

Modelo de diseño unifactorial completamente aleatorizado

Introducción y ejemplos

Este modelo es el más sencillo del diseño de experimentos, en el cual la variable respuesta puede depender de la influencia de un único factor, de forma que el resto de las causas de variación se engloban en el error experimental.

Se supone que el experimento ha sido aleatorizado por completo, es decir, todas las unidades experimentales han sido asignadas al azar a los tratamientos.

Vamos a considerar dos tipos de modelos: el de efectos fijos y el de efectos aleatorios. Se presentan ambos tipos mediante dos ejemplos:

Ejemplo 1. Una firma comercial desea conocer la influencia que tiene el nivel cultural de las familias en el éxito de una campaña publicitaria sobre cierto producto. Para ello, aprovecha los resultados de una encuesta anterior clasificando las respuestas en tantos grupos como niveles culturales ha establecido.

Estamos ante un modelo de un solo factor, ya que la firma sólo está interesada en averiguar si los distintos niveles culturales influyen o no de la misma manera sobre las ventas, no importándole la influencia del resto de los factores que pueden inducir a una mayor o menor tendencia a la compra. El modelo es de diseño fijo porque la firma aplicará los resultados de la investigación exclusivamente a los niveles culturales establecidos por ella, que pueden abarcar o no la gama completa de formación cultural.

Ejemplo 2. En una fábrica se han observado anomalías en la calidad de las piezas pro-

ducidas por un tipo de máquinas: por haber sido revisadas recientemente se piensa que los defectos puedan deberse a los trabajadores. Para contrastar esta hipótesis se toma una muestra aleatoria de trabajadores y se controla la calidad de las distintas piezas que cada uno obtiene.

Al igual que en el ejemplo anterior el modelo de comportamiento es de un solo factor, la calidad del trabajo de los trabajadores, pero al extender el resultado del análisis a toda la población de la que procede la muestra de obreros, el modelo es aleatorio, ya que de él deduciremos si los obreros que integran la población estudiada realizan un trabajo de la misma calidad o no.

En el Ejemplo 1, la firma tenía una gama de formaciones culturales muy amplia, pero sólo le interesaban unas determinadas. Para ella, la población de niveles estaba compuesta por los elegidos en el estudio, por lo cual los resultados sólo se pueden aplicar a ellos. En este caso, los niveles del factor se han elegido de forma determinista, basándose en datos históricos.

Por el contrario, en el Ejemplo 2, no interesa la calidad del trabajo de los trabajadores, sino poder atribuir la aparición de piezas defectuosas a todos los trabajadores o a las máquinas. Si del análisis se deduce que la muestra de trabajadores no presenta diferencias de calidades, se inferirá que en la población tampoco, por lo cual se pueden atribuir los fallos a las máquinas. En este caso, los niveles del factor se han elegido de forma aleatoria, pudiéndose inferir los resultados a toda la población de trabajadores.

Así, se pueden considerar dos posibles variantes de diseño unifactorial:

- (i) Los niveles del factor se seleccionan de modo específico por el experimentador. Esto constituye el llamado modelo de *efectos fijos*.
- (ii) Los niveles de un factor son una muestra aleatoria de una *población* mayor de tratamientos. Esto es el modelo de *efectos aleatorios*.

Modelo de efectos fijos

Sea Y la variable respuesta que deseamos analizar. Podemos resolver dos tipos de problemas:

1. Consideramos a poblaciones diferentes y comparamos la respuesta a un tratamiento, o único nivel de un factor. En la población i -ésima ($i = 1, \dots, a$) se toman n_i observaciones. La respuesta se cuantifica mediante y_{ij} , donde $i = 1, \dots, a$ se refiere a la población en estudio y $j = 1, \dots, n_i$ se refiere a la observación j -ésima.
2. Consideramos ahora un factor con a niveles, es decir, en total a tratamientos, y una única población. Se observa la respuesta y_{ij} del tratamiento i -ésimo a n_i observaciones de la población.

En cualquiera de los dos casos el modelo se puede expresar como:

$$y_{ij} = \mu_i + \varepsilon_{ij}$$

donde $i = 1, \dots, a$; $j = 1, \dots, n_i$ y $\sum_{i=1}^a n_i = N$, siendo μ_i el valor medio de Y , la variable respuesta, en la población o nivel i -ésimo, y ε_{ij} es el error aleatorio que incluye a todos los factores que influyen en la respuesta y no están incluidos en el modelo.

Alternativamente, se puede expresar de esta manera:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

donde $i = 1, \dots, a$; $j = 1, \dots, n$, suponiendo grupos de igual tamaño.

De este modo,

(i) y_{ij} es la observación (i, j) -ésima.

(ii) μ es la media global.

(iii) τ_i es el efecto del i -ésimo tratamiento.

(iv) ε_{ij} es el error aleatorio, tal que $\varepsilon_{ij} \sim N(0, \sigma^2)$ independientes entre sí, $E[\varepsilon_{ij}] = 0$ y $Var[\varepsilon_{ij}] = \sigma^2$.

Se supone, además, que las unidades experimentales están en un ambiente uniforme, lo cual lleva a un diseño completamente aleatorizado.

En el modelo de efectos fijos, los efectos de los tratamientos τ_i se definen como desviaciones respecto a la media general, por lo que

$$\begin{aligned} \sum_{j=1}^n \sum_{i=1}^a \tau_i &= 0 \\ \sum_{i=1}^a n\tau_i &= 0 \implies \\ \sum_{i=1}^a \tau_i &= 0 \end{aligned}$$

La esperanza del tratamiento i es

$$E[y_{ij}] = \mu + \tau_i$$

donde $i = 1, \dots, a$. De este modo es igual al término de la media general más el efecto del tratamiento i .

El problema que se trata de analizar es

$$H_0 \equiv \mu_1 = \mu_2 = \dots = \mu_a$$

$$H_1 \equiv \mu_i \neq \mu_j \text{ (para al menos un par)}$$

y esto es equivalente a

$$H_0 \equiv \tau_1 = \tau_2 = \dots = \tau_a$$

$$H_1 \equiv \tau_i \neq 0, \quad \exists i$$

El problema se puede resumir en la siguiente tabla:

Nivel	Observaciones	Totales	Promedios
1	y_{11} y_{12} \cdots y_{1n}	$y_{1\cdot}$	$\bar{y}_{1\cdot}$
2	y_{21} y_{22} \cdots y_{2n}	$y_{2\cdot}$	$\bar{y}_{2\cdot}$
\vdots	\cdots	\cdots	\cdots
a	y_{a1} y_{a2} \cdots y_{an}	$y_{a\cdot}$	$\bar{y}_{a\cdot}$
		$y_{\cdot\cdot}$	$\bar{y}_{\cdot\cdot}$

La idea es descubrir cómo se reparte la variabilidad total de la muestra. Una posible medida de variabilidad total es la suma de cuadrados, denominada *total*, o suma total de cuadrados corregida:

$$SCT = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{\cdot\cdot})^2$$

donde

$$\bar{y}_{\cdot\cdot} = \frac{1}{n \cdot a} \sum_{i=1}^a \sum_{j=1}^n y_{ij}$$

Se puede descomponer en dos partes esta suma total de cuadrados:

$$\begin{aligned} SCT &= \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{\cdot\cdot})^2 = \sum_{i=1}^a \sum_{j=1}^n ((\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}) + (y_{ij} - \bar{y}_{i\cdot}))^2 = \\ &= n \sum_{i=1}^a (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 + \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i\cdot})^2 = \\ &= SCTra + SCE. \end{aligned}$$

ya que

$$\begin{aligned} &2 \sum_{i=1}^a \sum_{j=1}^n (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}) (y_{ij} - \bar{y}_{i\cdot}) = \\ &= 2 \sum_{i=1}^a (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}) \sum_{j=1}^n (y_{ij} - \bar{y}_{i\cdot}) \end{aligned}$$

pero

$$\sum_{j=1}^n (y_{ij} - \bar{y}_{i\cdot}) = n\bar{y}_{i\cdot} - n\bar{y}_{i\cdot} = 0$$

y así los dobles productos se hacen 0.

Las diferencias entre los promedios observados de los tratamientos y el promedio general, da una medida de las diferencias entre los tratamientos.

Las diferencias de las observaciones dentro de los tratamientos con respecto al promedio del tratamiento, se considera error aleatorio.

Grados de libertad.

Se tiene un total de $a \cdot n$ observaciones y de a tratamientos.

SCT tiene $(an - 1)$ grados de libertad.

$SCTra$ tiene $(a - 1)$ grados de libertad.

SCE tiene $a(n - 1)$ grados de libertad, porque hay n réplicas dentro de cada tratamiento, es decir, se tienen $(n - 1)$ grados de libertad para estimar el error experimental. Al tener a tratamientos, se tiene un total de $a(n - 1)$ grados de libertad.

Observaciones.

Se tiene que

$$SCE = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 = \sum_{i=1}^a \left[\sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 \right].$$

Si el término entre paréntesis se divide entre $n - 1$, se obtiene la cuasivarianza del tratamiento i :

$$s_i^2 = \frac{1}{n - 1} \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2.$$

Se puede estimar la varianza poblacional combinando dichas varianzas por grupos:

$$\begin{aligned} \frac{(n - 1)s_1^2 + (n - 1)s_2^2 + \cdots + (n - 1)s_a^2}{(n - 1) + (n - 1) + \cdots + (n - 1)} &= \frac{\sum_{i=1}^a \left[\sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 \right]}{\sum_{i=1}^a (n - 1)} = \\ &= \frac{SCE}{N - a} \end{aligned}$$

donde $N = a \cdot n$.

Si no hay diferencias entre los a tratamientos, se puede estimar la varianza poblacional σ^2 como

$$\frac{SCTra}{a - 1} = \frac{n \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2}{a - 1}$$

cuando las medias de los tratamientos son iguales, ya que el término

$$\frac{\sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2}{a - 1}$$

sería un estimador de la varianza de la media muestral: σ^2/n .

Se dispone, así de dos posibles estimadores de σ^2 :

$$\begin{aligned} MCTra &= \frac{SCTra}{a - 1} \\ MCE &= \frac{SCE}{N - a} \end{aligned}$$

Cuando no existen diferencias entre las medias de los tratamientos, las estimaciones deben ser similares.

Si consideramos las medias de cuadrados anteriores, entonces, se puede demostrar, sustituyendo, que

$$\begin{aligned} E(MCE) &= \sigma^2 \\ E(MCTra) &= \sigma^2 + \frac{n \sum_{i=1}^a \tau_i^2}{a - 1}. \end{aligned}$$

De este modo, si para algún $\tau_i \neq 0$, entonces $E(MCTra) > \sigma^2$.

La idea básica es diseñar un contraste que tenga en cuenta estas diferencias entre los dos estimadores de σ^2 .

Como los errores ε_{ij} se distribuyen independientemente entre sí, según una $N(0, \sigma)$, entonces, por el lema de Fisher

$$\begin{aligned} \frac{SCE}{\sigma^2} &\sim \chi_{N-a}^2 \\ \frac{SCTra}{\sigma^2} &\sim \chi_{a-1}^2 \end{aligned}$$

siempre que $\tau_i = 0, \forall i$.

NOTA: *Teorema de Cochran:*

Sea $z_i \sim N(0, 1)$ independientes entre sí, para $i = 1, 2, \dots, v$ y sea

$$\sum_{i=1}^v z_i^2 = Q_1 + Q_2 + \dots + Q_s$$

donde $s \leq v$ y cada Q_i tiene v_i grados de libertad ($i = 1, 2, \dots, s$), entonces Q_1, Q_2, \dots, Q_s son v.a. independientes distribuidas como una chi cuadrado con v_1, v_2, \dots, v_s grados de libertad respectivamente, si y sólo si

$$v = v_1 + v_2 + \dots + v_s$$

Si se aplica el teorema de Cochran, se tiene que $\frac{SSE}{\sigma^2}$ y $\frac{SSTra}{\sigma^2}$ son independientes, por lo que si $\tau_i = 0, \forall i$, entonces

$$F_0 = \frac{\frac{SCTra}{a-1}}{\frac{SCE}{N-a}} = \frac{MCTra}{MCE}$$

se distribuye como una F de Snedecor, $F_{a-1, N-a}$.

Si algún $\tau_i \neq 0$, entonces $E(MSTra) > \sigma^2$ entonces el valor del estadístico F_0 es mayor, obteniéndose una región crítica superior, de modo que se rechaza, a nivel α , la hipótesis nula de igualdad de tratamientos, si

$$F_0 > F_{\alpha, a-1, N-a}$$

Resumen: Tabla ANOVA.

$$H_0 \equiv \tau_1 = \tau_2 = \dots = \tau_a$$

$$H_1 \equiv \tau_i \neq 0, \quad \exists i$$

F. Variación	S. Cuadrados	gl	M. Cuadrados	F₀
Factor	$SCTra = n \sum_{i=1}^a (\bar{y}_i - \bar{y}_{..})^2$	$a - 1$	$MCTra = \frac{SCTra}{a-1}$	$F_o = \frac{MCTra}{MCE}$
Error	$SCE = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2$	$N - a$	$MCE = \frac{SCE}{n-a}$	
Total	$SCT = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2$	$N - 1$		

Se rechaza H_0 a nivel α cuando $F_0 > F_{\alpha, a-1, N-a}$.

Estimación de los parámetros.

Dado el modelo

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

donde $i = 1, \dots, a$; $j = 1, \dots, n$, se pueden estimar los parámetros μ y τ_i por el método de los mínimos cuadrados, de modo que no se necesita suponer normalidad de los errores ε_{ij} .

la suma de los cuadrados de los errores es

$$L = \sum_{i=1}^a \sum_{j=1}^n \varepsilon_{ij}^2 = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \mu - \tau_i)^2,$$

de modo que los estimadores de μ y τ_i son los valores $\hat{\mu}$ y $\hat{\tau}_i$ que minimizan el funcional L .

Derivando respecto cada uno de los parámetros, se obtiene un total de $(a + 1)$ ecuaciones:

$$\begin{aligned} \frac{\partial L}{\partial \mu} &= 0 \implies -2 \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \hat{\mu} - \hat{\tau}_i) = 0 \\ \frac{\partial L}{\partial \tau_i} &= 0 \implies -2 \sum_{j=1}^n (y_{ij} - \hat{\mu} - \hat{\tau}_i) = 0, \quad i = 1, 2, \dots, a \end{aligned}$$

se obtiene

$$\begin{cases} N\hat{\mu} + n\hat{\tau}_1 + n\hat{\tau}_2 + \dots + n\hat{\tau}_a = y_{..} \\ \left\{ \begin{array}{cccccc} n\hat{\mu} & +n\hat{\tau}_1 & & & & = y_{1.} \\ n\hat{\mu} & & +n\hat{\tau}_2 & & & = y_{2.} \\ \vdots & \dots & \dots & \dots & \dots & \vdots \\ n\hat{\mu} & & & & +n\hat{\tau}_a & = y_{a.} \end{array} \right. \end{cases}$$

Estas se denominan *ecuaciones normales* de mínimos cuadrados. Si se suman las últimas a ecuaciones, se obtiene la primera ecuación, de modo que no forman un sistema independiente de ecuaciones y no existe solución única. Para evitar esto, se considera la restricción

$$\sum_{i=1}^a \hat{\tau}_i = 0,$$

obteniéndose, entonces, los estimadores

$$\hat{\mu} = \bar{y}_{..}$$

$$\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..}$$

para $i = 1, 2, \dots, a$.

