Diseño de Experimentos con varios factores

Jhon Jairo Padilla Aguilar, PhD.

Experimento Factorial

- En cada ensayo completo o réplica del experimento se investigan todas las combinaciones posibles de los niveles de los factores.

Experimento factorial

- Si hay dos factores A y B con *a* niveles del factor A y *b* niveles del factor B, entonces cada réplica incluye todas las combinaciones de tratamientos *ab*.
- Efecto de un factor:
  - Es el cambio en la respuesta que produce un cambio en el nivel del factor
  - Se le llama efecto principal (de un factor primario del estudio)

Ejemplo

- Experimento factorial con 2 factores (A, B), cada uno con dos niveles (low, high)

Table 14-1 A Factorial Experiment with Two Factors

	Factor B	
Factor A	B_{low}	B_{high}
A_{low}	10	20
A_{high}	30	40

Ejemplo

- Efecto principal de A es la diferencia de los promedios en los dos niveles (high y low) así:

$$A = \frac{30+40}{2} - \frac{10+20}{2} = 20$$

- Es decir, al cambiar el factor A del nivel bajo al nivel alto se produce un incremento de la respuesta media de 20 unidades
- El efecto principal de B será:

$$B = \frac{20+40}{2} - \frac{10+30}{2} = 20$$

Interacción entre factores

- La diferencia de la respuesta entre los niveles de un factor no es la misma en todos los niveles de los otros factores

Table 14-2 A Factorial Experiment with Interaction

Factor A	Factor B: B_{low}	Factor B: B_{high}
A_{low}	10	20
A_{high}	30	0

- En el nivel bajo del factor B, el efecto de A es A=30-10=20
- En el nivel alto del factor B, el efecto de A es A=0-20=-20
- El nivel de A depende del nivel elegido del factor B: Hay interacción entre A y B

Interacción entre factores

- Cuando la interacción es grande, los efectos principales correspondientes tienen muy poco significado práctico
- En el ejemplo de la tabla 2, se encuentra que el efecto principal de A es:

$$A = \frac{30+0}{2} - \frac{10+20}{2} = 0$$

- Y se podría estar tentado a concluir que el factor A no tiene efecto alguno
- Pero al examinar los efectos de A en los diferentes niveles de B, se encontró que no era cierto.
- Es muy útil el conocimiento de la interacción AB

Interacción entre factores

- Cuando está presente una interacción, los efectos principales de los factores que intervienen en la interacción quizá no tengan mucho significado.
- El efecto de la interacción AB es la diferencia de los promedios de las diagonales

Table 14-1 A Factorial Experiment with Two Factors

Factor A	Factor B: B_{low}	B_{high}
A_{low}	10	20
A_{high}	30	40

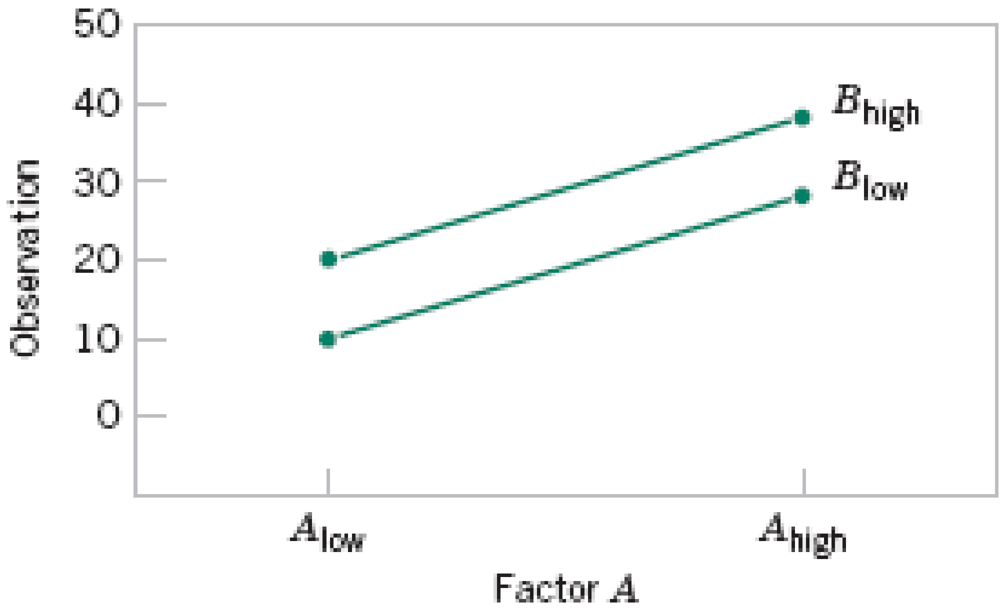
$$AB = \frac{20+30}{2} - \frac{10+40}{2} = 0$$

