PARTE 4

El error tipo II en el análisis de la significación de los efectos

CAPÍTULO 5

El error tipo II en el análisis de la significación de los efectos

5.1 Introducción

Algo que distingue el uso de las pruebas estadísticas clásicas es el contar con un procedimiento que permite decidir entre la hipótesis nula H_0 y la hipótesis alternativa H_1 en función del "riesgo" asociado con cometer un error Tipo I (rechazar una hipótesis nula verdadera). Al fijar el investigador el valor de de la probabilidad de error Tipo I, α, que desea o necesita usar, está consciente de que el procedimiento de prueba le garantiza que, de un gran número de veces que se emplee este procedimiento, el porcentaje de veces que la H_0 verdadera será rechazada erróneamente en favor de H_1 es α. Sin embargo no es común en la práctica que en estos procedimientos de decisión se cuestione acerca de la posibilidad de que el mecanismo de prueba acepte H_0 falsas (error Tipo II), y menos aún acerca del valor de la probabilidad de cometerlo (valor β). Quizá esta forma de proceder se geste en los mismos libros de texto, ya que una vez tratados los tipos de error, casi sin excepción, no volverán a mencionar el error tipo II en ninguno de los métodos que empleen contraste de hipótesis, y todo girará en torno a α. Sin embargo no hay que olvidar que en el enfoque clásico de Neyman-Pearson la decisión de aceptar o rechazar la hipótesis nula depende del costo asociado con cometer un error tipo I o un error tipo II.

Si se le pregunta a un ingeniero o investigador si prefiere tener en cuenta variables que no influyen en el proceso o desechar variables que si influyen, seguramente contestará que prefiere tomar en consideración variables que no influyen. A final de cuentas si se controla una variable que no influye, ésto no afectará al proceso, en cambio el no controlar una variable que si influye seguramente traerá contratiempos y disminuirá la calidad del producto. Paradójicamente cuando se estudia la significación de los efectos se hace hincapié en el porcentaje de veces que la prueba lleva a incluir erróneamente una variable inerte, y nada se dice del porcentaje de veces que la prueba desecha una variable activa de manera equivocada. Esto es particularmente cierto en los diseños de filtración (*screening designs*), cuyo objetivo primario es separar de todas las variables de interés aquellas que no tienen un efecto significativo en la respuesta. Aunque en estos estudios se asume implícitamente el principio de escasez de factores (*sparsity factor*), con lo cual solo esperamos un número reducido de factores activos, es factible que el proceso de selección tenga valores grandes de la tasa de error Tipo II, razón por la cual seguramente dejaremos de considerar variables relevantes para nuestros estudios posteriores.

El problema de la falta de atención al error tipo II en parte puede deberse a que generalmente no existe un valor único de β, y este hecho marca una diferencia sustancial con el empleo del valor α. Si la hipótesis alterna es una hipótesis simple, por ejemplo H_1: $\theta = 43$, se tendrá, para un valor fijo de α, un valor de β; no obstante lo usual es que la hipótesis alterna sea una hipótesis compuesta[1], tal vez del tipo H_1: $\theta > 40$, obligando a que exista un valor de β para cada posible valor del parámetro mayor que 40. Así al no rechazar H_0 el investigador se ve precisado a especular posibles valores plausibles de θ y para cada uno de ellos encontrar el riesgo que se está asumiendo en esta prueba al aceptar la hipótesis nula. Esto genera una "imprecisión" en el riesgo que se tiene al no rechazar H_0.

Otro aspecto que quizá ha contribuido también a soslayar el uso de β es el hecho de que las pruebas comúnmente empleadas son las "mejores", en el sentido de minimizar los valores de β (en realidad se busca maximizar la

---

[1] Formalmente se dice que la hipótesis es compuesta si la hipótesis de interés especifica más de una distribución posible para la muestra. Casella y Berger (1990).

potencia[2]) sujetos a una cota en α, y por tanto aunque no sabemos sus valores sabemos que no hay otra prueba que nos dé valores menores de β para un α fijo. También es conocido que para un tamaño de muestra fijo se cumple que en la medida que disminuye la tasa de un error se incrementa la tasa del otro error, pero que en muchos casos es posible disminuir ambas tasas incrementando el tamaño de la muestra, esto último lleva a la recomendación de tener el máximo tamaño de muestra posible, garantizando la disminución de β pero dejando desconocido su valor.

La determinación de la tasa de error tipo II en los diseños factoriales demanda contar con una estimación de la varianza de los efectos. En el caso de tener experimentos factoriales con réplicas se puede recurrir a las técnicas formales de contraste de hipótesis para juzgar la significación de los efectos estimados, y al tener una estimación de la varianza es posible evaluar sin mayor problema una tasa de error β para un valor de α predeterminado.

Una situación diferente se presenta cuando no hay réplicas. En los diseños factoriales sin réplicas es común emplear el método de la representación de los efectos en papel probabilístico normal, propuesto por Daniel (1959), para discriminar los efectos significativos. Al no ser el método de Daniel una prueba estadística "formal" no se cuenta con una estimación de la desviación estándar ni con grados de libertad para el error, lo que a primera vista imposibilita el cálculo de la tasa de error tipo II[3]. Empero paquetes

---

[2] La función potencia $Pot(\theta)$ está definida como

$$Pot(\theta) = \begin{cases} \alpha & si\ \theta \in \Theta_0 \\ 1-\beta(\theta) & si\ \theta \in \Theta_A \end{cases}$$

donde Θ_0 es el subconjunto del espacio paramétrico definido por H_0 i.e. H_0: $\theta \in \Theta_0$.

