

CONTROL DE TAMAÑOS Y POTENCIAS EN PRUEBAS DE HIPÓTESIS

E. CASTILLO, A. LUCEÑO, E. MORA, J. RODRÍGUEZ

En este trabajo se analizan los grados de libertad de que dispone el investigador para fijar o acotar los tamaños de primera y segunda especie y las potencias garantizadas en aceptación y rechazo, asociados a una prueba óptima de un contraste de hipótesis. En especial se estudia la influencia del tamaño de muestra fijo o libre, la proximidad de las hipótesis nula y alternativa y el dar al investigador una tercera opción consistente en dudar (no decidirse ni por las hipótesis nula ni por la alternativa).

Finalmente se incluyen tablas que muestran qué tamaños o potencias garantizadas, o cotas de los mismos, pueden darse independientemente y las restricciones a que están sometidos los restantes parámetros.

1. INTRODUCCION

El problema de los contrastes se presenta al investigador /1,2/ como una elección entre dos hipótesis incompatibles y exhaustivas H_0 y H_1 , llamadas nula y alternativa respectivamente. Para decidir entre estas dos hipótesis se dispone de una cierta información a la vista de la cual el investigador puede inclinarse por: Rechazar H_0 (aceptar H_1) o aceptar H_0 (rechazar H_1). Combinando las dos posibilidades: certeza de H_0 o certeza de H_1

con estas dos posibles decisiones, se obtienen las cuatro situaciones que se muestran en la Tabla 1, dos de las cuales conducen a una decisión correcta y las otras dos a una decisión falsa. Cada prueba o criterio de decisión que el investigador elija, lleva asociadas unas probabilidades para cada una de esas cuatro situaciones. Lo ideal sería diseñar una prueba que permitiera tomar siempre decisiones correctas, es decir, que las probabilidades fueran 1 para las situaciones correctas y 0 para las erróneas. Sin embargo,

Tabla 1 - Riesgos de las decisiones en pruebas de hipótesis

HIPOTESIS CIERTA	DECISION	PROBABILIDAD DE ADOPTAR ESA DECISION BAJO LA HIPOTESIS CIERTA Y EL TEST ELEGIDO	SUPREMOS DE LOS RIESGOS E INFIMOS DE LAS POTENCIAS BAJO LAS HIPOTESIS CIERTAS PARA EL TEST ELEGIDO
H_0	RECHAZAR H_0 (decisión errónea)	RIESGO DE PRIMERA ESPECIE	TAMAÑO DE PRIMERA ESPECIE (α)
	ACEPTAR H_0 (decisión correcta)	POTENCIA EN ACEPTACION	POTENCIA GARANTIZADA EN ACEPTACION (p_a)
H_1	RECHAZAR H_0 (decisión correcta)	POTENCIA EN RECHAZO	POTENCIA GARANTIZADA EN RECHAZO (p_r)
	ACEPTAR H_0 (decisión errónea)	RIESGO DE SEGUNDA ESPECIE	TAMAÑO DE SEGUNDA ESPECIE (β)

E. Castillo, A. Luceño, E. Mora, J. Rodríguez. Escuela de Ingenieros de Caminos, Canales y Puertos. Departamento de Matemáticas aplicadas a la Ingeniería. Av. de Castros, s/n. Santander

Article rebut el juny de 1981.

como la información disponible no es total, hay que contentarse con maximizar las probabilidades de decidir correctamente, que reciben el nombre de potencias, y minimizar las de decidir erróneamente, que reciben el nombre de riesgos (ver columna 3 de la tabla 1)

Además, si las hipótesis son compuestas, pueden existir más de un riesgo y una potencia de cada tipo, por lo que se introducen los conceptos de tamaños y potencias garantizadas, que respectivamente acotan superior e inferiormente los riesgos y las potencias (columna 4 de la tabla 1). De esta forma, acotando tamaños y potencias garantizadas, quedan automáticamente acotados los riesgos y las potencias.

Sería de desear que el investigador pudiese controlar independientemente los cuatro valores (dos tamaños y dos potencias garantizadas) o al menos acotarlos independientemente. Sin embargo, ello no siempre es posible. En este trabajo va a analizarse qué restricciones se oponen a dicho deseo y en qué medida afectan a estas restricciones el tener un tamaño de muestra fijo o libre, la proximidad de las hipótesis H_0 y H_1 y el dar al investigador una tercera opción que es la de dudar (no decidirse ni por H_0 ni por H_1).

