

ANÁLISIS GRÁFICO DE BLAND-ALTMAN

Cristina Gil Martínez (cristina_gil_m@hotmail.com)

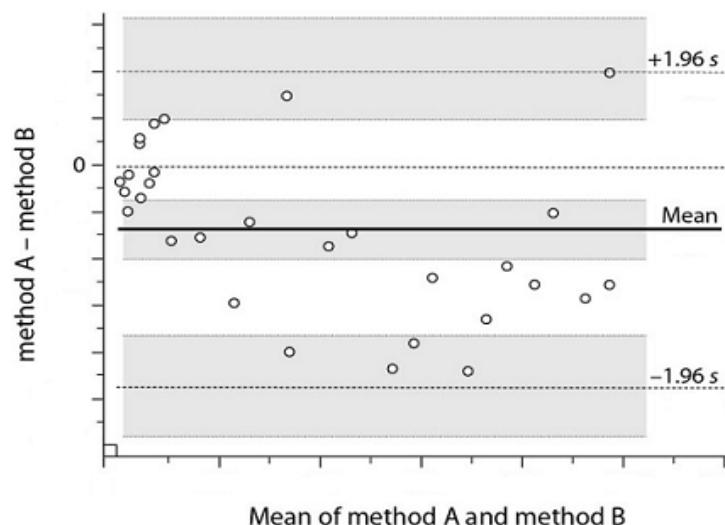
Noviembre, 2017

1. INTRODUCCIÓN

El método de **Bland-Altman** es un método gráfico que permite comparar dos técnicas de medición sobre una misma variable cuantitativa. Es aplicable, por ejemplo, a casos en los que se requiere medir la diferencia entre un nuevo método con respecto a uno ya establecido o validado, y así poder validar o no el nuevo en función de si resulta fiable y reproducible para el uso previsto. Para ello, el método de Bland-Altman cuantifica la diferencia media entre ambos métodos (bias o sesgo) y un rango de confianza, entre los cuales se espera que se incluyan el 95% de las diferencias entre una técnica de medida y la otra. El análisis de Bland-Altman también puede ser útil para evaluar la reproducibilidad de un solo método. El gráfico podría ser útil en este caso para determinar si la variabilidad o precisión del método en cuestión se relaciona con el tamaño de la característica que se está midiendo. En este caso particular, dado que se usa siempre el mismo método, el bias sería 0. Para determinar si ambos métodos pueden usarse indistintamente, han de establecerse *a priori* los límites de las diferencias máximas aceptables en base a criterios biológicos/analíticos.

2. ELEMENTOS E INTERPRETACIÓN

A continuación, se muestra un ejemplo de la estructura de un gráfico de Bland-Altman (Giavarina, 2015):



En esta representación, el eje Y corresponde a las diferencias entre los valores pareados de los métodos A y B ($A-B$), mientras que el eje X representa el respectivo valor de la media de ambos ($(A+B)/2$), como mejor estimador del valor real de la variable. De manera alternativa se pueden graficar los datos en función de uno de los métodos que se considere como referencia. También podrían graficarse las diferencias en formato de porcentajes o ratios.

Tres son las líneas paralelas que se representan:

- **Limit of agreement** superior: diferencia media + $1,96 \cdot SD$.
- **Diferencia media**: valor medio determinado por A – valor medio determinado por B. Refleja el error sistemático, en contraposición con la precisión (SD y CI) que refleja el error aleatorio.
- **Limit of agreement** inferior: diferencia media - $1,96 \cdot SD$.

Si el método A y B obtienen de media valores similares, entonces la diferencia media se situará en cero o próxima a cero. Si se encontrara lejos de este valor, significaría que los dos métodos producen resultados diferentes (el nuevo sub- o sobreestima el valor del método validado).

La representación de los *limits of agreement* o límites de concordancia permiten juzgar visualmente la concordancia entre ambos métodos. Estos límites establecen el rango en el que se encontrarán aproximadamente un 95% de las veces, las diferencias en los datos de una técnica y de la otra. Cuanto menor sea el rango entre los límites, mejor será el “agreement”. Un rango alto entre ambos límites de concordancia implicaría una baja precisión de uno o ambos métodos. Si estos límites no exceden la diferencia máxima aceptable entre métodos, se considera que los dos están en concordancia y pueden ser usados indistintamente. Es importante tener en cuenta que un cambio en la escala de medida no afecta la correlación entre los datos, pero sí afecta la concordancia.