Table 14-2 A Factorial Experiment with Interaction

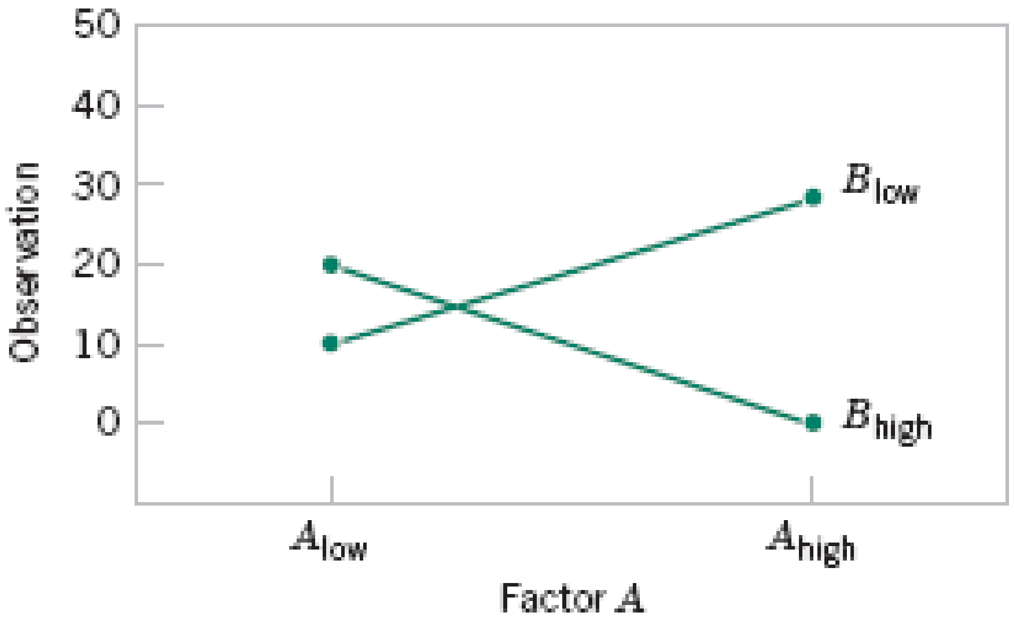
Factor A	Factor B: B_{low}	B_{high}
A_{low}	10	20
A_{high}	30	0