2. CONTRASTES FISCALMENTE MAL PLANTEADOS

En muchos casos prácticos se plantean contrastes cuyas hipótesis nula, H_0 , y alternativa, H_1 , están infinitamente próximas, en el sentido de que pueden encontrarse dos hipótesis simples, pertenecientes una a cada una de aquellas, prácticamente iguales; así, por ejemplo, al contrastar la hipótesis nula de que la media μ de una variable aleatoria normal, con varianza 1, es menor que 4 frente a la alternativa de que es mayor o igual que 4, se pueden encontrar las hipótesis $\mu=4-\epsilon$ ($\epsilon>0$) y $\mu=4$, pertenecientes respectivamente a H_0 y H_1 , que, con tal de elegir ϵ suficientemente pequeño, están tan próximas como se desee. Intuitivamente se comprende que resulta prácticamente imposible su discernimiento, lo cual, matemáticamente, se traduce en que, para la prueba óptima, la potencia garantizada en rechazo, p_r , coincide con el tamaño de primera especie de la prueba, α , y en que la potencia garantizada en acepta-

ción, p_a , coincide con el tamaño de segunda especie, β , es decir:

$$p_r = \alpha \quad (1)$$

$$p_a = \beta \quad (2)$$

lo cual implica que α y p_r por una parte, y β y p_a por otra, no pueden elegirse independientemente, con el agravante de que estas pruebas tienen forzosamente o tamaños grandes o potencias garantizadas pequeñas, ya que al mejorar los tamaños se empeoran las potencias garantizadas y viceversa.

Estas consideraciones conducen a la definición siguiente: Se dice que un contraste está físicamente mal planteado si sus pruebas óptimas en un cierto sentido (maximin, uniformemente más potentes, etc.) verifican las relaciones (1) y (2).

Es digno de ser destacado el hecho de que la mayor parte de las pruebas utilizadas en la práctica son de este tipo.

Volviendo al ejemplo anterior, una forma de conseguir un contraste no físicamente mal planteado consiste en separar las hipótesis nula y alternativa; así, podría cambiarse la hipótesis nula, H_0 , por $\mu>4-\epsilon$, donde $\epsilon>0$ es un valor que, según el criterio del investigador, da lugar a hipótesis claramente diferenciadas. Es inmediato comprobar que la prueba óptima no está ya sometida a las restricciones (1) y (2), y que pueden mejorarse simultáneamente los tamaños y las potencias garantizadas sin más que aumentar el tamaño de la muestra.

3. PRUEBAS CON OPCION DE DUDA

Las pruebas de hipótesis que normalmente se utilizan en la práctica plantean dos opciones: rechazar H_0 y no rechazar H_0 . Sin embargo, en muchas ocasiones se identifica no rechazar H_0 con aceptar H_0 , lo cual lleva a que se verifiquen las relaciones:

$$\alpha + p_a = 1 \quad (3)$$

$$\beta + p_r = 1 \quad (4)$$

por lo que no pueden elegirse independientemente α y p_a por un lado y β y p_r por otro. Ahora bien, estas restricciones son menos graves que las (1) y (2), puesto que al mejorar los tamaños mejoran simultáneamente las potencias garantizadas y viceversa. Por

tanto, se pueden conseguir, al mismo tiempo, tamaños pequeños y potencias garantizadas -- grandes.

Una forma de eliminar las restricciones (3) y (4) es mediante la inclusión de una tercera opción, la de dudar (no decidirse ni por H_0 ni por H_1), que será elegida por el investigador cuando los datos no suministren suficiente información para inclinarse por H_0 o por H_1 con los tamaños y potencias garantizadas deseados. Esta opción es la que fundamenta los tests o pruebas secuenciales /3/.

