Diseños Factoriales 2^2

En este tipo de diseños se estudia el efecto de dos factores considerando dos niveles en cada uno. Cada réplica de este diseño consiste de $\mathbf{2} \times \mathbf{2} = \mathbf{4}$ combinaciones o tratamientos denotados de diversas maneras, como se puede ver en la siguiente tabla:

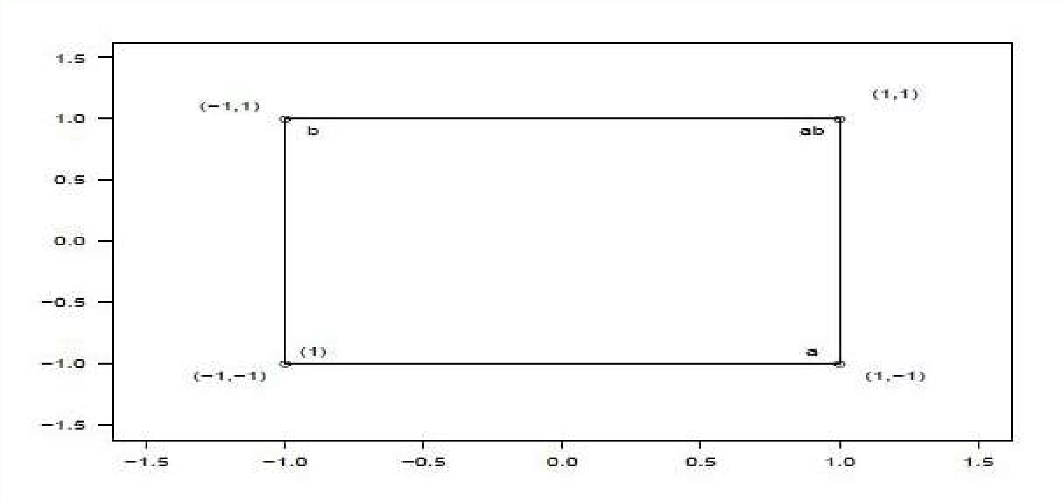
Tratamientos	A	B	A	B	A	B	A	B	A	B	A	B	Yates
Trat 1	bajo	bajo	A_1	B_1	A^-	B^-	-	-	0	0	-1	-1	(1)
Trat 2	alto	bajo	A_2	B_1	A^+	B^-	+	-	1	0	+1	-1	a
Trat 3	bajo	alto	A_1	B_2	A^-	B^+	-	+	0	1	-1	+1	b
Trat 4	alto	alto	A_2	B_2	A^+	B^+	+	+	1	+1	1	+1	ab

La notación de Yates $((1), a, b, ab)$ representa el total o suma de las observaciones en cada tratamiento:

- (1): Suma de los datos obtenidos en los niveles bajos de ambos factores $(-1, -1)$.
- a : Suma de los datos obtenidos en el nivel alto de A y bajo de B $(1, -1)$.
- b : Suma de los datos obtenidos en el nivel bajo de A y alto de B $(-1, 1)$.
- ab Suma de los datos obtenidos en los niveles altos de ambos factores $(+1, +1)$.

...

La representación geométrica de un Diseño $\mathbf{2^2}$ son los vértices de un cuadrado, como se ve a continuación:



El área limitada por el cuadrado se conoce como la región experimental y las conclusiones obtenidas del diseño sólo son válidas en esta región.