Si se asume que los errores están distribuidos según una normal, entonces cada $\bar{y}_{i.} \sim N(\mu_i, \sigma^2/n)$. De este modo, cuando σ^2 es desconocida un intervalo de confianza al $100(1 - \alpha)\%$ es

$$\left[\bar{y}_{i.} \pm t_{\frac{\alpha}{2}, N-a} \sqrt{\frac{MCE}{n}} \right].$$

De la misma manera,

$$\left[(\bar{y}_{i.} - \bar{y}_{..}) \pm t_{\frac{\alpha}{2}, N-a} \sqrt{\frac{2MCE}{n}} \right].$$

Diseño desequilibrado.

Si el número de observaciones es diferente según cada tratamiento i : n_i donde $i = 1, 2, \dots, a$, las expresiones previas son iguales salvo que se sustituye n por n_i :

$$\begin{aligned} SCT &= \sum_{i=1}^a \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^a \sum_{j=1}^{n_i} y_{ij}^2 - \frac{y_{..}^2}{N} \\ SCTra &= \sum_{i=1}^a \sum_{j=1}^{n_i} (\bar{y}_{i.} - \bar{y}_{..})^2 = \sum_{i=1}^a \frac{y_{i.}^2}{n_i} - \frac{y_{..}^2}{N} \\ SCE &= SCT - SCTra \end{aligned}$$

Para resolver las ecuaciones normales se considera la restricción

$$\sum_{i=1}^a n_i \hat{\tau}_i = 0$$

y se resuelve del mismo modo.

Si el diseño es no balanceado o desequilibrado, aumenta la sensibilidad del análisis unifactorial a la falta de igualdad entre las varianzas de cada grupo (heterocedasticidad).

Ejemplo 1 Un ingeniero de desarrollo de productos está interesado en maximizar la resistencia a la tensión de una nueva fibra sintética que se empleará en la manufactura de tela para camisas de hombre. El ingeniero sabe por experiencia que la resistencia está influida por el porcentaje de algodón presente en la fibra. Además, sospecha que el contenido de algodón debe estar aproximadamente entre un 10 y 40 % para que la tela resultante tenga otras características de calidad que se desean (como la capacidad de recibir un tratamiento de planchado permanente). El ingeniero decide probar muestras a cinco niveles de porcentaje de algodón: 15, 20, 25, 30 y 35 %. Asimismo, decide ensayar cinco muestras a cada nivel de contenido de algodón. Las 25 observaciones deben asignarse al azar. Para ilustrar la forma en que puede aleatorizarse el orden de ejecución, supóngase que las observaciones se numeran como sigue:

% algodón					
15	1	2	3	4	5
20	6	7	8	9	10
25	11	12	13	14	15
30	16	17	18	19	20
35	21	22	23	24	25

Ahora se elige al azar un número entre 1 y 25. Supongamos que es el 8, entonces la observación 8^a se ejecuta primero (es decir, a un 20% de algodón). A continuación se elige un número al azar entre 1 y 25, quitando el 8. Supongamos que es el 4, entonces la observación 4^a se ejecuta en segundo lugar (a un 15 % de algodón). Se repite el proceso hasta completar las 25 observaciones.

Esta secuencia de prueba aleatorizada es necesaria para evitar que los resultados se contaminen por los efectos de variables desconocidas que pueden salir de control durante el experimento. Para ilustrar esto, supóngase que se ejecutan las 25 muestras de prueba en el orden no aleatorizado original (esto es, las 5 muestras con un 15 % de algodón se prueban primero, luego las 5 muestras con un 20 % de algodón, y así sucesivamente). Si la máquina que prueba la resistencia a la tensión presenta un efecto de calentamiento tal que

a mayor tiempo de funcionamiento menores lecturas de resistencia a la tensión, entonces dicho efecto contaminará los datos de resistencia e invalidará el experimento.

Supóngase ahora que el ingeniero ejecuta la prueba en el orden aleatorio que hemos determinado. Las observaciones obtenidas acerca de la resistencia a la tensión son:

% de algodón	Observaciones	Suma	Media
15	7 7 15 11 9	49	9.8
20	12 17 12 18 18	77	15.4
25	14 18 18 19 19	88	17.6
30	19 25 22 19 23	108	21.6
35	7 10 11 15 11	54	10.8
		376	15.04

Representamos el diagrama de dispersión para la resistencia frente al porcentaje de algodón, y el diagrama de cajas para la resistencia a la tensión a cada nivel de porcentaje de algodón.

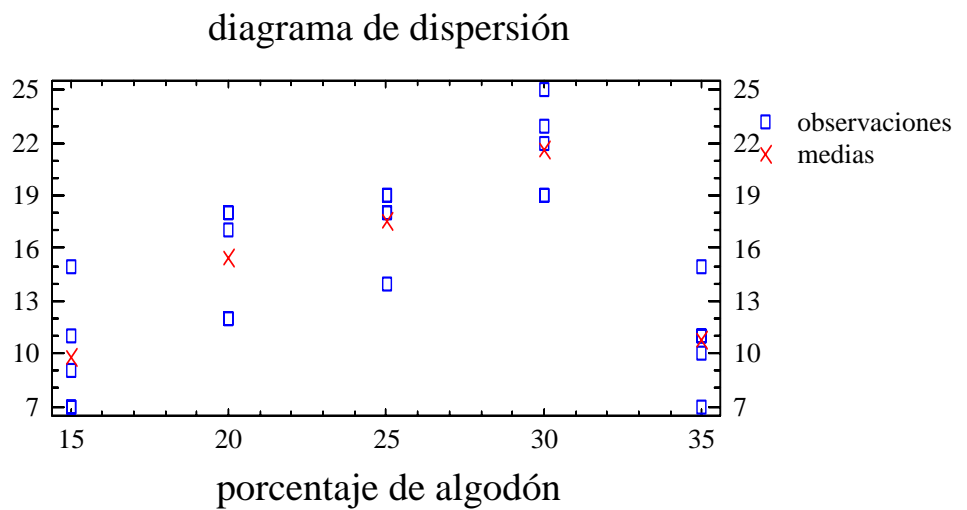
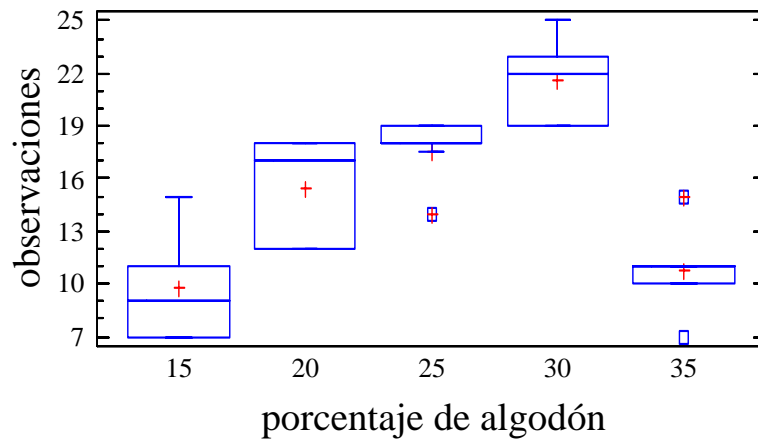


Diagrama de cajas



Ambas gráficas indican que la resistencia a la tensión aumenta con el contenido de algodón hasta el 30%. Mas allá del 30% ocurre un notable decrecimiento en la resistencia.

No hay una fuerte evidencia que sugiera que la variabilidad en la resistencia alrededor de la media dependa del porcentaje de algodón. Se sospecha, no obstante, que el porcentaje de algodón influye en la resistencia a la tensión.

Se disponen los datos en una tabla como esta:

Observaciones	Sumas	Medias
y_{11}, \dots, y_{1n_1}	$y_{1\cdot}$	$\bar{y}_{1\cdot}$
\vdots	\vdots	\vdots
y_{I1}, \dots, y_{In_I}	$y_{I\cdot}$	$\bar{y}_{I\cdot}$
	$y_{\cdot\cdot}$	$\bar{y}_{\cdot\cdot}$

A) Hipótesis del modelo

Las principales hipótesis del modelo son:

- Normalidad: ε_{ij} sigue una distribución normal.
- Linealidad: $E(\varepsilon_{ij}) = 0$
- Homocedasticidad: $Var(\varepsilon_{ij}) = \sigma^2$

- Independencia: ε_{ij} son independientes entre sí.

Estas hipótesis son equivalentes a las siguientes:

- Normalidad: y_{ij} sigue una distribución normal.
- $E(y_{ij}) = \mu_i$
- Homocedasticidad: $Var(y_{ij}) = \sigma^2$
- Independencia: y_{ij} son independientes entre sí.

B) Metodología

En nuestro análisis vamos a seguir los siguientes pasos:

- Estimar los parámetros del modelo.
- Contrastar si el factor influye en la respuesta, es decir, si los valores medios de Y son diferentes al cambiar el nivel del factor.
- Si el factor influye en la variable respuesta, es decir, las medias no son iguales, buscar las diferencias entre poblaciones (o niveles del factor).
- Diagnóstico del modelo: comprobar si las hipótesis del modelo son ciertas mediante el análisis de los residuos.

C) Estimación de los parámetros

En este ejemplo, $a = 5$, $n_i = 5$ y $N = 25$. Las estimaciones puntuales de los parámetros son las siguientes:

$$\hat{\mu}_1 = \bar{y}_{1.} = 9,8$$

$$\hat{\mu}_2 = \bar{y}_{2.} = 15,4$$

$$\hat{\mu}_3 = \bar{y}_{3.} = 17,6$$

$$\hat{\mu}_4 = \bar{y}_{4.} = 21,6$$

$$\hat{\mu}_5 = \bar{y}_{5.} = 10,8$$

Por ejemplo, el intervalo de confianza para μ_1 , al nivel $(1 - \alpha) = 0,95$, es:

$$\begin{aligned} & \left[\bar{y}_{i.} \pm t_{\frac{\alpha}{2}, N-a} \sqrt{\frac{MCE}{n}} \right] = \\ & = \left[9,8 \pm t_{0,025,20} \sqrt{\frac{8,06}{5}} \right] = \\ & \quad [7,1515, 12,4485] \end{aligned}$$

D) Análisis de la varianza

El contraste de hipótesis que vamos a abordar es el siguiente:

$$\begin{cases} H_0 : \mu_1 = \dots = \mu_a \text{ (el factor no influye)} \\ H_1 : \text{algún factor es diferente (el factor influye)} \\ \text{nivel de significación } \alpha \end{cases}$$

FV	SC	GL	F
Tratamiento	$SCTra = \sum_{i=1}^a n_i (\bar{y}_{i.} - \bar{y}_{..})^2$	$a - 1$	$F_0 = \frac{SCTra/(a-1)}{SCE/(N-a)}$
Error	$SCE = \sum_{i=1}^a \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2$	$N - a$	
Total	$SCT = \sum_{i=1}^a \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2$	$N - 1$	

siendo $FV = Fuente de variación$, $SC = Suma de Cuadrados$, $GL = Grados de libertad$.

Las sumas de cuadrados también se pueden calcular de la siguiente forma:

$$\begin{aligned} SCT &= \sum \sum y_{ij}^2 - n\bar{y}_{..}^2 \\ SCTra &= \sum n_i \bar{y}_{i.}^2 - n\bar{y}_{..}^2 \\ SCE &= SCT - SCTra \end{aligned}$$

Cuando sólo hay dos poblaciones (un factor con dos niveles), este contraste es idéntico al contraste de la t para comparar las medias de dos poblaciones normales e independientes con la misma varianza.

Analizamos a continuación la tabla de análisis de la varianza del ejemplo 1

$$\begin{aligned} SCT &= \sum \sum y_{ij}^2 - n\bar{y}_{..}^2 = \\ &= 7^2 + 7^2 + 15^2 + \dots + 15^2 + 11^2 - 25 \times 15,04^2 = \\ &= 636,96 \end{aligned}$$

$$\begin{aligned} SCTra &= \sum n_i \bar{y}_i^2 - n\bar{y}_{..}^2 = \\ &= 5(9,8^2 + \dots + 10,8^2) - 25 \times 15,04^2 = \\ &= 475,76 \end{aligned}$$

$$SCE = SCT - SCTra = 636,96 - 475,76 = 161,2$$

La tabla ANOVA es:

F.V.	S.C.	G.L.	M.C.	F
Tratamiento	475.76	4	118.94	14.76
Error	161.2	20	8.06	
Total	636.96	24		

$$F_{4,20;0,1} = 2,2489$$

$$F_{4,20;0,05} = 2,8661$$

$$F_{4,20;0,01} = 4,4307$$

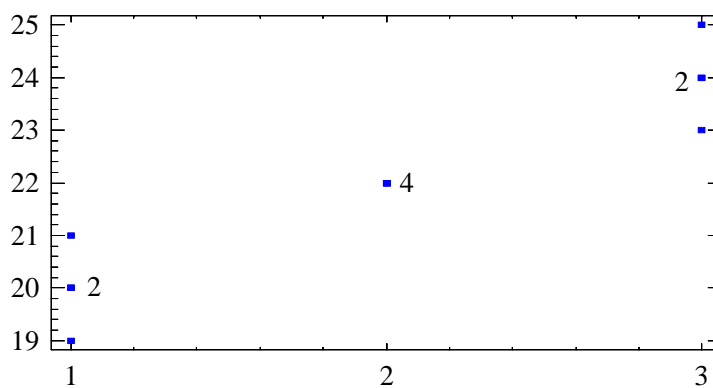
Por lo tanto, rechazamos H_0 a los niveles anteriores y concluimos que hay diferencias entre los tratamientos.

Ejemplo 2. Analizaremos los siguientes conjuntos de datos:

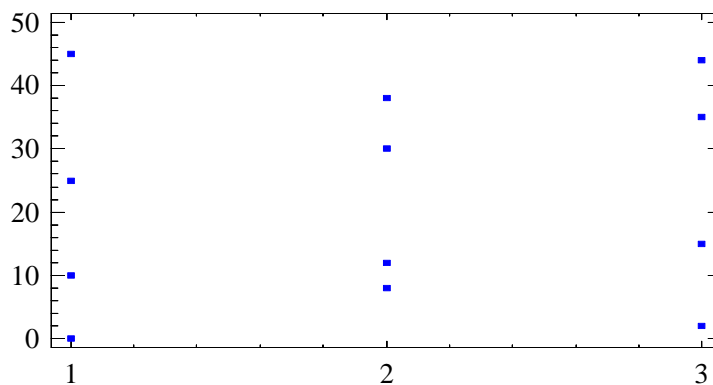
Primer caso						Segundo caso					
				Sumas	Medias					Sumas	Medias
20	19	20	21	80	20	45	0	10	25	80	20
22	22	22	22	88	22	8	30	38	12	88	22
24	24	23	25	96	24	15	44	2	35	96	24
				264	22					264	22

Las medias son iguales en los dos casos, con lo cual la diferencia de medias debería ser igual en ambos casos. Los diagramas de puntos, considerando en abscisas los grupos y en ordenadas las observaciones, son:

Primer caso



Segundo caso



Debido a las diferentes dispersiones (varianzas) que existen en los dos casos, la impresión visual es muy distinta. En el segundo caso no se aprecia diferencia entre los tres grupos (el factor no parece influir), mientras que en el primer caso, la cosa no está tan clara. Entonces, no es suficiente sólo con comparar las medias de cada grupo, la variabilidad también influye. Lo que vamos a hacer es comparar la variabilidad entre las medias con la variabilidad dentro de cada grupo, mediante el análisis de la varianza.

