

DEPARTAMENTO DE ESTADÍSTICA DE LA UNIVERSIDAD DE BARCELONA

PROGRAMA DE PROBABILITATS I ESTADÍSTICA (BIENNIO 87/89)

# Optimización en estudios de Monte Carlo en Estadística: Aplicaciones al Contraste de Hipótesis.

Memoria presentada para optar  
al título de Doctor en Biología  
por la Universidad de Barcelona,  
por

Esteban Vegas Lozano

Visto bueno  
El Director,

Dr. Jordi Ocaña Rebull,  
Catedrático de Universidad.  
Departamento de Estadística.  
Universidad de Barcelona.

Barcelona, 20 de Junio de 1996

### 4.3.5 Gráficos

Se presenta un conjunto de gráficos asociados a las anteriores tablas de resultados.

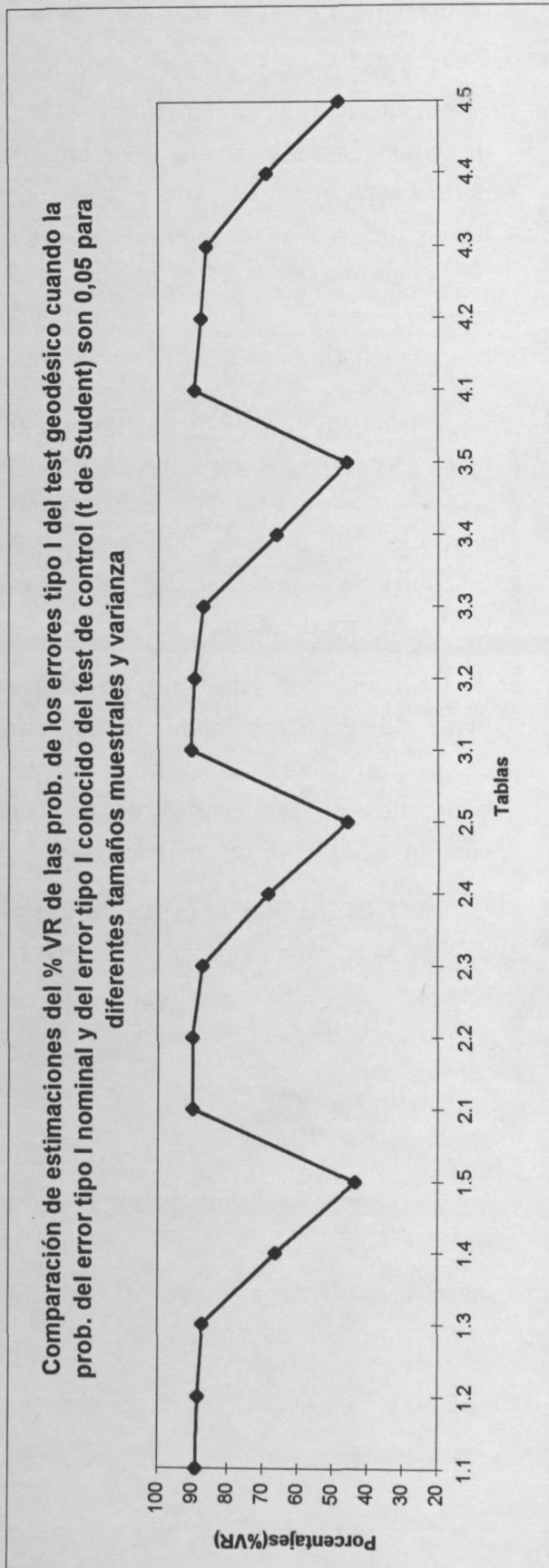
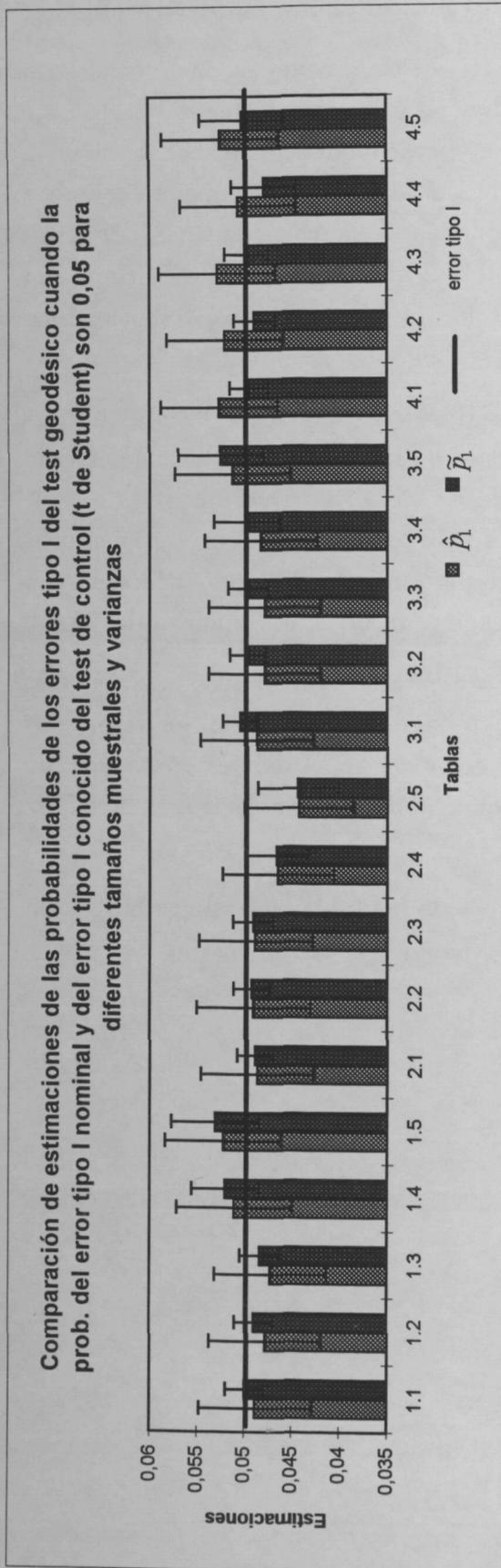
Hay dos tipos de gráficos. El primero es el que tiene como eje de ordenadas las estimaciones de  $p_1$ , mientras que el segundo tipo son estimaciones del porcentaje de reducción de la varianza. En ambos, la numeración del eje de abscisas sigue la notación "número de tabla". "posición de la subtabla", de esta manera se indica las condiciones experimentales que corresponden a esa subtabla. Así, por ejemplo, la numeración 3.2 es la tabla 3 subtabla 2 que corresponde a la condición experimental del estudio de simulación de la potencia del nuevo test geodésico con  $n = 5000$ ,  $n_1 = n_2 = 61$ ,  $\sigma_1^2 = 1.05$   $\sigma_2^2 = 0.95$ .

En el primer gráfico se presenta dos series de estimaciones de  $p_1$ . La primera serie está formada por las estimaciones de la potencia del test de estudio sólo obtenida a partir de la frecuencia relativa de  $Y$ ,  $\hat{p}_1$ , y el margen de error asociado a un nivel de confianza del 95%. Y la segunda serie son también estimaciones de la potencia del test de estudio pero usando el nuevo estimador,  $\tilde{p}_1$ , y el correspondiente margen de error asociado a un nivel de confianza del 95%.

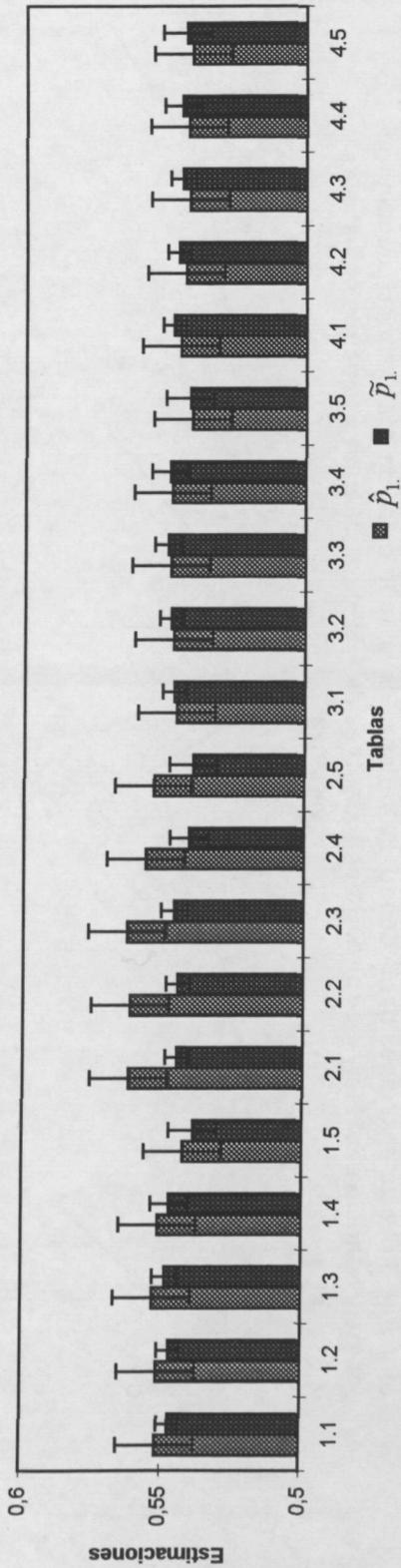
Utilizando estos dos tipos de gráficos se presentan las tablas 1 hasta la 4 en tres parejas de gráficos, donde en cada pareja (o página) se muestran los resultados de las estimaciones de la probabilidad de error de tipo I del test geodésico, de la potencia del test geodésico cuando la potencia del test de control es de 0.55 y cuando es de 0.92, respectivamente.

De igual manera se presentan los gráficos de las tablas 4 hasta la 8. Con la diferencia de que la potencia que se está evaluando es la del test de Cochran y Cox.

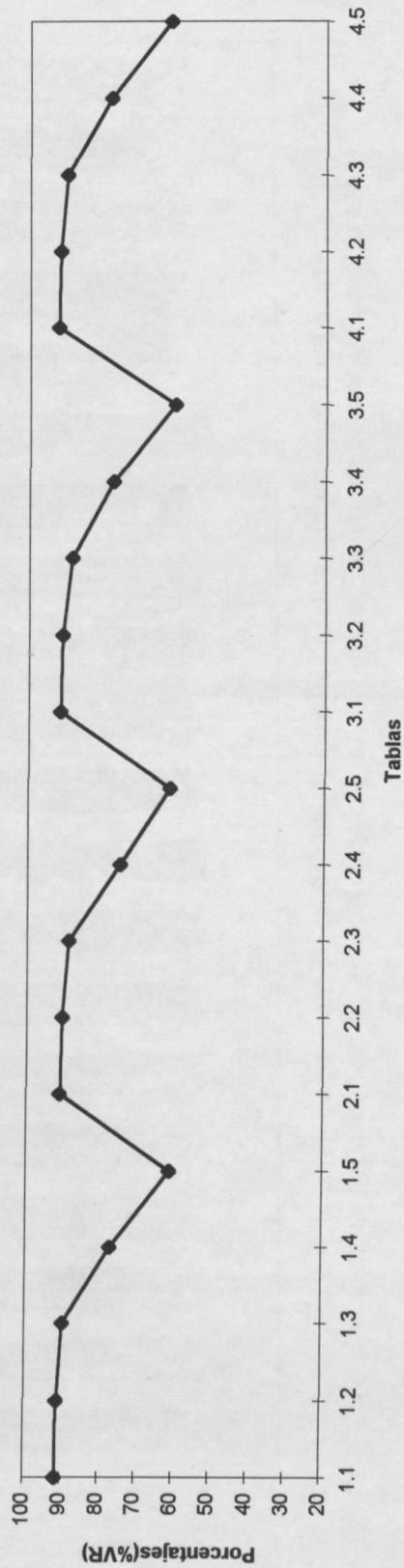
4.3.5.1 Gráficos de los resultados del estudio simulación del test geodésico



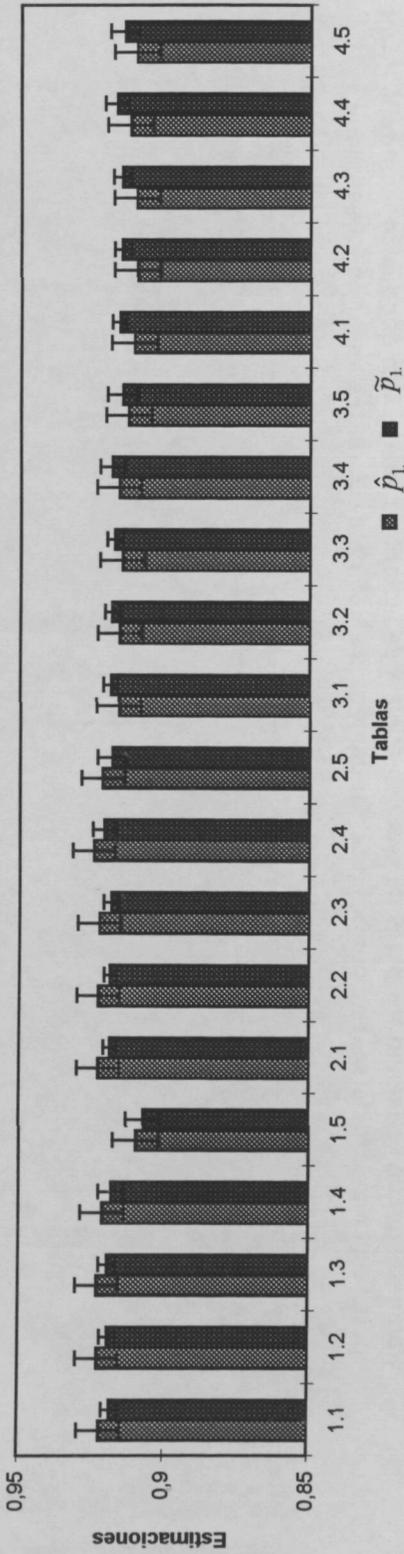
Comparación de estimaciones de las potencias del test geodésico cuando la potencia conocida del test de control (t de Student) es de 0,55 para diferentes tamaños muestrales y varianzas



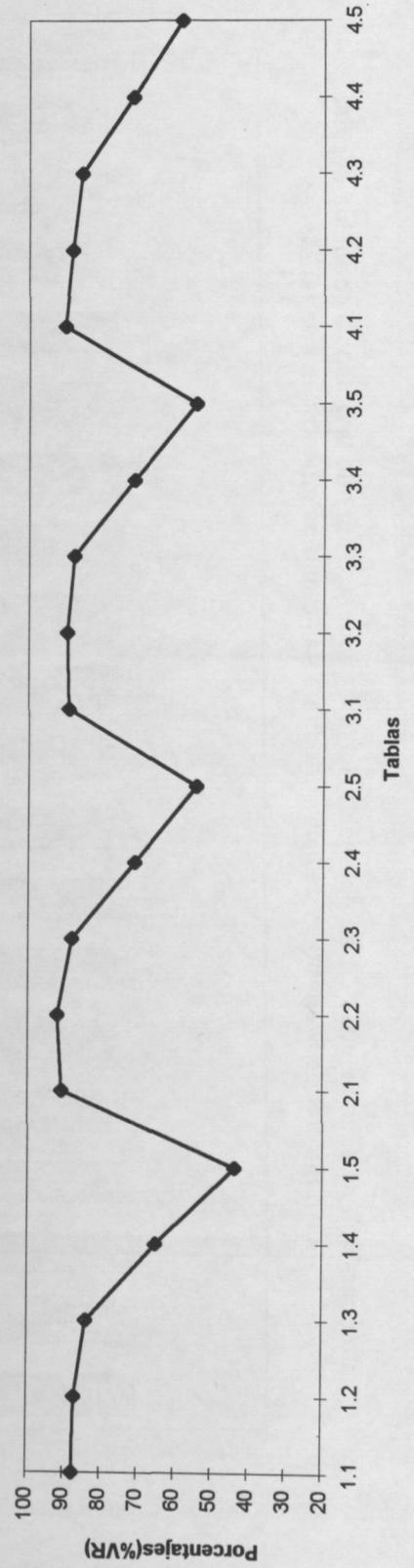
Comparación de estimaciones del % de red. de varianza de las potencias del test geodésico cuando la potencia conocida del test de control (t de Student) es de 0,55 para diferentes tamaños muestrales y varianzas