Así pues, las pruebas con opción de duda no identifican el no rechazo de H_0 con la aceptación de H_0 y permiten las tres opciones:

- a) Rechazar H_0 , que equivale a aceptar H_1
- b) Aceptar H_0 , que equivale a rechazar H_1
- c) Dudar

De esta forma, las ecuaciones (3) y (4) se transforman en las inecuaciones:

$$\alpha + p_a \leq 1 \quad (3')$$

$$\beta + p_r \leq 1 \quad (4')$$

4. RESTRICCIONES IMPUESTAS POR LAS PRUEBAS OPTIMAS

4.1 TAMAÑO DE MUESTRA FIJO

Cuando se trabaja con pruebas óptimas (uniformemente más potentes, maximin, etc.) y -- con tamaño de muestra, n, fijo, aparecen restricciones del tipo:

$$\alpha = g_n(p_r) \quad (5)$$

$$\beta = h_n(p_a) \quad (6)$$

Así, para el ejemplo de contraste propuesto: $H_0: \mu < 4 - e; \sigma = 1$ frente a $H_1: \mu \geq 4; \sigma = 1$ la prueba uniformemente más potente viene definida por (figura 1):

$$\left. \begin{array}{l} \text{si } \bar{x} \geq K_2 \text{ se rechaza } H_0 \\ \text{si } \bar{x} \leq K_1 \text{ se acepta } H_0 \\ \text{si } K_1 < \bar{x} < K_2 \text{ se duda} \end{array} \right\} \quad (7)$$

donde \bar{x} es la media muestral y K_1 y K_2 están ligados con los tamaños y potencias garantizadas mediante las relaciones:

$$\alpha = 1 - F_{N(4-e, 1/n)}(K_2) \quad (8)$$

$$\beta = F_{N(4, 1/n)}(K_1) \quad (9)$$

$$p_r = 1 - F_{N(4, 1/n)}(K_2) \quad (10)$$

$$p_a = F_{N(4-e, 1/n)}(K_1) \quad (11)$$

donde $F_{N(\mu, 1/n)}(x)$ es la función de distribución de la variable aleatoria normal de media μ y varianza $1/n$.

Eliminando K_1 y K_2 en (8), (9), (10) y (11) resultan:

$$\alpha = 1 - F_{N(4-e, 1/n)}(F_{N(4, 1/n)}^{-1}(1 - p_r)) \quad (12)$$

$$\beta = F_{N(4, 1/n)}(F_{N(4-e, 1/n)}^{-1}(p_a)) \quad (13)$$

que son las relaciones del tipo (5) y (6) -- para este contraste, y que se ilustran en la figura 2.

En el caso general, dada una prueba, es decir, dados los criterios para aceptación, -- rechazo y duda, quedan unívocamente determinados α , p_r , β y p_a . Si, en vez de una prueba, se da una familia de pruebas cuyo criterio de rechazo de H_0 depende de un parámetro (α o p_r) y cuyo criterio de aceptación de H_0 depende de otro parámetro (β o p_a), -- como ocurre con la familia de las pruebas -- óptimas con tamaños α y β dados, existirán dos relaciones del tipo (5) y (6).

Además, las funciones $h_n(x)$ y $g_n(x)$ que aparecen en (5) y (6) están ligadas mediante:

$$1 - x = g_n(1 - h_n(x)) \quad 0 \leq x \leq 1 \quad (14)$$

que también puede ponerse en la forma:

$$1 - x = h_n(1 - g_n(x)) \quad 0 \leq x \leq 1 \quad (15)$$

Es trivial comprobar, sin más que utilizar las expresiones (12) y (13), de las que resultan

$$g_n(x) = 1 - F_{N(4-e, 1/n)}(F_{N(4, 1/n)}^{-1}(1 - x)) \quad (16)$$

$$h_n(x) = F_{N(4, 1/n)}(F_{N(4-e, 1/n)}^{-1}(x)) \quad (17)$$

que se verifican estas dos relaciones en el caso del ejemplo.

La demostración de que las relaciones (14) y (15) son ciertas en el caso de pruebas -- sin región de duda resulta de considerar -- las ecuaciones (3), (4), (5) y (6). En efecto, despejando α y p_r de (3) y (4) respectivamente e introduciéndolas en (5) resulta:

$$1 - p_a = g_n(1 - \beta) \quad (18)$$

y sustituyendo el valor de β dado por (6) -- resulta finalmente

$$1 - p_a = g_n(1 - h_n(p_a)) \quad 0 \leq p_a \leq 1 \quad (19)$$

que es la expresión (14).

