

CI0131 Diseño de experimentos - Laboratorio - I Ciclo 2024

Laboratorio: ANOVA de una vía

Integrantes:

- A. Badilla Olivas B80874
- Brandon Mora Umaña C15179

Introducción

El presente informe tiene como objetivo analizar la eficiencia de tres insecticidas diferentes. Para ello, se utilizará un análisis de la varianza (ANOVA) de una vía. El ANOVA es una prueba estadística que permite comparar las medias de dos o más grupos. En este caso, se compararán las medias del número de insectos muertos por parcela luego de la aplicación de cada uno de los tres insecticidas.

Metodología

Los datos utilizados en este análisis se encuentran en el archivo "Insect.csv". Este archivo contiene 36 lecturas de conteos de insectos muertos por parcela luego de la aplicación de uno de los 3 insecticidas. La variable de respuesta es la columna "contador", que tiene el conteo de los insectos muertos recolectados.

El análisis se llevará a cabo en R, utilizando el paquete "tidyverse".

Análisis

A. Cargar los datos, Revisar Primeras 6 Filas

Cargamos los datos, revimos los primeros datos y vemos también un resumen de estos.

```
# Cargar los datos
Insect <- read.csv("Insect.csv", header = TRUE, encoding = "UTF-8")
```

```
# Convertir la variable "producto" a factor
Insect$producto <- as.factor(Insect$producto)
# Revisar las primeras 6 filas
head(Insect)
```

```
##   contador producto orden
## 1      10  Clorban     1
## 2      11 Bioquim     1
## 3      21 Diazinon    1
## 4       7  Clorban     2
## 5      17 Bioquim     2
## 6      20 Diazinon    2
```

```
# Verificar que la variable "producto" es un factor
summary(Insect)
```

```
##      contador      producto      orden
## Min.   : 7.00   Bioquim :12   Min.    : 1.00
## 1st Qu.:13.00   Clorban :12   1st Qu.: 3.75
## Median :16.50   Diazinon:12   Median  : 6.50
## Mean   :16.42                      Mean    : 6.50
## 3rd Qu.:20.00                      3rd Qu.: 9.25
## Max.   :26.00                      Max.    :12.00
```

