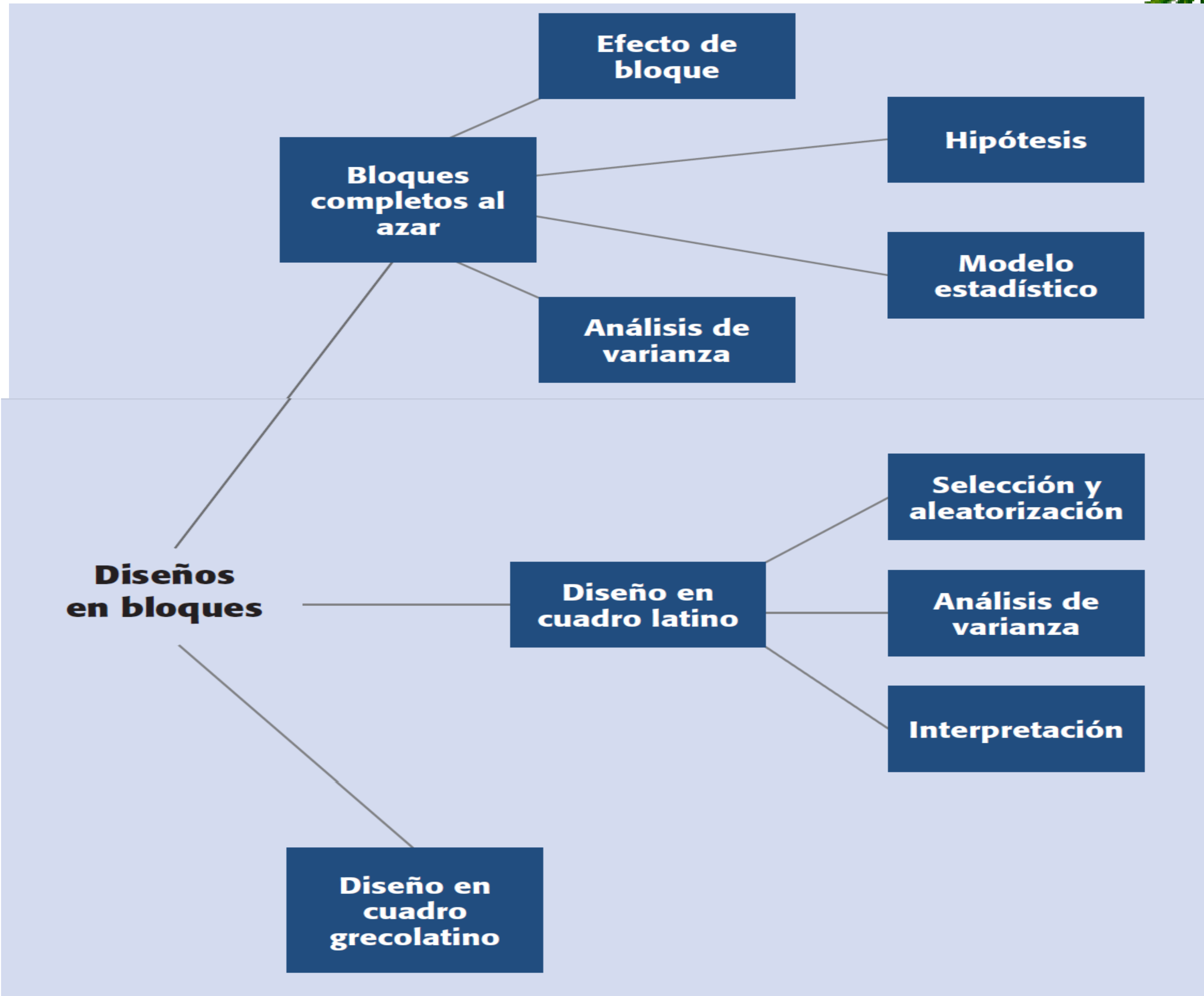

Diseños en bloques
Bloques completos al azar
Efecto de bloque
Hipótesis
Modelo estadístico
Análisis de varianza
Diseño en cuadro latino
Selección y aleatorización
Análisis de varianza
Interpretación
Diseño en cuadro grecolatino

Factores de bloque

- Son las variables adicionales al factor de interés que se incorporan de manera explícita en un experimento comparativo para **no sesgar la comparación.**
- Por ejemplo, en el caso de comparar cuatro máquinas que son manejadas por cuatro operadores, es pertinente incluir explícitamente al factor operadores (bloques) para lograr el propósito del estudio.