DIAGRAMA 1

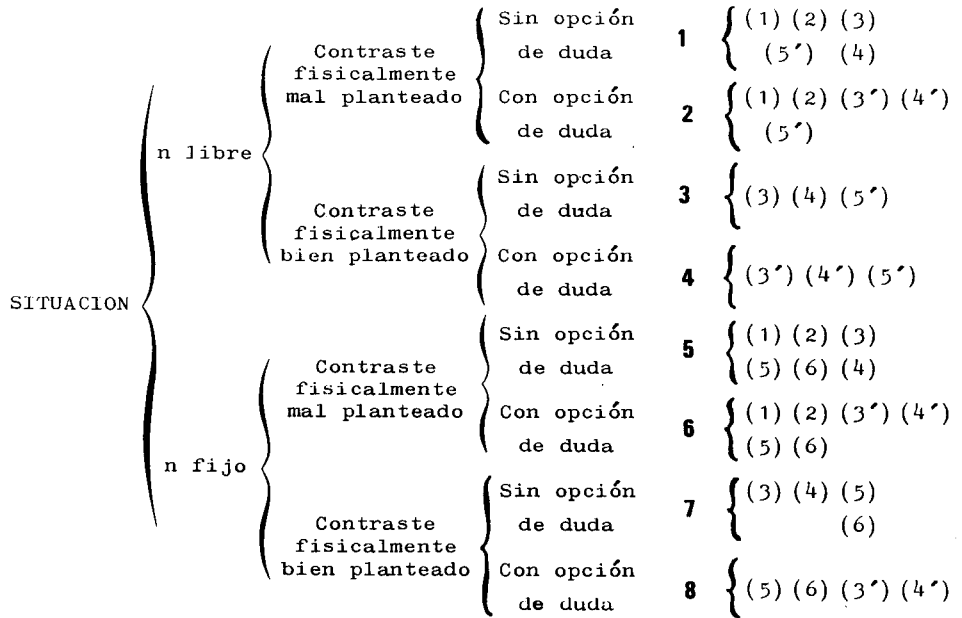


TABLA 2

SITUACIONES	PARAMETROS QUE PUEDEN FIJARSE LIBREMENTE	RESTRICCIONES PARA EL RESTO DE PARAMETROS
1 5	uno cualquiera de α, β, p_r y p_a	$p_r = \alpha, p_a = \beta, \alpha + p_a = 1$
2 6	uno cualquiera de α, β, p_r y p_a	$p_r = \alpha, p_a = \beta, \alpha + p_a \leq 1$
3	uno cualquiera de α, β, p_r y p_a	$\alpha + p_a = 1, \beta + p_r = 1$ $(\alpha, \beta) \in \{ (g_n(p_r), h_n(p_a)), n \in \mathbb{N} \}$
4	uno cualquiera de α, β, p_r y p_a	$\alpha + p_a \leq 1, \beta + p_r \leq 1$ $(\alpha, \beta) \in \{ (g_n(p_r), h_n(p_a)), n \in \mathbb{N} \}$
7	uno cualquiera de α, β, p_r y p_a	$\alpha + p_a = 1, \beta + p_r = 1$ $\alpha = g_n(p_r)$
8	uno cualquiera de α, β, p_r y p_a	$\alpha + p_a \leq 1, \beta + p_r \leq 1$ $\alpha = g_n(p_r), \beta = h_n(p_a)$

TABLA 3

SITUACION	COTAS QUE PUEDEN FIJARSE INDEPENDIEMENTE	RESTRICCIONES PARA EL RESTO DE COTAS
1 5	Cualquiera de las parejas (α^*, p_a^*) o (β^*, p_r^*)	$\alpha^* \geq p_r^*$; $\beta^* \geq p_a^*$ $p_r^* + p_a^* \leq 1$; $\alpha^* + \beta^* \geq 1$
2 6	Cualquiera de las parejas (α^*, p_a^*) , (β^*, p_r^*) , (α^*, β^*)	$\alpha^* \geq p_r^*$; $\beta^* \geq p_a^*$ $p_a^* + p_r^* \leq 1$
3 4	α^* , β^* , p_r^* y p_a^*	—
7	Cualquiera de las parejas (α^*, p_a^*) o (β^*, p_r^*)	$\alpha^* \geq \varepsilon_n(p_r^*)$; $\beta^* \geq h_n(p_a^*)$ $\alpha^* \geq \varepsilon_n(1-\beta^*)$; $\varepsilon_n(p_r^*) + p_a^* \leq 1$
8	Cualquiera de las parejas (α^*, p_a^*) , (β^*, p_r^*) , (α^*, β^*)	$\alpha^* \geq \varepsilon_n(p_r^*)$; $\beta^* \geq h_n(p_a^*)$ $\varepsilon_n(p_r^*) + p_a^* \leq 1$