B. Media, Desviación Estandar y Diagrama de Cajas

Ahora exploramos la desviación estandar y la media de cada uno de los insecticidas. Así mismo, realizamos un diagrama de cajas para ver las diferencias.

```
# Cargar el paquete dplyr
library(dplyr)

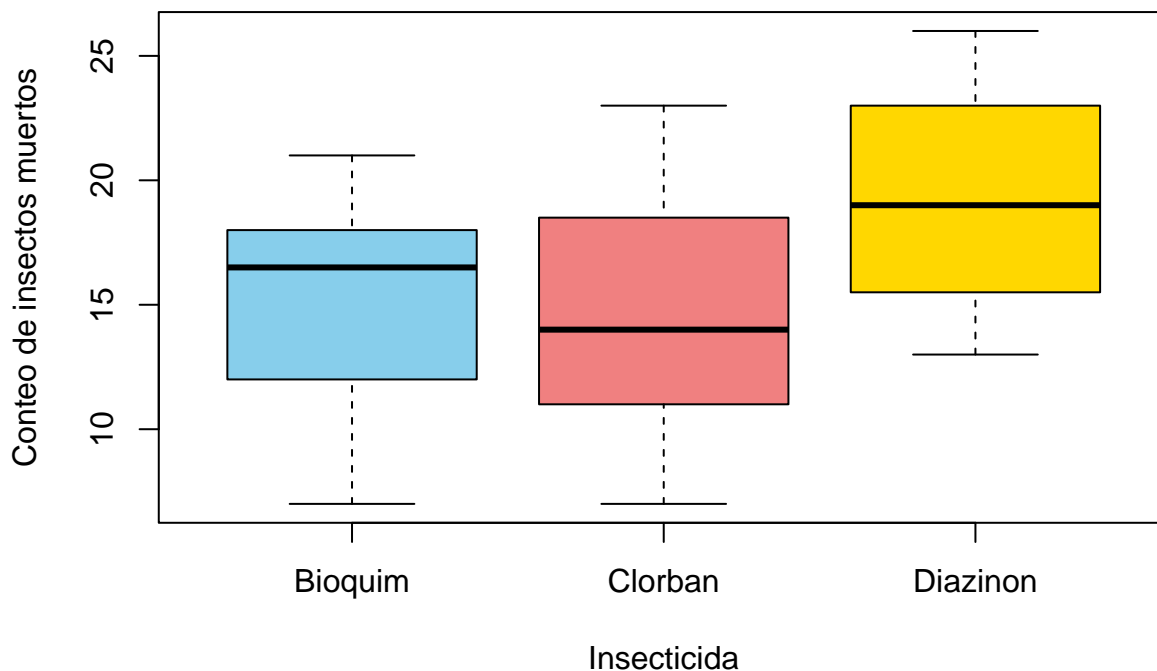
# Calcular la media y la desviación estándar para cada insecticida
Insect %>%
  group_by(producto) %>%
  summarise(media = mean(contador),
            desviacion_estandar = sd(contador))

## # A tibble: 3 x 3
##   producto media desviacion_estandar
##   <fct>     <dbl>             <dbl>
## 1 Bioquim    15.3              4.27
## 2 Clorban    14.5              4.72
## 3 Diazinon   19.4              4.44

# Crear un diagrama de cajas para cada insecticida
colores_insecticidas <- c("skyblue", "lightcoral", "gold")

boxplot(contador ~ producto,
        data = Insect,
        main = "Conteo de insectos muertos por insecticida",
        xlab = "Insecticida",
        ylab = "Conteo de insectos muertos",
        col = colores_insecticidas, # Usamos nuestro vector de colores
        border = "black")
```

Conteo de insectos muertos por insecticida



C. Ajustar el modelo ANOVA de una vía

Proseguimos a ajustar el modelo anova de una vía a los datos de los insecticidas. Específicamente, establecemos la relación contador ~ producto.

```
# Ajustar el modelo ANOVA de una vía
modelo <- aov(contador ~ producto, data = Insect)
```

```
# Ver el resumen del modelo
summary(modelo)
```