- En un diseño en bloques completos al azar (DBCA) se consideran tres fuentes de variabilidad: el factor de tratamientos, el factor de bloque y el error aleatorio.
- Los factores de bloqueo que aparecen en la práctica son: turno, lote, día, tipo de material, línea de producción, operador, máquina, método, etc.

DBCA

Es un diseño de dos factores, uno denominado **factor tratamiento** y el otro **factor bloque**. Consideramos el diseño sin repetición, es decir para cada bloque y tratamiento sólo hay una observación. Ambos factores son de efectos fijos.

Términos técnicos

- **Respuesta**
  - Variable cuantitativa métrica dependiente, que es afectada por los factores en el proceso.
  - Ejemplos: Rendimiento, peso, precio de un producto, etc.

- **Factor Tratamiento**
  - Variable categórica independiente del que interesa conocer su influencia en la respuesta.

- **Factor Bloque**
  - Variable categórica independiente en la que no se está interesado en conocer su influencia en la respuesta pero se supone que ésta existe y se quiere controlar para disminuir la variabilidad residual (es decir, eliminar su efecto en el análisis).

- **Tratamientos**
  - Son los diferentes niveles o categorías del factor tratamiento. Son, por tanto, las condiciones experimentales que se desean comparar en el experimento.

Fuente: Ulima, Probabilidad y estadistica II, https://www.slideshare.net/walterperez716/estadistica-1-62033559

Modelo estadístico de un DBCA

- **Modelo teórico:**

$$Y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij} = \mu_{ij} + \varepsilon_{ij} \qquad i = 1,2,\ldots,b \quad j = 1,2,\ldots,a$$

Y_{ij} = *respuestas del j - ésimo tratamiento para el i - ésimo bloque*

b = *número de bloques* $\quad$ a = *número de tratamientos*

$\mu \quad \beta_i \quad \tau_j$

Efecto del tratamiento j en la respuesta

Efecto del bloque i en la respuesta

Media poblacional de la variable respuesta

ε_{ij} = *error aleatorio por observación*

Variación causada por los factores no controlados

Fuente: Ulima, Probabilidad y estadistica II, https://www.slideshare.net/walterperez716/estadistica-1-62033559

2. Formulación de hipótesis

- **Se prueba la igualdad de los efectos de los tratamientos.**

$$H_0 : \tau_j = 0 \quad j = 1,2,\ldots,a \quad (\mu_{\circ 1} = \mu_{\circ 2} = \ldots = \mu_{\circ a})$$

(El factor tratamiento NO influye sobre la respuesta)

$$H_a : \tau_j \neq 0 \quad \text{algún } j \quad (\mu_{\circ j} \neq \mu_{\circ k} \text{ algún } j \neq k)$$

(El factor tratamiento SI influye sobre la respuesta)

- **Prueba de importancia del bloqueo.**

$$H_0 : \beta_i = 0 \quad i = 1,2,\ldots,b \quad (\mu_{1\circ} = \mu_{2\circ} = \ldots = \mu_{b\circ})$$

(El factor bloque NO influye sobre la respuesta)

$$H_a : \beta_i \neq 0 \quad \text{algún } i \quad (\mu_{i\circ} \neq \mu_{k\circ} \text{ algún } i \neq k)$$

(El factor bloque SI influye sobre la respuesta)

OJO. Si se rechaza Ho, entonces hay evidencia de que el factor bloque fue importante en el diseño.