[3] Es notable como se logró imponer este método en la práctica experimental a pesar de la advertencia que Daniel hace al comentar las fallas de la graficación Half-Normal "Es innecesario alertar a los estadísticos experimentados que el uso del gráfico Half-Normal, sugerido aquí, está todavía lleno de sesgos subjetivos, que no se ofrece como un substituto general para el análisis de la varianza y que su uso de forma rutinaria puede ser catastrófico". Daniel enfatiza que se debe esperar que

estadísticos como Minitab emplean un procedimiento para estimar la desviación estándar de los efectos (seudo error estándar, PSE) desarrollado por Lenth (1989) y con base en él determinan los efectos que en el gráfico en papel probabilístico normal (y también en el diagrama de Pareto) se destacarán como significativos. Esto se hace a través de una prueba convencional, esto es, los efectos significativos son aquellos que en valor absoluto superan PSE*$t_{\alpha/2,\, m/3}$, donde m es el número de efectos.

Los diseños factoriales a dos niveles han tenido su propio desarrollo, con nomenclatura específica y métodos especiales que los hace parecer, para muchos usuarios, al margen de los modelos lineales cuando en realidad no es así. El uso de los métodos de selección de variables para la definición del modelo permite en muchos casos contar con una estimación de la varianza del error cuando no se tiene réplicas, esta estimación puede ser mejor que la que proporciona el método de Lenth. En el anexo 7 se comenta el método de Lenth y se hacen algunas reflexiones acerca de su uso en la representación de los efectos en papel probabilístico normal y acerca del empleo de métodos de selección de variables para estimar la varianza.

Otra posibilidad para el análisis de la significación de los efectos, al no tener réplicas, consiste en agrupar las interacciones de orden superior para estimar la varianza. Esta es por ejemplo la opción por defecto que tiene el software Statistica.

Ya sea con la estimación de la varianza del método de Lenth, con la que se derive al emplear algún método de selección de variables, o con la que se tiene al agrupar las interacciones de orden superior, es posible contar con la tasa de error tipo II en los diseños factoriales a dos niveles sin réplicas.

---

sucedan errores de ambos tipos al usar su procedimiento y señala que una línea sin grandes irregularidades en el gráfico puede dar lugar a sobreestimar la magnitud del error y por tanto perder un cierto número de efectos reales al determinar los efectos significativos.

5.2 El gráfico Alfa-Beta

Ya Box (1982) destacó que las decisiones que más influyen en los resultados de una investigación no dependen realmente de criterios estadísticos. Estas decisiones conciernen con la elección de la región de interés, la selección de las variables que participarán, sus rangos y localizaciones y las transformaciones necesarias. Si realmente, como parece ser, es más relevante en la industria no omitir variables activas que incluir variables inertes se necesita que se modifique la primacía del error tipo I a favor del error tipo II, esto demandará de los usuarios el acostumbrarse a fijar los niveles aceptables de ambas tasas de error, y de ser necesario condicionar el valor de α al valor deseado de β.

Una forma de abordar el problema de la "imprecisión" en el riesgo que se tiene al no rechazar H_0 consiste en fijar una cota inferior del valor del efecto en la hipótesis alternativa. Esto es, será necesario que el investigador defina la magnitud del **efecto mínimo**, *em*, que se desea que no pase desapercibida en el momento de la discriminación de los efectos significativos. A partir de este valor se deberá seleccionar la tasa de error tipo II que se está dispuesto a asumir.

Consideremos por ejemplo que un efecto real de 4 unidades ya es relevante para los ingenieros (*em* = 4) y que estos quisieran que la prueba lo detectara un 80% de las veces, o equivalentemente aceptarían una tasa de error tipo II del 20%. En el gráfico Alfa-Beta de la Figura 5.1 se tiene los valores β, bajo el supuesto de normalidad, que se generan cuando se tiene H_0: Efecto = 0 vs H_1: Efecto = 4 y diferentes valores de α, la estimación de la desviación estándar de los efectos corresponde a la que se obtiene con el método de Lenth, para el ejemplo de Myers, Kim y Griffiths (1997).

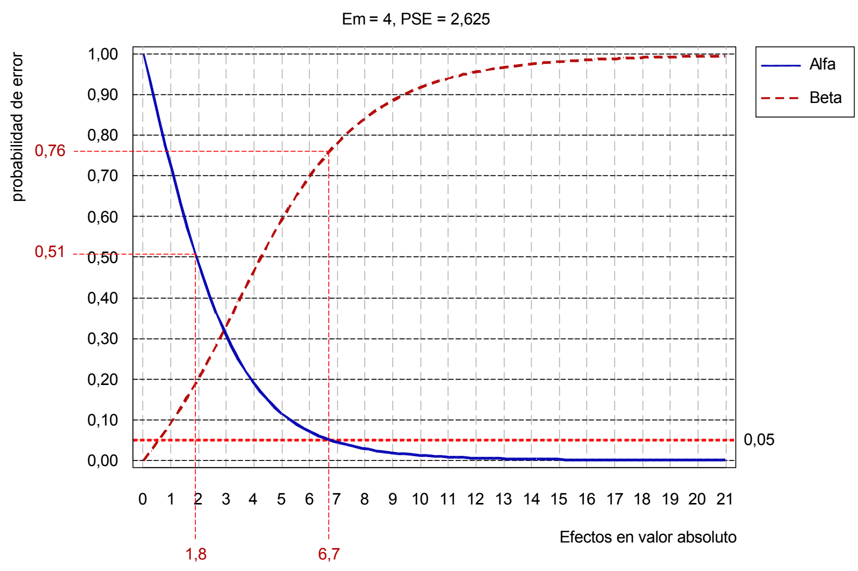

Figura 5.1 Gráfico Alfa-Beta para *em* = 4 con una estimación del error de los efectos de 2,625 y gl = 5.

La Figura 5.1 muestra que si se tiene un valor de *em* igual a 4 y se atiene al valor de alfa del 5% para discriminar los efectos significativos, entonces se produce una tasa de error tipo II del 76%, en cambio, si uno se preocupa por percibir un 80% de las veces un efecto de magnitud 4 (β=0,20) esto condiciona a que el valor alfa sea de 0,51.

