ST 437/537: Applied Multivariate and Longitudinal Data Analysis

# Multiple Treatments Comparison

***Arnab Maity***
*NCSU Department of Statistics*
*SAS Hall 5240 919-515-1937 amaity[at]ncsu.edu*

## Introduction

The ideas of the previous lecture **[One sample case] (https://www.stat.ncsu.edu/people/maity/courses/st537-S2019/Lecture05_OneSample.html)** can be used to develop procedures for comparison of two or more population mean vectors. We will first study such comparisons when the data arise from a paired design (measurements before/after), then when the data arises from multiple treatments.

## Paired comparison

**Key assmuption**

**Each item/individual in the sample are measured twice**: before and after a treatment/condition on the same attributes. Thus for individual $i$, we have a **pair** of vectors, $\boldsymbol{X}_i$ and $\boldsymbol{Y}_i$, corresponding to measurements before and after treatment, where $\boldsymbol{X}_i$ and $\boldsymbol{Y}_i$ are $p \times 1$ vectors.

Suppose that $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ form a random smaple with mean $\boldsymbol{\mu}_1$, and similarly, $\boldsymbol{Y}_1, \ldots, Y_n$ form a random smaple with mean $\boldsymbol{\mu}_2$. We want to answer the following questions:

- what are the posible values of $\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1$?
- How to test $H_0 : \boldsymbol{\mu}_2 - \boldsymbol{\mu}_1 = \delta_0$ vs. $H_0 : \boldsymbol{\mu}_2 - \boldsymbol{\mu}_1 \neq \delta_0$ for some known vector $\delta_0$?

As an illustrative eaxmaple, consider **[the wastewater monitoring] (data/T6-1.dat)** (Table 6.1 in the Johnson and Wichern textbook). Municipal wastewater treatment plants are required by law to monitor their discharge into rivers and streams on a regular basis. Concerns about the reliability of data from one of these self-monitoring programs led to a study in which samples of effuent were divided and sent to labs for testing. One half of each sample was sent to the Wisconsin State Laboratory of Hygiene and one half was sent to a private commercial laboratory routinely used in the monitoring program. Measurements of biochemical oxygen demand (BOD) and suspended solids (SS) were obtained for $n = 11$ sample splits from the two laboratories.

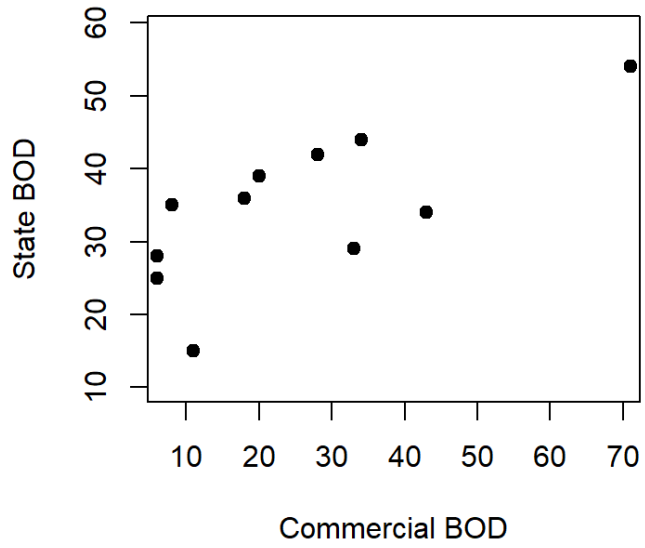
```
dat <- read.table("data/T6-1.dat")
colnames(dat)  = c("Commercial_BOD", "Commercial_SS",
                "StateLab_BOD", "StateLab_SS")
dat
```

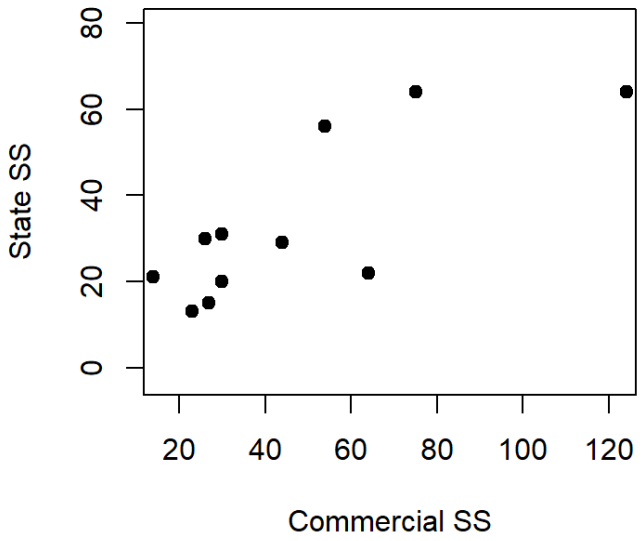
```
##    Commercial_BOD Commercial_SS StateLab_BOD StateLab_SS
## 1               6            27           25          15
## 2               6            23           28          13
## 3              18            64           36          22
## 4               8            44           35          29
## 5              11            30           15          31
## 6              34            75           44          64
## 7              28            26           42          30
## 8              71           124           54          64
## 9              43            54           34          56
## 10             33            30           29          20
## 11             20            14           39          21
```

```
colMeans(dat)
```

```
## Commercial_BOD  Commercial_SS   StateLab_BOD    StateLab_SS
##       25.27273       46.45455       34.63636       33.18182
```

```
library(MASS)
par(mfrow = c(1,2))
eqscplot(dat$Commercial_BOD, dat$StateLab_BOD, pch=19, xlab = "Commercial BOD", ylab =
"State BOD")
eqscplot(dat$Commercial_SS, dat$StateLab_SS, pch=19, xlab = "Commercial SS", ylab = "Sta
te SS")
```