Fuente: Ulima, Probabilidad y estadistica II, https://www.slideshare.net/walterperez716/estadistica-1-62033559

Supuestos del modelo DBCA

- **Los efectos de tratamientos y bloques son aditivos. Es decir no hay interacción entre tratamientos y bloques.**
- **Los datos deben ser descritos adecuadamente por el modelo:**

$$Y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij}$$

donde:

$$\varepsilon_{ij} \rightarrow N(0, \sigma^2)$$

- **Los errores deben satisfacer lo siguiente:**
  - Son v.a con distribución normal con media cero (**normalidad** de residuos)
  - Tienen varianza constante (**homocedásticidad**)
  - Son variables aleatorias independientes (**no autocorrelación** de errores)

Fuente: Ulima, Probabilidad y estadistica II, https://www.slideshare.net/walterperez716/estadistica-1-62033559

3. Obtener Suma de Cuadrados de Bloques

$$SC\ Bloques = \sum_{j=1}^{r} Y_{.j}^{2}/t - F.C. \qquad Y_{.j} = \sum_{i=1}^{t} y_{ij}$$

4. Obtener Suma de Cuadrados de Tratamientos

$$SC\ Trat = \frac{\sum_{i=1}^{t} Y_{i.}^{2}}{r} - F.C. \qquad Y_{i.} = \sum_{j=1}^{r} y_{ij}$$

5. Obtener Suma de Cuadrados del Error

SC Error=SC Total- SC Bloques- SC Tratamientos

6. Obtener Grados de Libertad

g.l. Tratamientos = t-1 g.l. Bloques = r-1

g.l. Error = (t-1) x (r-1) g.l. Total = (t x r)-1

7. Obtener Cuadrados Medios

$$CMBloques = \frac{SCBloques}{g.l.bloques}$$

$$CMTrat = \frac{SCTrat}{g.l.Trat}$$

$$CMError = \frac{SC\ Error}{g.l.\ Error}$$

8. Obtener Valores de F

$$FBloques = \frac{CM\ Bloques}{CM\ Error}$$

$$Ftrat = \frac{CM\ trat}{CM\ Error}$$

9. Obtener el Coeficiente de Variación

$$CV(\%) = \left(\frac{\sqrt{CM\ Error}}{\bar{Y}_{..}}\right) \times 100 \quad \bar{Y}_{..} = Media\ general$$

10. Obtener el coeficiente de determinación R^2

$$R^2 = \frac{SC\ Bloques\ + SC\ Tratamientos}{SC\ Total}$$

- R^2 indica la proporción de la suma de cuadrados total que es explicada por la variación entre bloques y entre tratamientos.
- Conforme el valor de R^2 se aproxima a 1.0 esto indicará que los datos analizados tuvieron un mejor ajuste del modelo lineal.

Fuente: https://slideplayer.es/slide/10510003/

Tabla 4.2 ANOVA para un diseño en bloques completos al azar.

Fuente de variabilidad	Suma de cuadrados	Grado de libertad	Cuadrado medio	F_0	Valor-p
Tratamientos	SC_{TRAT}	$k-1$	CM_{TRAT}	$F_0 = \frac{CM_{TRAT}}{CM_E}$	$P(F > F_0)$
Bloques	SC_B	$b-1$	CM_B	$F_0 = \frac{CM_B}{CM_E}$	$P(F > F_0)$
Error	SC_E	$(k-1)(b-1)$	CM_E		
Total	SC_T	$N-1$			

Ejemplo 4.1

En el ejemplo 3.1, donde se planteó la comparación de cuatro métodos de ensamble, ahora se va a controlar activamente en el experimento a los operadores que realizarán el ensamble, lo que da lugar al siguiente diseño en bloques completos al azar.

Método	Operador			
	1	2	3	4
A	6	9	7	8
B	7	10	11	8
C	10	16	11	14
D	10	13	11	9

Recordemos que la variable de respuesta son los minutos en que se realiza el ensamble. Para comparar los cuatro métodos se plantea la hipótesis:

$$H_0 : \mu_A = \mu_B = \mu_C = \mu_D = \mu$$
$$H_A : \mu_i \neq \mu_j \text{ para algún } i \neq j = A, B, C, D$$

- **Modelo teórico:**

$$Y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij} = \mu_{ij} + \varepsilon_{ij} \qquad i = 1,2,\ldots,b \quad j = 1,2,\ldots,a$$

Y_{ij}: variable respuesta, el tiempo (min) de ensamblaje.

μ: es la media de minutos de ensamble.