Con respecto a los datos, es importante observar si la variabilidad es consistente a lo largo del rango de valores y en torno a la línea de la diferencia media, o si por el contrario se detectan anomalías, como por ejemplo: los datos en torno a la línea de bias siguen una cierta tendencia conforme el valor de la media entre ambas técnicas aumenta (bias proporcional); todos o la mayoría de los datos se encuentran por encima o por debajo de esta línea.

3. EJEMPLO: ANÁLISIS DE BLAND-ALTMAN EN R

Para el presente ejemplo se emplea un set de datos hipotético obtenido de Giavarina (2015) para evaluar la concordancia entre un método A y un método B, con 30 observaciones cada uno.

Paso 1: Importar datos desde archivo Excel / Generar la matriz de datos en R.

```
metodo_A <- c (1, 5, 10, 20, 50, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000)
metodo_B <- c (8, 16, 30, 24, 39, 54, 40, 68, 72, 62, 122, 80, 181, 259, 275, 380, 320, 434, 479, 587, 626, 648, 738, 766, 793, 851, 871, 957, 1001, 960)
diferencia <- metodo_A - metodo_B
media <- (metodo_A + metodo_B)/2
porcentaje <- ((diferencia/media)*100)
datos <- data.frame(metodo_A, metodo_B, diferencia, media, porcentaje)
```

```
head(datos)
```

```
##  metodo_A metodo_B diferencia media porcentaje
## 1         1         8         -7   4.5 -155.55556
## 2         5        16        -11  10.5 -104.76190
## 3        10        30        -20  20.0 -100.00000
## 4        20        24         -4  22.0  -18.18182
## 5        50        39         11  44.5   24.71910
## 6        40        54        -14  47.0  -29.78723
```

```
summary(datos)
```

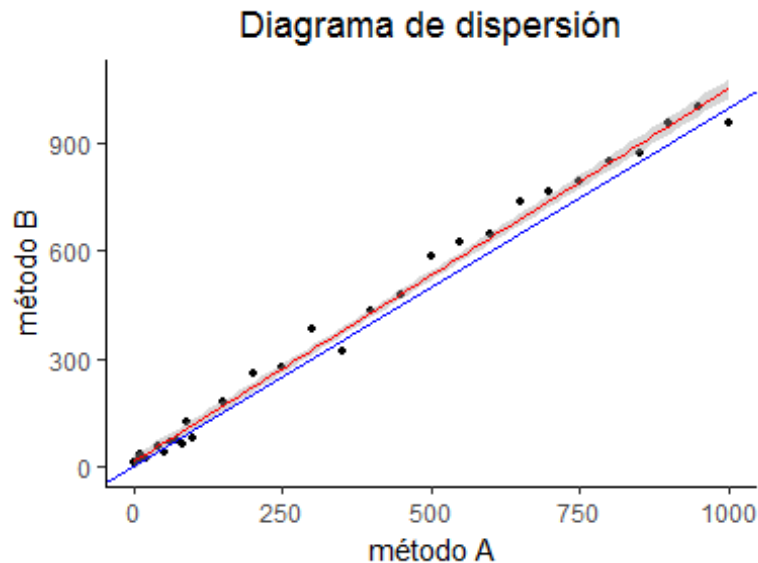
```
##      metodo_A      metodo_B      diferencia      media
## Min.   :  1.0   Min.   :  8.0   Min.   : -88.00   Min.   :  4.50
## 1st Qu.: 62.5   1st Qu.: 63.5   1st Qu.: -51.00   1st Qu.: 65.75
## Median :275.0   Median :297.5   Median : -27.00   Median :298.75
## Mean   :364.2   Mean   :391.4   Mean   : -27.17   Mean   :377.78
## 3rd Qu.:637.5   3rd Qu.:715.5   3rd Qu.:  -4.75   3rd Qu.:676.50
## Max.   :1000.0   Max.   :1001.0   Max.   :  40.00   Max.   :980.00
## porcentaje
## Min.   : -155.556
## 1st Qu.: -18.594
## Median :  -8.579
## Mean   : -17.400
## 3rd Qu.:  -3.420
## Max.   :  25.352
```

Paso 2: Análisis gráfico de los datos.