Our question of interest is whether the average (BOD, SS) measures are same between state and commercial labs.

## Univariate case: paired $t$-test and intervals

Let us first review the univariate case. Specifically, for the $i$-th subject, we observe $(X_i, Y_i)$. Here $\mu_1 = E(X_i)$ and $\mu_2 = E(Y_i)$. We want to perform inference on $\mu_2 - \mu_1$.

We first define the difference $D_i = Y_i - X_i$. Note that

$$E(D_i) = \mu_2 - \mu_1 .$$

We assume that $\{D_1, \ldots, D_n\}$ is a random sample from a $N(\mu_2 - \mu_1, \sigma_D^2)$ distribution. Thus a point estimator of $\mu_2 - \mu_1$ is the sample mean of differences, $\bar{D} = \sum_{i=1}^{n} D_i/n$. An point estimator of $\sigma_D^2$ is the sample variance of the differences, $s_D^2$.

***Hypothesis testing:*** *To test $H_0 : \mu_2 - \mu_1 = \delta_0$ vs. $H_0 : \mu_2 - \mu_1 \neq \delta_0$ we use a one-sample t-test: define the test statistic*

$$T = \frac{\bar{D} - \delta_0}{s_D/\sqrt{n}}.$$

*Given a significance level $\alpha$, we reject $H_0$ if*

$$|T| \geq t_{n-1}(\alpha/2).$$

*This test is often called a **paired** $t$**-test**.*

***Confidence interval:*** *We can create a confidence interval for $\mu_2 - \mu_1$ using the t-distribution as well. Specifically, a $(1-\alpha)100\%$ confidence interval for $\mu_2 - \mu_1$ is*

$$\left(\bar{D} \pm t_{n-1}(\alpha/2)\frac{s_D}{\sqrt{n}}\right).$$

When the sample size is large ($n \geq 40$), we can obtain approximate/large sample $z$-interval and $z$-test by replacing $t_{n-1}(\alpha/2)$ by $z(\alpha/2)$.

Consider the watershade data, but only consider the variable `BOD`.

```
# extract the BOD variable for state and commercial labs
state_BOD <- dat$StateLab_BOD
com_BOD <- dat$Commercial_BOD

# Differences
D_BOD <- state_BOD - com_BOD

# 95% confidence Interval and 5% t-test
t.test(D_BOD)
```

```
##
##  One Sample t-test
##
## data:  D_BOD
## t = 2.2001, df = 10, p-value = 0.05243
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.119457 18.846730
## sample estimates:
## mean of x
##  9.363636
```

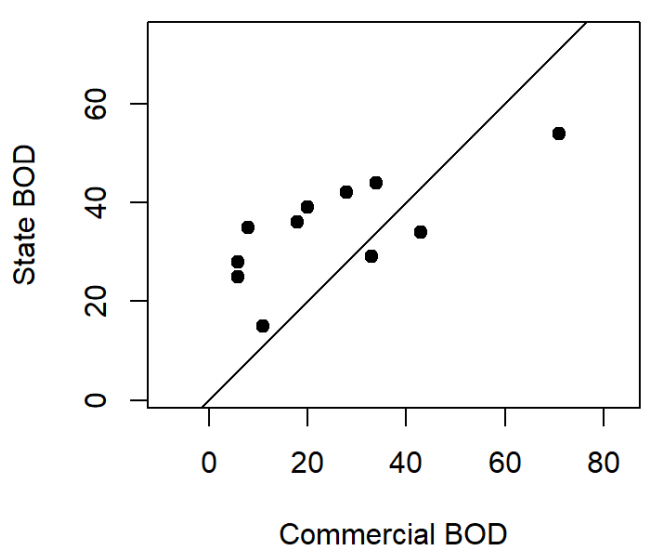
The 95% confidence interval is $(-0.119, 18.847)$. Also the p-value is larger than $0.05$. We can conclude that the average BOD measurements of the state lab is the same as commercial labs.