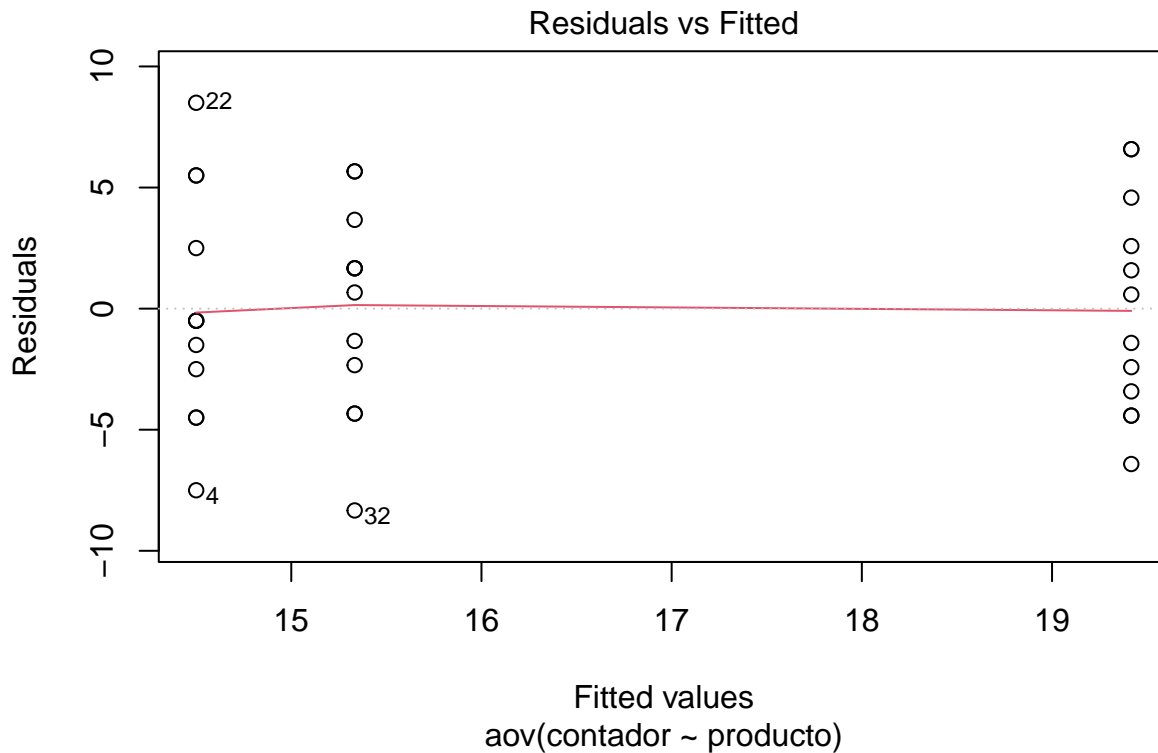
```
##              Df Sum Sq Mean Sq F value Pr(>F)
## producto      2  166.2   83.08    4.138  0.0249 *
## Residuals    33  662.6   20.08
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

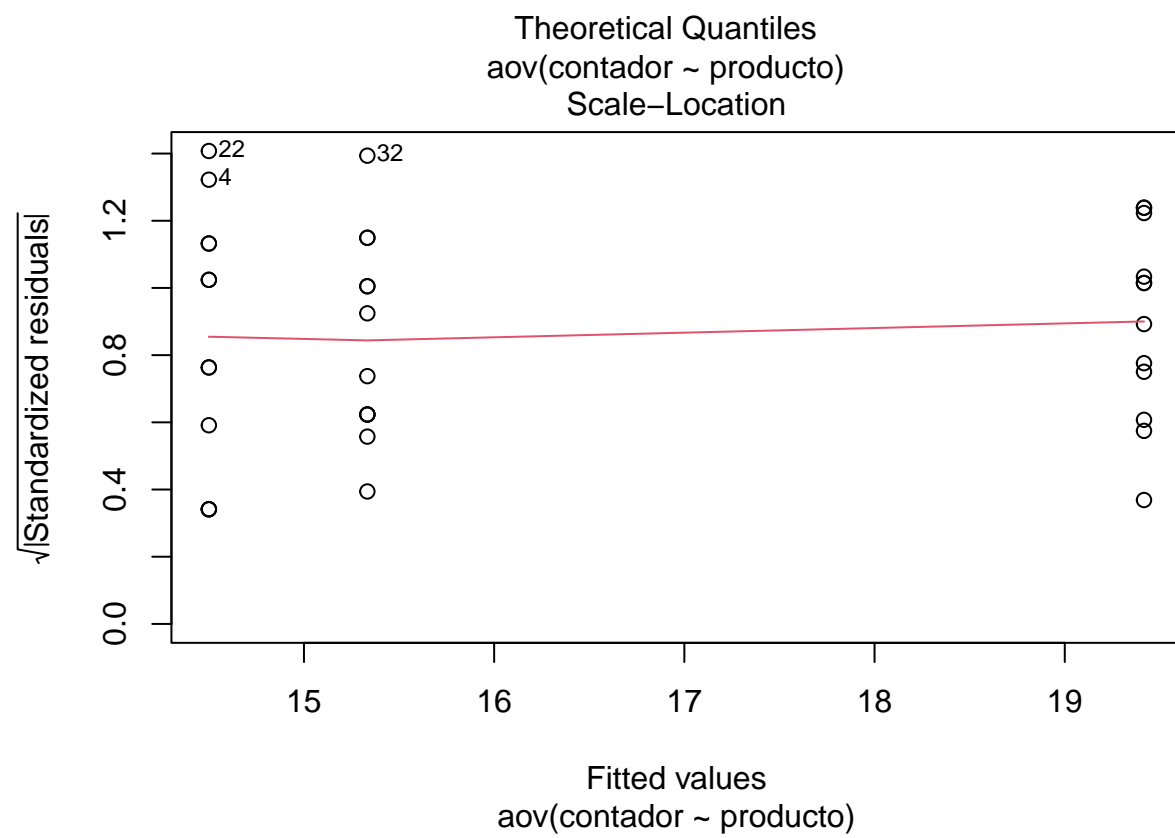