En la Tabla 5.1 se evaluó los valores p de los efectos del ejemplo de Myers, Kim y Griffiths (1997), en él se separan los efectos que se deben considerar activos si deseamos detectar un *em* de 4 con un valor de β del 20%. Dado que tenemos un valor mayor de alfa que el empleado por Myers et al. se aceptaría como efectos significativos hasta la interacción X_1X_2Z, cinco términos más que en la situación convencional de alfa igual al 5%. Esto resulta lógico, ya que si se quiere una potencia grande (de 80% en lugar de 24%) para aceptar como significativo un efecto de solamente 4 unidades, se debe ser más laxo al momento de considerar que tan grande es un efecto para asumirlo activo.

Tabla 5.1 Valores p con PSE = 2,625.

Términos		Efecto	valor p
Considerados activos	Z	21,62	0,000
	X_2Z	-18,13	0,001
	X_3Z	16,63	0,001
	X_3	14,62	0,003
	X_2	9,87	0,013
	X_1X_3Z	4,13	0,176
	X_1	3,12	0,288
	$X_1X_2X_3$	-2,62	0,364
	X_1X_2	2,38	0,406
	X_1X_2Z	1,87	0,508
Considerados inertes	X_2X_3Z	-1,63	0,562
	$X_1X_2X_3Z$	1,37	0,624
	X_2X_3	-1,13	0,685
	X_1X_3	-0,37	0,893
	X_1Z	0,12	0,965

Una ventaja del gráfico Alfa-Beta es que presenta los valores de los efectos en valor absoluto que se considerarían significativos a partir del valor alfa o beta seleccionado, por lo cual no es necesario para este fin que el usuario conozca los correspondientes valores p. Por ejemplo, en la Figura 5.1 se aprecia que si se desea un nivel α de 0,05 entonces serán significativos los efectos con un valor absoluto superior a 6,7, si se acepta que α sea 0,51 (β = 0,20) lo serán los que tengan un valor superior a 1,8, Tabla 5.1.

Tabla 5.2 Alfas necesarias para tener la potencia requerida para un *em* = 4 y número de términos que se incluirían en el modelo. Se usó PSE = 2,625.

Potencia	Alfa	Términos incluidos
0,25	0,053	5
0,50	0,167	5
0,60	0,244	6
0,70	0,352	7
0,80	0,519	10
0,90	0,812	13
0,91	0,856	13
0,92	0,909	14
0,93	0,962	14
0,9363	1,000	15

Para poder detectar un efecto real de valor 4 con potencias superiores al 80% se requiere de valores muy grandes de α y por tanto de aceptar un gran número de efectos, muchos de los cuales seguramente inertes, como se aprecia en la Tabla 5.2. Surge un compromiso entre tener una potencia alta, con la posible inclusión de factores inertes, y tener un valor de α más tradicional. El problema de tener un modelo que no sea parsimonioso puede solventarse parcialmente eliminando términos a través del uso de criterios técnicos.

Una manera de eliminar el inconveniente de usar valores grandes de alfa es tener réplicas, por ejemplo para la misma estimación de la desviación estándar de los efectos, PSE = 2,625, el mismo valor del efecto mínimo de 4 y un valor de alfa del 5% se tiene una potencia de 0,848 con 4 réplicas y una potencia de 0,958 con 6. La opción de tener un valor pequeño para ambos tipos de error a través de un número adecuado de réplicas es claramente deseable, aunque lamentablemente impráctico en un gran número de aplicaciones en la industria.

Como se sabe, para un valor fijo de la potencia se puede contar con valores de alfa menores si disminuye el valor de la estimación de la desviación estándar del error, esto se puede lograr en ocasiones a través de un mejor proceso de estimación. Al aplicar el método de la regresión paso a paso con los valores por defecto de Minitab a los datos de Myers et al. (1997), y al usar el método de los mejores subconjuntos se pueden considerar en el modelo los términos X_2, X_3, Z, X_2Z, X_3Z y X_1X_3Z; con ello se tendría un valor de S = 3,76 y 9 grados de libertad. Esto disminuye la estimación de la desviación estándar de los efectos de 2,625 a 1,66.

El gráfico Alfa-Beta de la Figura 5.2 se construyó con la nueva estimación de la varianza, en él se aprecia que gracias a esta reducción si se toma el criterio de $\alpha = 0{,}05$ se produce una disminución en el valor de beta de 0,77 a 0,44. Si se opta por un valor de $\beta = 0{,}20$ únicamente se necesita que alfa sea del 16% en lugar del 52% y los efectos que se consideran significativos son los que tienen valores mayores a 2,6 en valor absoluto. Con este último criterio entrarían al modelo los términos indicados por la regresión paso a paso además de X_1 y X1X2X3, Tabla 5.3.

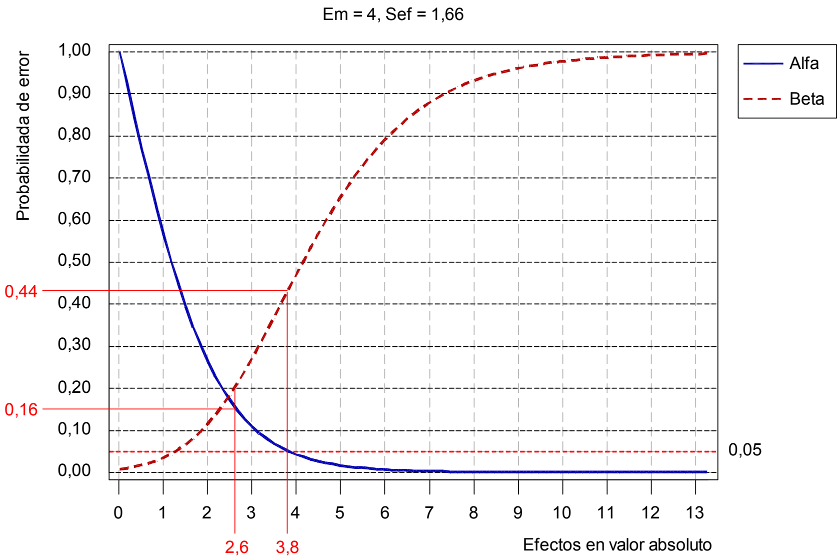

Figura 5.2 Gráfico Alfa-Beta para *em* = 4 con una estimación del error de los efectos de 1,66 y gl = 8.