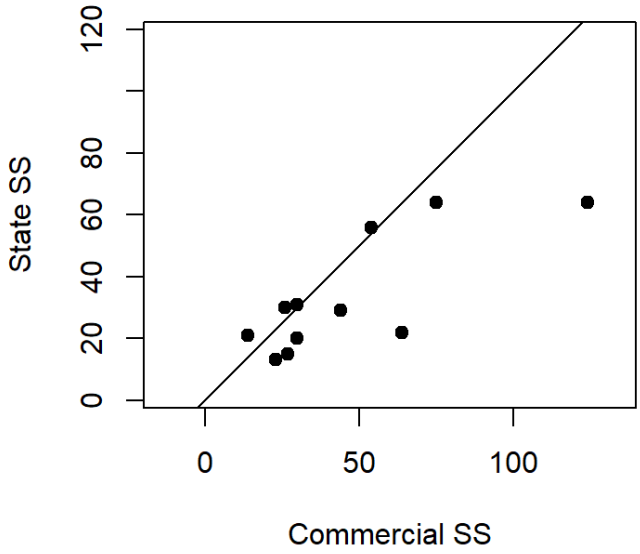
We can run the same procedure on the `ss` variable; the results are show below.

```
##
##  One Sample t-test
##
## data:  D_SS
## t = -2.1515, df = 10, p-value = 0.05692
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -27.0180504   0.4725958
## sample estimates:
## mean of x
## -13.27273
```

It seems that for `ss` as well, we can conclude that the average SS measurements of the state labs is the same as commercial labs.

Looking at the results together, it seems state labs BOD measurements tend to be higher than that of commercil labs (95% CI captures mosty positive values). For `ss`, the situation is opposite: state labs SS measurements tend to be higher than that of commercil labs (95% CI captures mosty negative values). We display the data again, but using the same scale and a diagonal line.




Now let us consider the two variables (SS, BOD) together. The correlation matrix of the data is shown below.

```
round( cor(dat), 3)
```

```
##                Commercial_BOD Commercial_SS StateLab_BOD StateLab_SS
## Commercial_BOD          1.000         0.781        0.723       0.789
## Commercial_SS           0.781         1.000        0.677       0.790
## StateLab_BOD            0.723         0.677        1.000       0.604
## StateLab_SS             0.789         0.790        0.604       1.000
```

Notice the high correlation among the variables. It seems a multivariate analysis is well suited in this case.

## Multivariate inference

In thi situation, we first commpute the **difference vector**

$$\boldsymbol{D}_i = \boldsymbol{Y}_i - \boldsymbol{X}_i.$$

Note that $E(\boldsymbol{D}_i) = \boldsymbol{\mu}_2 - \boldsymbol{\mu}_1$. We assume that the difference vectors $\{D_1, \ldots, D_n\}$ is a random sample from a multivariate normal distribution $N(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_D)$.

A point estimator of $\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1$ is the sample mean vector $\bar{\boldsymbol{D}}$; we can also estimate $\boldsymbol{\Sigma}_D$ by the sample covariance matrix of the difference vectors, $\boldsymbol{S}_D$.

***Hypothesis testing:*** *To test the null hypothesis* $H_0 : \boldsymbol{\mu}_2 - \boldsymbol{\mu}_1 = \delta_0$ *vs* $H_1 : \boldsymbol{\mu}_2 - \boldsymbol{\mu}_1 \neq \delta_0$*, we calculate the* ***Hotelling's*** $T^2$ ***test*** *statistic*

$$T^2 := \frac{(n-p)n}{(n-1)p} \bar{\boldsymbol{D}}^T \boldsymbol{S}_D^{-1} \bar{\boldsymbol{D}};$$

*here* $\bar{\boldsymbol{D}}$ *and* $\boldsymbol{S}_d$ *are the sample mean and sample covariance of* $d_i$*'s. If the null hypothesis is true then* $T^2 \sim F_{p,n-p}$*. Thus we reject* $H_0$ *at level* $\alpha$ *if the observed value of* $T^2$ *exceeds the critical value* $F_{p,n-p}(\alpha)$*.*

Define $\delta_k$ to be the $k$-th component of $\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1$. Also denote $\bar{\boldsymbol{D}} = (\bar{D}_1, \ldots, \bar{D}_p)^T$.

***Simultaneous considence interval*** *The* $100(1-\alpha)\%$ *simultaneous confidence intervals of* $\delta_k$*,* $k = 1, \ldots, p$ *are:*

$$\bar{D}_k \pm \sqrt{\frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha)} \sqrt{S_{D,kk}/n}, k = 1, \ldots, p;$$

*here* $\bar{D}_k$ *is the* $k$*-th element of the vector* $\bar{D}$*, and* $S_{d,kk}$ *is the* $k$*th element of the diagonal of* $\boldsymbol{S}_D$*.*

When the sample size is large, then $\frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha)$ is replaced by $\chi^2_p(\alpha)$.