Comparación de estimaciones de las potencias del test geodésico cuando la potencia conocida del test de control (t de Student) es de 0,92 para diferentes tamaños muestrales y varianzas

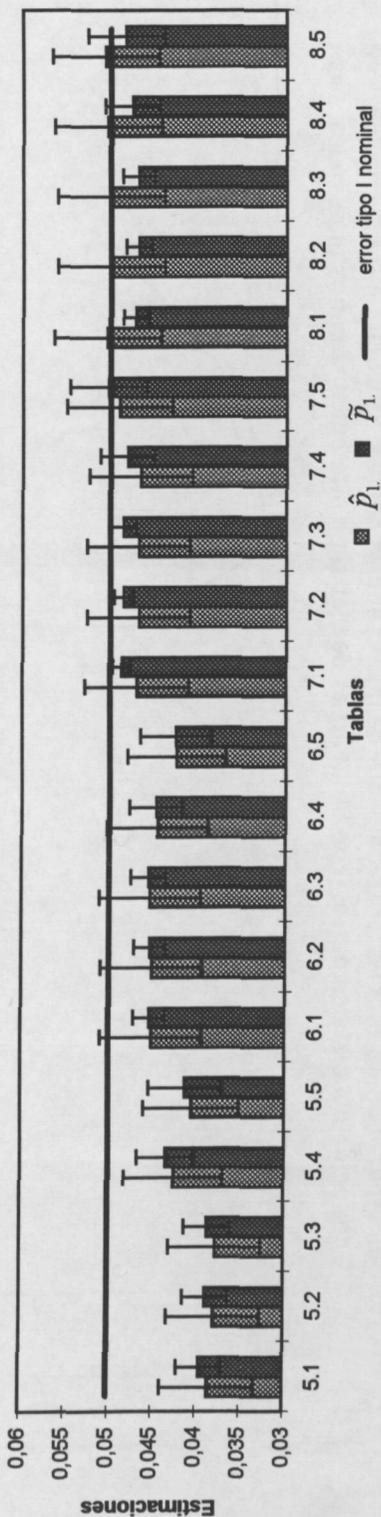


Comparación de estimaciones del % de red. de varianza de las potencias del test geodésico cuando la potencia conocida del test de control (t de Student) es de 0,92 para diferentes tamaños muestrales y varianzas

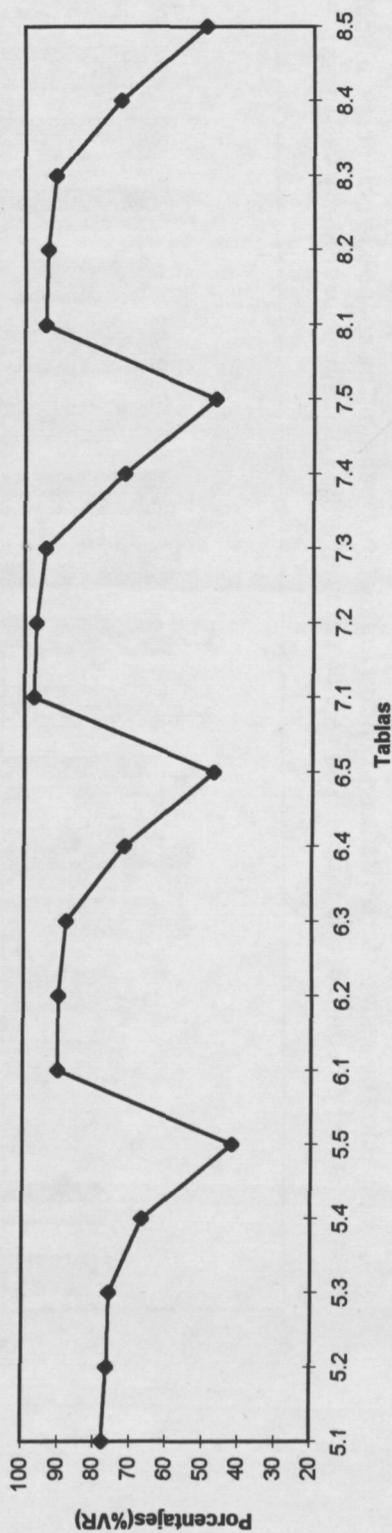


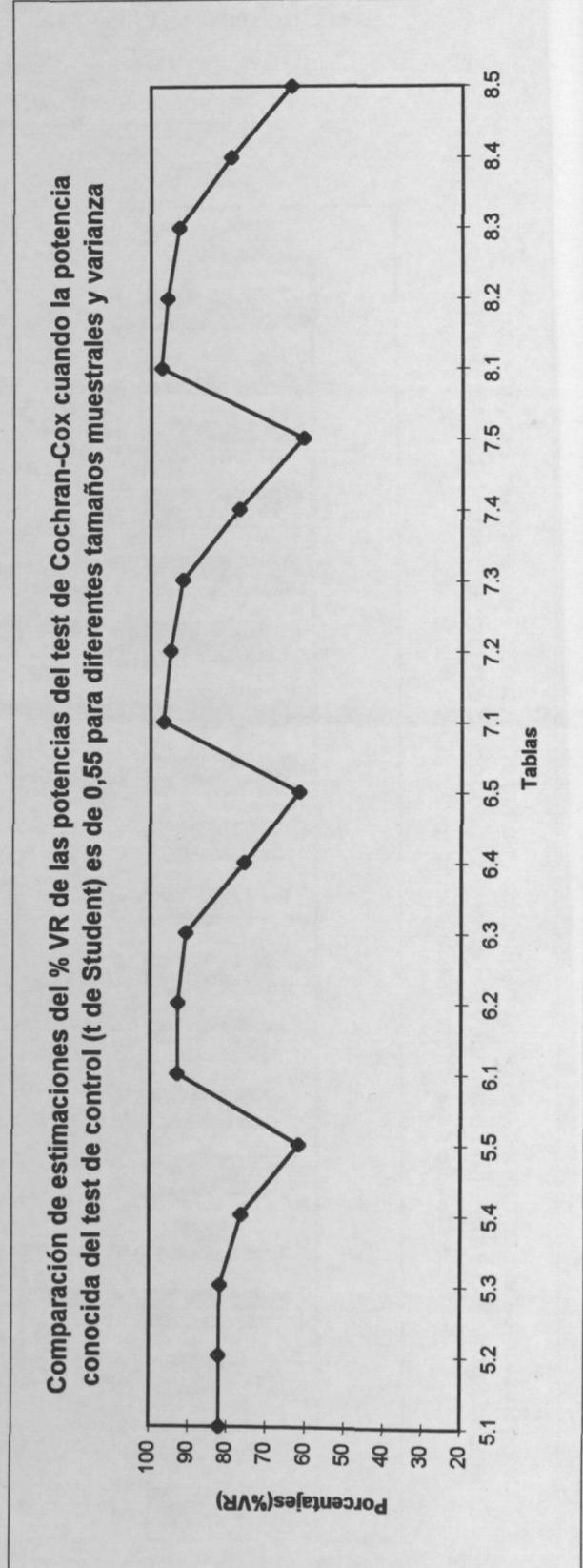
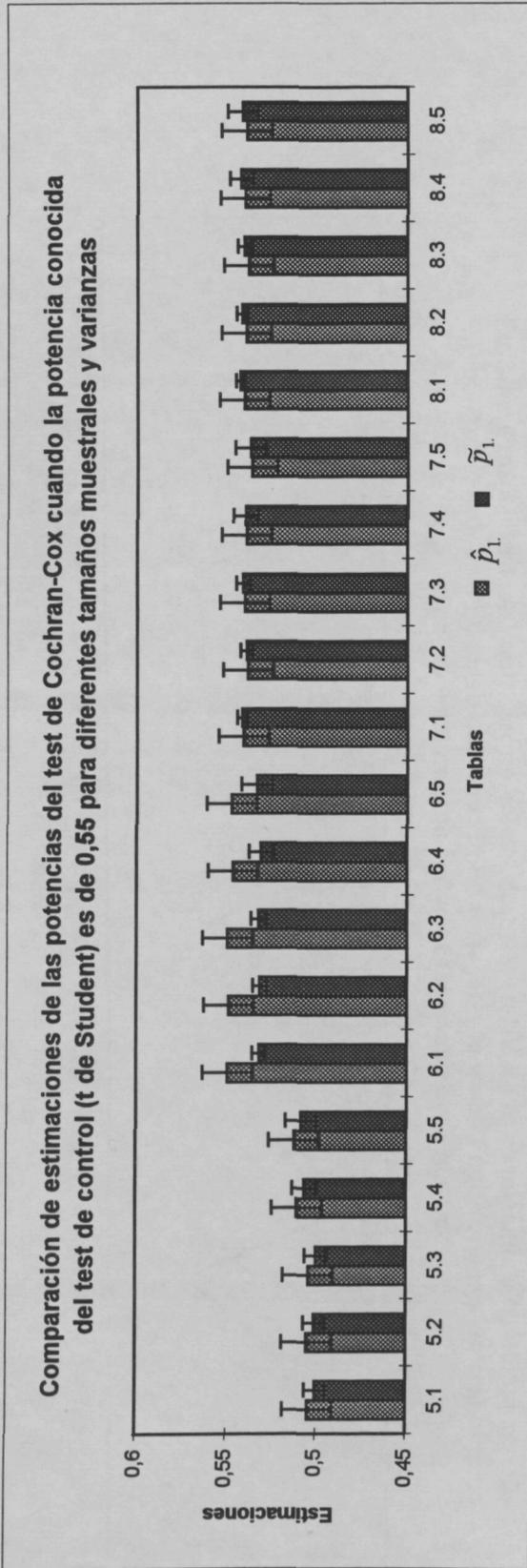
4.3.5.2 Gráficos de los resultados del estudio de simulación del test de Cochran y Cox

Comparación de estimaciones de las probabilidades de los errores tipo I del test de Cochran-Cox cuando la prob. del error tipo I nominal y del error tipo I conocido del test de control (t de Student) son 0,05 para diferentes tamaños muestrales y varianzas

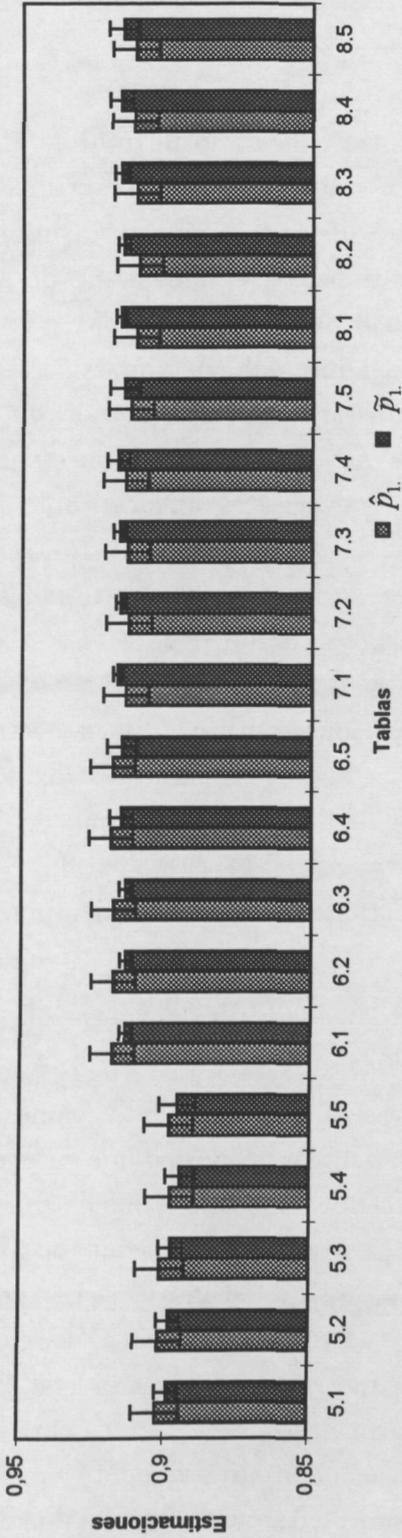


Comparación de estimaciones del % VR de las prob. de los errores tipo I del test de Cochran-Cox cuando la prob. del error tipo I nominal y del error tipo I conocido del test de control (t de Student) son 0,05 para diferentes tamaños muestrales y varianzas

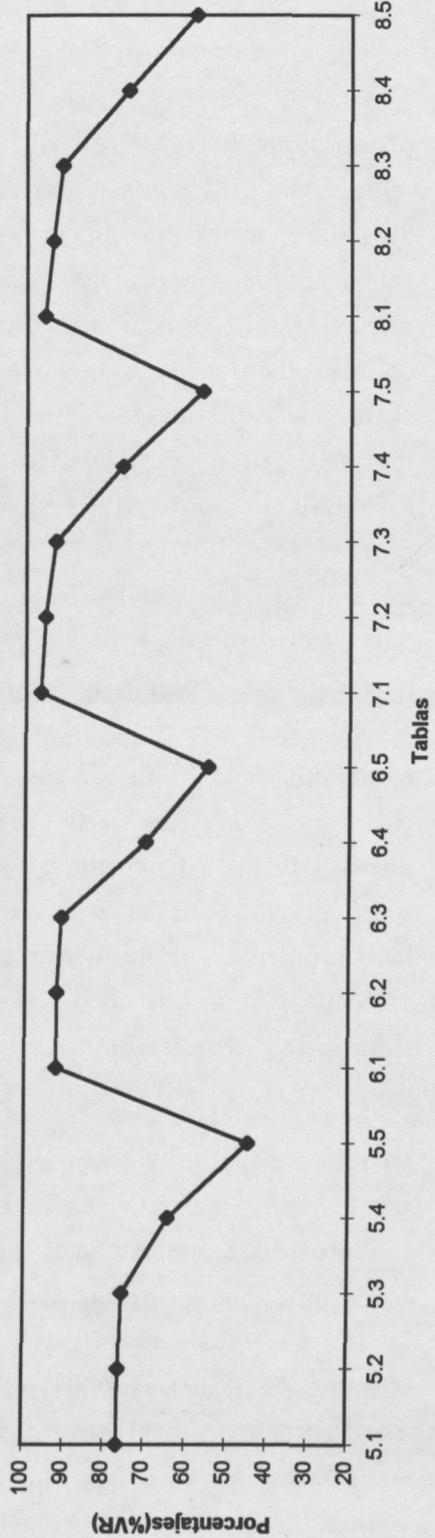




Comparación de estimaciones de las potencias del test de Cochran-Cox cuando la potencia conocida del test de control (t de Student) es de 0,92 para diferentes tamaños muestrales y varianzas



Comparación de estimaciones del % VR de las potencias del test de Cochran-Cox cuando la potencia conocida del test de control (t de Student) es de 0,92 para diferentes tamaños muestrales y varianzas



### 4.3.6 Conclusiones

El uso del estimador definido en el capítulo 3 como:

$$\tilde{p}_1 = p_0 \frac{n_{10}}{n_{00} + n_{10}} + p_1 \frac{n_{11}}{n_{01} + n_{11}} \quad (4.27)$$

proporciona una buena reducción de la varianza con un coste mínimo en el incremento de la complejidad del diseño de la simulación para la primera serie de simulaciones (varianzas iguales o no muy diferentes) que va disminuyendo tanto como aumenta las diferencias entre varianzas (segunda serie de simulaciones). Esto se refleja en las tablas 1 hasta 8 por la columna  $\widehat{\%VR}$  y por la columna  $\widehat{\%MER}$  (porcentaje de reducción del intervalo de confianza), el cual es menos espectacular pero tal vez más realista como una medida del incremento de precisión. Dentro de la primera serie de simulaciones la estimación más baja del porcentaje de reducción de la varianza es un aceptable 75.68% para el test de Cochran y Cox, cuando  $n_1 = n_2 = 11$ ,  $\Delta\mu = 1.50755596$  y  $\sigma_1^2 = 1.10$   $\sigma_2^2 = 0.90$ . Todas las reducciones de las varianzas estimadas están entre este mínimo y un valor máximo de 97.28%, también para el test de Cochran y Cox, cuando  $n_1 = n_2 = 121$ ,  $\Delta\mu = 0.268181860$  y  $\sigma_1^2 = \sigma_2^2 = 1.0$ . En la segunda serie de simulaciones las reducciones de las varianzas estimadas están entre un valor mínimo de 41.87% para el test Cochran y Cox, cuando  $n_1 = n_2 = 11$ ,  $\Delta\mu = 0$  y  $\sigma_1^2 = 1.6$   $\sigma_2^2 = 0.4$ , y un valor máximo de 79.42% para el test de Cochran y Cox, cuando  $n_1 = n_2 = 121$ ,  $\Delta\mu = 0.268181860$  y  $\sigma_1^2 = 1.333333333$   $\sigma_2^2 = 0.666666667$ . Esta tendencia a la disminución de las reducciones de las varianzas estimadas queda plasmada en las gráficas al incrementar la amplitud del intervalo de confianza de  $\tilde{p}_1$ , cuando se aumenta las diferencias entre varianzas para un mismo tamaño muestral.

