
CI0131 Diseño de experimentos - Laboratorio 3 - I Ciclo 2024

T-test: Comparación de medias independientes

Integrantes:

- A. Badilla Olivas B80874
- Brandon Mora Umaña C15179

Introducción

En este laboratorio, vamos a analizar el conjunto de datos `nacim.csv` para determinar si existen diferencias significativas en el peso de los bebés al nacer en función del sexo del bebé y si el bebé fue prematuro. Para esto, realizaremos un t-test de dos muestras para cada caso.

A. Peso del bebé en función del sexo

1. Cargar datos y librerías

1.A Instalacion y carga de librerías

```
install.packages(c(
  "dplyr",
  "car",
  "effsize",
  "tidyverse",
  "ggplot2",
  "gridExtra",
  "tidyr",
  "reshape2",
  "RColorBrewer",
  "ggrepel",
  "BSDA"
))
```

```
## Installing packages into '/home/brandonmu/R/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
```

```
library(tidyverse)
library(car)
library(effsize)
library(dplyr)
library(ggplot2)
library(gridExtra)
library(tidyr)
library(reshape2)
library(RColorBrewer)
library(ggrepel)
library(rmarkdown)
library(BSDA)
options(width = 80)
```

Primero, cargamos el conjunto de datos `nacim.csv`.

```
nacim <- read.csv("nacim.csv", header=TRUE, stringsAsFactors=TRUE)
head(nacim, 10)
```

```
##      Column1 f_age m_age weeks premature visits gained weight sex_baby  smoke
## 1          1    31    30    36    premie    13      1   6.88    male   smoker
## 2          2    34    36    39 full term     5    35   7.69    male nonsmoker
## 3          3    36    35    40 full term    12    29   8.88    male nonsmoker
## 4          4    41    40    40 full term    13    30   9.00    female nonsmoker
## 5          5    42    37    40 full term    NA    10   7.94    male nonsmoker
## 6          6    37    28    40 full term    12    35   8.25    male   smoker
## 7          7    35    35    28    premie     6    29   1.63    female nonsmoker
## 8          8    28    21    35    premie     9    15   5.50    female   smoker
## 9          9    22    20    32    premie     5    40   2.69    male   smoker
## 10         10    36    25    40 full term    13    34   8.75    female nonsmoker
```

```
summary(nacim)
```

```
##      Column1      f_age      m_age      weeks
##  Min.   : 1.00   Min.   :17.00   Min.   :15.0   Min.   :26.00
## 1st Qu.: 38.25   1st Qu.:25.00   1st Qu.:22.0   1st Qu.:37.00
##  Median : 75.50   Median :29.00   Median :25.0   Median :39.00
##  Mean   : 75.50   Mean   :29.78   Mean   :26.6   Mean   :38.33
## 3rd Qu.:112.75   3rd Qu.:34.00   3rd Qu.:31.0   3rd Qu.:40.00
##  Max.   :150.00   Max.   :47.00   Max.   :41.0   Max.   :44.00
##
##      NA's   :31
##      premature      visits      gained      weight      sex_baby
## full term:120   Min.   : 2.0   Min.   : 0.00   Min.   : 1.630   female:68
## premie   : 30   1st Qu.:10.0   1st Qu.:22.75   1st Qu.: 6.455   male  :82
##
##      Median :12.0   Median :30.50   Median : 7.310
##
##      Mean   :11.5   Mean   :32.45   Mean   : 7.046
##
##      3rd Qu.:14.0   3rd Qu.:40.00   3rd Qu.: 8.000
##
##      Max.   :19.0   Max.   :85.00   Max.   :10.130
##
##      NA's   :1     NA's   :2
##
##      smoke
## nonsmoker:100
## smoker   : 50
##
##
##
##
```

2. Hipótesis

Establecemos las hipótesis nula y alternativa:

- **H0:** No hay diferencia en el peso promedio al nacer entre bebés de sexo masculino y femenino.
- **HA:** Existe una diferencia en el peso promedio al nacer entre bebés de sexo masculino y femenino.