Ejemplo:

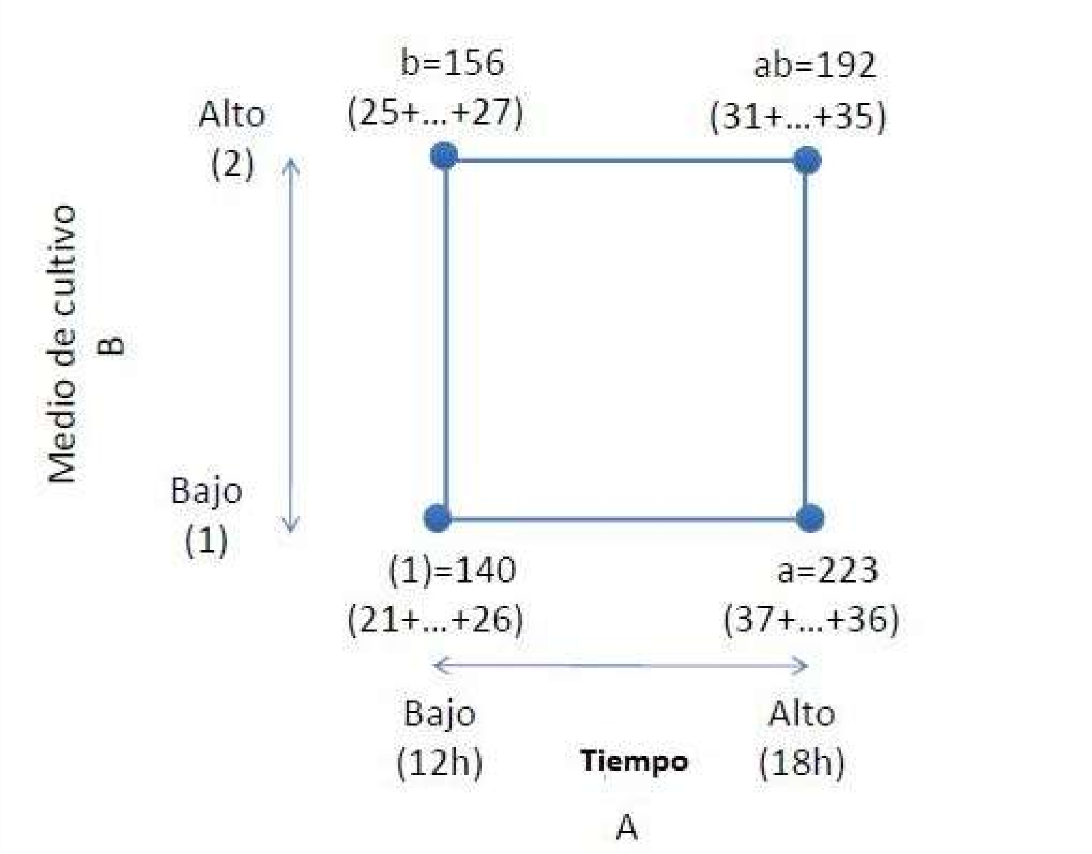
En un estudio se esta interesado en los efectos de dos medios de cultivo diferentes y dos tiempos diferentes sobre el crecimiento de un virus particular. Se realizan **seis** réplicas y se obtienen los siguientes resultados:

	Medio de cultivo			
Tiempo (h)	1		2	
12	21	22	25	26
	23	28	24	25
	20	26	29	27
18	37	39	31	34
	38	38	29	33
	35	36	30	35

Este es un ejemplo de un diseño factorial $\mathbf{2^2}$, con **seis** réplicas por tratamiento.
Sea el tiempo el factor A con niveles de interés de $12\ h$ y $18\ h$. Sea el medio del cultivo el factor B, con niveles de interés denotados por 1 y 2.

Representación gráfica del diseño

Este diseño se representa gráficamente como sigue:



Descripción de la Figura anterior

- En la figura se observa que el nivel alto de cualquiera de los factores en una combinación de tratamientos se denota por la letra minúscula correspondiente y que el nivel bajo de un factor en una combinación de tratamientos se denota por la ausencia de la letra respectiva.
- La letra $\mathbf{a}$ representa la combinación de tratamientos con $\mathbf{A}$ en el nivel alto y $\mathbf{B}$ en el nivel bajo.
- La letra $\mathbf{b}$ representa la combinación de tratamientos con $\mathbf{A}$ en el nivel bajo y $\mathbf{B}$ en el nivel alto.
- Las letras $\mathbf{ab}$ representa la combinación de tratamientos con ambos factores en el nivel alto.
- Por convención, se usa $(\mathbf{1})$ para denotar la combinación de tratamientos con ambos factores en el nivel bajo.
- Así mismo, los símbolos, $(\mathbf{1})$, $\mathbf{ab}$, $\mathbf{a}$ y $\mathbf{b}$, representan el total de las $\mathbf{n}$ réplicas hechas con la combinación de los tratamientos.