Como primer paso, inspeccionamos nuestros datos mediante la representación de un diagrama de dispersión:

```
library(ggplot2)
#Diagrama de dispersión con La línea de identidad (bisectriz) y modelo Li
neal de los datos
```

```
ggplot(data = datos, mapping = aes(x = metodo_A, y = metodo_B)) +
  geom_point(color = "black", size = 1) +
  labs(title = "Diagrama de dispersión", x = "método A", y = "método B") +
  geom_smooth(method = "lm", se = TRUE, color = "red", lwd = 0.5) +
  geom_abline(intercept = 0, slope = 1, lwd = 0.7, col = "blue") +
  theme_bw() + theme(plot.title = element_text(hjust = 0.5)) +
  theme(axis.line = element_line(colour = "black"), panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.border = element_blank(), panel.background = element_blank())
```



#Coeficiente de correlación

```
cor.test(datos$metodo_A, datos$metodo_B, method = "pearson")
```

```
## Pearson's product-moment correlation
##
## data:  datos$metodo_A and datos$metodo_B
## t = 57.561, df = 28, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.9910929 0.9980231
## sample estimates:
##      cor
## 0.9958011
```

El coeficiente de correlación de Pearson muestra una fuerte asociación positiva entre los valores obtenidos por el método A y B ($r = 0.996$, 95% CI = 0.991-0.998, $p\text{-value} < 0.001$). Sin embargo, una alta correlación no es necesariamente sinónimo de concordancia entre métodos, puesto que evalúa la relación y no la diferencia.

```

#Modelo lineal de Los datos
modelo_lineal <- lm (formula = metodo_B~metodo_A, data= datos)
summary(modelo_lineal)

##
## Call:
## lm(formula = metodo_B ~ metodo_A, data = datos)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -93.615 -11.908   0.052  12.314  55.504
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  12.0163     8.8667   1.355   0.186
## metodo_A      1.0416     0.0181  57.561 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 32.49 on 28 degrees of freedom
## Multiple R-squared:  0.9916, Adjusted R-squared:  0.9913
## F-statistic: 3313 on 1 and 28 DF,  p-value: < 2.2e-16

#Intervalos de confianza (CI) de Los parámetros del modelo lineal
confint(modelo_lineal, level= 0.95)

##              2.5 %    97.5 %
## (Intercept) -6.146185 30.178863
## metodo_A      1.004532  1.078666

```

En este ejemplo, la recta de regresión $y = 1.04x + 12.02$ tiene una pendiente de 1.04, lo cual podría evaluarse como una buena concordancia entre métodos. En este caso la pendiente indica que el método B sobreestima al A. Además, los datos muestran una dispersión muy cercana a la bisectriz (pero no sobre ella), lo cual también sugiere la existencia de un pequeño nivel de discrepancia entre los métodos.

Paso 3: Obtención del gráfico Bland-Altman

MÉTODO MANUAL:

```

BA_plot <- ggplot(data= datos, aes(x=media, y = diferencia))+
  geom_point(pch= 1, size=1.5, col= "black")+
  labs(title = "Bland-Altman plot", x = "media método A y B", y = "método
A - método B") +
  ylim(mean(datos$diferencia)-4*sd(datos$diferencia), mean(datos$diferencia)+4*sd(diferencia))+
  #Línea de bias

