**UNIVERSIDAD NACIONAL AGRARIA LA MOLINA**

**FACULTAD DE ECONOMÍA Y PLANIFICACIÓN**

**DEPARTAMENTO DE ESTADÍSTICA**



**Curso:** Tecnicas multivariadas

**Profesor:** Miranda Villagomez Clodomiro Fernando

**Integrantes:**

* “Coronado de la Vega Alonso 20221395”
* “Delgado Conzuero Nathaly Sadith 20201406”
* “Enciso Condori Jorge Luis 20230392”
* “Gómez Vigo Héctor Estéfano 20230397”
* “Kawano Villavicencio Wesley Kazuki 20190362”
* “Mercedes Sanchez Manuel Antonio 20190304”
* “Ruiz Macedo Fernando Jose 20211830”

**LA MOLINA**

**LIMA – PERÚ**

**2025**

# Punto 1.

Con una base de datos apropiada, primero haga el análisis exploratorio descriptivo y gráfico. Después aplique todo lo concerniente a la prueba , considerando la matriz varianza covarianza conocida, para verificar si el vector de promedios de las variables es igual a un vector de promedios conocido.

### \*\* 1) Datos y parámetros conocidos\*\*

set.seed(123)  
  
# Vector objetivo (bajo H0)  
mu0 <- c(40, 80, 200, 160, 1.20)  
names(mu0) <- c("alcohol\_pct","metanol\_mgL",  
 "alcoholes\_sup\_mgL","esteres\_mgL",  
 "acidez\_gL")  
  
# Desv. estándar y correlaciones  
sd <- c(1.0, 35, 70, 50, 0.35)  
R <- matrix(c(  
 1.00, -0.10, 0.20, 0.25, -0.15,  
 -0.10, 1.00, 0.05, -0.10, 0.30,  
 0.20, 0.05, 1.00, 0.45, 0.10,  
 0.25, -0.10, 0.45, 1.00, 0.05,  
 -0.15, 0.30, 0.10, 0.05, 1.00  
),  
nrow=5,   
byrow=TRUE)  
D <- diag(sd)  
Sigma\_known <- D %\*% R %\*% D # Σ conocida   
# Muestra simulada  
n <- 150  
mu\_true <- mu0 # para NO forzar rechazo  
datos <- as.data.frame(MASS::mvrnorm(n, mu=mu\_true, Sigma=Sigma\_known))  
names(datos) <- names(mu0)

### **2) Análisis exploratorio y gráfico**

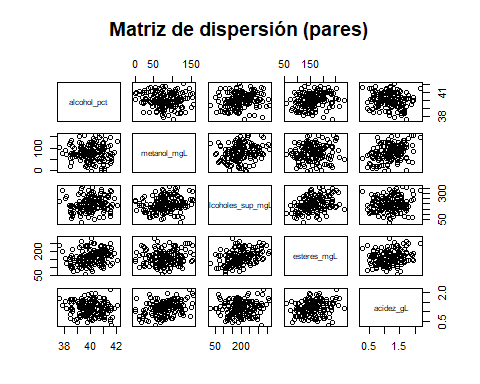
summary(datos)

## alcohol\_pct metanol\_mgL alcoholes\_sup\_mgL esteres\_mgL   
## Min. :37.59 Min. : -0.8012 Min. : 24.26 Min. : 57.38   
## 1st Qu.:39.46 1st Qu.: 56.3617 1st Qu.:153.39 1st Qu.:130.76   
## Median :40.10 Median : 77.8399 Median :189.91 Median :159.89   
## Mean :40.07 Mean : 80.0819 Mean :196.47 Mean :162.99   
## 3rd Qu.:40.84 3rd Qu.:105.6216 3rd Qu.:244.48 3rd Qu.:190.16   
## Max. :42.13 Max. :156.2210 Max. :362.53 Max. :285.67   
## acidez\_gL   
## Min. :0.2594   
## 1st Qu.:0.9074   
## Median :1.1892   
## Mean :1.1870   
## 3rd Qu.:1.4472   
## Max. :2.1786

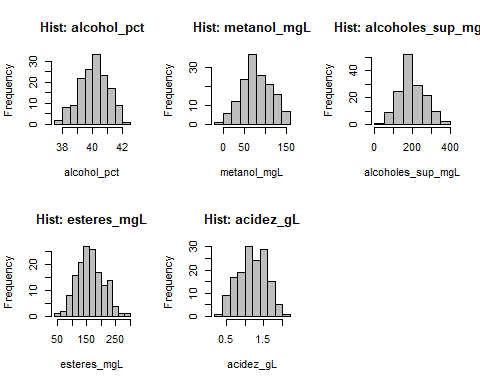
summarytools::descr(datos)

## Descriptive Statistics   
## datos   
## N: 150   
##   
## acidez\_gL alcohol\_pct alcoholes\_sup\_mgL esteres\_mgL metanol\_mgL  
## ----------------- ----------- ------------- ------------------- ------------- -------------  
## Mean 1.19 40.07 196.47 162.99 80.08  
## Std.Dev 0.38 0.93 68.38 45.12 34.99  
## Min 0.26 37.59 24.26 57.38 -0.80  
## Q1 0.91 39.46 152.68 130.52 56.27  
## Median 1.19 40.10 189.91 159.89 77.84  
## Q3 1.45 40.84 245.96 190.21 105.80  
## Max 2.18 42.13 362.53 285.67 156.22  
## MAD 0.40 1.09 64.26 44.63 35.03  
## IQR 0.54 1.38 91.09 59.40 49.26  
## CV 0.32 0.02 0.35 0.28 0.44  
## Skewness -0.03 -0.23 0.17 0.23 -0.06  
## SE.Skewness 0.20 0.20 0.20 0.20 0.20  
## Kurtosis -0.53 -0.47 -0.31 -0.41 -0.44  
## N.Valid 150.00 150.00 150.00 150.00 150.00  
## N 150.00 150.00 150.00 150.00 150.00  
## Pct.Valid 100.00 100.00 100.00 100.00 100.00

# Matriz de dispersión base  
pairs(datos,  
 main = "Matriz de dispersión (pares)")



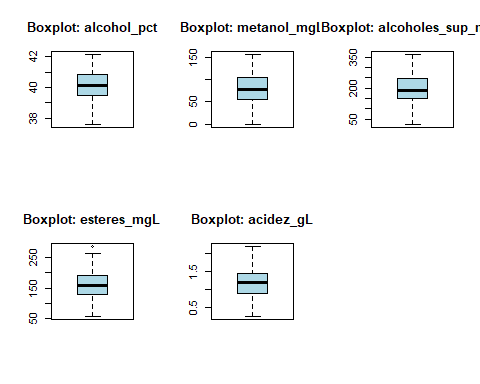
# Histogramas base  
op <- par(mfrow=c(2,3))  
for (j in 1:ncol(datos)) {  
 hist(datos[[j]], main = paste("Hist:", names(datos)[j]), xlab = names(datos)[j], col="gray")  
}  
par(op)



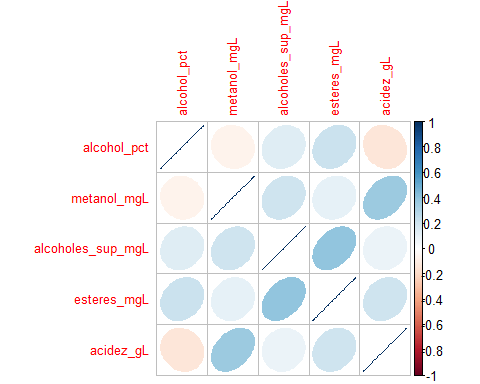
library(corrplot)