*The Bonferroni* $100(1-\alpha)\%$ *simultaneous confidence intervals for* $\delta_k$*'s are:*

$$\bar{D}_k \pm t_{n-1}\left(\frac{\alpha}{2p}\right) \sqrt{S_{D,kk}/n}, \quad k = 1, \ldots, p;$$

*here* $S_{D,kk}$ *is the* $k$*th element of the diagonal of* $\boldsymbol{S}_D$*.*

In the watershade data example, we first perform the Hotelling's $T^2$ test.

```
library(ICSNP)

# BOD differences
D_BOD <- dat$StateLab_BOD - dat$Commercial_BOD
# SS differences
D_SS <- dat$StateLab_SS - dat$Commercial_SS

# Obtain the vectors of difference between state and commercial
# Each ROW is one difference vector
Dmat <- cbind(D_BOD, D_SS)

# Hotelling's T2
HotellingsT2(Dmat)
```

```
##
##  Hotelling's one sample T2-test
##
## data:  Dmat
## T.2 = 6.1377, df1 = 2, df2 = 9, p-value = 0.02083
## alternative hypothesis: true location is not equal to c(0,0)
```

We can do this test manually as well.

```
n <- nrow(dat)
p <- 2

# Dbar, average of the observed differences
Dbar <- colMeans(Dmat)
# covariance
S_d <- cov(Dmat)
# T2
T2 <- (n - p)*n/(p*(n-1)) * t(Dbar) %*% solve(S_d) %*% Dbar
# critical value
crit <- qf(0.05, p, n-p, lower.tail = F)
# p-value
pv <- pf(c(T2), p, n-p, lower.tail = F)

data.frame(T2 = T2, critical = crit, pvalue = pv)
```

```
##        T2 critical     pvalue
## 1 6.13769 4.256495 0.02082779
```

From the results (p-value less than 0.05), we reject the null hypothesis that the means of the commercial and state lab measurements are same. Recall that we did not catch this difference using individual paired t-tests.

We compute the 95% simultaneous confidence for the mean differences below.

```
h <- sqrt((n-1)*p*crit/(n-p))* sqrt(diag(S_d)/n)
lower = Dbar -  h
upper = Dbar +  h
cbind(lower, upper)
```

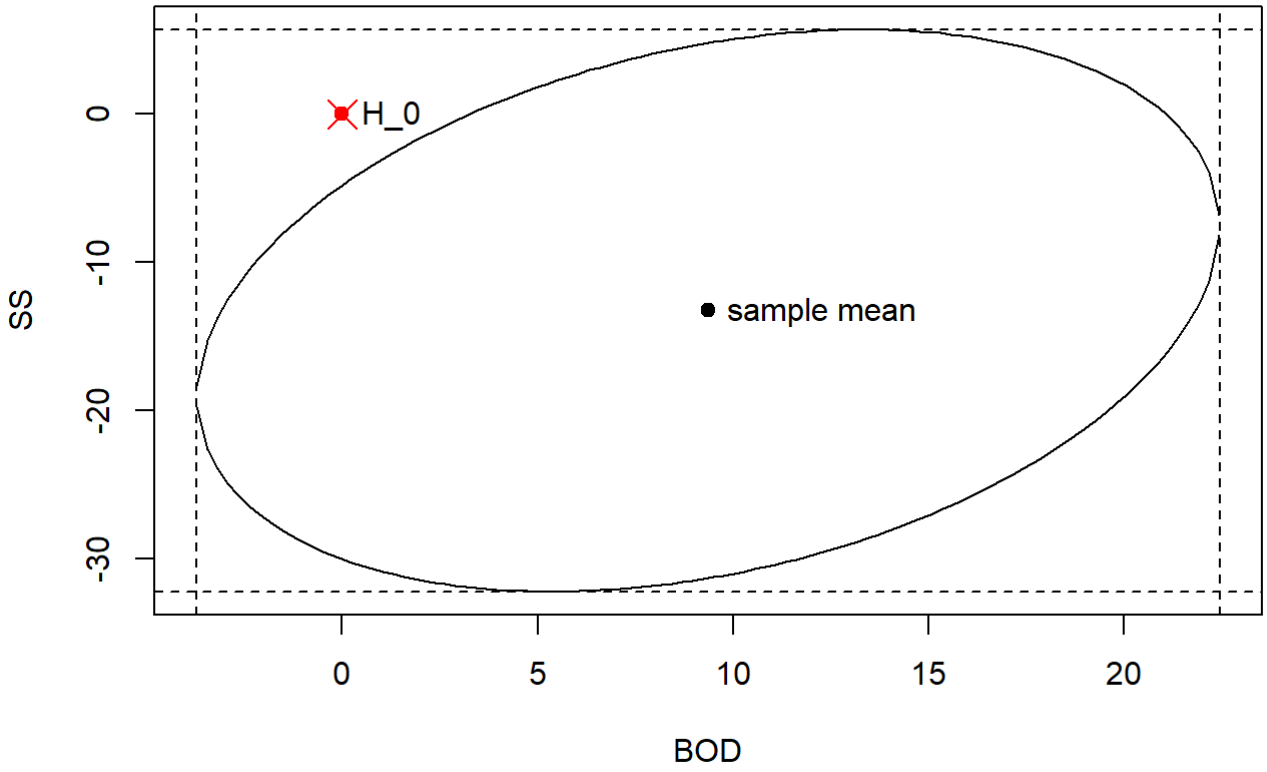
```
##           lower     upper
## D_BOD  -3.72600 22.453272
## D_SS  -32.24557  5.700119
```

Note that the simultaneous confidence intervals are wide, and as such they **should not be used for hypothesis testing**.

