Multinormal Distribution

- Distribution and density function
  - cdf and pdf of a random vector
  - Conditional distribution
  - Expectation and covariance matrix
  - Density plots
- Multinormal distribution
  - Density function
  - Standardization
  - Transformation
  - Sampling distribution and central limit theorem
  - Some properties
- Testing
  - One population
  - Comparison of two means

cdf and pdf of a random vector

Let $\mathbf{X} = (X_1, X_2, \ldots, X_\mathrm{p})'$ be a random vector

- CDF: $\mathrm{F}(\mathbf{x}) = \mathrm{P}(\mathbf{X} \le \mathbf{x}) = \mathrm{P}(X_1 \le x_1, X_2 \le x_2, \ldots, X_\mathrm{p} \le x_\mathrm{p})$
- Joint pdf: $f(\mathbf{x}) = f(X_1 = x_1, X_2 = x_2, \ldots, X_\mathrm{p} = x_\mathrm{p})$
- Independent $\Rightarrow$ joint pdf $= f(X_1 = x_1, X_2 = x_2, \ldots, X_\mathrm{p} = x_\mathrm{p})$

$$= f(X_1 = x_1)\, f(X_2 = x_2) \cdots f(X_p = x_p)$$
$$= \Pi_i\ f(X_i = x_i)$$

- Marginal cdf
- Marginal pdf

Conditional distribution

For bivariate random vector $(X_1, X_2)'$

- $f(x_2|x_1) = f(x_1,x_2)/f(x_1)$

Example: $f(x_1, x_2) = \begin{cases} \frac{1}{2}x_1 + \frac{3}{2}x_2 & 0 \le x_1, x_2 \le 1, \\ 0 & \text{otherwise} \end{cases}$

- Is $f(x_1, x_2)$ a density function?
- Marginal densities?
- Conditional density $f(x_2|x_1)$?

Expectation and Covariance

- Expectation ($\mathbf{X}$ is a random vector, $\mathbf{C}$ and $\mathbf{D}$ are constant matrix)
  - $E(\mathbf{X}) = ?$
  - $E(\mathbf{CX} + \mathbf{D}) = ?$

- Variance and Covariance ($\mathbf{X}$ and $\mathbf{Y}$ are random vectors, $\mathbf{C}$ and $\mathbf{D}$ are constant matrix)
  - $\mathrm{Var}(\mathbf{X}) = ?$
  - $\mathrm{Var}(\mathbf{CX} + \mathbf{D}) = ?$
  - $\mathrm{Var}(\mathbf{CX} + \mathbf{DY}) = ?$
  - $\mathrm{Cov}(\mathbf{CX}, \mathbf{DY}) = ?$