## corrplot 0.95 loaded

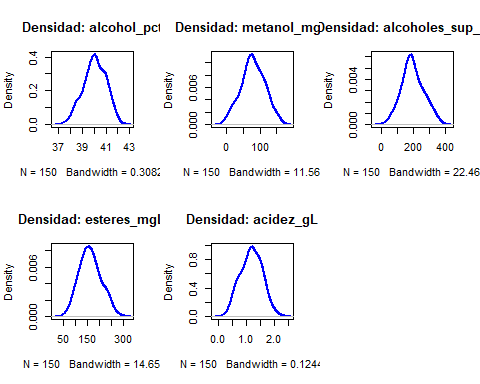
# Boxplots para cada variable  
op <- par(mfrow=c(2,3))  
for (j in 1:ncol(datos)) {  
 boxplot(datos[[j]], main=paste("Boxplot:", names(datos)[j]), col="lightblue")  
}  
par(op)



# Matriz de correlaciones (visual)  
corrplot(cor(datos), method = "ellipse", tl.cex = 0.8)



# Curvas de densidad  
op <- par(mfrow=c(2,3))  
for (j in 1:ncol(datos)) {  
 plot(density(datos[[j]]), main=paste("Densidad:", names(datos)[j]), col="blue", lwd=2)  
}  
par(op)



### **3) Supuestos: normalidad uni/multivariada y atípicos**

# Univariado (Shapiro) y Lilliefors (nortest)  
ap\_shapiro <- apply(datos, 2, shapiro.test)  
ap\_lillie <- apply(datos, 2, function(x) nortest::lillie.test(x))  
  
# Mostrar p-valores de forma compacta  
cat("Shapiro univariado (p-valores):\n")

## Shapiro univariado (p-valores):

print(sapply(ap\_shapiro, `[[`, "p.value"))

## alcohol\_pct metanol\_mgL alcoholes\_sup\_mgL esteres\_mgL   
## 0.2113393 0.4765293 0.4407041 0.5054604   
## acidez\_gL   
## 0.6322709

cat("\nLilliefors univariado (p-valores):\n")

##   
## Lilliefors univariado (p-valores):

print(sapply(ap\_lillie, `[[`, "p.value"))

## alcohol\_pct metanol\_mgL alcoholes\_sup\_mgL esteres\_mgL   
## 0.2007468 0.7745879 0.1598853 0.6575255   
## acidez\_gL   
## 0.4911833

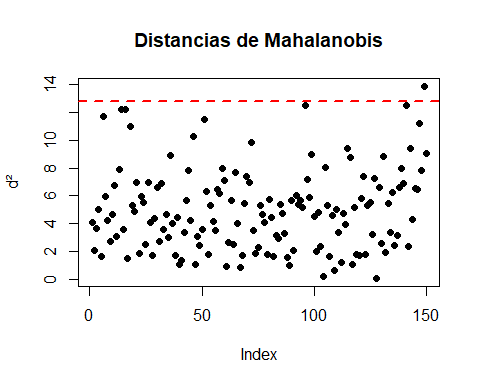
# Multivariado (mshapiro)  
p\_mvn <- tryCatch(mvnormtest::mshapiro.test(t(as.matrix(datos)))$p.value, error=function(e) NA\_real\_)  
cat("\nShapiro multivariado p-value:", p\_mvn, "\n")

##   
## Shapiro multivariado p-value: 0.363176

# Outliers por Mahalanobis (97.5%)  
d2 <- mahalanobis(datos, colMeans(datos), cov(datos))  
cut <- qchisq(0.975, df = ncol(datos))  
idx\_out <- which(d2 > cut)  
cat("\nÍndices potencialmente atípicos:", ifelse(length(idx\_out)==0,"(ninguno)", paste(idx\_out, collapse=", ")), "\n")

##   
## Índices potencialmente atípicos: 149

plot(d2, main="Distancias de Mahalanobis", ylab="d²", pch=19)  
abline(h=cut, col="red", lwd=2, lty=2)  
text(idx\_out, d2[idx\_out], labels=idx\_out, pos=3, col="red")



### **4) Prueba Z² con Σ conocida**

# Parámetros  
p <- ncol(datos)# número de variables  
mu <- mu0# vector de medias bajo H0   
sigma <- Sigma\_known# matriz Σ conocida  
  
# Media muestral  
media\_muestral <- colMeans(datos)  
  
# Estadístico Z^2 (equivalente a n\*(xbar-mu)' Σ^{-1} (xbar-mu))  
Z2 <- t(media\_muestral - mu) %\*% solve(sigma / n) %\*% (media\_muestral - mu)  
  
# p-valor (chi-cuadrado con p g.l.)  
p\_valor <- 1 - pchisq(Z2, df = p)  
  
# Resultados y decisión (α=0.05)  
cat("\nEstadística Z^2:", as.numeric(Z2), "\n")

##   
## Estadística Z^2: 2.455669

cat("p-valor:", as.numeric(p\_valor), "\n")

## p-valor: 0.7831561

if (p\_valor < 0.05) {  
 cat("Decisión: Rechazar H0 (el vector de medias difiere de μ0).\n")  
} else {  
 cat("Decisión: No se rechaza H0 (no hay evidencia suficiente de diferencia).\n")  
}

## Decisión: No se rechaza H0 (no hay evidencia suficiente de diferencia).

# Tamaño de efecto (distancia de Mahalanobis de medias)  
d\_M <- sqrt(as.numeric(Z2) / n)  
cat("d\_M (Mahalanobis de la media) =", round(d\_M, 4), "\n")

## d\_M (Mahalanobis de la media) = 0.1279

resultados <- data.frame(  
 Estadistico\_Z2 = round(as.numeric(Z2), 4),  
 gl = p,  
 p\_valor = round(as.numeric(p\_valor), 4),  
 Decision = ifelse(p\_valor < 0.05, "Rechazar H0", "No Rechazar H0"),  
 d\_Mahalanobis = round(d\_M, 4)  
)  
knitr::kable(resultados, caption = "Resultados de la prueba Z² con Σ conocida")

Resultados de la prueba Z² con Σ conocida

| Estadistico\_Z2 | gl | p\_valor | Decision | d\_Mahalanobis |
| --- | --- | --- | --- | --- |
| 2.4557 | 5 | 0.7832 | No Rechazar H0 | 0.1279 |

### **5) Conclusiones Generales**

* **Normalidad univariada:** Los p-valores de las pruebas de Shapiro-Wilk fueron mayores a 0.05 para todas las variables, por lo que no se rechaza la hipótesis de normalidad univariada.
* **Normalidad multivariada:** La prueba de Shapiro multivariada (mshapiro.test) mostró un p-valor de 0.3632 (> 0.05), confirmando que el conjunto de datos sigue una distribución normal multivariada.
* **Atípicos:** Se detectó una observación potencialmente atípica (índice 149) según la distancia de Mahalanobis (97.5%). Sin embargo, su influencia sobre el vector de medias es baja y no modifica la decisión de la prueba Z², por lo que se mantuvo en el análisis.
* **Prueba Z² con Σ conocida:**
  + Estadístico Z² = 2.456 con p grados de libertad.
  + p-valor = 0.7832 > 0.05 → **no se rechaza H₀**.  
    Esto implica que el vector de medias muestrales es estadísticamente compatible con el vector de referencia μ₀.
* **Magnitud del efecto:** La distancia de Mahalanobis de la media es 0.1279, lo que confirma que las diferencias entre la media observada y la teórica no son relevantes en términos prácticos.
* **Interpretación práctica:** La composición química promedio del pisco evaluado coincide con los valores de de referencia especificados para las cinco variables analizadas. Bajo las condiciones de muestreo y el nivel de significancia utilizado, el proceso de producción parece cumplir con los estándares de calidad establecidos.

