

# Capítulo 16

## **Análisis e interpretación estadística de la experimentación in vitro**

J. A. Izquierdo\*

Y. López F.\*\*

---

\* Producción Vegetal, FAO/RLAC, Santiago, Chile.

\*\* Facultad de Ciencias Agropecuarias, Universidad Nacional de Colombia, Palmira, Colombia.

## **Introducción**

El valor de una tecnología depende de la capacidad que tenga la investigación para mantenerla en avance continuo. En lo que concierne al cultivo de células y tejidos vegetales, la obtención de soluciones técnicas implica la utilización empírica del método de prueba y error; la investigación debe validar las hipótesis sugeridas por experiencias anteriores, por observaciones y por el estado actual de la información en áreas específicas, acción que es esencial para la generación de tecnología en este campo del cultivo de tejidos.

## **Necesidad de la Evaluación Estadística**

Con relativa frecuencia, los investigadores —y muy especialmente los que realizan su trabajo en las condiciones del laboratorio— se preguntan: ¿Es la evaluación estadística una verdadera necesidad para este tipo específico de experimentación?

La falta de criterios elementales para que el investigador pueda contestar esta pregunta ha hecho que, frecuentemente, los resultados presentados en la literatura relacionada con el cultivo de tejidos no estén caracterizados por la consistencia de los datos y por la repetibilidad de los protocolos experimentales.

Para responder esa pregunta, es necesario tener en cuenta, en primer lugar, la variabilidad de los materiales biológicos que debe manipular el investigador, ya que la variación fenotípica (F) de los individuos es el resultado del efecto conjunto de la carga genética (G) por el efecto de los factores ambientales (A), según la expresión:

$$F = G \times A$$

Es posible deducir, por tanto, que las diferencias entre los individuos son debidas a diferencias genéticas y ambientales que se hallan fuera del control razonable del experimentador. El reconocimiento de la existencia de esta variabilidad implica reconocer la dificultad de evaluar una práctica que ha sido efectuada sobre una sola unidad experimental. Un experimento con una sola repetición suministra siempre una medida muy pobre del efecto de la práctica o tratamiento aplicado; cuando no hay al menos dos unidades experimentales tratadas similarmente, no es posible la medición del error experimental. Este es, precisamente, la medida de la variabilidad que se halla fuera del control del investigador.

Por esa razón, el tratamiento estadístico de una colección de datos experimentales permite hacer la estimación sin sesgo del efecto de los tratamientos sobre el material investigado, y la evaluación de las diferencias obtenidas entre los tratamientos mediante pruebas de significancia basadas en la medida del error experimental. Estas pruebas valoran la probabilidad de que las diferencias halladas entre los tratamientos ocurran únicamente como efecto del azar, es decir, de la parte no controlada del experimento.

Por consiguiente, se acepta que la forma razonable de conocer la magnitud de los efectos de ciertos factores sobre el material biológico aislado en forma de unidad (llámese ésta parcela, tubo de ensayo o caja Petri) y de medir el efecto del azar sobre esas unidades, es planear un experimento según criterios que fijen de antemano las hipótesis formuladas sobre él y los métodos específicos de llevarlo a cabo. Estos criterios determinarán las características de lo que se llama el diseño del experimento.

## **Principios Inherentes al Diseño de Experimentos**

Los investigadores consideran, en general, tres principios fundamentales en el diseño de un experimento, los cuales cobran una importancia particular en las áreas del cultivo de tejidos y de la biotecnología.

### **A. Repeticiones**

Significa que un tratamiento debe ser repetido, o asignado a dos unidades experimentales diferentes por lo menos, con el fin de proveer un estimador del error experimental y permitir una medida más precisa del efecto de los tratamientos.

#### **1. Número de repeticiones**

¿Cuántas veces debe aplicarse un tratamiento a la unidad experimental? En otras palabras: ¿Cuántas unidades experimentales por tratamiento hay que incluir en un diseño para que la estimación de la variación debida al error experimental sea lo más precisa posible?

Tal pregunta puede responderse de manera un tanto diferente según el campo al que pertenezca el experimento:

- Experimentos agrícolas: Se han escrito textos excelentes en donde se resuelve el problema del número de repeticiones de este tipo de experimentos; aunque ese número varía según las características de los materiales ensayados, las pautas encontradas para calcularlo por los biometristas y expertos en investigación de cultivos coinciden.
- Experimentos de biotecnología: La naturaleza de los materiales empleados y las condiciones físicas en las cuales se desarrolla la investigación en esta ciencia deben considerarse antes de establecer el número de repeticiones de estos experimentos.

**Características de los materiales.** Por su naturaleza, los materiales biológicos ensayados con más frecuencia (explantos, células aisladas, protoplastos) son, normalmente, de tamaño pequeño; por consiguiente, son muy sensibles a los pequeños cambios en los niveles de los factores externos. Una pequeña masa de células, por ejemplo, puede ser más sensible a un cambio minúsculo en la cantidad de minerales, vitaminas u hormonas de su medio que una planta entera que crece en un medio variable pero normal. Por tal razón, es muy probable que ese pequeño conjunto de células exprese características que, por causa de la regulación cibernética, un organismo completo más grande no pueda expresar ni con la misma rapidez ni en la misma magnitud.

Las altas variaciones que se presentan en este tipo de investigaciones tienen así una explicación, y deben tenerse en cuenta a la hora de determinar el número de repeticiones con que se diseña un experimento particular.

**Condiciones físicas de los experimentos.** a) *En los cuartos de crecimiento:* Se supone que estos cuartos están dotados de condiciones controladas; no obstante, están expuestos a condiciones locales y a gradientes de baja intensidad que pueden afectar la uniformidad con que cada factor ambiental llega a cada uno de los materiales ensayados en cada una de las unidades experimentales.

Así, por ejemplo, a veces es difícil garantizar plenamente que las fuentes de luz están colocadas de tal forma que la intensidad y la calidad de la luz son las mismas para todas las unidades experimentales; con frecuencia se detectan gradientes de iluminación que pueden alterar la respuesta de los materiales en estudio.