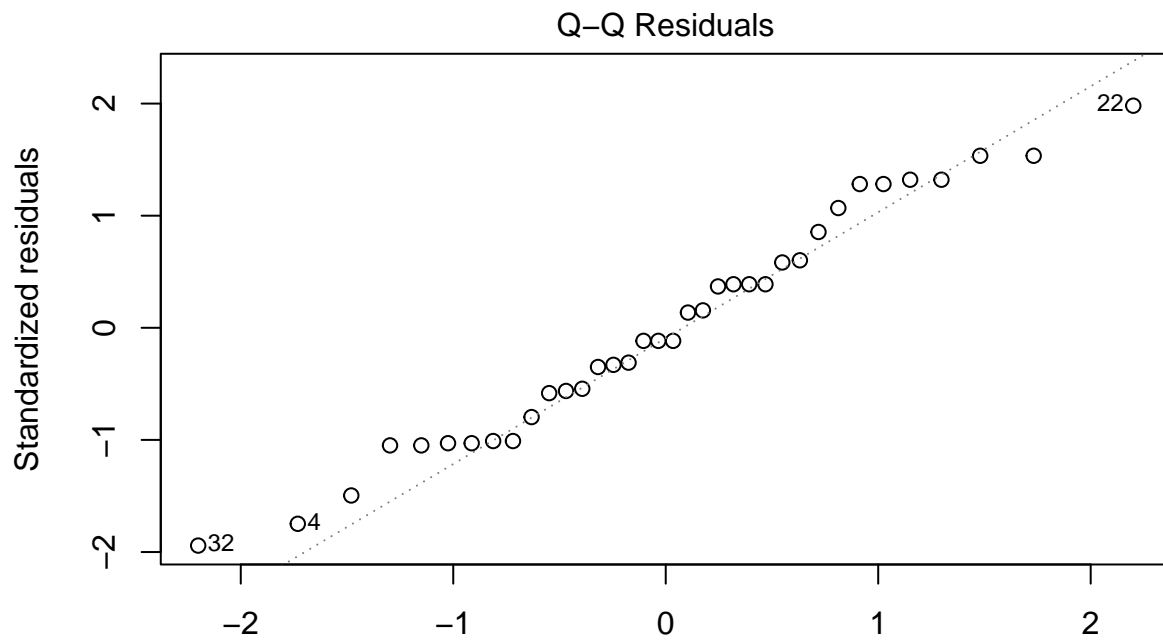
Se puede ver que el valor p es bastante bajo. Sería significativo para 0.05 como lo indica el modelo anova al imprimirlo con summary.

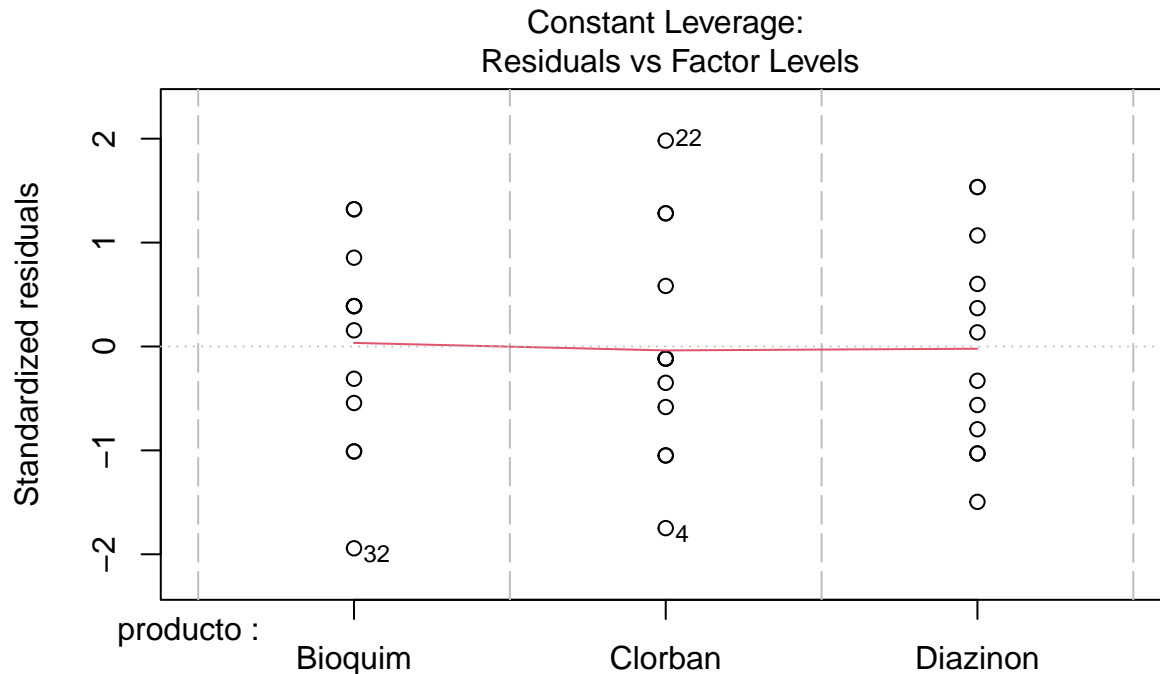
D. Comprobar los Supuestos del Modelo

En esta sección revisamos los supuestos del modelo, para ello primero utilizamos la función plot sobre el modelo para ver los gráficos de fitted vs Residuals y el gráfico QQ. Así también, hacemos una prueba Levene para verificar la igualdad de varianzas y un test Shapiro-Wilk para verificar la normalidad de los datos.