$$AB = \frac{20+30}{2} - \frac{10+0}{2} = 20$$

Gráficas de Interacción de dos factores



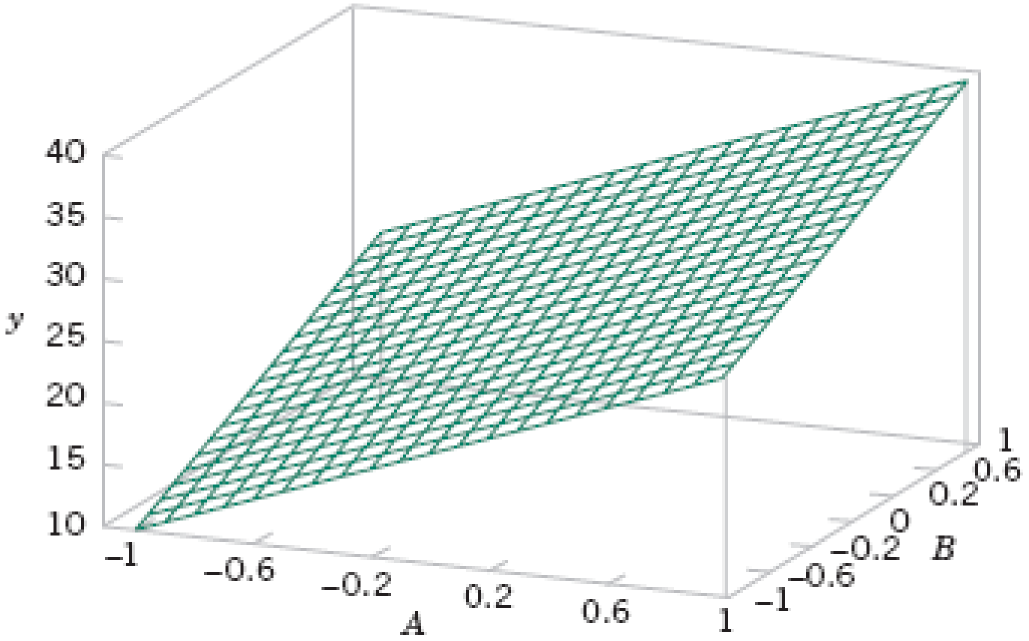
Experimento Factorial sin interacción

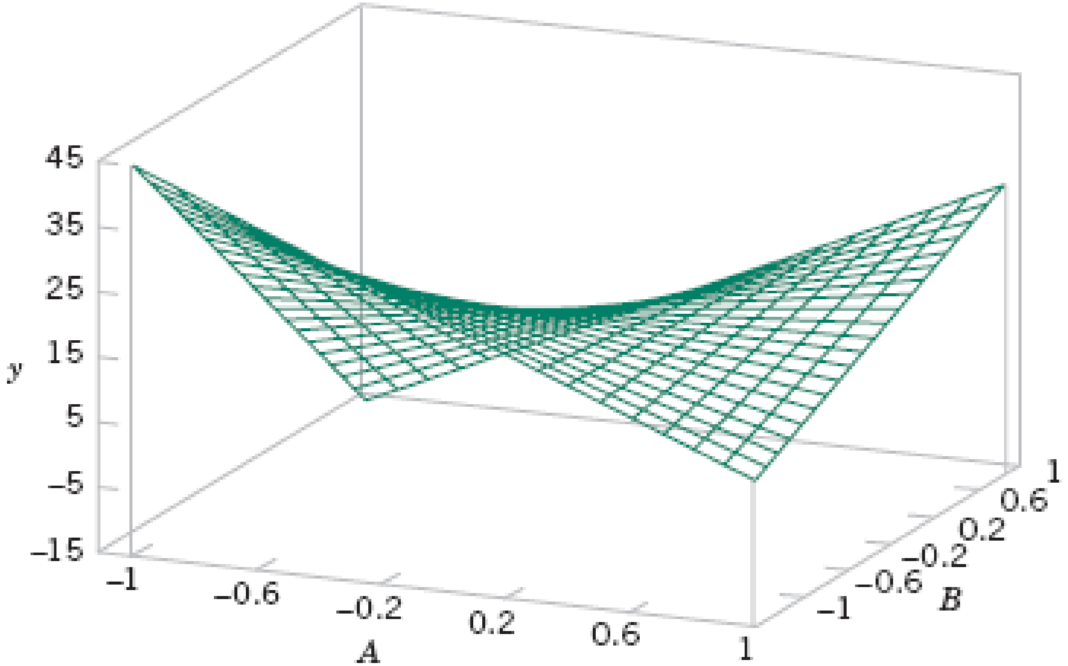


Experimento Factorial con interacción

Estas gráficas son útiles para presentar los resultados de experimentos. Muchos programas de computadora que se usan para analizar datos de experimentos diseñados construirán estas gráficas automáticamente.

