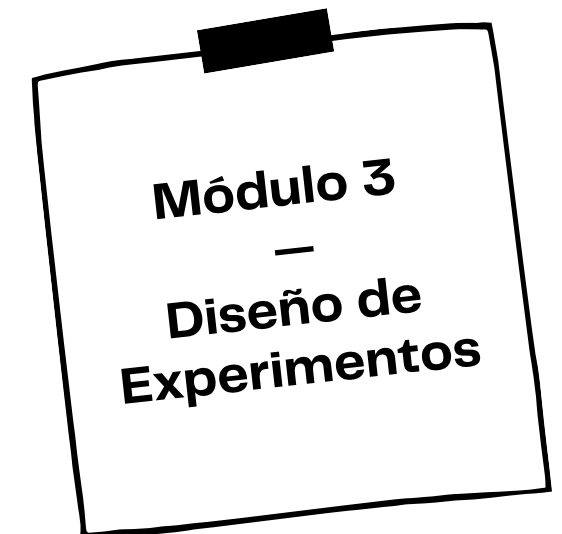


Diseño completamente aleatorizado



Q Agenda de hoy

- | | | | |
|---|--|---|-----------------------------------|
| 1 | Diseños para comparar dos o más tratamientos | 5 | Diagramas de caja |
| 2 | Diseño completamente aleatorizado | 6 | Comparaciones o pruebas múltiples |
| 3 | Análisis de Varianza (ANOVA) | 7 | Validación de los supuestos |
| 4 | Supuestos para el Análisis de Varianza | 8 | Elección del tamaño de muestra |

Diseños para comparar dos o más tratamientos

Recordemos que los diseños experimentales se clasifican en diferentes tipos dependiendo de su objetivo. Vamos a comenzar aprendiendo sobre los **diseños para comparar dos o más tratamientos**



Diseños para comparar dos o más tratamientos

- Comparar cuatro dietas alimenticias para determinar si alguna es mejor
- Comparar cuatro métodos de ensamble para determinar con cual se logra el menor tiempo de ensamble
- Comparar tres marcas diferentes de spray para matar insectos
- Comparar dos medicamentos para determinar cual actúa de forma más eficaz en el alivio del dolor de cabeza

Diseños para comparar dos o más tratamientos

Por lo general, en este tipo de experimentos, el interés del experimentador está centrado en comparar los tratamientos respecto a sus medias poblacionales. Por lo tanto, la hipótesis fundamental a probar cuando se comparan varios tratamientos es:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_K$$

H_a : No todas las medias poblacionales son iguales

¡Terminología!

Población: colección o totalidad de posibles individuos u objetos sobre los que se lleva a cabo un estudio.

Finita o pequeña



Es posible medir todos los individuos y tener un conocimiento exacto de las características (**parámetros**) de la población

¡Terminología!