```

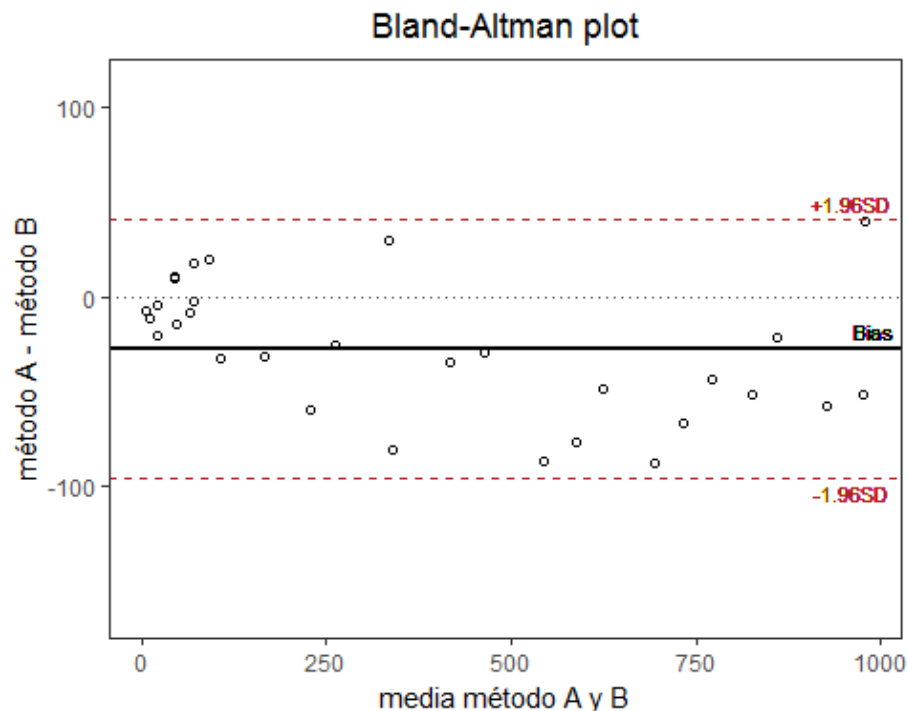
```

geom_hline(yintercept = mean(datos$diferencia), lwd= 1)+
#Línea en y=0
geom_hline(yintercept = 0, lty= 3, col="grey30")+
#Limits of Agreement
geom_hline(yintercept = mean(datos$diferencia)+1.96*sd(datos$diferencia
), lty=2, col= "firebrick")+
geom_hline(yintercept = mean(datos$diferencia)-1.96*sd(datos$diferencia
), lty=2, col= "firebrick")+
theme_bw()+theme(plot.title = element_text(hjust = 0.5))+
theme(panel.grid.major = element_blank(), panel.grid.minor = element_b
lank())+
geom_text(label = "Bias", x = 991, y = -18, size = 3, colour = "black")
+
geom_text(label = "+1.96SD", x = 960, y = 50, size = 3, colour = "fireb
rick")+
geom_text(label = "-1.96SD", x = 960, y = -103, size = 3, colour = "fir
ebrick")

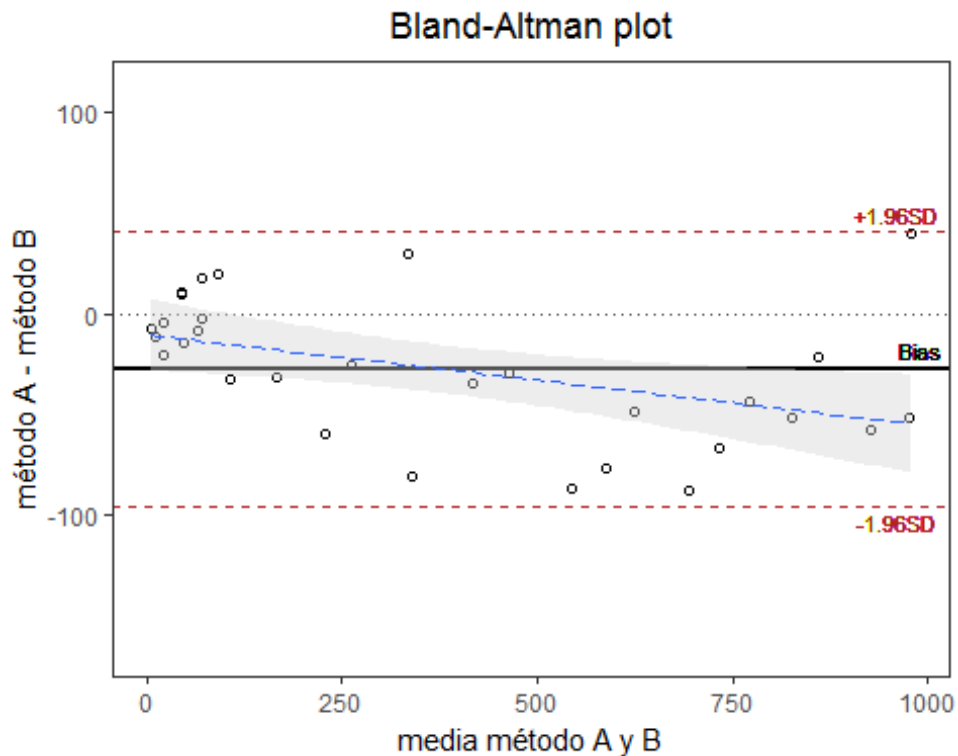
```