Vamos a construir la tabla ANOVA:

Caso 1:

$$\begin{aligned}
 SCT &= \sum_{i=1}^a \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2 = 36. \\
 SCTra &= \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2 = 32 \\
 SCE &= SCT - SCTra = 36 - 32 = 4
 \end{aligned}$$

La tabla ANOVA es:

F.V.	S.C.	G.L.	M.C.	F
Tratamiento	32	2	16	36
Error	4	9	0.444	
Total	36	11		

Como $F_{2,9;0,05} = 4,2565$, rechazamos la hipótesis nula y concluimos que el factor influye en la respuesta.

Caso 2:

$$\begin{aligned}
 SCT &= \sum \sum y_{ij}^2 - n\bar{y}_{..}^2 = 8692 - 12 \times 22^2 = 2884. \\
 SCTra &= \sum n_i \bar{y}_{i.}^2 - n\bar{y}_{..}^2 = 5840 - 12 \times 22^2 = 32 \\
 SCE &= SCT - SCTra = 2884 - 32 = 2852.
 \end{aligned}$$

La tabla ANOVA es:

F.V.	S.C.	G.L.	M.C.	F
Tratamiento	32	2	16	0.05
Error	2852	9	316.889	
Total	2884	11		

Como $F_{2,9;0,05} = 4,2565$, no rechazamos la hipótesis nula y concluimos que el factor no influye en la respuesta al nivel $\alpha = 0,05$.

Comparaciones entre medias

Una vez obtenidas diferencias significativas entre los tratamientos, conviene estudiar por qué se rechaza la igualdad entre medias, comparando todos los pares de medias, porque puede ser que se rechace la igualdad de medias porque haya un par de medias diferentes entre sí. Se considera, entonces, los siguientes contrastes:

$$H_0 \equiv \mu_i = \mu_j, \quad i \neq j$$

$$H_0 \equiv \mu_i \neq \mu_j, \quad i \neq j$$

Los métodos generales son las comparaciones múltiples y los tests de recorrido studentizado.

Comparaciones múltiples.

LSD de Fisher (Least significant difference)

Se contrasta $\mu_i = \mu_j$, para todo $i \neq j$, ($i, j = 1, \dots, a$).

Se tiene que se distribuye como una t de Student:

$$\frac{(\bar{y}_{i\cdot} - \bar{y}_{j\cdot}) - (\mu_i - \mu_j)}{\hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}} \sim t_{N-a}$$

Así, un Intervalo de Confianza para $(\mu_i - \mu_j)$ a nivel α es

$$[(\bar{y}_{i\cdot} - \bar{y}_{j\cdot}) \pm LSD_\alpha]$$

y se denomina

$$LSD_\alpha = t_{N-a, \frac{\alpha}{2}} \hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$$

1. Si $|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > LSD_\alpha \implies$ Se rechaza que $\mu_i = \mu_j$ a nivel α .
2. Si $|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| < LSD_\alpha \implies$ Se acepta que $\mu_i = \mu_j$ a nivel α .

Definición. (*Distribución de recorrido estudentizada*)

Si

$$Z_1, \dots, Z_a \sim N(0, 1)$$

$$U \sim \chi_m^2$$

independientemente, entonces,

$$Q = \max_{i \neq j} \frac{|Z_i - Z_j|}{\sqrt{\frac{U}{m}}} = \frac{Z_{(a)} - Z_{(1)}}{\sqrt{\frac{U}{m}}} \sim q_{a,m}$$

se distribuye con una distribución de recorrido estudentizado de parámetros a y m .

Método de Tukey

Se requiere que $n_i = n$, $i = 1, \dots, a$. Si esto no se cumple, entonces se toma $n = \min_i \{n_i\}$.

1. Si $|\bar{y}_i - \bar{y}_j| > q_{a, N-a; \alpha} \hat{\sigma} \sqrt{\frac{1}{n}} \implies$ Se rechaza que $\mu_i = \mu_j$ a nivel α .
2. Si $|\bar{y}_i - \bar{y}_j| < q_{a, N-a; \alpha} \hat{\sigma} \sqrt{\frac{1}{n}} \implies$ Se acepta que $\mu_i = \mu_j$ a nivel α .

Método de Bonferroni

En este criterio se rechaza $\mu_i = \mu_j$ ($i \neq j$) si

$$|\bar{y}_i - \bar{y}_j| > t_{N-a, \frac{\alpha}{2p}} \hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$$

donde p es el número de comparaciones que se pueden obtener: $1 \leq p \leq \binom{a}{2}$.

Se puede aproximar $t_{N-a, \frac{\alpha}{2p}}$ por una normal:

$$t_{v, \alpha} = z_\alpha + \frac{1}{4v} (z_\alpha^3 - z_\alpha),$$

siendo $z_\alpha \sim N(0, 1)$.

Ejemplos.

En el problema de comparación del porcentaje de algodón en las prendas, las medias muestrales eran:

$\bar{y}_i.$	$\bar{y}_1.$	$\bar{y}_2.$	$\bar{y}_3.$	$\bar{y}_4.$	$\bar{y}_5.$
	9,8	15,4	17,6	21,6	10,8

Se tiene que

$$a = 5$$

$$n = 5$$

$$N = 25$$

$$\hat{\sigma}^2 = 8,06$$

$$N - a = 20.$$

– *LSD de Fisher*

$$LSD_\alpha = t_{N-a, \frac{\alpha}{2}} \hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}} = t_{20, 0,025} \sqrt{8,06 \cdot \frac{2}{5}} = 3,745$$

– *Método de Tuckey*

$$HSD_\alpha = \frac{\hat{\sigma}}{\sqrt{n}} q_{a, N-a, \alpha} = \sqrt{\frac{8,06}{5}} q_{5, 20, 0,05} = 1,269 \cdot 4,24 = 5,38$$

– *Método de Bonferroni*

Como

$$p = \binom{5}{2} = 10$$

luego hay 10 posibles comparaciones:

$$B_\alpha = t_{N-a, \frac{\alpha}{2p}} \hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}} = t_{20, \frac{0,05}{20}} \sqrt{8,06 \cdot \frac{2}{5}}$$

Como

$$\begin{aligned} t_{20, \frac{0,05}{20}} &= t_{20, 0,0025} \approx z_{0,0025} + \frac{1}{4 \cdot 20} (z_{0,0025}^3 - z_{0,0025}) = \\ &= 2,81 + \frac{1}{80} (2,81^3 - 2,81) = 3,052 \end{aligned}$$

luego

$$B_\alpha = t_{20, \frac{0,05}{20}} \sqrt{8,06 \cdot \frac{2}{5}} = 3,052 \cdot \sqrt{8,06 \cdot \frac{2}{5}} = 5,48.$$

Así la tabla de diferencias es:

(i, j)	$(\bar{y}_i. - \bar{y}_j.)$	$LSD_\alpha = 3,745$	$HSD_\alpha = 5,38$	$B_\alpha = 5,48$
(1,2)	5,6	\neq	\neq	\neq
(1,3)	7,8	\neq	\neq	\neq
(1,4)	11,8	\neq	\neq	\neq
(1,5)	1,0	$=$	$=$	$=$
(2,3)	2,2	$=$	$=$	$=$
(2,4)	6,2	\neq	\neq	\neq
(2,5)	4,6	\neq	$=$	$=$
(3,4)	4	\neq	$=$	$=$
(3,5)	6,8	\neq	\neq	\neq
(4,5)	10,8	\neq	\neq	\neq

Tests de recorrido studentizado

En estos tests, se requiere que $n_i = n$, $i = 1, \dots, a$. Si esto no se cumple, entonces se toma la media armónica:

$$n = a \left(\frac{1}{n_1} + \dots + \frac{1}{n_a} \right)^{-1}$$

Los tests principales son:

- El test de Duncan
- El test de Newman-Keuls

En ambos tests se siguen los siguientes pasos:

(i) Se ordenan de manera creciente las medias muestrales a comparar:

$$\bar{y}_{(1)} < \bar{y}_{(2)} < \dots < \bar{y}_{(a)}.$$

(ii) Se comparan las diferencias entre dos medias separadas por p posiciones con $p = a, a - 1, \dots, 2$ usando los siguientes puntos críticos:

■ **Duncan.**

$$d_p = \frac{\hat{\sigma}}{\sqrt{n}} r_{p, N-a, \alpha}$$

donde $r_{p, N-a, \alpha}$ se obtiene a partir de la tabla de intervalos significativos de Duncan.

■ **Newman-Keuls.**

$$NK_p = \frac{\hat{\sigma}}{\sqrt{n}} q_{p,N-a,\alpha}$$

donde $q_{p,N-a,\alpha}$ se obtiene a partir de la tabla de la distribución de recorrido studentizado.

Por ejemplo, para $p = a$ se contrasta si

$$\begin{aligned} |\bar{y}_{(a)\cdot} - \bar{y}_{(1)\cdot}| &> r_{a,N-a;\alpha} \hat{\sigma} \sqrt{\frac{1}{n}} \\ |\bar{y}_{(a)\cdot} - \bar{y}_{(1)\cdot}| &> q_{a,N-a;\alpha} \hat{\sigma} \sqrt{\frac{1}{n}} \end{aligned}$$

Para $p = a - 1$ se contrasta si

$$\begin{aligned} |\bar{y}_{(a)\cdot} - \bar{y}_{(2)\cdot}| &> r_{a-1,N-a;\alpha} \hat{\sigma} \sqrt{\frac{1}{n}} \\ |\bar{y}_{(a-1)\cdot} - \bar{y}_{(1)\cdot}| &> r_{a-1,N-a;\alpha} \hat{\sigma} \sqrt{\frac{1}{n}} \end{aligned}$$

$$\begin{aligned} |\bar{y}_{(a)\cdot} - \bar{y}_{(2)\cdot}| &> q_{a-1,N-a;\alpha} \hat{\sigma} \sqrt{\frac{1}{n}} \\ |\bar{y}_{(a-1)\cdot} - \bar{y}_{(1)\cdot}| &> q_{a-1,N-a;\alpha} \hat{\sigma} \sqrt{\frac{1}{n}} \end{aligned}$$

(iii) Se van *declarando* diferentes o no a las parejas de medias. Si no se declaran diferentes, se conectan con una línea base. Al final sólo se declaran diferentes las medias que no están conectadas por ninguna línea.

(iv) Si un grupo de medias no es significativamente diferente, ningún subgrupo de ellas lo es.

Se tiene que la relación entre ambas tablas es la siguiente:

$$r_{p,N-a;\alpha} = q_{p,N-a;1-(1-\alpha)^{p-1}}$$

Si comparamos los respectivos puntos críticos con $N - a = 20$, por ejemplo:

p	2	3	4	5	6	7	8
$r_{p,20,0'01}$	4.02	4.22	4.33	4.4	4.47	4.53	4.58
$q_{p,20,0'01}$	4.02	4.64	5.02	5.29	5.51	5.69	5.84

Se observa que $r_{p,20,0'01} \leq q_{p,20,0'01}$, con lo cual se tiene que

$$d_{\alpha,p} = \frac{\hat{\sigma}}{\sqrt{n}} r_{p,N-a,\alpha} < NK_{\alpha,p} = \frac{\hat{\sigma}}{\sqrt{n}} q_{p,N-a,\alpha},$$

es decir, el test de Newman-Keuls es más conservador que el de Duncan, de modo que si se rechaza la H_0 aplicando el test de Newman-Keuls, también se rechaza aplicando el test de Duncan.

Ejemplo.

En el problema de comparación del porcentaje de algodón en las prendas, se ordenan las medias muestrales de menor a mayor:

\bar{y}_i	\bar{y}_1	\bar{y}_5	\bar{y}_2	\bar{y}_3	\bar{y}_4
	9,8	10,8	15,4	17,6	21,6

Test de Newman-Keuls:

$$\begin{aligned} p = 5 & \quad q_{5,20,0'05} = 4,24 & \quad NK_5 = 5,38 \\ p = 4 & \quad q_{4,20,0'05} = 3,96 & \quad NK_4 = 5,03 \\ p = 3 & \quad q_{3,20,0'05} = 3,58 & \quad NK_3 = 4,54 \\ p = 2 & \quad q_{2,20,0'05} = 2,95 & \quad NK_2 = 3,74 \end{aligned}$$

Test de Duncan:

$$\begin{aligned} p = 5 & \quad r_{5,20,0'05} = 3,25 & \quad d_5 = 4,12 \\ p = 4 & \quad r_{4,20,0'05} = 3,18 & \quad d_4 = 4,04 \\ p = 3 & \quad r_{3,20,0'05} = 3,10 & \quad d_3 = 3,93 \\ p = 2 & \quad r_{2,20,0'05} = 2,95 & \quad d_2 = 3,74 \end{aligned}$$

De este modo,

p	$ \bar{y}_i - \bar{y}_j $	<i>Newman-Keuls</i>	<i>Duncan</i>
5	$ \bar{y}_1 - \bar{y}_4 = 11,8$	$> 5,38$	$> 4,12$
4	$ \bar{y}_1 - \bar{y}_3 = 7,8$ $ \bar{y}_5 - \bar{y}_4 = 10,8$	$> 5,03$ $> 5,03$	$> 4,03$ $> 4,03$
3	$ \bar{y}_1 - \bar{y}_2 = 5,6$ $ \bar{y}_5 - \bar{y}_3 = 6,8$ $ \bar{y}_2 - \bar{y}_4 = 6,2$	$> 4,54$ $> 4,54$ $> 4,54$	$> 3,93$ $> 3,93$ $> 3,93$
2	$ \bar{y}_1 - \bar{y}_5 = 1$ $ \bar{y}_5 - \bar{y}_2 = 4,6$ $ \bar{y}_3 - \bar{y}_2 = 2,2$ $ \bar{y}_3 - \bar{y}_4 = 4$	$< \mathbf{3,74}$ $> 3,74$ $< \mathbf{3,74}$ $> 3,74$	$< \mathbf{3,74}$ $> 3,74$ $< \mathbf{3,74}$ $> 3,74$

Como conclusión, se obtiene que $\mu_4 > \mu_i$ para $i = 1, 2, 3, 5$ de manera significativa según ambos criterios.

Y el resultado es:

$$\underbrace{\mu_1 \quad \mu_5}_{\text{-----}} \quad \underbrace{\mu_2 \quad \mu_3}_{\text{-----}} \quad \mu_4$$

Contrastes Ortogonales (método de Scheffé)

En general, un contraste entre k medias poblacionales se puede definir como una combinación lineal

$$\Psi = c_1\mu_1 + c_2\mu_2 + \dots + c_a\mu_a$$

tales que c_1, \dots, c_a son constantes de suma nula: $\sum_{j=1}^a c_j = 0$.

Un estimador de Ψ es

$$\hat{\Psi} = c_1\bar{x}_1 + c_2\bar{x}_2 + \dots + c_a\bar{x}_a$$

y la estimación de la varianza de Ψ es

$$\hat{\sigma}_{\Psi}^2 = MCE \left(\frac{c_1^2}{n_1} + \dots + \frac{c_a^2}{n_a} \right).$$

En particular, por ejemplo, la diferencia entre dos medias cualesquiera equivale a un contraste

$$\hat{\Psi} = c_i\mu_i + c_j\mu_j$$

con $c_i = 1$ y $c_j = -1$ y cero para el resto de términos.