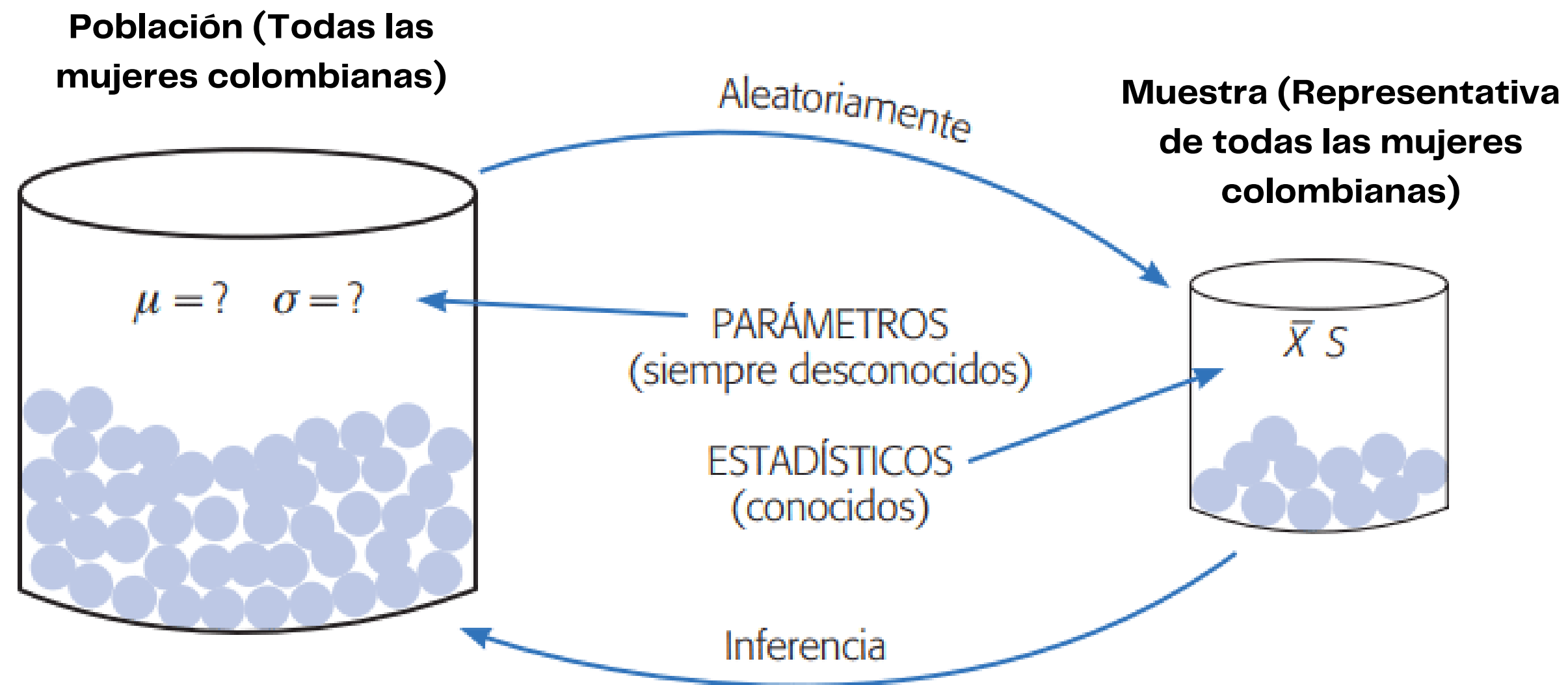
Población: colección o totalidad de posibles individuos u objetos sobre los que se lleva a cabo un estudio.

Infinita o grande

2

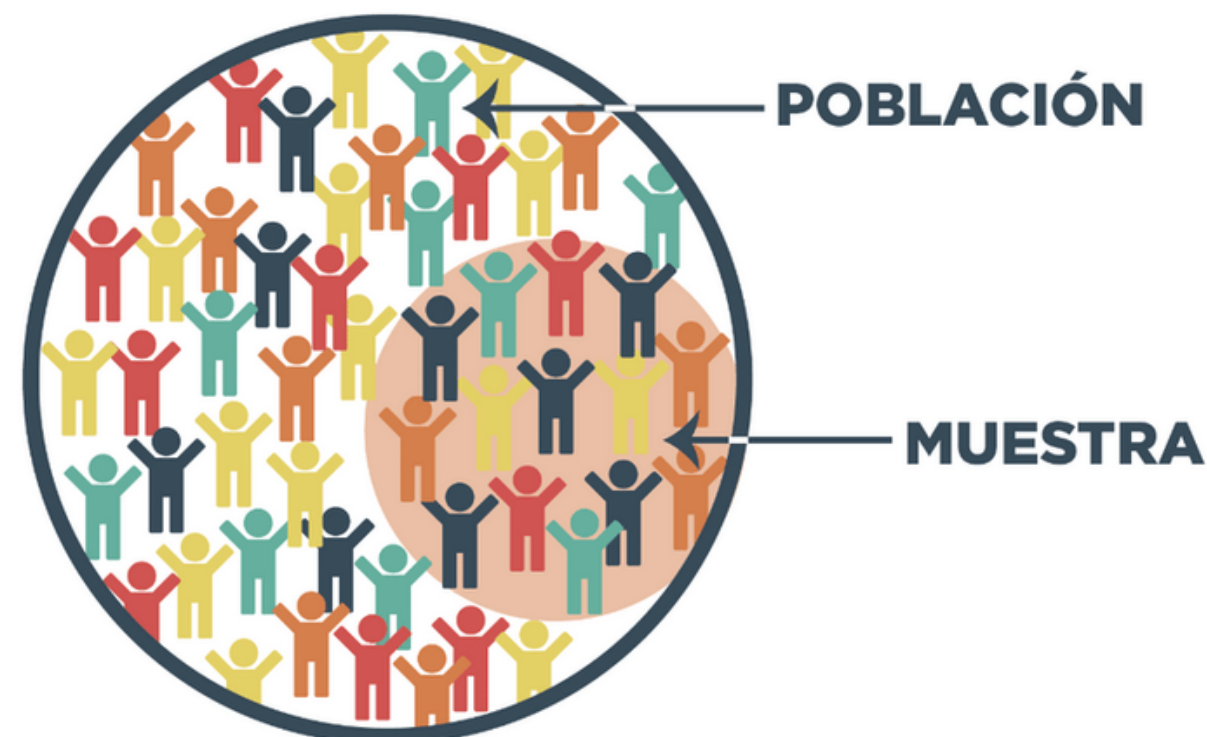
Es imposible medir a todos los individuos y por lo tanto, se debe extraer una muestra representativa de la población y con base en sus características (**estadísticos**) se pueden hacer afirmaciones acerca de los parámetros de la población.