BA_plot

En este ejemplo, el bias entre métodos es de -27.17 unidades, lo que significa que en promedio el método B mide 27.17 unidades más que el método A (24/30 = 80% de los datos se sitúan por debajo de 0). Además, este sesgo negativo parece deberse a medidas por encima de las 200 unidades, puesto que, para unidades medias menores a 200, los datos parecen encontrarse más próximos entre sí y más próximos a una diferencia media de cero.



```
BA_plot+
geom_smooth(method = "lm", se = TRUE, fill= "lightgrey", lwd= 0.1, lty=5)
```



La línea de regresión calculada para las diferencias denota un bias sistemático proporcional (no constante), con una tendencia negativa de las diferencias conforme aumenta la magnitud de la variable medida. En este caso podemos calcular el % de error (proporción entre la magnitud y el error de la medida):

```
sd <- sd(datos$diferencia)
bias <- mean(datos$diferencia)
LoA.superior <- bias+1.96*sd
LoA.inferior <- bias-1.96*sd
#Porcentaje de error tomando como referencia el método A
error <- (LoA.superior-LoA.inferior)/mean(metodo_A)*100

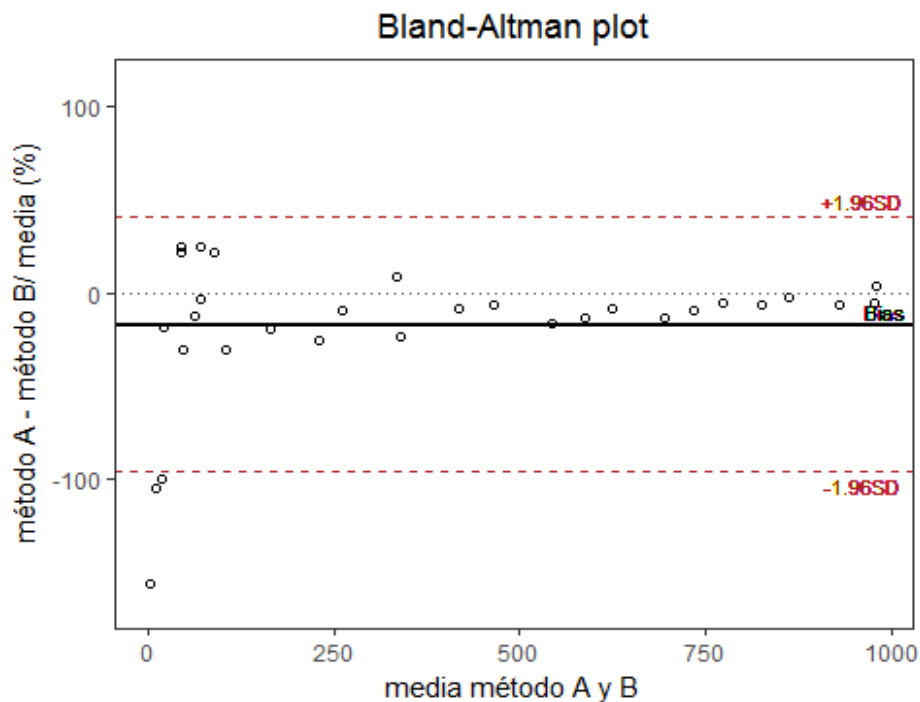
error

## [1] 37.46274
```