Density plots

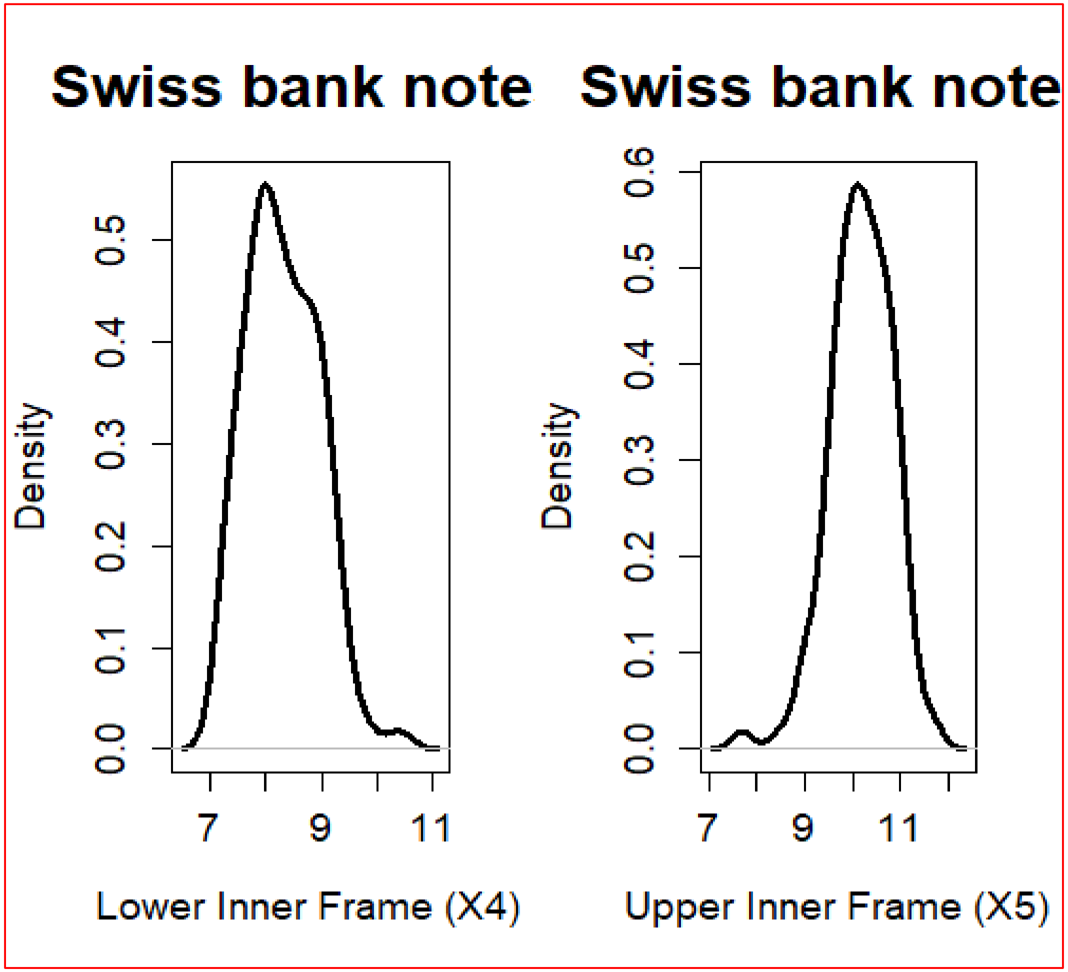
Univariate (Data: Swiss Bank Notes)

- X_1: Length of the bank note
- X_2: Height of the bank note, measured on the left
- X_3: Height of the bank note, measured on the right
- X_4: Distance of inner frame to the lower border
- X_5: Distance of inner frame to the upper border
- X_6: Length of the diagonal

R-code (MVAdenbank2)

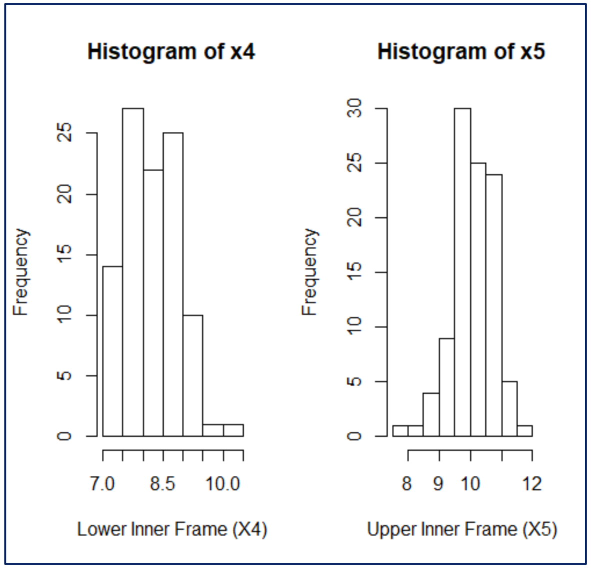
```
# clear variables and close windows
rm(list = ls(all = TRUE))
graphics.off()

# Observation: Density estimates are different because R uses a 'gaussian' kernel
as default, whereas Xplore uses a Quartic Kernel.

# load data
x = read.table("bank2.dat")

x4 = x[1:100, 4]
x5 = x[1:100, 5]

f4 = density(x4)
f5 = density(x5)

# plot
par(mfrow = c(1, 2))
plot(f4, type = "l", lwd = 3, xlab = "Lower Inner Frame (X4)", ylab = "Density",
main = "Swiss bank notes",
    cex.lab = 1.2, cex.axis = 1.2, cex.main = 1.8)
plot(f5, type = "l", lwd = 3, xlab = "Upper Inner Frame (X5)", ylab = "Density",
main = "Swiss bank notes",
    cex.lab = 1.2, cex.axis = 1.2, cex.main = 1.8)
```