β: factor bloque (operador)

τ: factor del tratamiento (método de ensamble)

ε: error experimental

	Operador				
Método	**1**	**2**	**3**	**4**	**Total por tratamiento**
A	6	9	7	8	$Y_{1\bullet} = 30$
B	7	10	11	8	$Y_{2\bullet} = 36$
C	10	16	11	14	$Y_{3\bullet} = 51$
D	10	13	11	9	$Y_{4\bullet} = 43$
Total	$Y_{\bullet 1} = 33$	$Y_{\bullet 2} = 48$	$Y_{\bullet 3} = 40$	$Y_{\bullet 4} = 39$	Total global $Y_{\bullet\bullet} = 160$

Con estos totales las sumas de cuadrados se obtienen fácilmente como:

$$
\begin{aligned}
SC_T &= \sum_{j=1}^{b} \sum_{i=1}^{k} Y_{ij}^2 - \frac{Y_{\bullet\bullet}^2}{N} = \left(6^2 + 7^2 + \ldots + 9^2\right) - \frac{160^2}{16} = 108 \\
SC_{TRAT} &= \sum_{i=1}^{k} \frac{Y_{i\bullet}^2}{b} - \frac{Y_{\bullet\bullet}^2}{N} = \frac{30^2 + 36^2 + 51^2 + 43^2}{4} - \frac{160^2}{16} = 61.5
\end{aligned} \tag{4.5}
$$

Tabla 4.3 ANOVA para el ejemplo 4.1.

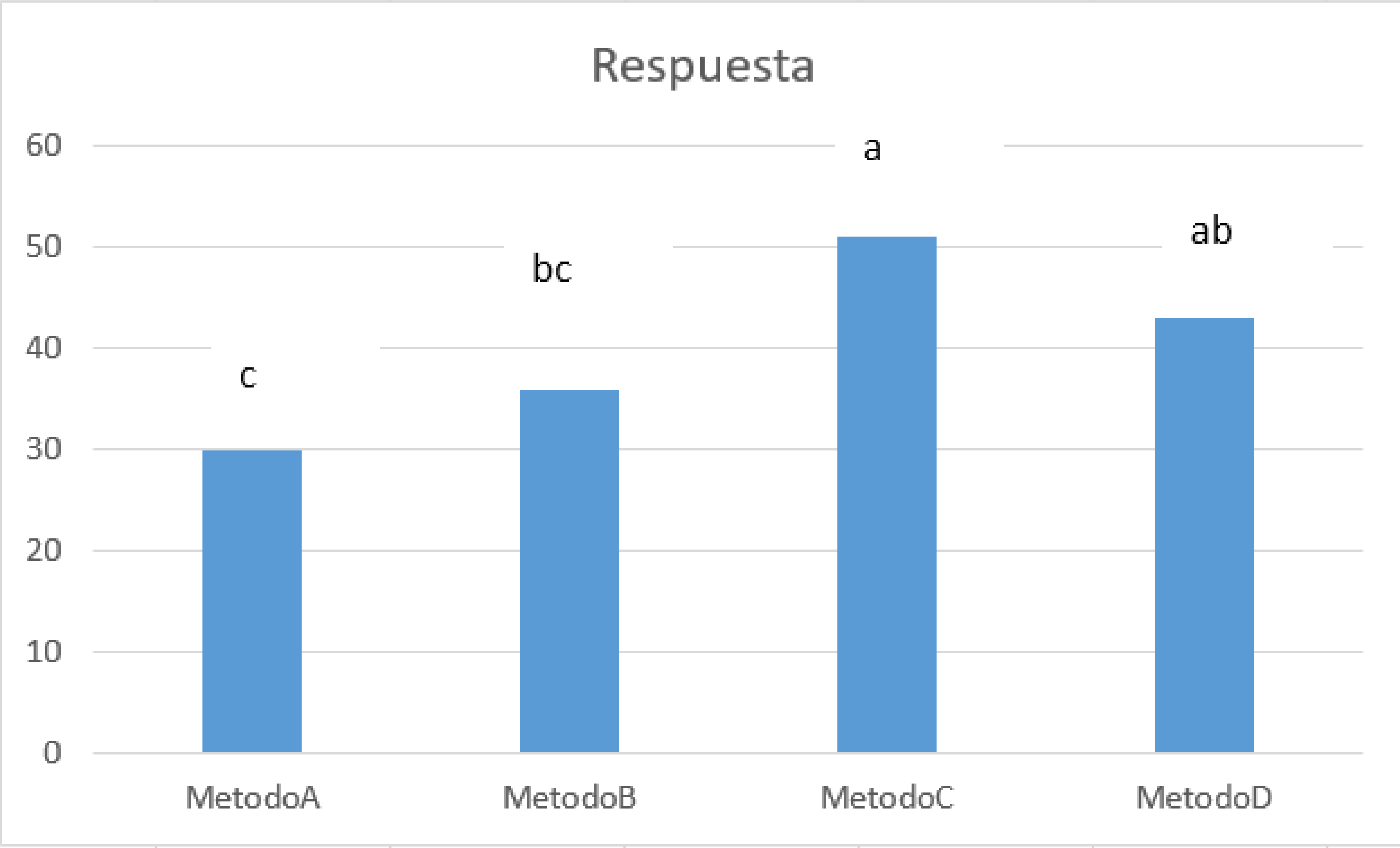
Fuente de variabilidad	Suma de cuadrados	Grado de libertad	Cuadrado medio	F_0	Valor-*p*
Métodos	61.5	3	20.5	10.25	0.003
Operadores	28.5	3	9.5	4.75	0.030
Error	18.0	9	2.0		
Total	108.0	15			