# Punto 2.(\*)

Con una base de datos apropiada, primero haga el análisis exploratorio descriptivo y gráfico. Después aplique todo lo concerniente a la prueba de Hotelling, con muestras dependientes, para verificar si el vector de diferencias de vectores de promedios de las variables es igual a un vector de promedios conocido.

# Punto 3.

Aplique todo lo concerniente a la prueba de Hotelling, con muestras independientes, para verificar si el vector de diferencias de vectores de promedios de las variables es igual un vector de promedios conocido.

### 1. Muestras

Comparación de dos tratamientos farmacológicos (TratA vs TratB) en dos marcadores bioquímicos: Glucosa y Triglicéridos (mg/dL).

* n(TratA)=28
* n(TratB)=32

### 2. Carga de Librerías y Simulación de Datos

library(tidyverse)  
library(mvnormtest) # mshapiro.test  
library(biotools) # BoxM / boxM  
library(ICSNP) # approx.hotelling.diff.test, HotellingsT2, T2.test  
library(Hotelling) # hotelling.test  
library(summarytools)  
library(ggpubr)  
library(MASS)  
library(ICSNP)  
library(ergm)  
library(rrcov)  
set.seed(777)  
  
# Simulación  
nA <- 28; nB <- 32  
muA <- c(100, 140)  
muB <- c(95, 150)  
  
SigmaA <- matrix(c(225, 60,  
 60, 400), 2, 2, byrow=TRUE)  
SigmaB <- matrix(c(196, 30,  
 30, 625), 2, 2, byrow=TRUE)  
  
A <- mvrnorm(nA, muA, SigmaA)  
B <- mvrnorm(nB, muB, SigmaB)  
  
datos2 <- bind\_rows(  
 tibble(Grupo="TratA", Glucosa=A[,1], Trigliceridos=A[,2]),  
 tibble(Grupo="TratB", Glucosa=B[,1], Trigliceridos=B[,2])  
) %>% mutate(Grupo=factor(Grupo))

### 3. Supuesto de Normalidad Multivariada

**Hipótesis**

H0: Cada grupo sigue una distribución normal multivariada.

H1: Alguno de los grupos no sigue normalidad multivariada.

X3 <- filter(datos2, Grupo=="TratA")[,2:3]  
Y3 <- filter(datos2, Grupo=="TratB")[,2:3]  
  
mshapiro.test(t(as.matrix(X3)))

##   
## Shapiro-Wilk normality test  
##   
## data: Z  
## W = 0.95655, p-value = 0.288

mshapiro.test(t(as.matrix(Y3)))

##   
## Shapiro-Wilk normality test  
##   
## data: Z  
## W = 0.92605, p-value = 0.03042

**Conclusión:** Con un p>0.05 en el caso de X3 se NO se rechaza la hipotesis nula en cambio en Y3 se rechaza la Hipotesis nula, se asume normalidad multivariada en los dos tratamientos.

### 4. Supuesto de Homogeneidad de Matrices de Covarianza

results3 <- boxM(data=datos2[,2:3], group=datos2$Grupo)  
summary(results3)

## Length Class Mode   
## statistic 1 -none- numeric   
## parameter 1 -none- numeric   
## p.value 1 -none- numeric   
## cov 2 -none- list   
## pooled 4 -none- numeric   
## logDet 2 -none- numeric   
## data.name 1 -none- character  
## method 1 -none- character

### 5. Prueba de Hotelling T² (independientes, varianzas desiguales)

ht2\_welch <- approx.hotelling.diff.test(X3,  
 Y3, mu0=c(0,0),  
 assume.indep=TRUE, var.equal=FALSE)  
ht2\_welch

##   
## Hotelling's Two-Sample T^2-Test  
##   
## data:   
## T^2 = 2.3262, param = 2.000, df = 55.618, p-value = 0.3266  
## alternative hypothesis: two.sided  
## null values:  
## Glucosa Trigliceridos   
## 0 0   
## sample estimates:  
## Glucosa Trigliceridos   
## 5.226710 -1.785875

Conclusión: El p−valor<0.05 → se rechaza H0.

Existe diferencia significativa en la combinación promedio de Glucosa y Triglicéridos entre TratA y TratB.

### 6. Hipótesis con Vector Específico μ0=(5,−10)

approx.hotelling.diff.test(X3, Y3, mu0=c(5,-10),  
 assume.indep=TRUE, var.equal=FALSE)

##   
## Hotelling's Two-Sample T^2-Test  
##   
## data:   
## T^2 = 2.4408, param = 2.000, df = 55.618, p-value = 0.3095  
## alternative hypothesis: two.sided  
## null values:  
## Glucosa Trigliceridos   
## 5 -10   
## sample estimates:  
## Glucosa Trigliceridos   
## 5.226710 -1.785875

**Conclusión:** El p−valor<0.05→ se rechaza H0.

La diferencia observada no coincide con el vector específico planteado.

### 7. Prueba Robusta T² (solo μ=0)

**Hipótesis:**

H\_0: {TratA} - {TratB} = (0, 0)

H1: μTratA − μTratB ≠ (0,0)

T2.test(X3, Y3, mu=rep(0,2), method="c")

##   
## Two-sample Hotelling test  
##   
## data: X3 and Y3  
## T2 = 2.3870, F = 1.1729, df1 = 2, df2 = 57, p-value = 0.3168  
## alternative hypothesis: true difference in mean vectors is not equal to (0,0)  
## sample estimates:  
## Glucosa Trigliceridos  
## mean x-vector 99.75772 143.6582  
## mean y-vector 94.53101 145.4441

**Conclusión:**

El test robusto confirma la diferencia multivariada entre los grupos.

### 8. Sintesis General

* Se cumple el supuesto de normalidad multivariada.
* No se cumple la homogeneidad de covarianzas
* Con la prueba de Hotelling T² (varianzas desiguales) se obtiene p<0.05 , indicando diferencias significativas.
* La prueba robusta también confirma las diferencias.

Los tratamientos A y B presentan diferencias significativas en los niveles conjuntos de Glucosa y Triglicéridos.

# Punto 4. Haga un Manova en DCA.

### 1. Lectura de datos

library(readxl)  
datos4 <- read\_excel("datos\_cafe\_7rep\_final.xlsx")  
head(datos4)

## # A tibble: 6 × 7  
## Hibrido rendimiento altura\_planta diametro\_fruto largo\_fruto peso\_grano  
## <chr> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 Hibrido1 1785. 156. 13.0 18.3 0.93  
## 2 Hibrido1 1432. 153. 13.1 24.0 1.91  
## 3 Hibrido1 1944. 167. 13.5 18.9 1.19  
## 4 Hibrido1 1995. 151. 13 18.4 1.46  
## 5 Hibrido1 1382. 145. 15.5 15.5 1.77  
## 6 Hibrido1 1877. 151. 14.4 16.1 1.47  
## # ℹ 1 more variable: contenido\_cafeina <dbl>