Para tamaños muestrales pequeños ( $n_i = 11$ ), todas las reducciones de las varianzas estimadas son mejores en el estudio del test geodésico que en el estudio del test de Cochran y Cox. Esta situación se invierte para tamaños muestrales mayores. Tanto en un caso como otro, las diferencias se reducen cuando aumenta la desigualdad entre varianzas (segunda serie de simulaciones). Así se refleja el hecho que para tamaños muestrales pequeños el test de Cochran y Cox es marcadamente más conservativo y menos potente que el test de control t de Student. Para tamaños muestrales más grandes, por otra parte, el test de Cochran y Cox se acerca al test t de Student y ambos se convierten progresivamente equivalentes a un test basado en las tablas normal, y por tanto altamente correlacionados. Sin

embargo, en cualquier caso, la técnica de reducción de la varianza se comporta de forma similar y aceptablemente bien en el estudio de ambos tests.

El test de Cochran y Cox es conocido por ser conservativo. Su verdadera extensión tiende a ser menor que el nivel de significación nominal (Lee and Gurland, 1975) y esta tendencia es más fuerte en el caso de igualdad de varianzas (o, más exactamente, cuando  $n_1\sigma_1^2 = n_2\sigma_2^2$ ). Esta propiedad puede ayudarnos a contrastar la sensibilidad de las pruebas de simulación y a señalar como se mejora por la técnica de reducción de la varianza. Al hacer 5000 réplicas de simulación, usando la frecuencia relativa para estimar la extensión del test ( $\hat{p}_1$ ), es muy difícil detectar esta tendencia conservativa. Como así se muestra en las tablas 5 hasta 8 cuando  $\Delta\mu = 0$ , esto es, cuando las simulaciones son realizadas bajo la hipótesis de igualdad de medias. En estas situaciones, la mayoría de los valores estimados  $\hat{p}_1$  no difieren significativamente del valor nominal de 0.05 ya que el intervalo de confianza aproximado del 95% alrededor de  $\hat{p}_1$ , que aparece en la columna etiquetada por  $\pm 1.959964\hat{\sigma}$  en las tablas 5 hasta 8, cubre erróneamente el valor 0.05 en muchos casos. Esta circunstancia se visualiza por el cruce de la línea  $p = 0.05$  por muchos de estos intervalos de confianza en la gráfica (pág. 107) que corresponde al test de Cochran y Cox para  $\Delta\mu = 0$ .

Las únicas excepciones al principio precedente corresponde a casos muy claros de tamaños muestrales pequeños  $n_1 = n_2 = 11$  y a un caso aislado con tamaños muestrales  $n_1 = n_2 = 31$ . Estos casos significativos se etiquetan con un signo † en las tablas. Incluso la hipótesis nula más fácil de rechazar,  $p_1 = 0.05$  versus la alternativa  $p_1 < 0.05$ , se acepta en las mismas situaciones cuando el test se basa en  $\hat{p}_1$ . Por otro lado, cuando  $\tilde{p}_1$  se usó en el estudio del test de Cochran y Cox, la hipótesis  $p_1 = 0.05$  se rechaza en todos los casos en que las varianzas son iguales o semejantes y se acepta cuando las varianzas son muy diferentes sólo en aquellos casos de tamaños muestrales grandes (61 y 121). Así queda expuesto en las tablas 5 hasta 8 (si hay diferencias significativas de nuevo se marca con †) y en la gráfica que corresponde al test de Cochran y Cox para  $\Delta\mu = 0$ , donde el intervalo de confianza basado en  $\tilde{p}_1$  no incluye el valor 0.05 en la primera serie de simulaciones mientras que esto no sucede en la segunda serie (para  $n_1 = n_2 = 61$  y 121) ya que el intervalo de confianza basado en  $\tilde{p}_1$  se hace cada vez más amplio. La falta de rechazo en la segunda serie de simulaciones está en concordancia con el hecho que el test de Cochran y Cox pierde progresivamente su carácter conservador cuando las varianzas se convierten gradualmente en desiguales. Por otra parte, se puede

observar como el carácter conservador de este test también se pone de manifiesto para las potencias de 0.55 y 0.92 en muestras pequeñas,  $n_1 = n_2 = 11$ .

Por tanto, se presenta un caso claro en el cual un rasgo cualitativo se puede detectar gracias al uso de la reducción de la varianza.

En contraste con la tendencia observada para el test de Cochran y Cox, ninguna de las extensiones estimadas para el test geodésico (tanto si las medimos usando  $\hat{p}_1$  como  $\tilde{p}_1$ ) difiere de manera significativa del valor nominal del 0.05 incluso para las varianzas más desiguales (excepto para un extraño caso de  $n_1 = n_2 = 31, \sigma_1^2 = 1.6, \sigma_2^2 = 0.4$ ). Así queda de manifiesto en la pérdida del signo † en las tablas 1 hasta 4 y en la gráfica que corresponde al test geodésico para  $\Delta\mu = 0$ , donde la línea  $p = 0.05$  corta el intervalo de confianza alrededor de la estimación de la extensión. En vista de lo anterior, se puede concluir que la verdadera extensión del test geodésico es igual al nivel de significación establecido o muy próximo a él.

Ambos tests bajo estudio, tienen potencias comparables pero significativamente diferentes para tamaños muestrales pequeños de  $n_1$  y  $n_2$ , a un nivel del 0.05, pero paulatinamente van disminuyendo sus diferencias al aumentar la desigualdad entre varianzas llegando al caso extremo de  $\sigma_1^2 = 1.6, \sigma_2^2 = 0.4$  donde solamente para  $\Delta\mu = 0$  existen diferencias significativas entre los tests. En estos casos las potencias estimadas del test de Cochran y Cox son siempre menores que las potencias estimadas del test geodésico. Estas diferencias también se atenúan con el incremento progresivo del tamaño muestral. Esta tendencia queda reflejada en las tablas por la dispersión gradual del signo ‡. De nuevo, estas diferencias son más fáciles de detectar bajo el uso de la reducción de la varianza, esto es, cuando  $\tilde{p}_1$  se usa en lugar de la frecuencia relativa  $\hat{p}_1$ , aunque al aumentar la desigualdad entre varianzas estas diferencias tienden a desaparecer. Se puede observar que cuando el tamaño muestral es  $n_1 = n_2 = 31$  ninguna de las diferencias basadas en  $\hat{p}_1$  es significativa, mientras que todas aquellas diferencias basadas en  $\tilde{p}_1$  de la primera serie de simulaciones son significativas. Este fenómeno no se repite para la segunda serie de simulaciones que tienen unas varianzas más desiguales.

En la segunda serie de simulaciones se observa, tanto en el test geodésico como en el test de Cochran y Cox y tanto si utilizamos  $\hat{p}_1$  como  $\tilde{p}_1$ , que no afecta, por lo menos de manera ostensible, el incremento de las diferencias entre las varianzas en la variación de las potencias. Tal como se expone en las tablas y en

sus gráficas asociadas.

Un efecto remarcable que produce la utilización de la técnica de reducción de la varianza es que si el valor medio de la variable de control  $C$ , en este caso la potencia del test  $t$  de Student para las 5000 réplicas de simulación, es inferior al valor esperado (0.05, 0.55 o 0.92) entonces, se producirá un incremento del valor obtenido en la variable respuesta  $Y$ , en nuestro caso el test geodésico o el test de Cochran y Cox, que será tanto mayor como sea la diferencia entre el valor medio y esperado de la variable de control  $C$ . Y de manera semejante ocurrirá en el caso contrario. Por tanto, los errores de muestreo para la obtención del valor medio de la variable de control servirán para corregir los correspondientes errores de muestreo en la variable de respuesta, influyendo en los cambios de valores de  $\hat{p}_1$  a  $\tilde{p}_1$  en los dos test de estudio.

En nuestras simulaciones, los valores medios de la potencia del test  $t$  de Student,  $\hat{p}_1$ , y del margen de error asociado a un nivel de confianza del 95%,  $\pm 1.96\hat{\sigma}_{ST}$ , fueron los siguientes:

$n = 5000$		Valores de potencia teórica					
		0.05		0.55		0.92	
$n_1 = n_2$		$\hat{p}_1$	$\pm 1.96\hat{\sigma}_{ST}$	$\hat{p}_1$	$\pm 1.96\hat{\sigma}_{ST}$	$\hat{p}_1$	$\pm 1.96\hat{\sigma}_{ST}$
11		0.0488	0.00597	0.5546	0.01378	0.9238	0.00735
31		0.0498	0.00603	0.5678	0.01373	0.9242	0.00734
61		0.0482	0.00594	0.5492	0.01379	0.9172	0.00764
121		0.0534	0.00623	0.5474	0.01379	0.9146	0.00775

En todos ellos, se observa que el intervalo de confianza del 95% para la potencia de la  $t$  de Student incluye el valor teórico esperado, como era de suponer. Si nos fijamos en la potencia teórica 0.05 de la  $t$  de Student, vemos que todos los valores medios son menores de 0.05 excepto para el tamaño muestral 121. Por consiguiente, los valores de la potencia del test de estudio, tanto el geodésico como el de Cochran y Cox, usando el nuevo estimador ( $\tilde{p}_1$ ) son mayores que sus respectivos valores en la potencia del test de estudio estimada como frecuencia relativa ( $\hat{p}_1$ ) excepto para  $n_1 = n_2 = 121$  que es menor. De la misma manera se puede razonar para las otras potencias teniendo en consideración el efecto corrector de la utilización de la técnica de reducción de la varianza.

De los párrafos anteriores se puede inferir que el nuevo (difícil de implementar) test geodésico es mejor que el test de Cochran y Cox, especialmente para muestras

pequeñas. Las diferencias se reducen tanto al aumentar la desigualdad entre varianzas, como al aumentar el tamaño muestral. La extensión del test geodésico es siempre igual o muy próxima al nivel de significación nominal y es más potente que el test de Cochran y Cox. Estos buenos resultados en el test geodésico son, de hecho, de limitado interés práctico. Nuestro estudio de simulación no es una comparación completa de las diferentes soluciones prácticas al problema de Behrens-Fisher. Muchas de estas soluciones dan mejor resultado que el test de Cochran y Cox (y se comportan de forma similar al test geodésico) y su uso está parcialmente limitado por su comparativamente mayor dificultad, la cual es extrema para el test geodésico.

Por otro lado, los resultados concernientes a la técnica de reducción de varianza son buenos cuando el grado de correlación con el test de control, en este caso, test *t* de Student, es alto. Así, cuando las diferencias entre las varianzas no es grande (primera serie de simulaciones) en todos los casos es de un valor superior al 75% para  $\widehat{\%VR}$  y de un valor superior al 50% para  $\widehat{\%MER}$ . Además, la utilización de esta técnica supone un mínimo esfuerzo en la complejidad del diseño de la simulación.

FIG.1: ALGORITMO PARA EL ESTUDIO DE LA POTENCIA DEL TEST GEODÉSICO

Introducción de parámetros

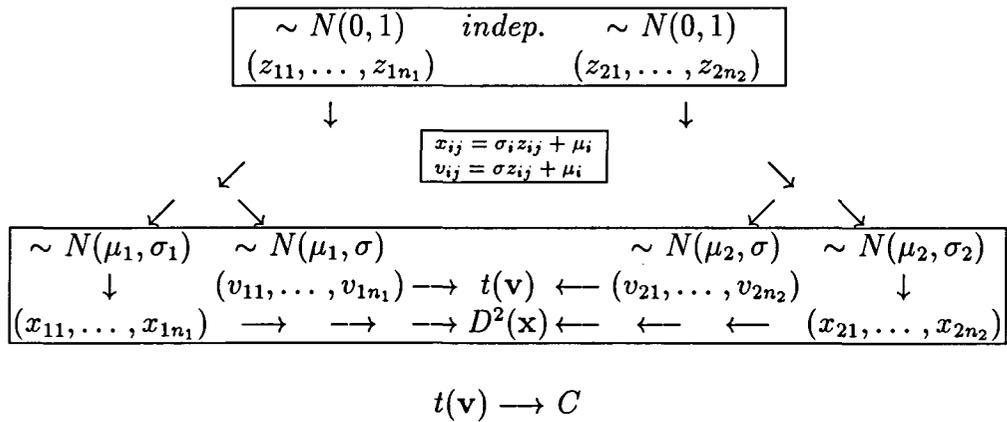
$$n_1, n_2, \sigma_1, \sigma_2, \beta_c, \alpha, n, B$$

↓

$$\begin{matrix} \mu_1 = 0 \\ \mu_2 = \phi_{\beta_c} \sigma \sqrt{\frac{2(n_1 + n_2)}{n_1 n_2}} \end{matrix} \quad \text{donde } \sigma^2 = \frac{n_1 \sigma_1^2 + n_2 \sigma_2^2}{n_1 + n_2} \text{ y } \phi_{\beta_c} \text{ es obtenido en el nomograma} \\ \text{que determina la función potencia de la t de Student}$$

Repite desde j=1 hasta n:

Generación de las muestras



Bootstrap paramétrico sobre x

$$\left. \begin{matrix} \sim N(\bar{x}, s_1) & \sim N(\bar{x}, s_2) \\ (x_{11}^*, \dots, x_{1n_1}^*)_1 & (x_{21}^*, \dots, x_{2n_2}^*)_1 \\ \dots & \dots \\ (x_{11}^*, \dots, x_{1n_1}^*)_B & (x_{21}^*, \dots, x_{2n_2}^*)_B \end{matrix} \right\} \rightarrow \begin{matrix} D_{*(1)}^2 \\ \dots \\ D_{*(B)}^2 \end{matrix}$$

$$\frac{\#[D_*^2 \geq D^2(\mathbf{x})] + 1}{B + 1} \geq \alpha \rightarrow Y = 0$$

$$< \alpha \rightarrow Y = 1$$

Fin Repite

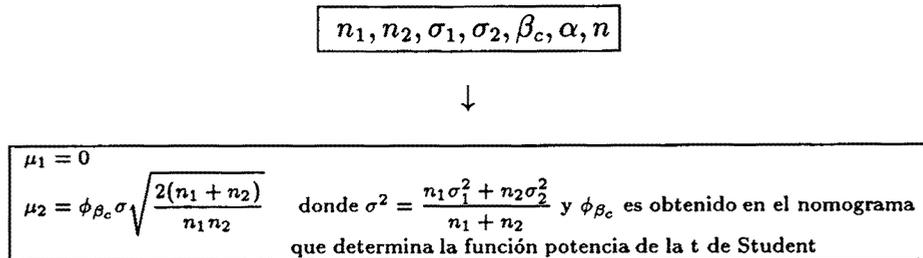
Del proceso anterior:

Reducción de la Varianza

$$\left. \begin{matrix} (Y_1, C_1) \\ \dots \\ (Y_n, C_n) \end{matrix} \right\} \rightarrow (\tilde{p}_1, \hat{\sigma}^2, \%VR, \dots)$$

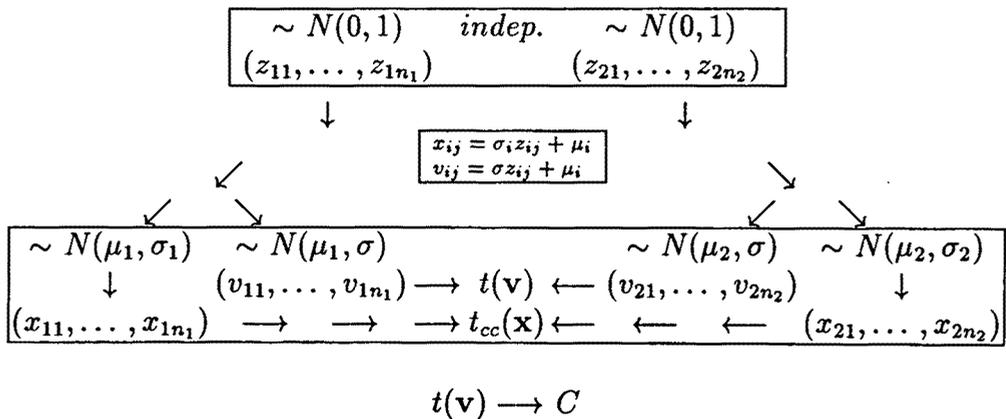
FIG.2: ALGORITMO PARA EL ESTUDIO DE LA POTENCIA DEL TEST DE COCHRAN Y COX

Introducción de parámetros



Repite desde  $j=1$  hasta  $n$ :

Generación de las muestras



Test de Cochran y Cox

$$t_{cc}(\mathbf{X}) = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(\hat{s}_1^2/n_1) + (\hat{s}_2^2/n_2)}}$$

$$t'_{\alpha/2} = \frac{t(\alpha/2, n_1 - 1)(\hat{s}_1^2/n_1) + t(\alpha/2, n_2 - 1)(\hat{s}_2^2/n_2)}{(\hat{s}_1^2/n_1) + (\hat{s}_2^2/n_2)}$$

$$|t_{cc}(\mathbf{X})| \geq t'_{\alpha/2} \rightarrow Y = 1$$

$$< t'_{\alpha/2} \rightarrow Y = 0$$

Fin repite

Del proceso anterior:

Reducción de la Varianza

$$\left. \begin{matrix} (Y_1, C_1) \\ \dots \\ (Y_n, C_n) \end{matrix} \right\} \rightarrow (\tilde{p}_1, \hat{\sigma}^2, \%VR, \dots)$$

## Capítulo 5

# ESTUDIO SOBRE LA ROBUSTEZ DEL TEST GEODÉSICO Y DEL TEST COCHRAN Y COX



**E**n los trabajos estadísticos es bastante común utilizar alguno de los tests mencionados en el capítulo 4 para resolver el problema de Behrens-Fisher sin verificar la normalidad de los datos. Esto conlleva un posible error en la curva de potencia que será mayor cuanto más se aleje de la normalidad.