Cálculo de los Efectos en un Diseño 2^2:

Los efectos principales de los factores considerados se definen como la diferencia de la respuesta media observada, cuando el factor estuvo en su nivel **alto** y la respuesta media observada cuando el factor estuvo en su nivel **bajo**:

El estimador del efecto principal del factor $\mathbf{A}$ es:

$$\left(\frac{\mathbf{ab}+\mathbf{a}}{\mathbf{2n}}\right)-\left(\frac{(\mathbf{1})+\mathbf{b}}{\mathbf{2n}}\right)=\frac{\mathbf{1}}{\mathbf{2n}}[\mathbf{ab}+\mathbf{a}-\mathbf{b}-(\mathbf{1})]$$

El estimador dle efecto principal del factor $\mathbf{B}$ es:

$$\left(\frac{\mathbf{ab}+\mathbf{b}}{\mathbf{2n}}\right)-\left(\frac{(\mathbf{1})+\mathbf{a}}{\mathbf{2n}}\right)=\frac{\mathbf{1}}{\mathbf{2n}}[\mathbf{ab}+\mathbf{b}-\mathbf{a}-(\mathbf{1})]$$

El estimador del efecto de Interacción de $\mathbf{A}$ con $\mathbf{B}$ es:

$$\left(\frac{\mathbf{ab}+(\mathbf{1})}{\mathbf{2n}}\right)-\left(\frac{\mathbf{a}+\mathbf{b}}{\mathbf{2n}}\right)=\frac{\mathbf{1}}{\mathbf{2n}}[\mathbf{ab}+(\mathbf{1})-\mathbf{a}-\mathbf{b}].$$

Este último es: Diferencia entre las medias de las diagonales del cuadrado.

Retomando el Ejemplo

El estimador del efecto principal del factor $\mathbf{A}$ es:

$$\text{Efecto de } \mathbf{A} \text{ estimado} = \frac{\mathbf{1}}{\mathbf{2n}}[\mathbf{ab}+\mathbf{a}-\mathbf{b}-(\mathbf{1})] = \frac{\mathbf{1}}{\mathbf{2}\times\mathbf{6}}[\mathbf{192}+\mathbf{223}-\mathbf{156}-\mathbf{140}] = \mathbf{9.92}$$

El estimador del efecto principal del factor $\mathbf{B}$ es:

$$\text{Efecto de } \mathbf{B} \text{ estimado} = \frac{\mathbf{1}}{\mathbf{2n}}[\mathbf{ab}+\mathbf{b}-\mathbf{a}-(\mathbf{1})] = \frac{\mathbf{1}}{\mathbf{2}\times\mathbf{6}}[\mathbf{192}+\mathbf{156}-\mathbf{223}-\mathbf{140}] = -\mathbf{1.25}$$

El estimador del efecto de Interacción de $\mathbf{A}$ con $\mathbf{B}$ es:

$$\text{Efecto de } \mathbf{AB} \text{ estimado} = \frac{\mathbf{1}}{\mathbf{2n}}[\mathbf{ab}+(\mathbf{1})-\mathbf{a}-\mathbf{b}] = \frac{\mathbf{1}}{\mathbf{2}\times\mathbf{6}}[\mathbf{192}+\mathbf{140}-\mathbf{223}-\mathbf{156}] = -\mathbf{3.92}$$

Interpretación de los Efectos:

1. El efecto de A, ie. del tiempo, es positivo, esto sugiere que al incrementar A del nivel bajo $(12h)$ al nivel alto $(18h)$, el crecimiento del virus aumenta.