Tabla 5.3 Valores p de los efectos con Sef = 1,66

Términos		Efecto	valor p
Considerados activos	Z	21,62	0,000
	X_2Z	-18,13	0,000
	X_3Z	16,63	0,000
	X_3	14,62	0,000
	X_2	9,87	0,000
	X1X3Z	4,13	0,038
	X_1	3,12	0,097
	$X_1X_2X_3$	-2,62	0,152
Considerados inertes	X_1X_2	2,38	0,190
	X_1X_2Z	1,87	0,291
	X_2X_3Z	-1,63	0,356
	$X_1X_2X_3Z$	1,37	0,432
	X_2X_3	-1,13	0,517
	X_1X_3	-0,37	0,827
	X_1Z	0,12	0,942

Es de esperar que si la magnitud de *em* que se desea detectar es mayor a 4 la posibilidad de cometer el error tipo II disminuya, los dos gráficos siguientes muestran que tanto se modifica, la estimación usada es PSE.

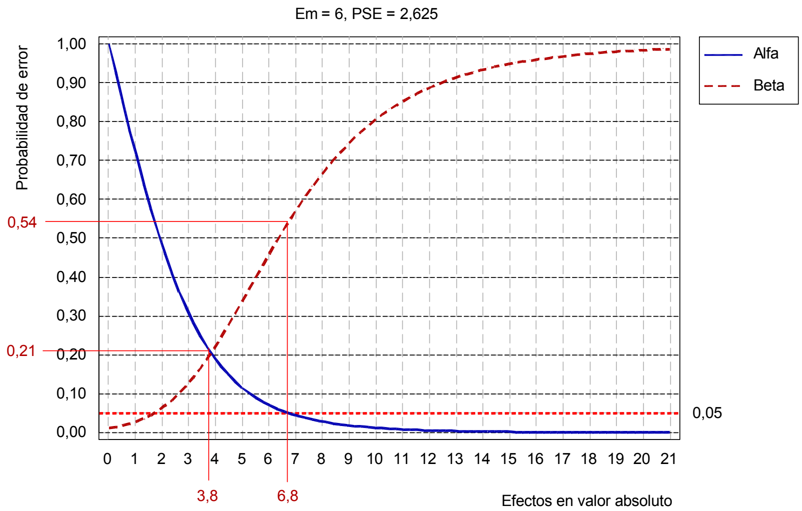

Figura 5.3 Gráfico Alfa-Beta para *em* = 6 con una estimación del error de los efectos de 2,625.

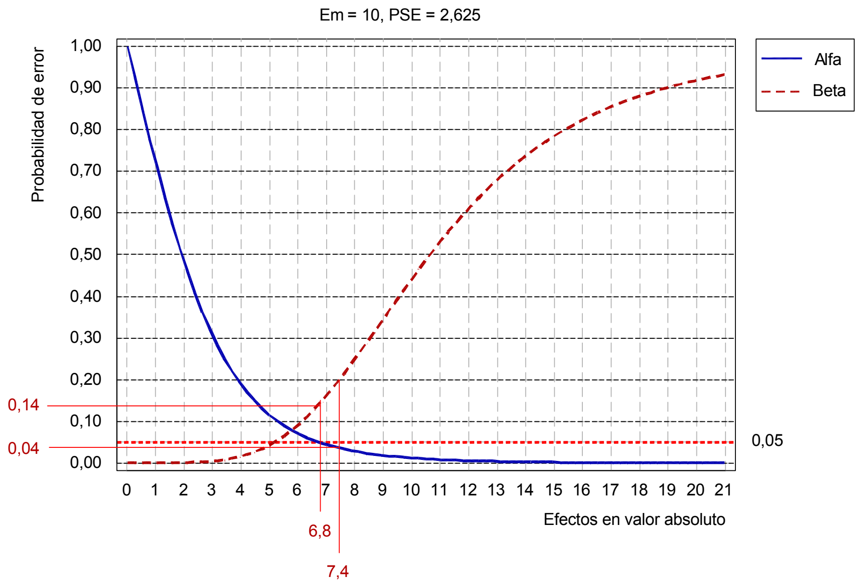

Figura 5.4 Gráfico Alfa-Beta para *em* = 10 con una estimación del error de los efectos de 2,625.

Si el valor de *em* es 6 entonces para α = 0,05 el valor de beta es de 0,54 en lugar de 0,77, disminuyendo hasta 0,14 si el valor del efecto es 10. Con la intención de tener una potencia del 80% el valor de alfa debe ser de 0,21 para el efecto mínimo de 6 y de 0,04 para el de 10.

El gráfico Alfa-Beta puede construirse con Minitab mediante una macro muy sencilla (ver Anexo 8). Como la curva correspondiente al riesgo α es independiente de la distancia significativa, se puede construir un gráfico que incluya las curvas de riesgo β para distintas distancias significativas, como en la Figura 5.5. De este gráfico se deduce que si *em* = 9 un valor crítico razonable podría ser igual a 5,25, con unos riesgos alfa y beta inferiores al 10 %. Sin embargo, si tenemos *em* = 5, la información que aporta el experimento no permite fijar un valor frontera con riesgos razonables.