Por tanto, es importante investigar la robustez del nuevo test, denominado geodésico. Así, se ha realizado un estudio de simulación de la potencia cuando cada población sigue una distribución lognormal y, de una manera más sucinta, cuando cada población sigue una distribución gamma. Ambas situaciones hipotéticas son fáciles encontrarlas en circunstancias reales.

Además, se han recorrido los mismos pasos en el test de Cochran y Cox para que posteriormente se puedan comparar sus resultados con los del test geodésico.

## 5.1 Estudio de la robustez para muestras originadas bajo la distribución lognormal

### 5.1.1 Introducción

Existen varias situaciones en las cuales la variabilidad de las observaciones muestrales se ajusta a una distribución lognormal siendo de particular interés en biología la distribución de los efectos de las dosis de fármacos. Así, en algunos artículos, en vez de verificar la lognormalidad suponen, sin comprobar, la normalidad de los datos y además, la igualdad de varianzas. Acto seguido, contrastan el valor medio del efecto de un fármaco con otro o con un placebo a partir de un test *t* de Student. Un caso concreto sobre esta situación se verá en el capítulo 6.

Por tanto, es razonable estudiar la robustez del nuevo test geodésico bajo lognormalidad. Así pues, se realiza un estudio de simulación semejante, para que sea lo más comparativo, al presentado en el capítulo 4.

### 5.1.2 Procedimiento

Se ha seguido la misma forma de proceder que la realizada para el estudio de simulación de la potencia del test geodésico. Además, también se ha estudiado la robustez para la misma situación en el test de Cochran y Cox, test bastante común para resolver el problema de Behrens-Fisher, que nos servirá de referencia a la hora de comparar los resultados.

En ambos tests, el esquema de la simulación se divide en tres partes (ver Fig. 1 y 2):

1. Entrada de parámetros.
2. Obtención de los dos vectores de resultados  $(Y_i, C_i)$   $i = 1, \dots, n$
3. Aplicación de la técnica de reducción de la varianza.

Es en la fase de generación de las muestras, incluida en el segundo ítem, donde aparece la diferencia con respecto al estudio de la potencia visto en el capítulo 4. Una vez obtenidos los  $n_1$  valores normales estándar *iid*,  $\mathbf{z}_1 = (z_{11}, \dots, z_{1n_1})$  y la otra serie independiente de  $n_2$  valores normales estándar *iid*,  $\mathbf{z}_2 = (z_{21}, \dots, z_{2n_2})$ , se utiliza la transformación  $t_{ij} = \tau_i z_{ij} + \nu_i$ ,  $i = 1, 2$ , donde

$$\nu_i = \log \left( \frac{\mu_i^2}{\sqrt{\sigma_i^2 + \mu_i^2}} \right) \text{ y } \tau_i^2 = \log \left( \frac{\sigma_i^2}{\mu_i^2} + 1 \right) \quad (5.1)$$

para conseguir una distribución normal  $N(\nu_i, \tau_i)$ ,  $i = 1, 2$ , respectivamente en cada una de las series. A continuación, se calcula la exponencial de cada uno de los  $n_1 + n_2$  valores obteniéndose una muestra  $\mathbf{x} = (x_{11}, \dots, x_{1n_1}; x_{21}, \dots, x_{2n_2})$  que seguirá una distribución Lognormal  $LN(\mu_i, \sigma_i)$ ,  $i = 1, 2$ , respectivamente, ya que

$$E(X_i) = e^{\left(\nu_i + \frac{\tau_i^2}{2}\right)} \text{ y } Var(X_i) = e^{2\nu_i} \left( e^{2\tau_i^2} - e^{\tau_i^2} \right). \quad (5.2)$$

De esta manera, gracias a la correlación inducida a partir del uso variables aleatorias comunes, se mantiene la correlación entre estos valores, con los  $n_1 + n_2$  valores que se obtienen con la transformación  $v_{ij} = \sigma z_{ij} + \mu_i$ ,  $i = 1, 2$ , donde  $\sigma^2$  (4.25), es el promedio ponderado por los tamaños muestrales de  $\sigma_1^2$  y  $\sigma_2^2$  que será usado para evaluar el test t de Student, utilizado como variable control.

El resto de pasos y la implementación es idéntica a la indicada en el estudio de la potencia del test geodésico y de Cochran y Cox.

### 5.1.3 Resultados

Como se ha mencionado al principio, los resultados deben ser lo más comparables con los obtenidos en el capítulo 4. Así, se ha trabajado con las mismas semillas para generar las normales tipificadas y realizar el bootstrap paramétrico, con el mismo número de simulaciones ( $n$ ) y remuestras ( $B$ ), etc.

Se ha pensado que era suficiente estudiar tres tipos de configuraciones de parejas de varianzas:  $(\sigma_1^2 = \sigma_2^2 = 1)$ ,  $(\sigma_1^2 = 1.10, \sigma_2^2 = 0.90)$  y  $(\sigma_1^2 = 1.6, \sigma_2^2 = 0.4)$ , en vez de las cinco configuraciones de parejas de varianzas usadas para el estudio de la potencia del test geodésico ya que abarca las condiciones extremas e intermedia. Por tanto, el número de simulaciones realizado es menor: Tamaños muestrales (11, 31, 61 y 121), potencias (0.05, 0.55 y 0.92) y las anteriores parejas de varianzas que nos dan un total de 36 simulaciones para cada test estudiado, el geodésico y el test de Cochran y Cox.

En las tablas 1 hasta 8 se muestran los resultados. Las cuatro primeras corresponden al estudio de robustez del test geodésico y las otras restantes al del test de Cochran y Cox. Cada tabla se divide de igual manera que las presentadas en el capítulo 4.

### 5.1.3.1 Resultados del estudio de simulación para la robustez del test geodésico en muestras generadas según una distribución Log-normal

**Tablas 1 y 2.** Nivel de significación, potencia y grado de reducción de la varianza en el estudio de la robustez del test geodésico definido mediante (4.15) para muestras obtenidas a partir de una distribución Lognormal.

† =significativamente diferente del nivel de significación nominal.

‡ =significativamente diferente del valor correspondiente en el test Cochran y Cox.

Tabla 1:  $n = 5000$ ,  $n_1 = n_2 = 11$ ,  $B = 1000$

$\Delta\mu$	$\sigma_1^2 = 1.0 \quad \sigma_2^2 = 1.0$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0434†‡	0.005648	0.0444†‡	0.002840	74.71	49.71
0.927147567	0.5684†	0.013730	0.5645†	0.007178	72.67	47.72
1.50755596	0.922†	0.007434	0.9031†	0.005890	46.99	27.19

$\Delta\mu$	$\sigma_1^2 = 1.10 \quad \sigma_2^2 = 0.90$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.044†‡	0.005685	0.045†‡	0.003005	72.06	47.14
0.927147567	0.5698†	0.013725	0.5660†	0.007850	67.29	42.81
1.50755596	0.8988†	0.008360	0.896†	0.006343	42.44	24.13

$\Delta\mu$	$\sigma_1^2 = 1.6 \quad \sigma_2^2 = 0.4$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0558†	0.006363	0.0565†‡	0.005431	27.14	14.64
0.927147567	0.5758†	0.013700	0.5727†	0.010065	46.03	26.53
1.50755596	0.8658	0.009449	0.8633	0.008149	25.63	13.76

Tabla 2:  $n = 5000$ ,  $n_1 = n_2 = 31$ ,  $B = 1000$

$\Delta\mu$	$\sigma_1^2 = 1.0 \quad \sigma_2^2 = 1.0$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0476	0.005902	0.0478	0.003451	65.81	41.52
0.538815856	0.5674†	0.013734	0.5526	0.007651	68.96	44.29
0.866595685	0.9182	0.007597	0.9148	0.004797	60.13	36.86

$\Delta\mu$	$\sigma_1^2 = 1.10 \quad \sigma_2^2 = 0.90$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0476	0.005902	0.0478	0.003589	63.02	39.19
0.538815856	0.5704†	0.013722	0.5560	0.008094	65.21	41.02
0.866595685	0.9136	0.007788	0.9102	0.005135	56.53	34.07

$\Delta\mu$	$\sigma_1^2 = 1.6 \quad \sigma_2^2 = 0.4$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0502	0.006053	0.0503	0.004853	35.72	19.83
0.538815856	0.572	0.013716	0.5601	0.010239	44.28	25.35
0.866595685	0.891	0.008639	0.8881	0.007004	34.26	18.92

**Tablas 3 y 4.** Nivel de significación, potencia y grado de reducción de la varianza en el estudio de la robustez del test geodésico definido mediante (4.15) para muestras obtenidas a partir de una distribución Lognormal.

† =significativamente diferente del nivel de significación nominal.

‡ =significativamente diferente del valor correspondiente en el test Cochran y Cox.

Tabla 3:  $n = 5000, n_1 = n_2 = 61, B = 1000$

$\Delta\mu$	$\sigma_1^2 = 1.0 \quad \sigma_2^2 = 1.0$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0484	0.005949	0.0499	0.003443	66.51	42.13
0.377708852	0.5524	0.013784	0.5532	0.007806	67.93	43.37
0.608175159	0.9084	0.007996	0.9107	0.004983	61.17	37.69

$\Delta\mu$	$\sigma_1^2 = 1.10 \quad \sigma_2^2 = 0.90$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0482	0.005937	0.0496	0.003669	61.81	38.20
0.377708852	0.5506	0.013789	0.5514	0.008054	65.88	41.59
0.608175159	0.9032	0.008197	0.9054	0.005507	54.86	32.81

$\Delta\mu$	$\sigma_1^2 = 1.6 \quad \sigma_2^2 = 0.4$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0506	0.006076	0.0517	0.004937	33.98	18.75
0.377708852	0.5528	0.013783	0.5535	0.010185	45.40	26.11
0.608175159	0.8848	0.008850	0.8867	0.007132	35.05	19.41

Tabla 4:  $n = 5000, n_1 = n_2 = 121, B = 1000$

$\Delta\mu$	$\sigma_1^2 = 1.0 \quad \sigma_2^2 = 1.0$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0528	0.006199	0.0501	0.003681	64.74	40.61
0.268181860	0.5448	0.013805	0.547	0.007336	71.76	46.86
0.4318118128	0.9108	0.007901	0.9153	0.004579	66.42	42.05

$\Delta\mu$	$\sigma_1^2 = 1.10 \quad \sigma_2^2 = 0.90$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0534	0.006233	0.0507	0.003644	65.81	41.53
0.268181860	0.5448	0.013805	0.547	0.007663	69.18	44.49
0.4318118128	0.9064	0.008074	0.9109	0.004897	63.22	39.35

$\Delta\mu$	$\sigma_1^2 = 1.6 \quad \sigma_2^2 = 0.4$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0544	0.006287	0.0524	0.004990	36.99	20.63
0.268181860	0.544	0.013807	0.5458	0.009867	48.93	28.53
0.4318118128	0.8924	0.008590	0.8962	0.006545	41.94	23.80

### 5.1.3.2 Resultados del estudio de simulación para la robustez del test de Cochran y Cox en muestras generadas según una distribución Lognormal

Tablas 5 y 6. Nivel de significación, potencia y grado de reducción de la varianza en el estudio de la robustez del test de Cochran y Cox para muestras obtenidas a partir de una distribución Lognormal.

† =significativamente diferente del nivel de significación nominal.

‡ =significativamente diferente del valor correspondiente en el test geodésico.

Tabla 5:  $n = 5000$ ,  $n_1 = n_2 = 11$

$\Delta\mu$	$\sigma_1^2 = 1.0 \quad \sigma_2^2 = 1.0$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\bar{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0362†‡	0.005178	0.0371†‡	0.002924	68.11	43.52
0.927147567	0.5256‡	0.013842	0.5216‡	0.007143	73.37	48.40
1.50755596	0.8858‡	0.008817	0.8826‡	0.006287	49.16	28.70

$\Delta\mu$	$\sigma_1^2 = 1.10 \quad \sigma_2^2 = 0.90$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\bar{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0356†‡	0.005136	0.0364†‡	0.003067	64.35	40.29
0.927147567	0.5316‡	0.013833	0.5278‡	0.007617	69.68	44.94
1.50755596	0.8808‡	0.008982	0.8777‡	0.006671	44.84	25.73

$\Delta\mu$	$\sigma_1^2 = 1.6 \quad \sigma_2^2 = 0.4$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\bar{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.047‡	0.005867	0.0476‡	0.005024	26.67	14.36
0.927147567	0.55‡	0.013791	0.5468‡	0.009957	47.87	27.80
1.50755596	0.8576	0.009687	0.855	0.008349	25.72	13.81

Tabla 6:  $n = 5000$ ,  $n_1 = n_2 = 31$

$\Delta\mu$	$\sigma_1^2 = 1.0 \quad \sigma_2^2 = 1.0$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\bar{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0444	0.00571	0.0446†	0.003119	70.16	45.37
0.538815856	0.5518‡	0.013786	0.5369	0.007632	69.35	44.64
0.866595685	0.913	0.007813	0.9094	0.004731	63.33	39.44

$\Delta\mu$	$\sigma_1^2 = 1.10 \quad \sigma_2^2 = 0.90$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\bar{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0438†	0.005673	0.044†	0.003201	68.17	43.58
0.538815856	0.556 ‡	0.013773	0.5416	0.008116	65.28	41.08
0.866595685	0.9078	0.008020	0.9043	0.005166	58.50	35.58

$\Delta\mu$	$\sigma_1^2 = 1.6 \quad \sigma_2^2 = 0.4$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\bar{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0462	0.005819	0.0463	0.004598	37.56	20.98
0.538815856	0.5644	0.013745	0.5524	0.010164	45.32	26.05
0.866595685	0.8882	0.008735	0.8853	0.007108	33.80	18.63

Tablas 7 y 8. Nivel de significación, potencia y grado de reducción de la varianza en el estudio de la robustez del test de Cochran y Cox para muestras obtenidas a partir de una distribución Lognormal.

† =significativamente diferente del nivel de significación nominal.

‡ =significativamente diferente del valor correspondiente en el test geodésico.