2. El efecto de B, ie. del medio de cultivo, es negativo, esto sugiere que al pasar del medio de cultivo tipo 1 al medio de cultivo tipo 2, el crecimiento del virus disminuye. La magnitud del efecto sugiere además que este efecto es pequeño.

3. En general, en este tipo de diseños, se analiza la magnitud y la dirección de los efectos de los factores a fin de determinar las variables que son de posible importancia.

Evaluación de la significancia de los efectos

Una vez hallados los efectos se necesita determinar si estos son o no significativos desde el punto de vista estadístico, a partir de la tabla ANOVA, para lo cual se procede como sigue:

Se definen los siguientes contrastes:

1. Contraste para $\mathbf{A}$:

$$\text{Contraste para } \mathbf{A} = \mathbf{ab} + \mathbf{a} - \mathbf{b} - (\mathbf{1})$$

2. Contraste para $\mathbf{B}$:

$$\text{Contraste para } \mathbf{B} = \mathbf{ab} + \mathbf{b} - \mathbf{a} - (\mathbf{1})$$

3. Contraste para $\mathbf{AB}$:

$$\text{Contraste para } \mathbf{AB} = \mathbf{ab} + (\mathbf{1}) - \mathbf{a} - \mathbf{b}$$

(El valor efecto se obtiene dividiendo el respectivo contraste por $\mathbf{2n}$).

...

Para calcular los contrastes se puede hacer uso de la siguiente tabla de signos:

A	B	AB	Yates
-	-	+	(1)
+	-	-	a
-	+	-	b
+	+	+	ab

Contraste B: $b + ab - (a + (1))$ Contraste AB: $(1) + ab - (a + b)$.

Las respectivas hipótesis a probar son:

$$\begin{cases} H_0 : \text{Efecto de } A = 0 \\ H_a : \text{Efecto de } A \neq 0 \end{cases}, \quad \begin{cases} H_0 : \text{Efecto de } B = 0 \\ H_a : \text{Efecto de } B \neq 0 \end{cases}, \quad \begin{cases} H_0 : \text{Efecto de } AB = 0 \\ H_a : \text{Efecto de } AB \neq 0 \end{cases}$$

La SS-asociada a cada contraste es:

$$\mathbf{SS_A} = \frac{(\sum \mathbf{c_i Y_i})^2}{\sum \mathbf{n_i c_i^2}} = \frac{(\sum \mathbf{c_i Y_i})^2}{\sum \mathbf{n_i}} = \frac{\mathbf{C^2}}{\mathbf{n2^2}}$$

y cada $\mathbf{SS}$ de los contrastes tiene asociado un grado de libertad.

...

Luego, se tiene que:

$$SSA = \frac{1}{2^2 n}[a + ab - b - (1)]^2 \ (\text{gl } 1) \qquad SSB = \frac{1}{2^2 n}[b + ab - a - (1)]^2 \ (\text{gl } 1)$$

$$SSAB = \frac{1}{2^2 n}[ab + (1) - a - b]^2 \ (\text{gl } 1)$$

$$SST = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{l=1}^{n} Y_{ijl}^2 - \frac{1}{2^2 n}Y_{...}^2 \ (gl \ \ n2^2 - 1).$$

La estadística de prueba es:

$$\mathbf{F} = \frac{\mathbf{SS_A/gl}}{\mathbf{MSE}} = \frac{\mathbf{MS_A}}{\mathbf{MSE}} \sim \mathbf{F_{1,glE}} = \mathbf{F_{1,4(n-1)}}.$$

La tabla ANOVA es:

Fuente Variación	SS	gl	MS	F_0	VP
A	SSA	1	MSA	$\frac{MSA}{MSE}$	$P[F_{1,4(n-1)} > F_0^A]$
B	SSB	1	MSB	$\frac{MSB}{MSE}$	$P[F_{1,4(n-1)} > F_0^B]$
AB	SSAB	1	MSAB	$\frac{MSAB}{MSE}$	$P[F_{1,4(n-1)} > F_0^{AB}]$
Error	SSE	4(n-1)	MSE		
Total	SST	$n2^2 - 1$			

Algunas anotaciones

NOTA: Si el $\mathbf{VP} < \alpha$ entonces se concluye que el efecto correspondiente es estadísticamente diferente de cero, es decir, tal efecto está **ACTIVO** o influye de manera significativa en la respuesta y mientras más pequeño sea el $\mathbf{VP}$, más importante es tal EFECTO.