También sería útil en este caso representar un gráfico utilizando el porcentaje de las diferencias:

```
BA_plot2 <- ggplot(data= datos, aes(x=media, y = porcentaje))+
  geom_point(pch= 1, size=1.5, col= "black")+
  labs(title = "Bland-Altman plot", x = "media método A y B", y = "método
A - método B/ media (%)") +
  ylim(mean(datos$diferencia)-4*sd(datos$diferencia), mean(datos$diferencia)+4*sd(diferencia))+
  #Línea de bias
  geom_hline(yintercept = mean(datos$porcentaje), lwd= 1)+
  #Línea en y=0
  geom_hline(yintercept = 0, lty= 3, col="grey30")+
  #Limits of Agreement
  geom_hline(yintercept = mean(datos$diferencia)+1.96*sd(datos$diferencia
), lty=2, col= "firebrick")+
  geom_hline(yintercept = mean(datos$diferencia)-1.96*sd(datos$diferencia
), lty=2, col= "firebrick")+
  theme_bw()+theme(plot.title = element_text(hjust = 0.5))+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_b
lank())+
  geom_text(label = "Bias", x = 991, y = -10, size = 3, colour = "black")
+
  geom_text(label = "+1.96SD", x = 960, y = 50, size = 3, colour = "fireb
rick")+
  geom_text(label = "-1.96SD", x = 960, y = -103, size = 3, colour = "fir
ebrick")
```

BA_plot2



```
mean(datos$porcentaje)
```



```
## [1] -17.3999
```

El bias es de un -17.4%, casi constante para todo el rango de la variable, con la excepción de valores menores a 100 unidades. Los *limits of agreement* se sitúan del 41.05% al -95.39%.

MÉTODO AUTOMÁTICO:

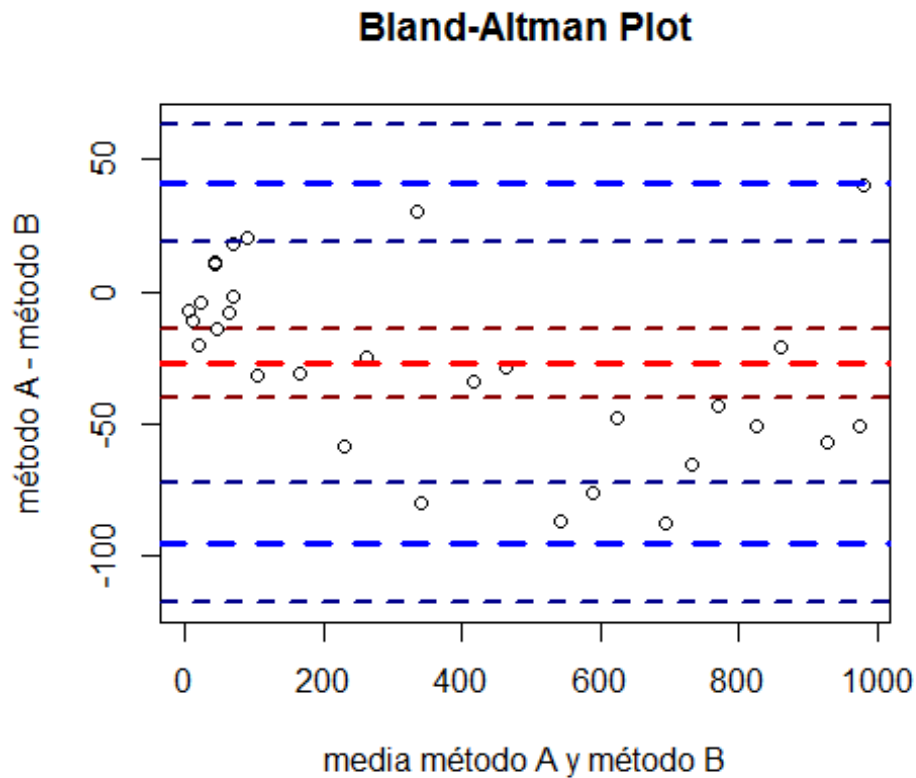
Entre otros, el paquete de R `BlandAltmanLeh` permite generar el gráfico de manera automática, además de calcular los estadísticos asociados, ofreciendo una opción más directa para el análisis aunque con menos flexibilidad para personalizar el gráfico resultante. Por un lado, la función `bland.altman.plot` de este paquete toma como argumentos principales:

- `group1`: vector correspondiente a los datos del primer método.
- `group2`: vector correspondiente a los datos del segundo método.
- `two`: líneas correspondientes a los *limits of agreement*. El valor por defecto de este argumento es de 1.96.
- `mode`: puede tomar valor 1 o 2. Si se ajusta a 1 (usado por defecto), se calcula la diferencia del primer método menos es segundo. El valor 2, calcula la diferencia del segundo método menos el primero.
- `graph.sys`: hace referencia al sistema de gráficos de R. Puede ajustarse a "base" o "ggplot2".
- `conf.int`: toma por defecto el valor 0, el cual excluye del gráfico los intervalos de confianza. Para que se muestren, ajustar este argumento a .95.