Density plots



<u>Histogram</u>

```
xx = read.table("bank2.txt")
x4 = xx[1:100,4]
x5 = xx[1:100,5]
hist(x4,xlab="Lower Inner Frame (X4)")
hist(x5,xlab="Upper Inner Frame (X5)")
```



Density plots

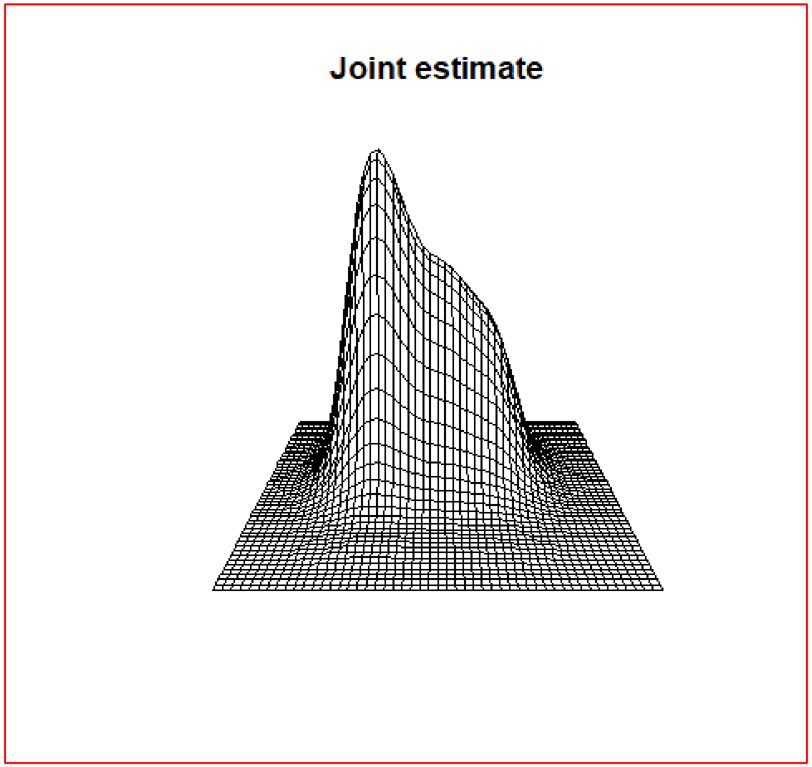
Bivariate (Data: Swiss Bank Notes)

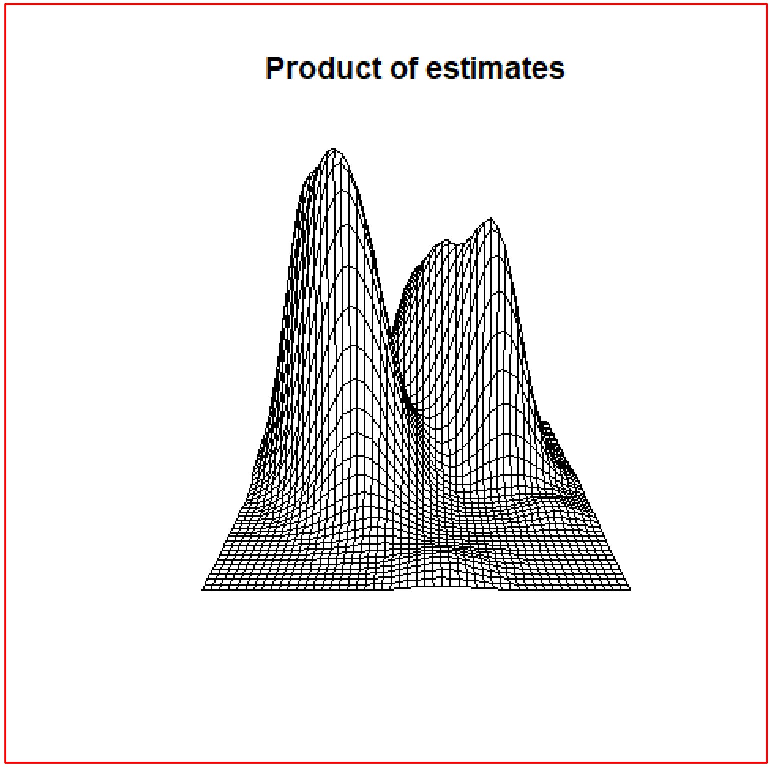
- X_1: Length of the bank note
- X_2: Height of the bank note, measured on the left
- X_3: Height of the bank note, measured on the right
- X_4: Distance of inner frame to the lower border
- X_5: Distance of inner frame to the upper border
- X_6: Length of the diagonal