¡Terminología!



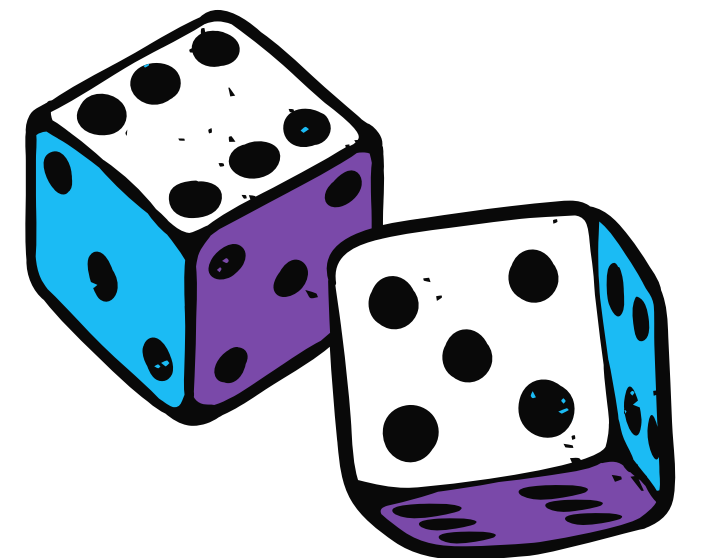
¡Terminología!

Estadística inferencial: consiste en hacer afirmaciones acerca de la población con base en la información de una muestra y se divide en dos procesos principales: **estimación** y **prueba de hipótesis**.



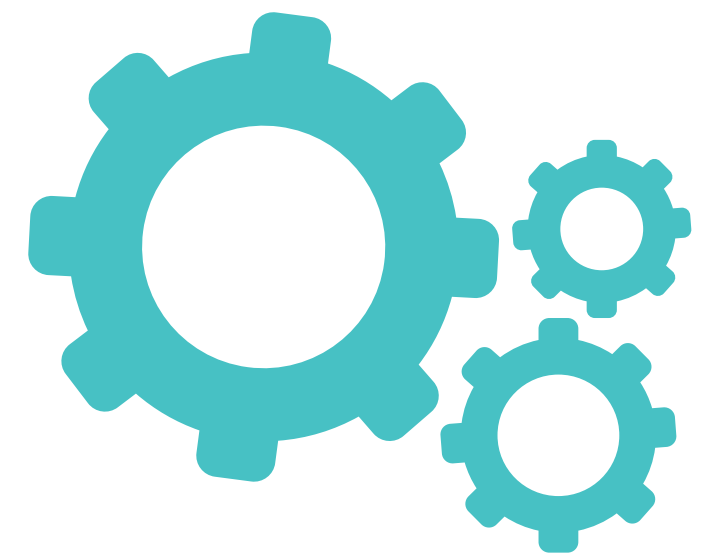
Diseño completamente aleatorizado

El diseño completamente aleatorizado es frecuentemente utilizado para comparar dos o más tratamientos. Con este diseño, todas las corridas experimentales o pruebas se realizan en orden aleatorio y solo se considera un factor que afecta a la variable respuesta



Ejemplo: comparación de métodos de ensamble

Un equipo de mejora continua de una compañía investiga el efecto de cuatro **métodos de ensamble** A, B, C y D de un producto sobre el **tiempo de ensamble** en minutos



Ejemplo: comparación de métodos de ensamble

1

Objetivo: Determinar a través de un experimento, cuál de los cuatro métodos de ensamble A, B, C y D tiene un mayor efecto sobre el tiempo de ensamble del producto.

2

Variable respuesta: Tiempo de ensamble de producto en minutos.

3

Factores a estudiar: Método de ensamble.

Ejemplo: comparación de métodos de ensamble

4 Niveles para los factores a estudiar: A, B, C y D.

5 Diseño experimental: diseño completamente aleatorizado. Se supone que, además del método de ensamble, no existe otro factor que influya de manera significativa sobre el tiempo de ensamble.