Por otro lado, la función `bland.altman.stats` puede utilizarse para el cálculo de los estadísticos.

```
library(BlandAltmanLeh)
```

```
bland.altman.plot(metodo_A, metodo_B, main="Bland-Altman Plot", xlab="med  
ia método A y método B", ylab="método A - método B", conf.int=.95)
```



```
BA_stats <- bland.altman.stats(metodo_A, metodo_B)
BA_stats$mean.diffs
## [1] -27.16667

BA_stats$lower.limit
## [1] -95.38632

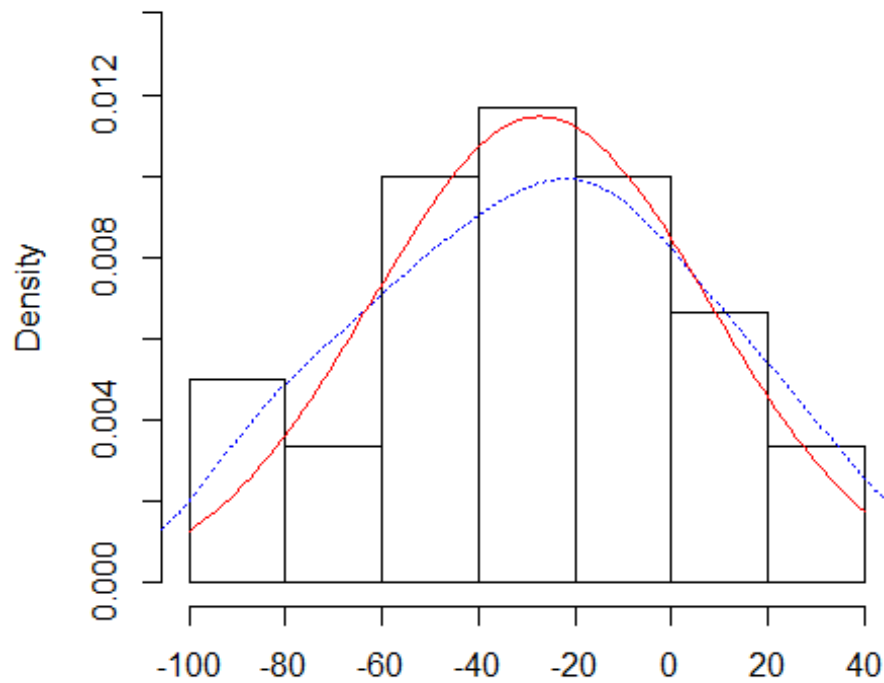
BA_stats$upper.limit
## [1] 41.05299
```