Gráfica de superficie tridimensional





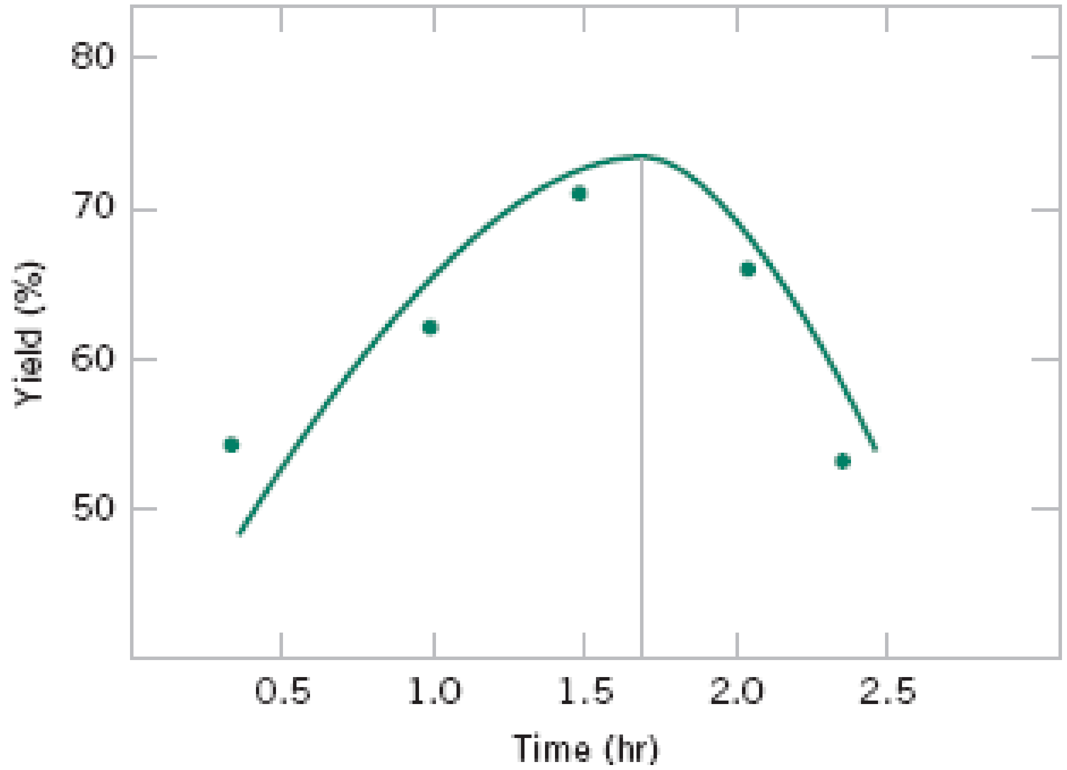
- Caso en que no hay interacción entre A y B.
- La superficie es un plano que se localiza arriba del espacio AB
- La pendiente del plano depende de A y B por aparte

- Caso en que hay interacción entre A y B
- El efecto es torcer el plano, generando una curvatura en la función de respuesta

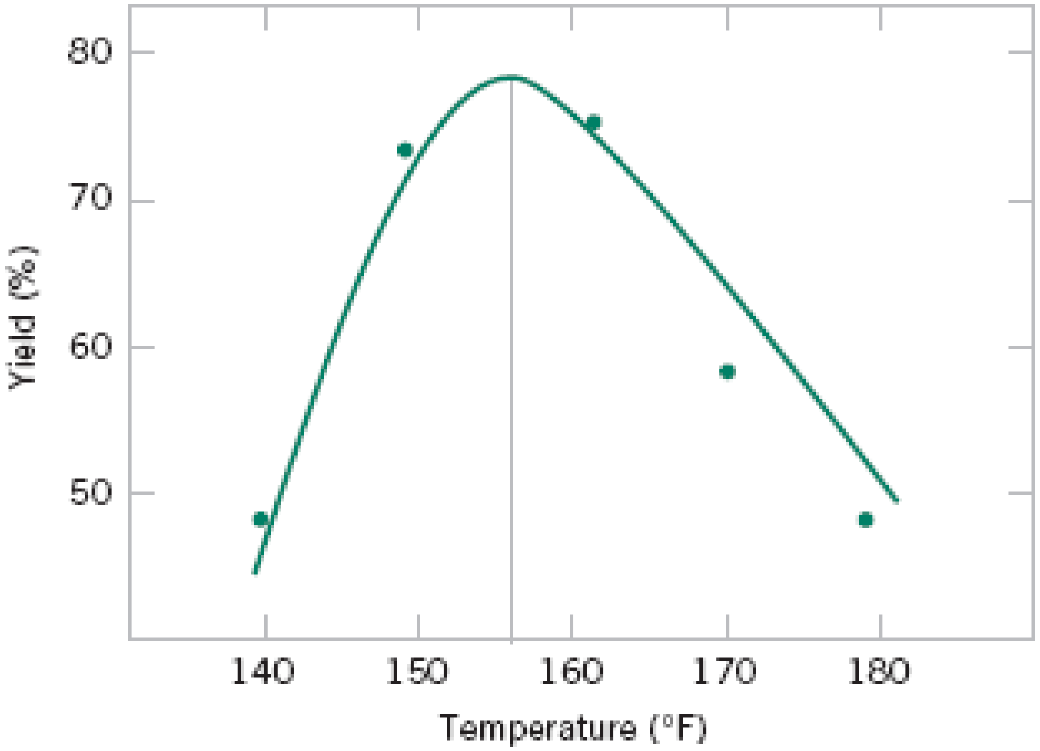
El problema de cambiar los factores uno a la vez