Tabla 7:  $n = 5000, n_1 = n_2 = 61$

$\Delta\mu$	$\sigma_1^2 = 1.0 \quad \sigma_2^2 = 1.0$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\overline{\%VR}$	$\overline{\%MER}$
0	0.047	0.005867	0.0485	0.003285	68.64	44.00
0.377708852	0.5468	0.013800	0.5476	0.007528	70.24	45.44
0.608175159	0.907	0.008051	0.9093	0.004960	62.04	38.39

$\Delta\mu$	$\sigma_1^2 = 1.10 \quad \sigma_2^2 = 0.90$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\overline{\%VR}$	$\overline{\%MER}$
0	0.0454	0.005771	0.0468	0.003406	65.17	40.98
0.377708852	0.5452	0.013804	0.5460	0.007957	66.77	42.36
0.608175159	0.9008	0.008286	0.9031	0.005419	57.23	34.60

$\Delta\mu$	$\sigma_1^2 = 1.6 \quad \sigma_2^2 = 0.4$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\overline{\%VR}$	$\overline{\%MER}$
0	0.0494	0.006007	0.0505	0.004846	34.91	19.33
0.377708852	0.5518	0.013786	0.5525	0.010140	45.89	26.44
0.608175159	0.8842	0.008870	0.8861	0.007117	35.62	19.76

Tabla 8:  $n = 5000, n_1 = n_2 = 121$

$\Delta\mu$	$\sigma_1^2 = 1.0 \quad \sigma_2^2 = 1.0$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\overline{\%VR}$	$\overline{\%MER}$
0	0.051	0.006099	0.0483	0.003559	65.94	41.64
0.268181860	0.542	0.013811	0.5442	0.007120	73.42	48.45
0.4318118128	0.911	0.007893	0.9155	0.004423	68.60	43.96

$\Delta\mu$	$\sigma_1^2 = 1.10 \quad \sigma_2^2 = 0.90$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\overline{\%VR}$	$\overline{\%MER}$
0	0.0516	0.006132	0.0489	0.003622	65.11	40.93
0.268181860	0.5446	0.013805	0.5468	0.007537	70.19	45.41
0.4318118128	0.9076	0.008028	0.9121	0.004780	64.55	40.46

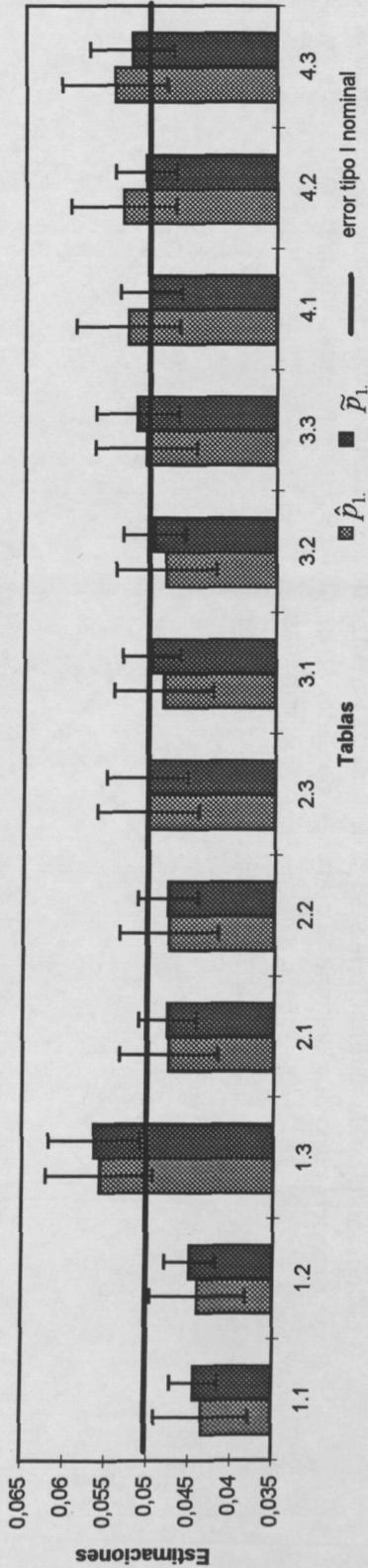
$\Delta\mu$	$\sigma_1^2 = 1.6 \quad \sigma_2^2 = 0.4$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\tilde{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\overline{\%VR}$	$\overline{\%MER}$
0	0.0516	0.006132	0.0496	0.004909	33.91	19.94
0.268181860	0.5434	0.013808	0.5452	0.009829	49.33	28.81
0.4318118128	0.8922	0.008597	0.896	0.006621	40.69	22.99

#### 5.1.4 Gráficos

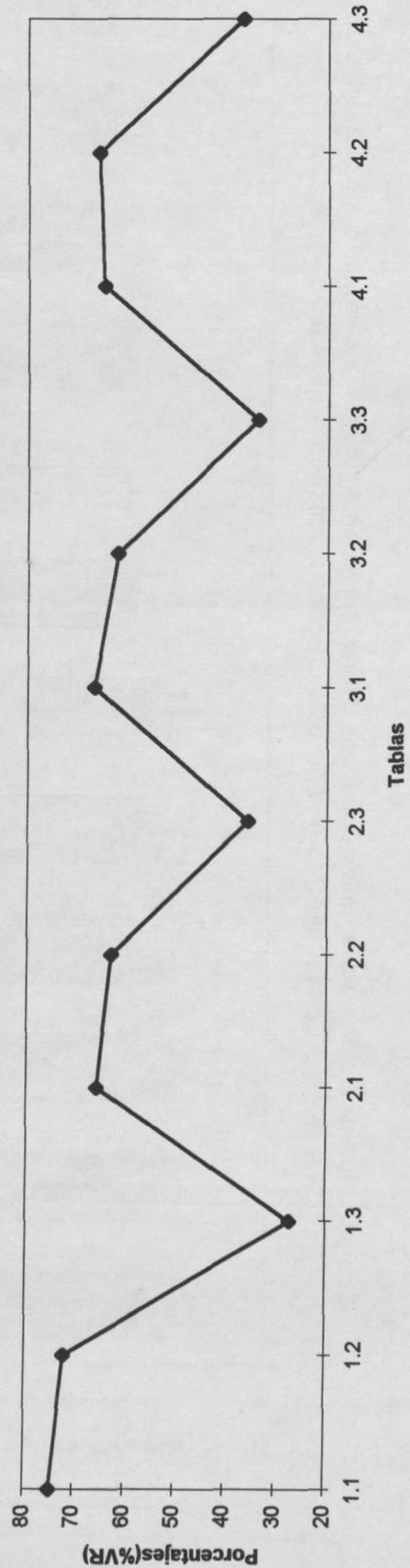
Son del mismo estilo que los presentados en el estudio de la potencia del test geodésico. Un tipo de gráfico indica las estimaciones de las potencias del test estudiado con ( $\tilde{p}_1$ ) o sin ( $\hat{p}_1$ ) aplicación de la técnica de reducción de la varianza y sus respectivos intervalos de confianza del 95% para la potencia. Mientras que el otro tipo de gráfico, muestra las estimaciones del porcentaje de reducción de varianza en cada situación. En cada uno de los dos test, geodésico y Cochran y Cox, se presentan tres gráficos de cada tipo.

5.1.4.1 Gráficos de los resultados del estudio de simulación de la robustez del test geodésico bajo lognormalidad

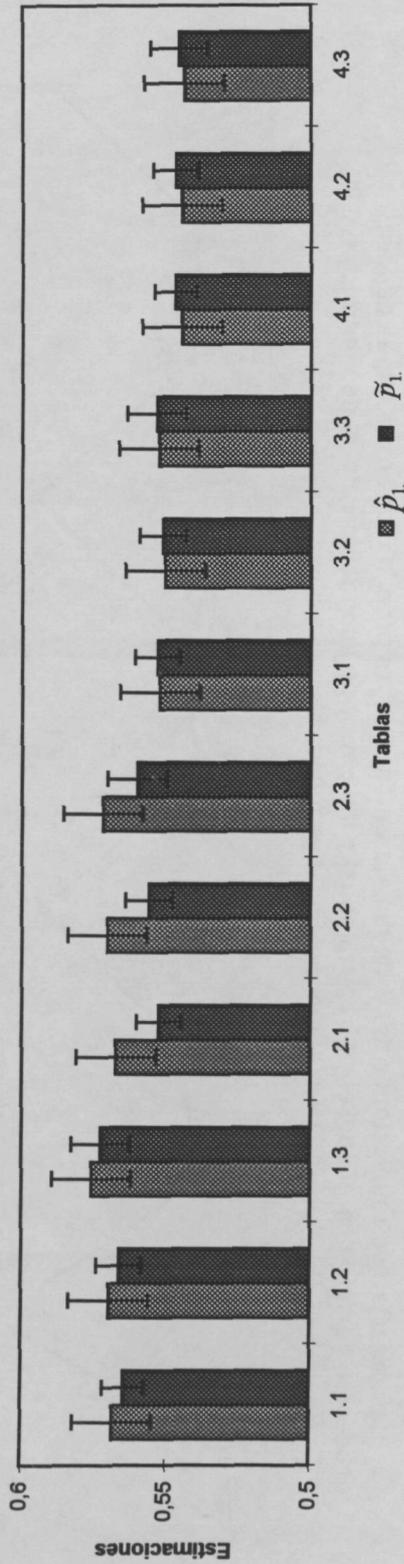
Comparación de estimaciones de las probabilidades de los errores tipo I del test geodésico cuando la prob. del error tipo I nominal y del error tipo I conocido del test de control (t de Student) son 0,05 para diferentes tamaños muestrales y varianzas



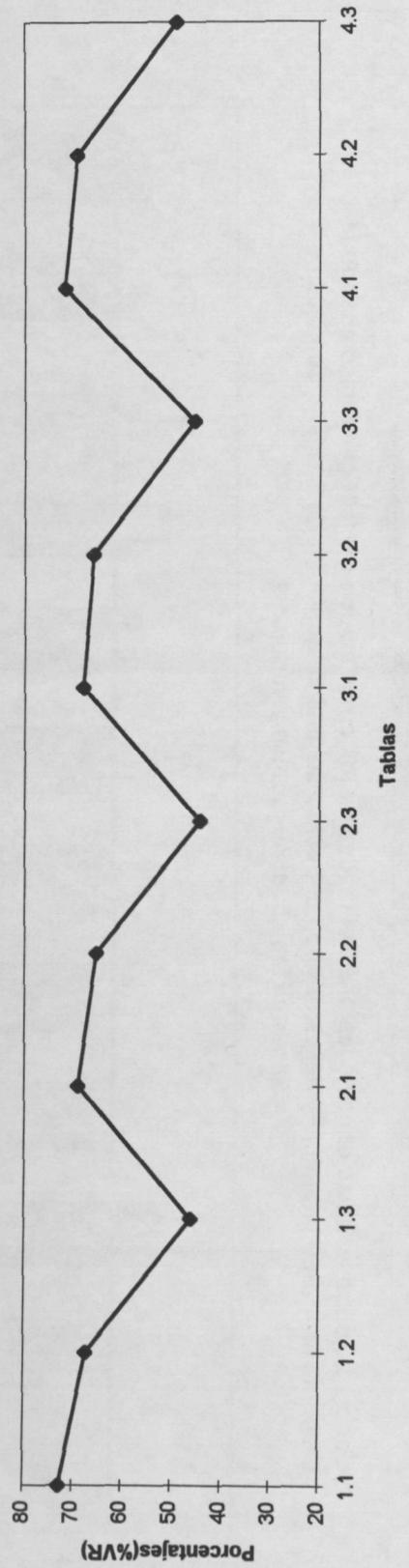
Comparación de estimaciones del % de red. de varianza de las prob. de los errores tipo I del test geodésico cuando la prob. del error tipo I nominal y del error tipo I conocido del test de control son 0,05 para diferentes tamaños muestrales y varianzas



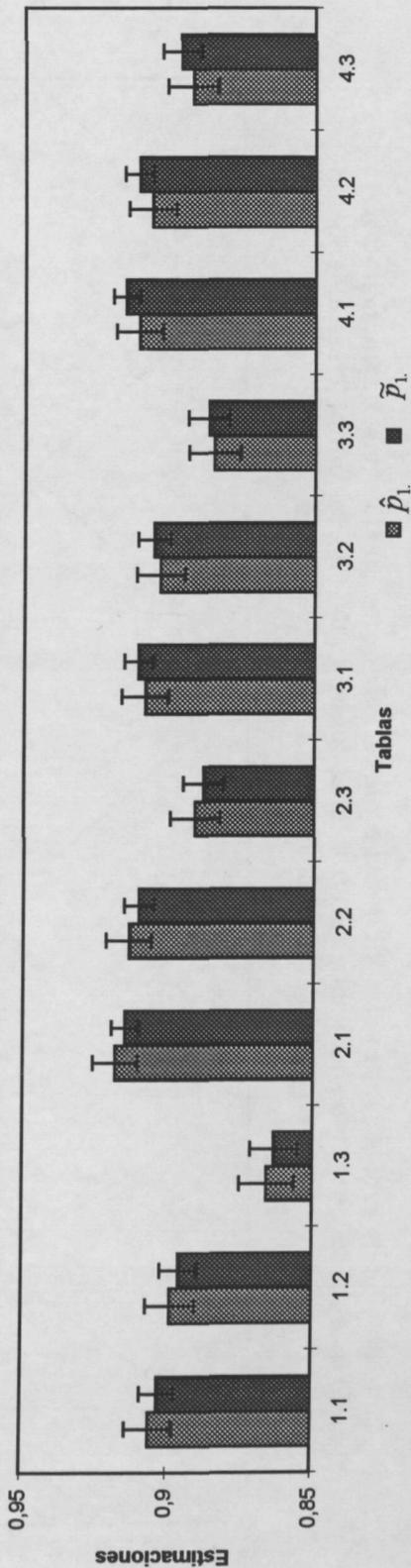
Comparación de estimaciones de las potencias del test geodésico cuando la potencia conocida del test de control (t de Student) es de 0,55 para diferentes tamaños muestrales y varianzas



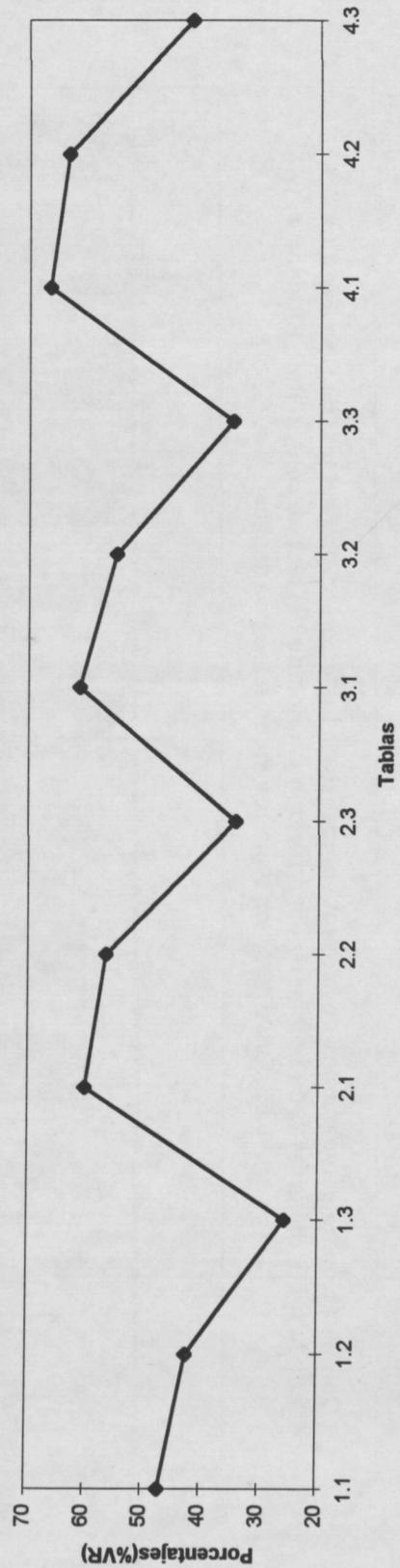
Comparación de estimaciones del % de red. de varianzas de las potencias del test geodésico cuando la potencia conocida del test de control (t de Student) es de 0,55 para diferentes tamaños muestrales y varianzas