6 Número de repeticiones para los diferentes niveles (tamaño de muestra): 4 repeticiones de cada nivel.

Ejemplo: comparación de métodos de ensamble

IMPORTANTE:



- El número de tratamientos **k** es determinado por el investigador y depende del problema que se está estudiando.
 - El número de observaciones o repeticiones por tratamiento **n** depende de la dispersión que muestren las mediciones (en la práctica se recomiendan entre 5 y 30 observaciones de cada tratamiento).
- También es importante tener en cuenta costos y tiempo.

Ejemplo: comparación de métodos de ensamble

Para determinar el orden en el que se llevarán a cabo las corridas experimentales y asegurar que se hagan de forma aleatoria se puede usar la función **sample()** incluida en el paquete base de R

*# corridas experimentales = # de niveles * # de repeticiones para cada nivel*

*# corridas experimentales = 4 * 4 = 16*

D C A B A D A B C D B A D C B C

Ejemplo: comparación de métodos de ensamble

Supongamos que el experimento se lleva a cabo y estos son los resultados obtenidos:

Método de ensamble			
A	B	C	D
6	7	11	10
8	9	16	12
7	10	11	11
8	8	13	9

Diseño balanceado:
mismo número
de repeticiones
en cada
tratamiento

Análisis de varianza (ANOVA)

El análisis de varianza (ANOVA) es la técnica central para el análisis de datos obtenidos mediante un experimento. Este procedimiento estadístico nos permite determinar si las diferencias observadas entre las medias muestrales de cada uno de los tratamientos son suficientemente grandes para rechazar la hipótesis nula planteada

Análisis de varianza (ANOVA)

Suposiciones para el análisis de varianza

- **La variable respuesta Y tiene una distribución normal:** El tiempo de ensamble tiene una distribución normal para todos los tratamientos (métodos de ensamble).
- **La varianza de la variable respuesta es constante:** la varianza en el tiempo de ensamble es la misma en cada tratamiento.
- **Las mediciones deben ser independientes:** el tiempo de ensamble del método A es independiente al tiempo de ensamble de los demás métodos.

Análisis de varianza (ANOVA)

Usemos la técnica del análisis de varianza para probar la igualdad de los tiempos promedio de los métodos de ensamble. **Esta prueba nos permitirá determinar si existe un efecto del método de ensamble en el tiempo de ensamble**

Análisis de varianza (ANOVA)

Prueba F

1 Definir las hipótesis

$$H_0: \mu_A = \mu_B = \mu_C = \mu_D$$

H_a: No todos los tiempos promedio de ensamble son iguales

Análisis de varianza (ANOVA)

Prueba F

2 Definir el estadístico de prueba

$$F = \frac{CM_{TRAT}}{CM_E}$$

Análisis de varianza (ANOVA)

Prueba F

3

Definir la regla de rechazo

Valor P

Rechazo H_0 si:

$$valor\ p \leq \alpha$$

Análisis de varianza (ANOVA)