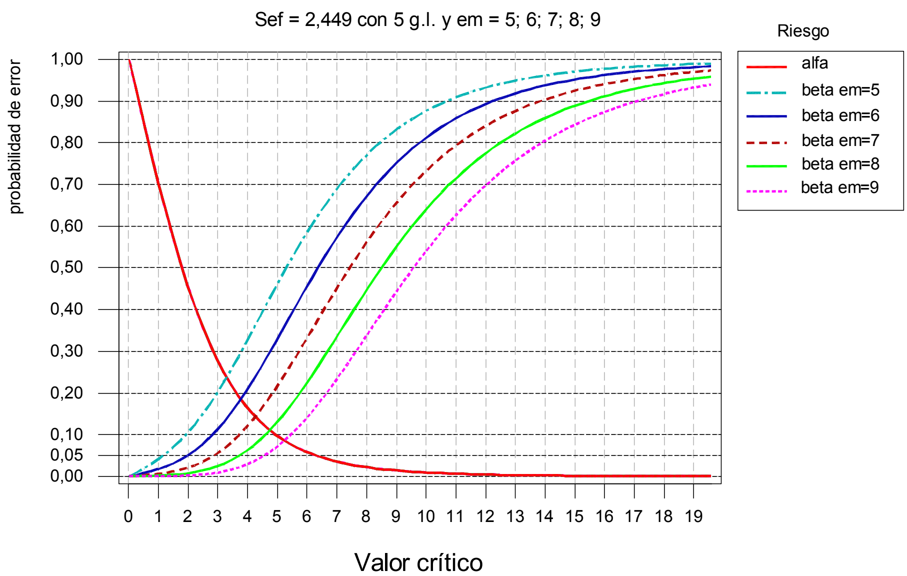

Figura 5.5 Gráfico Alfa-Beta donde se incluyen los riesgos beta para distintos valores del efecto mínimo.

5.3 El valor q

Si ya se estableció el valor deseado de *em* y se determinó los valores adecuados de α y β que se pueden asumir ¿qué se puede decir de la tasa de error tipo II cuando un efecto en concreto no se ha rechazado? Como se comentó anteriormente, una dificultad práctica para usar esta tasa estriba en sugerir el valor del parámetro de la hipótesis alternativa. Una determinación plausible para este valor es la propia estimación del efecto, con esta idea en mente se propone contar con un valor, *el valor q,* que sea dependiente de los datos y que contribuya a reflexionar en la tasa del error tipo II si el valor real del parámetro es el que señala su estimación.

Con base en el ejemplo de la resistencia al desgaste del capítulo 4 la Tabla 5.4 muestra los valores de los efectos estimados junto con los valores p y los valores q asociados. Para la obtención de estos valores se usó a PSE como estimación de la desviación estándar de los efectos.

Tabla 5.4 Efectos estimados de la resistencia al desgaste con los valores p y valores q asociados. Se usó $\alpha = 0{,}05$ y $S = 14{,}625$

Térm.	Efecto	val. p	val. q	Término	Efecto	val. p	val. q
B	58,75	0,002	0,048	ABCR	11,50	0,449	0,890
R	-57,25	0,003	0,058	ACRT	9,75	0,520	0,907
T	56,50	0,003	0,064	BT	-8,75	0,563	0,915
A	**-46,25**	**0,010**	**0,184**	BRT	-6,25	0,678	0,932
AR	32,75	0,048	0,472	AC	6,25	0,678	0,932
AB	-31,25	0,057	0,509	ABRT	6,00	0,690	0,934
BC	***21,25***	***0,176***	***0,738***	ABCT	5,25	0,727	0,938
CT	18,75	0,228	0,786	BCRT	4,00	0,790	0,943
ACR	18,75	0,228	0,786	ABC	-3,75	0,803	0,944
BCT	17,50	0,258	0,807	BR	-2,50	0,868	0,947
ACT	14,00	0,360	0,860	ABCRT	1,75	0,907	0,949
ART	14,00	0,360	0,860	ABR	-1,50	0,920	0,949
C	-13,75	0,369	0,863	CRT	-1,50	0,920	0,949
CR	-13,75	0,369	0,863	BCR	-1,00	0,947	0,950
ABT	-13,00	0,394	0,872	RT	-0,75	0,960	0,950
AT	11,75	0,440	0,887				

Los valores q corresponden, bajo el supuesto de normalidad, al valor β que se tiene cuando en la hipótesis alterna el efecto tiene el valor del efecto estimado. Por ejemplo, si para el efecto A se tiene como hipótesis alterna a H_1: $\tau_A = -46{,}25$ entonces el valor de β es 0,184 si $\alpha = 0{,}05$. A estos valores los hemos denotado valores q en analogía a los valores p y como ellos se determinan a posteriori con base en las estimaciones de los efectos.

El vínculo entre el valor p y la tasa de error tipo I está dado por la misma definición "El valor p para el punto muestral $\boldsymbol{x}$ es el menor valor de α para el cual este punto muestral conducirá al rechazo de H_0" (Casella, G. y Berger, R. (1990)). En tanto que la relación entre el valor q y la tasa de error tipo II es "El valor q para la estimación de un efecto es el valor de β si el valor propuesto en H_1 para el valor del efecto es la estimación, con un valor dado de α".

El uso del valor q claramente tiene sentido cuando no se rechaza la hipótesis nula. En el gráfico en papel probabilístico normal de la Figura 5.6 se observa un punto que no se marcó significativo y parece estar a igual distancia de la recta que el punto AD que si se marcó, es más este último punto se ve más cercano a los puntos que conforman la recta que el otro que no fue seleccionado, esto se aprecia más claramente en el diagrama de puntos. Este punto no marcado concierne al efecto de la interacción AB y su distancia a cero, 31,25, es bastante semejante a la de la interacción AD, 32,75, que si se marcó. El valor p de la interacción no considerada AB es de 5,7% y por tal razón no se marcó.

Si se considera inerte el efecto AB y se desea tener una idea de la magnitud del error que acarrea esta decisión en el caso de que no lo sea, entonces un valor razonable a sugerir para el posible valor real del efecto lo constituye la estimación del mismo. De esta manera, al calcular la tasa de error con dicho valor lo que se obtiene es su valor q, este valor resultas ser excesivamente grande (0,51) lo cual conforma un criterio más para considerar la inclusión de la interacción en el modelo.