También se puede concluir comparando el valor del $\mathbf{F}$-calculado con el respectivo valor crítico de la tabla. Se rechaza $\mathbf{H_0}$-si el $\mathbf{F}$-calculado es mayor que el $\mathbf{F}$-de la tabla.

Ejemplo

Se desea estudiar el efecto del tamaño de la broca (Factor A) y la velocidad de la broca (factor B) sobre la vibración de la ranura (respuesta). Para ello se decide utilizar un diseño factorial $\mathbf{2^2}$ con $\mathbf{4}$ réplicas, ie. $\mathbf{4}$-repeticiones por tratamiento. El total de corridas $\mathbf{n} \times \mathbf{2^2} = \mathbf{4} \times \mathbf{2^2} = \mathbf{16}$, las cuales se realizan en orden aleatorio. El tamaño de la broca utilizada es: $\mathbf{1/16}$ y $\mathbf{1/8}$ plg y la velocidad considerada es $\mathbf{40}$ y $\mathbf{90}$ rev/seg. Los datos obtenidos son:

A:Broca	B:Vel.	Orden				A	B	Vibr.				Total
1/16	40	4	8	12	14	-	-	18.2	18.9	12.9	14.4	(1): 64.4
1/8	40	1	6	10	13	+	-	27.2	24.0	22.4	22.5	a: 96.1
1/16	90	3	7	11	15	-	+	15.9	14.5	15.1	14.2	(b): 59.7
1/8	90	2	5	9	16	+	+	41.0	43.9	36.3	39.9	ab: 161.1

Preguntas a resolver

1. ¿La velocidad y el tamaño de la broca afectan la vibración de la ranura?
2. Si la afectan, ¿cómo es tal efecto y cuál es la combinación de velocidad y tamaño de broca que minimizan la vibración?
3. ¿Cuál es la vibración esperada en las condiciones ”óptimas”?
4. ¿Se cumplen los supuestos del modelo?

...

Efecto de A: $\frac{1}{2(4)}[96{,}1 + 161{,}1 - 59{,}7 - 64{,}4] = 16{,}64$

Efecto de B: $\frac{1}{2(4)}[59{,}7 + 161{,}1 - 96{,}1 - 64{,}4] = 7{,}54$

Efecto de AB: $\frac{1}{2(4)}[161{,}1 + 64{,}4 - 59{,}7 - 96{,}1] = 8{,}71$

De los efectos el más alto es el asociado al TAMAÑO DE BROCA.

Luego se tienen las sumas de cuadrados asociadas a cada contraste:

$$
\begin{aligned}
SSA &= \frac{1}{2^2 4}[16{,}64 * 2 * 4]^2 = 2^2 * \text{Efecto de A}^2 = 1107{,}22 \\
SSB &= \frac{1}{2^2 4}[7{,}54 * 2 * 4]^2 = 2^2 * \text{Efecto de B}^2 = 227{,}25 \\
SSAB &= \frac{1}{2^2 4}[8{,}71 * 2 * 4]^2 = 2^2 * \text{Efecto de AB}^2 = 303{,}63 \\
SST &= \sum\sum\sum Y_{ijl}^2 - \frac{1}{2^2 n} Y_{\cdots}^2 = 1709{,}83 \\
SSE &= SST - SSA - SSB - SSAB = 71{,}73(gl15 - 3 = 12)
\end{aligned}
$$

...