3. Parámetro estimado (estadístico)

Calculamos la diferencia entre las medias muestrales del peso de los bebés de sexo masculino y femenino.

```
male <- nacim %>% filter(sex_baby == "male") %>% pull(weight)
female <- nacim %>% filter(sex_baby == "female") %>% pull(weight)
```

```
mean(male) - mean(female)
```

```
## [1] 0.1839491
```

4. Validar las condiciones para aplicar un t-test

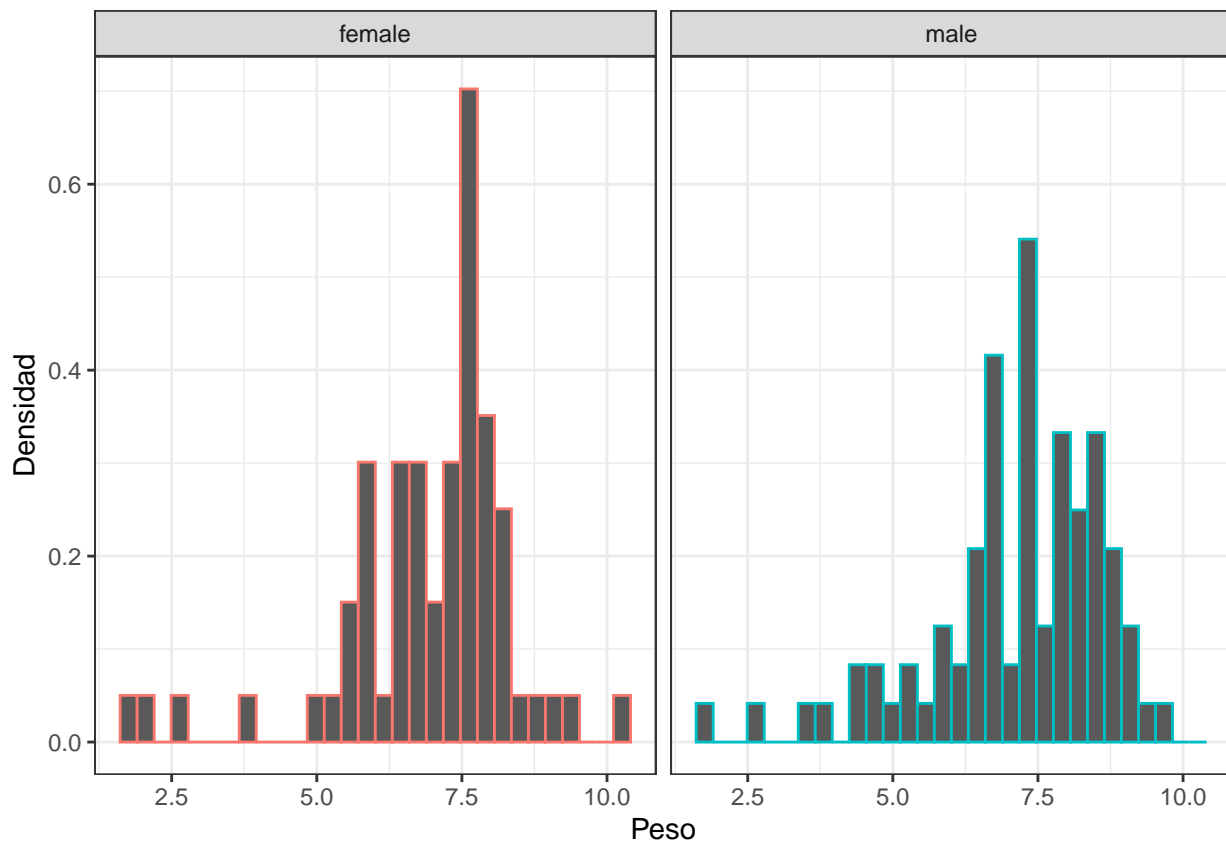
4.1. Independencia: Se asume que los datos son independientes ya que se trata de un muestreo aleatorio.

4.2. Normalidad: Evaluamos la normalidad de la variable `weight` para cada sexo utilizando histogramas, gráficos Q-Q y la prueba de Shapiro-Wilk.

```
ggplot(nacim,aes(x = weight)) +  
  geom_histogram(aes(y = ..density.., colour = sex_baby)) +  
  facet_grid(.~ sex_baby) +  
  theme_bw() + theme(legend.position = "none") + labs(x="Peso", y="Densidad")
```