El procedimiento de Scheffé se aplica de la siguiente forma:

(i) Especificar Ψ así como los coeficientes que determinan el contraste. Calcular $\hat{\Psi}$ reemplazando las medias muestrales por las poblacionales.

(ii) Estimar $\hat{\sigma}_{\Psi}^2$ y calcular la razón $\frac{\hat{\Psi}}{\hat{\sigma}_{\Psi}}$.

(iii) Si la razón

$$\frac{|\hat{\Psi}|}{\hat{\sigma}_{\Psi}} > \sqrt{(a-1)F_{a-1, N-a, \alpha}},$$

se rechaza la hipótesis $H_0 \equiv \Psi = 0$ al nivel α .

Ejemplo.

Supongamos que se trata de contrastar

$$\Psi_1 = \mu_1 + \mu_3 - \mu_4 - \mu_5$$

$$\Psi_2 = \mu_1 - \mu_4$$

Las estimas son

$$\hat{\Psi}_1 = \bar{x}_1 + \bar{x}_3 - \bar{x}_4 - \bar{x}_5 = 9,8 + 17,6 - 21,6 - 10,8 = 5,0$$

$$\hat{\Psi}_2 = \bar{x}_1 - \bar{x}_4 = 9,8 - 21,6 = -11,8$$

$$\hat{\sigma}_{\Psi_1} = \sqrt{MCE \sum_{i=1}^5 \frac{c_i^2}{n_i}} = \sqrt{8,06 \frac{4}{5}} = 2,54$$

$$\hat{\sigma}_{\Psi_2} = \sqrt{8,06 \frac{2}{5}} = 1,8$$

y como $F_{a-1, N-a, \alpha} = F_{4, 20, 0,01} = 4,43$, de modo que $\sqrt{(a-1)F_{a-1, N-a, \alpha}} = \sqrt{4 \cdot 4,43} = 4,21$

$$\frac{|\hat{\Psi}_1|}{\hat{\sigma}_{\Psi_1}} = \frac{5,0}{2,54} = 1,97 < 4,21$$

se acepta la H_0 .

$$\frac{|\hat{\Psi}_2|}{\hat{\sigma}_{\Psi_2}} = \frac{11,8}{1,8} = 6,56 > 4,21$$

se rechaza la H_0 .

Aunque el método de Scheffé permite plantear muchas posibles comparaciones entre medias, cuando se estudian sólo diferencias entre medias resulta menos eficaz que los tests específicos para diferencias de pares de medias.

Estudio de la adecuación del modelo

La mayor parte de las desviaciones de las hipótesis básicas del modelo, se pueden estudiar a través de los residuos:

$$e_{ij} = y_{ij} - \hat{y}_{ij}$$

al igual que se hace, habitualmente en Regresión.

Así, el dibujo de los errores de modo secuencial a como aparecen las observaciones permite detectar correlaciones entre los mismos y, de este modo, se observa si se cumple la hipótesis de independencia. Si no es así, es un problema difícil de corregir en la práctica, como no sea repitiendo el experimento y aleatorizando de modo conveniente.

También se puede considerar la gráfica de los errores frente a los valores predichos $\hat{y}_{ij} = \bar{y}_i$, que no debería presentar tendencias en cuanto a su aspecto. Si lo hace, es un signo de la existencia de varianzas no constante o heterocedasticidad.

Cuando $n_i = n$, para $i = 1, \dots, a$ la existencia de varianzas heterogéneas entre los grupos, apenas afecta al contraste de la F . Sin embargo, si los tamaños muestrales son desiguales, la probabilidad de cometer error de tipo I puede ser diferente al valor α prefijado.

Para comprobar este supuesto, se puede considerar el test de Levene o el test de Barlett. Sin embargo, la prueba de Levene tiene la ventaja de que no se ve afectada por la falta de normalidad del modelo, y se puede aplicar a tamaños muestrales desiguales.

El estadístico de contraste de la prueba de Levene es

$$F_0 = \frac{\sum_{i=1}^a \sum_{j=1}^{n_i} (d_{ij} - \bar{d}_i) / (N - a)}{\sum_{i=1}^a n_i (\bar{d}_i - \bar{d}_{..}) / (a - 1)}$$

donde

$$\begin{aligned} d_{ij} &= |y_{ij} - \bar{y}_i|, \\ \bar{d}_i &= \frac{\sum_{j=1}^{n_i} d_{ij}}{n_i}, \\ \bar{d}_{..} &= \frac{\sum_{i=1}^a \sum_{j=1}^{n_i} d_{ij}}{N}. \end{aligned}$$

Si las varianzas son homogéneas, entonces este estadístico F_0 se distribuye como una F de Snedecor, $F_{a-1, N-a, \alpha}$, siendo α el nivel de significación elegido.

Transformaciones para conseguir homocedasticidad

Cuando se presenta homocedasticidad se debe a menudo a que la varianza cambia cuando lo hace la media. Si μ_i es la media del grupo i -ésimo y σ_i su desviación típica, entonces $\sigma_i = f(\mu_i)$ para alguna función f .

Para estabilizar la varianza se busca una función T tal que $T(x)$ tenga varianza constante. En la práctica se usa

$$T(x_{ij}) = \begin{cases} x_{ij}^\lambda & \lambda \neq 0 \\ \log(x_{ij}) & \lambda = 0 \end{cases}$$

En particular se suele considerar una función f de la forma

$$f(\mu_i) = k\mu_i^\alpha$$

de modo que

$$\sigma_i = k\mu_i^\alpha$$

Se puede demostrar que para conseguir la homocedasticidad, se debe usar la transformación con parámetro $\lambda = 1 - \alpha$.

Para estimar λ se usan los diagramas *rango-media*. Se asume el ajuste a una ecuación del tipo $\sigma = k\mu^\alpha$ y así

$$\log(\sigma) = \log(k) + \alpha \log(\mu)$$

α será la pendiente de la recta que pasa por los puntos del gráfico, esto es, de la correspondiente recta de regresión. Una vez estimado α se calcula $\lambda = 1 - \alpha$.

Observaciones

(i) Las transformaciones estabilizadoras de la varianza se definen sólo para conjuntos de datos positivos. En caso contrario, hay que sumar una constante a los datos.

- (*ii*) En general se considera una rejilla de valores de λ y se va probando con múltiplos de $\frac{1}{2}$.
- (*iii*) Frecuentemente la transformación no sólo estabiliza la varianza sino que normaliza los datos, cuando estos no se distribuyen como una normal.

Modelo de efectos aleatorios

Si el número de niveles del factor no está fijado de antemano, sino que es una muestra aleatoria de una población de niveles, entonces se tiene un modelo de efectos aleatorios.

El modelo se expresa igual que antes

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

donde $i = 1, 2, \dots, a$ y $j = 1, 2, \dots, n$, siendo, en este caso, τ_i y ε_{ij} variables aleatorias.

Si se asume que τ_i y ε_{ij} son independientes, y que τ_i tiene como varianza σ_τ^2 , entonces la varianza de una observación dada es

$$\text{Var}(y_{ij}) = \sigma_\tau^2 + \sigma^2.$$

Se denomina a σ_τ^2 y a σ^2 como los *componentes de la varianza* y se supone que

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

$$\tau_i \sim N(0, \sigma_\tau^2)$$

independientemente entre sí.

Ahora carece de sentido contrastar hipótesis basadas en tratamientos individuales, por lo que se contrasta:

$$H_0 \equiv \sigma_\tau^2 = 0$$

$$H_1 \equiv \sigma_\tau^2 > 0$$

Todos los tratamientos serán iguales si $\sigma_\tau^2 = 0$. Sin embargo, si $\sigma_\tau^2 > 0$ existe variabilidad entre los tratamientos.

En este caso, si H_0 es cierta, $\sigma_\tau^2 = 0$, entonces

$$F_0 = \frac{\frac{SCTra}{a-1}}{\frac{SCE}{N-a}} = \frac{MCTra}{MCE} \sim F_{a-1, N-a}$$

Si se consideran los valores esperados de las medias de cuadrados, entonces

$$\begin{aligned} E[MCTra] &= \frac{1}{a-1} E[SCTra] = \\ &= \frac{1}{a-1} E \left[\sum_{i=1}^a \frac{y_i^2}{n} - \frac{y_{..}^2}{N} \right] = \sigma^2 + n\sigma_\tau^2. \end{aligned}$$

Del mismo modo, se obtiene que

$$E[MCE] = \sigma^2.$$

Si la hipótesis alternativa es cierta, entonces el valor esperado del numerador en F_0 es mayor que el esperado del denominador. Así, se rechaza H_0 para valores altos de F_0 , con lo cual, la región crítica es unilateral superior, rechazándose si

$$F_0 > F_{a-1, N-a, \alpha}$$

El procedimiento de cálculo es igual que en el modelo de efectos fijos, aunque las conclusiones se aplican a toda la población de tratamientos.

Estima de los componentes de la varianza

Si se igualan los valores esperados de las medias de cuadrados con los valores observados, se obtiene

$$\begin{aligned} MCTra &= \sigma^2 + n\sigma_\tau^2 \\ MCE &= \sigma^2 \end{aligned}$$

de donde

$$\begin{aligned} \hat{\sigma}^2 &= MCE \\ \hat{\sigma}_\tau^2 &= \frac{MCTra - MCE}{n} \end{aligned}$$

NOTA:

Si n_i , para $i = 1, \dots, a$ son distintos entre sí, se sustituye en la expresión anterior n por

$$n_0 = \frac{1}{a-1} \left[\sum_{i=1}^a n_i - \frac{\sum_{i=1}^a n_i^2}{\sum_{i=1}^a n_i} \right].$$

Ejemplo.

Una fábrica de maquinillas de afeitar utiliza una gran cantidad de máquinas en la producción. Se desea que las máquinas sean homogéneas para producir objetos de la misma calidad. Para investigar si existen variaciones significativas entre las máquinas, se seleccionan 4 al azar y se mide el porcentaje de un cierto componente de la hoja. El experimento se realiza con orden aleatorio.

						$y_{i\cdot}$
<i>Máquina 1</i>	98	97	99	96		390
<i>Máquina 2</i>	91	90	93	92		366
<i>Máquina 3</i>	96	95	97	95		383
<i>Máquina 4</i>	95	96	99	98		388
						$y_{\cdot\cdot} = 1527$

Se obtiene la siguiente tabla ANOVA:

F.V.	S.C.	G.L.	M.C.	F₀
<i>Explicada</i>	89.19	3	29.73	15.68
<i>Residual</i>	22.75	12	1.90	
<i>Total</i>	11.94	15		

Como

$$F_{3,12,0.05} = 3,49 < 15,68$$

Se rechaza $H_0 \equiv \sigma_\tau = 0$.

Estimación de los componentes de la varianza:

$$\begin{aligned}\hat{\sigma}^2 &= MCE = 1,90 \\ \hat{\sigma}_\tau^2 &= \frac{MCTra - MCE}{n} = \frac{29,73 - 1,90}{4} = 6,96.\end{aligned}$$

La estimación de la varianza de cualquier observación de la muestra es

$$\hat{\sigma}^2 + \hat{\sigma}_\tau^2 = 1,90 + 6,96 = 8,86$$

y la mayor parte de la variabilidad se debe a diferencias entre las máquinas.

Intervalos de confianza para los componentes de la varianza

El intervalo de confianza para σ^2 al $100(1 - \alpha)\%$ es

$$\frac{(N - a)MCE}{\chi_{N-a, \frac{\alpha}{2}}^2} \leq \sigma^2 \leq \frac{(N - a)MCE}{\chi_{N-a, 1-\frac{\alpha}{2}}^2}$$

El intervalo de confianza para σ_τ^2 no se puede calcular de modo exacto, dado que depende de una combinación lineal de χ^2 's. Por tanto se calcula el intervalo para el cociente

$$\frac{\sigma_\tau^2}{\sigma^2 + \sigma_\tau^2}.$$

Se denomina

$$l_1 = \frac{1}{n} \left(\frac{MCTra}{MCE} \frac{1}{F_{a-1, N-a, \frac{\alpha}{2}}} - 1 \right)$$

$$l_2 = \frac{1}{n} \left(\frac{MCTra}{MCE} \frac{1}{F_{a-1, N-a, 1-\frac{\alpha}{2}}} - 1 \right)$$

entonces el intervalo de confianza al $100(1 - \alpha)\%$ es

$$\frac{l_1}{1 + l_1} \leq \frac{\sigma_\tau^2}{\sigma^2 + \sigma_\tau^2} \leq \frac{l_2}{1 + l_2}.$$

Ejemplo.

En el caso de la fábrica de maquinillas de afeitar,

$$F_{a-1, N-a, \frac{\alpha}{2}} = F_{3, 12, 0, 025} = 4,47$$

$$F_{a-1, N-a, 1-\frac{\alpha}{2}} = F_{3, 12, 0, 975} = \frac{1}{F_{12, 3, 0, 025}} = 0,070.$$

De este modo

$$l_1 = \frac{1}{n} \left(\frac{MCTra}{MCE} \frac{1}{F_{a-1, N-a, \frac{\alpha}{2}}} - 1 \right) = 0,625$$

$$l_2 = \frac{1}{n} \left(\frac{MCTra}{MCE} \frac{1}{F_{a-1, N-a, 1-\frac{\alpha}{2}}} - 1 \right) = 54,883$$

de modo que

$$\begin{aligned}\frac{l_1}{1+l_1} &\leq \frac{\sigma_\tau^2}{\sigma^2 + \sigma_\tau^2} \leq \frac{l_2}{1+l_2} \\ \frac{0,625}{1,625} &\leq \frac{\sigma_\tau^2}{\sigma^2 + \sigma_\tau^2} \leq \frac{54,883}{55,883} \\ 0,39 &\leq \frac{\sigma_\tau^2}{\sigma^2 + \sigma_\tau^2} \leq 0,98\end{aligned}$$

Esto es, la variabilidad de las máquinas justifica entre el 40 % y el 98 % de la variabilidad total.

Test de Kruskal-Wallis

Cuando no está justificado asumir normalidad, se puede utilizar la metodología no paramétrica. El test de Kruskal-Wallis propone como hipótesis nula que los a tratamientos son iguales, frente a la hipótesis alternativa de que algunas observaciones son mayores que otras entre los tratamientos. Se puede considerar que este test es adecuado para contrastar la igualdad entre las medias.

Procedimiento.

Se calculan rangos de cada una de las observaciones y_{ij} de manera creciente y se reemplaza por su rango R_{ij} , donde la menor observación tendría el valor 1. En caso de empates, se asigna a todas las observaciones empatadas el valor medio de sus correspondientes rangos.

Se denota como R_i la suma de los rangos del i -ésimo tratamiento de modo que el estadístico es

$$H = \frac{1}{S^2} \left[\sum_{i=1}^a \frac{R_i^2}{n_i} - \frac{N(N+1)^2}{4} \right]$$

donde n_i es el número de observaciones que hay en el tratamiento i , N es el número total de observaciones y

$$S^2 = \frac{1}{N-1} \left[\sum_{i=1}^a \sum_{j=1}^{n_i} R_{ij}^2 - \frac{N(N+1)^2}{4} \right].$$

Se puede observar que S^2 es simplemente la varianza de los rangos. Si no hay empates, entonces $S^2 = \frac{N(N+1)}{12}$ y el test se simplifica, quedando el estadístico

$$H = \frac{12}{N(N+1)} \sum_{i=1}^a \frac{R_i^2}{n_i} - 3(N+1).$$

Para valores $n_i > 5$, H se distribuye aproximadamente como una χ_{a-1}^2 si la hipótesis nula es cierta. Por tanto, si $H > \chi_{a-1, \alpha}^2$ se rechaza la hipótesis nula a un nivel α .