## Geometry

The Hotelling's $T^2$ test essentially draws an ellipse (recall data ellipses in the multivariate normal lecture) around the sample mean and checks whether the proposed value (the value proposed in $H_0$) lies in the ellipse or not. We reject $H_0$ if the poposed value does not fall inside the ellipse. This ellipse is often refered to as a *confidence region*.

For our dataset, the ellipse is shown below. We can see that the null value $(0, 0)$ (marked by the red x) falls outside the confidence region. This matches will the result we obtained from the hypothesis test, where we rejected $H_0 : \boldsymbol{\mu}_2 - \boldsymbol{\mu}_1 = 0$.




The vertical and horizontal lines representing the projection of the ellipse onto the x- and y-axis are infact the corresponding simultaneous confidence intervals we constructed before.

It is interesting to see that the simultaneous intervals does contain zero. This is because the simultaneous confidence coefficient applies to the entire set of intervals that could be constructed for all possible linear combinations of the form $a_1\delta_1 + a_2\delta_2$, where $\delta_1$ is the mean difference in `SS` and $\delta_2$ is the mean difference in `BOD`. The particular intervals obtained using ($a_1 = 1, a_2 = 0$) and ($a_1 = 0, a_2 = 1$) contain zero. However, there are other for choices of $a_1$ and $a_2$ that will produe intervals not containing zero.

In other words, if we were to not reject $H_0 : \delta = 0$, then **all** simultaneous intervals (corresponding to all possible choiches of $a_1$ and $a_2$) would contain zero.

To see an example (purely for demonstration, not interpretation), consider the linear combination $1.5\delta_1 - 0.5\delta_2$. A point estimator of this parameter is $1.5\bar{D}_1 - 0.5\bar{D}_2$, and the variance can be estimated by $(1.5, -0.5)\boldsymbol{S}_D(1.5, -0.5)^T/n$. the 95% simultaneous confidence interval for this particular choice of $a_1 = 1.5$ and $a_2 = -0.5$ is shown below.

```
## [1]  1.666067 39.697570
```

Clearly, for this choice of $a_1$ and $a_2$, the interval does not contain zero. If $H_0$ was indeed true, every interval (that is, for all choices for $a_1$ and $a_2$) would contain zero.

# Multiple treatments comparison

Let us consider the case where each sampling unit is subjected to more than two treatments. In this course, we will only consider the case where only one variable is measured for each subject under each treatment.

As an illustrative example, we consider the **[anesthesia dataset] (data/T6-2.dat)** in Table 6.2 in the Johnson and Wichern textbook. The dataset measures anesthetizing effect of $CO_2$ and halothane in dogs.

"Improved anesthetics are often developed by first studying their effects on animals. In one study 19 dogs were initially given the drug pentobarbitol. Each dog was then administered carbon dioxide $CO_2$ at each of two pressure levels. Next, halothane (H) was added and then administration of $CO_2$ was repeated. The response, milliseconds between heartbeats was measured for the four treatment combinations."

The data are displayed in Table 6.2 in the Johnson and Wichern textbook. The treatment combinations are shown below.

| Treatment | $CO_2$ pressure | Halothane |
|---|---|---|
| 1 | high | present |
| 2 | low | present |
| 3 | high | absent |
| 4 | low | absent |

```
dat <- as.matrix( read.table("data/T6-2.dat") )
colnames(dat)  = c("trt 1", "trt 2","trt 3", "trt 4")
head(dat)
```

```
##      trt 1 trt 2 trt 3 trt 4
## [1,]   426   609   556   600
## [2,]   253   236   392   395
## [3,]   359   433   349   357
## [4,]   432   431   522   600
## [5,]   405   426   513   513
## [6,]   324   438   507   539
```

Define $X_{ij}$ = the response for the $i$-th dog under $j$-th treatment. We define

$$\mu_j = \text{mean response under } j\text{-th treatment} = E(X_{ij}).$$

We might be interested in testing the following hypotheses:

1. Null effect/no treatment effect: $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

or equivalently, testing the following jointly:

2. No effect of halothane. $H_0 : (\mu_1 + \mu_2) - (\mu_3 + \mu_4) = 0$
3. No effect of $CO_2$. $H_0 : (\mu_1 + \mu_3) - (\mu_2 + \mu_4) = 0$
4. No interaction of $CO_2$ and halothane. $H_0 : (\mu_1 + \mu_4) - (\mu_2 + \mu_3) = 0$