La humedad está asociada con diversos problemas patológicos difíciles de controlar en una cámara de crecimiento y que pueden afectar notablemente el curso y los resultados de los experimentos; esos problemas están ligados a fenómenos respiratorios en que las respuestas de los materiales a las concentraciones de oxígeno inducen un cambio diferencial en las

concentraciones parciales de otros gases como el  $\text{CO}_2$ . El efecto de este gas en la cámara de crecimiento también debe tenerse en cuenta. Aunque se supone que estas condiciones son uniformes para todas las unidades experimentales, en realidad se establecen coeficientes de variación que, sin tener en cuenta la discusión anterior, serían inexplicables.

Se dispone normalmente de cuartos de crecimiento y de cámaras controladas de pequeñas dimensiones; por consiguiente, el número de unidades experimentales está restringido por el factor espacio. Esta limitación debe considerarse cuando se proyecta un experimento y se calculan tanto el número de factores que se investigará como el número de repeticiones necesario para hacer una buena estimación del efecto de un factor dado.

b) *En el invernadero y la casa de malla*: Los materiales que se investigan en estos espacios son, casi siempre, de tamaño mayor que los materiales considerados en la sección anterior; se esperaría, por tanto, que sus respuestas a los efectos de los tratamientos fueran un tanto menos variables. Sin embargo, en un invernadero se encuentran también, con frecuencia, altos índices de variación que, como en el caso anterior, se explican por gradientes de luz, de temperatura o de humedad que pueden medirse con relativa facilidad. El tamaño relativamente grande de las unidades experimentales reduce las limitaciones de la escogencia tanto del número de tratamientos como del número de repeticiones a la hora de hacer el diseño del experimento.

## 2. Solución general al problema del número de repeticiones

La solución está contenida en la siguiente expresión:

$$N = C V^2 / \sigma^2 \quad (1)$$

donde:

V = coeficiente de variación

$\sigma$  = diferencia entre medias extremas

C = parámetro que es función del sesgo de error y de la distribución de las medias.

Los valores de V,  $\sigma$  y C deben encontrarse empíricamente, es decir, el empleo de la fórmula implica el conocimiento previo de la variabilidad de los materiales que se investigarán, y éste se obtiene de un ensayo diseñado para tal propósito.

Se puede afirmar, entonces, que el número de repeticiones de un experimento depende de la magnitud de las diferencias que se quieran detectar y

de la variabilidad de los datos con los cuales se trabaja. En la práctica, esta afirmación es una limitación importante, ya que para determinar el número de repeticiones se debe gastar tiempo, trabajo y materiales, es decir, aumentar el costo del ensayo, condición que muchas instituciones aceptan con dificultad. El número de repeticiones queda así limitado por la dificultad de conocer de antemano la variabilidad de los materiales ante las condiciones del ensayo.

Como se deduce de la expresión (1), otra limitación de este problema es el hecho de que en los ensayos de laboratorio se encuentran con frecuencia coeficientes de variabilidad muy grandes. Siendo el número de repeticiones directamente proporcional al cuadrado del coeficiente de variación, el valor de  $N$  será muy alto (hasta 30 ó 40 repeticiones) de tal modo que, siguiendo sólo este criterio, los experimentos se harían impracticables.

En estos casos, es indispensable el criterio biológico del investigador quien debe conocer, tanto como sea posible, las características biológicas de sus materiales y especialmente lo que se espera de ellos no sólo respecto a su variabilidad sino a su respuesta biológica.

## **B. Aleatorización**

Este principio permite obtener estimaciones sin riesgo del error experimental y de las medias de los tratamientos. Es muy importante en el diseño de experimentos de biotecnología porque en éstos hay grandes variaciones.

La aleatorización consiste en la asignación de los tratamientos a las unidades experimentales de tal forma que todas ellas tengan una oportunidad equitativa de recibir el tratamiento. Para tal efecto se acude a tablas de números aleatorios, a conteos sencillos con fichas, o a tablas generadas por un computador.

## **C. Control local**

La gran mayoría de los experimentos que se llevan a cabo en biotecnología tienen diseños factoriales en arreglos hechos completamente al azar; su ventaja es la sencillez pero su desventaja es que no consideran las limitaciones impuestas por los gradientes antes mencionados y esto afecta la magnitud del error experimental.

El control local, mediante restricciones en la aleatorización del diseño, significa una reducción en el error experimental. Este control se hace por el

'sistema de bloques' que consiste en la asignación sistemática de las repeticiones en bloques según el sentido del gradiente; de este modo se garantiza que las repeticiones de cada tratamiento reciban efectos del factor considerado (luz, humedad, temperatura) que tengan la misma magnitud. La práctica de los bloques permite entonces extraer el efecto del gradiente del total del error experimental.

## **Método Científico**

Es importante recordar algunas nociones relacionadas con la planeación del experimento. En general, se admite que el método científico es el procedimiento empleado para hacer investigación, y que comprende cuatro actividades:

- a. Observación de los hechos o fenómenos.
- b. Elaboración de hipótesis que sirvan como marco de explicación de tales fenómenos.
- c. Experimentación.
- d. Interpretación de los resultados y confrontación de éstos con las hipótesis.

Por consiguiente, el experimento es una herramienta que el investigador utiliza para descubrir algo desconocido o para someter a prueba un principio o una hipótesis.

### **Características de un buen experimento**

Un experimento bien planeado ostenta cuatro características básicas:

**Simplicidad.** Con frecuencia se pretende responder a demasiadas preguntas con un solo experimento; esta actitud complica el diseño y el manejo del experimento de tal modo que la información obtenida es confusa y difícil de interpretar desde el punto de vista biológico. Lo razonable es diseñar experimentos en que los tratamientos y el arreglo experimental sean consistentes con pocos objetivos bien definidos.

**Grado de precisión.** Se garantiza que el experimento puede medir diferencias entre los tratamientos con el grado de precisión que requiere el investigador. Por ello, el manejo mismo del experimento, en su aspecto operativo, es de suma importancia.

**Ausencia de error sistemático.** Se garantiza esta característica mediante la aleatorización y el sistema de bloques, los cuales reducen tanto el sesgo

debido a una asignación sistemática de los tratamientos como el efecto sistemático de los gradientes.