Notación

Y_{ij} = j – ésima observación en el tratamiento i

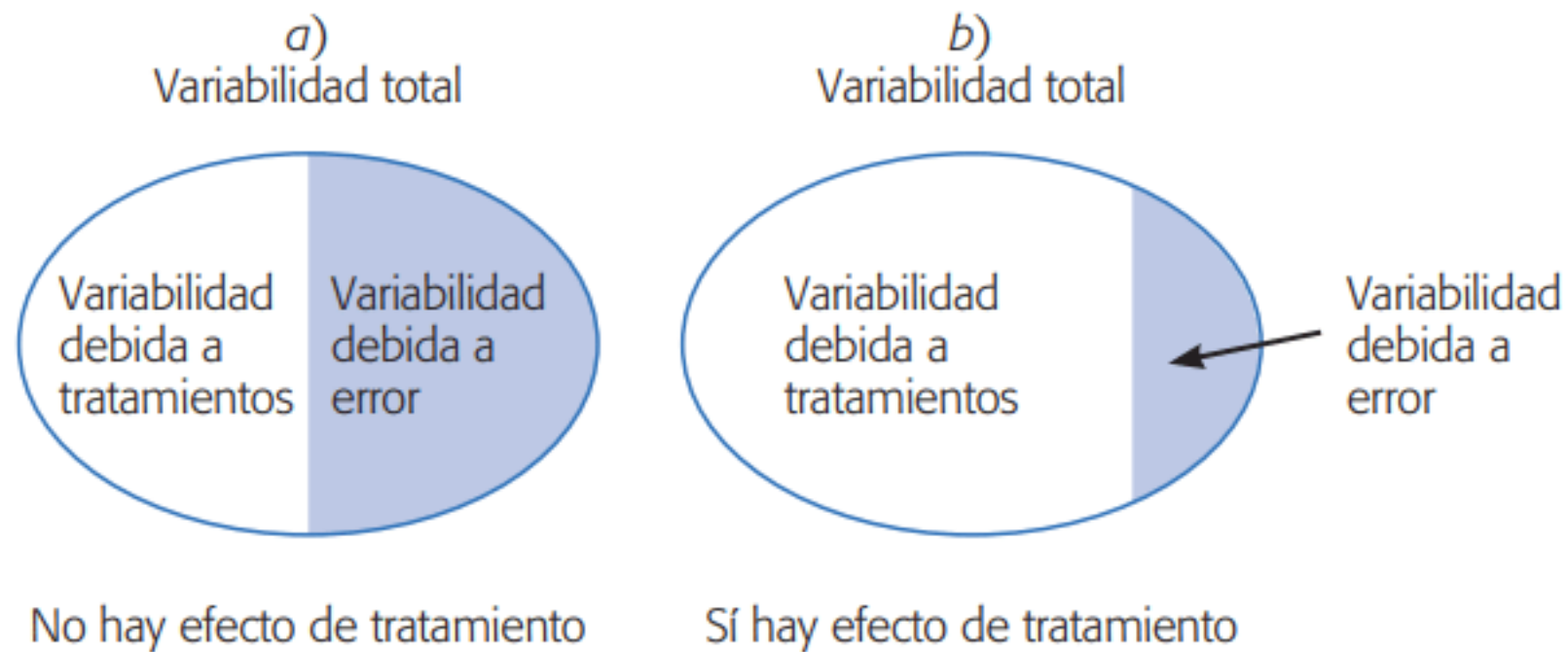
$Y_{i.}$ = Suma de las observaciones del tratamiento i

$\bar{Y}_{i.}$ = Promedio de las observaciones del tratamiento i

$Y_{..}$ = Suma de todas las observaciones

$\bar{Y}_{..}$ = Promedio de todas las observaciones

Análisis de varianza (ANOVA)



Suma de cuadrados de tratamientos

$$SC_{TRAT} = \sum_{i=1}^k n_i (\bar{Y}_{i.} - \bar{Y}_{..})^2$$

Suma de cuadrados del error

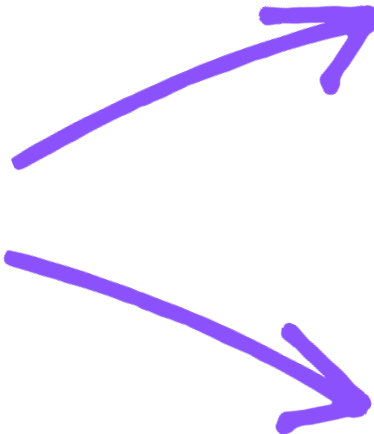
$$SC_E = \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i.})^2$$

Análisis de varianza (ANOVA)

Suma de cuadrados total

$$SC_T = SC_{TRAT} + SC_E$$

Análisis de varianza (ANOVA)

$$F = \frac{CM_{TRAT}}{CM_E}$$

$$CM_{TRAT} = \frac{SC_{TRAT}}{k - 1}$$
$$CM_E = \frac{SC_E}{N - k}$$

Análisis de varianza (ANOVA)

Si llevamos a cabo estos cálculos, esta es la ANOVA resultante:

	Suma de cuadrados	Grados de libertad	Cuadrados medios	F	Valor-p
Tratamientos	69.5	3	23.17	9.42	0.0018
Error	29.5	12	2.46		
Total	99.0	15			

Análisis de varianza (ANOVA)

Utilicemos R para llevar a cabo el análisis de varianza....

Análisis de varianza (ANOVA)

Conclusión: Como el valor- p es menor al nivel de significancia de 0.05, se rechaza la hipótesis nula y se concluye que al menos dos de los tiempos promedio de ensamble son diferentes

Diagramas de cajas

Una forma de comparar descriptivamente los tratamientos es mediante diagramas de caja