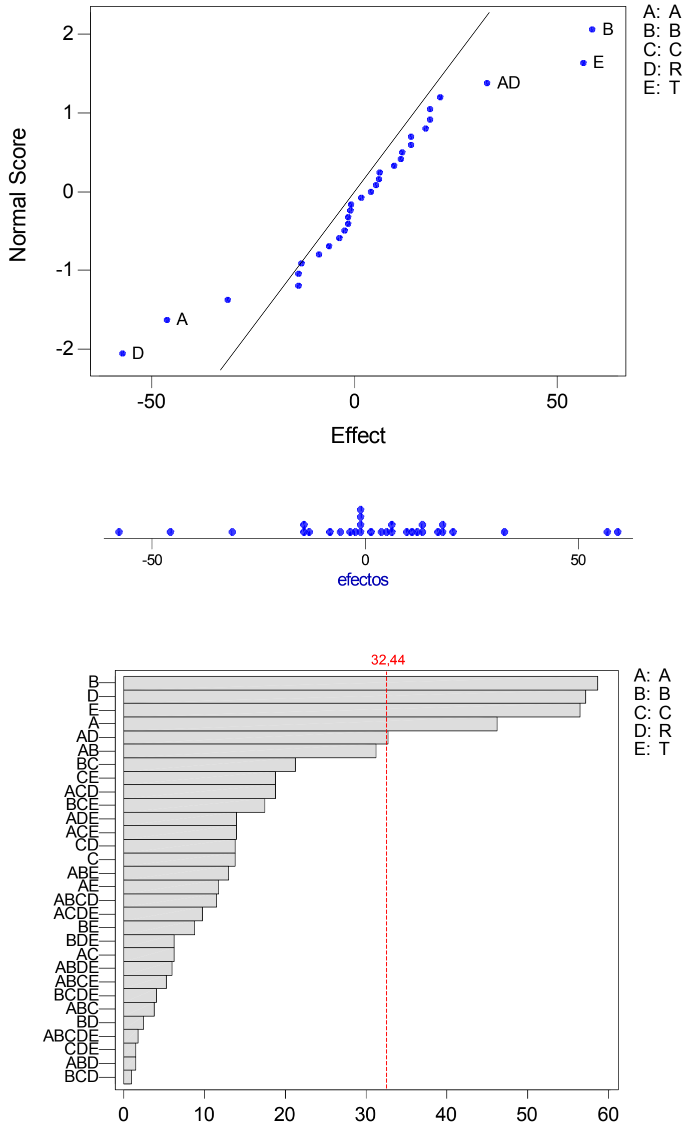

Figura 5.6 Representación de los efectos para el problema de la resistencia al desgaste.

La Figura 5.7 ilustra la relación entre el valor p y el valor q para la estimación del efecto de la interacción BC (en cursiva en la Tabla 5.4), se considera un valor de $\alpha = 0{,}05$ para la determinación de las región de aceptación y del valor q. La cantidad 21,25 corresponde a la estimación del efecto y en la Figura 5.7 está sin estandarizar para simplificar la interpretación. El valor q es función del valor del efecto estimado, del nivel de significación α y de la estimación del la desviación estándar de los efectos.

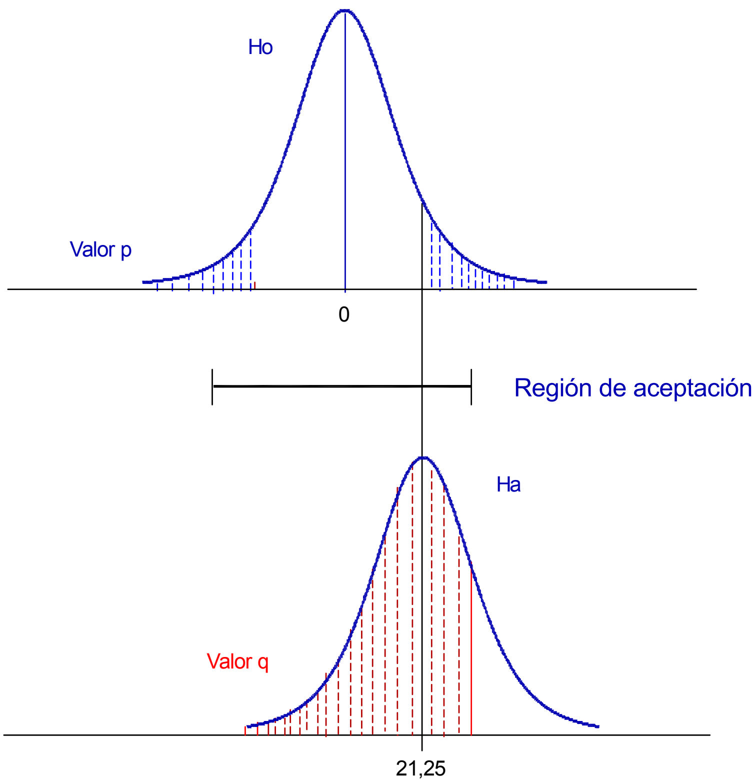

Figura 5.7 Relación entre el valor p y el valor q de la interacción BC**.** El valor de BC igual a 21,25 está sin estandarizar.

Con los datos de la tasa de filtración que se trató en el capítulo 4 se ha construido el gráfico en papel probabilístico normal y el diagrama de puntos de los efectos; en la Figura 5.8, Myers, Kim y Griffiths (1997) indican que la presión (X_1) no impacta la tasa de filtración y que el modelo de la respuesta puede ser:

$$\hat{y} = 70,06 + 4,94x_2 + 7,31x_3 + 10,81z - 9,06x_2z + 8,31x_3z \qquad (5.1)$$

La Tabla 5.5 contiene los valores de los efectos estimados así como los valores p y los valores q asociados; para calcular estos valores se usó como estimador de la desviación estándar de los efectos el PSE.

