



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**Programa de Maestría y Doctorado en Ciencias Matemáticas y de la  
Especialización en Estadística Aplicada**

**TAMAÑO DE MUESTRA EN DISEÑO DE EXPERIMENTOS**

**TESINA QUE PARA OPTAR POR EL GRADO DE  
Especialista en Estadística Aplicada**

**PRESENTA:  
Carlos Gabriel Sánchez Lordméndez**

**TUTOR PRINCIPAL  
M. en C. Patricia Romero Mares  
Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas, UNAM**

Ciudad de Universitaria, CD. MX. Diciembre 2016.



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## Contenido

Objetivos y delimitaciones. ....	1
1. Conceptos importantes relacionados a la determinación del tamaño de muestra. ....	2
1.1. Importancia de la determinación del tamaño de muestra en diseño de experimentos. ....	2
1.2. Algunas distribuciones de probabilidad.....	2
1.3. Conceptos relacionados a pruebas de hipótesis.....	4
2. Determinación del tamaño de muestra en un diseño con un factor fijo y dos niveles (diferencia de medias)...8	
2.1. Caso en el que se conocen las varianzas.....	8
2.1.1. Determinación del tamaño de muestra en R. ....	13
2.1.2. Ejemplo de determinación del tamaño de muestra. ....	14
2.2. Caso en el que se desconocen las varianzas pero pueden suponerse iguales. ....	15
2.2.1. Determinación del tamaño de muestra en R. ....	15
2.2.2. Ejemplo de determinación del tamaño de la muestra en el caso de una prueba sobre diferencia de medias. ....	20
3. Determinación del tamaño de muestra en el diseño de un factor fijo con más de dos niveles. ....	22
3.1. Ejemplo de determinación del tamaño de muestra. ....	26
3.2. Determinación del tamaño de muestra en R.....	30
4. Determinación del tamaño de muestra en diseños de bloques completos al azar un factor varios niveles. ....	34
4.1. Ejemplo de determinación del tamaño de muestra en diseño de bloques completos al azar. ....	37
5. Determinación del tamaño de muestra en el diseño factorial de dos factores completamente al azar. ....	39
5.1. Ejemplo de determinación del tamaño de muestra en un diseño factorial con dos factores.....	41
Conclusiones. ....	43
Anexo. Uso de software para determinar el tamaño de muestra. ....	44
Aplicación de <i>G*Power</i> a la determinación del tamaño de muestra para un estudio de diferencia de medias con varianzas iguales. ....	44
Aplicación de <i>Piface</i> a la determinación del tamaño de muestra para un estudio de diferencia de medias con varianzas iguales. ....	47
Aplicación de <i>G*Power</i> a la determinación del tamaño de muestra en el caso de un factor fijo. ....	48
Aplicación de <i>Piface</i> para determinar el tamaño de muestra en un diseño factorial de dos factores completamente al azar. ....	51
Bibliografía .....	54

## Objetivos y delimitaciones.

El objetivo principal del presente trabajo es estudiar la determinación del tamaño de muestra en algunos de los diseños experimentales estudiados durante el curso de la especialización en estadística aplicada. Se buscó tener un enfoque sencillo y conceptual, sin requerir muchas herramientas matemáticas, para que el trabajo pudiera ser una referencia sencilla a personas de diversas formaciones con conocimientos de estadística. Se persiguió clarificar y transparentar la metodología que se encontró para la determinación del tamaño de muestra usando siempre software libre. Cabe aclarar que este texto de carácter monográfico no es un trabajo exhaustivo sobre el tema pues sólo se aborda la determinación del tamaño de muestra a través de la potencia<sup>\*</sup>. El documento no constituye una investigación de corte estrictamente aplicada, pues si bien sí se dan ejemplos, estos se han trabajado para hacer énfasis y aclarar la parte conceptual sobre cómo determinar el tamaño de muestra y no se ahonda más allá en las posibles aplicaciones.

En el anexo se muestra de manera superficial el uso de un par de programas especializados en el estudio de la potencia y determinación del tamaño de muestra, llamados *G\*Power*, de origen alemán y *Piface* de origen estadounidense. Ambos son de uso libre y constituyen herramientas muy interesantes que ofrecen la posibilidad de hacer análisis detallados, se invita a los lectores interesados a revisarlos.

---

<sup>\*</sup> Adcock (Adcock, 2008), realiza un análisis del tema más amplio y por ejemplo clasifica a los métodos para determinar el tamaño de muestra en frecuentistas y bayesianos.

## 1. Conceptos importantes relacionados a la determinación del tamaño de muestra.

### 1.1. Importancia de la determinación del tamaño de muestra en diseño de experimentos.

Según Ryan (Ryan, 2013), el diseño de experimentos es por lo general costoso y debido a ello ha habido gran interés en determinar el tamaño de muestra adecuado para ellos. Un experimento con un tamaño de muestra mayor al necesario se traduce en un derroche innecesario de recursos, mientras que un experimento con un tamaño de muestra debajo del adecuado corre el riesgo de no producir resultados útiles o concluyentes, lo cual finalmente también sería un derroche de recursos. En experimentos que involucran animales o personas, la determinación del tamaño adecuado de muestra también implica cuestiones éticas (Russell V., 2001). Además, el número de réplicas en un estudio de investigación afecta la precisión de las estimaciones de las medias de los tratamientos y la potencia de las pruebas estadísticas para detectar las diferencias entre las medias de los grupos en tratamiento (Kuehl, 2000). Todas estas razones hacen que el estudio del tamaño de la muestra en diseño de experimentos sea importante.

Como se verá más adelante, para determinar el tamaño de muestra en algunos diseños experimentales se utiliza la función potencia de una prueba de hipótesis derivada de la tabla de ANOVA. Este enfoque si bien no es el único, sí es uno de los más empleados (Russell V., 2001). En dicha prueba de hipótesis se involucra la distribución  $F$  y la probabilidad del error tipo II. Debido a esto, se comenzará describiendo las funciones de densidad relacionadas a la distribución  $F$  y posteriormente los conceptos relacionados a pruebas de hipótesis, de manera breve en ambos casos.

### 1.2. Algunas distribuciones de probabilidad.

*Distribución normal.*

La función de densidad normal con media  $\mu$  y varianza  $\sigma^2$  se denota como  $N(\mu, \sigma^2)$  y su función de densidad es:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

$$-\infty < x < \infty$$

$$-\infty < \mu < \infty$$

$$\sigma^2 > 0$$

Para indicar que una variable  $Z$  se distribuye de manera normal con media  $\mu$  y varianza  $\sigma^2$ , se escribe  $Z \sim N(\mu, \sigma^2)$ . A la distribución normal con media cero y varianza uno se le denomina distribución normal estándar y se le denota como  $N(0,1)$ .

#### *Distribución ji cuadrada.*

Si se define la variable aleatoria:

$$W = Z_1^2 + \dots + Z_k^2$$

En donde  $Z_i \sim N(\mu_i, \sigma^2)$  y cada variable  $Z_i$  es estadísticamente independiente de las otras, lo cual se expresa como  $Z_i \perp Z_j$  si  $i \neq j$ , entonces la variable  $W$  se distribuye ji cuadrada con  $k$  grados de libertad y parámetro de no centralidad  $\lambda = \sum_{i=1}^k \mu_i^2$ , lo cual se denota como  $\chi_{k,\lambda}^2$ . En el caso en el que  $Z_i \sim N(0,1) \forall i$ , la variable  $W$  se distribuye de manera ji cuadrada central o simplemente ji cuadrada, lo cual se escribe como  $\chi_k^2$ .

#### *Distribución t de Student.*

Si se define la variable aleatoria:

$$t = \frac{Z}{\sqrt{\frac{X}{n}}}$$

En donde  $Z \sim N(0,1)$ ,  $X \sim \chi_n^2$  y además  $Z \perp X$ , entonces se dice que la variable  $t$  se distribuye t de Student con  $n$  grados de libertad y se escribe  $t \sim T_n$ .

### *Distribución F de Fisher.*

Si se define la variable aleatoria:

$$F = \frac{\frac{W_1}{v_1}}{\frac{W_2}{v_2}}$$

En donde  $W_1 \sim \chi_{v_1}^2$  y  $W_2 \sim \chi_{v_2}^2$  y además  $W_1 \perp W_2$ , entonces la variable  $F$  tiene una distribución  $F$  de Fisher con  $v_1$  grados de libertad en el numerador y  $v_2$  grados de libertad en el denominador, lo cual se escribe como  $F_{v_1, v_2}$ . Si  $W_1 \sim \chi_{v_1, \lambda}^2$ , entonces la variable  $F$  tiene una distribución  $F$  de Fisher no central con  $v_1$  grados de libertad en el numerador y  $v_2$  grados de libertad en el denominador y parámetro de no centralidad  $\lambda$ , lo cual se denota por  $F_{v_1, v_2, \lambda}$ .

### 1.3. Conceptos relacionados a pruebas de hipótesis.

Con frecuencia, el objetivo de una investigación no es estimar un parámetro sino decidir en cuál de dos conjuntos  $\Theta_0$  o  $\Theta_1$  que no se traslapan es más asequible que el parámetro poblacional se encuentre. Una hipótesis estadística es una aseveración sobre el valor de un solo parámetro, sobre los valores de varios parámetros o sobre la forma de una distribución de probabilidad completa (Devore, 2008). La afirmación de que el parámetro se encuentra en el conjunto  $\Theta_0$  se llama hipótesis nula y suele denotarse por  $H_0$ , mientras que el hecho de asumir que se encuentra en  $\Theta_1$  se llama hipótesis alternativa y se suele escribir como  $H_1$ . En una prueba de hipótesis se define una regla para decidir si se acepta la hipótesis nula o se rechaza en favor de la hipótesis alternativa (Rincón, 2007).

La decisión de rechazar o no la hipótesis nula se basa en el valor de una variable  $T$  a la cual se le denomina estadístico de prueba y que es obtenida a partir de una muestra de la población; se espera que el estadístico de prueba  $T$  tenga una distribución conocida bajo el supuesto de que se cumple la hipótesis nula. Para tomar la decisión de rechazar o no la hipótesis nula  $H_0$ , se define un conjunto de valores  $C$ , dentro del conjunto de los posibles valores del estadístico  $T$ , tal que si

el valor del estadístico de prueba  $T$  se encuentra en  $C$ , entonces se decide en favor de  $H_1$  y se rechaza  $H_0$ ; de manera contraria, si el valor del estadístico de prueba no está en  $C$ , entonces no se rechazará  $H_0$  y se rechazará  $H_1$ . Al conjunto  $C$  se le denomina región de rechazo o región crítica de la prueba. Una vez que se tiene especificado el estadístico de prueba y la región de rechazo, la prueba de hipótesis está completamente especificada.

A manera de ejemplo sencillo y sin entrar en muchos detalles, supóngase que se está estudiando el peso promedio de las lagartijas de cierta región, se quiere probar la hipótesis de que pesan en promedio 50 gramos y se toma el peso en una muestra de treinta lagartijas resultando su peso promedio de  $\bar{X} = 75$  gramos con una desviación estándar de  $s = 10$  gramos. En este caso la hipótesis nula sería  $H_0: \mu = 50$ , mientras que la hipótesis alternativa sería  $H_1: \mu \neq 50$ ; el estadístico de prueba es el peso promedio muestral  $\bar{X}$ , el cual se sabe que se distribuye como t de Student luego de aplicarle la operación  $(\bar{X} - \mu)/(s/\sqrt{n})$ . Si se define como región crítica al conjunto  $C = \{\bar{X} > 60\}$ , entonces en este caso el valor del estadístico de prueba se encontraría en la región crítica y por ende se rechazaría  $H_0$ .

Cabe mencionar que el rechazo de la hipótesis nula es una conclusión fuerte de que  $\theta_0$  no parece ser consistente con los datos observados en la muestra, mientras que no rechazar  $H_0$  es una conclusión débil que debe ser interpretada como evidencia de que  $\theta_0$  sí fue consistente con los datos obtenidos de la muestra (Ross, 2010). Al realizar una prueba de hipótesis, hay cuatro posibles resultados, los cuales se resumen en la Tabla 1 (Walpole, Myers, Myers, & Keying, 2007).



Tabla 1. Situaciones posibles al probar una hipótesis estadística.

		Realidad	
		$H_0$ es verdadera	$H_0$ es falsa
Decisión estadística	No rechazar $H_0$	Decisión correcta	Error tipo II
	Rechazar $H_0$	Error tipo I	Decisión correcta

A la probabilidad de ocurrencia del error tipo I se le denomina tamaño de la prueba y se le denota por  $\alpha^\dagger$ ; en este sentido, debe entenderse a  $\alpha$  como la probabilidad de rechazar la hipótesis nula cuando esta es cierta. Mientras que, a la probabilidad de cometer el error tipo II, es decir de no rechazar la hipótesis nula cuando esta es falsa, se le denota por  $\beta$ . Entonces  $1 - \beta$  corresponde a la probabilidad de rechazar correctamente la hipótesis nula y también se le denomina potencia de la prueba. Considerando la región de rechazo  $C$ , su complemento  $C^c$  y el valor del estadístico de prueba  $t$ , el valor de  $\alpha$  y  $\beta$  pueden escribirse como:

$$\alpha = P(t \in C | H_0 \text{ cierta})$$

$$\beta = P(t \in C^c | H_1 \text{ cierta})$$

Para ejemplificar los conceptos anteriores, considérese una prueba de hipótesis en donde la distribución de la estadística de prueba bajo la hipótesis nula  $H_0$  es una distribución  $F$  central, mientras que la distribución de la estadística de prueba bajo la hipótesis alternativa  $H_1$  es una distribución  $F$  no central (Giesbrecht & Gumpertz, 2004), como se muestra en la Figura 1. En la Figura 1 a manera de ejemplo el punto de corte es el valor  $t = 2$ , es decir, a partir de él se

---

<sup>†</sup> De manera más precisa, si  $C$  es la región de rechazo y se desean probar las hipótesis  $H_0: \theta \in \Theta_0$  vs  $H_1: \theta \in \Theta_1$ , el tamaño de la prueba se define como:

$$\sup_{\theta \in \Theta_0} P(t \in C | \theta) = \alpha$$

En el caso de que las hipótesis sean del tipo  $H_0: \theta = a$  vs  $H_1: \theta = b$ , lo cual se define como hipótesis simples, el tamaño de la muestra corresponde a:

$$P(t \in C | H_0 \text{ es cierta}) = \alpha$$

Este último caso es el que se está abordando en la descripción.

decidirá en favor o en contra de la hipótesis nula; un valor mayor a 2 del estadístico de prueba conllevará a rechazar  $H_0$ , mientras que un valor más pequeño hará que no se rechace  $H_0$ . Entonces el área a la derecha de  $t = 2$  y bajo la distribución  $F$  central representa la probabilidad de rechazar  $H_0$  cuando ésta es cierta, es decir corresponde a  $\alpha$ , la probabilidad del error tipo I. Mientras que el área a la izquierda de  $t = 2$  y debajo de la distribución  $F$  no central representa la probabilidad de elegir  $H_0$  siendo  $H_1$  la distribución correcta, es decir representa  $\beta$ , la probabilidad de cometer el error tipo II.