**Validez de las conclusiones.** En un experimento bien planeado se pretende lograr la universalidad de las conclusiones, es decir, que tengan un rango de validez tan amplio como sea posible. Una forma de conseguirlo es distribuyendo las repeticiones en el tiempo y en el espacio; de este modo se sabe si el efecto de los factores sobre la unidad experimental varía de un mes a otro, de un semestre a otro, de la época seca a la lluviosa, o de un sitio a otro.

Tres normas sencillas permiten pues adquirir las características mencionadas: a) diseñar repeticiones; b) hacer la aleatorización de los tratamientos; y c) solicitar la ayuda del estadístico o del biometrista si surgen dificultades de diseño o de interpretación estadística de los resultados. La mayor parte de los pasos que se dan en el manejo de un experimento no son de orden estadístico; no obstante, el manejo estadístico constituye una parte muy importante de la experimentación.

Los criterios anteriores son especialmente importantes en el desarrollo de la experimentación en el cultivo de células y tejidos. Aunque a menudo se considera que las condiciones *in vitro* ofrecen un mayor control de la variabilidad externa, la utilización cuidadosa de la estadística, en este caso, es una condición inherente a la credibilidad actual y futura de esta nueva biotecnología. La consistencia y la repetibilidad de los resultados deben estar bajo el amparo estricto del rigor estadístico; estos criterios faltan a menudo en publicaciones de trabajos sobre cultivo de tejidos.

El presente capítulo sugiere a los investigadores criterios útiles para la interpretación estadística de los resultados en el cultivo de células y tejidos. No pretende ser una contribución a la tan poblada área de la estadística ya que existen numerosos y apropiados libros y revisiones sobre estadística experimental agrícola (Cochran, 1971; Cochran et al., 1950; Cox, 1958; Draper et al., 1966; Little et al., 1975; Steel et al., 1960), los cuales se pueden y deben utilizar como guía en el diseño de los experimentos y en los análisis de los resultados.

Este capítulo está dedicado a investigadores que comprenden la necesidad del criterio probabilístico y que, habiendo tomado algún curso básico de estadística experimental, aceptan que el análisis de varianza (ANOVA) es el mecanismo de prueba de hipótesis.

## Interpretación de las Respuestas a las Dosis

El propósito obvio de un experimento es resolver las preguntas planteadas en las hipótesis; sin embargo, muchas veces los resultados se interpretan y presentan con poca o ninguna referencia a la pregunta original.

En el cultivo de tejidos, una de las preguntas que se formulan más comúnmente es la que se refiere a la dosis óptima de un regulador del crecimiento u otro constituyente del medio de cultivo; también es frecuente interrogarse sobre los factores ambientales (temperatura, humedad, etc). El siguiente es un ejemplo acerca de la relación dosis-respuesta.

Se trata de un experimento cuyo objetivo es determinar el efecto del  $Al^{+++}$  en la reducción del crecimiento de callos originados a partir de cotiledones de alfalfa; se hicieron 7 tratamientos consistentes en 7 dosis de  $Al^{+++}$  aplicado al medio de cultivo, y la toxicidad se evaluó en términos de peso fresco (variable 1) y de diámetro del callo (variable 2), a los 30 días (Cuadro 16.1).

Cuadro 16.1. Respuesta<sup>a</sup> de los callos de cotiledones de alfalfa a varias dosis de aluminio, en términos de peso fresco y diámetro.

$Al^{+++}$ (dosis)	Variable 1 (mg de peso fresco)	Variable 2 (mm de $\Phi$ )
0	222 a	9 d
1	202 ab	12 cd
2	205 ab	15 bc
3	186 bc	22 a
4	164 cd	19 ab
5	156 cd	17 b
6	147 d	11 cd

a. Respuestas a los 30 días de iniciado el tratamiento.

La pregunta simple, útil y directa que el investigador debería haberse planteado era: ¿Qué relación existe entre la dosis de  $Al^{+++}$  y la respuesta del tejido? Sin embargo, a menudo predomina otra pregunta menos útil y que además es mucho más difícil de responder: de los 21 posibles pares (o combinaciones) de tratamientos, ¿cuál o cuáles son diferentes respecto a los demás?

En forma general, en las publicaciones sobre cultivos de tejidos se observa una tendencia a responder el segundo tipo de pregunta, con resultados similares a los que presenta el Cuadro 16.1. Es obvio que la primera pregunta no ha sido contestada en el caso de la variable 1 (peso fresco) a pesar de que es inherente al objetivo del experimento; para obtener esa respuesta es necesario hacer la distribución de la suma de cuadrados (Cuadro 16.2).

Cuadro 16.2. Distribución de la suma de cuadrados (S.C.) de los tratamientos en cuanto al peso fresco (variable 1).

Fuente de variación	g.l.	S.C.	Distribución (% del total)
Tratamiento	6	23805	—
Efecto lineal	1	22885	96
Efecto residual	5	920	4

La situación que presenta el Cuadro 16.2 puede ilustrarse en la Figura 16.1, en donde se observa una relación lineal altamente significativa que explica el 96% de la variación de la respuesta. En estas condiciones la segunda pregunta, sobre cuáles tratamientos son significativamente diferentes de los demás, es irrelevante; estamos frente a una relación lineal y, dentro del rango probado, todos los niveles (dosis) de tratamiento producen efectos siempre diferentes. El mejor estimador de los efectos de los tratamientos está constituido por los puntos de la ecuación de regresión.

Existen otras relaciones entre dosis y respuestas. La variable 2 (diámetro) presenta una respuesta curvilínea (Figura 16.2) y la interpretación de estos resultados puede ser aún más confusa en términos de la identificación de los tratamientos con respuestas significativamente diferentes. Sin embargo, la participación de los efectos de los tratamientos muestra que los componentes lineales y cuadráticos explican el 85% de la variabilidad en las respuestas (Cuadro 16.3). Por lo tanto, se podría resumir simplemente diciendo que la respuesta encontrada en este caso es una curva de segundo grado.

Los ejemplos anteriores tienen el objetivo de mostrar que siempre que se establezca una serie de dosis se debe intentar un esfuerzo por encontrar la relación dosis-respuesta.