- Suponga que en un proceso químico, un ingeniero tiene interés en encontrar los valores de temperatura y tiempo que maximizan el rendimiento.

Procedimiento con un factor a la vez

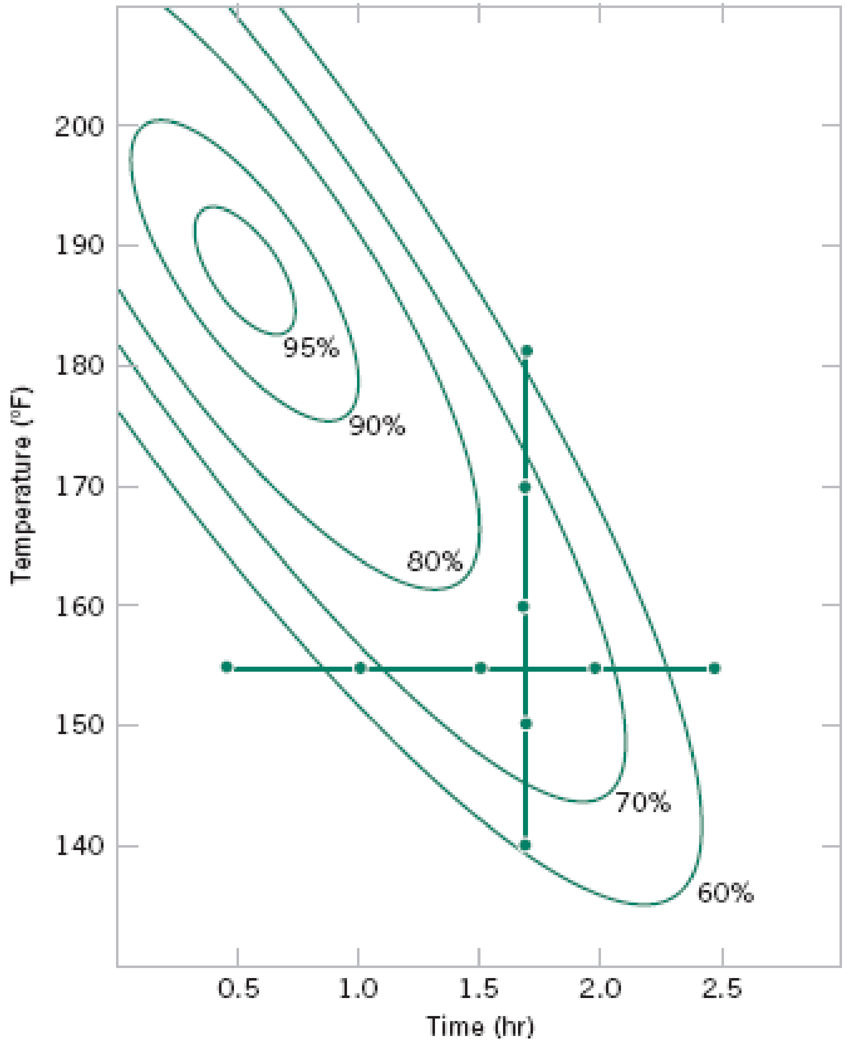


Primero se varía el tiempo con una temperatura constante de 155°F. Se determina que el tiempo que da mejor rendimiento es 1.7 horas.