Diagramas de cajas

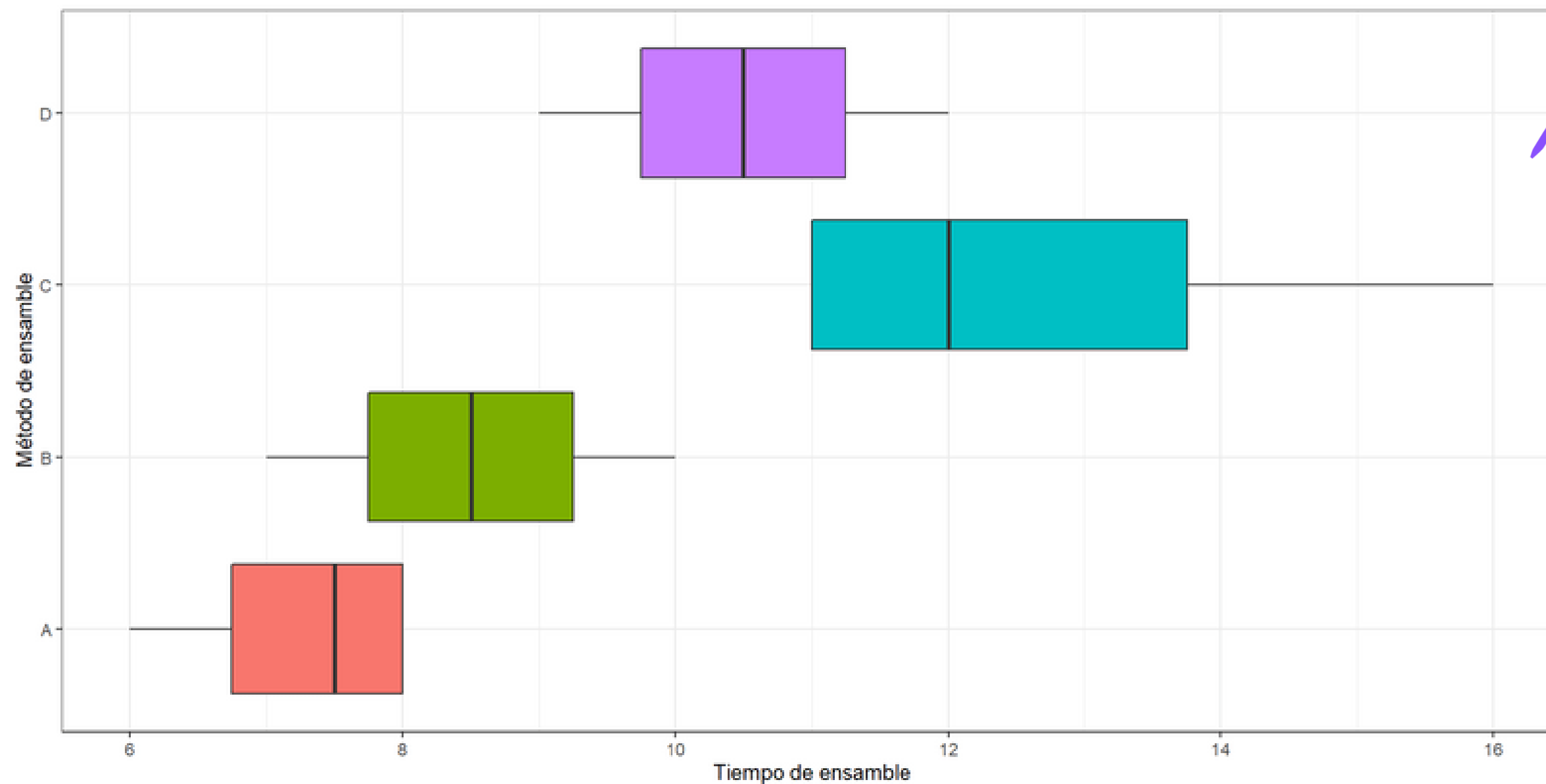


Diagrama de
caja

Diagramas de cajas

- Cuando los diagramas no se traslapan, es probable que los tratamientos sean diferentes entre sí (entre mayor sea el número de datos, mayor es esta probabilidad).
- Cuando se traslapan un poco, es probable que haya o no diferencias significativas y por lo tanto, se debe llevar a cabo una prueba estadística para determinar si en efecto hay diferencias significativas.


Comparaciones o pruebas múltiples

Cuando se emplea el análisis de varianza, solo es posible concluir si las medias poblacionales son iguales o no. Sin embargo, en muchos casos es necesario saber cuáles tratamientos son diferentes

Comparaciones o pruebas múltiples

Método LSD (Diferencia Mínima Significativa)

El objetivo de esta prueba es probar la igualdad de todos los **posibles pares de medias**.


$$k(k - 1)/2$$

LSD: Diferencia mínima significativa para considerar que dos tratamientos son diferentes

Comparaciones o pruebas múltiples

Método LSD (Diferencia Mínima Significativa)

- 1 Definir la hipótesis

$$H_0: \mu_i = \mu_j$$

$$H_a: \mu_i \neq \mu_j$$

Comparaciones o pruebas múltiples

Método LSD (Diferencia Mínima Significativa)

2

Definir el estadístico de prueba

$$t = \frac{\bar{Y}_{i.} - \bar{Y}_{j.}}{\sqrt{CME \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}}$$

Comparaciones o pruebas múltiples

Método LSD (Diferencia Mínima Significativa)

3

Definir la regla de rechazo

Valor P

Rechazo H_0 si:

$$valor\ p \leq \alpha$$

Comparaciones o pruebas múltiples

Método LSD (Diferencia Mínima Significativa)

Otra forma de rechazar la hipótesis nula...