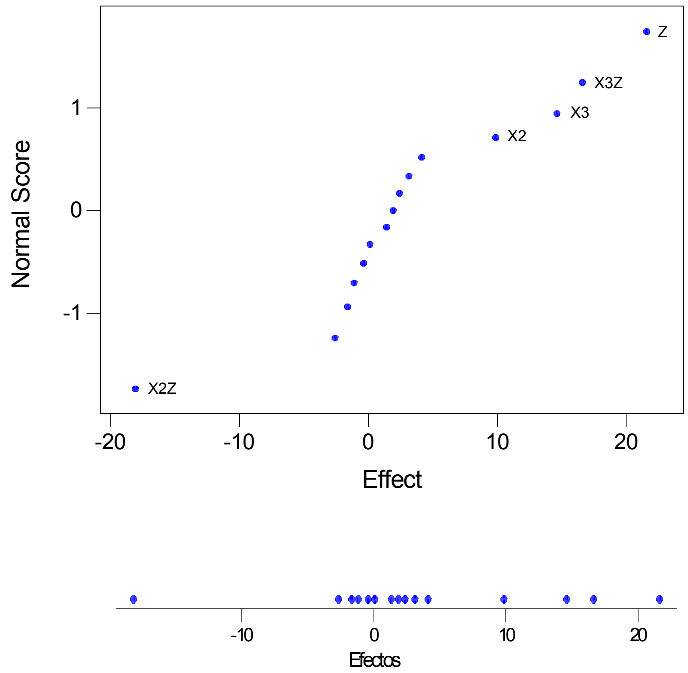

Figura 5.8 Representación de los efectos en papel probabilístico normal y en diagrama de puntos.

La variable presión que se descartó para el modelo tiene un efecto estimado de 3,12, lo que genera un valor p de 0,29 y un valor q de 0,84, claramente el valor p es demasiado grande como para pensar en incluir esta variable, no obstante si se tiene dudas sobre su inclusión entonces el valor q nos alerta de la tasa tan grande que se tiene de cometer el error tipo II, en las condiciones indicadas.

Tabla 5.5 Efectos de la tasa de filtración con sus valores p y sus valores q calculados a partir de PSE = 2,625 y α = 0,05.

Términos		Efecto	valor p	valor q
Considerados activos	Z	21,62	0,000	0,000
	X_2Z	-18,13	0,001	0,000
	X_3Z	16,63	0,001	0,002
	X_3	14,62	0,003	0,009
	X_2	9,87	0,013	0,152
Considerados inertes	X_1X_3Z	4,13	0,177	0,750
	X_1	3,12	0,287	0,836
	$X_1X_2X_3$	-2,62	0,363	0,870
	X_1X_2	2,38	0,407	0,884
	X_1X_2Z	1,87	0,507	0,910
	X_2X_3Z	-1,63	0,563	0,920
	$X_1X_2X_3Z$	1,37	0,623	0,929
	X_2X_3	-1,13	0,686	0,935
	X_1X_3	-0,37	0,892	0,948
	X_1Z	0,12	0,964	0,950

La representación en un diagrama de puntos de los valores de los efectos estimados se da en la Figura 5.9; en ella se delimita con las líneas sólidas la zona de aceptación de la hipótesis nula, que corresponde al intervalo [-6,75, 6,75], los puntos corresponden a los efectos.

Como se observa en la Figura 5.9, para un valor de α fijo, al aumentar el valor absoluto de un efecto disminuye su valor p y también su valor q, aunque no de forma equivalente. Todos los efectos seleccionados para el modelo por Myers et al. tienen un valor p inferior al 5%, y por tanto los valores q deben tener valores pequeños, aún así el factor X2 con un efecto significativo de 9,87 aunque tiene un valor p bastante pequeño de 0,013 tiene un valor q cinco veces mayor.

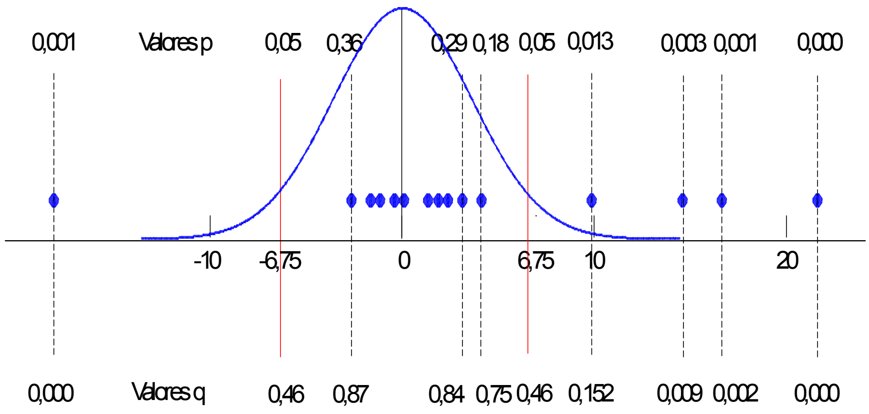

Figura 5.9 Comportamiento del valor p y el valor q, el valor de alfa es 5% y el estimador de la desviación estándar de los efectos es PSE. Los valores en la parte superior corresponden a los valores p, los de la parte media a los efectos sin estandarizar y en la parte inferior se encuentran los valores q.

Esta disparidad entre el valor p y el valor q se aprecia en la Figura 5.10, en ella se ve que la interacción X_2Z con un valor de -18,13 tiene el mayor efecto negativo estimado, y que tanto el valor p como el q son prácticamente cero, en cambio para el mayor efecto positivo no significativo, asociado a la interacción X_1X_3Z, el valor p es de 0,177 y el valor q es de 0,750.