Se fija el tiempo de reacción en 1.7 horas y se varía la temperatura.
El mejor rendimiento se obtiene en 155°F aproximadamente

Resultados con un Factor a la vez



- En el gráfico con niveles de rendimiento se observa que los cambios que se hicieron están en una zona lejana del máximo posible del rendimiento.
- El experimento se hizo en una zona donde el máximo posible era del 75%

Concluyendo

- El enfoque de un factor a la vez falló porque no puede detectar la interacción entre la temperatura y el tiempo
- Los experimentos factoriales son la única manera de detectar interacciones
- El método de un factor a la vez es ineficiente porque:
  - Requerirá un número mayor de experimentos que el método factorial
  - NO existe la seguridad de que producirá los resultados correctos!

Experimentos Factoriales con dos factores

- El experimento factorial más sencillo es aquel en que sólo intervienen dos factores (A y B)
- Hay *a* niveles del factor A y b niveles del factor B
- El experimento contiene n réplicas
- Cada réplica contiene todas las combinaciones de tratamientos *ab*
- La observación de la celda ij-ésima para la réplica k-ésima se denota por y_{ijk}
- Cuando se realiza el experimento, las *abn* observaciones se corren en orden aleatorio (Diseño completamente aleatorizado)

Arreglo de datos para un experimento de dos factores

Table 14-3 Data Arrangement for a Two-Factor Factorial Design

		Factor B					
		1	2	$\cdots$	b	Totals	Averages
Factor A	1	$y_{111}, y_{112}, \ldots, y_{11n}$	$y_{121}, y_{122}, \ldots, y_{12n}$		$y_{1b1}, y_{1b2}, \ldots, y_{1bn}$	$y_{1\cdot\cdot}$	$\bar{y}_{1\cdot\cdot}$
	2	$y_{211}, y_{212}, \ldots, y_{21n}$	$y_{221}, y_{222}, \ldots, y_{22n}$		$y_{2b1}, y_{2b2}, \ldots, y_{2bn}$	$y_{2\cdot\cdot}$	$\bar{y}_{2\cdot\cdot}$
	$\vdots$						
	a	$y_{a11}, y_{a12}, \ldots, y_{a1n}$	$y_{a21}, y_{a22}, \ldots, y_{a2n}$		$y_{ab1}, y_{ab2}, \ldots, y_{abn}$	$y_{a\cdot\cdot}$	$\bar{y}_{a\cdot\cdot}$
Totals		$y_{\cdot 1\cdot}$	$y_{\cdot 2\cdot}$		$y_{\cdot b\cdot}$	$y_{\cdot\cdot\cdot}$	
Averages		$\bar{y}_{\cdot 1\cdot}$	$\bar{y}_{\cdot 2\cdot}$		$\bar{y}_{\cdot b\cdot}$		$\bar{y}_{\cdot\cdot\cdot}$

Modelo estadístico lineal

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases}$$

- μ es el efecto promedio global
- τ_i es el efecto del nivel i-ésimo del factor A
- β_j es el efecto del nivel j-ésimo del factor B
- $(\tau\beta)_{ij}$ es el efecto de la interacción entre A y B
- ε_{ijk} es un componente de error aleatorio que tiene una distribución normal con media cero y varianza σ^2.