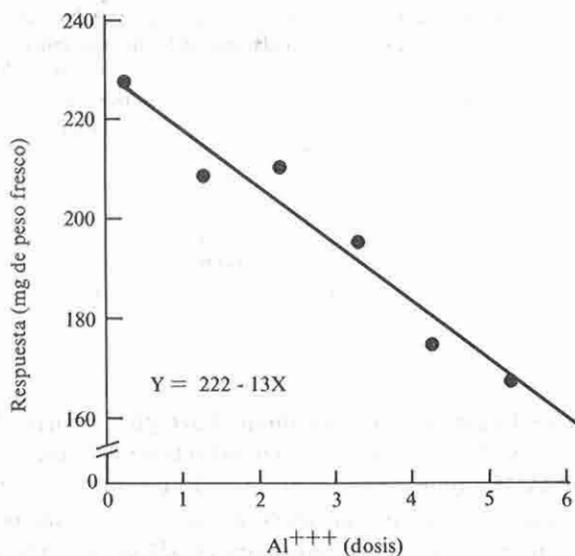


Figura 16.1. Regresión lineal que representa la respuesta del callo (mg de peso fresco) a las diferentes dosis de  $Al^{+++}$  del medio de cultivo.

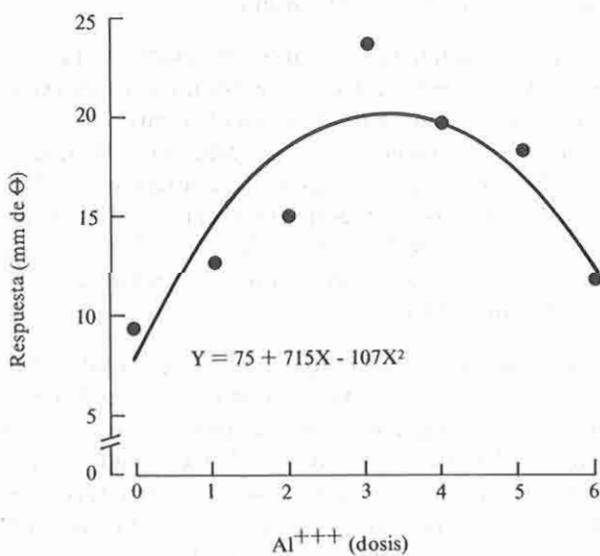


Figura 16.2. Curva de segundo grado que muestra la relación entre la dosis de  $Al^{+++}$  y la respuesta de los callos (mm de Ø)

Cuadro 16.3. Distribución de las sumas de cuadrados (S.C.) de los tratamientos respecto al diámetro del callo (variable 2).

Fuentes de variación	g.l.	S.C.	Distribución (% del total)
Tratamiento	6	650.0	—
Efectos:			
Lineal	1	71.5	11
Cuadrático	1	482.0	74
Cúbico	1	40.8	6
Residual	3	55.7	9

En la experimentación *in vitro* es común el arreglo factorial de tratamientos. En este arreglo se utilizan dos o más factores a distintas concentraciones y en todas las combinaciones posibles. Es un enfoque diferente de los anteriores; en éstos se quería saber cómo era la relación dosis-respuesta. Las preguntas lógicas ahora son: ¿Cuál es el efecto de cada factor? ¿Hay interacciones entre los factores? ¿Depende la respuesta de un factor del nivel del otro?

El Cuadro 16.4 corresponde a un ejemplo en el cual se obtuvieron las medias de respuestas a cada tratamiento o combinación de tres factores con dos niveles cada uno (arreglo factorial  $2^3$ ).

Las letras con que se intenta clasificar los promedios explican confusamente los efectos de los tratamientos. En cambio, si se realiza el reparto de la suma de cuadrados debida a los tratamientos entre los efectos de los principales de éstos y los efectos de las interacciones (Cuadro 16.5), se encuentra que el factor A tiene el efecto más importante y explica más del 83% del total de la variación; el factor B cuenta por el 12% de la suma de cuadrados, y no existe evidencia de efecto del factor C o de las interacciones. En este caso entonces sólo es importante presentar los efectos de los factores A y B (Cuadro 16.6).

En un nuevo ejemplo, combinando el análisis factorial y el estudio de tendencias en las respuestas de los tratamientos, se pretende ilustrar cómo se pueden interpretar y presentar las interacciones. Si se toman los niveles  $2 \times 2 \times 4$  en los tres factores respectivos A, B y C, siendo C el tiempo de exposición, obviamente las preguntas que se deben formular son las siguientes: ¿Cuáles son los efectos de los factores A y B? ¿Existe relación entre la respuesta y el tiempo de exposición? ¿Existe alguna interacción significativa entre los tres factores? Estas preguntas no se pueden responder con una presentación como la del Cuadro 16.7.

Cuadro 16.4. Respuestas a los tratamientos de un arreglo factorial 2<sup>3</sup>.

Tratamientos (combinación de factores)	Respuestas
A B C	14.4 b
A B c	8.5 b
A b C	15.2 ab
A b c	13.0 ab
a B C	21.4 a
a B c	18.6 ab
a b C	22.6 a
a b c	22.7 a

Cuadro 16.5. Distribución de la suma de cuadrados (S.C.) del arreglo factorial 2<sup>3</sup>.

Fuente de variación	g.l.	S.C.	Efecto relativo (% del total)
Tratamientos	7	877	—
Factor A	1	730	83
Factor B	1	107	12
Factor C	1	23	3
A x B	1	8	} 2
A x C	1	1	
B x C	1	3	
A x B x C	1	5	

Cuadro 16.6. Efecto promedio de los factores.

Factor	Efecto promedio <sup>a</sup>
A	11.8
a	21.3
B	14.7
b	18.4

a. Se expresa en las unidades de los datos originales o de las respuestas indicadas en el Cuadro 16.4.

Cuadro 16.7. Respuesta a la interacción de los factores A, B y C.

Nivel de A	Nivel de B	Respuestas según nivel de C			
		0	1	2	3
1	1	74 c	92 c	108 b	134 b
1	2	48 b	108 d	156 c	292 d
2	1	12 a	18 a	76 a	92 a
2	2	18 a	58 b	90 ab	162 c

Si se distribuyen los 15 grados de libertad de los tratamientos, se puede encontrar una explicación acerca de los efectos de varios factores y de sus interacciones (Cuadro 16.8). Se observa que el 92% de la variación en los tratamientos se debe a cuatro efectos: el principal de A, el principal de B, la tendencia lineal del factor C, y la interacción entre el factor B y la tendencia lineal del factor C. Estos resultados se pueden presentar en forma gráfica (Figuras 16.3 y 16.4).