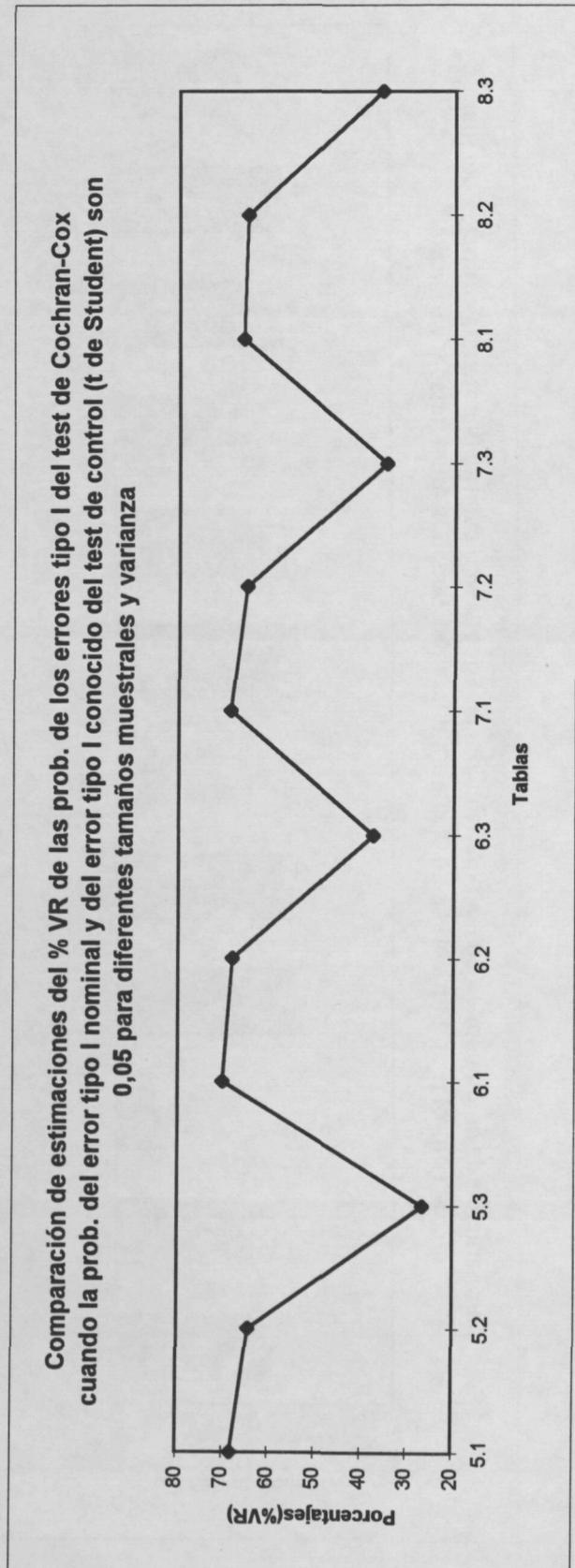
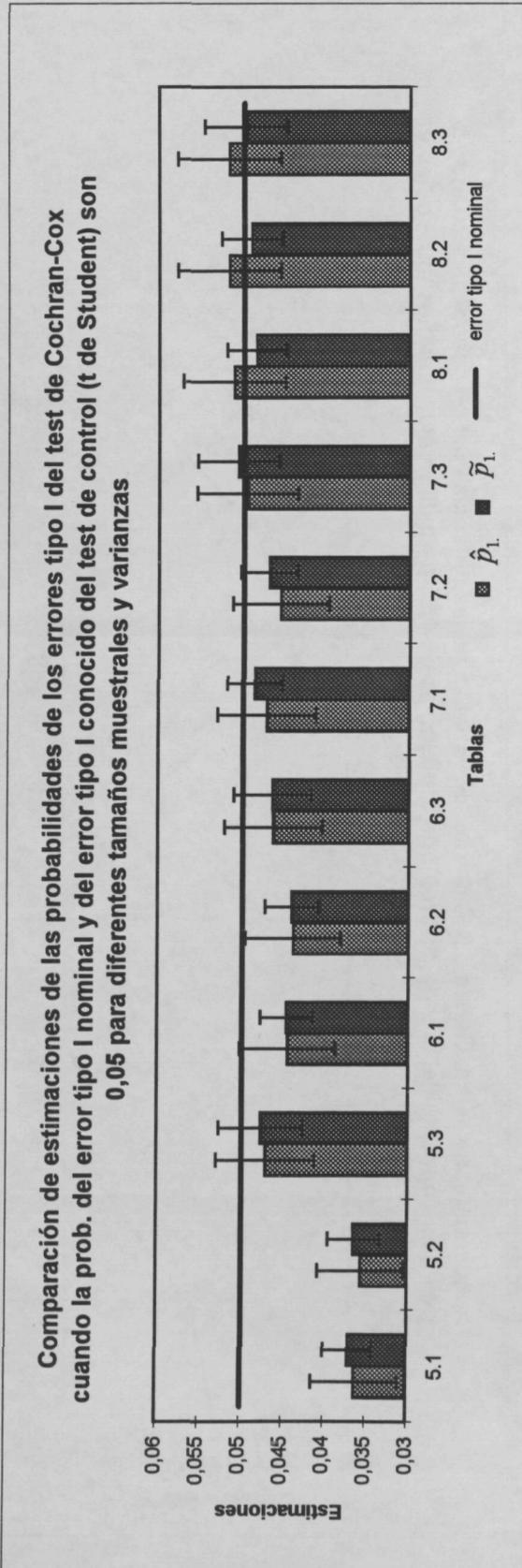
Comparación de estimaciones de las potencias del test geodésico cuando la potencia conocida del test de control (t de Student) es de 0,92 para diferentes tamaños muestrales y varianzas



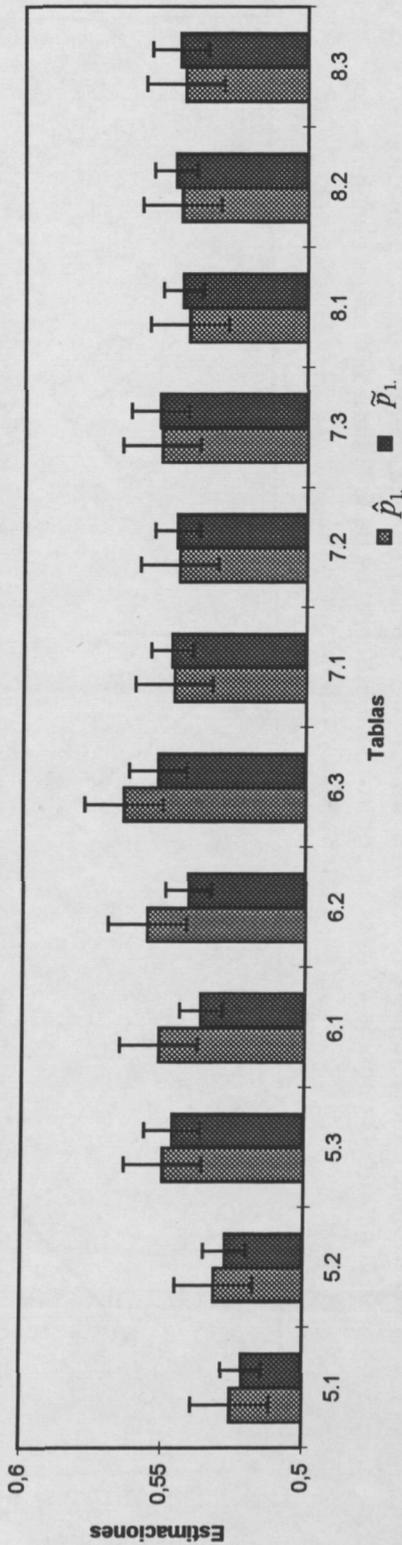
Comparación de estimaciones del % de red. de varianzas de las potencias del test geodésico cuando la potencia conocida del test de control (t de Student) es de 0,92 para diferentes tamaños muestrales y varianzas



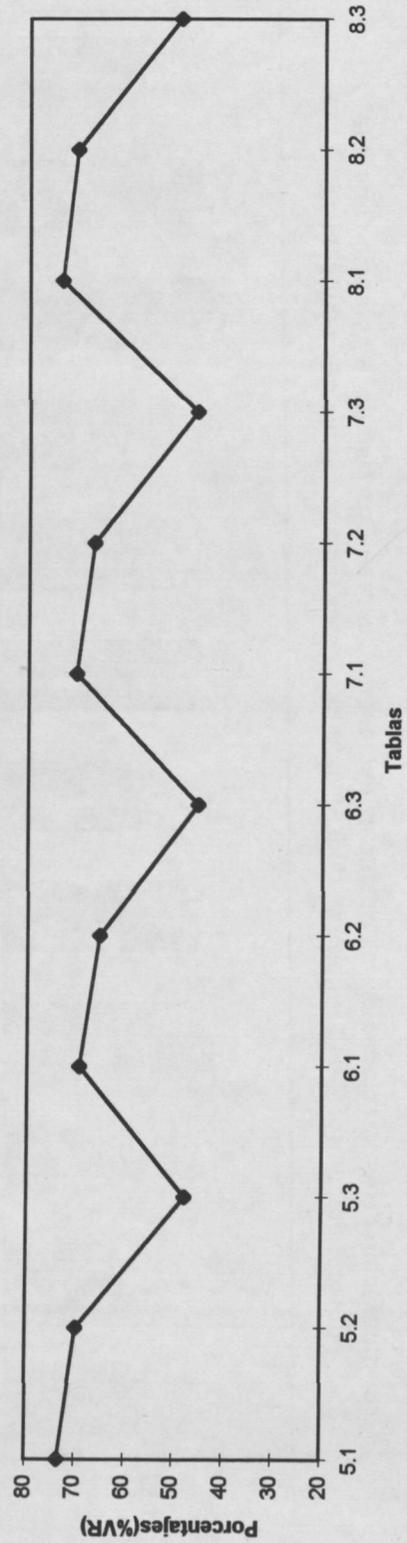
5.1.4.2 Gráficos de los resultados del estudio de simulación de la robustez del test de Cochran y Cox bajo lognormalidad



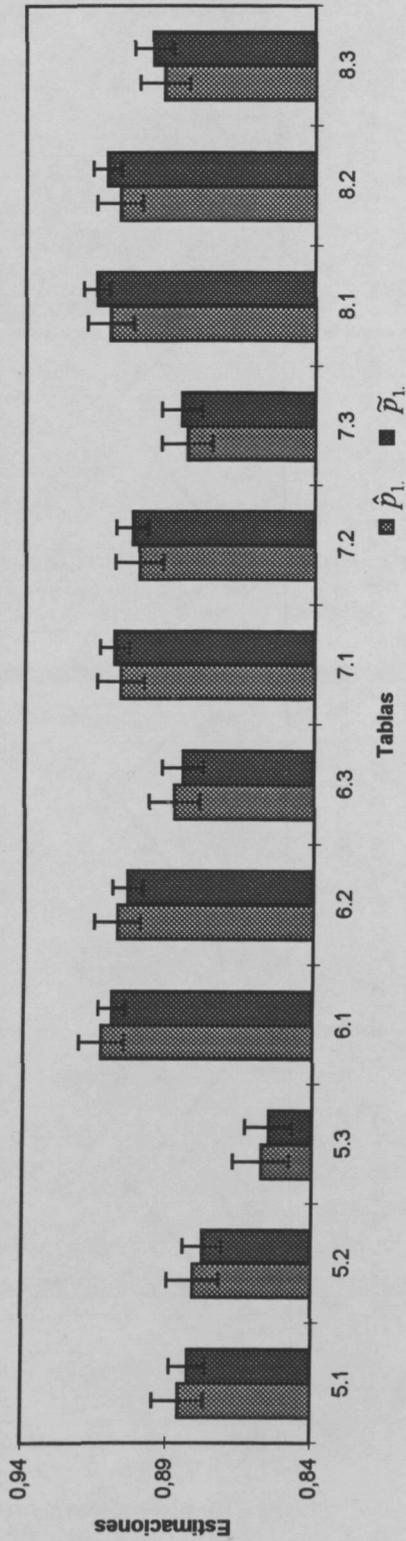
Comparación de estimaciones de las potencias del test de Cochran-Cox cuando la potencia conocida del test de control (t de Student) es de 0,55 para diferentes tamaños muestrales y varianzas



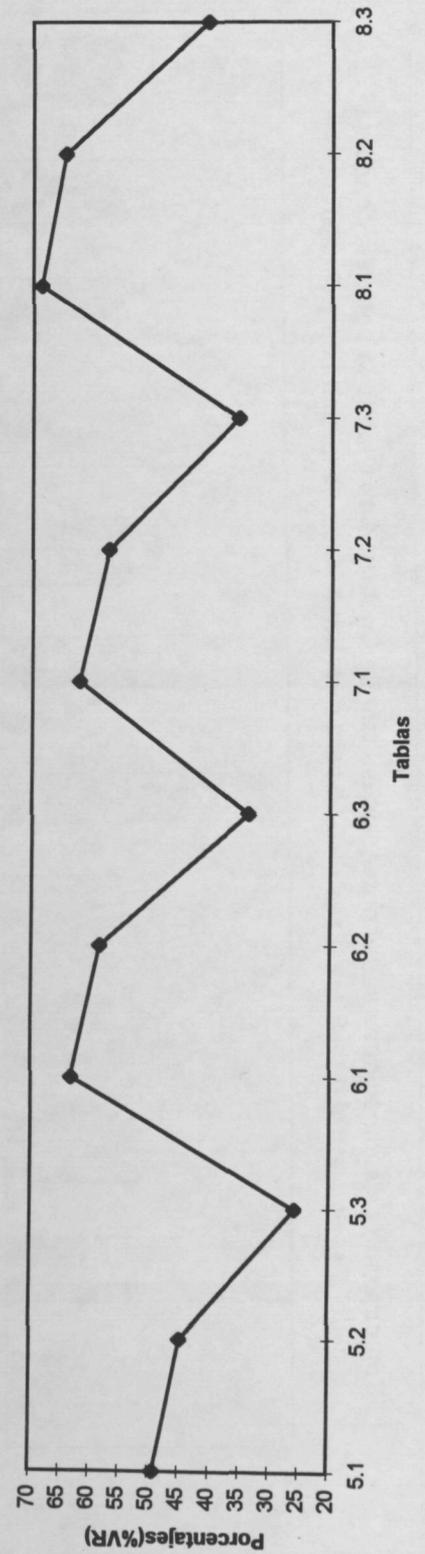
Comparación de estimaciones del % de red. de varianzas de las potencias del test de Cochran-Cox cuando la potencia conocida del test de control (t de Student) es de 0,55 para diferentes tamaños muestrales y varianzas



Comparación de estimaciones de las potencias del test de Cochran-Cox cuando la potencia conocida del test de control (t de Student) es de 0,92 para diferentes tamaños muestrales y varianzas



Comparación de estimaciones del % de red. de varianza de las potencias del test de Cochran-Cox cuando la potencia conocida del test de control (t de Student) es de 0,92 para diferentes tamaños muestrales y varianzas



### 5.1.5 Conclusiones

Se hará la comparación de las estimaciones de las potencias obtenidas bajo normalidad con sus correspondientes bajo lognormalidad en cada test por separado. A continuación se examinará las diferencias existentes en la robustez entre el test geodésico y el test de Cochran y Cox. Por último, se observará como afecta en la robustez las estimaciones de las potencias sin la técnica de reducción de la varianza ( $\hat{p}_{1.}$ ) o con ella ( $\tilde{p}_{1.}$ ).