Ejemplo.

En el ejemplo de las camisas fabricadas según su porcentaje de algodón, se tenían los siguientes datos:

% de algodón	Observaciones
15	7 7 15 11 9
20	12 17 12 18 18
25	14 18 18 19 19
30	19 25 22 19 23
35	7 10 11 15 11

Si se calculan los correspondientes rangos, se obtiene:

Rangos		Suma
R_{1j}	2 2 12.5 7 4	27.5
R_{2j}	9.5 14 9.5 16.5 16.5	66
R_{3j}	11 16.5 16.5 20.5 20.5	85
R_{4j}	20.5 25.5 23 20.5 24	113
R_{5j}	2 5 7 12.5 7	33.5

Así, calculando

$$S^2 = \frac{1}{N-1} \left[\sum_{i=1}^a \sum_{j=1}^{n_i} R_{ij}^2 - \frac{N(N+1)^2}{4} \right] =$$

$$\frac{1}{24} \left[5497,79 - \frac{25 \cdot 26^2}{4} \right] = 53,03$$

$$H = \frac{1}{S^2} \left[\sum_{i=1}^a \frac{R_{i.}^2}{n_i} - \frac{N(N+1)^2}{4} \right] =$$

$$= \frac{1}{53,03} \left[52,45 - \frac{25 \cdot 26^2}{4} \right] = 19,25.$$

Como $H > \chi_{4,0,01}^2 = 13,28$, entonces se rechaza la hipótesis nula obteniéndose la misma conclusión que en el caso de usar el test clásico paramétrico.

Test de aleatorización y test Bootstrap

Se pueden realizar tests de aleatorización sobre los errores para contrastar medias. El algoritmo es

1. Calcular el estadístico F_0 del modo habitual sobre los datos originales.
2. Calcular el *residuo* para cada observación, como la diferencia entre cada observación y la media de todas las observaciones dentro de su grupo correspondiente.
3. Asignar aleatoriamente los residuos en los grupos del mismo tamaño sumándolos a las medias de cada grupo, y calcular F_1 , el estadístico de la F de Snedecor obtenido sobre los nuevos datos generados.
4. Repetir el paso (3) un número N de veces para generar los valores F_1, F_2, \dots, F_N
5. Declarar que F_0 es significativo a un nivel α si es mayor que el valor correspondiente al percentil $(1 - \alpha)$ de los valores F_1, F_2, \dots, F_N .

Se puede modificar este algoritmo, cambiando el paso (3) remuestreando los residuos *con reemplazamiento* para producir nuevos conjuntos de datos. Este método da un test Bootstrap de significación que tiene propiedades similares al anterior. Sin embargo, hay una diferencia entre ambos métodos:

El test basado en aleatorización, se basa en la idea de que los residuos aparecen orden aleatorio, mientras que el método Bootstrap se basa en una aproximación a la distribución F de Snedecor que se obtendría remuestreando de las poblaciones de donde vienen los datos originales.

Selección del tamaño de una muestra

En diseño de experimentos un problema importante es el de determinar el número de réplicas que se tienen que realizar para cada tratamiento.

Una técnica frecuentemente empleada se basa en fijar el error de tipo II.

Observaciones.

Se puede cometer error de tipo I:

$$\alpha = P \{ \text{Rechazar } H_0 | H_0 \text{ es cierta} \}$$

o bien

$$\beta = P \{ \text{No Rechazar } H_0 | H_0 \text{ es falsa} \}$$

Se llama potencia de un test a la $P \{ \text{Rechazar } H_0 | H_0 \cup H_1 \}$, de modo que

$$1 - \beta = P \{ \text{Rechazar } H_0 | H_0 \text{ es falsa} \}$$

coincide con la potencia del test cuando H_0 es falsa.

Se trata de construir contrastes que tengan un tamaño α fijo y una potencia máxima (es decir un valor β pequeño) cosa que cumplen, por ejemplo, los test *UMP* (de uniformemente máxima potencia).

En este caso

$$\beta = 1 - P \{ F_0 > F_{a-1, N-a, \alpha} | H_0 \text{ es falsa} \}.$$

Para calcular esta probabilidad, se necesita conocer la distribución de $F_0 = \frac{MCTra}{MCE}$ cuando la hipótesis nula es falsa. Se puede demostrar que en ese caso, se distribuye como una F no centrada con $a - 1$ y $N - a$ grados de libertad y un cierto parámetro de centralidad.

Se utilizan curvas características que dibujan la probabilidad de error de tipo II (β) frente a un parámetro Φ donde

$$\Phi^2 = \frac{n \sum_{i=1}^a \tau_i^2}{a\sigma^2} = \frac{\sum_{i=1}^a n_i \tau_i^2}{a\sigma^2}$$

La cantidad Φ^2 está relacionada con el parámetro de centralidad, y se presentan habitualmente curvas para $\alpha = 0,05$ y $\alpha = 0,01$.

El parámetro anterior, depende de

1. Los valores τ_1, \dots, τ_a o bien μ_1, \dots, μ_a para los que se consideran medias *distintas*, ya que obviamente dichos valores no son conocidos previamente.
2. El valor de σ^2 , que al ser también desconocido, se suele usar el valor que se obtiene mediante una muestra *piloto*.
3. El número de réplicas por tratamiento.

Así, fijados los valores de τ_i y el valor de σ^2 se debe determinar n para que la potencia sea $(1 - \beta)$. Una manera de hacerlo es buscando en las tablas de curvas características de operación.

Ejemplo.

Supongamos que en el ejemplo de las prendas el experimentador está interesado en rechazar la igualdad entre los tratamientos con una probabilidad mínima de 0,9 (error de tipo II: $\beta = 0,1$).

Se asumen unas medias poblacionales por grupo igual a

$$\mu_1 = 11, \quad \mu_2 = 12, \quad \mu_3 = 15, \quad \mu_4 = 18, \quad \mu_5 = 19$$

de modo que la media total es $\mu = \frac{11+12+15+18+19}{5} = 15$.

Supongamos una estimación previa (mediante e.g. una muestra piloto) de $\hat{\sigma}^2 = 9$ y que el nivel α elegido es 0,01. Se tiene que $\tau_i = \mu_i - \mu$, de manera que

$$\tau_1 = 11 - 15 = -4$$

$$\tau_2 = 12 - 15 = -3$$

$$\tau_3 = 15 - 15 = 0$$

$$\tau_4 = 18 - 15 = 3$$

$$\tau_5 = 19 - 15 = 4$$

Entonces

$$\Phi^2 = \frac{n \sum_{i=1}^5 \tau_i^2}{5\hat{\sigma}^2} = \frac{n(16 + 16 + 9 + 9)}{5 \cdot 9} = 1,11 \cdot n$$

Se construye una tabla, dando distintos valores a n :

n	Φ^2	Φ	$g.l. (a - 1, a(n - 1))$	β	$(1-\beta)$ Potencia
4	4,44	2,11	(4, 15)	0,3	0,7
5	5,55	2,36	(4, 20)	0,15	0,85
6	6,66	2,58	(4, 25)	0,04	0,96

Por tanto es necesario realizar, al menos, 6 réplicas.

Lectura de las Curvas de Operación

- (i) Se elige la curva de operación. Para ello se calculan los grados de libertad: $a - 1 = 5 - 1 = 4$, y se elige la curva con $v_1 = 4$.
- (ii) Se fija el haz de curvas correspondiente al valor de α elegido: en el ejemplo, sería $\alpha = 0,01$.
- (iii) Se elige la curva correspondiente a $v_2 = a(n - 1)$. Por ejemplo, si $n = 4$, se tomaría $v_2 = 15$.
- (iv) En el eje X se busca el valor del parámetro Φ y se fija la ordenada para ese valor de Φ que muestra la curva de operación elegida en (iii). Por ejemplo, para $n = 4$, $v_2 = 15$, el valor está cerca de 0,30.
- (v) El valor de la probabilidad de error de tipo II está en la ordenada. En el ejemplo es 0.30, de manera que la potencia es $(1 - \beta) = 0,70$.

A menudo resulta difícil seleccionar las medias para cada tratamiento que se quieren usar, para determinar el tamaño de la muestra. Una alternativa consiste en considerar el valor de la máxima diferencia posible entre las medias: D .

Se puede demostrar que el valor mínimo de Φ^2 es

$$\Phi^2 = \frac{D^2 n}{2a\sigma^2}.$$

Como es el valor mínimo, entonces se obtiene el tamaño muestral adecuado para obtener como mínimo la potencia especificada.

Modelo de efectos aleatorios

En este modelo, se contrasta

$$H_0 \equiv \sigma_\tau = 0$$

$$H_1 \equiv \sigma_\tau > 0$$

En este caso, si H_1 es cierta, entonces

$$F_0 = \frac{MCTra}{MCE} \sim F_{a-1, N-a}$$

de manera que se pueden usar las tablas habituales de la F de Snedecor para determinar el tamaño muestral.

También se pueden usar curvas de operación característica, donde aparecen las gráficas del error de tipo II, frente al parámetro

$$\lambda = \sqrt{1 + \frac{n\sigma_\tau^2}{\sigma^2}}$$

Los términos σ_τ^2 y σ^2 al ser desconocidos se fijan dependiendo de la sensibilidad deseada para el experimento.

Aplicación con R

Se puede usar la librería *Rcmdr* de R, y ejecutar las siguientes sentencias en la ventana de *arriba* de Rcmdr:

```
library(Rcmdr)
Datos <- read.table("C:/CursoCIII/Disenno/Practic06/dat1Fac.txt",
header=TRUE, sep="", na.strings="NA", dec=".", strip.white=TRUE)

Datos$grupo <- factor(Datos$grupo,
labels=c('15%', '20%', '25%', '30%', '35%'))

tapply(Datos$medida, Datos$grupo, var, na.rm=TRUE)
levene.test(Datos$medida, Datos$grupo)

tapply(Datos$medida, Datos$grupo, var, na.rm=TRUE)
bartlett.test(medida ~ grupo, data=Datos)

anova(lm(medida ~ grupo, data=Datos))

tapply(Datos$medida, Datos$grupo, mean, na.rm=TRUE) # means
tapply(Datos$medida, Datos$grupo, sd, na.rm=TRUE) # std. deviations
tapply(Datos$medida, Datos$grupo, function(x) sum(!is.na(x))) # counts
tapply(Datos$medida, Datos$grupo, median, na.rm=TRUE)

kruskal.test(medida ~ grupo, data=Datos)

plotMeans(Datos$medida, Datos$grupo, error.bars="conf.int",
level=0.95)

boxplot(medida~grupo, ylab="medida", xlab="grupo", data=Datos)

#.....
```

Alternativamente, se puede hacer lo mismo con R pero mediante sentencias:

```
# Con Sintaxis

setwd("c:/Curso/... ")
datos <- read.table("dat1Fac.txt", header=T)
attach(datos)

elgrupo <- factor(grupo, labels=c('15%', '20%', '25%', '30%', '35%'))

# Para ver transformaciones de Box-Cox:
# Se busca el maximo de la funcion de verosimilitud
library(MASS)
boxcox(medida ~ grupo, data=datos, lambda=seq(-3, 3))

# De modo artesanal:
premedias <- lapply(1:5,function(eso){mean(medida[grupo==eso])})
predesv <- lapply(1:5,function(eso){sqrt(var(medida[grupo==eso]))})

medias <- NULL
medias <- for (i in 1: 5) {medias <- c(medias,premedias[[i]]) }

desv <- NULL
desv <- for (i in 1: 5) {desv <- c(desv,predesv[[i]]) }
```

```

lmedias <- log(medias)
ldesv <- log(desv)
mod <- lm(ldesv~lmedias)
summary(mod)

# El coeficiente de la transformacion es (1-pendiente)
# redondeado al valor mas proximo a multiplos de 0.5
lambda <- 1-mod$coefficients[[2]]

boxplot(medida ~ elgrupo, main="Distribución de medidas por grupos")

fac1 <- aov(medida ~ elgrupo)
summary(fac1)

coefficients(fac1)

# Graficas por defecto de aov
par(mfrow=c(2,2))
plot(fac1)

#.....

# Graficos de ajuste varios

# Grafica de ajuste a normalidad
qqnorm(fac1$res)
qqline(fac1$res)

plot(fac1$fit,fac1$res,xlab="Valores Ajustados",ylab="Residuos",
main="Residuos frente a niveles")
abline(h=0,lty=2)

# Analizo los residuos para verificar que cumple con las hipótesis
plot(fitted.values(fac1),rstandard(fac1),
xlab="Valores Ajustados", ylab="Residuos Estandarizados",pch=20)

plot(jitter(fac1$fit),fac1$res,xlab="Fitted",ylab="Residuos",
main="Grafico Jittered")

#.....

Se puede considerar un test no paramétrico, de modo alternativo:

# test de no parametrico de Kruskal-Wallis
krus <- kruskal.test(medida,elgrupo)
krus

```

Para comparaciones múltiples se pueden considerar el test de LSD, el de Bonferroni y el test de Tukey. El test de LSD hay que programarlo:

```
# test de LSD
n1 <- sum(fac1$model$grupo=="1")
n4 <- sum(fac1$model$grupo=="4")
s <- sqrt(sum((fac1$residuals)^2)/fac1$df.residual)
tcrit <- qt(0.025, fac1$df.residual, lower.tail=F)
LSD <- tcrit*s*sqrt((1/n1)+(1/n4))
LSD

# Metodo de Bonferroni
library(stats)
pairwise.t.test(medida,elgrupo,p.adjust.method="bonferroni")

# test de Tukey
TukeyHSD(aov(medida ~ elgrupo))
```

Aplicación con SAS

```
options ls=75 nodate nonumber;
title 'ANOVA UNIFACTORIAL DE EFECTOS FIJOS';
data anol;
input grupo medida;
cards;
1 7
1 7
1 15
1 11
1 9
2 12
2 17
2 12
2 18
2 18
3 14
3 18
3 18
3 19
3 19
4 19
4 25
4 22
4 19
4 23
5 7
5 10
5 11
5 15
5 11
;
proc anova;
class grupo;
model medida=grupo;
means grupo /duncan snk lsd tukey;
run;
```

ANOVA UNIFACTORIAL DE EFECTOS FIJOS

The ANOVA Procedure

Class Level Information

Class	Levels	Values
grupo	5	1 2 3 4 5

Number of observations 25

Dependent Variable: medida

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	475.7600000	118.9400000	14.76	<.0001
Error	20	161.2000000	8.0600000		
Corrected Total	24	636.9600000			

R-Square	Coeff Var	Root MSE	medida Mean
0.746923	18.87642	2.839014	15.04000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
grupo	4	475.7600000	118.9400000	14.76	<.0001

t Tests (LSD) for medida

NOTE: This test controls the Type I comparisonwise error rate, not the experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of t	2.08596
Least Significant Difference	3.7455

Means with the same letter are not significantly different.

t Grouping	Mean	N	grupo
A	21.600	5	4
B	17.600	5	3
B	15.400	5	2
C	10.800	5	5
C	9.800	5	1

Duncan's Multiple Range Test for medida

NOTE: This test controls the Type I comparisonwise error rate, not the experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06

Number of Means	2	3	4	5
Critical Range	3.745	3.931	4.050	4.132

Means with the same letter are not significantly different.