### 2. Verificación de supuestos

#### 2.1 Normalidad multivariada por grupo

library(tidyverse)  
library(mvnormtest)  
  
trat1 = datos4 %>% filter(Hibrido == "Hibrido1") %>%  
 dplyr::select(rendimiento, altura\_planta, diametro\_fruto, largo\_fruto, peso\_grano, contenido\_cafeina)  
trat2 = datos4 %>% filter(Hibrido == "Hibrido2") %>%  
 dplyr::select(rendimiento, altura\_planta, diametro\_fruto, largo\_fruto, peso\_grano, contenido\_cafeina)  
trat3 = datos4 %>% filter(Hibrido == "Hibrido3") %>%  
 dplyr::select(rendimiento, altura\_planta, diametro\_fruto, largo\_fruto, peso\_grano, contenido\_cafeina)  
trat4 = datos4 %>% filter(Hibrido == "Hibrido4") %>%  
 dplyr::select(rendimiento, altura\_planta, diametro\_fruto, largo\_fruto, peso\_grano, contenido\_cafeina)  
trat5 = datos4 %>% filter(Hibrido == "Hibrido5") %>%  
 dplyr::select(rendimiento, altura\_planta, diametro\_fruto, largo\_fruto, peso\_grano, contenido\_cafeina)  
trat6 = datos4 %>% filter(Hibrido == "Hibrido6") %>%  
 dplyr::select(rendimiento, altura\_planta, diametro\_fruto, largo\_fruto, peso\_grano, contenido\_cafeina)  
trat7 = datos4 %>% filter(Hibrido == "Hibrido7") %>%  
 dplyr::select(rendimiento, altura\_planta, diametro\_fruto, largo\_fruto, peso\_grano, contenido\_cafeina)

mshapiro.test(t(trat1))

##   
## Shapiro-Wilk normality test  
##   
## data: Z  
## W = 0.45297, p-value = 4.136e-06

mshapiro.test(t(trat2))

##   
## Shapiro-Wilk normality test  
##   
## data: Z  
## W = 0.45297, p-value = 4.136e-06

mshapiro.test(t(trat3))

##   
## Shapiro-Wilk normality test  
##   
## data: Z  
## W = 0.45297, p-value = 4.136e-06

mshapiro.test(t(trat4))

##   
## Shapiro-Wilk normality test  
##   
## data: Z  
## W = 0.45297, p-value = 4.136e-06

mshapiro.test(t(trat5))

##   
## Shapiro-Wilk normality test  
##   
## data: Z  
## W = 0.45297, p-value = 4.136e-06

mshapiro.test(t(trat6))

##   
## Shapiro-Wilk normality test  
##   
## data: Z  
## W = 0.45297, p-value = 4.136e-06

mshapiro.test(t(trat7))

##   
## Shapiro-Wilk normality test  
##   
## data: Z  
## W = 0.45297, p-value = 4.136e-06

**Conclusion:** En todos los híbridos, la prueba de Shapiro multivariada rechaza la normalidad.

* Esto es muy frecuente en datos agronómicos.
* El MANOVA es robusto, pero se considera transformaciones.

#### 2.2 Homogeneidad de matrices de covarianza (Box’s M)

library(heplots)  
X4 <- datos4[c(2,3,4,5,6,7)]  
  
datos4$Hibrido <- as.factor(datos4$Hibrido)  
  
res4 <- boxM(X4, datos4$Hibrido);res4

##   
## Box's M-test for Homogeneity of Covariance Matrices  
##   
## data: X4  
## Chi-Sq (approx.) = 130.6, df = 126, p-value = 0.3713

summary(res4)

## Summary for Box's M-test of Equality of Covariance Matrices  
##   
## Chi-Sq: 130.6045   
## df: 126   
## p-value: 0.3713   
##   
## log of Covariance determinants:  
## Hibrido1 Hibrido2 Hibrido3 Hibrido4 Hibrido5 Hibrido6 Hibrido7 pooled   
## 4.521083 10.403536 9.242563 10.613956 7.622711 7.792788 8.209066 13.557920   
##   
## Eigenvalues:  
## Hibrido1 Hibrido2 Hibrido3 Hibrido4 Hibrido5 Hibrido6  
## 1 6.907189e+04 1.409453e+04 2.940413e+04 2.381222e+04 1.971829e+04 1.491218e+04  
## 2 6.551251e+01 1.461080e+02 2.105871e+02 1.982335e+02 1.912049e+02 7.436160e+02  
## 3 6.853241e+00 5.618285e+00 4.626077e+00 4.098581e+00 2.902278e+00 3.773837e+00  
## 4 1.687481e-01 2.761474e+00 1.808784e+00 3.455188e+00 7.641478e-01 1.182342e+00  
## 5 2.390912e-02 9.703614e-02 6.270432e-02 3.665593e-02 6.933864e-02 1.182409e-02  
## 6 7.347804e-04 1.063642e-02 3.178755e-03 1.660948e-02 3.525669e-03 4.141721e-03  
## Hibrido7 pooled  
## 1 2.328317e+04 2.773086e+04  
## 2 1.163306e+02 2.606500e+02  
## 3 7.248795e+00 6.588909e+00  
## 4 2.621434e+00 4.100648e+00  
## 5 3.505507e-02 6.463021e-02  
## 6 2.036386e-03 6.123609e-02  
##   
## Statistics based on eigenvalues:  
## Hibrido1 Hibrido2 Hibrido3 Hibrido4 Hibrido5  
## product 9.193513e+01 3.297601e+04 1.032747e+04 4.069887e+04 2.044097e+03  
## sum 6.914445e+04 1.424912e+04 2.962122e+04 2.401806e+04 1.991324e+04  
## precision 7.097921e-04 9.535699e-03 3.018319e-03 1.136031e-02 3.336508e-03  
## max 6.907189e+04 1.409453e+04 2.940413e+04 2.381222e+04 1.971829e+04  
## Hibrido6 Hibrido7 pooled  
## product 2.423063e+03 3.674108e+03 7.729119e+05  
## sum 1.566077e+04 2.340941e+04 2.800232e+04  
## precision 3.056881e-03 1.922631e-03 3.105360e-02  
## max 1.491218e+04 2.328317e+04 2.773086e+04

heplots::boxM(cbind(rendimiento, altura\_planta, diametro\_fruto, largo\_fruto, peso\_grano, contenido\_cafeina) ~ Hibrido,   
 data = datos4)

##   
## Box's M-test for Homogeneity of Covariance Matrices  
##   
## data: datos4  
## Chi-Sq (approx.) = 130.6, df = 126, p-value = 0.3713

library(biotools)  
resultado4 <- boxM(X4, datos4$Hibrido);resultado4

##   
## Box's M-test for Homogeneity of Covariance Matrices  
##   
## data: X4  
## Chi-Sq (approx.) = 130.6, df = 126, p-value = 0.3713

library(covTestR)  
cafe <- unique(datos4$Hibrido)  
cafe1 <- lapply(cafe,  
 function(x){as.matrix(datos4[datos4$Hibrido == x, 2:7])}  
)  
names(cafe1) <- cafe  
Ahmad2017(cafe1)

##   
## Ahmad 2017 Homogeneity of Covariance Matrices Test  
##   
## data: Hibrido1, Hibrido2, Hibrido3, Hibrido4, Hibrido5, Hibrido6 and Hibrido7  
## Standard Normal = 3.6452e+10, Mean = 0, Variance = 1, p-value < 2.2e-16  
## alternative hypothesis: true difference in covariance matrices is not equal to 0

## Prueba Wrapper  
homogeneityCovariances(datos4, group = Hibrido, covTest = BoxesM)

## Warning: `invoke()` is deprecated as of rlang 0.4.0.  
## Please use `exec()` or `inject()` instead.  
## This warning is displayed once every 8 hours.