TABLA ANOVA

Fuente Variación	SS	gl	MS	F_0	VP
A: Broca	1107,22	1	1107,22	185,25	$< 0{,}0001$
B: Veloc.	227,25	1	227,25	38.02	$< 0{,}0001$
AB	303.63	1	303,63	50.80	$< 0{,}0001$
Error	71,71	12	5.98		
Total	1709,83	15			

En conclusión: Los tres efectos son significativos.

Interpretación y Conclusiones:

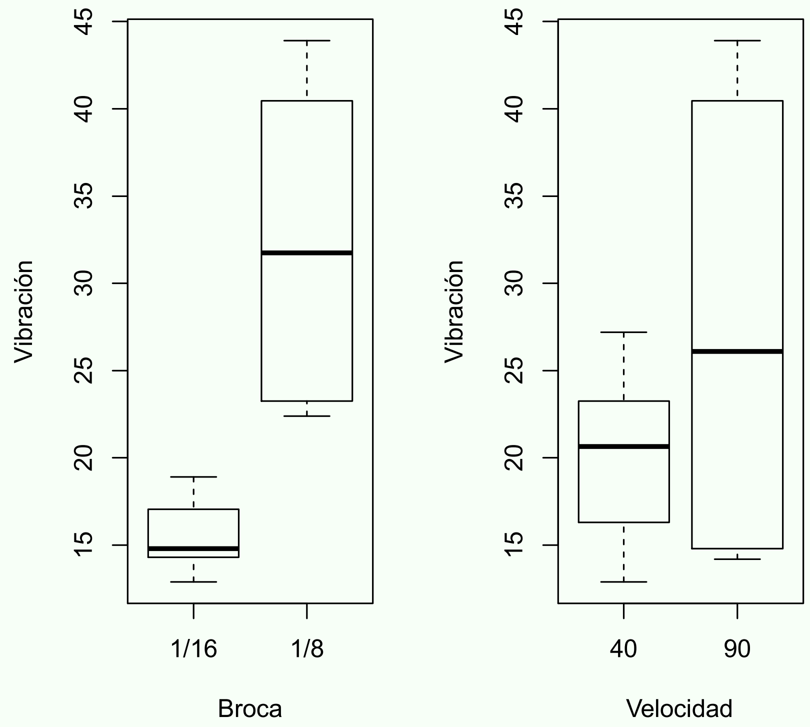
El objetivo era minimizar la vibración de la ranuradora y las interacciones en este caso tiene prioridad respecto a los efectos principales.

Cuando la broca se encuentra en su nivel más bajo, la velocidad no afecta de manera significativa la vibración.

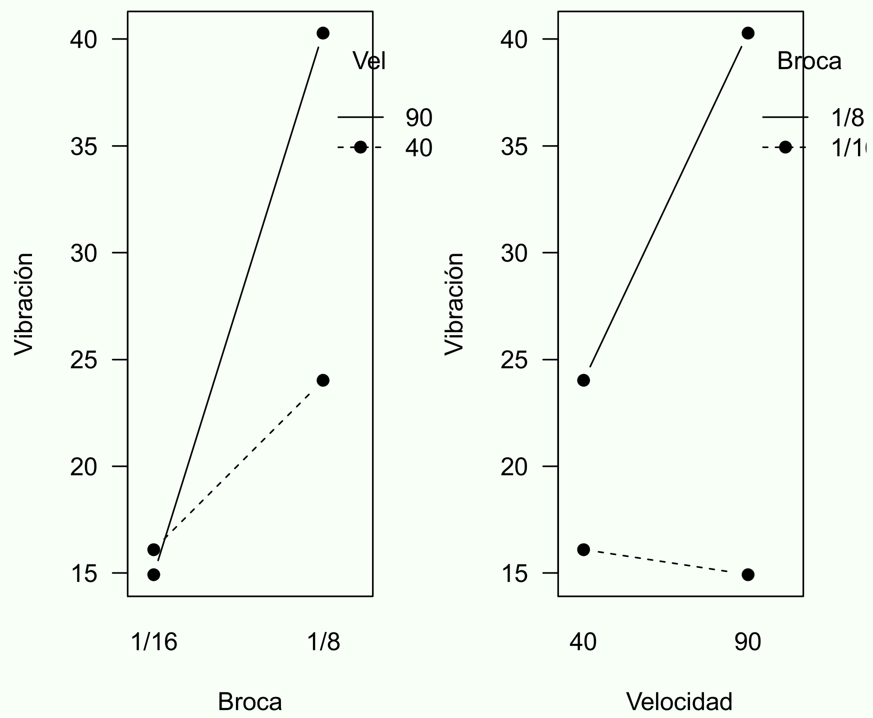
Cuando la broca se encuentra en su nivel alto, la velocidad tiene un efecto considerable sobre la vibración.