In general, all the hypotheses presented above can be written as

$$H_0 : \boldsymbol{C}\boldsymbol{\mu} = 0,$$

where $\boldsymbol{C}$ is $q \times 4$ matrix of contrasts. For example, the first hypothesis (1.) of no treatment effect $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ can be writtten as $H_0 : \mu_1 - \mu_2 = 0, \mu_1 - \mu_3 = 0, \mu_1 - \mu_4 = 0$. Thus we can write $H_0$ as

$$\underbrace{\begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}}_{\boldsymbol{C}} \underbrace{\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix}}_{\boldsymbol{\mu}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Similarly, for the hypotheses in (2. – 4.) can be written as

$$\underbrace{\begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}}_{\boldsymbol{C}} \underbrace{\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix}}_{\boldsymbol{\mu}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Let us collect the data for the $i$-th dog (in general subject/item) as $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{i4})^T$. Thus $\boldsymbol{\mu} = E(\boldsymbol{X}_i)$. We know that a point estimator of $\boldsymbol{\mu}$ is $\bar{\boldsymbol{X}}$ and a point estimator of $\boldsymbol{\Sigma} = cov(\boldsymbol{X})$ is the sample covariance matrix $\boldsymbol{S}$. Hence we have

Point estimator of $\boldsymbol{C\mu}$ is $\boldsymbol{C\bar{X}}$

Covariance matrix of $\boldsymbol{C\bar{X}}$ is $\boldsymbol{CSC}^T/n$

***Hotelling's $T^2$ test:*** *Testing $H_0 : \boldsymbol{C\mu} = 0$ is done using the Hotelling's $T^2$ test statistic:*

$$T^2 := \frac{(n-q)n}{(n-1)q}(\boldsymbol{C\bar{X}})^T(\boldsymbol{CSC}^T)^{-1}\boldsymbol{C\bar{X}};$$

*here $\bar{\boldsymbol{X}}$ and $\boldsymbol{S}$ are the sample mean and covariance of $X_i$'s. If the null hypothesis is true then $T^2 \sim F_{q,n-q}$. Thus we reject $H_0$ at level $\alpha$ if the observed value of $T^2$ exceeds the critical value $F_{q,n-q}(\alpha)$.*

**Note: in the formula above, $q$ denotes the number of rows/contrasts in $C$ that are being tested.**

Let us perform the hypothesis testing for the dataset at hand. We start with the eauality of the treatment means $H_0 : \mu_1 = \ldots = \mu_4$. Recall the contrast matrix for this hypothesis is

$$\boldsymbol{C} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}$$

A function for testing multiple contrasts is shown below.

```
T2.contrast <- function(data.matrix, contrast.matrix, alpha = 0.05){
  # Input args
  #   data.matrix: n x p matrix, each row is one subject, each col is one treatment
  #   contrast.matrix: q x p matrix C, each row is one contrast
  #   alpha: significance level, default 0.05

  dat <- data.matrix
  C <- contrast.matrix

  # sample mean vector
  xbar <- colMeans(dat)
  # sample covariance matrix
  S <- cov(dat)

  # parameters
  n <- nrow(dat)
  q <- nrow(C)

  # Intermediate quantities
  invCSC <- solve(C%*%S %*% (t(C)))
  Cxbar <- C %*% xbar
  # test statistic
  T2 <-  n*(n-q)/((n-1)*q) * (t(Cxbar)) %*% invCSC %*% (Cxbar)
  # critical value
  critical_value_F = qf(p = 0.05, df1 = q, df2 = n-q, lower.tail = F)
  # p-value
  pv <- pf(T2, df1 = q, df2 = n-q, lower.tail = F)

  # display the results
  results <- data.frame(T2 = T2, critical = critical_value_F,
                        df1 = q, df2 = n-q, pvalue = pv)

  return(results)
}
```

Let us use this function to test the hypothesis of equality of mean.

```
# contrast matrix for testing equality of means
C <- cbind(c(1,1,1), -diag(1, 3))
C
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    1   -1    0    0
## [2,]    1    0   -1    0
## [3,]    1    0    0   -1
```