Duncan Grouping	Mean	N	grupo
A	21.600	5	4
B	17.600	5	3
B	15.400	5	2
C	10.800	5	5
C	9.800	5	1

Student-Newman-Keuls Test for medida

NOTE: This test controls the Type I experimentwise error rate under the complete null hypothesis but not under partial null hypotheses.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06

Number of Means	2	3	4	5
Critical Range	3.7454539	4.5427095	5.0256316	5.3729604

Means with the same letter are not significantly different.

SNK Grouping	Mean	N	grupo
A	21.600	5	4
B	17.600	5	3
B	15.400	5	2
C	10.800	5	5
C	9.800	5	1

Tukey's Studentized Range (HSD) Test for medida

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of Studentized Range	4.23186
Minimum Significant Difference	5.373

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	grupo
A	21.600	5	4
A			
B A	17.600	5	3
B			
B C	15.400	5	2
C			
D C	10.800	5	5
D			
D	9.800	5	1

```

options ls=75 nodate nonumber;
title 'ANOVA UNIFACTORIAL DE EFECTOS ALEATORIOS';
data anol;
input caja peso;
cards;
1 48
1 49
2 46
2 49
2 49
3 51
3 50
3 50
3 52
3 49
4 51
4 51
4 52
4 53
5 52
5 50
5 53
6 50
6 50
6 51
6 49
;
proc glm;
class caja;
model peso=caja;
random caja/ test;

proc varcomp method=typel;
class caja;
model peso=caja;
run;

```

ANOVA UNIFACTORIAL DE EFECTOS ALEATORIOS

The GLM Procedure

Class Level Information

Class	Levels	Values
caja	6	1 2 3 4 5 6

Number of observations 21

Dependent Variable: peso

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	36.69285714	7.33857143	5.21	0.0057
Error	15	21.11666667	1.40777778		
Corrected Total	20	57.80952381			

R-Square Coeff Var Root MSE peso Mean
 0.634720 2.361750 1.186498 50.23810

Source	DF	Type I SS	Mean Square	F Value	Pr > F
caja	5	36.69285714	7.33857143	5.21	0.0057

Source	DF	Type III SS	Mean Square	F Value	Pr > F
caja	5	36.69285714	7.33857143	5.21	0.0057

Source Type III Expected Mean Square
 caja Var(Error) + 3.4476 Var(caja)

Tests of Hypotheses for Random Model Analysis of Variance

Dependent Variable: peso

Source	DF	Type III SS	Mean Square	F Value	Pr > F
caja	5	36.692857	7.338571	5.21	0.0057
Error: MS(Error)	15	21.116667	1.407778		

Variance Components Estimation Procedure

Class Level Information

Class	Levels	Values
caja	6	1 2 3 4 5 6

Number of observations 21
 Dependent Variable: peso

Type 1 Analysis of Variance

Source	DF	Sum of Squares	Mean Square
caja	5	36.692857	7.338571
Error	15	21.116667	1.407778
Corrected Total	20	57.809524	.

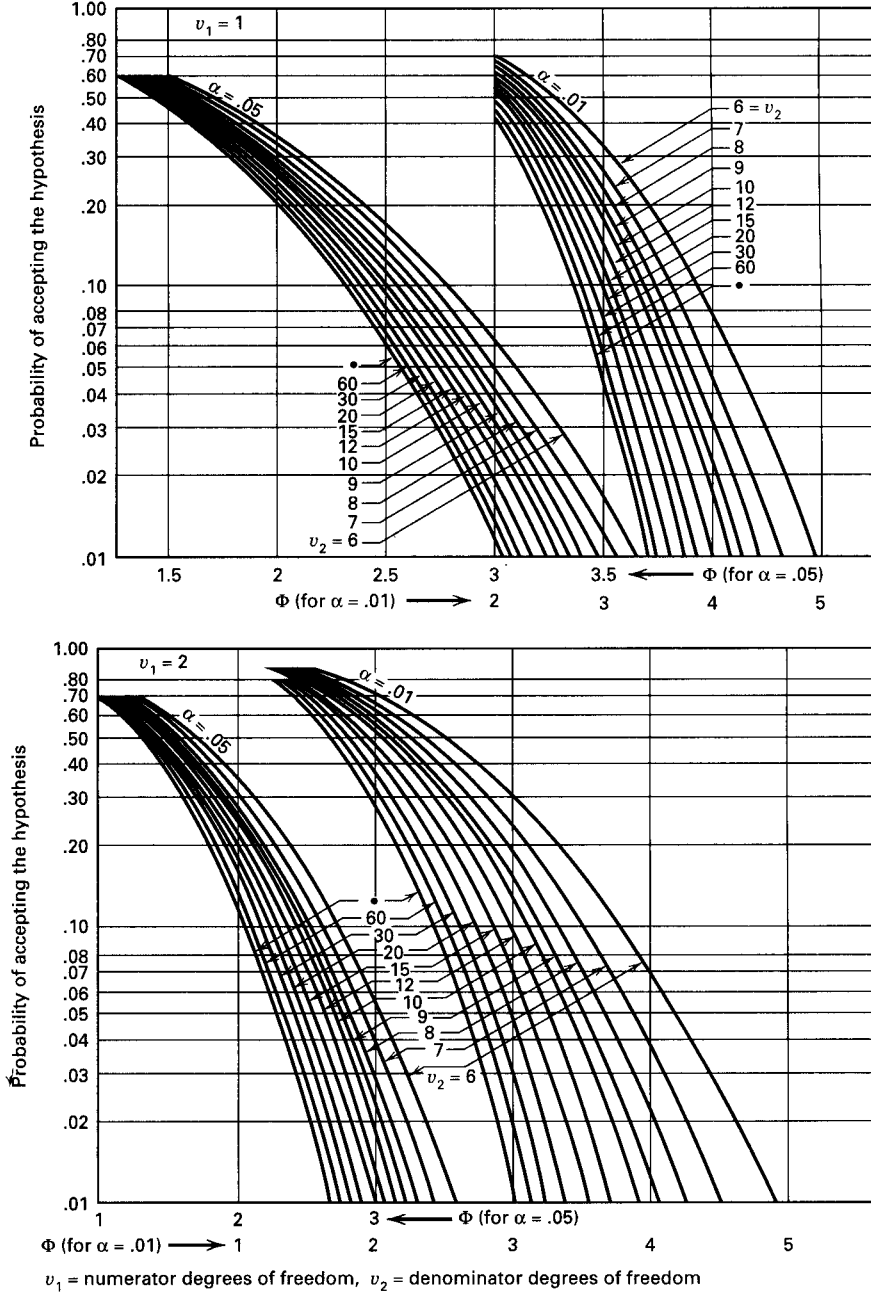
Type 1 Analysis of Variance

Source	Expected Mean Square
caja	$\text{Var}(\text{Error}) + 3.4476 \text{ Var}(\text{caja})$
Error	$\text{Var}(\text{Error})$
Corrected Total	.

Type 1 Estimates

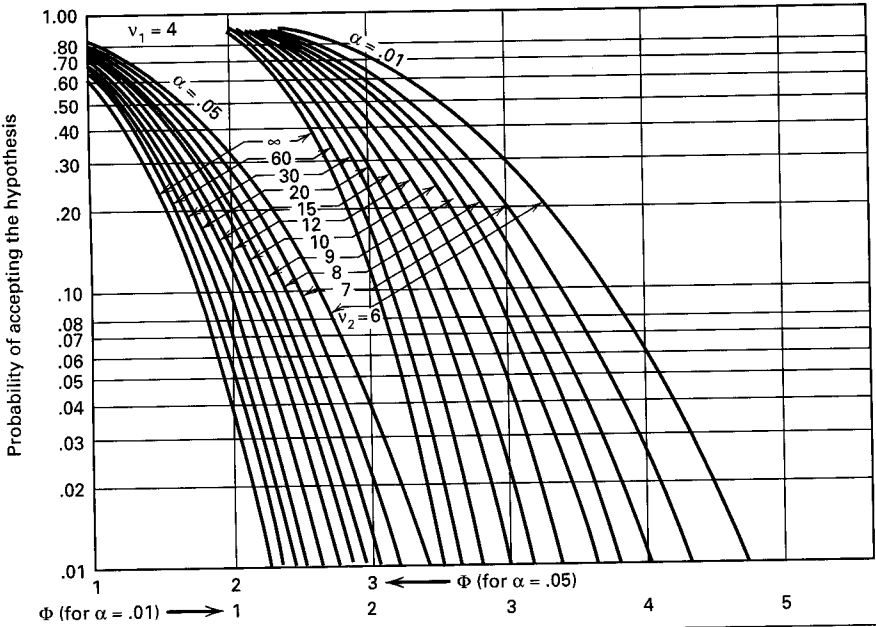
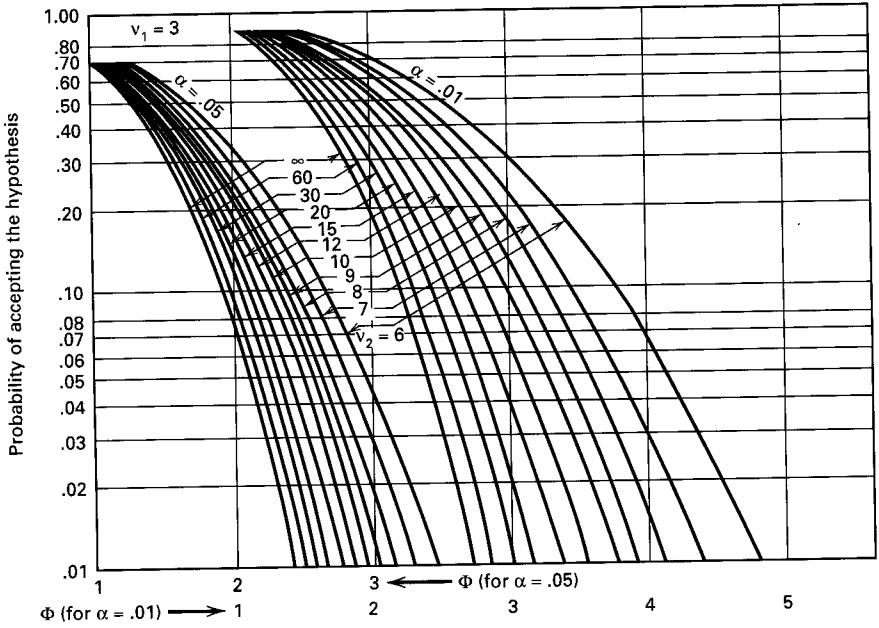
Variance Component	Estimate
Var(caja)	1.72026
Var(Error)	1.40778

V. Operating Characteristic Curves for the Fixed Effects Model Analysis of Variance^a

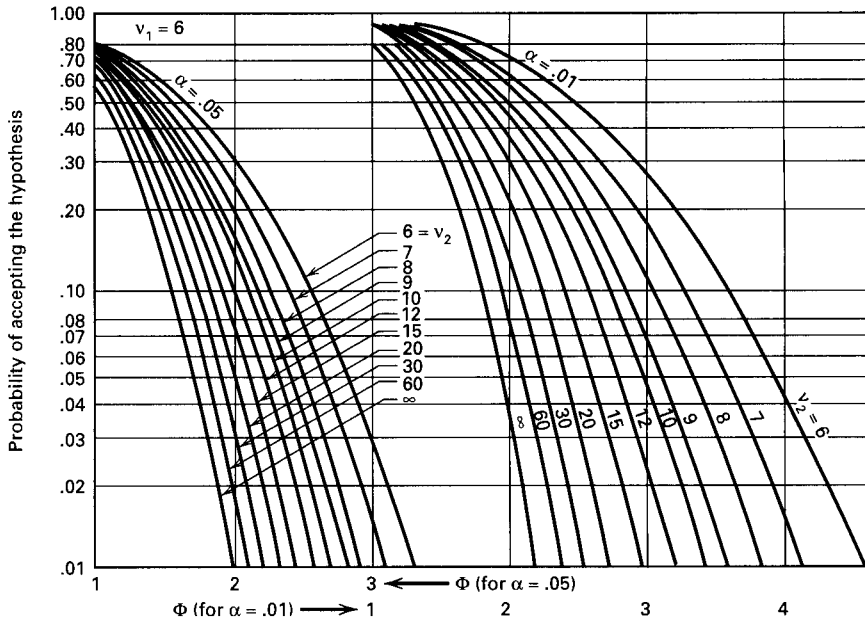
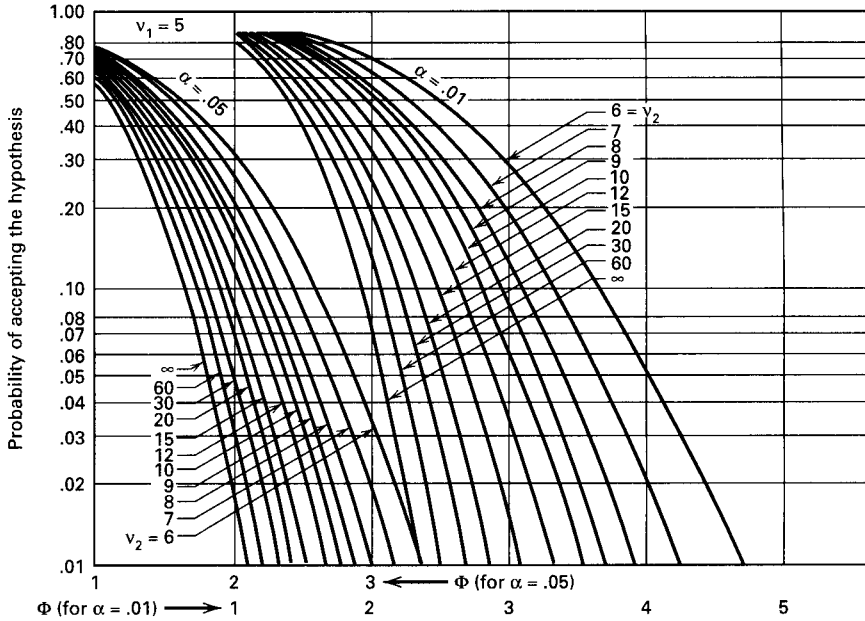


^a Adapted with permission from *Biometrika Tables for Statisticians*, Vol. 2, by E. S. Pearson and H. O. Hartley, Cambridge University Press, Cambridge, 1972.