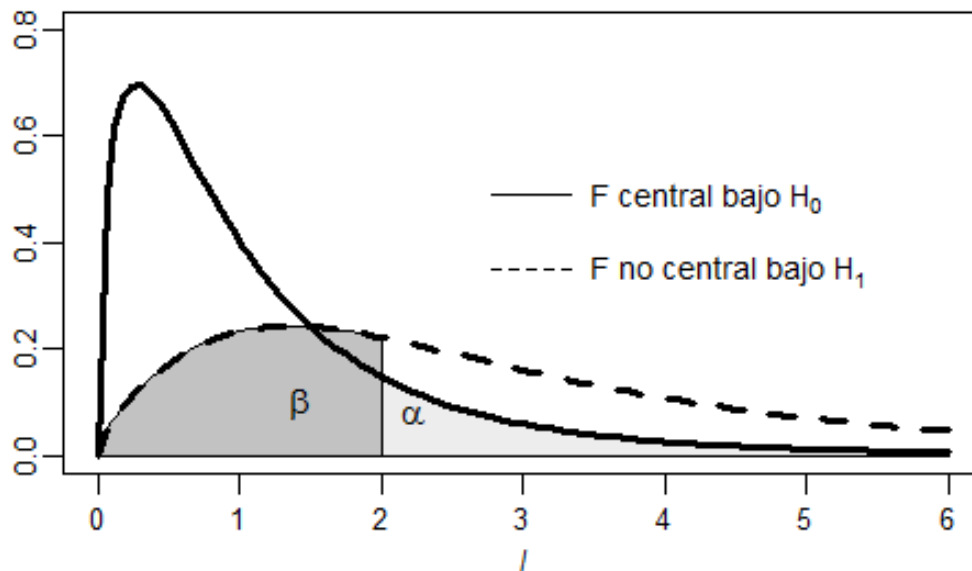


Figura 1. Prueba de hipótesis con distribución  $F$  y errores tipo I y II.

## 2. Determinación del tamaño de muestra en un diseño con un factor fijo y dos niveles (diferencia de medias).

En los problemas de investigación experimental es frecuente que sea necesario comparar dos tratamientos diferentes, es decir, dos condiciones diferentes del proceso, como por ejemplo dos tipo de dietas, dos formas de cultivar (Gutiérrez Pulido & Vara Salazar, 2012). En estos casos, se extraen dos muestras aleatorias independientes, del mismo tamaño, y se comparan sus medias dando lugar a la siguiente prueba de hipótesis:

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

En donde:

- $\mu_1$  es la media de la población 1.
- $\mu_2$  es la media de la población 2.

Suponiendo que las muestras provienen de poblaciones que siguen una distribución normal, pueden encontrarse fórmulas y procedimientos para determinar el tamaño de muestra en caso de conocer las varianzas  $\sigma_1^2$  y  $\sigma_2^2$ , de desconocerlas pero poderlas asumir iguales y cuando las varianzas  $\sigma_1^2$  y  $\sigma_2^2$  son desconocidas y distintas. Por sencillez se abordarán a continuación los dos primeros casos; puede encontrarse en el libro de Desu (Desu & Raghavarao, 1990) una propuesta para determinar el tamaño de muestra en el tercer caso.

### 2.1. Caso en el que se conocen las varianzas.

Cuando se conoce la varianza de ambas poblaciones, la distribución muestral de  $\bar{X}_1 - \bar{X}_2$  se distribuye de manera normal con error estándar  $\sqrt{(\sigma_1^2 + \sigma_2^2)/n}$  (Johnson & Kuby, 2012) y puede encontrarse una expresión para determinar el tamaño de muestra (Walpole, Myers, Myers, & Keying, 2007):

$$n = \frac{(z_{\alpha/2} + z_{\beta})^2 (\sigma_1^2 + \sigma_2^2)}{\delta^2} \dots (2.1)$$

En donde:

- $\alpha$ : es el nivel de significancia de la prueba.
- $\beta$ : es la probabilidad del error tipo II
- $z_{\alpha/2}$ : es el cuantil superior  $\alpha/2$  de la distribución normal estándar.
- $z_{\beta}$ : es el cuantil superior  $\beta$  de la distribución normal estándar.
- $\sigma_i^2$ : es la varianza de la población  $i = 1,2$ .
- $\delta$ : es la diferencia entre las medias que se desea detectar.

Una posible deducción de la fórmula (2.1) para el caso de una hipótesis alternativa específica  $\mu_1 - \mu_2 = \delta$  puede facilitarse con el apoyo de la Figura 2 (Walpole, Myers, Myers, & Keying, 2007).

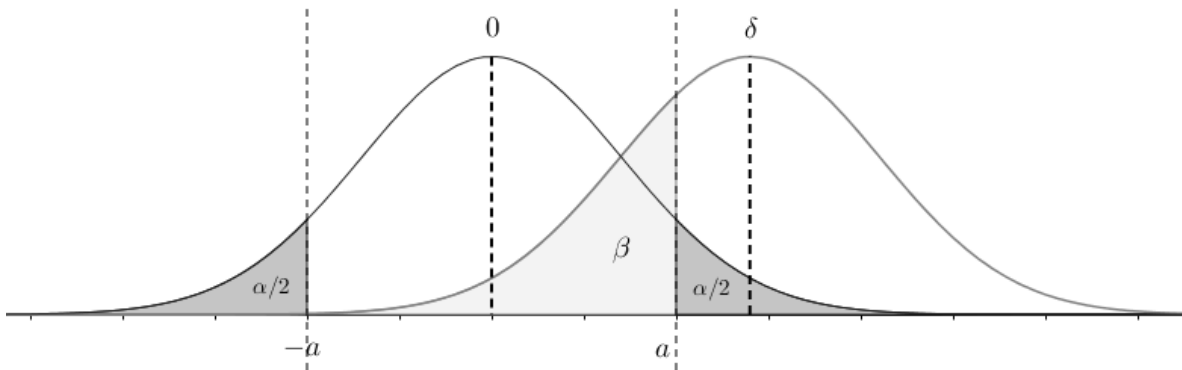


Figura 2. Errores tipo I y II en la prueba en la que  $H_0: \mu_1 - \mu_2 = 0$  vs  $H_1: \mu_1 - \mu_2 = \delta$ .

En este caso las dos distribuciones son normales y puede escribirse la probabilidad del error tipo II como:

$$\beta = P\left(-a < \bar{X}_1 - \bar{X}_2 < a \mid \bar{X}_1 - \bar{X}_2 \sim N\left(\delta, \frac{\sigma_1^2 + \sigma_2^2}{n}\right)\right)$$

$$\beta = P\left(\frac{-a - \delta}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}}} < Z < \frac{a - \delta}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}}} \mid \bar{X}_1 - \bar{X}_2 \sim N\left(\delta, \frac{\sigma_1^2 + \sigma_2^2}{n}\right)\right)$$

Considerando además que

$$-z_{\alpha/2} = \frac{-a}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}}} \quad y \quad z_{\alpha/2} = \frac{a}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}}}$$

Puede escribirse:

$$\beta = P\left(-z_{\alpha/2} - \frac{\delta}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}}} < Z < z_{\alpha/2} - \frac{\delta}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}}} \mid \bar{X}_1 - \bar{X}_2 \sim N\left(\delta, \frac{\sigma_1^2 + \sigma_2^2}{n}\right)\right)$$

Al considerar la situación mostrada en la Figura 2, puede observarse la siguiente relación entre el cuantil  $z_\beta$  de la distribución normal con media  $\delta$  y la parte derecha de la desigualdad en la expresión anterior que involucra al cuantil  $z_{\alpha/2}$ :

$$-z_\beta \approx z_{\alpha/2} - \frac{\delta}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}}}$$

Y finalmente al despejarse  $n$  de la última expresión puede obtenerse la fórmula mostrada en (2.1):

$$\frac{\delta^2}{\frac{\sigma_1^2 + \sigma_2^2}{n}} = (z_{\alpha/2} + z_\beta)^2$$

$$n = \frac{(z_{\alpha/2} + z_\beta)^2 (\sigma_1^2 + \sigma_2^2)}{\delta^2} \dots (2.1)$$

La expresión (2.1) también es útil para hacer explícita la relación del tamaño de muestra con el valor de  $\alpha$ , de  $\beta$ , la varianza  $\sigma_i^2$  y la diferencia en los efectos a detectar  $\delta$ . A través de la expresión (2.1) puede observarse que el tamaño de

muestra resulta ser proporcional a  $z_{\alpha/2}$ ,  $z_{\beta}$  y a las varianzas  $\sigma_i^2$ . Por otra parte, como a valores mayores de  $z_{\alpha/2}$  y  $z_{\beta}$  corresponden menores valores de  $\alpha$  y  $\beta$ , puede observarse que el tamaño de muestra resulta ser inversamente proporcional a la diferencia  $\delta$  que se desea detectar, a la probabilidad del error tipo I (el valor de  $\alpha$ ) y a la probabilidad del error tipo II (el valor de  $\beta$ ). De la última observación puede desprenderse que como el tamaño de muestra es inversamente proporcional al valor de  $\beta$  (disminuye cuando el valor de  $\beta$  aumenta pues  $z_{\beta}$  disminuye y viceversa), entonces el tamaño de muestra resultará directamente proporcional al valor de  $1 - \beta$ , es decir de la potencia. Estas relaciones siguen manteniéndose para otros diseños experimentales como se verá más adelante.

Para hacer más claras las relaciones descritas, en la Figura 3 se grafica el tamaño de muestra necesario para un valor de  $\alpha$  y  $\beta$  fijos; puede observarse cómo aumenta el tamaño de muestra cuando las varianzas aumentan o disminuye el valor de la diferencia a detectar. Por otra parte, en la Figura 4 se mantienen fijos los valores de  $\alpha$  y la suma de las varianzas  $\sum_{i=1}^2 \sigma_i^2$  mientras que varía el valor de la diferencia a detectar y de  $\beta$ ; puede observarse que mientras más grande es el valor de  $\beta$ , más chico es el valor de la potencia  $1 - \beta$  y también que para un valor determinado de  $\delta^2$ , a mayor valor de la potencia se requiere mayor tamaño de muestra.

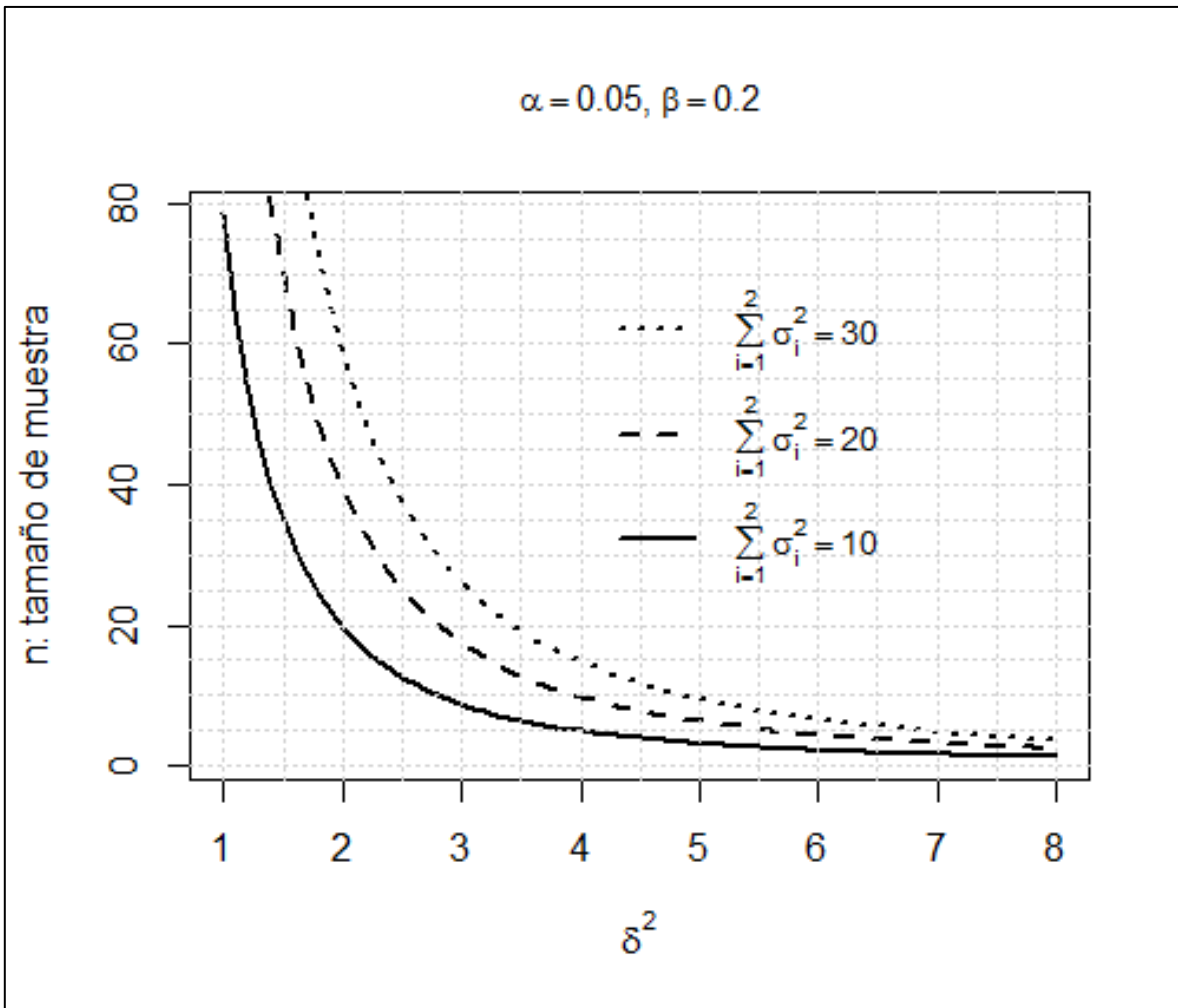


Figura 3. Tamaño de muestra en un diseño de diferencia de medias cuando varían la suma de varianzas y la diferencia a detectar.