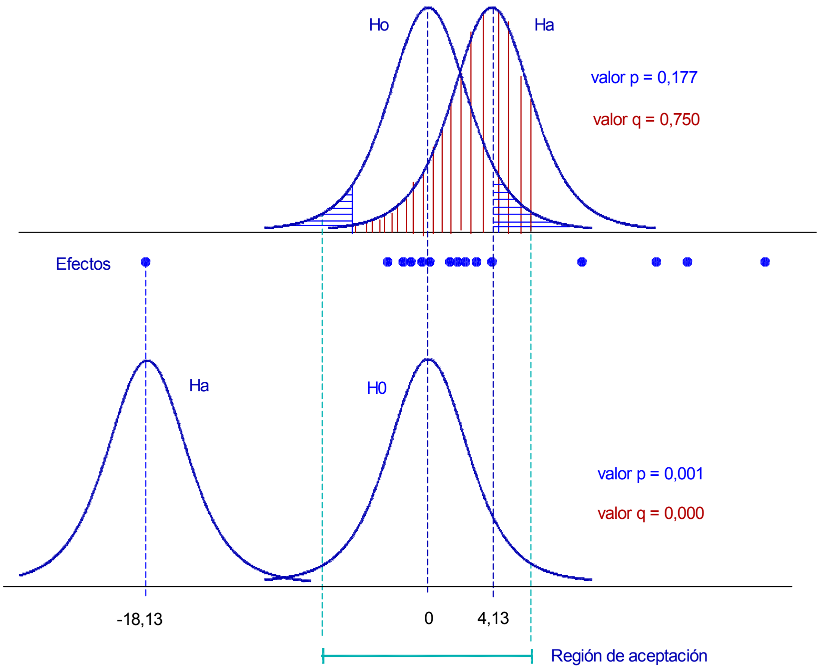

Figura 5.10 Comportamiento desigual entre el valor p y el valor q. La curva de la izquierda está centrada en X_2Z y la de la derecha en X_1X_3Z. El área con líneas verticales está asociada al valor q y con las líneas horizontales al valor p.

La potencia de la prueba es un concepto asociado a la tasa de error tipo II, que resulta más fácil de interpretar para algunos investigadores. Dado que el valor q corresponde a la tasa de error tipo II cuando en la hipótesis alterna el efecto tiene el valor del efecto estimado, entonces uno menos el valor q será la potencia relacionada con esta hipótesis alterna.

La Figura 5.11 muestra el comportamiento del valor p y la potencia así considerada para los efectos estimados de la tasa de filtración, este comportamiento como es de esperar resulta muy desigual, si el efecto toma valores absolutos entre 6,75 y 20 su valor p solo disminuye de 0,05 a 0,0, en cambio su potencia aumenta de 0,713 a 1,0.

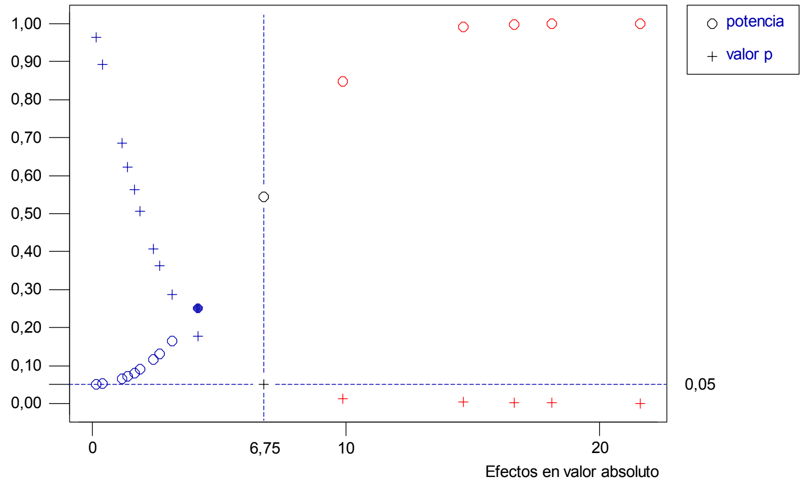

Figura 5.11 Tendencia del valor p y la potencia de la tasa de filtración. Los símbolos a la derecha de la línea vertical están asociados a los efectos considerados significativos. El punto sólido es la potencia de $X_1X_3Z = 4{,}13$ (el mayor efecto no considerado significativo)

El gráfico Alfa-Beta y los valores q son herramientas que coadyuvan en la elección de los efectos significativos, fomentan la percepción del error tipo II en la toma de decisiones, modifican la predisposición a asignar rutinariamente un nivel de significación del cinco por ciento y por su simplicidad de cálculo y elaboración representan una herramienta muy accesible y fácil de comprender y de aplicar.

5.4 Conclusiones

Con base en lo comentado se considera que:

- Es conveniente conocer los errores tipo II que se cometen en el análisis de la significación de los efectos.

- El ingeniero debe tener presente el efecto mínimo, *em*, que es relevante para su estudio y la tasa de error de tipo II (o potencia) que está dispuesto a asumir.

- El valor de la tasa de error tipo I por consecuencia se debe condicionar al valor de beta.

- El uso de los valores q y del gráfico Alfa-Beta complementan el proceso de discriminación de los efectos significativos.

- Si consideramos el valor beta en nuestros criterios de selección, es probable que alfa no tome un valor pequeño y por tanto se incluyan posibles efectos inertes como significativos, por esta razón es conveniente usar criterios técnicos adicionales para eliminarlos.

5.6 Aportaciones

- Se elaboró el gráfico Alfa-Beta que permite conocer la tasa de error tipo II cuando se desea detectar un valor específico en los efectos, con un nivel de significación dado. De forma equivalente, provee el valor de alfa que se debe usar para elegir los efectos significativos cuando se desea tener una potencia determinada. El gráfico presenta los valores de los efectos, en valor absoluto, que se considerarán significativos a partir del valor alfa o beta seleccionado, por lo cual no es necesario para este fin que el usuario conozca los correspondientes valores p.

- Se ha construido el valor q para evidenciar la posibilidad del error tipo II y ayudar en el proceso de selección de los efectos que se considerarán significativos.