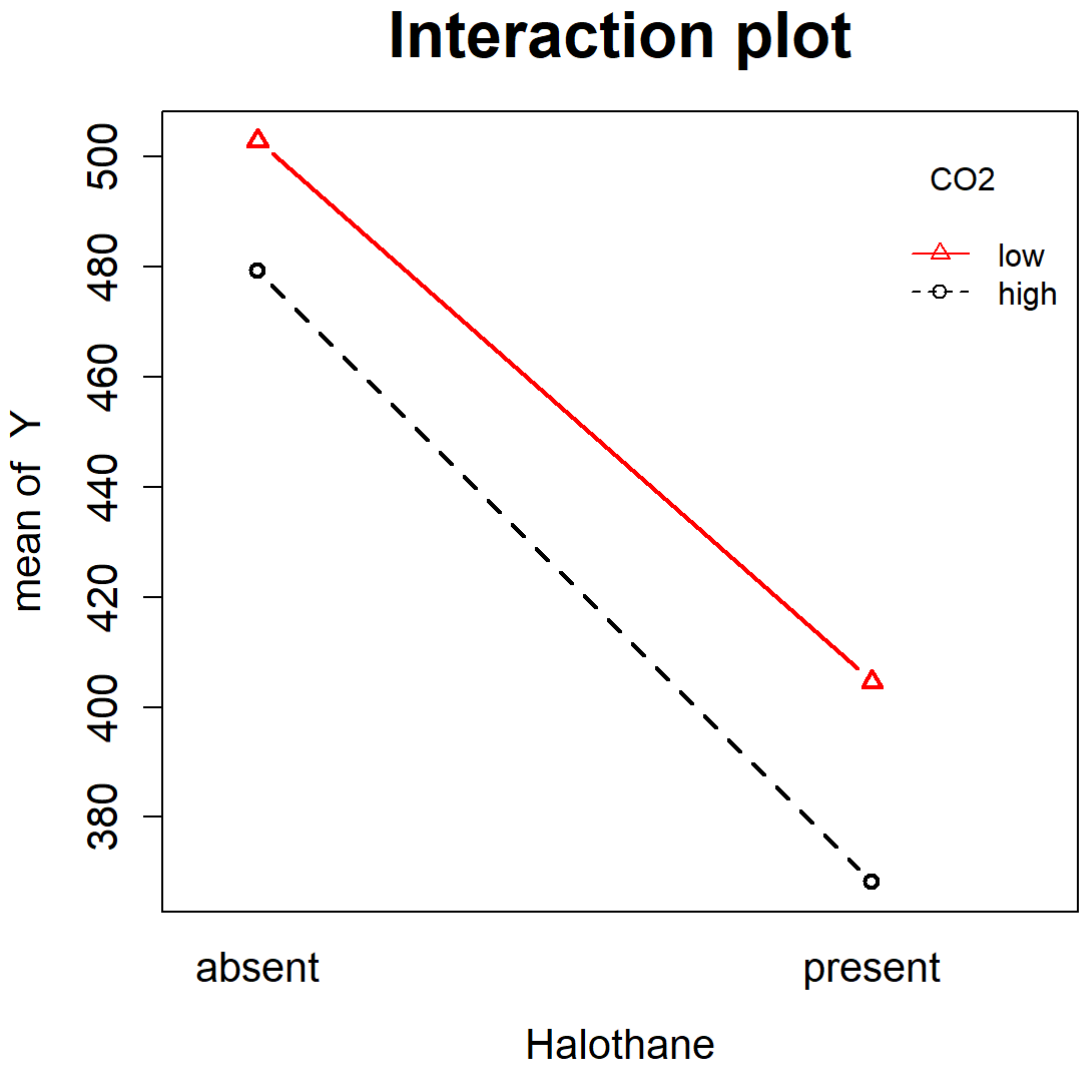
```
# Test
T2.contrast(dat, C)
```

```
##         T2 critical df1 df2       pvalue
## 1 34.37521 3.238872   3  16 3.317767e-07
```

Clearly, there is strong evidence that the four treatment means are not equal.

To investigate further, we can test individual hypotheses 2. – 4. above to identify the intercation and main effects of halothane and CO2. We first start with testing the interaction. Let us look at the interaction plot. We can use the function `interaction.plot()` to do this.

```
# vectorize the data for interaction plot
# this creates a vector by concatinating columns
Y = c(dat)

# Define the treatment combinations
CO2 = c(rep("high", 19), rep("low", 19), rep("high", 19), rep("low", 19))
Halothane = c(rep("present", 19), rep("present", 19), rep("absent", 19), rep("absent", 1
9))

# interaction plot
interaction.plot(Halothane, CO2, Y, lwd=2, col=c(1,2), cex.axis=1.3, cex.lab=1.3, main=
"Interaction plot", cex.main=2, type="b", pch=1:2)
```




Let us fomally test the hypothesis of no interaction, $H_0 : (\mu_1 + \mu_4) - (\mu_2 + \mu_3) = 0$

```
# contrast for no interaction
C <- matrix(c(1, -1, -1, 1), nrow = 1)

# test
T2.contrast(dat, C)
```

```
##          T2 critical df1 df2    pvalue
## 1 0.4112318 4.413873   1  18 0.5294265
```

It is evident that there is no interaction effect. We can similarly test for the main effects of halothane and CO2.

## Simultaneous confidence intervals

We can also create simultaneous intervals for any set of given contrasts $\boldsymbol{c}_1^T\boldsymbol{\mu}, \ldots, \boldsymbol{c}_k^T\boldsymbol{\mu}$ (you can think of $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k$ as rows of $\boldsymbol{C}$).