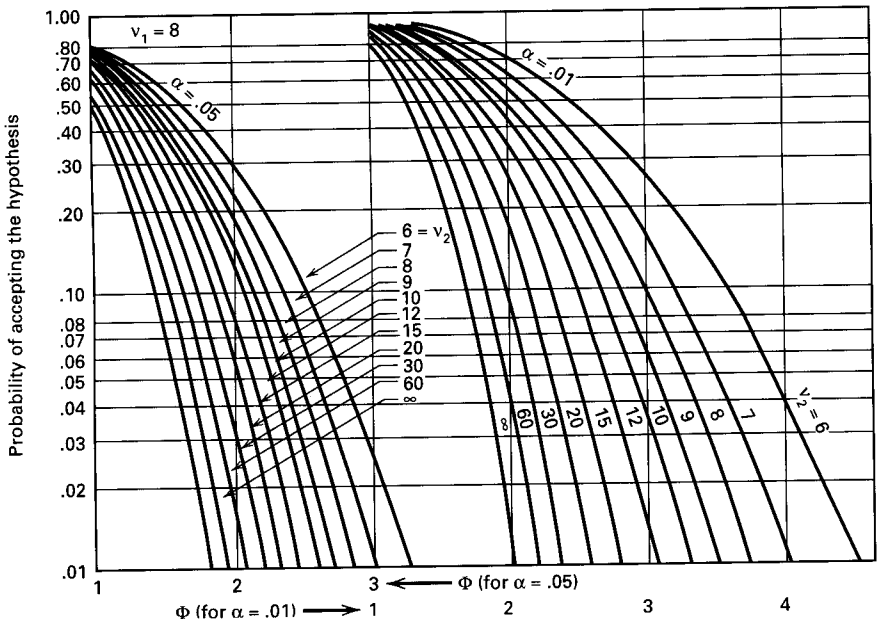
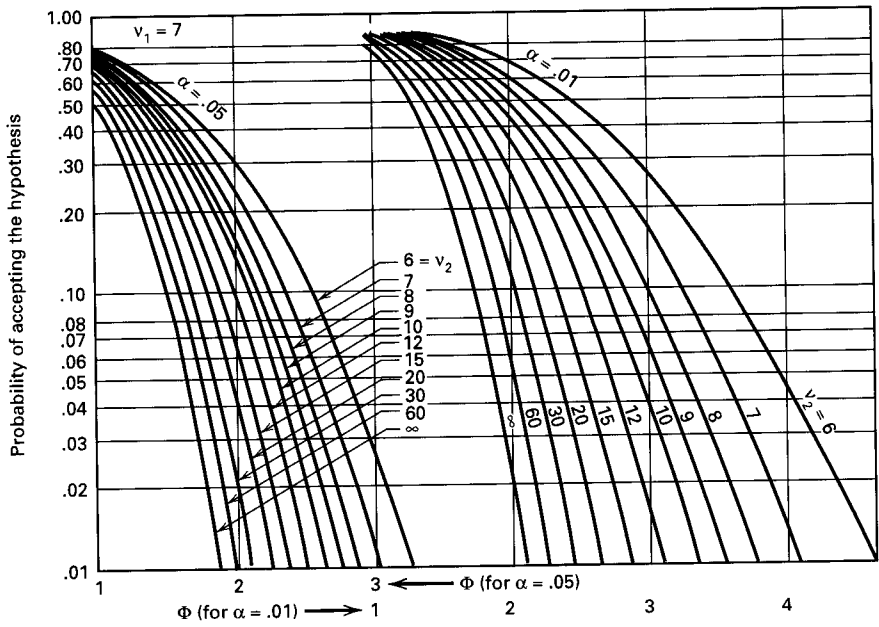
V. Operating Characteristic Curves for the Fixed Effects Model Analysis of Variance (continued)



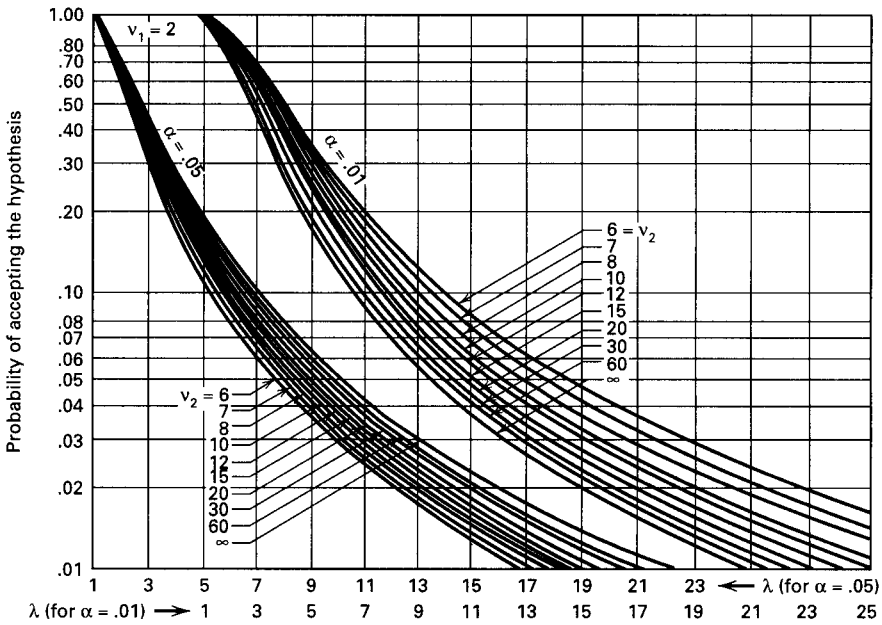
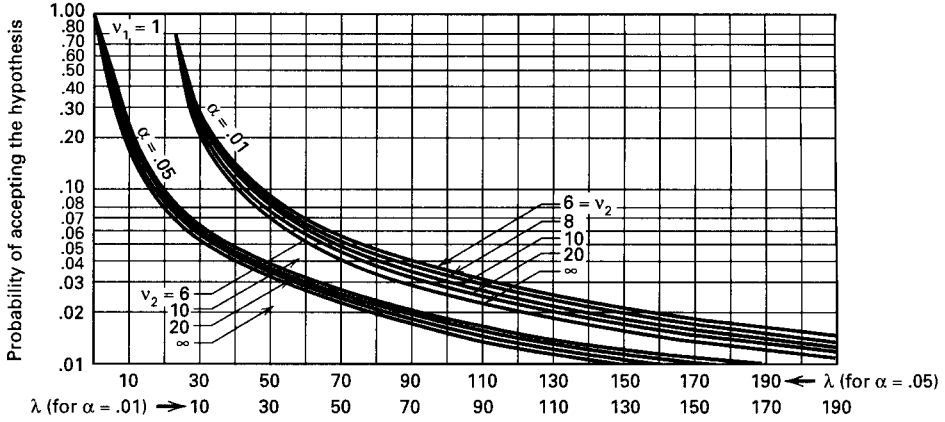
V. Operating Characteristic Curves for the Fixed Effects Model Analysis of Variance (continued)



V. Operating Characteristic Curves for the Fixed Effects Model Analysis of Variance (continued)

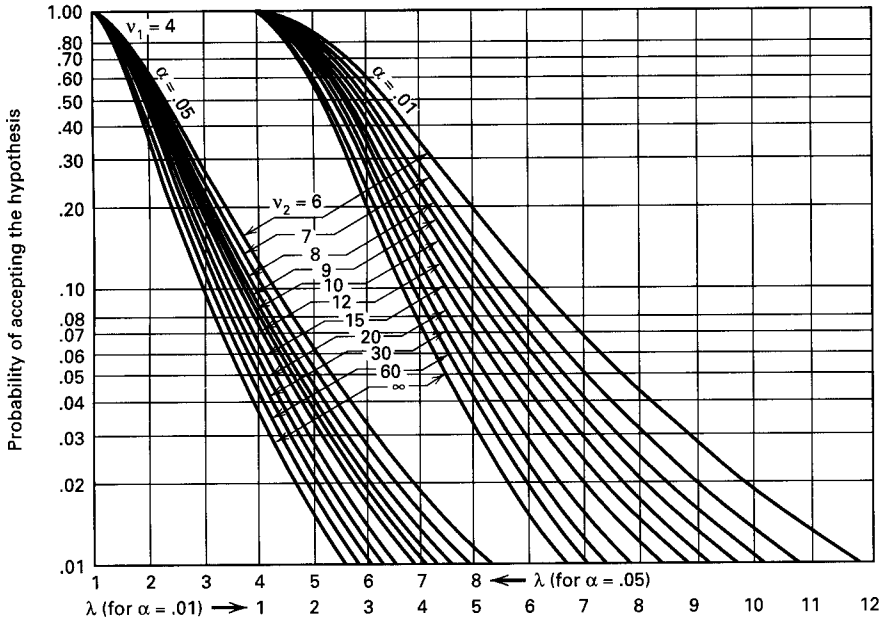
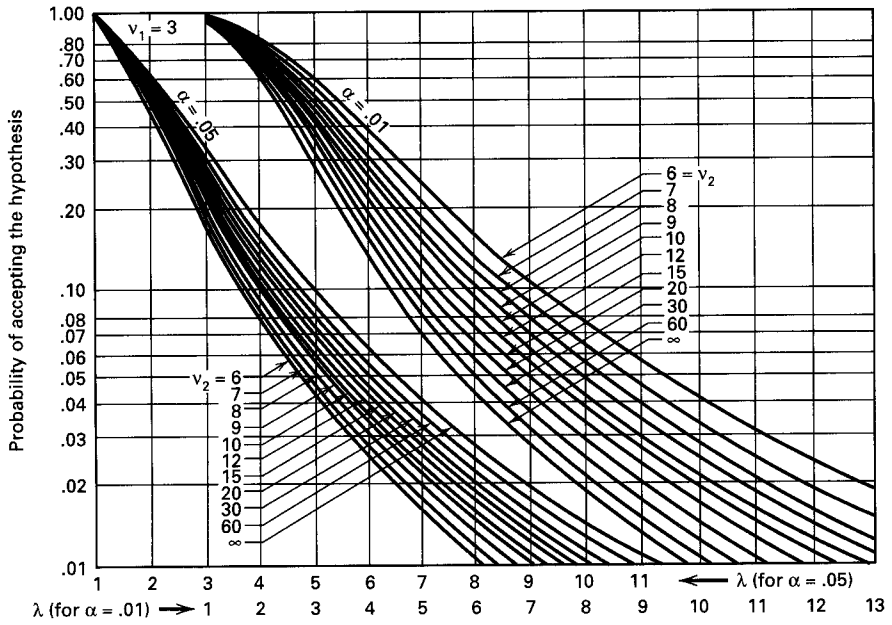


VI. Operating Characteristic Curves for the Random Effects Model Analysis of Variance^a

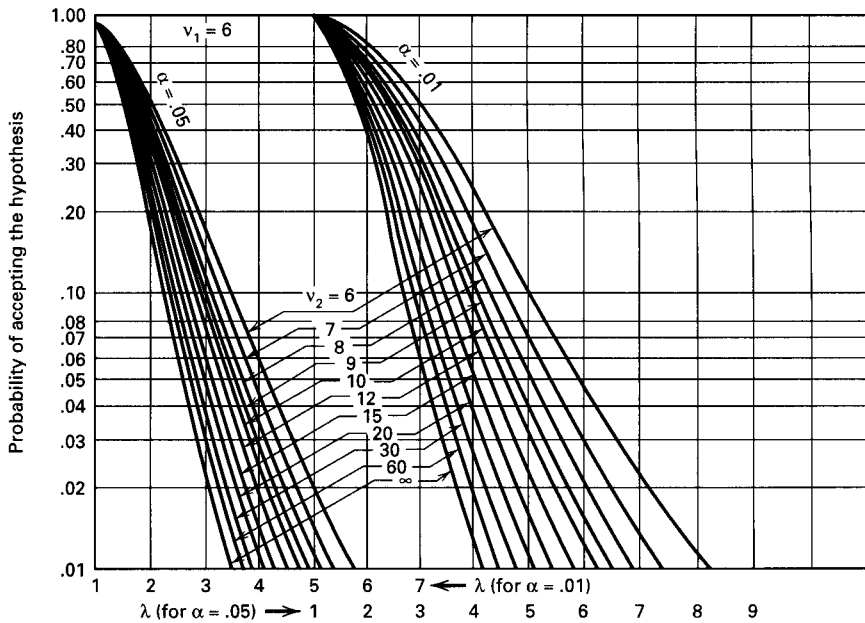
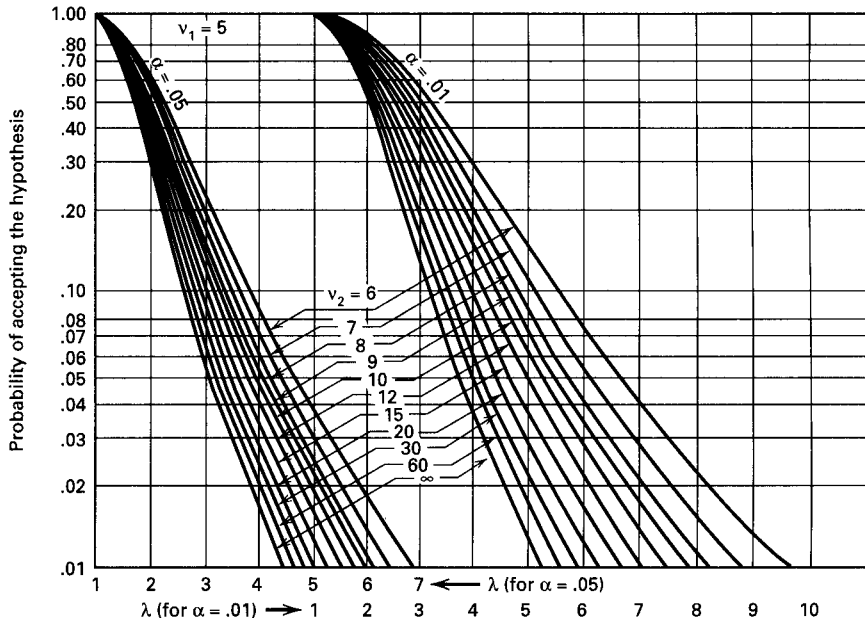


^a Reproduced with permission from *Engineering Statistics*, 2nd edition, by A. H. Bowker and G. J. Lieberman, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1972.