En el test geodésico se produce una disminución de la estimación de la probabilidad de error de tipo I para el tamaño muestral menor (11) cuando las varianzas son semejantes, tan marcada es la disminución que es significativamente diferente del nivel de significación nominal del 0.05 (marcado en las tablas con el signo †). Sin embargo, cuando ( $\sigma_1^2 = 1.6 \sigma_2^2 = 0.4$ ) es algo superior al 0.05 y semejante al valor obtenido en las condiciones de normalidad. Así, se observa una tendencia creciente a aumentar la probabilidad de error de tipo I al incrementarse la diferencias entre las varianzas. Esta tendencia va disminuyendo al aumentar el tamaño muestral. Y las diferencias en la estimación de la probabilidad de error de tipo I entre normalidad y lognormalidad en los tamaños muestrales superiores (31, 61, 121) es cada vez menor. En el caso de la potencia del test de control, t de Student, igual a 0.55 hay un incremento en la estimación de la potencia del test geodésico con respecto a la situación de normalidad. Este incremento es menos notorio al aumentar el tamaño muestral. Por otro lado, se observa un incremento en la estimación de la potencia al aumentar la diferencia entre varianzas fijado un tamaño muestral pero esta diferencia se va difuminando al aumentar el tamaño muestral siendo semejante a las estimaciones de potencias bajo normalidad. Para la potencia de 0.92 en el t de Student se observa una disminución de la estimación de la potencia con respecto a la normalidad. Como pasaba antes, la diferencia mayor se produce en el tamaño muestral menor, que al aumentar esta se va menguando la diferencia. Además, se produce una tendencia a disminuir la estimación de la potencia al aumentar la diferencia entre varianzas pero, cuando se incrementa el tamaño muestral esta tendencia es cada vez más suave.

En el test de Cochran y Cox también se produce una disminución de la estimación de la probabilidad de error de tipo I para  $n_1 = n_2 = 11$  cuando las varianzas son semejantes y un incremento cuando las varianzas son muy diferentes. Esto queda reflejado en una tendencia a aumentar la estimación de la probabilidad de error de tipo I al aumentar la diferencia entre varianzas. Esta tendencia se va reduciendo

al aumentar el tamaño muestral y cada vez existen menos diferencias con las estimaciones obtenidas bajo normalidad. El carácter conservativo se observa en menos ocasiones que bajo normalidad, sólo para tamaños muestrales pequeños (11 y 31) y con la estimación obtenida aplicando la técnica de reducción de la varianza ( $\tilde{p}_1$ ) en el caso de  $n_1 = n_2 = 31$ . Cuando la potencia de la  $t$  de Student es de 0.55 se produce un aumento en la estimación de la potencia que se va reduciendo al incrementarse el tamaño muestral. Al aumentar la diferencia entre varianzas también se produce un incremento de la potencia que se va reduciendo hasta ser nula al aumentar el tamaño muestral. Para una potencia de 0.92 de la  $t$  de Student se produce una disminución de la estimación de la potencia del test de Cochran y Cox, mayor cuanto más grande sea la diferencia entre las varianzas. Esta tendencia a la disminución de la estimación de la potencia se va suavizando al aumentar el tamaño muestral.

El efecto que produce en la variación de la estimación de las potencias que las muestras sean originadas bajo la distribución lognormal es semejante en ambos tests. Así, en el caso de la estimación de la probabilidad de error de tipo I y cuando el test de control tiene una potencia intermedia (0.55) se produce un incremento en la estimación al aumentar diferencia entre varianzas que se suaviza al aumentar el tamaño muestral. En la primera situación (estimación de la prob. de error de tipo I) los valores obtenidos están por debajo de sus correspondientes estimaciones bajo la distribución normal cuando las varianzas son semejantes, aunque al aumentar el tamaño muestral estas diferencias desaparecen. En cambio, en la segunda situación (potencia del test  $t$  de Student igual a 0.55) la estimación está por encima de sus correspondientes estimaciones bajo la distribución normal y, también tienden a desaparecer estas diferencias al aumentar el tamaño muestral. Cuando el test de control tiene una potencia alta (0.92) se observa una reducción en la estimación de la potencia al aumentar la diferencia entre varianzas. Sin embargo, esta diferencia disminuye al aumentar el tamaño muestral. Aquí también, las estimaciones de las potencias eran inferiores a sus correspondientes bajo normalidad pero esta diferencia también se mengua al incrementar el tamaño muestral. Por tanto, la diferencias observadas son originadas sólo por los valores iniciales de las estimaciones de potencia bajo normalidad en cada test, así, pues, la distribución lognormal no produce ningún efecto distinto en los dos test estudiados.

Si ahora nos centramos en los márgenes de error de  $\hat{p}_1$  y  $\tilde{p}_1$ , etiquetados respec-

tivamente como  $\pm 1.959964\hat{\sigma}$  y  $\pm 1.959964\hat{\sigma}_{GS}$  se puede advertir que el primero será siempre más amplio aunque se irá igualando cuanto mayor sea la diferencia entre varianzas. Por otro lado, nos encontramos con un incremento del margen de error de  $\tilde{p}_1$  en el test de Cochran y Cox mayor que en el test geodésico con respecto a sus correspondientes resultados en la estimación de las potencias bajo normalidad. En cambio, el margen de error de  $\hat{p}_1$ , tanto en un test como en otro, puede aumentar como disminuir con respecto a los valores obtenidos en la estimación de las potencias bajo normalidad. Por otro lado, el carácter conservativo del test de Cochran y Cox es difícil de ver si se utiliza la frecuencia relativa ( $\hat{p}_1$ ) como estimación de la probabilidad de error de tipo I. Al aumentar la diferencia entre las varianzas fijado el tamaño muestral y la potencia del test t de Student se produce un decremento del porcentaje estimado de reducción de la varianza ( $\widehat{\%VR}$ ) y por tanto también en el porcentaje estimado de reducción del intervalo de confianza ( $\widehat{\%MER}$ ) no habiendo diferencias notorias entre ambos tests. En cuanto a cifras, la estimación del porcentaje de reducción de la varianza para ( $\sigma_1^2 = 1, \sigma_2^2 = 1$ ) oscila entre un 70%, para ( $\sigma_1^2 = 1.10, \sigma_2^2 = 0.90$ ) entre un 60% y para ( $\sigma_1^2 = 1.4, \sigma_2^2 = 0.6$ ) entre un 35%. Desde otro punto de vista, si comparamos las cifras de  $\widehat{\%VR}$  y de  $\widehat{\%MER}$  del estudio de la potencia obtenidas bajo lognormalidad con sus respectivas bajo normalidad, tanto en el test geodésico como en el test de Cochran y Cox, se advierte que el  $\widehat{\%VR}$  y el  $\widehat{\%MER}$  bajo lognormalidad están alrededor de un 70-75% del  $\widehat{\%VR}$  y de un 70% del  $\widehat{\%MER}$  bajo normalidad, respectivamente. Por tanto, es una disminución de un orden del 30% con respecto a los valores obtenidos de  $\widehat{\%VR}$  y  $\widehat{\%MER}$  bajo normalidad.

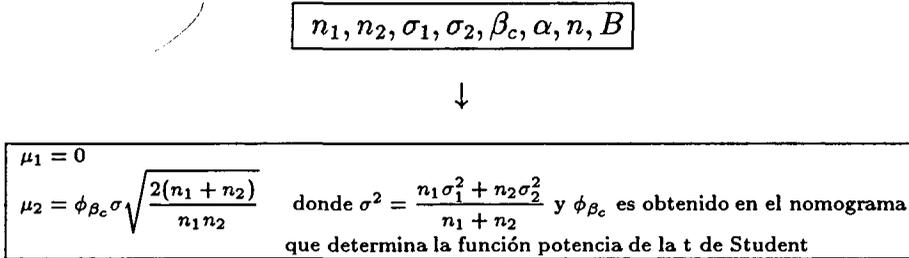
En resumen, el test geodésico es preferible para tamaños muestrales pequeños, aunque se puede dar para varianzas semejantes unas estimaciones de la probabilidad de error de tipo I inferior al nivel nominal mientras que no hay excesivas diferencias con el test de Cochran y Cox para tamaños muestrales mayores. En cualquier caso, es mejor utilizar la estimación de la potencia aplicando la técnica de reducción de la varianza ( $\tilde{p}_1$ ) que ( $\hat{p}_1$ ) aunque esta mejora se va diluyendo al aumentar la diferencia entre varianzas ya que disminuye la correlación existente entre la potencia del test de estudio, geodésico o Cochran y Cox, con la potencia del test de control, t de Student.

En cualquier caso, ambos tests son apreciablemente robustos frente a alejamientos de la normalidad de los datos. De hecho, tal como se verá en el capítulo 6 parece

claro que es más grave la no consideración de posibles diferencias entre varianzas al comparar medias que el hecho de no tener en cuenta posibles alejamientos de la normalidad.

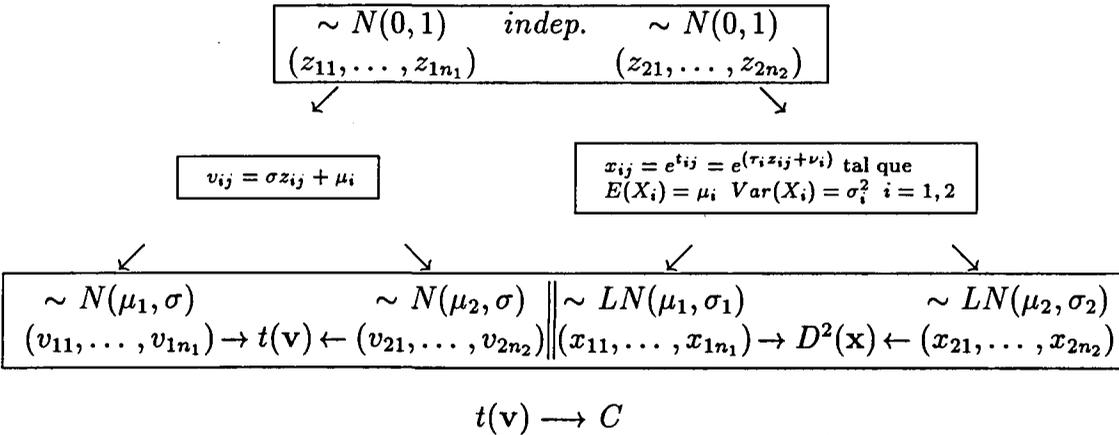
FIG.1: ALGORITMO PARA EL ESTUDIO DE LA ROBUSTEZ DEL TEST GEODÉSICO SUPONIENDO QUE LOS VALORES MUESTRALES PROVIENEN DE DISTRIBUCIONES LOGNORMALES

Introducción de parámetros

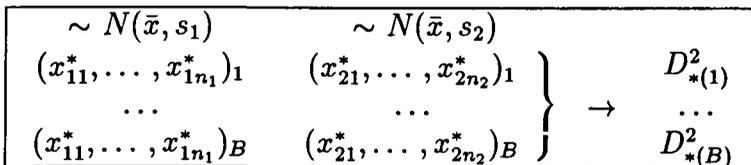


Repite desde j=1 hasta n:

Generación de las muestras



Bootstrap paramétrico sobre x



$$\frac{\#[D_*^2 \geq D^2(\mathbf{x})] + 1}{B + 1} \geq \alpha \rightarrow Y = 0$$

$$< \alpha \rightarrow Y = 1$$

Fin Repite

Del proceso anterior:

Reducción de la Varianza

$$\left. \begin{matrix} (Y_1, C_1) \\ \dots \\ (Y_n, C_n) \end{matrix} \right\} \rightarrow (\hat{p}_1, \hat{\sigma}^2, \%VR, \dots)$$

FIG.2: ALGORITMO PARA EL ESTUDIO DE LA ROBUSTEZ DEL TEST DE COCHRAN Y COX SUPONIENDO QUE LOS VALORES MUESTRALES PROVIENEN DE DISTRIBUCIONES LOGNORMALES

Introducción de parámetros

$$n_1, n_2, \sigma_1, \sigma_2, \beta_c, \alpha, n$$

↓

$$\begin{array}{l} \mu_1 = 0 \\ \mu_2 = \phi_{\beta_c} \sigma \sqrt{\frac{2(n_1 + n_2)}{n_1 n_2}} \end{array} \quad \text{donde } \sigma^2 = \frac{n_1 \sigma_1^2 + n_2 \sigma_2^2}{n_1 + n_2} \text{ y } \phi_{\beta_c} \text{ es obtenido en el nomograma} \\ \text{que determina la función potencia de la } t \text{ de Student}$$

Repite desde  $j=1$  hasta  $n$ :

Generación de las muestras

$$\begin{array}{cc} \sim N(0,1) & \text{indep.} & \sim N(0,1) \\ (z_{11}, \dots, z_{1n_1}) & & (z_{21}, \dots, z_{2n_2}) \end{array}$$

$$v_{ij} = \sigma z_{ij} + \mu_i$$

$$x_{ij} = e^{t_{ij}} = e^{(r_i z_{ij} + v_i)} \text{ tal que } \\ E(X_i) = \mu_i \quad \text{Var}(X_i) = \sigma_i^2 \quad i = 1, 2$$

$$\begin{array}{cccc} \sim N(\mu_1, \sigma) & \sim N(\mu_2, \sigma) & \sim LN(\mu_1, \sigma_1) & \sim LN(\mu_2, \sigma_2) \\ (v_{11}, \dots, v_{1n_1}) \rightarrow t(\mathbf{v}) \leftarrow (v_{21}, \dots, v_{2n_2}) & \parallel & (x_{11}, \dots, x_{1n_1}) \rightarrow D^2(\mathbf{x}) \leftarrow & (x_{21}, \dots, x_{2n_2}) \end{array}$$

$$t(\mathbf{v}) \rightarrow C$$

Test de Cochran y Cox

$$\begin{array}{l} t_{cc}(\mathbf{x}) = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(\hat{s}_1^2/n_1) + (\hat{s}_2^2/n_2)}} \\ t'_{\alpha/2} = \frac{t(\alpha/2, n_1 - 1)(\hat{s}_1^2/n_1) + t(\alpha/2, n_2 - 1)(\hat{s}_2^2/n_2)}{(\hat{s}_1^2/n_1) + (\hat{s}_2^2/n_2)} \end{array}$$

$$\begin{array}{l} |t_{cc}(\mathbf{x})| \geq t'_{\alpha/2} \rightarrow Y = 1 \\ < t'_{\alpha/2} \rightarrow Y = 0 \end{array}$$

Fin repite

Del proceso anterior:

Reducción de la Varianza

$$\left. \begin{array}{l} (Y_1, C_1) \\ \dots \\ (Y_n, C_n) \end{array} \right\} \rightarrow (\tilde{p}_1, \hat{\sigma}^2, \%VR, \dots)$$

## 5.2 Diseño del estudio por simulación de la robustez para muestras originadas bajo la distribución gamma

### 5.2.1 Introducción

Más que plantear un estudio de la robustez bajo la distribución gamma lo que se presenta es un ejemplo bajo el cual la técnica de reducción de la varianza no funciona adecuadamente y cómo puede ser solucionado este problema. Este mal funcionamiento es debido a la falta de correlación entre la variable respuesta y la variable control. Por tanto, hay que ir con precaución en la utilización de esta técnica ya que aunque se intente inducir la correlación a partir del uso de variables aleatorias comunes puede ocurrir que debido a la técnica utilizada para la generación de la distribución, en este caso la gamma, se puede perder la monotonía en el valor de la semilla utilizada en cada paso para generar cada uno de los valores, tanto de la distribución gamma como de la distribución normal, y con ello no alcanzar la correlación deseada, y por ende, no conseguir una mejor estimación con la técnica de reducción de la varianza. Como posible solución de este problema se plantea implementar una rutina para generar los valores de la gamma teniendo como base, en la mayoría de los casos, una adecuada transformación de valores distribuidos normalmente (Ahrens and Dieter, 1982).

Existen muchas situaciones en las cuales la variabilidad de la muestra se explica de una forma óptima con una distribución gamma, por ejemplo, en el estudio del tiempo transcurrido hasta la presencia de  $k$  acontecimientos, que no con una distribución normal. Sin embargo, se suele soslayar este hecho y además se supone tácitamente normalidad. Es por lo tanto lógico que alguien plantee la necesidad de un estudio de robustez, por ejemplo para el test geodésico y el de Cochran y Cox. Ahora veremos como se puede realizar este estudio por simulación y las dificultades con las que se toparía.

### 5.2.2 Procedimiento

Para ilustrar el problema y la posible solución se ha realizado el mismo proceso (simulación) dos veces. En el primer caso, se genera los valores que siguen una distribución gamma a partir de un paquete estadístico o numérico (opción más

común y la que provoca que la técnica de reducción de la varianza no sea satisfactoria). En particular, es la rutina G05FFF de las NAG (Mark 16 version) con un valor de  $\alpha > 1$  la cual emplea el algoritmo propuesto por Best (1978). Está basado en el uso de la distribución t de Student con dos grados de libertad como distribución base en la aplicación del método de rechazo. El segundo caso, que muestra la posible solución, implementa el algoritmo de Ahrens and Dieter para la generación de la distribución gamma que utiliza los valores anteriormente generados de la distribución normal necesarios para realizar la t de Student.