```
# Graficar los residuos del modelo
plot(modelo)
```







Factor Level Combinations

```
# Realizar la prueba de Levene para la igualdad de varianzas
```

```
library(car)
```

```
leveneTest(contador ~ producto, data = Insect)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
```

```
##      Df F value Pr(>F)
```

```
## group 2  0.0741 0.9287
```

```
##      33
```

```
# Realizar la prueba de Shapiro-Wilk para la normalidad de los residuos
```

```
shapiro.test(modelo$residuals)
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data:  modelo$residuals
```

```
## W = 0.97166, p-value = 0.4725
```

El valor de la pruebas llevan a no rechazar las hipotesis nulas de cada prueba, es decir, se cumple igualdad de varianzas y que los datos sean normales.

E. Realizar pruebas post hoc TukeyHSD y graficarlo

Realizamos pruebas post hoc para ver que grupos de tratamiento tienen diferencias entre sí. Para ello utilizamos TukeyHSD sobre el modelo anova ya ajustado. Además, graficamos el modelo TukeyHSD para observar los intervalos de confianza.

```
# Realizar la prueba de Tukey para comparaciones múltiples
```

```
TukeyHSD(modelo, conf.level = 0.95)
```

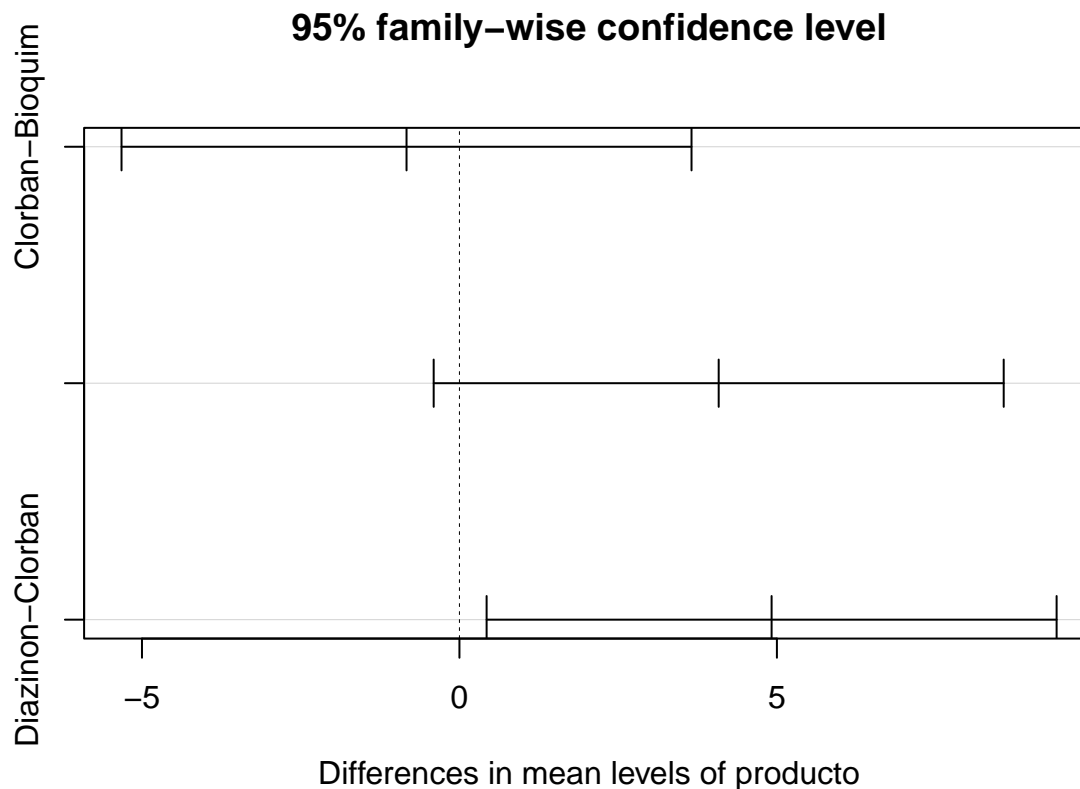
```
## Tukey multiple comparisons of means
```

```
## 95% family-wise confidence level
```

```
##
```

```
## Fit: aov(formula = contador ~ producto, data = Insect)
##
## $producto
##              diff          lwr          upr          p adj
## Clorban-Bioquim -0.8333333 -5.3220878  3.655421  0.8922890
## Diazinon-Bioquim  4.0833333 -0.4054211  8.572088  0.0804812
## Diazinon-Clorban  4.9166667  0.4279122  9.405421  0.0293305