donde b es el número de bloques, que hace las veces de número de réplicas, y $(k - 1)(b - 1)$ son los grados de libertad del CM_E. De aquí que en el ejemplo que nos ocupa, como $t_{0.025, 9} = 2.26$, entonces,

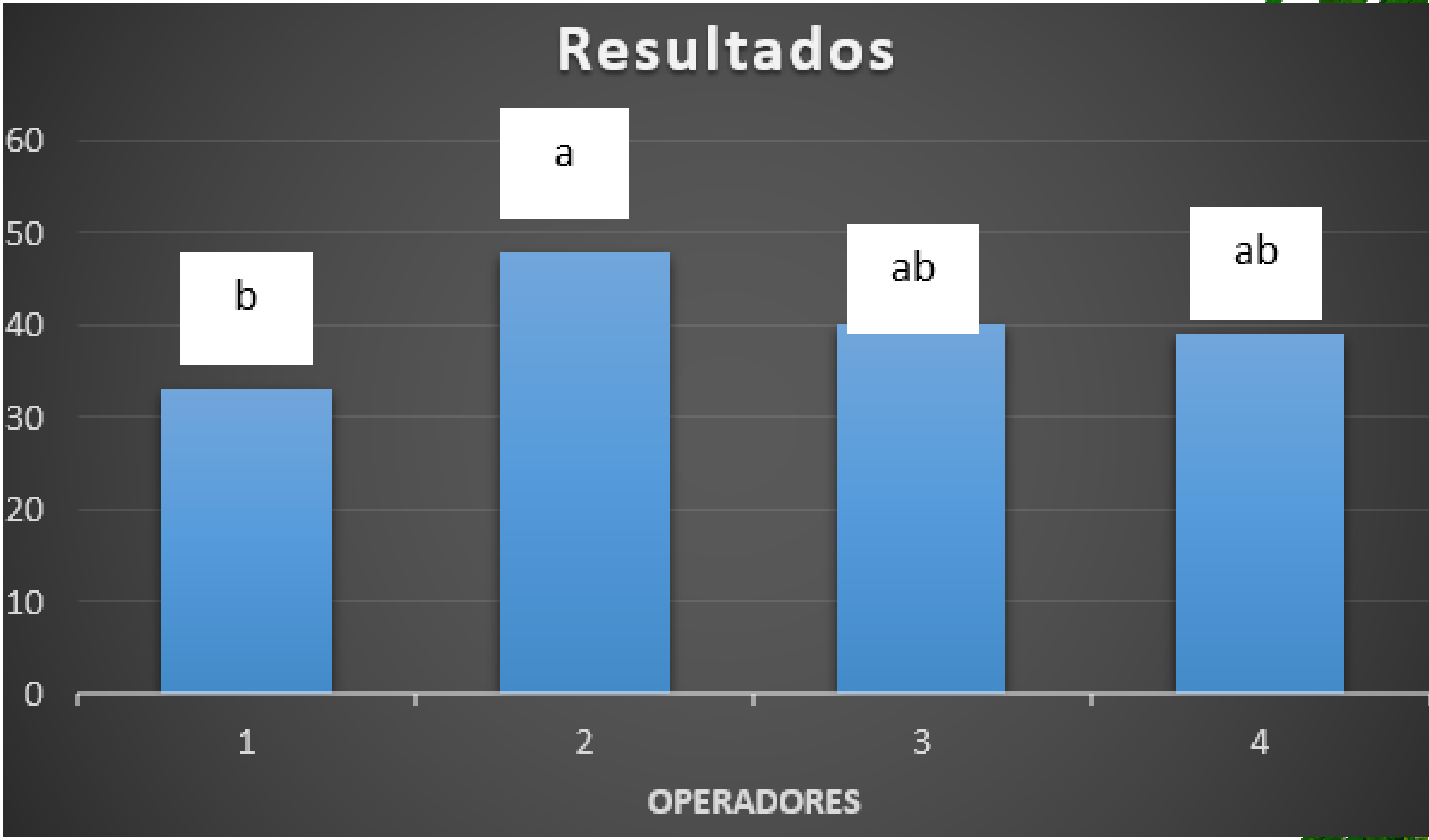
$$LSD = 2.26\sqrt{2 \times 2 / 4} = 2.26$$