El algoritmo de Ahrens and Dieter aprovecha la propiedad siguiente: una adecuada transformación de la raíz cuadrada de una variable aleatoria gamma con  $\alpha \geq 1$  produce una función de densidad próxima a la normal estándar. Por tanto, una modificación de la técnica de rechazo que comience por el muestreo de valores de una distribución normal, será capaz de aceptar y transformar, rápidamente, la observación normal inicial, al menos en un 85 % de las veces, en un valor con distribución gamma.

El algoritmo para generar la distribución gamma estándar,  $G(\alpha,1)$ :

$$f(r) = \frac{1}{\Gamma(\alpha)} e^{-r} r^{\alpha-1} \quad \alpha > 1, r > 0$$

propuesto por Ahrens and Dieter es el siguiente:

**Algoritmo:**

0. Poner  $\alpha' \leftarrow 0$  y  $\alpha'' \leftarrow 0$  (en tiempo de compilación)
1. Si  $\alpha \neq \alpha'$  poner  $\alpha' \leftarrow \alpha$ ,  $s_2 \leftarrow \alpha - 1/2$ ,  $s \leftarrow \sqrt{s_2}$  y  $d \leftarrow 4\sqrt{2} - 12s = 5.65685424949228 - 12s$ .
2. Generar T (valor normal estándar). Poner  $X \leftarrow s+T/2$ . Si  $T \geq 0$  devolver  $X^2$ .
3. Generar U (valor Uniforme(0,1)). Si  $dU \leq T^3$  devolver  $X^2$ .
4. Si  $\alpha \neq \alpha'$  poner  $\alpha'' \leftarrow \alpha$ , y calcular  $q_0$ ,  $b$ ,  $\sigma$  y  $c$  como sigue:  
 $q_0 \leftarrow \sum q_k \alpha^{-k}$  (en lugar de su definición  $\ln \sqrt{2\pi} - \ln \Gamma(\alpha) - s_2 + s_2 \ln s_2$ );  
 Si  $1 \leq \alpha \leq 3.686$ :  $b \leftarrow 0.463 + s + 0.178s_2$ ,  $\sigma \leftarrow 1.235$ ,  
 $c \leftarrow 0.195/s - 0.079 + 0.16s$   
 Si  $3.686 < \alpha \leq 13.022$ :  $b \leftarrow 1.654 + 0.0076s_2$ ,  $\sigma \leftarrow 1.68/s + 0.275$ ,  $c \leftarrow 0.062/s + 0.024$   
 Si  $13.022 < \alpha \leq \infty$ :  $b \leftarrow 1.77$ ,  $\sigma \leftarrow 0.75$ ,  $c \leftarrow 0.1515/s$
5. Si  $X \leq 0$  ir al paso 8.
6. Poner  $V \leftarrow T/(s + s)$  y calcular Q como sigue:  
 Si  $|V| > 1/4$ :  $Q \leftarrow q_0 - sT + T^2/4 + (s_2 + s_2) \ln(1 + V)$   
 Si  $|V| \leq 1/4$ :  $Q \leftarrow q_0 + (T^2/2) \sum a_k V^k$
7. Si  $\ln(1 - U) \leq Q$  devolver  $X^2$ .
8. Generar E (valor exponencial estándar) y U (valor Uniforme(0,1)). Poner  $U \leftarrow U+U-1$  y  $T \leftarrow b+E\sigma \text{sign}(U)$
9. Si  $T \leq \tau_1 = -0.71874483771719$  ir al paso 8.
10. Poner  $V \leftarrow T/(s + s)$  y calcular Q como en el paso 6.
11. Si  $Q \leq 0$  o si  $c|U| > (\exp(Q) - 1) \exp(E - T^2/2)$  ir al paso 8. (si  $Q \leq 1/2$  el factor  $\exp(Q) - 1$  se calcula como  $\sum e_k Q^k$ )
12. Poner  $X \leftarrow s + T/2$  y devolver  $X^2$ .

Se puede dividir el algoritmo en tres partes. Una de inicialización, que sólo se ejecuta cuando existe un cambio del valor del parámetro  $\alpha$ , provocando que se deban de recalcular las cantidades  $s_2$ ,  $s$ ,  $d$ ,  $q_0$ ,  $b$ ,  $\sigma$  y  $c$  (pasos 1 y 4). La evaluación de  $q_0$  en el paso 4 nunca se hace a partir de su definición directamente, ya que sería lento y poco preciso, sino a partir de una aproximación polinómica. Los coeficientes utilizados en la implementación están en la tabla A (además, se incluyen otros coeficientes que son necesarios en el paso 6 y 11). La segunda parte, corresponde a los pasos 2 y 3. La mayor parte de los valores generados salen de estos pasos que utilizan el cuadrado de una transformación lineal del valor normal generado para obtener el valor con la distribución gamma. La última parte, más complicada, menos rápida y a la que se llega en menos ocasiones, corresponde a los pasos 5 hasta 12.

Tabla A. Coeficientes de la aproximación polinómica.

$ \epsilon  < 2.6 \times 10^{-10}$		$ \epsilon  < 2.1 \times 10^{-9}$		$ \epsilon  < 3.1 \times 10^{-11}$	
$q_1$	0.0416666664	$a_1$	0.333333333	$e_1$	1.000000000
$q_2$	0.0208333723	$a_2$	-0.249999949	$e_2$	0.499999994
$q_3$	0.0079849875	$a_3$	0.199999867	$e_3$	0.166666848
$q_4$	0.0015746717	$a_4$	-0.166677482	$e_4$	0.041664508
$q_5$	-0.0003349403	$a_5$	0.142873973	$e_5$	0.008345522
$q_6$	0.0003340332	$a_6$	-0.124385581	$e_6$	0.001353826
$q_7$	0.0006053049	$a_7$	0.110368310	$e_7$	0.000247453
$q_8$	-0.0004701849	$a_8$	-0.112750886		
$q_9$	0.0001710320	$a_9$	0.104089866		

Una vez obtenido un valor con distribución gamma estándar,  $r$ , es muy fácil de obtener cualquier valor con distribución gamma,  $G(\alpha, \beta)$ :

$$f(n) = \frac{1}{\Gamma(\alpha)\beta^\alpha} e^{-r/\beta} r^{\alpha-1} \quad \alpha > 1, \beta > 0, n > 0$$

a partir de la sencilla multiplicación  $n = \beta r$ .

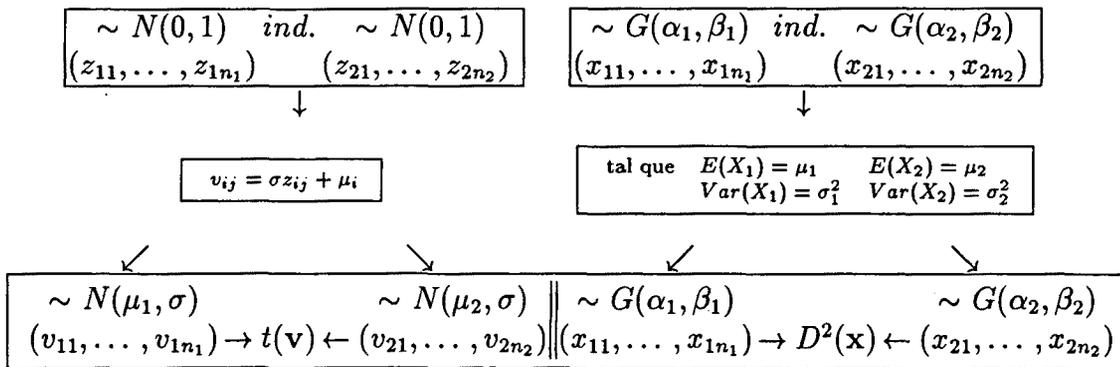
El procedimiento seguido es similar al indicado en el anterior estudio de robustez bajo la suposición de lognormalidad. La diferencia está en la generación de las muestras (ver fig. 3). En ambos casos, rutina G05FFF de las NAG o implementación del algoritmo de Ahrens and Dieter, se debe de conseguir aquellos valores de los parámetros  $\alpha$  y  $\beta$  de la distribución gamma tales que  $E(X_i) = \mu_i$  y  $Var(X_i) = \sigma_i^2$  para  $i = 1, 2$ . Por otro lado, la consecución de los valores normales

que serán utilizados para determinar el test t de Student de control se hará de la manera habitual; a partir del vector  $(z_{11}, \dots, z_{1n_1}; z_{21}, \dots, z_{2n_2})$  de  $n_1 + n_2$  valores normales estándar utilizando la transformación  $v_{ij} = \sigma z_{ij} + \mu_i$ ,  $i = 1, 2$  donde  $\sigma^2$  sera el promedio ponderado de las varianzas con respecto de los tamaños muestrales.

FIG.3: GENERACIÓN DE LOS VECTORES MUESTRALES PARA EL ESTUDIO DE LA ROBUSTEZ DEL TEST GEODÉSICO Y EL DE COCHRAN Y COX BAJO LA DISTRIBUCIÓN GAMMA

Repite desde  $j=1$  hasta  $n$ :

Generación de las muestras



### 5.2.3 Resultados

Se muestran las simulaciones para el estudio de la potencia del test geodésico y del test de Cochran y Cox con el tamaño muestral menor,  $n_1 = n_2 = 11$ , con tres tipos de parejas varianzas:  $(\sigma_1^2 = \sigma_2^2 = 1)$ ,  $(\sigma_1^2 = 1.10, \sigma_2^2 = 0.90)$  y  $(\sigma_1^2 = 1.6, \sigma_2^2 = 0.4)$ , y tres potencias (0.05, 0.55 y 0.92) para comparar los resultados obtenidos con la rutina G05FFF y la implementación del algoritmo de Ahrens and Dieter. En ambos casos, se parten de las mismas semillas.

Las tablas 9 y 10 corresponden al estudio de la potencia del test geodésico y las tablas 11 y 12 al test de Cochran y Cox. Las primeras tablas, 9 y 11, se generan los valores con distribución gamma a partir del algoritmo de Ahrens and Dieter mientras que las otras tablas, 10 y 12, se generan a partir del algoritmo de Best (rutina G05FFF).

Tablas 9 y 10. Nivel de significación, potencia y grado de reducción de la varianza en el estudio de la robustez del test geodésico definido mediante (4.15) para muestras obtenidas a partir de una distribución Gamma. Datos de la simulación:  $n = 5000$ ,  $n_1 = n_2 = 11$ ,  $B = 1000$ .

† =significativamente diferente del nivel de significación nominal.

‡ =significativamente diferente del valor correspondiente en el test Cochran y Cox.

Tabla 9: Utilizando el algoritmo propuesto por Ahrens y Dieter.

$\Delta\mu$	$\sigma_1^2 = 1.0$		$\sigma_2^2 = 1.0$		%VR	%MER
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$		
0	0.0492†	0.005996	0.04978†	0.005284	22.32	11.87
0.927147567	0.5638†	0.013747	0.56102†	0.010954	36.51	20.32
1.50755596	0.9062†	0.008082	0.90419†	0.007128	22.22	11.81

$\Delta\mu$	$\sigma_1^2 = 1.10$		$\sigma_2^2 = 0.90$		%VR	%MER
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$		
0	0.0406††	0.005471	0.0411††	0.004896	19.91	10.51
0.927147567	0.5632†	0.013750	0.5606†	0.011331	32.09	17.59
1.50755596	0.9006†	0.008294	0.89869†	0.007464	19.01	10.01

$\Delta\mu$	$\sigma_1^2 = 1.6$		$\sigma_2^2 = 0.4$		%VR	%MER
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$		
0	0.0562†	0.006384	0.05656††	0.006142	7.44	3.79
0.927147567	0.5548†	0.013777	0.55271†	0.012283	20.52	10.85
1.50755596	0.8604	0.009607	0.85875	0.009090	10.48	5.39

Tabla 10: Utilizando un paquete de cálculo numérico (rutina G05FFF de las NAG).

$\Delta\mu$	$\sigma_1^2 = 1.0$		$\sigma_2^2 = 1.0$		%VR	%MER
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$		
0	0.0418††	0.005548	0.04178††	0.005547	0.03	0.01
0.927147567	0.5512†	0.013788	0.55106†	0.013785	0.04	0.02
1.50755596	0.9114†	0.007877	0.91142†	0.007878	-0.02	-0.01

$\Delta\mu$	$\sigma_1^2 = 1.10$		$\sigma_2^2 = 0.90$		%VR	%MER
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$		
0	0.0464†	0.005831	0.0464†	0.005833	-0.05	-0.02
0.927147567	0.5558†	0.013774	0.55584†	0.013776	-0.03	-0.02
1.50755596	0.8936†	0.008548	0.89369†	0.008544	0.09	0.04

$\Delta\mu$	$\sigma_1^2 = 1.6$		$\sigma_2^2 = 0.4$		%VR	%MER
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$		
0	0.0624††	0.006705	0.06241††	0.006707	-0.06	-0.03
0.927147567	0.5432†	0.013809	0.54325†	0.013811	-0.03	-0.01
1.50755596	0.8562	0.009727	0.85626	0.009726	0.02	0.01

Tablas 11 y 12. Nivel de significación, potencia y grado de reducción de la varianza en el estudio de la robustez del test Cochran y Cox para muestras obtenidas a partir de una distribución Gamma. Datos de la simulación:  $n = 5000$ ,  $n_1 = n_2 = 11$ .

† =significativamente diferente del nivel de significación nominal.

‡ =significativamente diferente del valor correspondiente en el test geodésico.

Tabla 11: Utilizando el algoritmo propuesto por Ahrens y Dieter.

$\Delta\mu$	$\sigma_1^2 = 1.0 \quad \sigma_2^2 = 1.0$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\bar{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0386†‡	0.005340	0.03908†‡	0.004796	19.35	10.19
0.927147567	0.5108‡	0.013857	0.50793‡	0.010876	38.41	21.52
1.50755596	0.8872‡	0.00877	0.88488‡	0.007566	25.57	13.73

$\Delta\mu$	$\sigma_1^2 = 1.10 \quad \sigma_2^2 = 0.90$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\bar{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0322†‡	0.004894	0.03262†‡	0.004456	17.09	8.94
0.927147567	0.5192‡	0.013850	0.51646‡	0.011152	35.17	19.48
1.50755596	0.8822‡	0.008937	0.88012‡	0.008008	19.71	10.39

$\Delta\mu$	$\sigma_1^2 = 1.6 \quad \sigma_2^2 = 0.4$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\bar{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0462‡	0.005819	0.04652‡	0.005618	6.78	3.45
0.927147567	0.5258‡	0.013842	0.52365‡	0.012257	21.59	11.45
1.50755596	0.8498	0.009904	0.84802	0.009312	11.59	5.97

Tabla 12: Utilizando un paquete de cálculo numérico (rutina G05FFF de las NAG).

$\Delta\mu$	$\sigma_1^2 = 1.0 \quad \sigma_2^2 = 1.0$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\bar{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0332†‡	0.004967	0.03319†‡	0.004966	0.01	0.01
0.927147567	0.5058‡	0.013860	0.50565‡	0.013856	0.06	0.03
1.50755596	0.8898‡	0.008681	0.88982‡	0.008681	-0.02	-0.01

$\Delta\mu$	$\sigma_1^2 = 1.10 \quad \sigma_2^2 = 0.90$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\bar{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0368†‡	0.005219	0.03679†‡	0.005219	0.00	0.00
0.927147567	0.5094‡	0.013858	0.50943‡	0.013861	-0.04	-0.02
1.50755596	0.8760‡	0.009136	0.87611‡	0.009132	0.10	0.05

$\Delta\mu$	$\sigma_1^2 = 1.6 \quad \sigma_2^2 = 0.4$					
	$\hat{p}_1$	$\pm 1.959964\hat{\sigma}$	$\bar{p}_1$	$\pm 1.959964\hat{\sigma}_{GS}$	$\%VR$	$\%MER$
0	0.0546‡	0.006298	0.05462‡	0.006300	-0.06	-0.03
0.927147567	0.5210‡	0.013848	0.52105‡	0.013850	-0.02	-0.01
1.50755596	0.8436	0.010069	0.84368	0.010068	0.03	0.02