", "time"))$sd^2

#alg.BC(ns=barleySGT.ns, ybars=barleySGT.means,s2=barleySGT.s2, a=2, b=2, c=7, L=5000)
#p=0.9976, drop
#alg.C(ns=barley.ns, ybars=barley.means,s2=barley.s2, a=2, b=2, c=7, L=5000) #GST
#p=0, time has signif efffect, same conclusion as F-test

#alg.C(ns=barleyTSG.ns, ybars=barleyTSG.means, s2=barleyTSG.s2, a=7, b=2, c=2, L=5000)
#p=0.9996 no signif effect of genotype


##site?
barleyGTS.ns <- summarySE(barley_ex, "wt", c("genotype", "time", "site"))$N
barleyGTS.means <- summarySE(barley_ex, "wt", c("genotype", "time", "site"))$wt
barleyGTS.s2 <- summarySE(barley_ex, "wt", c("genotype", "time", "site"))$sd^2

#alg.C(ns=barleyGTS.ns, ybars=barleyGTS.means, s2=barleyGTS.s2, a=2, b=7, c=2, L=5000)
#p=0.9998, site is NS
```

```
#multiple comparisons
#this function tests all pairwise comparisons of the levels of factor A,
# so we use the TSG order
Q.Amc(L=5000, ns=barleyTSG.ns, means=barleyTSG.means, s2=barleyTSG.s2,
alpha=0.05, a=7, b=2, c=2)

#Demonstrate the method on unbalanced data by collapsing time into L, M, H
#barley_ex$time2 <- "M"
#barley_ex$time2 <- ifelse(as.numeric(barley_ex$time) <=2, "L", barley_ex$time2)
#barley_ex$time2 <- ifelse(as.numeric(barley_ex$time) >=6, "H", barley_ex$time2)
#barley_ex$time2 <- as.factor(barley_ex$time2)

#still pretty balanced, separate the lowest level
#barley_ex$time2 <- ifelse(as.numeric(barley_ex$time) <2, "LL", barley_ex$time2)
#barley_ex$time2 <- as.factor(barley_ex$time2)

#summarySE(barley_ex, "wt", c("genotype", "time2", "site"))
#two of the bigger groups N=12 have larger var
#anova(lm(wt ~genotype*time2*site, data=barley_ex))
#still looks like time2 the only sig factor

#library(lmtest)
#bptest(lm(wt ~genotype*time2*site, data=barley_ex)) #violates

#mod.un <- lm(wt ~genotype*time2*site, data=barley_ex)
#plot(mod.un$fit, mod.un$resid)
#qqnorm(mod.un$resid)

#boxcox(lm(wt ~genotype*time2*site, data=barley_ex), lambda=seq(-6, -4, by=0.1))
#lambda = -4.5


#the above transformations didn't work so just try the untransformed data
#the three way interaction term was not significant
#anova(lm(wt ~genotype+ time2 + site + genotype:time2 + genotype:site + time2:site, data=barley_ex))
#anova(lm(wt ~genotype+ time2 + site + genotype:time2 + genotype:site, data=barley_ex))
#anova(lm(wt ~genotype+ time2 + site + genotype:time2, data=barley_ex))
#anova(lm(wt ~genotype+ time2 + site, data=barley_ex))
#anova(lm(wt ~genotype+ time2, data=barley_ex))
#anova(lm(wt ~time2, data=barley_ex))

#TukeyHSD(aov(wt ~time2, data=barley_ex)) #all pairs significantly different


#PB methods
#barleyGST2.ns <- summarySE(barley_ex, "wt", c("genotype", "site", "time2"))$N
#barleyGST2.means <- summarySE(barley_ex, "wt", c("genotype", "site", "time2"))$wt
#barleyGST2.s2 <- summarySE(barley_ex, "wt", c("genotype", "site", "time2"))$sd^2


#alg.ABC(ns=barleyGST2.ns, ybars=barleyGST2.means,s2=barleyGST2.s2, a=2, b=2, c=4, L=5000)
#p=0.9734, can drop 3way term
```

```
#alg.BC(ns=barleyGST2.ns, ybars=barleyGST2.means,s2=barleyGST2.s2, a=2, b=2, c=4, L=5000)
#p=0.94, can drop site:time2