Hipótesis a probar

- El interés se centra en probar las hipótesis de que no hay efecto principal para el factor A, que no hay efecto principal para el factor B, y que no hay efecto de interacción AB.
- Se usa el Análisis de varianza para probar la hipótesis (Análisis de varianza de dos direcciones)

Análisis estadístico con el modelo de efectos fijos

- Suponga que el experimentador elije específicamente los *a* niveles del factor A y los b niveles del factor B.
- Las inferencias se restringen entonces a estos niveles.
- Los efectos del factor A, B y la interacción AB se consideran desviaciones de la media, por lo que

$$\sum_{i=1}^{a}\tau_i = 0 \text{ y } \sum_{j=1}^{b}\beta_j = 0 \qquad \sum_{i=1}^{a}(\tau\beta)_{ij} = \sum_{j=1}^{b}(\tau\beta)_{ij} = 0$$

Definiciones necesarias

$$y_{i..} = \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk} \qquad \bar{y}_{i..} = \frac{y_{i..}}{bn} \qquad i = 1, 2, \ldots, a$$

$$y_{.j.} = \sum_{i=1}^{a} \sum_{k=1}^{n} y_{ijk} \qquad \bar{y}_{.j.} = \frac{y_{.j.}}{an} \qquad j = 1, 2, \ldots, b$$

$$y_{ij.} = \sum_{k=1}^{n} y_{ijk} \qquad \bar{y}_{ij.} = \frac{y_{ij.}}{n} \qquad \begin{array}{l} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{array}$$

$$y_{...} = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk} \qquad \bar{y}_{...} = \frac{y_{...}}{abn}$$

Hipótesis a probar

Las hipótesis que se probarán son las guientes:

1. H_0: $\tau_1 = \tau_2 = \cdots = \tau_a = 0$ Ningún efecto principal del factor A
   H_1: Al menos una $\tau_i \neq 0$
2. H_0: $\beta_1 = \beta_2 = \cdots = \beta_b = 0$ Ningún efecto principal del factor B
   H_1: Al menos una $\beta_j \neq 0$
3. H_0: $(\tau\beta)_{11} = (\tau\beta)_{12} = \cdots = (\tau\beta)_{ab} = 0$ Ninguna interacción
   H_1: Al menos una $(\tau\beta)_{ij} \neq 0$

Descomposición de la suma de cuadrados

La identidad de suma de cuadrados para ANOVA de dos factores es:

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{...})^2 = bn\sum_{i=1}^{a}(\bar{y}_{i..}-\bar{y}_{...})^2$$

$$+\ an\sum_{j=1}^{b}(\bar{y}_{.j.}-\bar{y}_{...})^2$$

$$+\ n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y}_{...})^2$$

$$+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij.})^2 \qquad (14\text{-}3)$$

O simbólicamente,

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E \qquad (14\text{-}4)$$

La prueba de hipótesis

Para probar H_0: $\tau_i = 0$, se usa la razón:

$$F_0 = \frac{MS_A}{MS_E}$$

Para probar H_0: $\beta_j = 0$, se usa la razón:

$$F_0 = \frac{MS_B}{MS_E}$$

Para probar H_0: $(\tau\beta)_{ij} = 0$, se usa la razón:

$$F_0 = \frac{MS_{AB}}{MS_E}$$

Fórmulas para el cálculo de los cuadrados:

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{\cdots}^2}{abn} \qquad (14\text{-}5)$$

$$SS_A = \sum_{i=1}^{a} \frac{y_{i\cdot\cdot}^2}{bn} - \frac{y_{\cdots}^2}{abn} \qquad (14\text{-}6)$$

$$SS_B = \sum_{j=1}^{b} \frac{y_{\cdot j\cdot}^2}{an} - \frac{y_{\cdots}^2}{abn} \qquad (14\text{-}7)$$

$$SS_{AB} = \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{y_{ij\cdot}^2}{n} - \frac{y_{\cdots}^2}{abn} - SS_A - SS_B \qquad (14\text{-}8)$$

$$SS_E = SS_T - SS_{AB} - SS_A - SS_B \qquad (14\text{-}9)$$

Resumen para ANOVA con dos factores y modelo de efectos fijos

Table 14-4 ANOVA Table for a Two-Factor Factorial, Fixed-Effects Model

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A treatments	SS_A	$a - 1$	$MS_A = \dfrac{SS_A}{a - 1}$	$\dfrac{MS_A}{MS_E}$
B treatments	SS_B	$b - 1$	$MS_B = \dfrac{SS_B}{b - 1}$	$\dfrac{MS_B}{MS_E}$
Interaction	SS_{AB}	$(a - 1)(b - 1)$	$MS_{AB} = \dfrac{SS_{AB}}{(a - 1)(b - 1)}$	$\dfrac{MS_{AB}}{MS_E}$
Error	SS_E	$ab(n - 1)$		
Total	SS_T	$abn - 1$	$MS_E = \dfrac{SS_E}{ab(n - 1)}$	