# Visualizar los intervalos de confianza del 95%
plot(TukeyHSD(modelo, conf.level = 0.95))
```



Los valores p ajustados indican que con un nivel de confianza de 95%, Diazinon y Clorban muestran diferencias significativas entre sí.

E.1. Reportando Resultados

Exploración de los datos

- La media del conteo de insectos muertos es mayor para el insecticida Diazinon, seguido por Bioquim y Clorban.
- La desviación estándar es similar para los tres insecticidas.
- El diagrama de cajas muestra que la distribución del conteo de insectos muertos es similar para los tres insecticidas, aunque para bioquim el rango intercuartil es el menor, lo que sugiere menos variabilidad que los otros dos insecticidas.

ANOVA

- El valor p del ANOVA es menor que 0.05, lo que indica que existe una diferencia estadísticamente significativa entre las medias del conteo de insectos muertos para los tres insecticidas.

Supuestos del modelo

- El gráfico de residuos vs. valores ajustados muestra que los residuos se distribuyen de manera uniforme alrededor de 0, lo que sugiere que se cumple el supuesto de homocedasticidad.
- Sin embargo, la prueba de Levene para la igualdad de varianzas no es significativa ($p > 0.05$), lo que indica que se cumple el supuesto de homocedasticidad.
- El gráfico Q-Q muestra que los residuos siguen aproximadamente una línea recta, lo que sugiere que se cumple el supuesto de normalidad.
- La prueba de Shapiro-Wilk para la normalidad de los residuos no es significativa ($p > 0.05$), lo que indica que se cumple el supuesto de normalidad.

Pruebas post hoc

- La prueba de Tukey para comparaciones múltiples muestra que existen diferencias estadísticamente significativas entre las medias del conteo de insectos muertos para los tres insecticidas.
- Diazinon tiene una media significativamente mayor que Clorban y Bioquim.
- No hay diferencia significativa entre Clorban y Bioquim.

F. Calcular el tamaño del efecto

Calculamos el valor Eta Cuadrado utilizando la biblioteca lsr y la función etaSquared.

```
# Cargar el paquete lsr
library(lsr)

# Calcular el tamaño del efecto eta cuadrado

etaSquared(modelo, anova = TRUE)
```

	eta.sq	eta.sq.part	SS	df	MS	F	p
## producto	0.2005028	0.2005028	166.1667	2	83.08333	4.13797	0.02491609
## Residuals	0.7994972	NA	662.5833	33	20.07828	NA	NA

El efecto del insecticida es grande.

G. Calcular la potencia de la prueba con 0.05

```
# Calcular la potencia de la prueba
power.anova.test(groups = 3, n = 12, between.var = 134.22, within.var = 34.72, sig.level = 0.05)

##
##      Balanced one-way analysis of variance power calculation
##
##      groups = 3
##      n = 12
##      between.var = 134.22
##      within.var = 34.72
##      sig.level = 0.05
##      power = 1
##
## NOTE: n is number in each group
```

H. Determinar el tamaño de muestra necesario para una potencia de 0.8

```
# Determinar el tamaño de muestra necesario para una potencia de 0.8
power.anova.test(groups = 3, between.var = 134.22, within.var = 34.72, power = 0.8, sig.level = 0.05)
```

```
##
##      Balanced one-way analysis of variance power calculation
##
##      groups = 3
##      n = 2.550626
##      between.var = 134.22
##      within.var = 34.72
##      sig.level = 0.05
##      power = 0.8
##
## NOTE: n is number in each group
```

Tamaño del efecto

- El tamaño del efecto η^2 es 0.2, lo que indica un efecto grande.

Potencia de la prueba

- La potencia de la prueba es 0.999, lo que indica que hay una alta probabilidad de detectar una diferencia significativa entre las medias si realmente existe una diferencia.

Tamaño de muestra necesario para una potencia de 0.8

- Se necesita un tamaño de muestra de 3 por grupo para lograr una potencia de 0.8.

Conclusión

Los resultados del ANOVA de una vía indican que existe una diferencia estadísticamente significativa entre las medias del conteo de insectos muertos para los tres insecticidas. Las pruebas post hoc muestran que Diazinon es el insecticida el más efectivo. El tamaño del efecto es grande, lo que indica que el tipo de insecticida tiene un impacto significativo en el conteo de insectos muertos. La potencia de la prueba es alta, lo que indica que hay una alta probabilidad de detectar una diferencia significativa si realmente existe una diferencia. Se necesita un tamaño de muestra de 3 por grupo para lograr una potencia de 0.8.