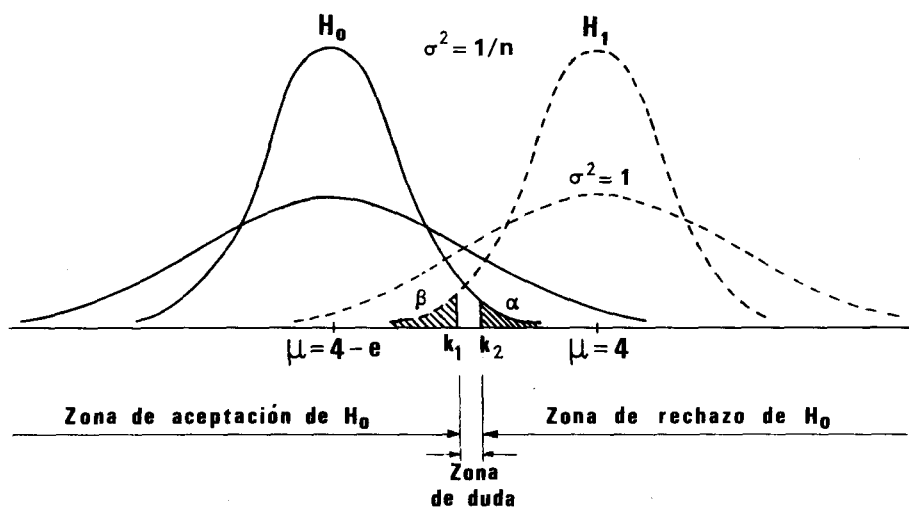


Figura 1

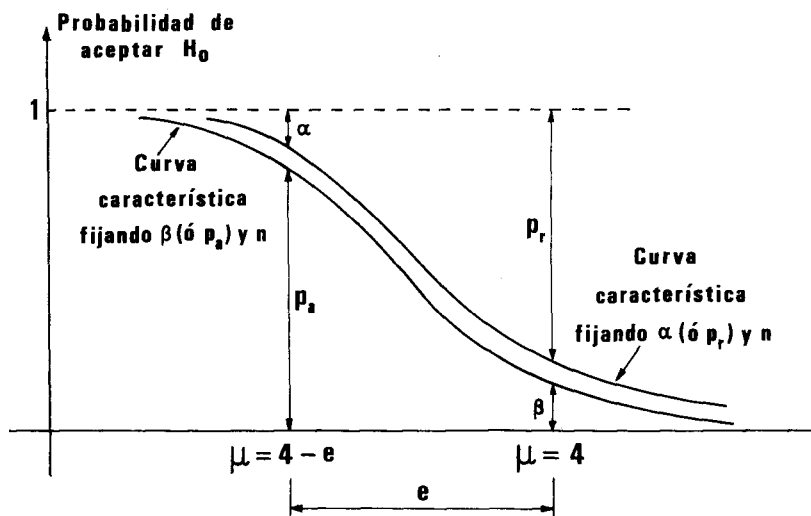


Figura 2

Análogamente, despejando β y p_a de (4) y (3) introduciéndolos en (6) y sustituyendo el valor de α de (5) se obtiene

$$1-p_r = h_n(1-g_n(p_r)), \quad 0 \leq p_r < 1 \quad (20)$$

que es la expresión (15).

Para el caso general, es decir, cuando se admite opción de duda, las relaciones (14) y (15) siguen siendo válidas, ya que la familia de las pruebas óptimas no depende de que se permita o no, esta opción.

Las restricciones (5) y (6) son normalmente mucho más débiles que las (1) y (2), especialmente para valores de n grandes.