Histogramas

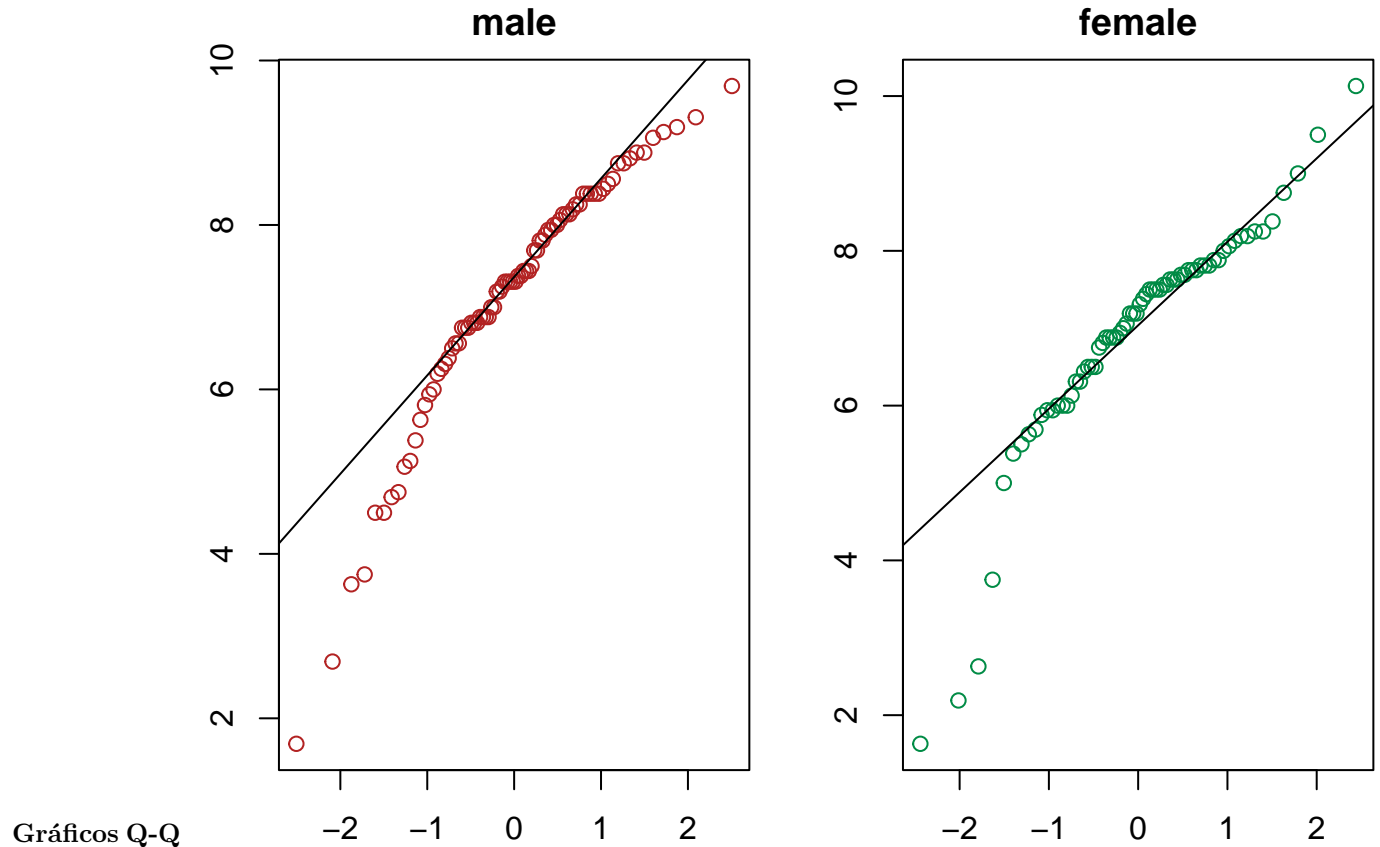
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
par(mar = c(2, 2, 2, 2))  
par(mfrow = c(1, 2))  
  
qqnorm(male, xlab = "", ylab = "",  
       main = "male", col = "firebrick")
```

```
qqline(male)

qqnorm(female, xlab = "", ylab = "",
       main = "female", col = "springgreen4")
qqline(female)
```



```
shapiro.test(male)
```

Prueba de Shapiro-Wilk

```
##
##  Shapiro-Wilk normality test
##
## data:  male
## W = 0.92418, p-value = 0.0001243
```