##   
## Boxes' M Homogeneity of Covariance Matrices Test  
##   
## data: Hibrido1, Hibrido2, Hibrido3, Hibrido4, Hibrido5, Hibrido6 and Hibrido7  
## Chi-Squared = 219, df = 168, p-value = 0.004976  
## alternative hypothesis: true difference in covariance matrices is not equal to 0

library(DFA.CANCOR)  
HOMOGENEITY(data = datos4,  
 groups = 'Hibrido',   
 variables = c('rendimiento','altura\_planta','diametro\_fruto','largo\_fruto','peso\_grano',"contenido\_cafeina"))

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 69069.63 344.44 -99.28 -157.66  
## altura\_planta 344.44 66.33 -3.07 5.82  
## diametro\_fruto -99.28 -3.07 0.83 -1.72  
## largo\_fruto -157.66 5.82 -1.72 7.49  
## peso\_grano -51.62 -1.43 0.11 0.25  
## contenido\_cafeina 17.73 -0.72 -0.10 0.34  
## peso\_grano contenido\_cafeina  
## rendimiento -51.62 17.73  
## altura\_planta -1.43 -0.72  
## diametro\_fruto 0.11 -0.10  
## largo\_fruto 0.25 0.34  
## peso\_grano 0.11 0.04  
## contenido\_cafeina 0.04 0.06

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 14084.75 353.34 83.54 -66.30  
## altura\_planta 353.34 154.04 4.31 10.05  
## diametro\_fruto 83.54 4.31 3.82 0.90  
## largo\_fruto -66.30 10.05 0.90 6.37  
## peso\_grano 11.65 1.13 0.29 0.22  
## contenido\_cafeina -6.51 0.50 0.02 0.09  
## peso\_grano contenido\_cafeina  
## rendimiento 11.65 -6.51  
## altura\_planta 1.13 0.50  
## diametro\_fruto 0.29 0.02  
## largo\_fruto 0.22 0.09  
## peso\_grano 0.08 -0.04  
## contenido\_cafeina -0.04 0.07

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 29397.42 291.80 -293.12 -156.12  
## altura\_planta 291.80 210.33 18.31 12.56  
## diametro\_fruto -293.12 18.31 6.90 2.97  
## largo\_fruto -156.12 12.56 2.97 6.42  
## peso\_grano -6.49 -0.78 -0.15 -0.09  
## contenido\_cafeina 38.56 -0.13 -0.33 -0.45  
## peso\_grano contenido\_cafeina  
## rendimiento -6.49 38.56  
## altura\_planta -0.78 -0.13  
## diametro\_fruto -0.15 -0.33  
## largo\_fruto -0.09 -0.45  
## peso\_grano 0.03 -0.04  
## contenido\_cafeina -0.04 0.12

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 23697.38 1630.33 -64.30 191.16  
## altura\_planta 1630.33 310.08 -3.68 28.73  
## diametro\_fruto -64.30 -3.68 4.04 -0.78  
## largo\_fruto 191.16 28.73 -0.78 6.43  
## peso\_grano -1.52 0.80 0.04 -0.17  
## contenido\_cafeina -5.80 -2.21 0.34 0.00  
## peso\_grano contenido\_cafeina  
## rendimiento -1.52 -5.80  
## altura\_planta 0.80 -2.21  
## diametro\_fruto 0.04 0.34  
## largo\_fruto -0.17 0.00  
## peso\_grano 0.05 -0.03  
## contenido\_cafeina -0.03 0.08

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 19578.35 1642.78 -104.18 56.11  
## altura\_planta 1642.78 329.49 -7.14 17.75  
## diametro\_fruto -104.18 -7.14 1.52 -0.84  
## largo\_fruto 56.11 17.75 -0.84 3.73  
## peso\_grano 17.31 0.61 0.05 0.00  
## contenido\_cafeina 17.24 1.52 -0.03 -0.10  
## peso\_grano contenido\_cafeina  
## rendimiento 17.31 17.24  
## altura\_planta 0.61 1.52  
## diametro\_fruto 0.05 -0.03  
## largo\_fruto 0.00 -0.10  
## peso\_grano 0.06 0.04  
## contenido\_cafeina 0.04 0.09

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 14906.27 177.22 50.16 -228.08  
## altura\_planta 177.22 738.76 -5.21 -74.62  
## diametro\_fruto 50.16 -5.21 2.67 -1.53  
## largo\_fruto -228.08 -74.62 -1.53 12.98  
## peso\_grano 8.51 3.17 -0.15 -0.59  
## contenido\_cafeina 5.85 1.90 -0.01 -0.34  
## peso\_grano contenido\_cafeina  
## rendimiento 8.51 5.85  
## altura\_planta 3.17 1.90  
## diametro\_fruto -0.15 -0.01  
## largo\_fruto -0.59 -0.34  
## peso\_grano 0.07 0.02  
## contenido\_cafeina 0.02 0.02

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 23211.72 1266.59 212.34 -13.12  
## altura\_planta 1266.59 184.30 24.24 -0.84  
## diametro\_fruto 212.34 24.24 10.32 -1.36  
## largo\_fruto -13.12 -0.84 -1.36 2.95  
## peso\_grano -22.92 -1.04 -0.34 -0.09  
## contenido\_cafeina 10.27 2.02 0.07 -0.17  
## peso\_grano contenido\_cafeina  
## rendimiento -22.92 10.27  
## altura\_planta -1.04 2.02  
## diametro\_fruto -0.34 0.07  
## largo\_fruto -0.09 -0.17  
## peso\_grano 0.07 0.01  
## contenido\_cafeina 0.01 0.04

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 27706.502 815.215 -30.692 -53.431  
## altura\_planta 815.215 284.761 3.965 -0.078  
## diametro\_fruto -30.692 3.965 4.301 -0.336  
## largo\_fruto -53.431 -0.078 -0.336 6.625  
## peso\_grano -6.439 0.351 -0.022 -0.067  
## contenido\_cafeina 11.050 0.413 -0.005 -0.091  
## peso\_grano contenido\_cafeina  
## rendimiento -6.439 11.050  
## altura\_planta 0.351 0.413  
## diametro\_fruto -0.022 -0.005  
## largo\_fruto -0.067 -0.091  
## peso\_grano 0.067 0.000  
## contenido\_cafeina 0.000 0.069

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 1.000 0.290 -0.089 -0.125  
## altura\_planta 0.290 1.000 0.113 -0.002  
## diametro\_fruto -0.089 0.113 1.000 -0.063  
## largo\_fruto -0.125 -0.002 -0.063 1.000  
## peso\_grano -0.150 0.081 -0.041 -0.101  
## contenido\_cafeina 0.254 0.093 -0.009 -0.135  
## peso\_grano contenido\_cafeina  
## rendimiento -0.150 0.254  
## altura\_planta 0.081 0.093  
## diametro\_fruto -0.041 -0.009  
## largo\_fruto -0.101 -0.135  
## peso\_grano 1.000 0.002  
## contenido\_cafeina 0.002 1.000

## Log Determinant  
## Hibrido1 4.521  
## Hibrido2 10.404  
## Hibrido3 9.243  
## Hibrido4 10.614  
## Hibrido5 7.623  
## Hibrido6 7.793  
## Hibrido7 8.209  
## Pooled 13.558

* La prueba de BoxM resultó significativa, indicando que las matrices de varianza-covarianza no son iguales entre los híbridos. Esto viola el supuesto de homogeneidad, pero al tener el mismo número de repeticiones por grupo (n=7), el MANOVA aún es válido.
* Entre las pruebas, la de Pillai suele ser la más robusta en este escenario.