4.2 TAMAÑO DE MUESTRA LIBRE

Si el tamaño de muestra, en vez de ser fijo, es libre, debe existir un tamaño de muestra n para el que se verifiquen las restricciones (5) y (6). Ello conduce a que los parámetros α , β , p_r y p_a están ligados en forma discreta mediante la relación

$$\exists n \text{ tal que } \alpha = g_n(p_r), \beta = h_n(p_a) \quad (5')$$

Esta relación significa que puede elegirse libremente un parámetro de cada pareja (α , p_r) y (β , p_a), y los dos parámetros restantes dentro de una sucesión del tipo (A_n, B_n) resultante de las ecuaciones (5) y (6). Así, por ejemplo, si se eligen α y β , la pareja (p_r, p_a) debe ser un término de la sucesión $(g_n^{-1}(\alpha), h_n^{-1}(\beta))$; si se elige α y p_a , la pareja (p_r, β) debe ser un término de la sucesión $(g_n^{-1}(\alpha), h_n(p_a))$ y así con el resto de posibilidades.

5. ELECCION INDEPENDIENTE DE TAMAÑOS Y POTENCIAS GARANTIZADAS

Para el investigador sería de mucho interés el disponer de la posibilidad de fijar a su libre elección los dos tamaños y las dos potencias garantizadas de la prueba que desea utilizar ante un contraste dado. En este apartado se analiza esta posibilidad en los casos que resultan de combinar factorialmente cada una de las siguientes situaciones:

- Tamaño de muestra fijo o libre
- Contrastes físicamente bien o mal planteados
- Pruebas con y sin opción de duda

En el Diagrama 1 se muestra el árbol resultante de estas combinaciones, que se numeran de 1 a 8, así como los números de las ecuaciones que restringen la libertad de elección de α , β , p_a y p_r en cada uno de los casos. Las ecuaciones que aparecen en las líneas inferiores de cada caso son combinaciones de las ecuaciones de las líneas superiores y por tanto no añaden restricción alguna. La Tabla 2 muestra las combinaciones de los parámetros α , β , p_r y p_a que pueden fijarse independientemente en cada caso, así como las restricciones que resultan para el resto de ellos.

De la Tabla 2 pueden sacarse las siguientes conclusiones:

- Sólo hay libertad total de elección para uno de los parámetros α , β , p_a y p_r , y eso ocurre en las ocho situaciones estudiadas.
- Los contrastes físicamente mal planteados imponen restricciones más fuertes que los bien planteados.
- El tamaño de muestra fijo impone restricciones más fuertes que el tamaño libre, salvo en el caso de los contrastes físicamente mal planteados (los casos 1 y 5 por un lado y los 2 y 6 por otro, resultan idénticos).
- La inclusión de la opción de duda da lugar a restricciones más débiles.

6. ELECCION INDEPENDIENTE DE COTAS PARA TAMAÑOS Y POTENCIAS GARANTIZADAS

En muchas ocasiones el investigador no necesitará trabajar con valores concretos de tamaños y potencias garantizadas, que fue el problema tratado en el apartado anterior, y se dará por satisfecho si puede acotar superiormente los tamaños e inferiormente las potencias y si, además, estas cotas pueden elegirlas independientemente.

En lo que sigue se analizan las combinaciones de las cotas α^* , β^* , p_a^* y p_r^* de los parámetros α , β , p_a y p_r , respectivamente, que pueden fijarse independientemente.

Las ecuaciones (1), (2), (3), (4), (3'), (4'), (5), (5') y (6) conducen, respectivamente, a las acotaciones:

$$\alpha^* > \alpha = p_r > p_r^* \quad (21)$$

$$\beta^* \geq \beta = p_a \geq p_a^* \quad (22)$$

$$\alpha^* \geq \alpha = 1 - p_a \leq 1 - p_a^* \quad (23)$$

$$\beta^* \geq \beta = 1 - p_r \leq 1 - p_r^* \quad (24)$$

$$\alpha^* \geq \alpha < 1 - p_a \leq 1 - p_a^* \quad (23')$$

$$\beta^* \geq \beta < 1 - p_r \leq 1 - p_r^* \quad (24')$$

$$\alpha^* \geq \alpha = g_n(p_r) \geq g_n(p_r^*) \quad (25)$$

$$\exists n \text{ tal que } \alpha^* \geq \alpha = g_n(p_r) \geq g_n(p_r^*) \\ \beta^* \geq \beta = h_n(p_a) \geq h_n(p_a^*) \quad (25')$$

$$\beta^* \geq \beta = h_n(p_a) \geq h_n(p_a^*) \quad (26)$$

donde se ha tenido en cuenta que las funciones $g_n(x)$ y $h_n(x)$ son no decrecientes.