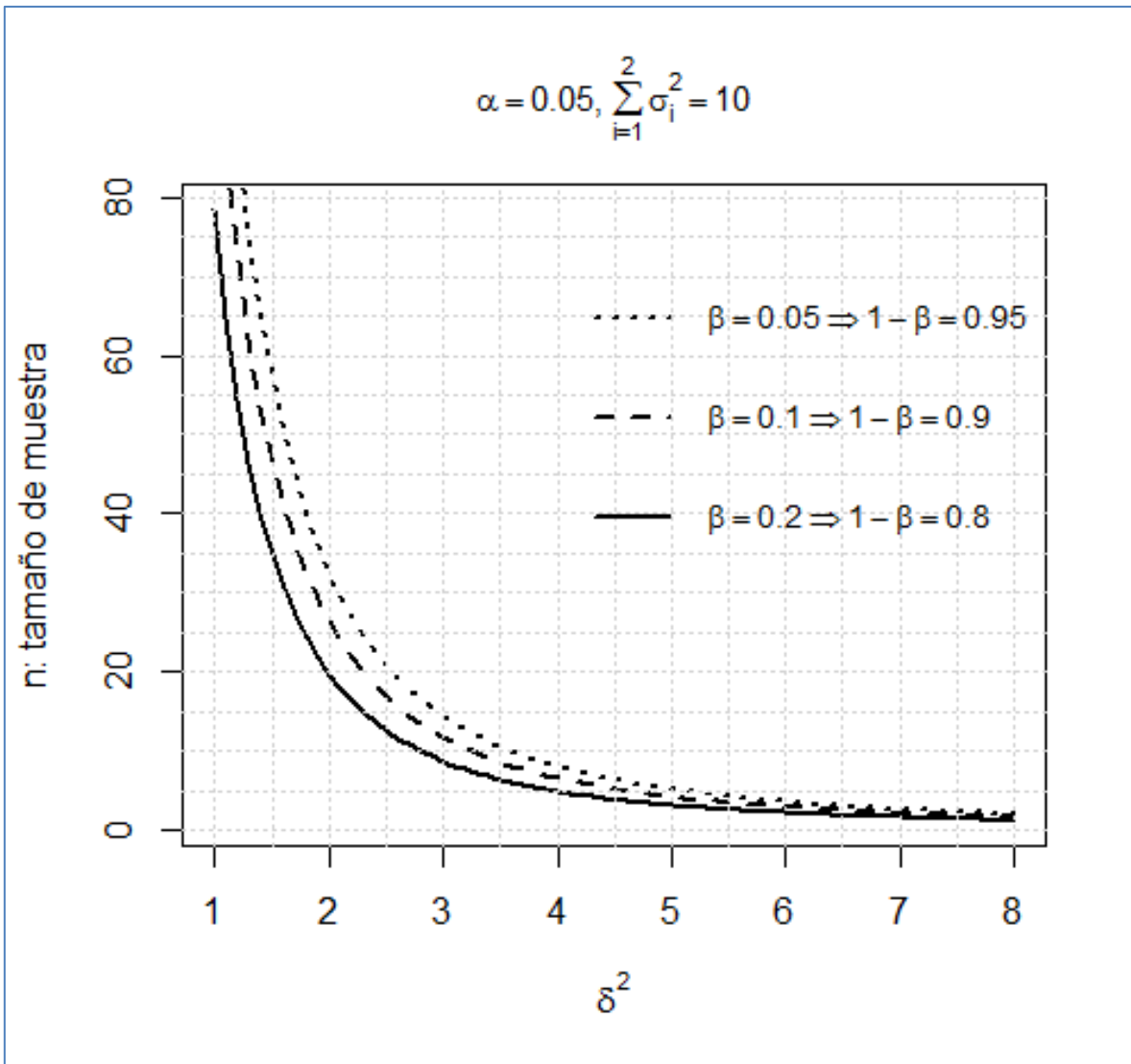


Figura 4. Tamaño de muestra en un diseño de diferencia de medias cuando varían  $\beta$ , la potencia y la diferencia a detectar.

### 2.1.1. Determinación del tamaño de muestra en R.

La fórmula para determinar el tamaño de muestra en el caso de que las varianzas sean conocidas, es decir la expresión (2.1), no presenta ninguna dificultad y una posible implementación de ella en R se muestra en la Figura 5.



```

difmed1<-function(alfa,beta,sigma1,sigma2,delta){
  za<-qnorm(alfa/2,lower.tail = F)
  zb<-qnorm(beta,lower.tail = F)
  n<-((za+zb)^2*(sigma1+sigma2))/delta^2
  return(ceiling(n))
}

```

Figura 5. Implementación de la expresión (2.1) para determinar el tamaño de muestra.

### 2.1.2. Ejemplo de determinación del tamaño de muestra.

*Ejemplo 1.* Se va a realizar un experimento en donde se van a comparar las medias de dos poblaciones independientes y se desea determinar el tamaño de muestra adecuado para tener  $\alpha = \beta = 0.05$  ¿Cuál sería el tamaño de muestra en caso de que las varianzas de las poblaciones fueran  $\sigma_1^2 = 1$ ,  $\sigma_2^2 = 1.2$  y se deseara lograr detectar una diferencia de 2 unidades?

Se recurre a la expresión (2.1):

$$n = \frac{(z_{\alpha/2} + z_{\beta})^2 (\sigma_1^2 + \sigma_2^2)}{\delta^2}$$

Con lo cual, el tamaño de muestra necesario para detectar una diferencia de 2 unidades con una potencia de 95% es:

$$n = \frac{(1.96 + 1.65)^2 (1 + 1.2)}{2^2} = 7.16 \therefore n = 8$$

Este mismo resultado se obtiene al llamar a la función mostrada en la Figura 5 a través del comando en R `difmed1(.05,.05,1,1.2,2)`.

## 2.2. Caso en el que se desconocen las varianzas pero pueden suponerse iguales.

En caso que se desconozca la varianza pero se pueda suponer igual en ambas poblaciones, puede encontrarse la siguiente expresión para el tamaño de la muestra (Desu & Raghavarao, 1990)<sup>‡</sup>:

$$n = \frac{\nu}{2} + 1 = 2 \left( \frac{\sigma}{\delta} \right)^2 \{t_{\alpha}(\nu) + t_{\beta}(\nu)\}^2 \dots (2.2)$$

En donde:

- $\alpha$ : es el nivel de significancia de la prueba.
- $\beta$ : es la probabilidad del error tipo II
- $t_{\alpha}(\nu)$ : es el cuantil superior  $\alpha$  de una variable t de Student con  $\nu$  grados de libertad.
- $t_{\beta}(\nu)$ : es el cuantil superior  $\beta$  de una variable t de Student con  $\nu$  grados de libertad.
- $\sigma$ : es la desviación estándar de ambas poblaciones.
- $\delta$ : es la diferencia entre las medias que se desea detectar.

A continuación se proporciona una implementación en R de la expresión (2.2), así como un ejemplo de la determinación del tamaño de muestra para este diseño experimental.

### 2.2.1. Determinación del tamaño de muestra en R.

La dificultad al trabajar con la expresión (2.2) está en la determinación de los grados de libertad adecuados. A partir del ejemplo proporcionado por Desu sobre el manejo de la expresión (2.2) (Desu & Raghavarao, 1990), puede deducirse el algoritmo mostrado en la Figura 6 para determinar el tamaño de muestra en el diseño en el cual se comparan las medias de dos muestras independientes.

---

<sup>‡</sup> En este texto se menciona que esta expresión corresponde a la determinación del tamaño de muestra cuando se desea realizar una prueba de hipótesis de una cola y que en caso de que la prueba sea de dos colas debe reemplazarse  $t_{\alpha}$  por  $t_{\alpha/2}$ .

(1) En la expresión:

$$v = 2 \left[ 2 \left( \frac{\sigma}{\delta} \right)^2 \{t_{\alpha}(v) + t_{\beta}(v)\}^2 - 1 \right] \dots (A)$$

Reemplazar  $t_{\alpha}(v)$  por  $z_{1-\alpha}$  y  $t_{\beta}(v)$  por  $z_{1-\beta}$  y obtener una estimación inicial de los grados de libertad  $v_{ini}$ .

(2) Con el valor  $v_{ini}$  calcular otro valor de los grados de libertad  $v_{new}$  con la expresión (A).

(3) Si  $v_{ini} < v_{new}$  hacer  $v_{ini} = v_{new}$  y repetir (2). En caso contrario parar.

(4) Obtener el tamaño de muestra con la expresión (2.2).

Figura 6. Algoritmo para encontrar el tamaño de muestra para un experimento de diferencia de medias en el caso de desconocer la varianza pero suponerla igual.

Si se tienen estimaciones para la desviación estándar  $\sigma$  o si se puede suponer que la diferencia entre las medias  $\delta$  es un múltiplo de  $\sigma$ , puede definirse el cociente  $D = |\delta| \div \sigma$  para el cual existe una tabla para la determinación del tamaño de muestra, que se muestra en la Figura 7 y la Figura 8 (Davies, 1960).

Table E.1

NUMBER OF OBSERVATIONS FOR *t*-TEST OF DIFFERENCE BETWEEN TWO MEANS

The entries in this table show the number of observations needed in a *t*-test of the significance of the difference between two means in order to control the probabilities of the errors of the first and second kinds at  $\alpha$  and  $\beta$  respectively.

Single-sided test Double-sided test		Level of <i>t</i> -test																							
		0.01					0.02					0.05					0.1								
		$\alpha = 0.005$ $\alpha = 0.01$					$\alpha = 0.01$ $\alpha = 0.02$					$\alpha = 0.025$ $\alpha = 0.05$					$\alpha = 0.05$ $\alpha = 0.1$								
		0.01	0.05	0.1	0.2	0.5	0.01	0.05	0.1	0.2	0.5	0.01	0.05	0.1	0.2	0.5	0.01	0.05	0.1	0.2	0.5				
Value of $D = \frac{\delta}{\sigma}$	0.05																				0.05				
	0.10																				0.10				
	0.15																				0.15				
	0.20																			137	0.20				
	0.25														124					88	0.25				
	0.30																				61	0.30			
	0.35																			102	45	0.35			
	0.40																			108	78	35	0.40		
	0.45																			108	86	62	28	0.45	
	0.50																			88	70	51	23	0.50	
	0.55																			112	73	58	42	19	0.55
	0.60																			89	61	49	36	16	0.60
	0.65																			76	52	42	30	14	0.65
	0.70																			66	45	36	26	12	0.70
	0.75																			57	40	32	23	11	0.75
	0.80																			50	35	28	21	10	0.80
	0.85																			45	31	25	18	9	0.85
	0.90																			40	28	22	16	8	0.90
	0.95																			36	25	20	15	7	0.95
	1.00																			33	23	18	14	7	1.00

TABLES

609

Figura 7. Tamaño de muestra para una prueba *t* de diferencia de dos medias con las desviaciones estándar desconocidas pero iguales (Davies, 1960).

Table E-1 (continued)

NUMBER OF OBSERVATIONS FOR *t*-TEST OF DIFFERENCE BETWEEN TWO MEANS

The entries in this table show the number of observations needed in a *t*-test of the significance of the difference between two means in order to control the probabilities of the errors of the first and second kinds at  $\alpha$  and  $\beta$  respectively.

610

Single-sided test Double-sided test  $\beta =$		Level of <i>t</i> -test																				
		0.01					0.02					0.05					0.1					
		$\alpha = 0.005$ $\alpha = 0.01$					$\alpha = 0.01$ $\alpha = 0.02$					$\alpha = 0.025$ $\alpha = 0.05$					$\alpha = 0.05$ $\alpha = 0.1$					
		0.01	0.05	0.1	0.2	0.5	0.01	0.05	0.1	0.2	0.5	0.01	0.05	0.1	0.2	0.5	0.01	0.05	0.1	0.2	0.5	
Value of $D = \frac{\delta}{\sigma}$	1.1	42	32	27	22	13	38	28	23	19	11	32	23	19	14	8	27	19	15	12	6	1.1
	1.2	36	27	23	18	11	32	24	20	16	9	27	20	16	12	7	23	16	13	10	5	1.2
	1.3	31	23	20	16	10	28	21	17	14	8	23	17	14	11	6	20	14	11	9	5	1.3
	1.4	27	20	17	14	9	24	18	15	12	8	20	15	12	10	6	17	12	10	8	4	1.4
	1.5	24	18	15	13	8	21	16	14	11	7	18	13	11	9	5	15	11	9	7	4	1.5
	1.6	21	16	14	11	7	19	14	12	10	6	16	12	10	8	5	14	10	8	6	4	1.6
	1.7	19	15	13	10	7	17	13	11	9	6	14	11	9	7	4	12	9	7	6	3	1.7
	1.8	17	13	11	10	6	15	12	10	8	5	13	10	8	6	4	11	8	7	5	1.8	
	1.9	16	12	11	9	6	14	11	9	8	5	12	9	7	6	4	10	7	6	5	1.9	
	2.0	14	11	10	8	6	13	10	9	7	5	11	8	7	6	4	9	7	6	4	2.0	
	2.1	13	10	9	8	5	12	9	8	7	5	10	8	6	5	3	8	6	5	4	2.1	
	2.2	12	10	8	7	5	11	9	7	6	4	9	7	6	5	8	6	5	4	2.2		
	2.3	11	9	8	7	5	10	8	7	6	4	9	7	6	5	7	5	5	4	2.3		
	2.4	11	9	8	6	5	10	8	7	6	4	8	6	5	4	7	5	4	4	2.4		
	2.5	10	8	7	6	4	9	7	6	5	4	8	6	5	4	6	5	4	3	2.5		
	3.0	8	6	6	5	4	7	6	5	4	3	6	5	4	4	5	4	3		3.0		
3.5	6	5	5	4	3	6	5	4	4		5	4	4	3	4	3			3.5			
4.0	6	5	4	4		5	4	4	3		4	4	3		4				4.0			

DESIGN AND ANALYSIS

Figura 8. Tamaño de muestra para una prueba *t* de diferencia de dos medias con las desviaciones estándar desconocidas pero iguales (Davies, 1960).

La tabla de la Figura 7 y la Figura 8 puede replicarse implementando en R el algoritmo mostrado en la Figura 6; una posible implementación se muestra en la Figura 9. Una parte de los resultados obtenidos de dicha implementación se presenta en la Figura 10 en donde se contrastan los resultados obtenidos en la tabla con los del algoritmo de la Figura 9; puede observarse que se presentan ligeras diferencias, las cuales pueden atribuirse al método de redondeo.

```
difmed2<-function(alpha,beta,cociente){
  vn=100;cont=1
  v=ceiling(2*(2*cociente^2*(qnorm(alpha,lower.tail = F)
    +qnorm(beta,lower.tail = F))^2-1))
  if(v==0) return(NA)
  vi=ceiling(2*(2*cociente^2*(qt(alpha,v,lower.tail = F)
    +qt(beta,v,lower.tail = F))^2-1))
  while(vi<vn){#se quiere comparar por lo menos una vez con los siguientes
    grados de libertad
    if(cont!=1) vi=vn#sólo la primer vez vi se queda igual
    vn=ceiling(2*(2*cociente^2*(qt(alpha,vi,lower.tail = F)
      +qt(beta,vi,lower.tail = F))^2-1))
    cont=cont+1
  }
  return(ceiling(vi/2+1))
}
```

Figura 9. Implementación para determinar el tamaño de muestra de un experimento sobre diferencia de medias en R.