R-code (MVAdenbank3)

```
# clear variables and close windows
rm(list = ls(all = TRUE))
graphics.off()

# install and load packages
libraries = c("KernSmooth", "graphics")
lapply(libraries, function(x) if (!(x %in% installed.packages())) {
install.packages(x)
})
lapply(libraries, library, quietly = TRUE, character.only = TRUE)

# load data
xx = read.table("bank2.dat")

# Compute a kernel density estimates
dj = bkde2D(xx[, 4:5], bandwidth = 1.06 * c(sd(xx[, 4]), sd(xx[, 5])) * 200^(-
1/5))
d1 = bkde(xx[, 4], gridsize = 51)
d2 = bkde(xx[, 5], gridsize = 51)
dp = (d1$y) %*% t(d2$y)

# plot
persp(d1$x, d2$x, dp, box = FALSE, main = "Joint estimate")
persp(dj$x1, dj$x2, dj$fhat, box = FALSE, main = "Product of estimates")
```

Density plots





Multinormal distribution

Density function

$\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}),\ \boldsymbol{\Sigma} > 0$

$$f(\boldsymbol{x}) = |2\pi\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\}$$

Standardization: $\boldsymbol{Z} = \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{x}-\boldsymbol{\mu}) \sim N(\boldsymbol{0}, \boldsymbol{I})$

Transformation: distribution of $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}$?

Multinormal distribution (Contours of constant density)

Contour ellipses

$\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$, dimension p

Contours of constant density for the p-dimensional normal distribution are ellipsoids defined by $\boldsymbol{x}$ such that

$$(\boldsymbol{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) = d^2$$

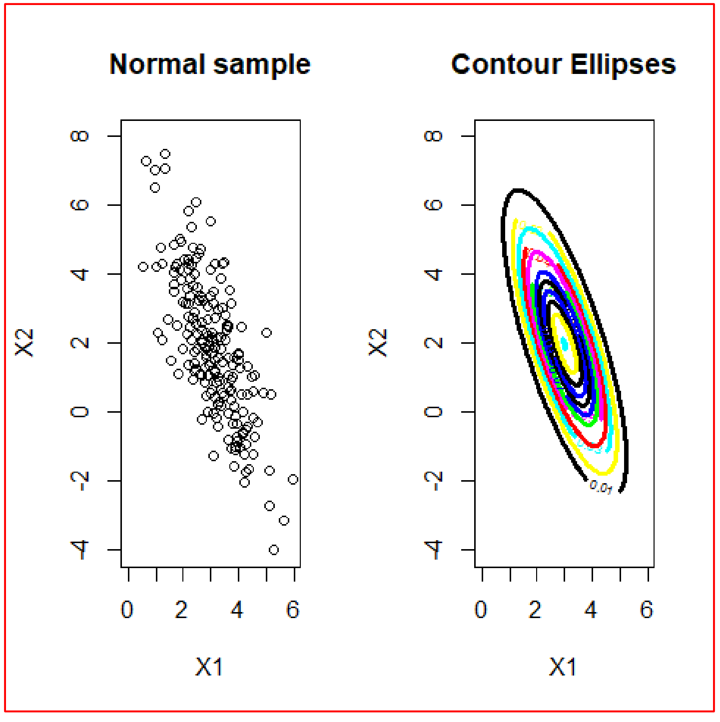
These ellipsoids are centered at $\boldsymbol{\mu}$ and have axes $\pm\sqrt{\lambda_i}\, \boldsymbol{e}_i$, where $\boldsymbol{\Sigma} \boldsymbol{e}_i = \lambda_i \boldsymbol{e}_i$, $i = 1,\ldots,p$.