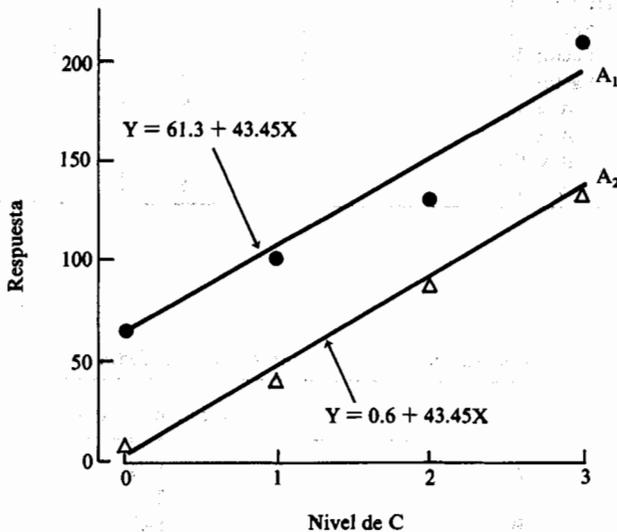


Figura 16.3. Regresiones lineales que representan la respuesta (Y) a los dos niveles del factor A en los cuatro niveles del C.

Cuadro 16.8. Distribución de la suma de cuadrados (S.C.) de los tratamientos.

Fuente de variación	g.l.	S.C.	Distribución (% del total)
<b>Tratamientos</b>	<b>15</b>	<b>216083</b>	<b>—</b>
<b>Factor A</b>	<b>1</b>	<b>44287</b>	<b>20.5</b>
<b>Factor B</b>	<b>1</b>	<b>19927</b>	<b>9.2</b>
<b>A x B</b>	<b>1</b>	<b>817</b>	<b>0.4</b>
<b>C lineal</b>	<b>1</b>	<b>113274</b>	<b>52.4</b>
<b>C cuadrático</b>	<b>1</b>	<b>2977</b>	<b>1.4</b>
<b>C cúbico</b>	<b>1</b>	<b>163</b>	<b>0.1</b>
<b>A x C lineal</b>	<b>1</b>	<b>1717</b>	<b>0.8</b>
<b>B x C lineal</b>	<b>1</b>	<b>21094</b>	<b>9.8</b>
<b>Residual<sup>a</sup></b>	<b>7</b>	<b>11827</b>	<b>5.5</b>

a. Corresponde a la S.C. debida a los tratamientos y no incluida en la descomposición anterior. Algunos autores llaman 'residuo' al error experimental.

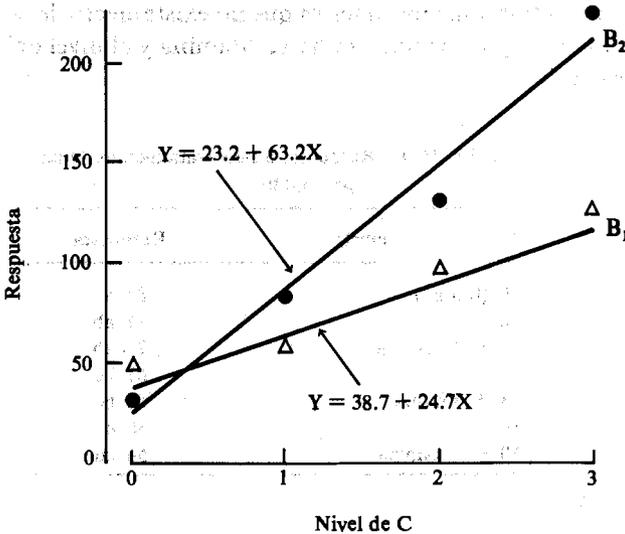


Figura 16.4. Regresiones lineales que representan la respuesta a los dos niveles del factor B en los cuatro niveles del C.

El experimento ilustra otro punto importante: que el significado de la F de los tratamientos con 15 grados de libertad es confuso; así, en el ejemplo se observa que F es el promedio de 4 grados de libertad altamente significativos y de 11 que no lo son. En estas circunstancias muchas veces los investigadores detienen el análisis estadístico y la interpretación cuando encuentran que el F de los tratamientos no es significativo, lo cual ha conducido a una interpretación incompleta e incorrecta de los experimentos. Por lo tanto, importa en estos casos distribuir los grados de libertad de los tratamientos, para visualizar los efectos de los factores principales y de las interacciones.

Al respecto se puede analizar otro ejemplo: hay siete tratamientos, uno de los cuales se considera como testigo, y se establecen tres niveles en la relación 2:5:10, aplicados con o sin adición de una vitamina. Los resultados se presentan en el (Cuadro 16.9).

En este ejemplo la factorización fue incompleta. Específicamente, se debió incluir un tratamiento con la vitamina sola; además, no existe aparentemente una justificación lógica para la serie particular de tratamientos. Es recomendable que, cuando se realicen estudios sobre la relación entre dosis y respuestas, se utilicen series de niveles establecidos en progresiones aritméticas. A pesar de lo anterior, la distribución de los efectos de los tratamientos muestra que no existe efecto de las vitaminas y que tampoco hay interacción entre la vitamina y el nivel del tratamiento (Cuadro 16.10).

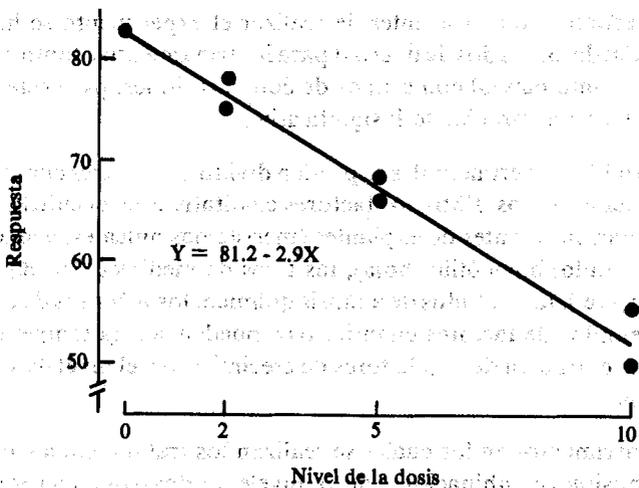
Cuadro 16.9. Respuesta a siete tratamientos de un experimento.