Paso 4: Verificar la distribución normal de los datos

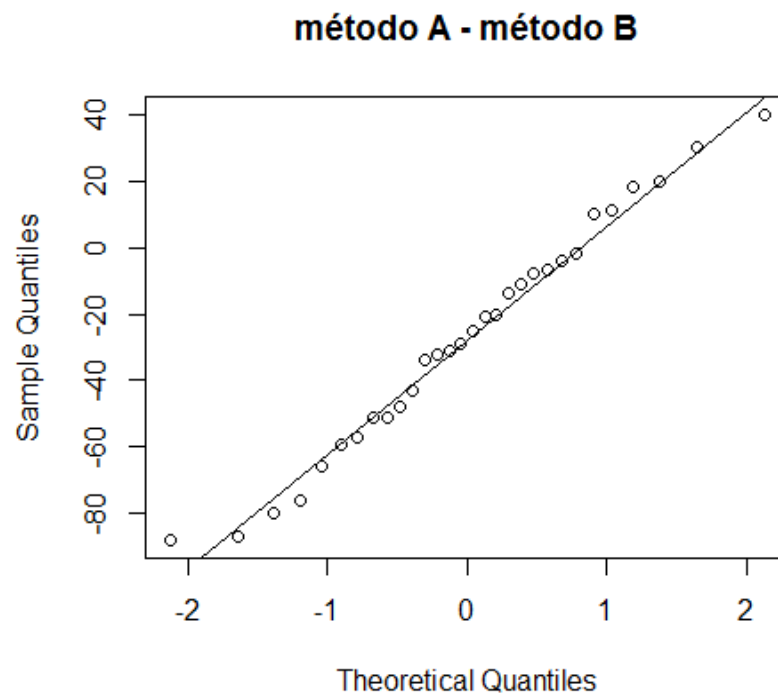
Para verificar esta condición gráficamente, puede elaborarse un histograma, acompañado del uso de tests estadísticos como el test de Kolmogorov-Smirnov o Shapiro-Wilk.

```
library(psych)
```

```
multi.hist(x = datos$diferencia, dcol = c("blue", "red"), dltty = c("dotted", "solid"), main="")
```



```
qqnorm(datos$diferencia, main = "método A - método B")
qqline(datos$diferencia)
```



```
shapiro.test(datos$diferencia)

##
## Shapiro-Wilk normality test
##
## data:  datos$diferencia
## W = 0.97956, p-value = 0.8137
```

El contraste de hipótesis confirma la distribución normal de los datos (p-value = 0.81), lo cual nos permite someterlos a tests paramétricos.

Si las diferencias no se distribuyeran normalmente, una transformación logarítmica de los valores originales podría llevarse a cabo. De manera alternativa, se puede optar por el uso de métodos no paramétricos. Estos son igualmente válidos pero más ineficientes en comparación con los paramétricos (y por tanto requieren una mayor cantidad de datos para el análisis).

Paso 5: Cálculo del 95% CI del bias y los *limits of agreement*

Tras asegurarnos que nuestras diferencias siguen una distribución gaussiana, podemos hacer uso del error estándar de la media (SEM) y la desviación estándar (SD) para obtener los intervalos de confianza correspondientes para saber cómo de precisas son las estimaciones del bias y los *limits of agreement*. Con la función `bland.altman.stats` podemos obtener los intervalos directamente:

```
bland.altman.stats(metodo_A, metodo_B)$CI.lines

## lower.limit.ci.lower lower.limit.ci.upper mean.diff.ci.lower
##          -117.89736          -72.87529          -40.16342
## mean.diff.ci.upper upper.limit.ci.lower upper.limit.ci.upper
##          -14.16991           18.54195           63.56403
```

El bias es significativo puesto que el intervalo de confianza para este valor (-40.16 a -14.17) no incluye el valor 0.

Tabla resumen de los estadísticos del análisis:

Parámetro	Unidad	SE	(95%CI) de	(95%CI) a
Tamaño muestra	30.00			
Media método A	364.2			
Media método B	391.37			
diferencia media (d)	-27.17	6.35	-40.16	-14.17
SD	34.81			
LoA sup (d+1.96SD)	41.05	11.01	18.54	63.56
LoA inf (d-1.96SD)	-95.39	11.01	-117.90	-72.87

Paso 6: Sacar conclusiones en función de los criterios establecidos a priori

4. Bibliografía

Giavarina D. Understanding Bland Altman analysis. *Biochemia Medica* 2015; 25(2): 141-151.

Hanneman SK. Design, analysis, and interpretation of method-comparison studies. *AACN Advanced Critical Care* 2008; 19(2):223-34.

Kaur P., Stoltzfus J. Bland-Altman plot: A brief overview 2017. *International Journal of Academic Medicine*; 3(1): 110-111.

Twomey PJ. How to use difference plots in quantitative method comparison studies. *Annals of clinical biochemistry* 2006; 43(2): 124-129.

<https://cran.r-project.org/web/packages/BlandAltmanLeh/vignettes/Intro.html>

<https://www.rdocumentation.org/packages/BlandAltmanLeh/versions/0.3.1/topics/bland.altman.plot>

<https://www.medcalc.org/manual/blandaltman.php>