Rechazo H_0 si:

$$|\bar{Y}_{i.} - \bar{Y}_{j.}| \geq LSD$$



$$LSD = t_{\frac{\alpha}{2}, n-k} \sqrt{CME \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

Pruebas múltiples en nuestro experimento

$$H_0 : \mu_A = \mu_B \text{ vs. } H_A : \mu_A \neq \mu_B$$

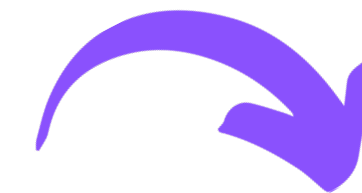
$$H_0 : \mu_A = \mu_C \text{ vs. } H_A : \mu_A \neq \mu_C$$

$$H_0 : \mu_A = \mu_D \text{ vs. } H_A : \mu_A \neq \mu_D$$

$$H_0 : \mu_B = \mu_C \text{ vs. } H_A : \mu_B \neq \mu_C$$

$$H_0 : \mu_B = \mu_D \text{ vs. } H_A : \mu_B \neq \mu_D$$

$$H_0 : \mu_C = \mu_D \text{ vs. } H_A : \mu_C \neq \mu_D$$



Posibles pares
de hipótesis

Pruebas múltiples en nuestro experimento

$$LSD = 2.42$$

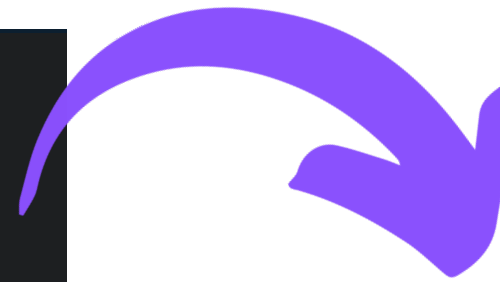
Prueba	Diferencia muestras de tiempos promedio	Decisión
$H_0: \mu_A = \mu_B$	1.25	No significativa
$H_0: \mu_A = \mu_C$	5.50	Significativa
$H_0: \mu_A = \mu_D$	3.25	Significativa
$H_0: \mu_B = \mu_C$	4.25	Significativa
$H_0: \mu_B = \mu_D$	2.00	No significativa
$H_0: \mu_C = \mu_D$	2.25	No significativa

Pruebas múltiples en nuestro experimento

Utilicemos R para llevar a cabo estas pruebas....

Pruebas múltiples en nuestro experimento

```
$groups
  tiempo groups
C   12.75    a
D   10.50   ab
B    8.50   bc
A    7.25    c
```



Los tratamientos que tienen diferentes caracteres en la columna grupos tienen medias significativamente diferentes