Multinormal distribution (Contour ellipses: Example)

Scatterplot of a normal sample and contour ellipses for

$$\mu = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \quad \Sigma = \begin{pmatrix} 1 & -1.5 \\ -1.5 & 4 \end{pmatrix}$$

Source: MVAcontnorm




```r
# install and load packages
libraries = c("MASS", "mnormt")
lapply(libraries, function(x) if (!(x %in% installed.packages())) {
install.packages(x)
})
lapply(libraries, library, quietly = TRUE, character.only = TRUE)

# parameter settings
n   = 200  # number of draws
mu  = c(3, 2)  # mean vector
sig = matrix(c(1, -1.5, -1.5, 4), ncol = 2)  # covariance matrix

# bivariate normal sample
set.seed(80)
y = mvrnorm(n, mu, sig, 2)

# bivariate normal density
xgrid = seq(from = (mu[1] - 3 * sqrt(sig[1, 1])), to = (mu[1] + 3 * sqrt(sig[1, 1])),
    length.out = 200)
ygrid = seq(from = (mu[2] - 3 * sqrt(sig[2, 2])), to = (mu[2] + 3 * sqrt(sig[2, 2])),
    length.out = 200)
z     = outer(xgrid, ygrid, FUN = function(xgrid, ygrid) {
    dmnorm(cbind(xgrid, ygrid), mean = mu, varcov = sig)
})

# Plot
par(mfrow = c(1, 2))
plot(y, col = "black", ylab = "X2", xlab = "X1", xlim = range(xgrid), ylim =
range(ygrid))
title("Normal sample")

# Contour ellipses
contour(xgrid, ygrid, z, xlim = range(xgrid), ylim = range(ygrid), nlevels = 10, col =
c("blue",
    "black", "yellow", "cyan", "red", "magenta", "green", "blue", "black"), lwd = 3,
    cex.axis = 1, xlab = "X1", ylab = "X2")
title("Contour Ellipses")
```

Sampling distribution and central limit theorem

Consider an iid sample of n random vectors $\boldsymbol{X}_i \sim (\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Let the sample mean be $\overline{\boldsymbol{X}}$

1. What is $\mathrm{E}(\overline{\boldsymbol{X}})$?
2. What is $\mathrm{Var}(\overline{\boldsymbol{X}})$?
3. What is $\mathrm{E}(\boldsymbol{S})$? [$\boldsymbol{S}$ is the sample covariance as defined in lecture notes 2]

- If $\boldsymbol{X}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\overline{\boldsymbol{X}} \sim N(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma})$

- Central limit theorem

Multinormal Distribution (some properties)

Partition: All subsets of a multivariate normal random vector is normal.

Example: $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{X} = (X_1, X_2, X_3, X_4)'$ and

$$\boldsymbol{\mu} = \begin{pmatrix} 2 \\ -3 \\ 6 \\ -1 \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 4.231 & 1 & 0.1 & 0.3 \\ 1 & 5 & 0.51 & 1.2 \\ 0.1 & 0.51 & 6.3 & 0 \\ 0.3 & 1.2 & 0 & 3 \end{pmatrix}$$

a) What is the distribution of $(X_1, X_3, X_4)'$?

b) What is the distribution of $(X_2, X_3)'$?