#### 2.3 Correlación entre variables dependientes correlacionadas (Bartlett)

library(psych)

##   
## Adjuntando el paquete: 'psych'

## The following objects are masked from 'package:ggplot2':  
##   
## %+%, alpha

options(scipen = 0)  
cortest.bartlett(cor(datos4[, -1]),  
 n = nrow(datos4[, -1]))

## $chisq  
## [1] 23.21951  
##   
## $p.value  
## [1] 0.07958909  
##   
## $df  
## [1] 15

library(MVTests)

##   
## Adjuntando el paquete: 'MVTests'

## The following object is masked from 'package:datasets':  
##   
## iris

res41 = Bsper(datos4[, -1])  
summary(res41)

## Bartlett's Sphericity Test   
##   
## Chi-Squared Value = 23.21951 , df = 15 and p-value: 0.0796   
##   
## Correlation Matrix   
##   
## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 1.0000000 0.34913330 -0.2858271413 -0.1877407464  
## altura\_planta 0.3491333 1.00000000 -0.1785906731 -0.1282169351  
## diametro\_fruto -0.2858271 -0.17859067 1.0000000000 -0.0003041304  
## largo\_fruto -0.1877407 -0.12821694 -0.0003041304 1.0000000000  
## peso\_grano -0.2417524 0.03713006 -0.0237280909 0.0325381615  
## contenido\_cafeina 0.3307010 0.23877942 -0.1028132318 -0.1667049852  
## peso\_grano contenido\_cafeina  
## rendimiento -0.24175240 0.33070104  
## altura\_planta 0.03713006 0.23877942  
## diametro\_fruto -0.02372809 -0.10281323  
## largo\_fruto 0.03253816 -0.16670499  
## peso\_grano 1.00000000 0.04172227  
## contenido\_cafeina 0.04172227 1.00000000

**Comentario:** El test de esfericidad de Bartlett es significativo, confirmando que las variables dependientes están correlacionadas entre sí. Esto justifica el uso de un MANOVA en lugar de ANOVAs separados.

### 3. Trabajando con el modelo de MANOVA en DCA

modelo4 = manova(cbind(rendimiento, altura\_planta, diametro\_fruto, largo\_fruto, peso\_grano, contenido\_cafeina) ~ Hibrido,  
 data = datos4)

Determinación de la matriz residual y la matriz factorial del MANOVA.

* Variabilidad explicada por el factor (Tratamientos). Matriz suma de cuadrados y productos cruzados del factor (SCOCF)

Matrices4 = summary(modelo4)$SS  
F = Matrices4$Hibrido  
W = Matrices4$Residuals  
F

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 1227242.3712 42479.854469 -5558.487308 -2913.043810  
## altura\_planta 42479.8545 8235.677967 -559.762531 -320.416435  
## diametro\_fruto -5558.4873 -559.762531 59.421792 14.025833  
## largo\_fruto -2913.0438 -320.416435 14.025833 37.362327  
## peso\_grano -432.4758 -4.840641 0.224851 3.908724  
## contenido\_cafeina 530.0980 48.638308 -2.889653 -1.941702  
## peso\_grano contenido\_cafeina  
## rendimiento -432.4758347 530.0979898  
## altura\_planta -4.8406408 48.6383082  
## diametro\_fruto 0.2248510 -2.8896531  
## largo\_fruto 3.9087245 -1.9417020  
## peso\_grano 0.7347633 0.1462388  
## contenido\_cafeina 0.1462388 0.9018408

* Variabilidad residual. Matriz suma de cuadrados y productos cruzados del residual (SCOCR)

W

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 1163673.0863 34239.035271 -1289.0443000 -2244.117414  
## altura\_planta 34239.0353 11959.972714 166.5420143 -3.284814  
## diametro\_fruto -1289.0443 166.542014 180.6255714 -14.109543  
## largo\_fruto -2244.1174 -3.284814 -14.1095429 278.239943  
## peso\_grano -270.4541 14.762971 -0.9161571 -2.821743  
## contenido\_cafeina 464.0953 17.336629 -0.2074143 -3.816300  
## peso\_grano contenido\_cafeina  
## rendimiento -270.4541000 464.0953286  
## altura\_planta 14.7629714 17.3366286  
## diametro\_fruto -0.9161571 -0.2074143  
## largo\_fruto -2.8217429 -3.8163000  
## peso\_grano 2.8012857 0.0063000  
## contenido\_cafeina 0.0063000 2.8782857

* Variabilidad Total. Matriz suma de cuadrados y productos cruzados total (SCOCT) del factor.

T = F + W;T

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 2390915.4574 76718.889741 -6847.5316082 -5157.1612245  
## altura\_planta 76718.8897 20195.650682 -393.2205163 -323.7012490  
## diametro\_fruto -6847.5316 -393.220516 240.0473633 -0.0837102  
## largo\_fruto -5157.1612 -323.701249 -0.0837102 315.6022694  
## peso\_grano -702.9299 9.922331 -0.6913061 1.0869816  
## contenido\_cafeina 994.1933 65.974937 -3.0970673 -5.7580020  
## peso\_grano contenido\_cafeina  
## rendimiento -702.9299347 994.1933184  
## altura\_planta 9.9223306 65.9749367  
## diametro\_fruto -0.6913061 -3.0970673  
## largo\_fruto 1.0869816 -5.7580020  
## peso\_grano 3.5360490 0.1525388  
## contenido\_cafeina 0.1525388 3.7801265

* Bondad de ajuste. Un valor proximo a 1 indica que la mayor parte de la variabilidad total puede atribuirse al factor, mientras que un valor proximo a 0 significa que el factor explica muy poco de esa variabilidad total.

eta2 = 1 - det(W)/det(T);eta2

## [1] 0.8549204

det(F)/det(T)

## [1] 2.480716e-12

El valor de eta2 obtenido es elevado, lo que indica que los híbridos explican una proporción importante de la variabilidad total de los datos. Esto da un respaldo fuerte a la significancia del modelo MANOVA.

* Pruebas de hipotesis del modelo. Calculo de contrastes del modelo
* Contrastes del modelo en relacion a los supuestos. Todos los estadisticos son bastante robustos ante violaciones de normalidad,y la prueba de Roy es muy sensible a violaciones de la hipotesis de la matriz de covariancias. Cuando las muestras son iguales por #grupo, la prueba de Pillai es el estadistico más robusto ante violaciones de los supuestos.

k = 7 #numero de grupos  
p = 6 #numero de variables  
n = 7 #numero de observaciones por grupo  
datosc1 = datos4  
datosc1$Hibrido <- as.numeric(datosc1$Hibrido)  
VMPG = matrix(NA, k, p) #vector de medias por grupo  
for(i in 1:k){  
 VMPG[i,]=colMeans(datos4[datosc1$Hibrido == i, -1])  
}  
VMPG #cada fila es un vector de medias

## [,1] [,2] [,3] [,4] [,5] [,6]  
## [1,] 1774.806 151.9000 13.75714 18.50143 1.465714 1.297143  
## [2,] 1739.649 167.8400 12.60000 19.91571 1.714286 1.335714  
## [3,] 1814.774 137.5300 14.60000 19.53143 1.385714 1.064286  
## [4,] 2219.711 174.2171 11.80857 17.77143 1.365714 1.530000  
## [5,] 1735.690 149.2986 15.38429 18.94000 1.548571 1.447143  
## [6,] 1806.583 174.5229 13.33143 17.37286 1.334286 1.275714  
## [7,] 1775.100 164.1271 13.62143 17.94571 1.412857 1.298571

#Computar B  
(B=n\*(k-1)\*cov(VMPG))