Tratamiento (nivel)	Respuesta
0 (testigo)	81 a
2	77 ab
2 + Vitamina	74 ab
5	67 bc
5 + Vitamina	66 bc
10	50 d
10 + Vitamina	56 cd

Los resultados descartan el efecto de la vitamina. El análisis de regresión (Figura 16.5) indica que en la distribución de las sumas de los cuadrado de los tratamientos, un 98% de la variación se explica por una respuesta de tipo lineal (Cuadro 16.11).

**Cuadro 16.10. Distribución de la suma de cuadrados (S.C.) de tratamientos de otro experimento.**

Fuente de variación	g.l.	S.C.	Distribución (% del total)
Tratamientos	6	3216	
Testigo vs. otros	1	915	28
Lineal vs. otros	1	2249	70
Vitamina	1	0	0
Residual	3	52	2



**Figura 16.5. Recta de regresión que representa la respuesta dada al nivel de la dosis.**

**Cuadro 16.11. Distribución de la suma de cuadrados de tratamientos.**

Fuente de variación	g.l.	S.C.	Distribución (% del total)
Tratamientos	6	3216	
Lineal	1	3164	98
Desviación del modelo lineal	5	52	2

Se ha mencionado anteriormente la distribución de los efectos de los tratamientos como ejemplo del manejo e interpretación de los resultados. Sin embargo esta técnica, altamente recomendada por los estadísticos, es a menudo desconocida o no se considera en los análisis estadísticos de los resultados que se publican abundantemente en nuestros días. Por otro lado, se nota un abuso en la utilización de gráficas y cuadros de valores promedio más o menos la desviación estándar. Por ello la técnica de la distribución de las sumas de cuadrados de los tratamientos merece una especial consideración.

## **Distribución de los Efectos de los Tratamientos**

Es muy recomendable que antes de realizar el experimento se haga una distribución de los grados de libertad para los tratamientos, tanto en forma individual como parcial con grupos de comparaciones; para este reparto no se requiere una prueba de F significativa.

Las variables experimentales se pueden dividir en factores cualitativos y factores cuantitativos. Entre los factores cualitativos en el cultivo in vitro pueden estar: las fuentes de explantes (meristemas apicales, yemas axilares, raíces, tallo, hipocótilo, hoja), los tipos de medios (MS, B5, White, Anderson, etc.), los métodos de análisis químico, los métodos de conteo, y otros. Ejemplos de factores cuantitativos pueden ser: la temperatura, el pH, la concentración de reguladores de crecimiento y el nivel de sacarosa, entre otros.

Los experimentos en los cuales se realizan los tratamientos a partir de todas las posibles combinaciones de los niveles de dos o más factores, tanto cualitativos como cuantitativos, son llamados experimentos factoriales. En este caso, el término factorial solamente describe la naturaleza del tratamiento y no el diseño del mismo.

Cuando se establecen factores cuantitativos, el análisis de regresión es la técnica más apropiada (Draper et al., 1966). En este caso, los grados de libertad de los tratamientos y la suma de cuadrados se deben distribuir entre los componentes debidos al efecto lineal, al efecto cuadrático, al efecto cúbico, y a otros polinomios de orden superior. Si hay suficiente conocimiento teórico para especificar una relación matemática entre las respuestas y el factor (relación logística, Mitscherlich, Gompertz, etc.), se debe aplicar la ecuación particular a los resultados.

Sin embargo, en la mayoría de la experimentación in vitro la relación matemática es tan compleja que todavía no se han encontrado buenas

aproximaciones, mediante modelos polinómicos apropiados, para explicar fenómenos tales como la diferenciación o la embriogénesis somática, ni aun sucesos tan comunes como la organogénesis adventicia. Si los niveles de los tratamientos están igualmente espaciados e igualmente repetidos, la computación para obtener los componentes lineales, cuadráticos, y de otro grado de las regresiones se simplifica por el uso de polinomios ortogonales (Little et al., 1975; Steel et al., 1960; Gómez et al. 1976).

## **Procedimientos de Comparación Múltiple**

Existen situaciones en las cuales los procedimientos de comparación múltiple, como es la prueba de rango triple de Duncan, resultan apropiados (Petersen, 1977); esta situación se presenta cuando se compara un arreglo al azar de genotipos o sustancias químicas. En estos casos, un número de pruebas estadísticas puede ayudar a la interpretación de los resultados de los experimentos. Sin embargo, los resultados que se han publicado, en cultivo de tejidos, fisiología, horticultura y otras áreas, denotan un abuso en la utilización de las comparaciones múltiples.

Uno de los objetivos, en el caso discutido, es poner alerta al investigador de un cultivo in vitro acerca del mal uso de esas técnicas. Las comparaciones múltiples no son apropiadas en los siguientes casos: a) en experimentos en los cuales los tratamientos están establecidos a partir de niveles de una variable cuantitativa; b) para combinaciones factoriales de dos o más factores, o dos o más niveles en cada factor; c) para tratamientos cuantitativos que fueron formulados en combinaciones lineales de particular interés (Petersen, 1977).

Existen muchos textos de estadística que contienen discusiones sobre este tema (Gómez et al., 1976; Little et al., 1975; Steel, et al., 1960) y también trabajos de revisión (Chew 1976 y 1980; Petersen, 1977); por lo tanto, sólo es importante delinear sucintamente las ventajas y desventajas de cada prueba.

### **Diferencia mínima significativa de Fisher**

La diferencia mínima significativa (DMS) consiste en aplicar la prueba t de Student a cualquier par de medias, pero se debe usar solamente si en el análisis de varianza el valor F del tratamiento es significativo; con esta protección, la DMS se transforma en una de las más poderosas técnicas de comparación múltiple (Chew, 1976; 1980).