> *The $100(1-\alpha)\%$ simultaneous confidence intervals} are*
>
> $$\boldsymbol{c}^T\bar{\boldsymbol{x}} \pm \sqrt{\frac{(n-1)q}{n-q}F_{q,n-q}(\alpha)}\sqrt{\frac{\boldsymbol{c}^T\boldsymbol{S}\boldsymbol{c}}{n}}$$
>
> *will contain $\boldsymbol{c}^T\boldsymbol{\mu}$ simultaneously for **all** $\boldsymbol{c}$ with probability $1-\alpha$.*

In our data set, let us create 95% simultaneous confidence intervals for

- effect of halothane: $(\mu_1 + \mu_2) - (\mu_3 + \mu_4)$
- effect of $CO_2$: $(\mu_1 + \mu_3) - (\mu_2 + \mu_4)$
- interaction of $CO_2$ and halothane: $(\mu_1 + \mu_4) - (\mu_2 + \mu_3)$.

As mentioned before, the contrast matrix is

$$\boldsymbol{C} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

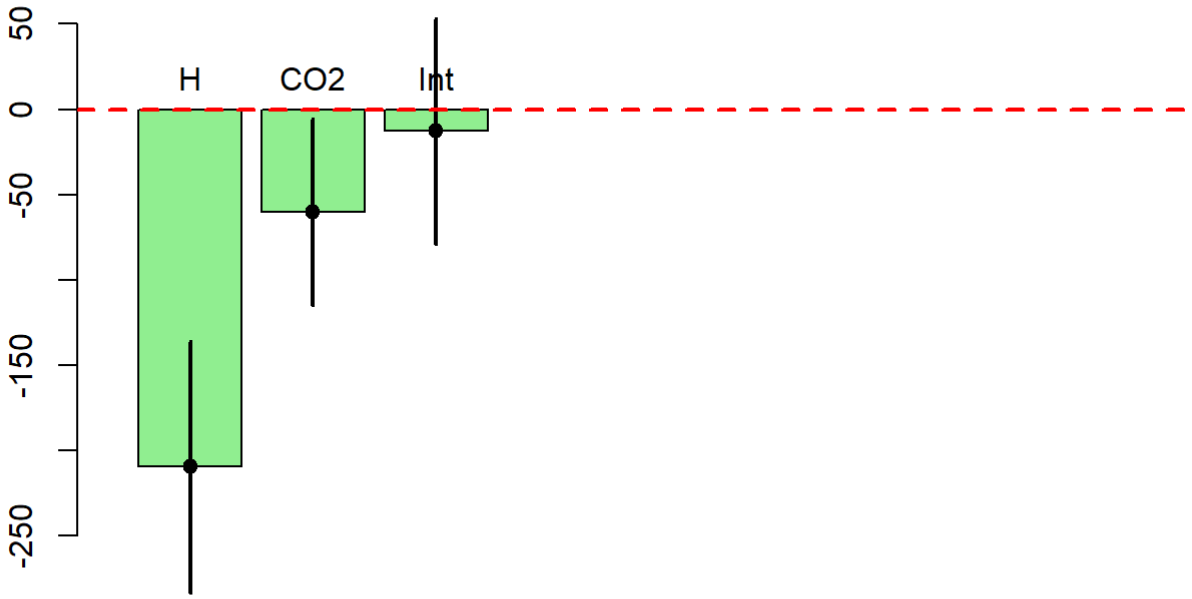
```
# contrast matrix
C <- matrix(c(1, 1, -1, -1, 1, -1, 1, -1, 1, -1, -1, 1), ncol=4,byrow=T)
rownames(C) <- c("Halothane (present - absent)", "CO2 (high - low)", "H-CO2 Interaction"
)
p <- ncol(C) # number of variables
q <- nrow(C) # number of contrasts
n <- nrow(dat) # sample size
C
```

```
##                              [,1] [,2] [,3] [,4]
## Halothane (present - absent)    1    1   -1   -1
## CO2 (high - low)                1   -1    1   -1
## H-CO2 Interaction               1   -1   -1    1
```

```
# sample mean vector
xbar = colMeans(dat)

# sample covariance matrix
S <- cov(dat)

# compute C xbar
mid <- C %*% xbar

# compute (C S C^T)
M <- C %*% S %*% t(C)

# critical value (scaled)
critical_value_F = sqrt(
  ( (n-1)*q/(n-q) ) * qf(p = 0.05, df1 = q, df2 = n-q, lower.tail = F)
)
# std errors
sdvec <- sqrt(diag(M)/n)

# Intervals with point estimates
result <- data.frame(Estimate = mid,
           Lower = mid - critical_value_F*sdvec,
           Upper = mid + critical_value_F*sdvec)
result
```

```
##                                Estimate      Lower      Upper
## Halothane (present - absent) -209.31579 -282.98128 -135.65030
## CO2 (high - low)              -60.05263 -114.72708   -5.37818
## H-CO2 Interaction             -12.78947  -78.72858   53.14964
```

While there is no interaction between halothane and CO2 effects, it seems the main effects of CO2 and Halothane are statistically significant. It seems that absence of halothane is associated with higher mean response, and high level of CO2 with lower mean response.




Main page: **ST 437/537: Applied Multivariate and Longitudinal Data Analysis (https://maityst537.wordpress.ncsu.edu/)**