```
shapiro.test(female)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  female
## W = 0.88002, p-value = 8.936e-06
```

Conclusión de Normalidad: Los gráficos Q-Q muestran que la distribución del peso de los bebés se desvía ligeramente de la normalidad, especialmente en el caso de las mujeres. La prueba de Shapiro-Wilk confirma esta desviación, ya que el valor p es menor que 0.05 en ambos casos, rechazando la hipótesis nula de

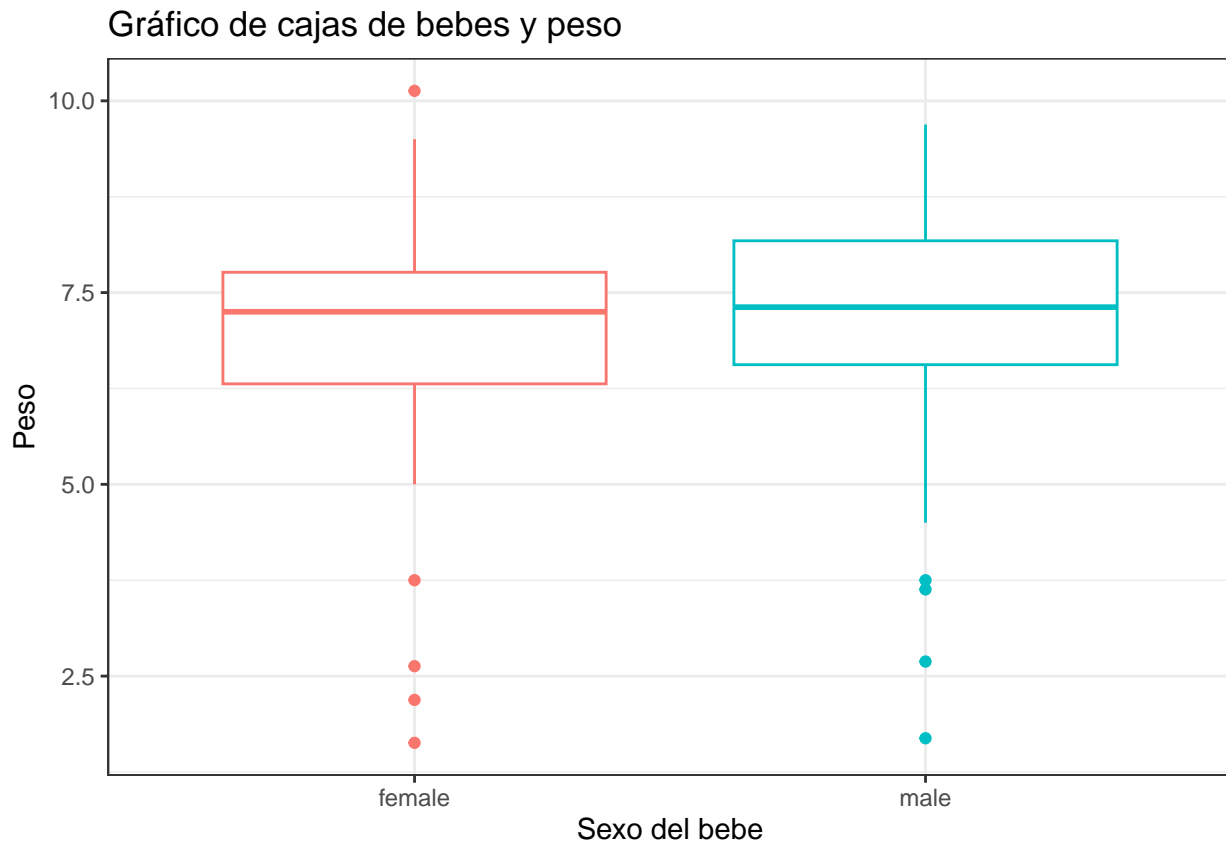
normalidad.

Nota: A pesar de que no se cumple la condición de normalidad, para efectos del ejercicio, continuaremos con el t-test. Es importante mencionar esta desviación en las conclusiones.

4.3. Igualdad de Varianza Evaluamos la homogeneidad de varianzas utilizando la prueba de Fligner-Killeen y la prueba de Levene.

Gráfico de cajas Primero observamos la distribución de los datos con un gráfico de cajas y bigotes.

```
ggplot(data = nacim) +  
  geom_boxplot(aes(x = sex_baby, y = weight, colour = sex_baby)) +  
  theme_bw() + theme(legend.position = "none") + labs(title='Gráfico de cajas de bebés y peso', x="Sexo del bebé")
```



```
fligner.test(weight ~ sex_baby, data = nacim)
```

Pruebas estadísticas

```
##  
## Fligner-Killeen test of homogeneity of variances  
##  
## data: weight by sex_baby  
## Fligner-Killeen: med chi-squared = 0.74732, df = 1, p-value = 0.3873  
leveneTest(weight ~ sex_baby, data = nacim, center = "median")  
  
## Levene's Test for Homogeneity of Variance (center = "median")  
## Df F value Pr(>F)
```

```
## group    1  0.2424 0.6232
##          148
```

Conclusión de Igualdad de Varianza: Ni la prueba de Fligner-Killeen ni la prueba de Levene encontraron evidencia significativa de que las varianzas sean diferentes entre ambos grupos.

5. Determinar el tipo de test

Utilizaremos un t-test de dos colas, ya que estamos interesados en determinar si hay una diferencia en el peso promedio al nacer, sin especificar una dirección.

6. Determinar el nivel de significancia

Establecemos el nivel de significancia en $\alpha = 0.05$.

7. Cálculo de p-value

Realizamos el t-test de dos muestras utilizando la función `t.test()`.