Considerando que esta comparación múltiple es ampliamente conocida por los investigadores de cultivos in vitro, solamente queda agregar que, aparte de ofrecer la protección necesaria, esta prueba permite utilizar tratamientos desigualmente repetidos y también puede servir para la estimación de intervalos.

### **Prueba de rango múltiple de Newman-Keuls**

Este método no requiere un valor  $F$  significativo en el análisis de varianza y es utilizable cuando los tratamientos están igualmente repetidos. Para aplicarlo se organizan los promedios en forma ascendente, pero en vez de comparar las diferencias entre cualquier par de medias contra un valor constante, como en la DMS, se utiliza una magnitud variable. A pesar de que esta prueba es intuitivamente más interesante que la DMS, no ha sido muy aceptada y, además, tiene una tasa de error bastante confusa (Chew, 1980).

### **Diferencia significativa de Tukey (HSD)**

Esta prueba requiere un número igual de repeticiones por tratamiento. Presenta la simplicidad de la DMS, en la cual una magnitud constante permite probar todas las diferencias entre todos los pares de tratamientos; dos tratamientos se declaran diferentes si la magnitud absoluta de la diferencia entre sus medias excede el HSD.

El HSD de Tukey se puede usar también para construir intervalos de confianza simultáneos para todos estos pares de tratamientos. En este caso, se denomina prueba de rango múltiple de Tukey, y el HSD tiene un criterio altamente conservador; sin embargo, ofrece la ventaja de su aplicabilidad.

### **Método de Scheffé**

El método de Scheffé, al igual que el HSD de Tukey, es aplicable a contrastes generales y no solamente a comparaciones apareadas. Puede tener una aplicación específica en el caso de que los promedios provengan de observaciones de diferentes tratamientos con varianzas desiguales.

A pesar de esta restricción, que inclusive involucra la prohibición de realizar el análisis general de varianza, el método de Scheffé podría ser utilizado en estos casos.

## **Prueba de rango múltiple de Duncan**

Esta prueba, que constituye el método de comparación múltiple más utilizado, supone que las medias de los tratamientos no están correlacionadas. El procedimiento es similar al de la prueba de Newman-Keuls excepto en el nivel de protección que éste necesita, ya que según Chew (1980) existe una mayor probabilidad de error del tipo I. Esto hace que la prueba de rango múltiple de Duncan sea muy poderosa, lo cual significa una mayor probabilidad de detectar las diferencias cuando ellas existen.

Los investigadores están a menudo más interesados en encontrar diferencias que en no encontrarlas, y por esta razón el procedimiento de Duncan ha recibido una amplia aceptación entre ellos.

Originalmente, esta prueba fue propuesta en el análisis de varianza sin la protección de un valor  $F$  significativo de tratamientos, pero ahora se exige la significancia de la prueba  $F$  como una condición necesaria para su aplicación. Una desventaja del método es que no posee capacidad para estimar los intervalos simultáneos.

En determinados sentidos, las pruebas DMS, Newman-Keuls y Tukey son casos particulares de la prueba de Duncan, y ésta se considera apropiada para comparaciones apareadas de medias no correlacionadas con igual número de repeticiones.

En contraste con la prueba DMS, en la cual se calcula solamente un valor que se usa en todas las comparaciones, la prueba de rango múltiple de Duncan establece un conjunto de diferencias significativas que van aumentando según la distancia entre las posiciones o el ordenamiento de las dos medias que se comparan. A pesar de que se necesitan operaciones adicionales para la aplicación de la prueba de Duncan, ésta se puede usar para probar diferencias entre pares de tratamientos sin considerar el número de los mismos.

Se han cometido abusos con la prueba de Duncan al mostrar diferencias sin la requerida protección de un valor  $F$  significativo en el ANDEVA, y eso ha llevado al descrédito de la misma. Sin embargo, estudios estadísticos detallados muestran que la prueba de rango múltiple de Duncan es una herramienta útil, si se utiliza apropiadamente.

## **Comparaciones con el testigo**

El procedimiento de Dunnett presenta una forma de establecer intervalos simultáneos o comparaciones múltiples del testigo con cada uno de los

otros tratamientos; en él se supone que los tratamientos tienen igual número de repeticiones.

Este procedimiento permite separar todos los tratamientos que son tan buenos como el testigo o mejores que él (Chew, 1976; Gómez et al., 1976). Tiene total validez cuando el investigador no desea ajustar una curva a las dosis, y le interesa comparar el testigo con cada uno de los otros niveles. Este es el caso de algunos estudios de toxicidad donde el objetivo del experimento es determinar la dosis más baja a la cual todavía existe una respuesta observable.

Hechas estas consideraciones, es importante recalcar que las pruebas de comparaciones múltiples son totalmente inapropiadas si existe la posibilidad de repartir los tratamientos en sus sumas cuadradas específicas. Esta alternativa siempre es factible si los tratamientos son de estructura factorial, o si constan de varios niveles de un factor cuantitativo. En este último caso, si existe una regresión significativa entre la respuesta y las dosis, todos los niveles (incluyendo los intermedios) no utilizados en el experimento son estadísticamente diferentes; se recomienda entonces utilizar comparaciones múltiples sólo en el caso de factores cualitativos.

Se han realizado esfuerzos para determinar el grado de precisión con que cada comparación múltiple realiza la separación de medias. Por ejemplo, se ha comprobado que la DMS de Bayes es casi tan poderosa como la DMS de Fischer, sin tener la alta probabilidad de errores de tipo I que caracteriza a este último.

## **Heterogeneidad in Vitro**

Generalmente, la heterogeneidad ambiental en experimentos in vitro se subestima aduciendo que existe un control de las condiciones (iluminación, temperatura, HR) bajo las cuales se obtienen las respuestas. Esto no está suficientemente respaldado en las publicaciones y, generalmente, no se indica la cantidad de variación explicada por los tratamientos (coeficiente de variación).

Se sabe que tubos adyacentes en los cuales se siembra simultáneamente el mismo tipo de explante con el mismo medio producen resultados diferentes; entre las causas de esa falta de consistencia figuran los gradientes de iluminación y de temperatura dentro de las cámaras y los cuartos de crecimiento. En estas condiciones es acertado caracterizar la heterogeneidad ambiental, lo que puede ser a veces un prerrequisito para la elección de técnicas experimentales de alta precisión.