La Tabla 3 muestra las combinaciones de las cotas α^* , β^* , p_r^* y p_a^* que pueden fijarse independientemente, así como las restricciones impuestas para el resto de ellas.

De la Tabla 3 pueden sacarse las siguientes conclusiones:

- Al menos pueden fijarse independientemente dos cotas de los parámetros α , β , p_r y p_a , pero, en general, no pueden ser -- dos cualesquiera.
- Sólo en el caso de tamaño de muestra libre y contrastes físicamente bien planteados pueden acotarse independientemente las cuatro cotas α^* , β^* , p_r^* y p_a^* .
- La opción de duda disminuye el número de restricciones excepto cuando el tamaño de la muestra es libre y el contraste es físicamente bien planteado. En este último caso, el tamaño de la muestra que se requiere es menor o igual si se admite la opción de duda que si no se admite.
- Los contrastes físicamente mal planteados imponen, en la elección de cotas, -- restricciones más fuertes que los bien planteados.
- El tamaño de muestra fijo impone restricciones más fuertes que el tamaño libre, salvo en el caso de los contrastes físicamente mal planteados, en los que las restricciones son idénticas.

7. CONCLUSIONES

Las conclusiones más importantes de este trabajo son:

- Los contrastes físicamente mal planteados conducen forzosamente a pruebas con tamaño de primera especie grande o potencia

en rechazo pequeña y tamaño de segunda especie grande o potencia en aceptación pequeña. Este inconveniente no puede evitarse ni siquiera aumentando el tamaño de la muestra. Por ello imponen restricciones más fuertes que los bien planteados, tanto en la elección de los parámetros (tamaños y potencias garantizadas) como en la de sus cotas -- (ver tablas 2 y 3).

- Desde un punto de vista práctico es mucho más adecuado plantear contrastes físicamente bien planteados. Ello exige que el investigador separe claramente -- las hipótesis nula y alternativa.
- Las pruebas sin opción de duda no permiten elegir independientemente el tamaño de primera especie y la potencia en aceptación, ni el tamaño de segunda especie y la potencia en rechazo.
- La inclusión de la opción de duda conduce a restricciones más débiles o idénticas, que sin opción de duda, en la elección de los tamaños y potencias garantizadas o de sus cotas. La identidad se consigue en el caso de tamaño de muestra libre y contraste físicamente bien planteado; sin embargo, el tamaño de muestra que se requiere en este caso es menor con opción de duda que sin ella.
- En el caso de las pruebas óptimas para un contraste dado, el tamaño de muestra fijo impone una relación entre el tamaño de primera especie y la potencia en rechazo, y otra relación entre el tamaño de segunda especie y la potencia en aceptación, por lo que las dos parejas de parámetros no pueden fijarse independientemente.
- En el caso de las pruebas óptimas para un contraste dado, aunque el tamaño de la muestra esté libre, aparece una relación de tipo discreto entre tamaños y potencias (expresión (5')) que conduce a no poder elegir independientemente esos parámetros.
- El tamaño de muestra fijo impone normalmente restricciones más fuertes que el tamaño libre. Sin embargo, en el caso de los contrastes físicamente mal planteados

teados las restricciones son idénticas.

8. En las ocho situaciones estudiadas puede fijarse independientemente el valor de sólo uno, cualquiera, de los cuatro parámetros: dos tamaños y dos potencias garantizadas, o al menos el de dos, no cualquiera (ver Tabla 3), de sus cotas. Las tablas 2 y 3 dan las restricciones a las que quedan sometidos el resto de los parámetros y de las cotas respectivamente.
9. De los casos estudiados, sólo los de tamaño de muestra libre y contraste físicamente bien planteado, permiten acotar independientemente los dos tamaños y las dos potencias garantizadas.

8. REFERENCIAS

- /1/ LEHMANN, E.L.: "Testing Statistical Hypotheses". John Wiley & Sons, Inc., New York, 1959.
- /2/ NEYMAN, J. y PEARSON, E.S.: "On the problems of the most efficient tests of Statistical hypotheses". Phil. Trans. Roy. Soc. Ser. A., Vol. 231, 1933, pp 289-337
- /3/ WALD, A.: "Sequential Analysis". John Wiley & Sons, Inc., New York, 1947.