Ejemplo

Aircraft primer paints are applied to aluminum surfaces by two methods: dipping and spraying. The purpose of the primer is to improve paint adhesion, and some parts can be primed using either application method. The process engineering group responsible for this operation is interested in learning whether three different primers differ in their adhesion properties.

A factorial experiment was performed to investigate the effect of paint primer type and application method on paint adhesion. For each combination of primer type and application method, three specimens were painted, then a finish paint was applied, and the adhesion force was measured. The data from the experiment are shown in Table 14-5. The circled numbers in the cells are the cell totals $y_{ij.}$. The sums of squares required to perform the ANOVA are computed as follows:

Observaciones realizadas

Table 14-5 Adhesion Force Data for Example 14-1

Primer Type	Dipping		Spraying		$y_{i\cdot\cdot}$
1	4.0, 4.5, 4.3	12.8	5.4, 4.9, 5.6	15.9	28.7
2	5.6, 4.9, 5.4	15.9	5.8, 6.1, 6.3	18.2	34.1
3	3.8, 3.7, 4.0	11.5	5.5, 5.0, 5.0	15.5	27.0
$y_{\cdot j\cdot}$	40.2		49.6		$89.8 = y_{\cdot\cdot\cdot}$

Cálculo de los cuadrados medios

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{...}^2}{abn}$$

$$= (4.0)^2 + (4.5)^2 + \cdots + (5.0)^2 - \frac{(89.8)^2}{18} = 10.72$$

$$SS_{\text{types}} = \sum_{i=1}^{a} \frac{y_{i..}^2}{bn} - \frac{y_{...}^2}{abn}$$

$$= \frac{(28.7)^2 + (34.1)^2 + (27.0)^2}{6} - \frac{(89.8)^2}{18} = 4.58$$

$$SS_{\text{methods}} = \sum_{j=1}^{b} \frac{y_{.j.}^2}{an} - \frac{y_{...}^2}{abn}$$

$$= \frac{(40.2)^2 + (49.6)^2}{9} - \frac{(89.8)^2}{18} = 4.91$$

Cálculo de los cuadrados medios

$$SS_{\text{interaction}} = \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{y_{ij\cdot}^2}{n} - \frac{y_{\cdots}^2}{abn} - SS_{\text{types}} - SS_{\text{methods}}$$

$$= \frac{(12.8)^2 + (15.9)^2 + (11.5)^2 + (15.9)^2 + (18.2)^2 + (15.5)^2}{3}$$

$$- \frac{(89.8)^2}{18} - 4.58 - 4.91 = 0.24$$

and

$$SS_E = SS_T - SS_{\text{types}} - SS_{\text{methods}} - SS_{\text{interaction}}$$

$$= 10.72 - 4.58 - 4.91 - 0.24 = 0.99$$

Análisis de varianza

Table 14-6 ANOVA for Example 14-1

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	f_0	P-Value
Primer types	4.58	2	2.29	28.63	2.7 × E-5
Application methods	4.91	1	4.91	61.38	4.7 × E-7
Interaction	0.24	2	0.12	1.50	0.2621
Error	0.99	12	0.08		
Total	10.72	17			

Análisis de varianza

- El experimentador ha decidido usar un α=0.05
- Puesto que $f_{0.05,2,12}$=3.89 y $f_{0.05,1,12}$=4.75, se concluye que los efectos principales del tipo de pintura tapaporo y del método de aplicación afectan la fuerza de adherencia
- Puesto que 1.5 < $f_{0.05,2,12}$, No hay indicios de interacción entre estos factores