#barleySGT2.ns <- summarySE(barley_ex, "wt", c("site","genotype",  "time2"))$N
#barleySGT2.means <- summarySE(barley_ex, "wt", c("site", "genotype", "time2"))$wt
#barleySGT2.s2 <- summarySE(barley_ex, "wt", c("site","genotype", "time2"))$sd^2

#alg.BC(ns=barleySGT2.ns, ybars=barleySGT2.means,s2=barleySGT2.s2, a=2, b=2, c=4, L=5000)
#p=0.9952, can drop genotype:time2


#barleyTSG2.ns <- summarySE(barley_ex, "wt", c("time2", "site","genotype"))$N
#barleyTSG2.means <- summarySE(barley_ex, "wt", c("time2","site", "genotype"))$wt
#barleyTSG2.s2 <- summarySE(barley_ex, "wt", c("time2","site","genotype"))$sd^2


#alg.BC(ns=barleyTSG2.ns, ybars=barleyTSG2.means,s2=barleyTSG2.s2, a=4, b=2, c=2, L=5000)
#p=0.9556, can drop site:genotype

#alg.C(ns=barleyGST2.ns, ybars=barleyGST2.means,s2=barleyGST2.s2, a=2, b=2, c=4, L=5000)
#p=0, time still has significant effect

#alg.C(ns=barleyTSG2.ns, ybars=barleyTSG2.means,s2=barleyTSG2.s2, a=4, b=2, c=2, L=5000)
#p=0.9716, genotype is not significant


#barleyTGS2.ns <- summarySE(barley_ex, "wt", c("time2","genotype", "site"))$N
#barleyTGS2.means <- summarySE(barley_ex, "wt", c("time2","genotype", "site"))$wt
#barleyTGS2.s2 <- summarySE(barley_ex, "wt", c("time2","genotype", "site"))$sd^2


#alg.C(ns=barleyTGS2.ns, ybars=barleyTGS2.means,s2=barleyTGS2.s2, a=4, b=2, c=2, L=5000)
#p=0.9904, site is not significant


##mult comparisons of factor A so we put time first
#Q.Amc(L=5000, ns=barleyTSG2.ns, means=barleyTSG2.means, s2=barleyTSG2.s2,
# alpha=0.05, a=4, b=2, c=2)
#all sig, agrees with Tukey's test

#TukeyHSD(aov(wt ~time2, data=barley_ex))
```

Q.Amc_oneway *PB multiple comparisons of factor A in one-way ANOVA*

Description

Using Parametric Bootstrap to simulate a distribution for multiple comparison in one-way ANOVA

Usage

```
Q.Amc_oneway(L,ns, means, s2, alpha)
```

Arguments

L	Number of simulated values for the distribution
ns	sample size for each group
means	sample mean for each group
s2	sample variance for each group
alpha	significant level

Value

the simulated p-value

D.crit: The (1 - alpha) percentile of the simulated distribution

res.df: The differences, confidence intervals for the difference, and p-values for comparisons of each two factor levels.

Examples

```
library(pbANOVA)

data(fedata)

fedata$depth <- factor(fedata$depth)

library(Rmisc)
summarySE(fedata, "Y", "depth")

feNs <- summarySE(fedata, "Y", "depth")$N
feYs <- summarySE(fedata, "Y", "depth")$Y
fes2 <- (summarySE(fedata, "Y", "depth")$sd)^2

anova(lm(Y~depth, data=fedata))  #F-test significant
#we saw in the dunnett's example that the equal variance assumption is violated

library(MASS) #need MASS for ginv function for all the interaction and main effects algorithms
alg.A1(ns=feNs, ybars=feYs, s2=fes2, a=6, L=5000)
#p=0.0038

#multiple comparisons
Q.Amc_oneway(L = 5000, ns=feNs, means=feYs, s2=fes2, alpha = 0.05)

#compare to Tukey's test
TukeyHSD(aov(Y~depth, data=fedata))

#results agree only for some levels.
```

Index

∗ **datasets**
barleyh20, 6
fedata, 9
potato, 10

alg.A1, 2
alg.ABC, 3
alg.BC, 3
alg.C, 4
alg.C.AB, 5

barleyh20, 6

dunnett.PB, 6

fedata, 9

potato, 10

Q.ABmc, 11
Q.Amc, 13
Q.Amc_oneway, 17