		$\alpha = 0.05$									
		Resultados mostrados por Davies					Resultados obtenidos con el algoritmo de la figura 9				
$\beta =$		0.01	0.05	0.1	0.2	0.5	0.01	0.05	0.1	0.2	0.5
Valor de $D = \frac{\delta}{\sigma}$	0.55	112	73	58	42	19	106	73	58	42	19
	0.6	89	61	49	36	16	89	62	49	36	17
	0.65	76	52	42	30	14	77	53	42	31	14
	0.7	66	45	36	26	12	66	46	36	27	13
	0.75	57	40	32	23	11	58	40	32	23	11
	0.8	50	35	28	21	10	51	35	28	21	10
	0.85	45	31	25	18	9	46	31	25	18	9
	0.9	40	28	22	16	8	41	28	23	17	8
	0.95	36	25	20	15	7	37	25	20	15	8
	1	33	23	18	14	7	33	23	19	14	7

Figura 10. Comparación de la tabla proporcionada por Davies (Davies, 1960) y los resultados obtenidos mediante el algoritmo de la Figura 9.

### 2.2.2. Ejemplo de determinación del tamaño de la muestra en el caso de una prueba sobre diferencia de medias.

*Ejemplo.* Se va a realizar un experimento en donde se comparen las medias de dos poblaciones independientes y se desea determinar el tamaño de muestra adecuado para tener  $\alpha = \beta = 0.05$  llevando a cabo una prueba de hipótesis de una cola ¿Cuál sería el tamaño de muestra en caso de que se desconocieran las varianzas pero se tuviera evidencia razonable de que son iguales y se deseara detectar una diferencia de dos desviaciones estándar?

Usando el algoritmo de la Figura 6 se obtienen las operaciones que se muestran a continuación para determinar el tamaño de muestra:

**Paso 1**

$$v_{ini} = 2 \left[ 2 \left( \frac{\sigma}{2\sigma} \right)^2 \{1.65 + 1.65\}^2 - 1 \right] = 8.89 \therefore v_{ini} = 9$$

**Paso 2**

$$v_{new} = 2 \left[ 2 \left( \frac{\sigma}{2\sigma} \right)^2 \{t_{\alpha}(9) + t_{\beta}(9)\}^2 - 1 \right] = 2 \left[ 2 \left( \frac{1}{2} \right)^2 \{1.83 + 1.83\}^2 - 1 \right] \therefore v_{new} = 12$$

**Paso 3**

Como  $v_{ini} < v_{new}$ , entonces  $v_{ini} = 12$ , además:

$$v_{new} = 2 \left[ 2 \left( \frac{\sigma}{2\sigma} \right)^2 \{t_{\alpha}(12) + t_{\beta}(12)\}^2 - 1 \right] = 2 \left[ 2 \left( \frac{1}{2} \right)^2 \{1.78 + 1.78\}^2 - 1 \right]$$
$$\therefore v_{new} = 11$$

Como  $v_{ini} > v_{new}$ , se para y se toma  $v = \max(v_{new}, v_{ini}) = 12$ .

**Paso 4**

$$\therefore n = \frac{v}{2} + 1 = \frac{12}{2} + 1 = 7$$

Otra alternativa para determinar el tamaño de muestra es usar la tabla de la Figura 8. Con estos fines se calcula el cociente  $D = |\delta| \div \sigma$  y se encuentra en este caso que  $D = \frac{2\sigma}{\sigma} = 2$ , con lo cual al situarse en el renglón donde  $D = 2$  y las columnas donde  $\alpha = \beta = 0.05$  en la Figura 8, puede determinarse que el tamaño de muestra necesario sería:

$$n = 7$$



### 3. Determinación del tamaño de muestra en el diseño de un factor fijo con más de dos niveles.

En el caso de un diseño de un factor fijo con  $a$  niveles y  $r$  repeticiones por nivel, el modelo estadístico que describe al diseño es:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, \dots, a \\ j = 1, \dots, r \end{cases}$$

Al cual se le conoce como modelo de efectos fijos, en donde:

- $y_{ij}$  es la respuesta observada en la unidad experimental en la repetición  $j$  y bajo el nivel  $i$  del tratamiento.
- $\mu$  es la media general de todas las observaciones.
- $\tau_i$  es el efecto del tratamiento  $i$ .
- $\varepsilon_{ij}$  es el error experimental bajo la repetición  $j$  y bajo el nivel  $i$  del tratamiento,  $\varepsilon_{ij} \sim N(0, \sigma^2)$  y los errores  $\varepsilon_{ij}$  son independientes.

Cabe mencionar que en este modelo, considerando  $\mu_i$ , la media del tratamiento  $i$ , puede establecerse la siguiente relación entre el efecto en el tratamiento  $i$  y la media en el tratamiento  $i$ :

$$\mu_i = \mu + \tau_i$$

Con lo cual, el modelo estadístico en este caso podría escribirse como:

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

$$i = 1, \dots, a; j = 1, \dots, r$$

Modelo al cual se le conoce como modelo de medias. Por otra parte, en este caso la tabla ANOVA se define como (Montgomery, 2004):

Tabla 2. Tabla ANOVA para un diseño de un factor fijo.

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	$F_0$
Entre los tratamientos	$SC_{trat} = r \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2$	$a - 1$	$CM_{trat} = \frac{SC_{trat}}{a - 1}$	$\frac{CM_{trat}}{CM_E}$
Error (dentro de los tratamientos)	$SC_E = \sum_{i=1}^a \sum_{j=1}^r (y_{ij} - \bar{y}_{i.})^2$	$ra - a$	$CM_E = \frac{SC_E}{ra - a}$	
Total	$SC_T = \sum_{i=1}^a \sum_{j=1}^r (y_{ij} - \bar{y}_{..})^2$	$ra - 1$		

Puede demostrarse además que  $SC_E/(ra - a) \sim \chi_{ra-a}^2$ , mientras que si  $\tau_1 = \dots = \tau_a = 0$ , entonces  $SC_{trat}/(a - 1) \sim \chi_{a-1}^2$ . Esto da pie al planteamiento de la siguiente prueba de hipótesis:

$H_0: \tau_1 = \dots = \tau_a = 0^{\S}$ $H_1: \tau_i \neq 0$ para al menos un $i$ .
--

En caso de cumplirse  $H_0$ , puede demostrarse que el estadístico  $F_0 = CM_{trat} \div CM_E$  tiene una distribución  $F$  central con  $v_1 = a - 1$  grados de libertad en el numerador y  $v_2 = ra - a$  grados de libertad en el denominador. Mientras que bajo la hipótesis alternativa  $H_1$ ,  $F_0$  tiene una distribución  $F$  no central con parámetro de desplazamiento  $\lambda = r \sum_{i=1}^a \tau_i^2 / \sigma^2$  (Kuehl, 2000) en donde  $r$  es el número de repeticiones en cada nivel, es decir el tamaño de muestra.

En este caso, la probabilidad de cometer el error tipo II es:

$$\beta = P(F_0 \leq F_{a-1, ra-a}^{1-\alpha} | H_0 \text{ falsa})$$

Considerando además que:

<sup>§</sup> Siguiendo el modelo de medias, esta hipótesis puede escribirse como  $H_0: \mu_1 = \dots = \mu_a = \mu$

$$P(F_0 \leq F_{a-1,ra-a}^{1-\alpha} | H_0 \text{ falsa}) + P(F_0 > F_{a-1,ra-a}^{1-\alpha} | H_0 \text{ falsa}) = 1$$

$$\beta + P(F_0 > F_{a-1,ra-a}^{1-\alpha} | H_0 \text{ falsa}) = 1$$

Entonces:

$$P(F_0 > F_{a-1,ra-a}^{1-\alpha} | H_0 \text{ falsa}) = 1 - \beta$$

Finalmente considerando que cuando  $H_0$  es falsa entonces  $F_0 \sim F_{a-1,ra-a,\lambda}$ , la potencia de la prueba se puede escribir como:

$$1 - \beta = P(F_0 > F_{a-1,ra-a}^{1-\alpha} | F_0 \sim F_{a-1,ra-a,\lambda}) \dots (3.1)$$

La función potencia involucra el tamaño de muestra en el parámetro de no centralidad  $\lambda$ , así que es factible obtener dicho tamaño a través de esta función. Sin embargo, debido a que especificar un tamaño de muestra puede conllevar a una potencia baja, para determinar el tamaño de muestra se fija más bien la potencia de la prueba, el nivel de significancia y una función  $\Phi$  del parámetro de no centralidad definido como  $\Phi = \sqrt{\lambda/a}$  y se construyen *curvas operativas*. Un ejemplo de una curva operativa se muestra en la Figura 11.

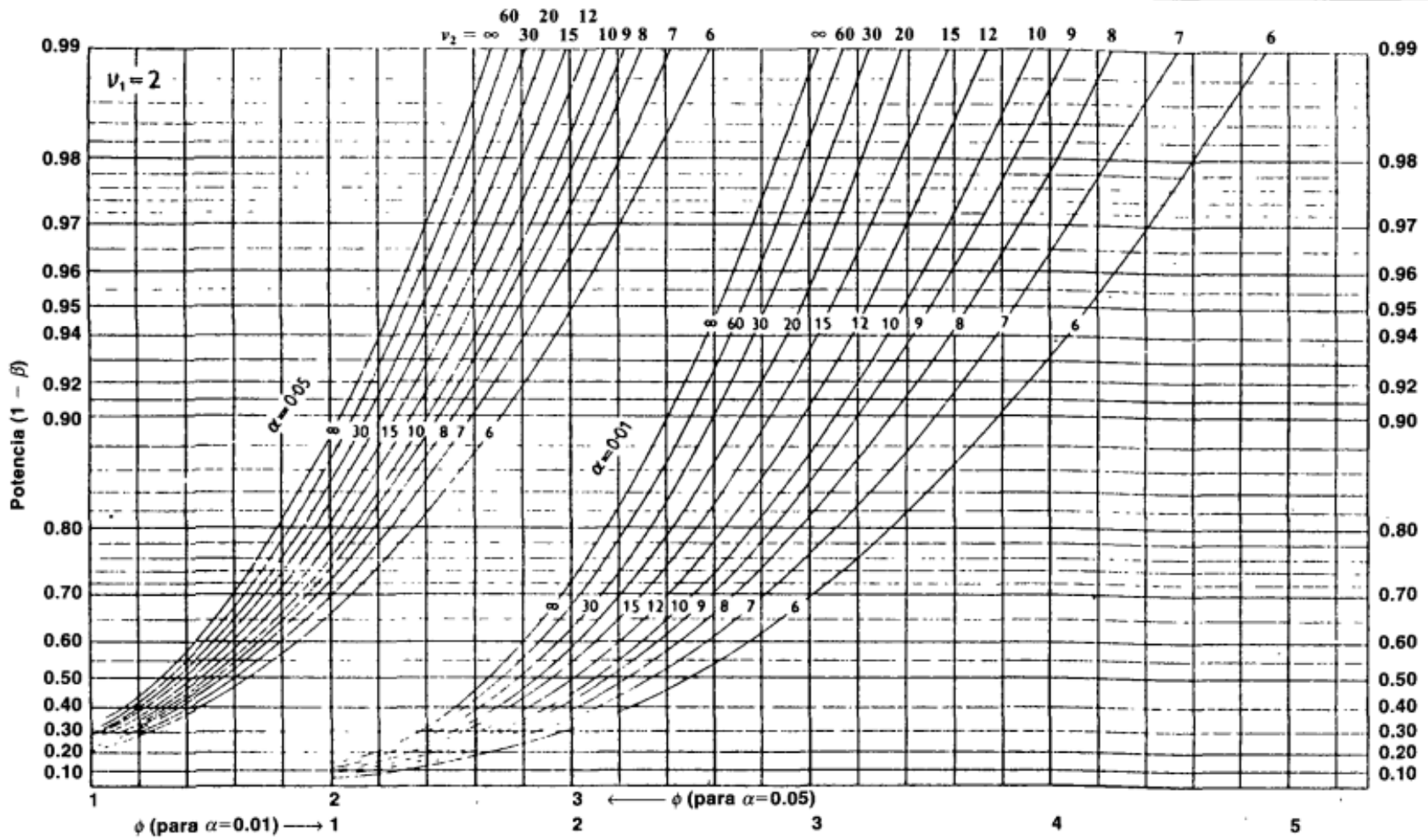


Figura 11. Ejemplo de curvas operativas (Kuehl, 2000).

Se presentan dos contratiempos al usar la función potencia y el parámetro de no centralidad para determinar el tamaño de muestra; el primero consiste en que el parámetro de no centralidad  $\lambda = r \sum_{i=1}^a \tau_i^2 / \sigma^2$  involucra el efecto de cada tratamiento  $\tau_i$  el cual no se conoce a menos de que se lleve a cabo el experimento. Es por esto que para determinar el tamaño de muestra se especifican los efectos  $\tau_i$  que se desearía detectar. El segundo contratiempo es que el valor de la varianza poblacional  $\sigma^2$  es generalmente desconocido; para remediar esto pueden considerarse tres posibilidades (Giesbrecht & Gumpertz, 2004): la primera es usar el conocimiento previo para establecer un valor razonable del valor de  $\sigma^2$ , la segunda posibilidad es establecer un rango de valores factibles de  $\sigma^2$  para lograr tener escenarios para el mejor y peor caso; mientras que la tercera posibilidad es realizar un estudio previo y usar el  $SC_E / (ra - a)$  como estimación de  $\sigma^2$ .

Para aclarar las ideas tratadas anteriormente y para ejemplificar la determinación del tamaño de muestra en este diseño, se proporcionan un par de ejemplos.

### 3.1. Ejemplo de determinación del tamaño de muestra.

*Ejemplo 1.* Suponga que se quiere llevar a cabo un experimento completamente al azar con un factor que tiene tres niveles. Se desea conocer el tamaño de muestra necesario para detectar efectos de  $\tau_1 = 0.21$ ,  $\tau_2 = 0.42$  y  $\tau_3 = -0.64^{**}$  en cada uno de los niveles, con nivel de significancia de 0.05 y con una potencia de 0.8. El conocimiento previo sugiere que un valor razonable para  $\sigma^2 = 0.2$ .

Primeramente se realiza la observación de que querer detectar estos efectos implicaría detectar las siguientes medias por tratamiento:

---

\*\* Para satisfacer el requerimiento del modelo estadístico en el cual  $\sum_{i=1}^3 \tau_i = 0$ , los valores reales de los efectos tendrían que ser:

$$\tau_1 = \frac{3}{14} \approx 0.21; \tau_2 = \frac{6}{14} \approx 0.42; \tau_3 = -\frac{9}{14} \approx -0.64$$

En los cálculos del ejemplo se trabajó con los valores reales no con las aproximaciones.

$$\tau_1 = 0.21 \Rightarrow \mu_1 = \mu + 0.21$$

$$\tau_2 = 0.42 \Rightarrow \mu_2 = \mu + 0.42$$

$$\tau_3 = -0.64 \Rightarrow \mu_3 = \mu - 0.64$$

Sin embargo, la media general  $\mu$  no se conoce a menos que se lleve a cabo el experimento. Es por eso que para determinar el tamaño de muestra se especificaron los efectos que se estaba interesado en detectar.