Multinormal Distribution (some properties)

Independent

If $\begin{pmatrix}\boldsymbol{X}_1\\ \boldsymbol{X}_2\end{pmatrix} \sim N\begin{bmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{bmatrix}$, then $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are independent if and only if $\boldsymbol{\Sigma}_{12} = \mathbf{0}$ and $\boldsymbol{\Sigma}_{21} = \mathbf{0}$.

Example: $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{X} = (X_1, X_2, X_3, X_4)'$ and

$$\boldsymbol{\Sigma} = \begin{pmatrix} 4.231 & 0 & 0 & 0.3 \\ 0 & 5 & 0.5 & 0 \\ 0 & 0.5 & 6.3 & 0 \\ 0.3 & 0 & 0 & 3 \end{pmatrix}$$

Are $(X_1, X_4)'$ and $(X_2, X_3)'$ independent?

Multinormal Distribution (some properties)

Conditional Distribution

If $\begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix} \sim N \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$, then conditional distribution of $\boldsymbol{X}_1$ given $\boldsymbol{X}_2 = \boldsymbol{x}_2$ is

$$N[\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \quad \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}]$$

Multinormal Distribution (Testing: Hotelling's T^2)

<u>One sample Mean Vector inference</u>

Null hypothesis: H$_0$: $\boldsymbol{\mu} = \boldsymbol{\mu}_0$
Alternative hypothesis: H$_1$: $\boldsymbol{\mu} \neq \boldsymbol{\mu}_0$

Let p be the dimension of the mean vector

Test Statistics: $T^2 = n(\overline{\boldsymbol{X}} - \boldsymbol{\mu}_0)'\boldsymbol{S}^{-1}(\overline{\boldsymbol{X}} - \boldsymbol{\mu}_0)$

Rejection Rule: reject the null hypothesis at level α if $T^2 > \frac{(n-1)p}{n-p} F_{p,n-p,\alpha}$

Remark: What is the confidence ellipsoid of $\boldsymbol{\mu}$?

Multinormal Distribution (Testing: Hotelling's T^2)

Example (Data: Swiss Bank Notes (number of observations = 200))

Use the first 100 case to find the mean and assume that to be $\boldsymbol{\mu}_0$

```
> xx = read.table("bank2.txt")
> top = xx[1:100,]
> bottom = xx[101:200]
> colMeans(top)
     V1      V2       V3      V4      V5      V6
214.969 129.943 129.720   8.305  10.168 141.517
> mu0 =  colMeans(top)
> # install package ICSNP
> library(ICSNP)
> HotellingsT2(bottom, mu = mu0)

        Hotelling's one sample T2-test

data:  bottom
T.2 = 1153.429, df1 = 6, df2 = 94, p-value < 2.2e-16
alternative hypothesis: true location is not equal to c(214.969,129.943,129.72,8.305,10.168,141.517)
```

Multinormal Distribution (Testing: Hotelling's T^2)

Comparison of two population means

Null hypothesis: H$_0$: $\boldsymbol{\mu_1} = \boldsymbol{\mu_2}$
Alternative hypothesis: H$_1$: $\boldsymbol{\mu_1} \neq \boldsymbol{\mu_2}$

Test Statistics: $T^2 = n(\overline{\boldsymbol{X}}_1 - \overline{\boldsymbol{X}}_2)'[\boldsymbol{S_p}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)]^{-1}(\overline{\boldsymbol{X}}_1 - \overline{\boldsymbol{X}}_2)$

Rejection Rule: reject the null hypothesis at level α if $T^2 > \frac{(n-1)p}{n-p} F_{p,n_1+n_2-p-1,\alpha}$

Multinormal Distribution (Testing: Hotelling's T^2)

Example (Data: Swiss Bank Notes (number of observations = 200))

Top (100) mean = bottom (100) mean?

```
> HotellingsT2(top, bottom)

Hotelling's two sample T2-test

data:  top and bottom
T.2 = 391.9217, df1 = 6, df2 = 193, p-value < 2.2e-16
alternative hypothesis: true location difference is not equal to c(0,0,0,0,0,0)
```