

```

## =====
## Probabilitat i Estadística
## Grupo 22, Bloque D, 5.12.2024
## Gráficos e IC para datos emparejados
## =====
## Notas:
## * Si usáis Mac, sustituir windows() por quartz()
## * Si usáis Linux, sustituir windows() por x11()

## Un conjunto de datos emparejados
## =====
load("Recollida_de_dades")
ls()
summary(Recollida_de_dades)
TWifi <- (Recollida_de_dades$'Wifi (ms)')
TEthernet <- (Recollida_de_dades$'Ethernet (ms)')
Tdiff <- TEthernet - TWifi
##-----

## Análisis descriptivo
## =====
summary(TWifi)
sd(TWifi)

summary(TEthernet)
sd(TEthernet)

summary(Tdiff)
sd(Tdiff)
##-----

## Intervalo de confianza del 95% de  $\mu_2 - \mu_1$  transformado(log)
## =====
TWifi_log <- log((Recollida_de_dades$'Wifi (ms)'))
TEthernet_log <- log((Recollida_de_dades$'Ethernet (ms)'))
t.test(TWifi_log, TEthernet_log, paired = TRUE, conf.level = 0.95)
##-----

## Gráficos para representar datos emparejados
## =====
## (i) Boxplots
windows(width = 12)
par(mfrow = 1:2, las = 1, font.lab = 2, font.axis = 2)
boxplot(Recollida_de_dades[2:3], ylab = "ms", col = c("steelblue", "tomato"),
        main = "Boxplots de los tiempos TWifi y TEthernet")
boxplot(Tdiff, ylab = "ms", xlab = "TEthernet-TWifi",
        main = "Diferencia de tiempos (TEthernet - TWifi)", col = 3)
abline(h = 0, lty = 2, lwd = 2)

```

```
##-----
```

```
## (ii) Gráfico de dispersión de TEthernet vs TWifi
```

```
## =====
```

```
windows()
par(las = 1, font.lab = 4, font.axis = 2)
plot(TEthernet ~ TWifi, Recollida_de_dades, pch = 19, xlim = c(0, max(TWifi, TEthernet)),
     ylim = c(0, max(TWifi, TEthernet)))
abline(0, 1, lwd = 2)
text(800, 7000, "TEthernet > TWifi", font = 2)
text(6500, 400, "TWifi > TEthernet", font = 2)
##-----
```

```
## (iii) Bland-Altman plot
```

```
## =====
```

```
# Graficar el Bland-Altman log-transformado
# Transformación logarítmica de los datos
TWifi_log <- log((Recollida_de_dades$'Wifi (ms)'))
TEthernet_log <- log((Recollida_de_dades$'Ethernet (ms)'))
```

```
# Cargar la librería BlandAltmanLeh
library(BlandAltmanLeh)
```

```
# Gráfico Bland-Altman con datos transformados
```

```
windows()
bland.altman.plot(
  TWifi_log, TEthernet_log,
  main = "Bland-Altman Plot (Transformación Logarítmica)",
  xlab = "Mitjana de TWifi i TEthernet (log(ms))",
  ylab = "Diferència (log(TWifi) - log(TEthernet))",
  pch = 19, col = "black"
)
```

```
# Gráfico con intervalos de confianza al 95%
```

```
bland.altman.plot(
  TWifi_log, TEthernet_log,
  conf.int = 0.95, # Intervalos de confianza
  main = "Bland-Altman Plot (Transformación Logarítmica con IC 95%)",
  xlab = "Mitjana de TWifi i TEthernet (log(ms))",
  ylab = "Diferència (log(TWifi) - log(TEthernet))",
  pch = 19, col = "blue"
)
```

```
##-----  
## (iv) LM  
## =====
```

```
Recollida_de_dades2$`Temps(ms)` <- log(Recollida_de_dades2$`Temps(ms)`)  
Recollida_de_dades2$`Mida(MB)` <- log(Recollida_de_dades2$`Mida(MB)`)
```

```
datos <- Recollida_de_dades2[, 2:4]  
wifi <- Recollida_de_dades2[, c(2,4)]  
ethernet <- Recollida_de_dades2[, c(2,4)]  
wifi <- wifi[1:60,]  
ethernet <- ethernet[61:120,]
```

```
#Para el wifi  
Wlmod <- lm(`Temps(ms)` ~ `Mida(MB)`, data = wifi)  
print(summary(Wlmod), signif.stars = FALSE)
```

```
windows(width = 12)  
par(las = 1, font.lab = 2, font.axis = 2, mfrow = 1:2)  
plot(Wlmod, 1:2, ask = FALSE, pch = 16)
```

```
windows(width = 8)  
par(las = 1, font = 2, font.axis = 2, font.lab = 2)  
with(wifi, plot(`Temps(ms)` ~ `Mida(MB)`, pch = 16,  
              xlab = "Mida(MB)", ylab = "Temps(ms)"))  
title("Temps que tarda en carregar una pàgina en funció del seu tamany amb Wi-fi")  
abline(6.5, 0.5, lwd = 3)
```

```
##Para el ethernet  
Elmod <- lm(`Temps(ms)` ~ `Mida(MB)`, data = ethernet)  
print(summary(Elmod), signif.stars = FALSE)
```

```
windows(width = 12)  
par(las = 1, font.lab = 2, font.axis = 2, mfrow = 1:2)  
plot(Elmod, 1:2, ask = FALSE, pch = 16)
```

```
windows(width = 8)  
par(las = 1, font = 2, font.axis = 2, font.lab = 2)  
with(ethernet, plot(`Temps(ms)` ~ `Mida(MB)`, pch = 16,  
                   xlab = "Mida(MB)", ylab = "Temps(ms)"))  
title("Temps que tarda en carregar una pàgina en funció del seu tamany amb Ethernet")  
abline(6.5, 0.5, lwd = 3)
```

```
##Para el wifi + ethernet
```

```
datos$`Wifi(1) / Ethernet(0)` <- as.factor(datos$`Wifi(1) / Ethernet(0)`)
WEImod <- lm(`Temps(ms)` ~ `Wifi(1) / Ethernet(0)` + `Mida(MB)`, data = datos)
print(summary(WEImod), signif.stars = FALSE)
```

```
windows(width = 12)
par(las = 1, font.lab = 2, font.axis = 2, mfrow = 1:2)
plot(WEImod, 1:2, ask = FALSE, pch = 16)
```

```
windows(width = 8)
par(las = 1, font.lab = 4, font.axis = 2, font = 2)
plot(`Temps(ms)` ~ `Mida(MB)`, data = datos, pch = 16, xlab = "Mida(MB)",
     ylab = "", col = as.numeric(`Wifi(1) / Ethernet(0)`) + 1)
title("El temps que es tarda en descarregar una pàgina web amb Wi-fi(1) i Ethernet(0)")
legend("topleft", levels(datos$`Wifi(1) / Ethernet(0)`), col = 2:3, bty = "n", pch = 16)
abline(lm(`Temps(ms)` ~ `Mida(MB)`, data = wifi),
       col = 1, lwd = 2)
abline(lm(`Temps(ms)` ~ `Mida(MB)`, data = ethernet),
       col = 1, lwd = 2)
```

#interval confiança (paired data) els dos donen exactament el mateix calcul

```
Twifi <- wifi[[1]]
Tethernet <- ethernet[[1]]
t.test(Tethernet, Twifi, paired = TRUE)
```

```
Tdiff <- Tethernet - Twifi
t.test(Tdiff)
```