Ejemplo - Continuación

Lo anterior se puede ver en las siguientes gráficas.

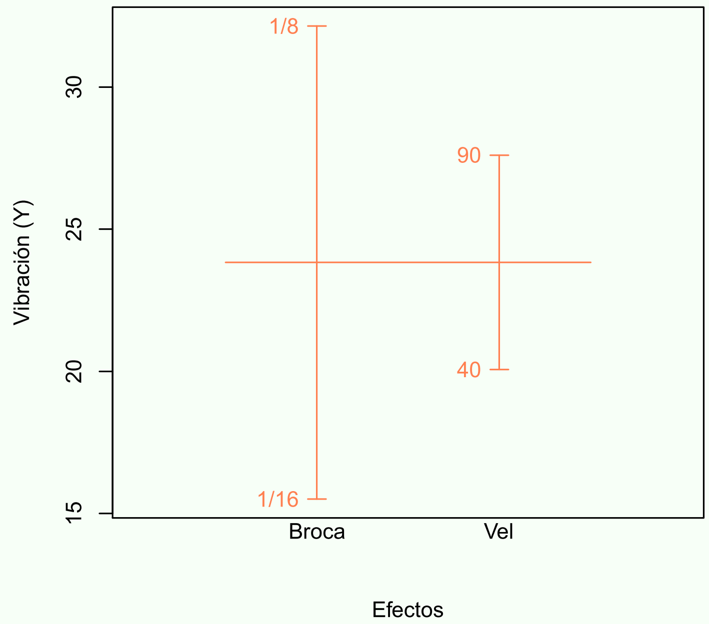


...



La mínima vibración se alcanza usando los tratamientos: (A^-, B^+) o (A^-, B^-).

...



A partir del gráfico de los efectos principales se observa que no se debe aumentar la velocidad ni el tamaño de la broca si se quiere minimizar la vibración, de donde la combinación a usar sería: (A^-, B^-).

Predicción de la respuesta media

Para el mejor tratamiento (ie. Respuesta media en: $(\mathbf{A}^-, \mathbf{B}^-)$ y $\mathbf{R}^2$:
Con los datos codificados de la forma: $\mathbf{x_1} = \pm\mathbf{1}(\mathbf{A})$, $\mathbf{x_2} = \pm\mathbf{1}(\mathbf{B})$, se ajusta el modelo de regresión lineal para describir el comportamiento de la vibración en cualquier punto del diseño, dicho modelo ajustado es:

$$\hat{\mathbf{Y}} = \mathbf{23.83} + \mathbf{8.32x_1} + \mathbf{3.77x_2} + \mathbf{4.36x_1x_2}$$

y se tiene un $\mathbf{R^2} = \mathbf{0.9476}$.
El valor de la respuesta media óptima ie. la respuesta media en: $(\mathbf{A}^-, \mathbf{B}^-)$ es:

$$\hat{\mathbf{Y}}_{\text{óptimo}} = \mathbf{23.83} + \mathbf{8.32}(-\mathbf{1}) + \mathbf{3.77}(-\mathbf{1}) + \mathbf{4.36}(-\mathbf{1})(-\mathbf{1}) = \mathbf{16.1}$$

NOTA:
En diseños $\mathbf{2^k}$, los coeficientes estimados se hallan dividiendo el respectivo efecto por 2, el intercepto $\hat{\beta}_0 = \mathbf{23.83125} = \overline{\mathbf{Y}}_{...}$- Vibración predicha en el centro de la región experimental.