Muchas veces el investigador se enfrenta a un inapropiado diseño del cuarto de cultivo por errores de construcción o de distribución de los equipos. Factores como la localización de puertas y unidades de aire acondicionado, el calor desprendido por tubos, lámparas o equipo auxiliar, el tipo de estanterías, el sistemas de iluminación, el aislamiento y las dimensiones de los cuartos influyen en la creación de gradientes en las condiciones básicas. Si tales factores son modificables a un costo y tiempo apropiados, el investigador no debe dudar en mejorar sus condiciones ambientales; si no es posible, es oportuno conocer, por lo menos, con qué heterogeneidad está trabajando.

El índice de heterogeneidad proporciona un valor de caracterización y es menos complicado que realizar una prueba de uniformidad. Dicho índice requiere sólo de datos de experimentos con repeticiones donde se puedan obtener distintos tamaños de unidades experimentales, y está basado en la relación que existe entre la varianza de la parcela y el tamaño de la misma (Gómez et al., 1976):

$$V_x = \frac{V_1}{x^b}; \log V_x = \log V_1 - b \log x \quad (2)$$

Por consiguiente, renombrando los términos:

$$\log V_x = a + b \log x \quad (3)$$

siendo  $b$  el coeficiente de heterogeneidad.

**Ejemplo:** Se trata de un experimento con un diseño de parcelas divididas, constituido por seis parcelas principales (niveles de citocinina) y 13 genotipos (meristemas apicales de cultivares de fresa) como subparcelas; se hicieron tres repeticiones. La unidad experimental consistió en 10 meristemas cultivados en un frasco de un litro, de sección cuadrada, y que ocupa un área de 80 cm<sup>2</sup>. La parcela principal ocupa un área de 1100 cm<sup>2</sup> (los frascos se ubican en fila dejando 5 cm entre uno y otro). El área de cada bloque (repetición) es de 6600 cm<sup>2</sup>.

Del análisis de varianza se obtienen los cuadrados medios (CM) correspondientes a los bloques ( $M_1$ ), a la parcela principal [error  $a$ : ( $M_2$ )] y a la subparcela [error  $b$ : ( $M_3$ )] Con estos CM se pueden calcular las varianzas entre parcelas de varios tamaños.

## Varianza

Las expresiones para calcular las varianzas son:

$$V'_1 = M_1 \quad (4)$$

$$V'_2 = \frac{r(a-1)M_2 + (v-1)M_1}{r \cdot a} \quad (5)$$

$$V'_3 = \frac{r \cdot a(b-1)M_3 + r(a-1)M_2 + (r-1)M_1}{(r \cdot a \cdot b) - 1} \quad (6)$$

El análisis se presenta numéricamente en el Cuadro 16.12.

Cuadro 16.12. Análisis de varianza para un diseño de parcelas divididas con 6 parcelas principales y 13 subparcelas.

Fuente de variación	g.l.	CM
Bloques ( $M_1$ )	2	1,141,670
Parcela principal (PP)	5	5,480,847
Error $a$ ( $M_2$ )	10	433,767
Subparcela (SP)	12	17,294,508
PP x SP	60	682,046
Error $b$ ( $M_3$ )	144	330,593

Las varianzas para las distintas parcelas serían:

$$V'_1 = 1,141,670$$

$$V'_2 = 488,325$$

$$V'_3 = 344,197$$

Sin embargo, antes de comparar esas varianzas es necesario convertirlas a otras que estén en relación con un área comparable, así:

$$V_1 = V'_1/78 = 14,637$$

$$V_2 = V'_2/13 = 37,563$$

$$V_3 = V'_3/1 = 344,197$$

Con la fórmula (3) se puede calcular una regresión lineal entre las varianzas  $V$  y el tamaño de la parcela, en términos del número de unidades

que la componen. Se obtuvieron un coeficiente de regresión  $b = 0.74$  y un coeficiente de determinación  $R^2 = 0.984^{**}$ .

El valor  $b$  es una medida cuantitativa de la heterogeneidad en el ambiente, y en valor absoluto indica el grado de correlación entre parcelas experimentales adyacentes. El valor de  $b$  varía de  $+1$  a  $0$ . Cuanto más grande (en términos absolutos) sea el valor de  $b$ , más baja será la heterogeneidad ambiental. En casos en que se detecten coeficientes bajos ( $b = 0.25$ ), ellos se pueden utilizar para reorientar los bloques, el tamaño de la parcela experimental o el número de repeticiones para obtener un mayor control experimental.

## Referencias

- Chew, V. 1976. Comparing treatment means: A compendium. *HortScience* 11:348-357.
- . 1980. Testing differences among means: Correct interpretations and some alternatives. *HortScience* 15:467-470.
- Cochran, W. G. 1971. Técnicas de muestreo. Consejo Económico Centroamericano (CEC), México.
- y Cox, G. N. 1950. *Experimental designs*. Wiley and Sons, Nueva York.
- Cox, D. R. 1958. *Planning of experiments*. Wiley and Sons, Nueva York.
- Draper, N. R. y Smith, H. 1966. *Applied regression analysis*. Wiley and Sons, Nueva York.
- Dudewicz, E. J. 1976. Ranking and selection procedures. En: *Introduction to statistics and probability*. Holt, Rinehart and Winston, Nueva York.
- Gómez, K. A. y Gómez, A. A. 1976. *Statistical procedures for agricultural research with emphasis on rice*. International Rice Research Institute (IRRI), Los Baños, Filipinas.
- Little, T. M. 1981. Interpretation and presentation of results. *HortScience* 16:637-640.
- . 1978. If Galileo published in *HortScience*. *HortScience* 13:504-506.
- y Hills, F. J. 1975. *Statistical methods in agricultural research*. University of California, Davis, California, E.U.
- Petersen, R. G. 1977. Use and misuse of multiple comparison procedures. *Agron. J.* 69:205-206.
- Steel, R. G. D. y Torrie, J. H. 1960. *Principles and procedures of statistics*. McGraw-Hill, Nueva York.