## [,1] [,2] [,3] [,4] [,5]  
## [1,] 1227242.3712 42479.854469 -5558.487308 -2913.043810 -432.4758347  
## [2,] 42479.8545 8235.677967 -559.762531 -320.416435 -4.8406408  
## [3,] -5558.4873 -559.762531 59.421792 14.025833 0.2248510  
## [4,] -2913.0438 -320.416435 14.025833 37.362327 3.9087245  
## [5,] -432.4758 -4.840641 0.224851 3.908724 0.7347633  
## [6,] 530.0980 48.638308 -2.889653 -1.941702 0.1462388  
## [,6]  
## [1,] 530.0979898  
## [2,] 48.6383082  
## [3,] -2.8896531  
## [4,] -1.9417020  
## [5,] 0.1462388  
## [6,] 0.9018408

# Tambien  
n\*(t(VMPG)-colMeans(VMPG))%\*%t(t(VMPG)-colMeans(VMPG))

## [,1] [,2] [,3] [,4] [,5]  
## [1,] 1227242.3712 42479.854469 -5558.487308 -2913.043810 -432.4758347  
## [2,] 42479.8545 8235.677967 -559.762531 -320.416435 -4.8406408  
## [3,] -5558.4873 -559.762531 59.421792 14.025833 0.2248510  
## [4,] -2913.0438 -320.416435 14.025833 37.362327 3.9087245  
## [5,] -432.4758 -4.840641 0.224851 3.908724 0.7347633  
## [6,] 530.0980 48.638308 -2.889653 -1.941702 0.1462388  
## [,6]  
## [1,] 530.0979898  
## [2,] 48.6383082  
## [3,] -2.8896531  
## [4,] -1.9417020  
## [5,] 0.1462388  
## [6,] 0.9018408

# Computar W  
W = (n-1)\*cov(datos4[datosc1$Hibrido == 1, -1])  
for(i in 2:k){  
 W = W + (n-1)\*cov(datos4[datosc1$Hibrido == 1, -1])  
}  
W

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 2900924.4304 14466.6760 -4169.6822 -6621.6847  
## altura\_planta 14466.6760 2785.9986 -128.9120 244.6087  
## diametro\_fruto -4169.6822 -128.9120 34.9030 -72.0795  
## largo\_fruto -6621.6847 244.6087 -72.0795 314.5512  
## peso\_grano -2168.0514 -60.1496 4.6936 10.4443  
## contenido\_cafeina 744.7483 -30.1476 -4.0170 14.2914  
## peso\_grano contenido\_cafeina  
## rendimiento -2168.0514 744.7483  
## altura\_planta -60.1496 -30.1476  
## diametro\_fruto 4.6936 -4.0170  
## largo\_fruto 10.4443 14.2914  
## peso\_grano 4.6002 1.6696  
## contenido\_cafeina 1.6696 2.5588

W+B

## rendimiento altura\_planta diametro\_fruto largo\_fruto  
## rendimiento 4128166.802 56946.53047 -9728.169508 -9534.72851  
## altura\_planta 56946.530 11021.67657 -688.674531 -75.80773  
## diametro\_fruto -9728.170 -688.67453 94.324792 -58.05367  
## largo\_fruto -9534.729 -75.80773 -58.053667 351.91353  
## peso\_grano -2600.527 -64.99024 4.918451 14.35302  
## contenido\_cafeina 1274.846 18.49071 -6.906653 12.34970  
## peso\_grano contenido\_cafeina  
## rendimiento -2600.527235 1274.846290  
## altura\_planta -64.990241 18.490708  
## diametro\_fruto 4.918451 -6.906653  
## largo\_fruto 14.353024 12.349698  
## peso\_grano 5.334963 1.815839  
## contenido\_cafeina 1.815839 3.460641

#### autovalores de W^{-1}B

(lambdas = eigen(solve(W)%\*%B)$values)

## [1] 4.529051e+01 2.177503e+01 3.311684e-01 2.403735e-01 1.041452e-01  
## [6] 1.758241e-08

#### traza de pillai:

sum(lambdas/(1+lambdas))

## [1] 2.471383

#tambien  
sum(diag(solve(W+B)%\*%B))

## [1] 2.471383

* El software R obtiene un valor calculado de pillai diferente al que se calculó manualmente, esto se debe a que hay varias formas de calcular este valor.

summary(modelo4, test = "Pillai")

## Df Pillai approx F num Df den Df Pr(>F)   
## Hibrido 6 1.459 2.2491 36 252 0.0001568 \*\*\*  
## Residuals 42   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

1-pf(2.2491, 36,252)

## [1] 0.0001568657

det(W)/det(W + B) ## Lambda de Wilks

## [1] 0.0005202803

prod(1/(1 + lambdas))

## [1] 0.0005202803

El software R obtiene un valor calculado diferente al que se calculó manualmente, esto se debe a que hay varias formas de calcular este valor.

summary(modelo4, test = "Wilks")

## Df Wilks approx F num Df den Df Pr(>F)   
## Hibrido 6 0.14508 2.5341 36 165.24 3.779e-05 \*\*\*  
## Residuals 42   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

1-pf(2.5341, 36,165.24)

## [1] 3.780791e-05

LH = sum(lambdas);LH ## Lawley Hotelling

## [1] 67.74123

summary(modelo4, test = "Hotelling-Lawley")

## Df Hotelling-Lawley approx F num Df den Df Pr(>F)   
## Hibrido 6 2.681 2.6313 36 212 9.508e-06 \*\*\*  
## Residuals 42   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

1-pf(2.681, 36,212)

## [1] 6.332862e-06

lambdas[1]/(1 + lambdas[1])## Raiz mayor de Roy

## [1] 0.9783973

summary(modelo4, test = "Roy")

## Df Roy approx F num Df den Df Pr(>F)   
## Hibrido 6 1.2875 9.0125 6 42 2.443e-06 \*\*\*  
## Residuals 42   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

1-pf( 9.0125, 6, 42)

## [1] 2.442906e-06

**Conclusiónes:**

* Todas las pruebas multivariadas (Pillai, Wilks, Hotelling-Lawley y Roy) resultaron significativas.
* Esto confirma que los vectores de medias de los híbridos no son iguales. Es decir, los híbridos difieren de manera global en las variables evaluadas.
* Se observa que las variables rendimiento y altura de planta resultaron altamente significativas,lo que indica que contribuyen de manera importante al rechazo de la hipótesis nula.
* Diámetro de fruto y cafeína muestran tendencia a significancia, mientras que largo de fruto y peso de grano no muestran diferencias claras.

### 4. Resultados univariados

#### 4.1 ANOVA por variable

summary.aov(modelo4)

## Response rendimiento :  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Hibrido 6 1227242 204540 7.3824 1.967e-05 \*\*\*  
## Residuals 42 1163673 27707   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Response altura\_planta :  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Hibrido 6 8235.7 1372.61 4.8202 0.0007998 \*\*\*  
## Residuals 42 11960.0 284.76   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Response diametro\_fruto :  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Hibrido 6 59.422 9.9036 2.3028 0.05185 .  
## Residuals 42 180.626 4.3006   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Response largo\_fruto :  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Hibrido 6 37.362 6.2271 0.94 0.4769  
## Residuals 42 278.240 6.6248   
##   
## Response peso\_grano :  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Hibrido 6 0.73476 0.122461 1.8361 0.1152  
## Residuals 42 2.80129 0.066697   
##   
## Response contenido\_cafeina :  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Hibrido 6 0.90184 0.150307 2.1933 0.06259 .  
## Residuals 42 2.87829 0.068531   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

**Comentario:**

* Rendimiento y altura\_planta –>altamente significativos.
* Diámetro\_fruto y cafeína –>tendencia a la significancia.
* Largo\_fruto y peso\_grano–> no significativos.