```
t.test(
  x = male,
  y = female,
  alternative = "two.sided",
  mu = 0,
  var.equal = TRUE,
  conf.level = 0.95
)

##
## Two Sample t-test
##
## data: male and female
## t = 0.74798, df = 148, p-value = 0.4557
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.3020371 0.6699352
## sample estimates:
## mean of x mean of y
## 7.129390 6.945441
```

7. 1. Prueba Z extra

```
z.test(
  x = male,
  y = female,
  alternative = "two.sided",
  mu = 0,
  conf.level = 0.95,
  sigma.x = sd(male),
  sigma.y = sd(female)
)

##
## Two-sample z-Test
##
## data: male and female
```

```
## z = 0.74943, p-value = 0.4536
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.2971252  0.6650234
## sample estimates:
## mean of x mean of y
## 7.129390  6.945441
```

8. Cálculo del tamaño del efecto

Calculamos la d de Cohen para evaluar el tamaño del efecto.

```
cohen.d(formula = weight ~ sex_baby, data = nacim, paired = FALSE)
```

```
##
## Cohen's d
##
## d estimate: -0.1226794 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.4470958  0.2017369
```

9. Conclusiones. Análisis de resultados

El p-value (0.4557) para la prueba t es mayor que alpha (0.05), por lo que no se dispone de evidencia suficiente para rechazar la hipótesis nula. Esto significa que no hay evidencia de una diferencia significativa en el peso promedio al nacer entre bebés de sexo masculino y femenino. También, dado que hay más de 30 muestras para cada grupo, la prueba z indica que la hipótesis nula no se rechaza con un p-value de (0.4536).

El tamaño del efecto, medido con la d de Cohen, es pequeño (0.18), lo que indica que la diferencia observada entre los grupos es pequeña, incluso si fuera estadísticamente significativa.

B. Peso del bebé en función de si fue prematuro

1. Hipótesis

Establecemos las hipótesis nula y alternativa:

- **H0:** No hay diferencia en el peso promedio al nacer entre bebés prematuros y no prematuros.
- **HA:** Existe una diferencia en el peso promedio al nacer entre bebés prematuros y no prematuros.

2. Parámetro estimado (estadístico)

Calculamos la diferencia entre las medias muestrales del peso de los bebés prematuros y no prematuros.

```
premature <- nacim %>% filter(premature == "premie") %>% pull(weight)
fullterm <- nacim %>% filter(premature == "full term") %>% pull(weight)

mean(fullterm) - mean(premature)
```

```
## [1] 2.28
```

3. Validar las condiciones para aplicar un t-test

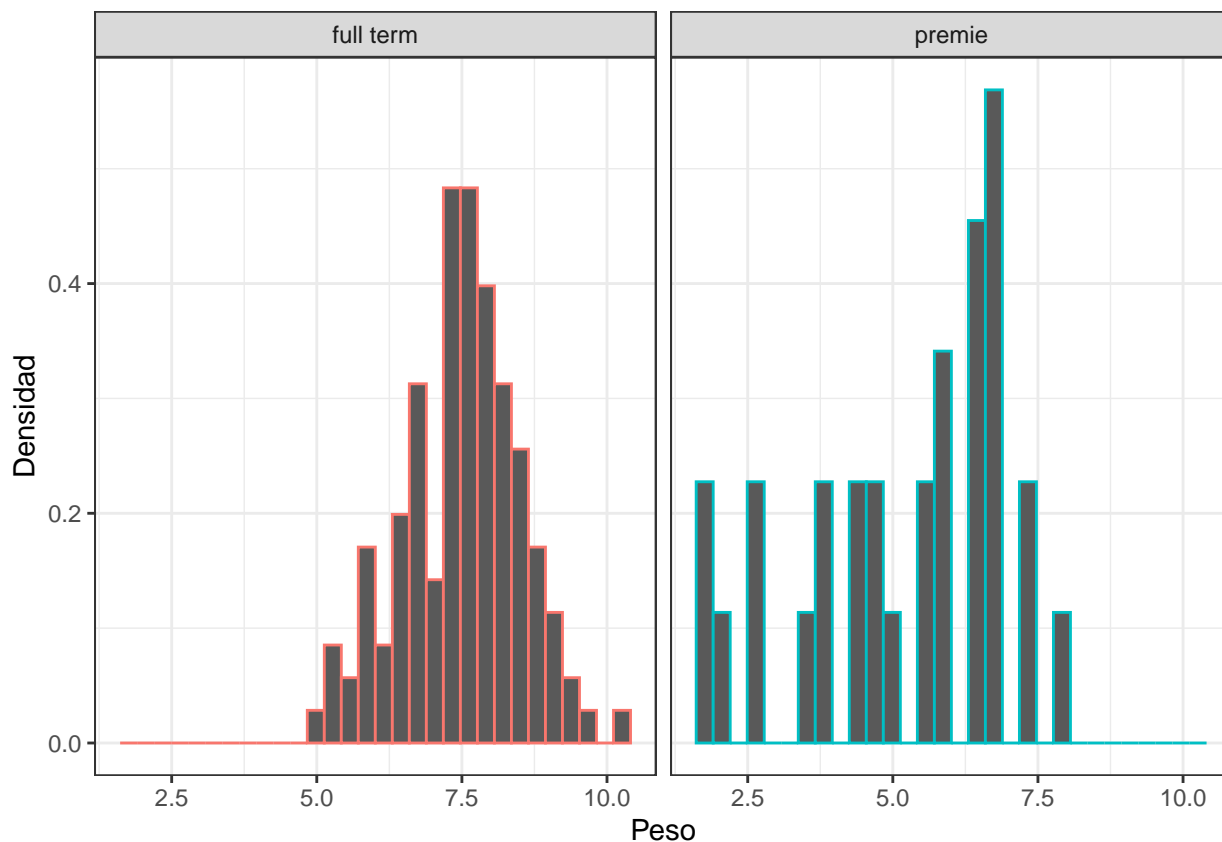
3.1. Independencia: Se asume que los datos son independientes ya que se trata de un muestreo aleatorio.

3.2. Normalidad: Evaluamos la normalidad de la variable **weight** para bebés prematuros y no prematuros utilizando histogramas, gráficos Q-Q y la prueba de Shapiro-Wilk.