Por otra parte, en este caso se tiene que:

$$v_1 = a - 1 = 3 - 1 = 2$$

$$\lambda = \frac{r \sum_{i=1}^3 \tau_i^2}{\sigma^2} = \frac{r(9/14)}{0.2} = \frac{45}{14}r \approx 3.21r$$

Así que:

$$\Phi = \sqrt{\frac{\lambda}{a}} = \sqrt{\frac{45r/14}{3}} = \sqrt{\frac{15}{14}r} \approx \sqrt{1.07r}$$

$$v_2 = 3(r - 1)$$

En este punto se hace una elección del número de réplicas por nivel para averiguar qué potencia involucran, por ejemplo, si se eligiera  $r = 3$  tendríamos:

$$\lambda = \frac{45}{14} \times 3 \approx 9.64$$

$$\Phi = \sqrt{\frac{15}{14} \times 3} \approx 1.7$$

$$v_2 = 3 \times 2 = 6$$

Para determinar la potencia alcanzada con estos valores, se emplea la tabla mostrada en la Figura 11. Para ello se utiliza el conjunto de curvas mostradas en la parte izquierda pues  $\alpha = 0.05$ , después se sitúa aproximadamente el valor de

$\Phi$  en el eje horizontal y se asciende hasta encontrar la intersección con la correspondiente curva con  $v_2 = 6$ . Una vez obtenida la intersección con la curva se recorre de manera paralela al eje horizontal para encontrar el valor de la potencia en el eje vertical. Realizando este procedimiento, se encuentra que la potencia de la prueba en este caso sería aproximadamente 0.52, lo cual no es suficiente para los requerimientos. Ahora se hace evidente la razón por la cual se fija la potencia deseada y no el número de réplicas ya que fijar el número de réplicas primero puede conllevar a una baja potencia.

Los resultados que se obtienen eligiendo  $r = 4$  y  $r = 5$  usando la tabla de la Figura 11 se muestran a continuación:

$r = 4$	$r = 5$
$v_2 = 3 \times 3 = 9$	$v_2 = 3 \times 4 = 12$
$\lambda \approx 12.85$	$\lambda = 16.07$
$\Phi = \sqrt{\frac{15}{14}} \times 3 \approx 2.0$	$\Phi = \sqrt{\frac{15}{14}} \times 3 \approx 2.3$
$1 - \beta \approx 0.73$	$1 - \beta \approx 0.90$

Así que el número de réplicas necesarias sería de 5.

*Ejemplo 2.* Para el mismo experimento del ejemplo 1 considere que los efectos que se quieren detectar son más pequeños para cada nivel, es decir de  $\tau_1 = -0.16, \tau_2 = -0.16$  y  $\tau_3 = 0.3^{\dagger\dagger}$ . Si se usa el nivel de significancia de 0.05 utilice

---

<sup>††</sup> Para satisfacer el requerimiento del modelo estadístico en el cual  $\sum_{i=1}^3 \tau_i = 0$ , los valores reales de los efectos tendrían que ser:

$$\tau_1 = \tau_2 = -\frac{1}{6}; \tau_3 = \frac{1}{3}$$

En los cálculos del ejemplo se trabajó con los valores reales no con las aproximaciones.

$r = 11$  y  $r = 21$ <sup>#</sup> así como las curvas operativas de la Figura 11 para determinar la potencia que se podría alcanzar.

Ahora se tiene que:

$$v_1 = 2$$

$$\lambda = \frac{r \sum_{i=1}^3 \tau_i^2}{\sigma^2} = \frac{r(1/6)}{0.2} = \frac{5}{6}r$$

Así que:

$$\Phi = \sqrt{\frac{(5r/6)}{3}} = \sqrt{\frac{5}{18}r} \approx \sqrt{0.27r}$$

$$v_2 = 3(r - 1)$$

Los resultados cuando  $r = 11$  y  $r = 21$  se muestran a continuación:

$r = 11$	$r = 21$
$v_2 = 30$	$v_2 = 60$
$\lambda \approx 9.2$	$\lambda = 17.5$
$\Phi \approx 1.7$	$\Phi \approx 2.4$
$1 - \beta \approx 0.73$	$1 - \beta \approx 0.97$

Como conclusión de los ejemplos se puede observar que un aumento en el número de réplicas, es decir, del tamaño de muestra conlleva un aumento en la potencia. Por otra parte, si se desean detectar efectos más pequeños, entonces se requerirá un tamaño de muestra mayor que el que se necesita para detectar efectos más grandes, con el mismo nivel de significancia y potencia. También se puede observar cómo el uso de las curvas operativas es impreciso.

<sup>#</sup> Se proponen usar estos valores para encontrar una línea en las curvas operativas de la Figura 11.



A continuación se ejemplifica cómo pueden implementarse estos cálculos y construir curvas operativas con la ayuda de R.

### 3.2. Determinación del tamaño de muestra en R.

En principio, la potencia mostrada en la ecuación (3.1) puede determinarse especificando los grados de libertad del numerador, denominador, nivel de significancia y parámetro de no centralidad. La potencia puede implementarse fácilmente en R como se muestra en la Figura 12.

```
potenciaL<-function(v1,v2,alpha,lambda){  
  c=qf(alpha,v1,v2,lower.tail = F)  
  return(pf(c,v1,v2,lambda,F))  
}
```

Figura 12. Implementación de la función potencia mostrada en la ecuación (3.1).

Los cálculos implicados para determinar la potencia que producen cierto número de réplicas se pueden también programar; una posible implementación de ellos se muestra en la Figura 13.

```
calculos<-function(r,alpha,taus,var,v1=2){  
  a<-length(taus)  
  sum(taus^2)*r/var->lambda  
  phi<-sqrt(lambda/a)  
  v2<-a*(r-1)  
  potenciaL(v1,v2,alpha,lambda)->pot  
  return(c(v2,lambda,round(phi,1),pot))  
}
```

Figura 13. Función que realiza los cálculos implicados en la determinación de la potencia.

El resultado de aplicar la función *calculos* de la Figura 13, a los casos obtenidos en el *ejemplo 1* se muestra a continuación en la Figura 14.

Resultados obtenidos con R	round(calculos(3,0.05,c(3/14,6/14,-9/14),0.2,2),2) [1] 6.00 9.64 1.80 0.56	round(calculos(4,0.05,c(3/14,6/14,-9/14),0.2,2),2) [1] 9.00 12.86 2.10 0.77	round(calculos(5,0.05,c(3/14,6/14,-9/14),0.2,2),2) [1] 12.00 16.07 2.30 0.89
Resultados obtenidos a mano.	$r = 3$ $v_2 = 3 \times 2 = 6$ $\lambda \approx 9.64$ $\Phi = \sqrt{\frac{15}{14}} \times 3 \approx 1.7$ $1 - \beta \approx 0.52$	$r = 4$ $v_2 = 3 \times 3 = 9$ $\lambda \approx 12.85$ $\Phi = \sqrt{\frac{15}{14}} \times 3 \approx 2.0$ $1 - \beta \approx 0.73$	$r = 5$ $v_2 = 3 \times 4 = 12$ $\lambda = 16.07$ $\Phi = \sqrt{\frac{15}{14}} \times 3 \approx 2.3$ $1 - \beta \approx 0.90$

Figura 14. Resultados de la utilización de R para determinar el tamaño de muestra.

El uso de R permite mayor precisión y además brinda la posibilidad de determinar resultados que con las curvas operativas no se podrían obtener, por ejemplo, usando la función de la Figura 13, puede determinarse que en el *ejemplo 2*, para obtener una potencia de 0.8, hubieran bastado realizar 13 repeticiones por nivel.

Una posible implementación de la función potencia a través del parámetro  $\Phi$  se muestra en la Figura 15 y una curva de operación a partir de esta función se muestra en la Figura 16.

```
potenciaPhi<-function(v1,v2,alpha,phi){
  a=v1+1
  c=qf(alpha,v1,v2,lower.tail = F)
  return(pf(c,v1,v2,phi^2*a,F))
}
```

Figura 15. Implementación de la función potencia usando el parámetro  $\Phi$ .

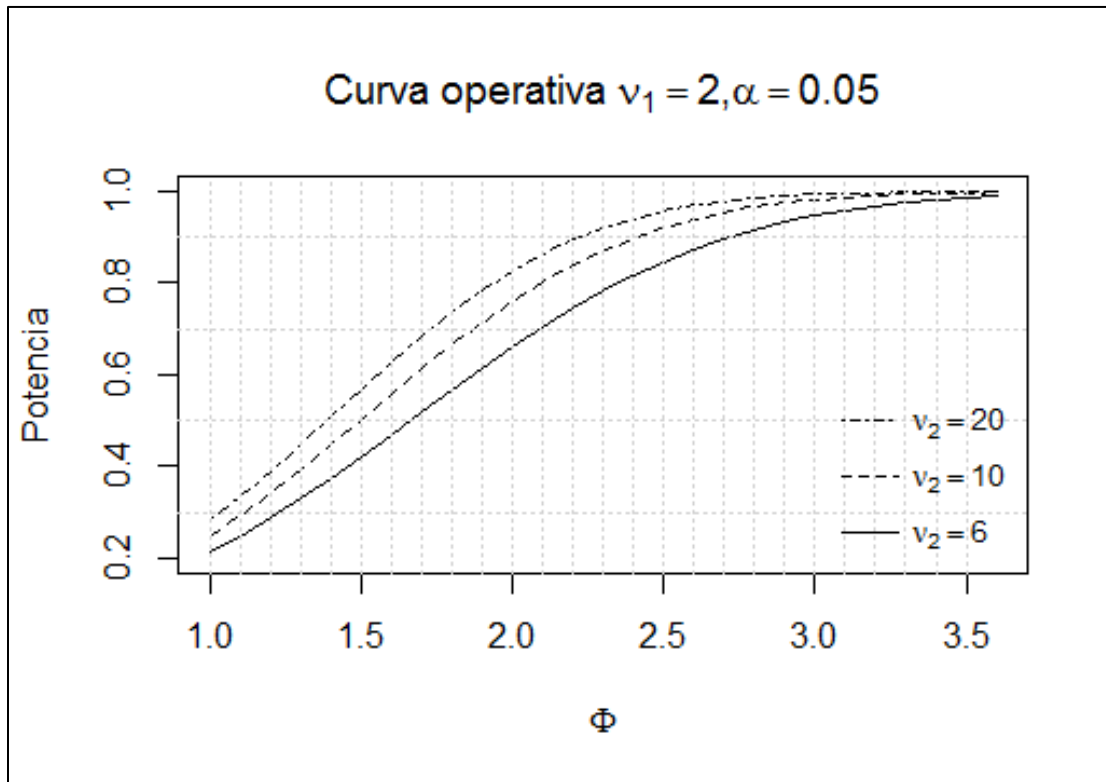


Figura 16. Curva operativa obtenida a través de la implementación de la función potencia en R.

Se puede hacer la observación de que el parámetro  $\Phi$  cumple sólo la función de reducir la escala en el eje horizontal de las curvas operativas. Además las curvas operativas como la mostrada en la Figura 11 tienen diferentes escalas en el eje vertical para enfatizar valores de la potencia más usuales. La gráfica de una curva de operación con respecto al parámetro de no centralidad  $\lambda$  se muestra en la Figura 17; debe observarse que a esta curva es similar a la que se obtiene usando el parámetro  $\Phi$ , con la diferencia de que la escala horizontal recorre más valores.

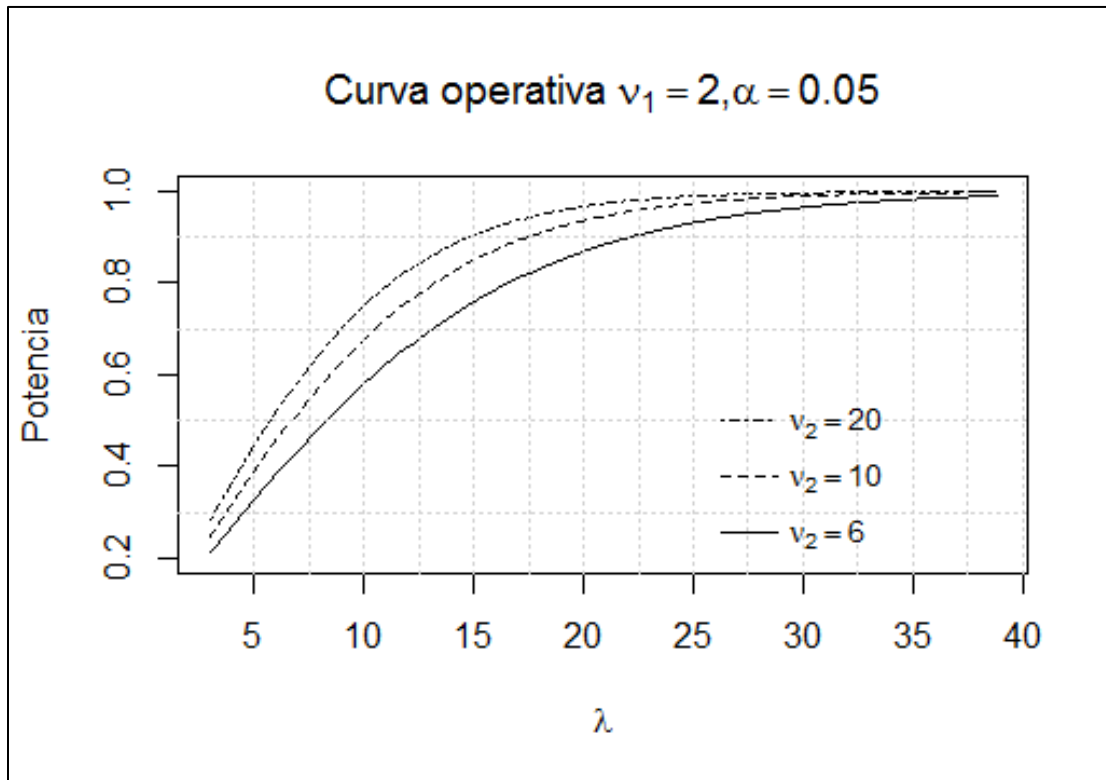


Figura 17. Curva operativa trazada a partir del parámetro de no centralidad  $\lambda$ . El código empleado para hacer la gráfica de la Figura 16 es análogo al empleado para hacer la gráfica de la Figura 17 y se muestra en la Figura 18.

```
#curva de operación con parámetro phi
v1=2;phi1=1;phi2=3.6
phival=seq(phi1,phi2,.1)#los valores de la función phi
plot(phival,potenciaPhi(v1,6,.05,phival),ylab = "Potencia",
      xlab=expression(Phi),type = "l",ylim=c(.2,1),
      main=expression(paste("Curva operativa ",nu[1]==2,",",alpha==0.05)))
points(phival,potenciaPhi(v1,10,.05,phival),type="l",lty=2)
points(phival,potenciaPhi(v1,20,.05,phival),type="l",lty=4)
abline(h=seq(.3,.9,.2),v=seq(1,3.5,.1),col="lightgray",lty="dotted")
legend(3,0.6, c(expression(nu[2]==20,nu[2]==10,nu[2]==6)),
      lty=c(4,2,1),cex=0.9, y.intersp=1.5,x.intersp=0.5,horiz=FALSE,bty="n")
```

Figura 18. Código empleado para generar curvas operativas.