Al comparar esta diferencia mínima significativa con los datos del ejemplo 4.1 se obtiene la siguiente tabla:

Diferencia poblacional	Diferencia muestral	Decisión
$\mu_A - \mu_B$	$\lvert -1.5 \rvert < 2.26$	No significativa
$\mu_A - \mu_C$	$\lvert -5.25 \rvert > 2.26$	Significativa
$\mu_A - \mu_D$	$\lvert -3.25 \rvert > 2.26$	Significativa
$\mu_B - \mu_C$	$\lvert -3.75 \rvert > 2.26$	Significativa
$\mu_B - \mu_D$	$\lvert -1.75 \rvert < 2.26$	No significativa
$\mu_C - \mu_D$	$2.00 < 2.26$	No significativa

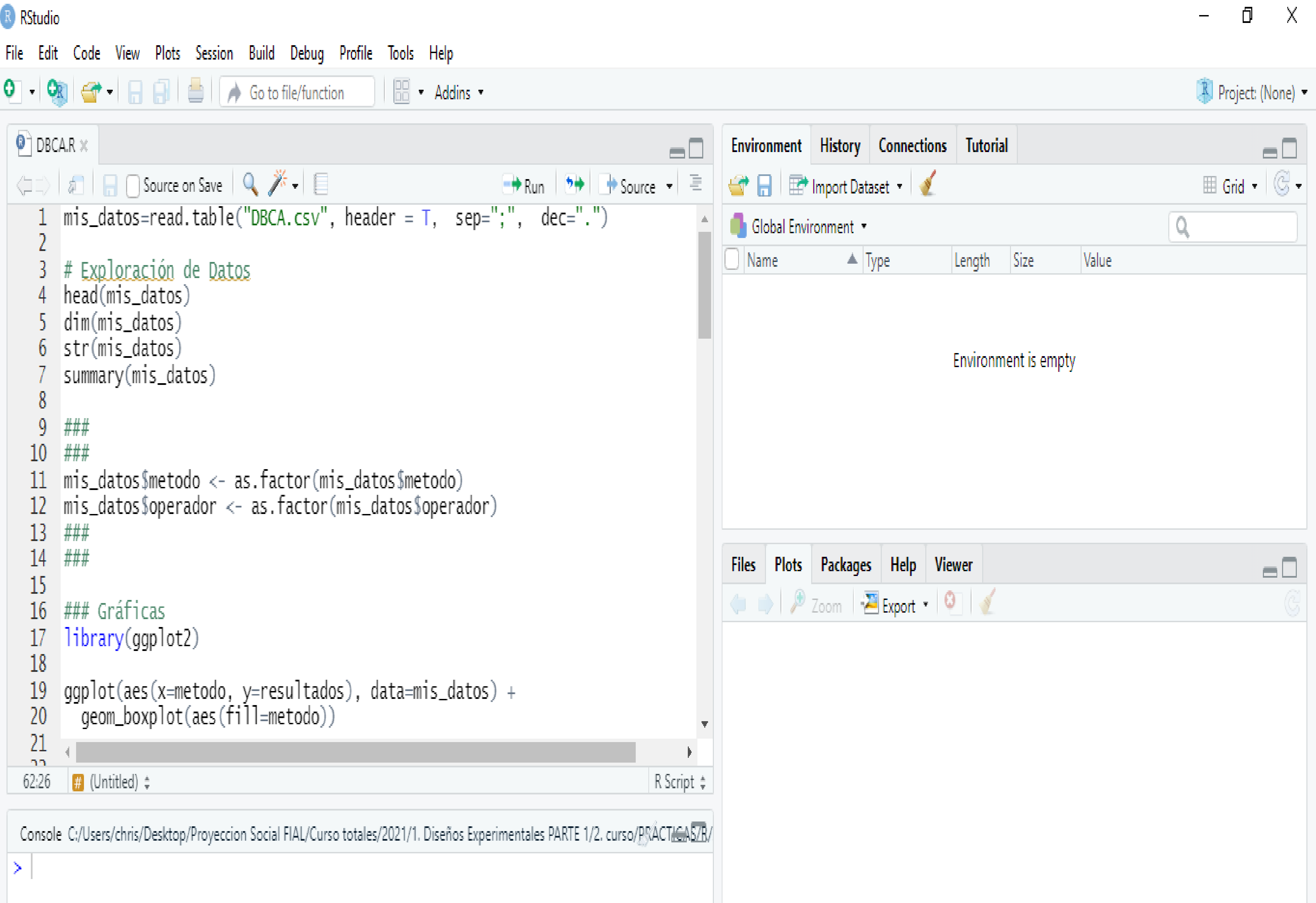


Respuesta
60
50
40
30
20
10
0
a
bc
ab
c
MetodoA
MetodoB
MetodoC
MetodoD


Resultados
60
50
40
30
20
10
0
a
b
ab
ab
1
2
3
4
OPERADORES

Se concluye que el tratamiento *A* es diferente de *C* y *D*, y que el tratamiento *B* es diferente de *C*. Las otras tres comparaciones (*A* con *B*, *B* con *D* y *C* con *D*) aceptan la hipótesis de igualdad. De acuerdo con esto, y dadas las respuestas medias muestrales $\bar{Y}_{A\bullet} = 7.5$, $\bar{Y}_{B\bullet} = 