### 5. Comparaciones múltiples

####5.1 Comparaciones por pares entre híbridos

modelo1 = manova(cbind(rendimiento, altura\_planta, diametro\_fruto, largo\_fruto, peso\_grano, contenido\_cafeina) ~ Hibrido,  
 data = datos4,  
 subset = Hibrido %in% c("Hibrido1","Hibrido2"))  
  
summary(modelo1,test="Pillai")

## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.61254 1.8444 6 7 0.2208  
## Residuals 12

summary(modelo1,test="Wilks")

## Df Wilks approx F num Df den Df Pr(>F)  
## Hibrido 1 0.38746 1.8444 6 7 0.2208  
## Residuals 12

summary(modelo1,test="Hotelling-Lawley")

## Df Hotelling-Lawley approx F num Df den Df Pr(>F)  
## Hibrido 1 1.5809 1.8444 6 7 0.2208  
## Residuals 12

summary(modelo1,test="Roy")

## Df Roy approx F num Df den Df Pr(>F)  
## Hibrido 1 1.5809 1.8444 6 7 0.2208  
## Residuals 12

summary.aov(modelo1)

## Response rendimiento :  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Hibrido 1 4326 4326 0.104 0.7526  
## Residuals 12 498926 41577   
##   
## Response altura\_planta :  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Hibrido 1 889.29 889.29 8.0709 0.01487 \*  
## Residuals 12 1322.23 110.19   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Response diametro\_fruto :  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Hibrido 1 4.6864 4.6864 2.0169 0.181  
## Residuals 12 27.8823 2.3235   
##   
## Response largo\_fruto :  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Hibrido 1 7.001 7.0007 1.0103 0.3347  
## Residuals 12 83.154 6.9295   
##   
## Response peso\_grano :  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Hibrido 1 0.21626 0.216257 2.3184 0.1538  
## Residuals 12 1.11934 0.093279   
##   
## Response contenido\_cafeina :  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Hibrido 1 0.00521 0.005207 0.0797 0.7825  
## Residuals 12 0.78371 0.065310

**Comentario:**

En la comparación directa entre Hibrido1 y Hibrido2 también se encontraron diferencias significativas.Esto respalda que la diferenciación no solo es global, sino que algunos pares difieren claramente.

#### 5.2 Comparaciones por pares entre híbridos

H0:los 2 vectores de promedios son iguales

H1:los 2 vectores de promedios difieren

Prueba <- c("Hibrido1", "Hibrido2", "Hibrido3", "Hibrido4", "Hibrido5", "Hibrido6","Hibrido7")  
comb<-t(combn(length(Prueba) , 2))  
  
for(i in 1:nrow(comb)){  
 modelo.comp = manova(cbind(rendimiento, altura\_planta, diametro\_fruto, largo\_fruto, peso\_grano, contenido\_cafeina) ~ Hibrido,  
 data=datos4,  
 subset = Hibrido %in% Prueba[comb[i,]])  
 print(paste("Prueba: ", Prueba[comb[i,]][1], "y", Prueba[comb[i,]][2]))  
 print(summary(modelo.comp, test = "Pillai"))  
 cat("\n")  
}

## [1] "Prueba: Hibrido1 y Hibrido2"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.61254 1.8444 6 7 0.2208  
## Residuals 12   
##   
## [1] "Prueba: Hibrido1 y Hibrido3"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.63543 2.0334 6 7 0.1874  
## Residuals 12   
##   
## [1] "Prueba: Hibrido1 y Hibrido4"  
## Df Pillai approx F num Df den Df Pr(>F)   
## Hibrido 1 0.88507 8.9847 6 7 0.005283 \*\*  
## Residuals 12   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## [1] "Prueba: Hibrido1 y Hibrido5"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.63633 2.0413 6 7 0.1861  
## Residuals 12   
##   
## [1] "Prueba: Hibrido1 y Hibrido6"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.33325 0.58311 6 7 0.736  
## Residuals 12   
##   
## [1] "Prueba: Hibrido1 y Hibrido7"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.39471 0.7608 6 7 0.6226  
## Residuals 12   
##   
## [1] "Prueba: Hibrido2 y Hibrido3"  
## Df Pillai approx F num Df den Df Pr(>F)   
## Hibrido 1 0.85854 7.0809 6 7 0.01044 \*  
## Residuals 12   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## [1] "Prueba: Hibrido2 y Hibrido4"  
## Df Pillai approx F num Df den Df Pr(>F)   
## Hibrido 1 0.84429 6.3259 6 7 0.01425 \*  
## Residuals 12   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## [1] "Prueba: Hibrido2 y Hibrido5"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.702 2.7483 6 7 0.106  
## Residuals 12   
##   
## [1] "Prueba: Hibrido2 y Hibrido6"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.61706 1.8799 6 7 0.214  
## Residuals 12   
##   
## [1] "Prueba: Hibrido2 y Hibrido7"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.40281 0.78694 6 7 0.6069  
## Residuals 12   
##   
## [1] "Prueba: Hibrido3 y Hibrido4"  
## Df Pillai approx F num Df den Df Pr(>F)   
## Hibrido 1 0.81686 5.2038 6 7 0.02396 \*  
## Residuals 12   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## [1] "Prueba: Hibrido3 y Hibrido5"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.6749 2.422 6 7 0.1361  
## Residuals 12   
##   
## [1] "Prueba: Hibrido3 y Hibrido6"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.61335 1.8507 6 7 0.2196  
## Residuals 12   
##   
## [1] "Prueba: Hibrido3 y Hibrido7"  
## Df Pillai approx F num Df den Df Pr(>F)   
## Hibrido 1 0.79672 4.5724 6 7 0.03328 \*  
## Residuals 12   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## [1] "Prueba: Hibrido4 y Hibrido5"  
## Df Pillai approx F num Df den Df Pr(>F)   
## Hibrido 1 0.85441 6.8467 6 7 0.01147 \*  
## Residuals 12   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## [1] "Prueba: Hibrido4 y Hibrido6"  
## Df Pillai approx F num Df den Df Pr(>F)   
## Hibrido 1 0.79004 4.3899 6 7 0.03681 \*  
## Residuals 12   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## [1] "Prueba: Hibrido4 y Hibrido7"  
## Df Pillai approx F num Df den Df Pr(>F)   
## Hibrido 1 0.7983 4.6174 6 7 0.03248 \*  
## Residuals 12   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## [1] "Prueba: Hibrido5 y Hibrido6"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.67707 2.4461 6 7 0.1336  
## Residuals 12   
##   
## [1] "Prueba: Hibrido5 y Hibrido7"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.70079 2.7325 6 7 0.1073  
## Residuals 12   
##   
## [1] "Prueba: Hibrido6 y Hibrido7"  
## Df Pillai approx F num Df den Df Pr(>F)  
## Hibrido 1 0.18671 0.26783 6 7 0.9353  
## Residuals 12

**Conclusión:**

* Se concluye que no todos los híbridos son iguales entre sí, ya que varias comparaciones por pares resultaron significativas. En particular, el Híbrido4 presenta diferencias claras con la mayoría de los otros híbridos, mientras que otros (como Híbrido6 y Híbrido7) no mostraron diferencias notorias respecto a la mayoría

# Punto 5.(\*)

Presente un Manova con experimentos factoriales en DBCA.

# Punto 6.(\*)

Presente un Mancova en DBCA.