Pruebas múltiples en nuestro experimento

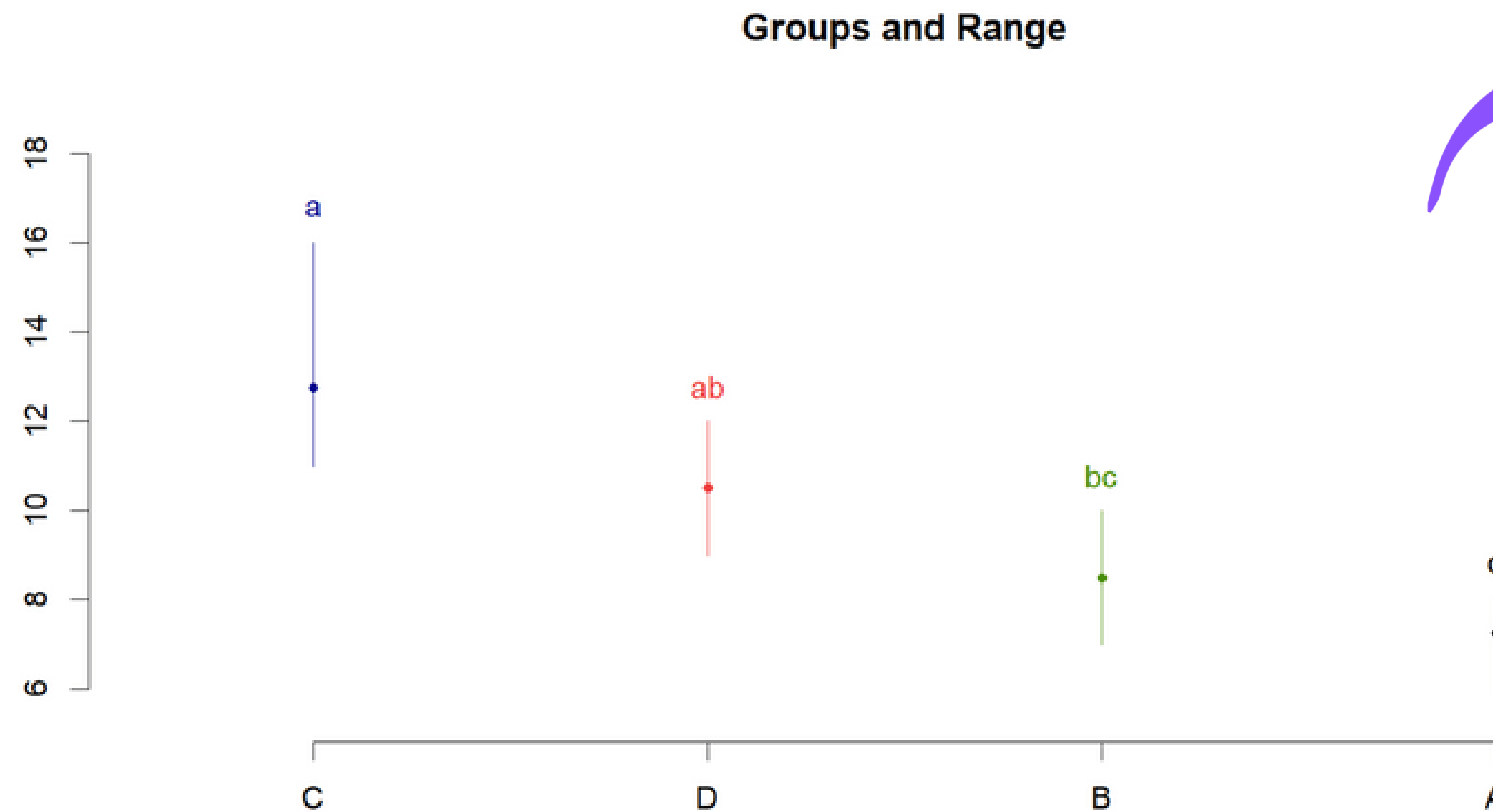
Intervalos de confianza para la media de los tratamientos con el método LSD

$$\bar{Y}_{i\cdot} \pm t_{\alpha/2, N-k} \sqrt{\frac{CM_E}{n_i}}$$

Gráfico de medias

Un gráfico de medias (means plot) permite hacer una comparación visual de las medias de los tratamientos. Este gráfico se construya a partir de los intervalos de confianza que se calculan con la prueba LSD

Gráfico de medias



Si los intervalos de confianza de dos tratamientos se traslapan, los tratamientos son estadísticamente iguales en cuanto a sus medias. En el caso contrario, son diferentes

Conclusiones del experimento

- A partir del experimento que se llevó a cabo y el análisis que se realizó de los datos obtenidos, se puede concluir que el método que logra el mejor desempeño es el método A ya que estadísticamente sus tiempos promedio de ensamble son menores a los de los métodos C y D.
- Por otro lado, al método A le sigue el método B ya que estadísticamente sus tiempos promedio de ensamble son menores que los del método C. Sin embargo, no es posible concluir si el método A es mejor que el B ya que su diferencia no es significativa estadísticamente.

Practiquemos

Se desea realizar un estudio sobre la efectividad de tres marcas diferentes de veneno para moscas. Para esto se lleva a cabo un experimento donde se aplica a un grupo de 100 moscas el producto y se cuenta el número de moscas muertas. Se repite el experimento 6 veces para cada veneno y estos son los resultados:

Marca de veneno	Repeticiones					
	1	2	3	4	5	6
1	72	65	67	75	62	73
2	55	59	68	70	53	50
3	64	74	61	58	51	69

Practiquemos

1. ¿Existe diferencia entre el número promedio de moscas que mata cada uno de los venenos?
2. ¿Hay alguna marca que sea mejor?
3. Construya un intervalo de confianza para el número de moscas muertas de cada una de las marcas
4. Dibuje las gráficas de medias y los diagramas de cajas e interpréte los.