```
ggplot(nacim,aes(x = weight)) +
  geom_histogram(aes(y = ..density.., colour = premature)) +
  facet_grid(~ premature) +
  theme_bw() + theme(legend.position = "none") + labs(x="Peso", y='Densidad')
```

Histogramas

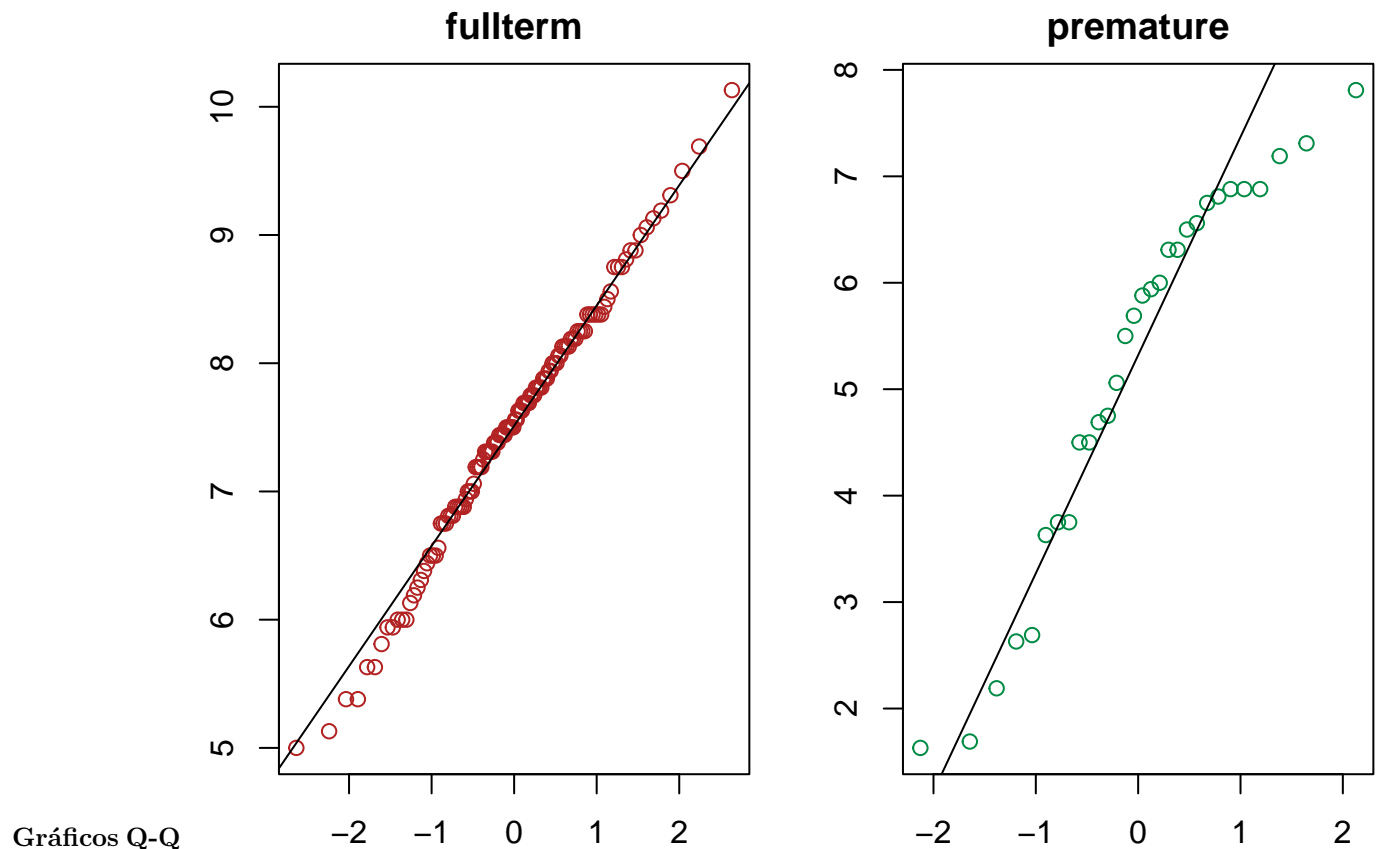
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.