#### 4. Determinación del tamaño de muestra en diseños de bloques completos al azar un factor varios niveles.

Cuando se va a llevar a cabo un experimento y se conocen de antemano diferencias importantes entre las unidades experimentales, es decir, fuentes de variación controlables, es aconsejable organizar las unidades experimentales en grupos llamados bloques. Si hay  $b$  bloques,  $a$  tratamientos y cada bloque tiene todos los tratamientos y estos a su vez se asignan de manera aleatoria, el diseño recibe el nombre de bloques completos al azar (RCBD) y su modelo estadístico es:

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

Donde:

- $\mu$  es la media general.
- $\tau_i$  es el efecto del  $i$ -ésimo tratamiento,  $i = 1, \dots, a$
- $\beta_j$  es el efecto de  $j$ -ésimo tratamiento,  $j = 1, \dots, b$
- $\varepsilon_{ij}$  es el error experimental, los cuales se suponen independientes y distribuidos  $N(0, \sigma^2)$ .

Además se asume que:

$$\sum_{i=1}^a \tau_i = \sum_{j=1}^b \beta_j = 0$$

En este caso, las sumas de cuadrados de los tratamientos, bloques, error y total,  $SC_{trat}$ ,  $SC_B$ ,  $SC_E$  y  $SC_T$  respectivamente, se definen como:

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

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

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

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

En donde se cumple la igualdad

$$SC_T = SC_{trat} + SC_B + SC_E$$

Y puede observarse que se hace una partición de la desviación del error experimental del diseño totalmente aleatorizado,  $y_{ij} - \bar{y}_{i.}$ , en dos componentes, el primer término se identifica con el criterio de bloques como  $(\bar{y}_{.j} - \bar{y}_{..})$ , el segundo término se identifica como un residual o error experimental  $y_{ij} - \bar{y}_{.j} - \bar{y}_{i.} + \bar{y}_{..}$  (Kuehl, 2000). La tabla ANOVA es como se muestra en la Tabla 3 (Montgomery, 2004).

Tabla 3. Tabla ANOVA para RCBD.

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	$F_0$
Tratamientos	$SC_{trat}$	$a - 1$	$CM_{trat} = \frac{SC_{trat}}{a - 1}$	$\frac{CM_{trat}}{CM_E}$
Bloques	$SC_B$	$b - 1$	$CM_B = \frac{SC_B}{b - 1}$	
Error	$SC_E$	$(a - 1)(b - 1)$	$CM_E = \frac{SC_E}{(a - 1)(b - 1)}$	
Total	$SC_T$	$ab - 1$		

En este caso, la atención se centra en averiguar si los efectos de los tratamientos son significativos; bajo el supuesto de normalidad para los errores, puede demostrarse que  $CM_{trat}$ ,  $CM_B$  y  $CM_E$  son variables aleatorias ji cuadradas independientes, por lo cual puede plantearse la siguiente prueba de hipótesis:

$$H_0: \tau_1 = \tau_2 = \dots = \tau_a = 0$$

$$H_1: \tau_i \neq 0 \text{ para al menos una } i$$

En caso de cumplirse  $H_0$ , puede demostrarse que el estadístico  $F_0 = CM_{trat} \div CM_E$  tiene una distribución  $F$  central con  $v_1 = a - 1$  grados de libertad en el numerador y  $v_2 = (a - 1)(b - 1)$  grados de libertad en el denominador. Mientras que bajo la hipótesis alternativa  $H_1$ ,  $F_0$  tiene una distribución  $F$  no central con parámetro de no centralidad (Giesbrecht & Gumpertz, 2004)

$$\lambda = b \sum_{i=1}^a (\tau_i - \bar{\tau})^2 / \sigma^2$$

En donde  $b$  es el número de bloques a usar, lo cual determina el tamaño de muestra y  $\bar{\tau} = \sum_{i=1}^a r_i \tau_i / \sum_{i=1}^a r_i$  donde  $r_i$  es el número de repeticiones en cada nivel del tratamiento, con lo cual, si  $r_i = r$  y apoyándose en el hecho de que  $\sum_{i=1}^a \tau_i = 0$ , puede concluirse que:

$$\lambda = \frac{b \sum_{i=1}^a \tau_i^2}{\sigma^2}$$

Y puede observarse que el parámetro de no centralidad tiene la misma estructura que para el caso de un diseño con un factor fijo, pero se sustituye  $r$  por  $b$ . Por otra parte, la potencia de la prueba puede escribirse como (Giesbrecht & Gumpertz, 2004):

$$1 - \beta = P \left( \frac{CM_{trat}}{CM_E} > F_{a-1, (a-1)(b-1)}^{1-\alpha} \mid \lambda \right)$$

De tal forma que el procedimiento para determinar el tamaño de muestra es muy similar al caso de un diseño con un factor fijo con más de un nivel, incluso Montgomery (Montgomery, 2004), sugiere usar las mismas curvas operativas con el parámetro  $\Phi$  definido como:

$$\Phi = \sqrt{\frac{b \sum_{i=1}^a \tau_i^2}{a\sigma^2}} = \sqrt{\frac{\lambda}{a}}$$

Así mismo, la determinación del tamaño de muestra mediante la potencia en el diseño RCBD presenta las mismas complicaciones que se tienen en el caso del diseño con un factor fijo en cuanto al desconocimiento de los efectos  $\tau_i$  y de la varianza  $\sigma^2$ .

A continuación se presenta un ejemplo de la determinación del tamaño de muestra en el caso de RCBD.

#### 4.1. Ejemplo de determinación del tamaño de muestra en diseño de bloques completos al azar.

*Ejemplo.* Suponga que se quiere llevar a cabo un experimento con bloques completos al azar con un factor y tres niveles. Se desea conocer el número de bloques necesario para detectar efectos de  $\tau_1 = 0.21$ ,  $\tau_2 = 0.42$  y  $\tau_3 = -0.64$ <sup>§§</sup>, con nivel de significancia de 0.05 y con una potencia de 0.8. El conocimiento previo sugiere que un valor razonable para  $\sigma^2 = 0.12$ .

Los cálculos se resumen a continuación:

---

<sup>§§</sup> Para satisfacer el requerimiento del modelo estadístico en el cual  $\sum_{i=1}^3 \tau_i = 0$ , los valores reales de los efectos tendrían que ser:

$$\tau_1 = \frac{3}{14} \approx 0.21; \tau_2 = \frac{6}{14} \approx 0.42; \tau_3 = -\frac{9}{14} \approx -0.64$$

En los cálculos del ejemplo se trabajó con los valores reales no con las aproximaciones.



Expresión	Evaluación
$v_1 = a - 1$	$v_1 = 3 - 1 = 2$
$\lambda = \frac{b \sum_{i=1}^3 \tau_i^2}{\sigma^2}$	$\lambda = \frac{b[(0.21)^2 + (0.42)^2 + (-0.64)^2]}{0.12} = 5.35b$
$\Phi = \sqrt{\frac{\lambda}{a}}$	$\Phi = \sqrt{\frac{5.35b}{3}} = \sqrt{1.78b}$
$v_2 = (a - 1)(b - 1)$	$v_2 = 2(b - 1)$

En este caso, para determinar la potencia alcanzada con determinado número de bloques, sólo se puede usar la curva operativa de la Figura 11 cuando el número de bloques es  $b = 4$  pues en este caso  $v_2 = 2(3) = 6$ , por lo cual, para calcular la potencia se empleó la función en R mostrada en la Figura 15, obteniéndose los siguientes resultados:

$b = 2$	$b = 3$	$b = 4$
$\Phi = 1.88$	$\Phi = 2.31$	$\Phi = 2.67$
$v_2 = 2$	$v_2 = 4$	$v_2 = 6$
$1 - \beta \approx 0.27$	$1 - \beta \approx 0.66$	$1 - \beta \approx 0.89$

Así que se deberían emplear 4 bloques para cumplir los requerimientos.

## 5. Determinación del tamaño de muestra en el diseño factorial de dos factores completamente al azar.

Una alternativa de diseño experimental cuando se estudia más de una variable o factor es el diseño factorial, en el cual en cada ensayo o réplica completa del experimento se investigan todas las combinaciones posibles de los niveles de los factores (Montgomery, 2004). Cuando hay solamente factores, el diseño se conoce como diseño factorial con dos factores. Para este tipo de diseño experimental, si uno de los factores es denotado por  $A$  y tiene  $a$  niveles, el otro es denotado por  $B$  con  $b$  niveles y hay  $n$  repeticiones, el modelo estadístico es:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$$

En donde:

- $\mu$ : es la media general.
- $\tau_i$ : es el efecto del nivel  $i$  del factor  $A$ ;  $i = 1, \dots, a$ .
- $\beta_j$ : es el efecto del nivel  $j$  del factor  $B$ ;  $j = 1, \dots, b$ .
- $(\tau\beta)_{ij}$ : es el efecto de la interacción entre los factores  $A$  y  $B$ .
- $\varepsilon_{ijk}$ : es el error experimental, los cuales se suponen independientes y distribuidos  $N(0, \sigma^2)$ .

Además se asume que los efectos cumplen con:

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

Las sumas de cuadrados son:

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

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

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

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

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

En donde se cumple la igualdad  $SC_T = SC_A + SC_B + SC_{AB} + SC_E$ . La tabla del análisis de varianza es como se muestra en la Tabla 4.

Tabla 4. Tabla ANOVA para un diseño factorial de dos factores.

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	$F_0$
Tratamientos de A.	$SC_A$	$(a - 1)$	$MS_A = \frac{SC_A}{a - 1}$	$\frac{MS_A}{MS_E}$
Tratamientos de B.	$SC_B$	$(b - 1)$	$MS_B = \frac{SC_B}{b - 1}$	$\frac{MS_B}{MS_E}$
Interacción.	$SC_{AB}$	$(a - 1)(b - 1)$	$MS_{AB} = \frac{SC_{AB}}{(a - 1)(b - 1)}$	$\frac{MS_{AB}}{MS_E}$
Error.	$SC_E$	$ab(n - 1)$	$MS_E = \frac{SC_E}{ab(n - 1)}$	
Total.	$SC_T$	$abn - 1$		

En este caso al haber dos factores, se pueden realizar pruebas de hipótesis sobre la nulidad de los efectos de cada factor y sobre la nulidad del efecto de interacción. En cada caso, de cumplirse la hipótesis nula, el estadístico de prueba tendrá una distribución  $F$  central, de lo contrario seguirá una distribución  $F$  no central. La Tabla 5 muestra el parámetro de no centralidad  $\lambda$  en cada caso. En este caso, la función  $\Phi^2$  está vinculada al parámetro de no centralidad de

manera que  $\Phi^2 = \frac{\lambda}{v_1+1}$  donde  $v_1$  son los grados de libertad del numerador (Kuehl, 2000).

Tabla 5. Parámetro de no centralidad  $\lambda$  para las diferentes pruebas de hipótesis en un experimento factorial (Montgomery, 2004) y (Kuehl, 2000).

Prueba sobre la nulidad del factor.	Parámetro de no centralidad $\lambda$ .	Grados de libertad numerador $v_1$ .	Grados de libertad denominador $v_2$ .
$A$	$\frac{bn}{\sigma^2} \sum_{i=1}^a \tau_i^2$	$a - 1$	$ab(n - 1)$
$B$	$\frac{an}{\sigma^2} \sum_{j=1}^b \beta_j^2$	$b - 1$	$ab(n - 1)$
$AB$	$\frac{n}{\sigma^2} \sum_{i=1}^a \sum_{j=1}^b (\tau\beta)_{ij}^2$	$(a - 1)(b - 1)$	$ab(n - 1)$

Debido a que se conoce el parámetro de no centralidad  $\lambda$  y la función  $\Phi$  del parámetro de no centralidad, se podrán utilizar las curvas operativas y los procedimientos empleados para determinar el tamaño de muestra en el caso de un diseño de un solo factor e incluso los programas descritos en la sección 3.2. Esto se ilustra a continuación.

### 5.1. Ejemplo de determinación del tamaño de muestra en un diseño factorial con dos factores.

*Ejemplo.* Se realizó un experimento factorial en ratones<sup>\*\*\*</sup> con dos factores donde el factor  $A$  tenía tres niveles y el factor  $B$  dos niveles, además  $\sum_i \tau_i^2 / \sigma^2 = 2$ . Si se desea tener un nivel de significancia de 0.05 y una potencia de 0.8, determinar el tamaño de muestra en el caso de su cálculo se base en la nulidad de los efectos del factor  $A$ .

<sup>\*\*\*</sup> Esta ejemplificación es una adaptación del ejemplo 2.3.1 citado por (Odeh & Fox, 1991) en la página 13, en él se describe mejor en qué consiste este experimento.

En este caso, los datos a emplear son:

$\alpha = 0.05$
$1 - \beta = 0.8$
$\sum_i \tau_i^2 / \sigma^2 = 2$
$a = 3$
$b = 2$

Con los datos se obtiene:

Expresión	Evaluación
$v_1 = a - 1$	$v_1 = 3 - 1 = 2$
$v_2 = ab(n - 1)$	$v_2 = 3(2)(n - 1) = 6(n - 1)$
$\lambda = \frac{bn}{\sigma^2} \sum_{i=1}^a \tau_i^2$	$\lambda = \frac{2n}{\sigma^2} \sum_{i=1}^3 \tau_i^2 = 2n(2) = 4n$
$\Phi = \sqrt{\frac{\lambda}{a}}$	$\Phi = \sqrt{\frac{4n}{3}}$

Utilizando el algoritmo mostrado en la Figura 15, se llegan a los resultados mostrados a continuación. Cabe mencionar que Odeh y Fox usan una metodología distinta para calcular la potencia de la prueba basándose en la determinación de los grados de libertad del denominador del estadístico  $F$  y llegan a la conclusión de que el valor crítico es  $F_{0.05}(2,18)$ , lo cual indica que han usado  $n = 4$  repeticiones.