Resultados obtenidos con R

```
lm(formula = vibracion ~ Broca * Vel)
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  23.8312     0.6112  38.991 5.22e-14 ***
Broca         8.3187     0.6112  13.611 1.17e-08 ***
Vel           3.7687     0.6112   6.166 4.83e-05 ***
Broca:Vel     4.3562     0.6112   7.127 1.20e-05 ***

Residual standard error: 2.445 on 12 degrees of freedom
Multiple R-squared:  0.9581,    Adjusted R-squared:  0.9476
F-statistic: 91.36 on 3 and 12 DF,  p-value: 1.569e-08
```

DISEÑOS 2^3

Diseños 2^3:
Diseños de tres factores, donde cada uno tiene **dos** niveles. Se obtiene un total de $\mathbf{2^3}$-Tratamientos y $\mathbf{n}$-réplicas en cada tratamiento.
Número total de observaciones: $\mathbf{n} \times \mathbf{2^3}$.
Región Experimental: Un cubo centrado en $(\mathbf{0}, \mathbf{0}, \mathbf{0})$ donde los vértices identifican los **ocho** tratamientos.

A	B	C	AB	AC	BC	ABC	Yates
-1	-1	-1	+1	+1	+1	-1	(1)
+1	-1	-1	-1	-1	+1	+1	a
-1	+1	-1	-1	+1	-1	+1	b
+1	+1	-1	+1	-1	-1	-1	ab
-1	-1	+1	+1	-1	-1	+1	c
+1	-1	+1	-1	+1	-1	-1	ac
-1	+1	+1	-1	-1	+1	-1	bc
+1	+1	+1	+1	+1	+1	+1	abc

OBSERVACIÓN:
Se estudian factores principales, efectos de interacciones dobles.
Interacciones de mayor orden (tres o más), generalmente no influyen de manera significativa, motivo por el cual no se recomienda estudiarlas.

Contrastes asociados al Diseño 2^3

$$\begin{aligned}
\text{contraste para } \mathbf{A} &= a+ab+ac+abc-(1)-b-c-bc\\
\text{contraste para } \mathbf{B} &= b+ab+bc+abc-(1)-a-c-ac\\
\text{contraste para } \mathbf{C} &= c+ac+bc+abc-(1)-b-a-ab\\
\text{contraste para } \mathbf{AB} &= (1)+ab+abc+c-a-b-ac-bc\\
\text{contraste para } \mathbf{BC} &= (1)+bc+abc+a-b-c-ab-ac\\
\text{contraste para } \mathbf{AC} &= (1)+ac+abc+b-a-c-ab-bc\\
\text{contraste para } \mathbf{ABC} &= abc+a+b+c-bc-ac-ab-(1)
\end{aligned}$$

Los efectos de cada **factor** o de **interacción** se calculan dividiendo el contraste respectivo por:

$$n2^{k-1} = n2^{3-1} = n2^2 = 4n.$$

Las sumas de cuadrados asociadas a los contrastes son:

$$SS_A = \frac{\text{Contraste de A}^2}{n \times 2^k} = \frac{\text{Contraste de A}^2}{n \times 2^3} = \frac{\text{Contraste de A}^2}{8n}.$$

Cada suma de cuadrados tiene asociado 1-gl.

...

Para definir los anteriores contrastes, se puede tener en cuenta la siguiente tabla de signos:

Treatment Combination	Factorial Effect							
	I	A	B	AB	C	AC	BC	ABC
(1)	+	−	−	+	−	+	+	−
a	+	+	−	−	−	−	+	+
b	+	−	+	−	−	+	−	+
ab	+	+	+	+	−	−	−	−
c	+	−	−	+	+	−	−	+
ac	+	+	−	−	+	+	−	−
bc	+	−	+	−	+	−	+	−
abc	+	+	+	+	+	+	+	+