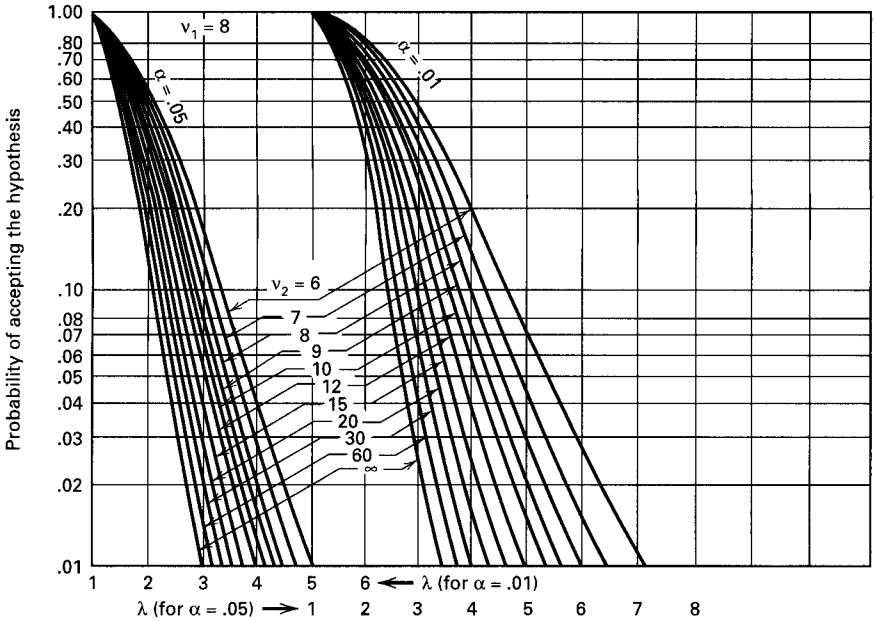
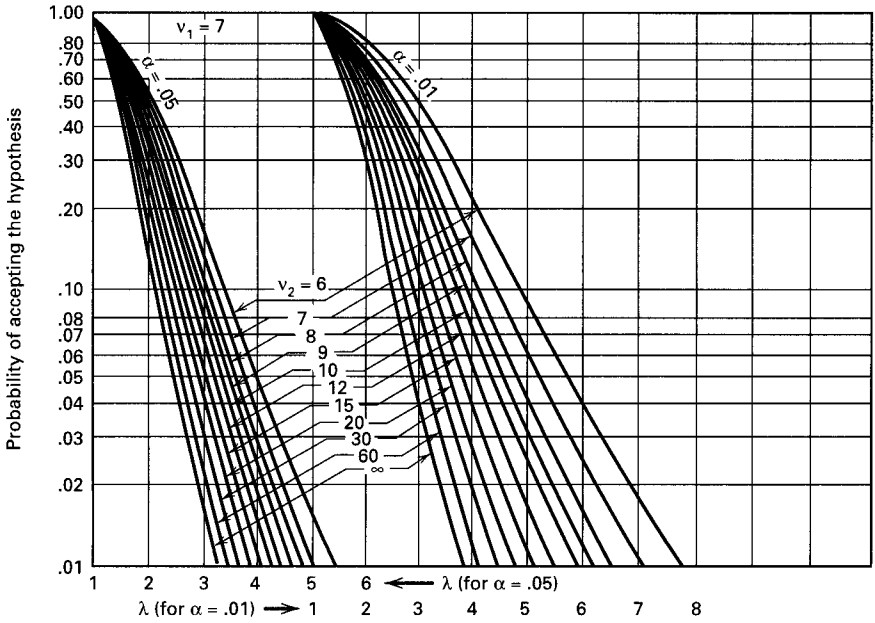
VI. Operating Characteristic Curves for the Random Effects Model Analysis of Variance (continued)



VI. Operating Characteristic Curves for the Random Effects Model Analysis of Variance (continued)



VI. Operating Characteristic Curves for the Random Effects Model Analysis of Variance (continued)



VII. Significant Ranges for Duncan's Multiple Range Test^a

$$r_{0.01}(p, f)$$

f	p											
	2	3	4	5	6	7	8	9	10	20	50	100
1	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0
2	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0
3	8.26	8.5	8.6	8.7	8.8	8.9	8.9	9.0	9.0	9.3	9.3	9.3
4	6.51	6.8	6.9	7.0	7.1	7.1	7.2	7.2	7.3	7.5	7.5	7.5
5	5.70	5.96	6.11	6.18	6.26	6.33	6.40	6.44	6.5	6.8	6.8	6.8
6	5.24	5.51	5.65	5.73	5.81	5.88	5.95	6.00	6.0	6.3	6.3	6.3
7	4.95	5.22	5.37	5.45	5.53	5.61	5.69	5.73	5.8	6.0	6.0	6.0
8	4.74	5.00	5.14	5.23	5.32	5.40	5.47	5.51	5.5	5.8	5.8	5.8
9	4.60	4.86	4.99	5.08	5.17	5.25	5.32	5.36	5.4	5.7	5.7	5.7
10	4.48	4.73	4.88	4.96	5.06	5.13	5.20	5.24	5.28	5.55	5.55	5.55
11	4.39	4.63	4.77	4.86	4.94	5.01	5.06	5.12	5.15	5.39	5.39	5.39
12	4.32	4.55	4.68	4.76	4.84	4.92	4.96	5.02	5.07	5.26	5.26	5.26
13	4.26	4.48	4.62	4.69	4.74	4.84	4.88	4.94	4.98	5.15	5.15	5.15
14	4.21	4.42	4.55	4.63	4.70	4.78	4.83	4.87	4.91	5.07	5.07	5.07
15	4.17	4.37	4.50	4.58	4.64	4.72	4.77	4.81	4.84	5.00	5.00	5.00
16	4.13	4.34	4.45	4.54	4.60	4.67	4.72	4.76	4.79	4.94	4.94	4.94
17	4.10	4.30	4.41	4.50	4.56	4.63	4.68	4.73	4.75	4.89	4.89	4.89
18	4.07	4.27	4.38	4.46	4.53	4.59	4.64	4.68	4.71	4.85	4.85	4.85
19	4.05	4.24	4.35	4.43	4.50	4.56	4.61	4.64	4.67	4.82	4.82	4.82
20	4.02	4.22	4.33	4.40	4.47	4.53	4.58	4.61	4.65	4.79	4.79	4.79
30	3.89	4.06	4.16	4.22	4.32	4.36	4.41	4.45	4.48	4.65	4.71	4.71
40	3.82	3.99	4.10	4.17	4.24	4.30	4.34	4.37	4.41	4.59	4.69	4.69
60	3.76	3.92	4.03	4.12	4.17	4.23	4.27	4.31	4.34	4.53	4.66	4.66
100	3.71	3.86	3.98	4.06	4.11	4.17	4.21	4.25	4.29	4.48	4.64	4.65
∞	3.64	3.80	3.90	3.98	4.04	4.09	4.14	4.17	4.20	4.41	4.60	4.68

f = degrees of freedom.

^a Reproduced with permission from "Multiple Range and Multiple F Tests," by D. B. Duncan, *Biometrics*, Vol. 1, No. 1, pp. 1-42, 1955.

$$r_{0.05}(p, f)$$

f	p											
	2	3	4	5	6	7	8	9	10	20	50	100
1	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0
2	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09
3	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50
4	3.93	4.01	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02
5	3.64	3.74	3.79	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83
6	3.46	3.58	3.64	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68
7	3.35	3.47	3.54	3.58	3.60	3.61	3.61	3.61	3.61	3.61	3.61	3.61
8	3.26	3.39	3.47	3.52	3.55	3.56	3.56	3.56	3.56	3.56	3.56	3.56
9	3.20	3.34	3.41	3.47	3.50	3.52	3.52	3.52	3.52	3.52	3.52	3.52
10	3.15	3.30	3.37	3.43	3.46	3.47	3.47	3.47	3.47	3.48	3.48	3.48
11	3.11	3.27	3.35	3.39	3.43	3.44	3.45	3.46	3.46	3.48	3.48	3.48
12	3.08	3.23	3.33	3.36	3.40	3.42	3.44	3.44	3.46	3.48	3.48	3.48
13	3.06	3.21	3.30	3.35	3.38	3.41	3.42	3.44	3.45	3.47	3.47	3.47
14	3.03	3.18	3.27	3.33	3.37	3.39	3.41	3.42	3.44	3.47	3.47	3.47
15	3.01	3.16	3.25	3.31	3.36	3.38	3.40	3.42	3.43	3.47	3.47	3.47
16	3.00	3.15	3.23	3.30	3.34	3.37	3.39	3.41	3.43	3.47	3.47	3.47
17	2.98	3.13	3.22	3.28	3.33	3.36	3.38	3.40	3.42	3.47	3.47	3.47
18	2.97	3.12	3.21	3.27	3.32	3.35	3.37	3.39	3.41	3.47	3.47	3.47
19	2.96	3.11	3.19	3.26	3.31	3.35	3.37	3.39	3.41	3.47	3.47	3.47
20	2.95	3.10	3.18	3.25	3.30	3.34	3.36	3.38	3.40	3.47	3.47	3.47
30	2.89	3.04	3.12	3.20	3.25	3.29	3.32	3.35	3.37	3.47	3.47	3.47
40	2.86	3.01	3.10	3.17	3.22	3.27	3.30	3.33	3.35	3.47	3.47	3.47
60	2.83	2.98	3.08	3.14	3.20	3.24	3.28	3.31	3.33	3.47	3.48	3.48
100	2.80	2.95	3.05	3.12	3.18	3.22	3.26	3.29	3.32	3.47	3.53	3.53
∞	2.77	2.92	3.02	3.09	3.15	3.19	3.23	3.26	3.29	3.47	3.61	3.67

VIII. Percentage Points of the Studentized Range Statistic^a
 $q_{\alpha,0.1}(p, f)$

f	p																			
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	90.0	135	164	186	202	216	227	237	246	253	260	266	272	272	282	286	290	294	298	
2	14.0	19.0	22.3	24.7	26.6	28.2	29.5	30.7	31.7	32.6	33.4	34.4	34.8	35.4	36.0	36.5	37.0	37.5	37.9	
3	8.26	10.6	12.2	13.3	14.2	15.0	15.6	16.2	16.7	17.1	17.5	17.9	18.2	18.5	18.8	19.1	19.3	19.5	19.8	
4	6.51	8.12	9.17	9.96	10.6	11.1	11.5	11.9	12.3	12.6	12.8	13.1	13.3	13.5	13.7	13.9	14.1	14.2	14.4	
5	5.70	6.97	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48	10.70	10.89	11.08	11.24	11.40	11.55	11.68	11.81	11.93	
6	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30	9.49	9.65	9.81	9.95	10.08	10.21	10.32	10.43	10.54	
7	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55	8.71	8.86	9.00	9.12	9.24	9.35	9.46	9.55	9.65	
8	4.74	5.63	6.20	6.63	6.96	7.24	7.47	7.68	7.87	8.03	8.18	8.31	8.44	8.55	8.66	8.76	8.85	8.94	9.03	
9	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.32	7.49	7.65	7.78	7.91	8.03	8.13	8.23	8.32	8.41	8.49	8.57	
10	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36	7.48	7.60	7.71	7.81	7.91	7.99	8.07	8.15	8.22	
11	4.39	5.14	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13	7.25	7.36	7.46	7.56	7.65	7.73	7.81	7.88	7.95	
12	4.32	5.04	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94	7.06	7.17	7.26	7.36	7.44	7.52	7.59	7.66	7.73	
13	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79	6.90	7.01	7.10	7.19	7.27	7.34	7.42	7.48	7.55	
14	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66	6.77	6.87	6.96	7.05	7.12	7.20	7.27	7.33	7.39	
15	4.17	4.83	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55	6.66	6.76	6.84	6.93	7.00	7.07	7.14	7.20	7.26	
16	4.13	4.78	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46	6.56	6.66	6.74	6.82	6.90	6.97	7.03	7.09	7.15	
17	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38	6.48	6.57	6.66	6.73	6.80	6.87	6.94	7.00	7.05	
18	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31	6.41	6.50	6.58	6.65	6.72	6.79	6.85	6.91	6.96	
19	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25	6.34	6.43	6.51	6.58	6.65	6.72	6.78	6.84	6.89	
20	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19	6.29	6.37	6.45	6.52	6.59	6.65	6.71	6.76	6.82	
24	3.96	4.54	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02	6.11	6.19	6.26	6.33	6.39	6.45	6.51	6.56	6.61	
30	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85	5.93	6.01	6.08	6.14	6.20	6.26	6.31	6.36	6.41	
40	3.82	4.37	4.70	4.93	5.11	5.27	5.39	5.50	5.60	5.69	5.77	5.84	5.90	5.96	6.02	6.07	6.12	6.17	6.21	
60	3.76	4.28	4.60	4.82	4.99	5.13	5.25	5.36	5.45	5.53	5.60	5.67	5.73	5.79	5.84	5.89	5.93	5.98	6.02	
120	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.38	5.44	5.51	5.56	5.61	5.66	5.71	5.75	5.79	5.83	
∞	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23	5.29	5.35	5.40	5.45	5.49	5.54	5.57	5.61	5.65	

f = degrees of freedom.

^a From J. M. May, "Extended and Corrected Tables of the Upper Percentage Points of the Studentized Range," *Biometrika*, Vol. 39, pp. 192-193, 1952. Reproduced by permission of the trustees, of *Biometrika*.

VIII. Percentage Points of the Studentized Range Statistic (continued)

f	$q_{0.05}(p, f)$																			
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	18.1	26.7	32.8	37.2	40.5	43.1	45.4	47.3	49.1	50.6	51.9	53.2	54.3	55.4	56.3	57.2	58.0	58.8	59.6	
2	6.09	8.28	9.80	10.89	11.73	12.43	13.03	13.54	13.99	14.39	14.75	15.08	15.38	15.65	15.91	16.14	16.36	16.57	16.77	
3	4.50	5.88	6.83	7.51	8.04	8.47	8.85	9.18	9.46	9.72	9.95	10.16	10.35	10.52	10.69	10.84	10.98	11.12	11.24	
4	3.93	5.00	5.76	6.31	6.73	7.06	7.35	7.60	7.83	8.03	8.21	8.37	8.52	8.67	8.80	8.92	9.03	9.14	9.24	
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32	7.47	7.60	7.72	7.83	7.93	8.03	8.12	8.21	
6	3.46	4.34	4.90	5.31	5.63	5.89	6.12	6.32	6.49	6.65	6.79	6.92	7.04	7.14	7.24	7.34	7.43	7.51	7.59	
7	3.34	4.16	4.68	5.06	5.35	5.59	5.80	5.99	6.15	6.29	6.42	6.54	6.65	6.75	6.84	6.93	7.01	7.08	7.16	
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18	6.29	6.39	6.48	6.57	6.65	6.73	6.80	6.87	
9	3.20	3.95	4.42	4.76	5.02	5.24	5.43	5.60	5.74	5.87	5.98	6.09	6.19	6.28	6.36	6.44	6.51	6.58	6.65	
10	3.15	3.88	4.33	4.66	4.91	5.12	5.30	5.46	5.60	5.72	5.83	5.93	6.03	6.12	6.20	6.27	6.34	6.41	6.47	
11	3.11	3.82	4.26	4.58	4.82	5.03	5.20	5.35	5.49	5.61	5.71	5.81	5.90	5.98	6.06	6.14	6.20	6.27	6.33	
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.40	5.51	5.61	5.71	5.80	5.88	5.95	6.02	6.09	6.15	6.21	
13	3.06	3.73	4.15	4.46	4.69	4.88	5.05	5.19	5.32	5.43	5.53	5.63	5.71	5.79	5.86	5.93	6.00	6.06	6.11	
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36	5.46	5.56	5.64	5.72	5.79	5.86	5.92	5.98	6.03	
15	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20	5.31	5.40	5.49	5.57	5.65	5.72	5.79	5.85	5.91	5.96	
16	3.00	3.65	4.05	4.34	4.56	4.74	4.90	5.03	5.15	5.26	5.35	5.44	5.52	5.59	5.66	5.73	5.79	5.84	5.90	
17	2.98	3.62	4.02	4.31	4.52	4.70	4.86	4.99	5.11	5.21	5.31	5.39	5.47	5.55	5.61	5.68	5.74	5.79	5.84	
18	2.97	3.61	4.00	4.28	4.49	4.67	4.83	4.96	5.07	5.17	5.27	5.35	5.43	5.50	5.57	5.63	5.69	5.74	5.79	
19	2.96	3.59	3.98	4.26	4.47	4.64	4.79	4.92	5.04	5.14	5.23	5.32	5.39	5.46	5.53	5.59	5.65	5.70	5.75	
20	2.95	3.58	3.96	4.24	4.45	4.62	4.77	4.90	5.01	5.11	5.20	5.28	5.36	5.43	5.50	5.56	5.61	5.66	5.71	
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01	5.10	5.18	5.25	5.32	5.38	5.44	5.50	5.55	5.59	
30	2.89	3.48	3.84	4.11	4.30	4.46	4.60	4.72	4.83	4.92	5.00	5.08	5.15	5.21	5.27	5.33	5.38	5.43	5.48	
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.74	4.82	4.90	4.98	5.05	5.11	5.17	5.22	5.27	5.32	5.36	
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73	4.81	4.88	4.94	5.00	5.06	5.11	5.15	5.20	5.24	
120	2.80	3.36	3.69	3.92	4.10	4.24	4.36	4.47	4.56	4.64	4.71	4.78	4.84	4.90	4.95	5.00	5.04	5.09	5.13	
∞	2.77	3.32	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55	4.62	4.68	4.74	4.80	4.84	4.98	4.93	4.97	5.01	