$n = 2$	$n = 3$	$n = 4$
$1 - \beta \approx 0.49$	$1 - \beta \approx 0.78$	$1 - \beta \approx 0.92$

## Conclusiones.

Con este trabajo se ha podido vislumbrar la importancia de la determinación del tamaño de muestra en tanto a que un tamaño muy pequeño puede llevar a manejar una potencia muy baja y en consecuencia una baja probabilidad de rechazar correctamente la hipótesis nula en un experimento.

El problema de la determinación del tamaño de muestra es amplio y profundo pues no sólo se encontraron textos antiguos sino también se fuentes y softwares más recientes; es además un tema que se sigue estudiando.

Se ha logrado observar que la potencia de la prueba se vincula al tamaño de la muestra y por lo tanto puede emplearse para determinar el tamaño de muestra. Se revisó que de manera general en diseños experimentales sencillos, el cálculo del tamaño de muestra se ve influenciado por los siguientes factores:

- La diferencia de efectos que se quiera detectar; mientras más pequeño sea el efecto más grande será el tamaño de muestra
- La varianza; mientras más grande sea la varianza más grande será el tamaño de muestra
- El nivel de significancia pues a menores niveles mayor será el tamaño de la muestra.
- La potencia, mientras mayor sea, se requerirá de una muestra más grande.

También pudo vislumbrarse la potencialidad que ofrecen los programas actuales en cuanto a la determinación del tamaño de muestra.

## Anexo. Uso de software para determinar el tamaño de muestra.

Se han encontrado dos programas desarrollados específicamente para el estudio de la potencia y el tamaño de muestra; ambos son de uso libre. Uno de ellos es *G\*Power* y el otro *Piface*. Ambos programas involucran teoría más específica y más profunda a la abordada en este trabajo y en consecuencia permiten más opciones o la variación de más parámetros para determinar la potencia o el tamaño de muestra. Se ha tratado de aplicar estos programas en la medida de lo posible a los ejemplos abordados a lo largo del texto. Los resultados obtenidos se detallan a continuación.

### Aplicación de *G\*Power* a la determinación del tamaño de muestra para un estudio de diferencia de medias con varianzas iguales.

Para este caso, se tienen los siguientes datos:

$$\alpha = \beta = 0.05$$
$$\frac{\delta}{\sigma} = 2$$

El procedimiento para la aplicación del programa fue:

1. Se selecciona la opción *t test* y la opción *Means: Difference between two independent means (two groups)*.
2. Se especifica el efecto de tamaño o *effect size*. Según el manual de este programa, se define como:

$$d = \frac{\mu_1 - \mu_2}{\sigma}$$

En este caso, si se establece  $\sigma = 1$ , se tiene que  $\delta = 2$ , lo cual establecería que la diferencia entre las medias es 2 y se puede usar la opción *Determine=>* para obtener el efecto de tamaño como se muestra en la figura A1. Una vez determinado este efecto se selecciona la opción *Calculate and transfer to main window*.

The image shows a dialog box in G\*Power software. It has two radio button options: 'n1 != n2' (unselected) and 'n1 = n2' (selected). Under 'n1 != n2', there are three input fields: 'Mean group 1' (0), 'Mean group 2' (1), and 'SD  $\sigma$  within each group' (0.5). Under 'n1 = n2', there are four input fields: 'Mean group 1' (0), 'Mean group 2' (2), 'SD  $\sigma$  group 1' (1), and 'SD  $\sigma$  group 2' (1). Below these is a 'Calculate' button and an 'Effect size d' field showing the value 2. At the bottom, there are 'Calculate and transfer to main window' and 'Close' buttons.

Figura A1. Determinación de *effect size* para diferencia de medias en G\*Power.

3. Se llenan los campos para la probabilidad del error tipo 1  $\alpha = 0.05$  y la potencia  $1 - \beta = 0.95$ . Como se suponen las muestras del mismo tamaño  $N_2/N_1 = 1$  como se muestra en la figura A2.



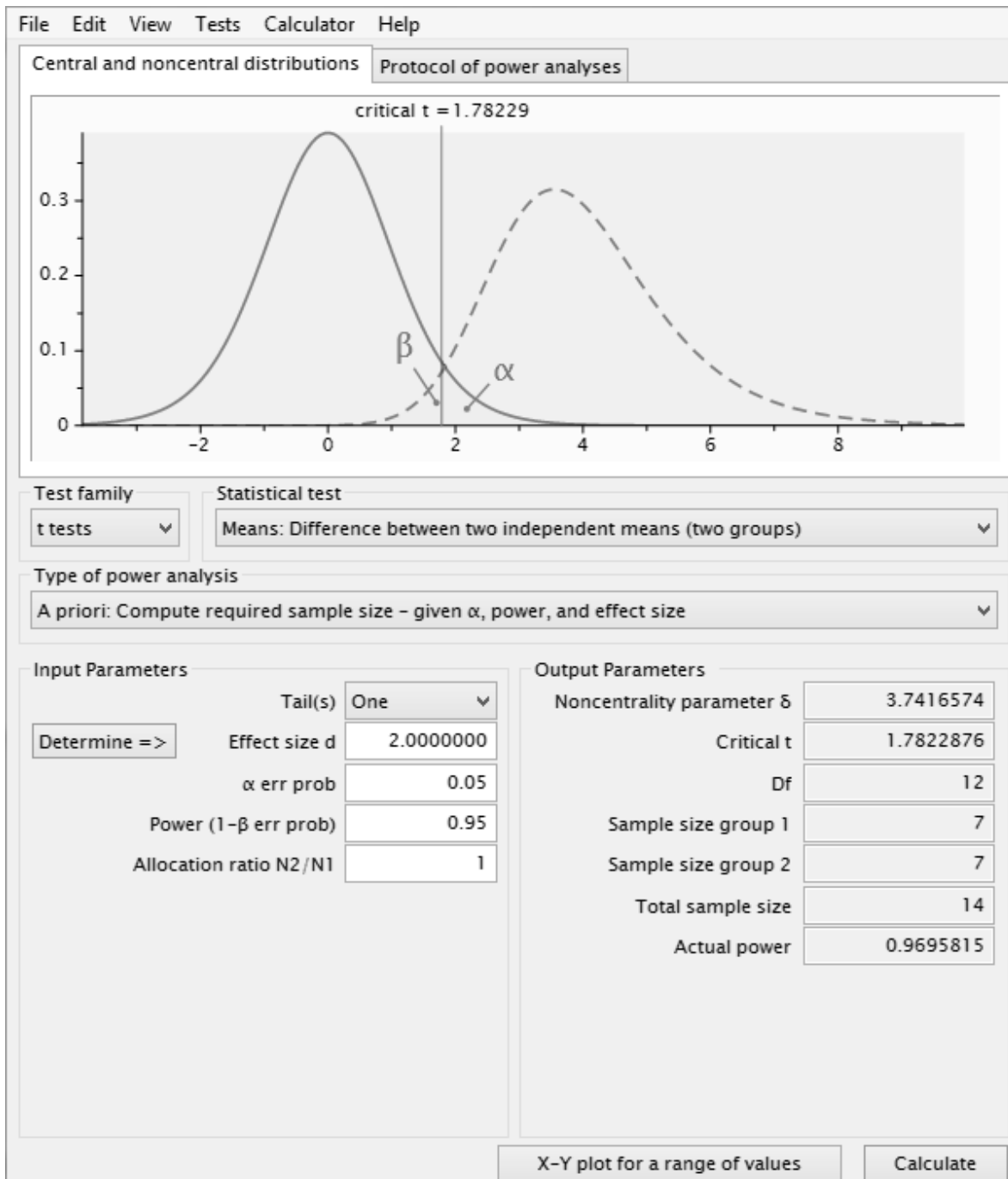


Figura A2. Resultado de la determinación del tamaño de muestra cuando para un experimento sobre diferencia de medias cuando  $\alpha = \beta = 0.05$  y  $\delta/\sigma = 2$  en G\*Power.

El resultado obtenido fue que el tamaño de muestra necesario fue de 7 unidades en cada grupo alcanzándose una potencia de 0.96. Cabe señalar que el tamaño

de muestra determinado sin el uso de este programa fue de 7 unidades por grupo.

### Aplicación de *Piface* a la determinación del tamaño de muestra para un estudio de diferencia de medias con varianzas iguales.

En este caso la aplicación es sencilla, sólo hay que seleccionar la opción *two-sample t test* y especificar los siguientes valores como se muestra en la figura A3:

- Equal sigmas
- Alpha: 0.05
- True difference of means: 2
- Power: 0.95

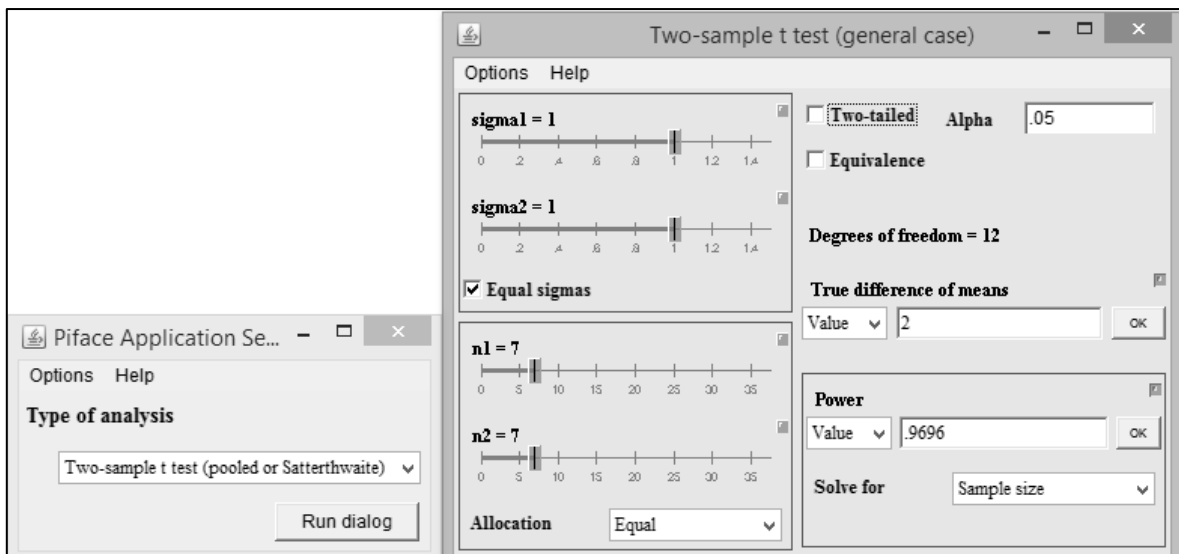


Figura A3. Aplicación de *Piface* para determinar el tamaño de muestra cuando  $\alpha = \beta = 0.05$  y  $\delta/\sigma = 2$ .

Una vez especificada la potencia, el programa realiza sus operaciones y determina que el tamaño de muestral necesario es de 7 unidades en cada grupo. Además determina que la potencia alcanzada es de 0.96.

## Aplicación de *G\*Power* a la determinación del tamaño de muestra en el caso de un factor fijo.

En este caso, los datos trabajados en el ejemplo del texto principal fueron:

$$\tau_1 = \frac{3}{14} \approx 0.21; \tau_2 = \frac{6}{14} \approx 0.42; \tau_3 = -\frac{9}{14} \approx -0.64$$
$$\sigma^2 = 0.2 \Rightarrow \sigma \approx 0.45$$
$$\alpha = 0.05$$
$$1 - \beta = 0.8$$

Los pasos a seguir para la aplicación de *G\*Power* fueron:

- (1) Se seleccionó *F test y ANOVA: Fixed effects, omnibus, one-way*.
- (2) Se determinó el valor de *effect size from means* de la siguiente manera (figura A4):

Select procedure  
Effect size from means

Number of groups 3

SD  $\sigma$  within each group 0.45

Group	Mean	Size
1	1.21	5
2	1.42	5
3	0.35	5

Equal n 5

Total sample size 15

Calculate Effect size f 1.028697

Calculate and transfer to main window

Close

Figura A4. Determinación de *effect size from means*.

- a. Se especifica el número de grupos que en este caso fue 3.
- b. Se especifica el valor de la desviación estándar que en este caso es  $\sigma = 0.45$ .
- c. Se especifica que todos los grupos son del mismo tamaño.
- d. Se especifican las medias tomando en cuenta que:

$$\tau_1 = 0.21 \Rightarrow \mu_1 = \mu + 0.21$$

$$\tau_2 = 0.42 \Rightarrow \mu_2 = \mu + 0.42$$

$$\tau_3 = -0.64 \Rightarrow \mu_3 = \mu - 0.64$$

Como el programa no acepta valores negativos, se propuso el valor  $\mu = 1$ , con lo cual:

$$\mu_1 = 1 + 0.21 = 1.21$$

$$\mu_2 = 1 + 0.42 = 1.42$$

$$\mu_3 = 1 - 0.64 = 0.35$$

- e. Se eligió la opción calcular y llevar a la ventana principal.
- (3) Se llenan los campos para el error tipo 1, la potencia y número de grupos como se muestra en la figura A5.