```
par(mar = c(2, 2, 2, 2))
par(mfrow = c(1, 2))

qqnorm(fullterm, xlab = "", ylab = "",
        main = "fullterm", col = "firebrick")
qqline(fullterm)

qqnorm(premature, xlab = "", ylab = "",
        main = "prematuro", col = "springgreen4")
qqline(premature)
```

```
shapiro.test(fullterm)
```

Prueba de Shapiro-Wilk

```
##
##  Shapiro-Wilk normality test
##
## data:  fullterm
## W = 0.99267, p-value = 0.7829
```

```
shapiro.test(premature)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  premature
## W = 0.92213, p-value = 0.03047
```

Conclusión de Normalidad: Los histogramas muestran una clara desviación de la normalidad en el grupo de bebés prematuros, con una cola más larga hacia la derecha. Los gráficos Q-Q también muestran desviaciones de la normalidad. La prueba de Shapiro-Wilk confirma esta desviación, ya que el valor p es menor que 0.05 en el caso de los bebés prematuros.

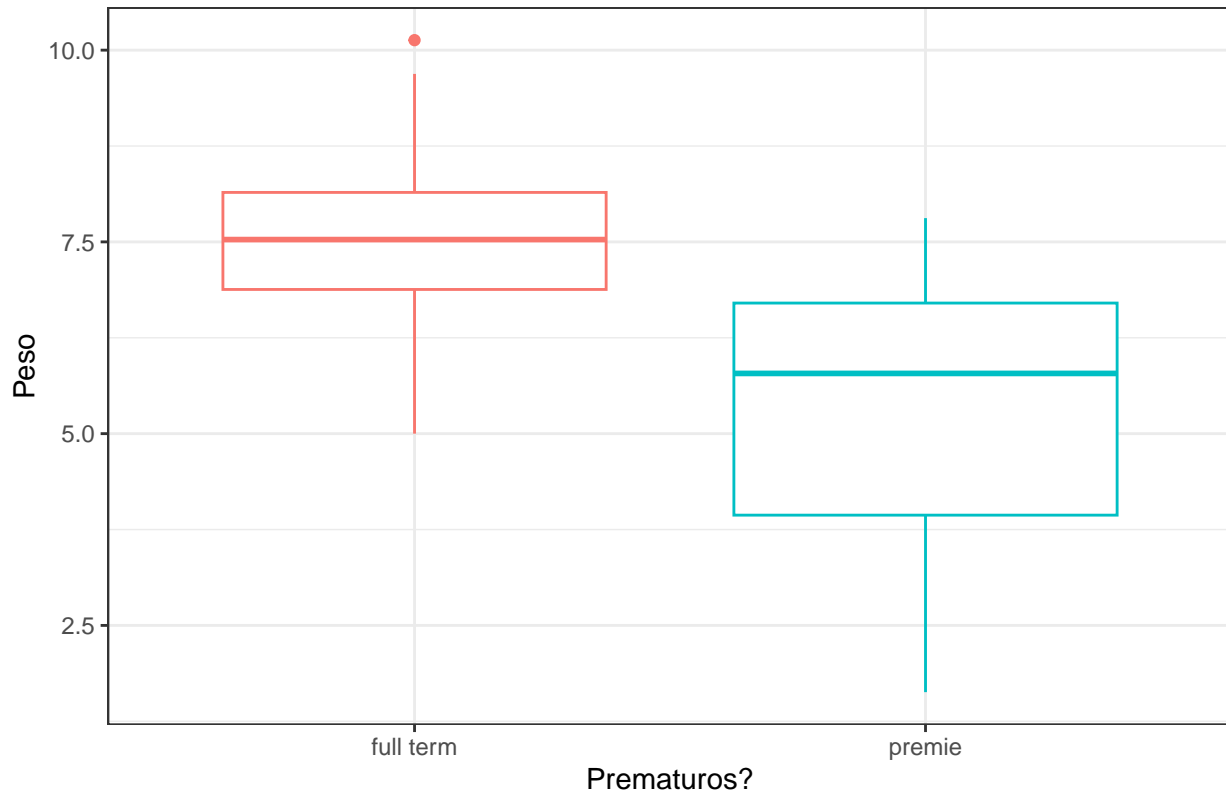
Nota: A pesar de que no se cumple la condición de normalidad, para efectos del ejercicio, continuaremos con el t-test. Es importante mencionar esta desviación en las conclusiones.

3.3. Igualdad de Varianza Evaluamos la homogeneidad de varianzas utilizando la prueba de Fligner-Killeen y la prueba de Levene.

Gráfico de cajas Primero observamos la distribución de los datos con un gráfico de cajas y bigotes.