9.0$, $\bar{Y}_{C\bullet} = 10.75$, $\bar{Y}_{D\bullet} = 12.75$, se concluye que el método *A* es mejor (requiere menos tiempo para el ensamble) que los métodos *C* y *D*, pero el método *A* no es mejor que el *B*.

TRABAJO EN “R”


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
DBCA.R
Source on Save
Run
Source
mis_datos=read.table("DBCA.csv", header = T, sep=";", dec=".")

# Exploración de Datos
head(mis_datos)
dim(mis_datos)
str(mis_datos)
summary(mis_datos)

###
###
mis_datos$metodo <- as.factor(mis_datos$metodo)
mis_datos$operador <- as.factor(mis_datos$operador)
###
###

Gráficas
library(ggplot2)

ggplot(aes(x=metodo, y=resultados), data=mis_datos) +
  geom_boxplot(aes(fill=metodo))
62:26
(Untitled)
R Script
Console C:/Users/chris/Desktop/Proyeccion Social FIAL/Curso totales/2021/1. Diseños Experimentales PARTE 1/2. curso/PRACTICAS/R/
>
Environment History Connections Tutorial
Import Dataset
Grid
Global Environment
Name Type Length Size Value
Environment is empty
Files Plots Packages Help Viewer
Zoom
Export