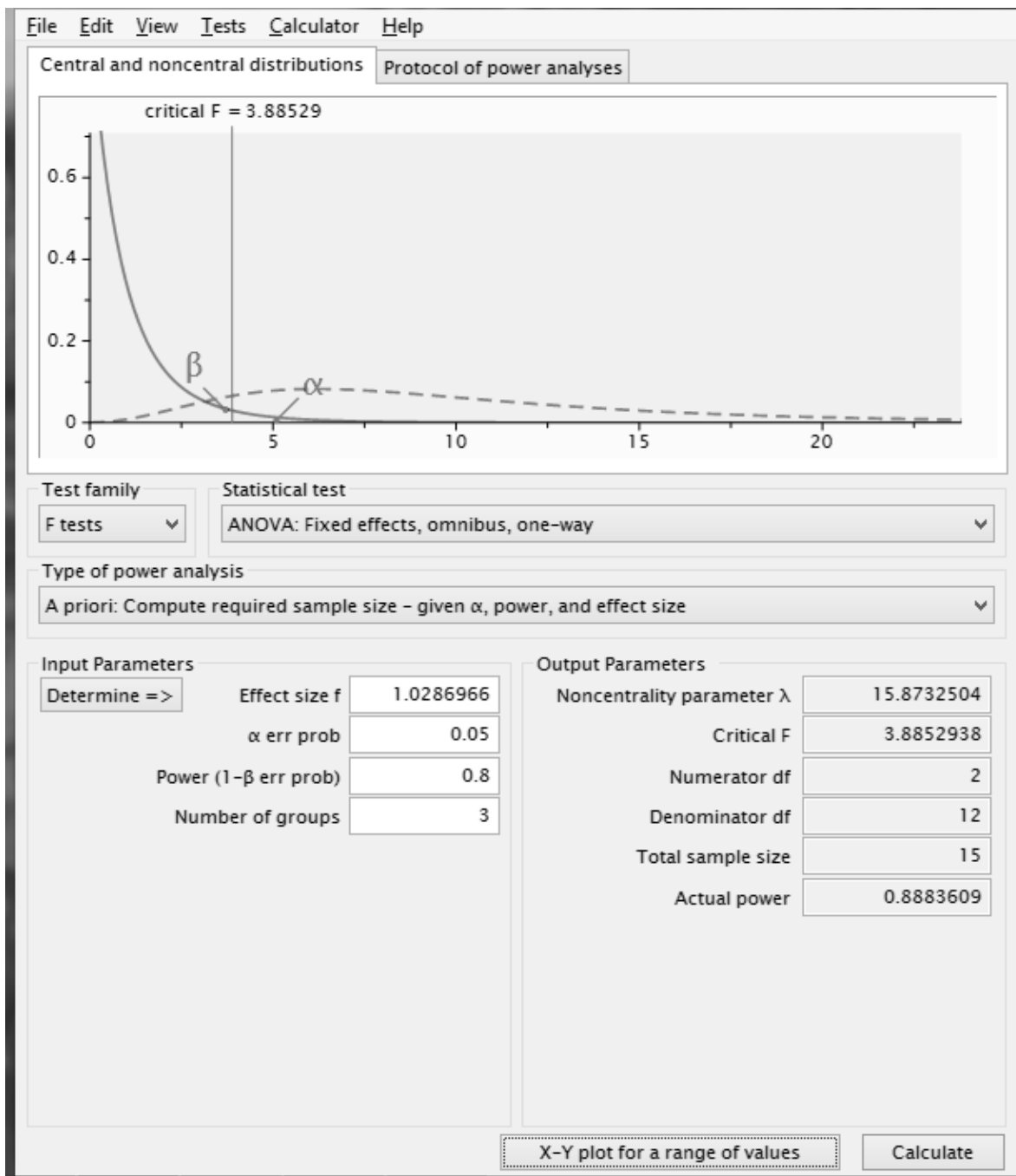


Figura A5. Resultados obtenidos con G\*Power para un diseño con un factor.

En este caso, el programa indica que el tamaño total de la muestra es de 15 unidades; considerando que se reparten de igual forma en cada nivel del factor, cada uno tendría 5 repeticiones alcanzándose una potencia de 0.88. Cabe señalar

que el resultado obtenido sin usar este programa fue de 5 unidades por nivel con una potencia aproximada de 0.90.

### Aplicación de Piface para determinar el tamaño de muestra en un diseño factorial de dos factores completamente al azar.

En este caso, se selecciona la opción *Two-way ANOVA* con el modelo *row|col*, se definen los niveles como *row 3 col 2* y se selecciona la opción *replicated within cells*, como se muestra en la figura A6. En la siguiente ventana se varía el tamaño de muestra. Al variar el número de repeticiones dentro de los tratamientos se obtienen los siguientes resultados:

$n = 2$	$1 - \beta \approx 0.48$
$n = 3$	$1 - \beta \approx 0.79$
$n = 3$	$1 - \beta \approx 0.92$

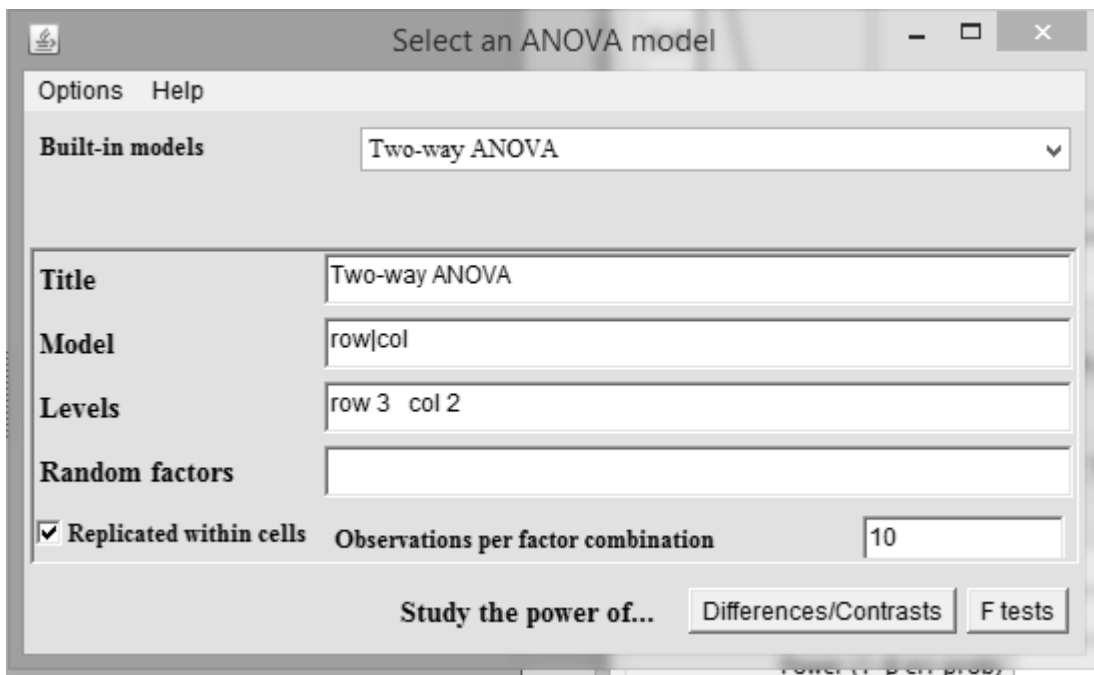


Figura A6. Ventana de diálogo para el diseño factorial de dos factores completamente al azar en Piface.

Las figuras A7, A8 y A9 muestran los resultados al variar número de repeticiones dentro de los tratamientos. Cabe mencionar que estos resultados coinciden con los obtenidos sin usar este software.

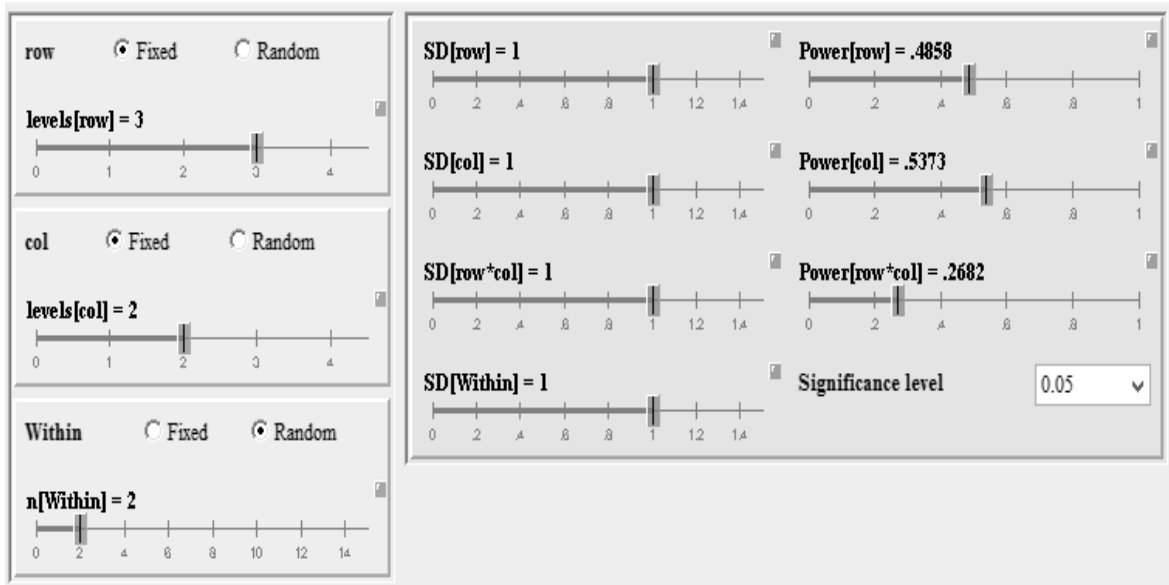


Figura A7. Resultados obtenidos en *Piface* cuando  $n = 2$ .

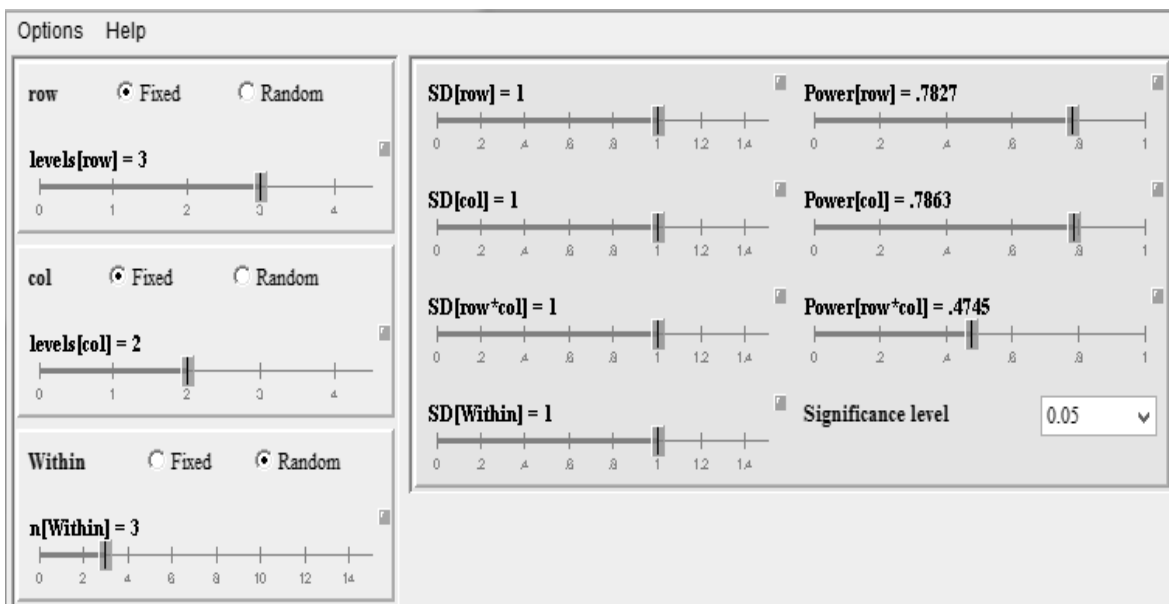


Figura A7. Resultados obtenidos en *Piface* cuando  $n = 3$ .

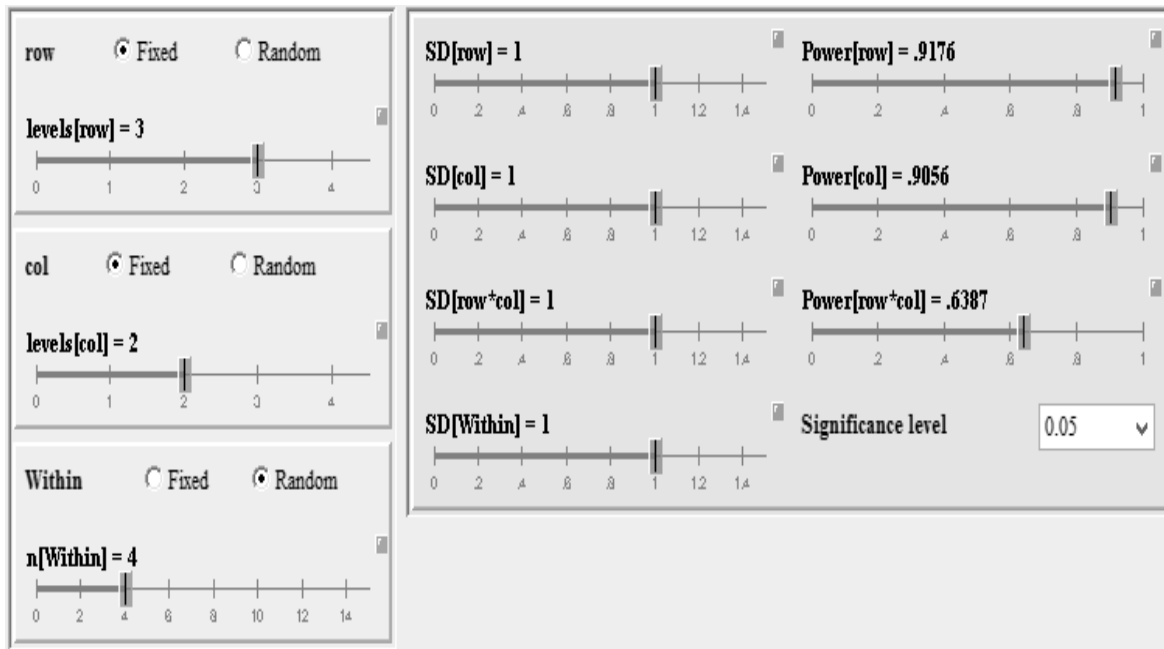


Figura A8. Resultados obtenidos en *Piface* cuando  $n = 4$ .



## Bibliografía

- Adcock, C. (2008, 06 10). *Sample Size Determination: A Review*. Retrieved from JSTOR: <http://www.jstor.org/stable/2988530>
- Davies, O. (1960). *The design and analysis of industrial experiments*. Londres: Imperial chemical industries by oliver and boyd.
- Desu, M., & Raghavarao, D. (1990). *Sample Size Methodology*. San Diego: Academic Press.
- Devore, J. (2008). *Probabilidad y estadística para ingeniería y ciencias*. Cengage Learning.
- Geza, S. (2007). *Introduction to Probability with Statistical Applications*. Boston: Birkhauser.
- Giesbrecht, F., & Gumpertz, M. (2004). *Planning, Construction, and Statistical Analysis of Comparative Experiments*. New Jersey: John Wiley & Sons.
- Gutiérrez Pulido, H., & Vara Salazar, R. (2012). *Análisis y diseño de experimentos*. México D.F.: McGraw-Hill Interamericana.
- Johnson, R., & Kuby, P. (2012). *Estadística elemental*. México: Cengage Learning.
- Kuehl, R. (2000). *Diseño de experimentos. Principios estadísticos de diseño y análisis de investigación*. México: Thomson Learning.
- Montgomery, D. (2004). *Diseño y análisis de experimentos*. México: Limusa Wiley.
- Odeh, R., & Fox, M. (1991). *Sample Size Choice. Charts for Experiments with Linear Models*. New York: CRC press.
- Rincón, L. (2007). *Curso elemental de probabilidad y estadística*. Ciudad de México.
- Ross, S. (2010). *Introductory statistics*. San Diego, California : Elsevier.
- Russell V., L. (2001). Some Practical Guidelines for Effective Sample Size Determination. *The American Statistician*, 187-193.
- Ryan, T. (2013). *Sample Size Determination and Power*. New Jersey: John Wiley & Sons.

Walpole, R., Myers, R., Myers, S., & Keying, Y. (2007). *Probabilidad y estadística para ingeniería y ciencias*. Pearson Educación.