```
ggplot(data = nacim) +  
  geom_boxplot(aes(x = premature, y = weight, colour = premature)) +  
  theme_bw() + theme(legend.position = "none") + labs(title="Gráfico de cajas peso en prematuros y no p
```

Gráfico de cajas peso en prematuros y no prematuros



```
fligner.test(weight ~ premature, data = nacim)
```

Pruebas estadísticas

```
##  
## Fligner-Killeen test of homogeneity of variances  
##  
## data: weight by premature  
## Fligner-Killeen:med chi-squared = 13.772, df = 1, p-value = 0.0002064
```

```
leveneTest(weight ~ premature, data = nacim, center = "median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")  
##      Df F value    Pr(>F)  
## group 1 18.983 2.453e-05 ***  
##      148  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Conclusión de Igualdad de Varianza: Tanto la prueba de Fligner-Killeen como la prueba de Levene rechazan la hipótesis nula de igualdad de varianzas.

4. Determinar el tipo de test

Utilizaremos un t-test de dos colas.

5. Determinar el nivel de significancia

Establecemos el nivel de significancia en $\alpha = 0.05$.

6. Cálculo de p-value

Realizamos el t-test de dos muestras, considerando que las varianzas no son iguales.

```
t.test(
  x = fullterm,
  y = premature,
  alternative = "two.sided",
  mu = 0,
  var.equal = FALSE,
  conf.level = 0.95
)

##
## Welch Two Sample t-test
##
## data: fullterm and premature
## t = 6.7599, df = 33.645, p-value = 9.523e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.59429 2.96571
## sample estimates:
## mean of x mean of y
##      7.502      5.222
```

6. 1. Prueba Z extra

```
z.test(
  x = fullterm,
  y = premature,
  alternative = "two.sided",
  mu = 0,
  conf.level = 0.95,
  sigma.x = sd(fullterm),
  sigma.y = sd(premature)
)

##
## Two-sample z-Test
##
## data: fullterm and premature
## z = 6.7599, p-value = 1.381e-11
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.618935 2.941065
## sample estimates:
## mean of x mean of y
##      7.502      5.222
```

7. Cálculo del tamaño del efecto

Calculamos la d de Cohen para evaluar el tamaño del efecto.

```
cohen.d(formula = weight ~ premature, data = nacim, paired = FALSE)

##
## Cohen's d
##
## d estimate: 1.917523 (large)
## 95 percent confidence interval:
##      lower      upper
## 1.458642 2.376405
```

8. Conclusiones. Análisis de resultados

El p-value (9.523×10^{-8}) es mucho menor que alpha (0.05), por lo que rechazamos la hipótesis nula. Esto significa que hay evidencia de una diferencia significativa en el peso promedio al nacer entre bebés prematuros y no prematuros. Para la prueba z el p-value (1.381×10^{-11}) sostiene estas mismas conclusiones.

El tamaño del efecto, medido con la d de Cohen, es grande (1.917523), lo que indica una diferencia sustancial en el peso promedio al nacer entre los dos grupos.

Nota: La validez de los resultados del t-test se ve afectada por la falta de normalidad en la distribución del peso de los bebés prematuros y la diferencia en las varianzas entre los grupos. Sería prudente considerar análisis adicionales, como pruebas no paramétricas o métodos de bootstrapping, para confirmar estos hallazgos.