

*D. Aguilar*

# Universidad Nacional Autónoma de México

FACULTAD DE CIENCIAS



---

LA SUPOSICION DE HOMOCEDASTICIDAD  
EN EL MODELO DE REGRESION LINEAL

T E S I S  
QUE PARA OBTENER EL TITULO DE:  
A C T U A R I O  
P R E S E N T A  
ALVARO AGUILAR GARCIA

NOVIEMBRE 81



Universidad Nacional  
Autónoma de México



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

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

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

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

## INDICE

	Pág.
PROLOGO	i
INTRODUCCION	ii
CAPITULO	
I. EFECTOS DE VIOLACION DE HOMOCEDASTICIDAD EN REGRESION LINEAL	1
El Modelo y las suposiciones básicas	1
Estimadores	7
Residuales	10
Métodos Gráficos para análisis de residuales	10
Efectos de violación a la suposición a homocedasticidad	17
II. PRUEBAS PARA LA DETECCION DE HETEROCEDASTICIDAD	20
Bartlett	22
Scheffé	24
Hartley	27
Cadwell	28
Levene	30
Goldfeld y Quandt	32
Miller	35
Glejser	40
Layard	43
Harvey y Phillips	45
Zroeter	48

CAPITULO	Pág.
III. SIMULACIONES	51
Resumen de resultados	68
LISTA DE TABLAS	73
BIBLIOGRAFIA	85

## PROLOGO

El por qué de tratar siempre de que se cumplan las suposiciones que se hacen o preguntarse ¿Qué sucede si se violan las suposiciones?, o ¿Cuáles son los efectos de violación a las suposiciones?, que finalmente se reflejan en los resultados de aplicar el análisis (o técnica) estadístico, fué lo que motivó el presente trabajo. Este se desarrolló sobre la suposición de igualdad de varianzas (homocedasticidad), un tópico dentro de los Modelos Lineales y en particular en Análisis de Regresión Lineal (la parte en que se desarrolla).

Por otra parte, puesto que en general, en cursos ordinarios de Análisis de Regresión o relativos a éste, se tratan muy pocas pruebas para detectar si existe violación de esta suposición (heterocedasticidad); se tuvo interés en saber si estas pruebas son realmente buenas o si habrían otras mejores o más fáciles para calcular. Por esto, se presenta a manera de informe lo que se pudo investigar.

Para terminar, se hace la aclaración que en todo el trabajo se presupone conocimiento mínimo de Análisis de Regresión.

## INTRODUCCION

Tanto en el Análisis de Regresión Lineal como en los Modelos Lineales en general, es indispensable hacer ciertas suposiciones las cuales validan todos los resultados e inferencias que se obtienen de aplicar tales técnicas estadísticas.

Dado el carácter de importancia que tienen las suposiciones, es obvio que en todo estudio que se lleve a cabo, debe uno tratar de que se cumplan, si no de forma estricta, al menos aproximadamente. Aún cuando las técnicas para verificar si se cumplen o no, son variadas. Sin embargo, no existe un método general para verificar todas las suposiciones al mismo tiempo. Estas van desde un análisis rápido (como son los métodos gráficos) hasta pruebas de significancia muy sofisticada pero que son mas confiables.

En caso de no cumplirse las suposiciones, lo recomendable es hacer transformaciones a los datos, para que estos tengan un comportamiento de acuerdo a los supuestos. Dado que este trabajo no trata a ningún nivel las transformaciones, se cita en él bibliografía para que el lector interesado pueda recurrir a ella.

Lo que aquí se presenta no es nada original, es mas bien una recopilación de información sobre los efectos de violación a homocedasticidad y como se puede detectar dicha violación. Se trató por lo tanto de seguir una secuencia lógica para la presentación de la misma. Con esto en mente, se procedió de la siguiente forma:

En el capítulo I, se hace un repaso del modelo de Regresión Lineal empezando por las suposiciones, se definen los residuales y se menciona brevemente la importancia de ellos dentro del análisis. Se da una explicación rápida de los Métodos Gráficos y finalmente, se exponen los efectos que acarrea la violación de la suposición de homocedasticidad. En el siguiente, capítulo II, se describen una serie de pruebas para detectar heterocedasticidad, la cual empieza por la prueba de Bartlett y la prueba de Scheffé (la razón de esto es porque éstas son las pruebas clásicas), las siguientes se citan en orden cronológico. Debe señalar que en algunas de las pruebas no se pudieron obtener los artículos originales. Sin embargo, se trató de darles una explicación clara.

En el capítulo III, se hizo una revisión de los experimentos (para comparar la potencia de las pruebas) que han realizado diferentes personas. Lo que se persigue con esto, es dar una idea del comportamiento de las pruebas y se incluye un resumen de los resultados obtenidos sobre las pruebas.

Finalmente se incluyen las tablas que se necesitan para llevar a efecto las pruebas y se cita la bibliografía para aquellos lectores interesados en un tratamiento mas formal de cada prueba.

## CAPITULO I

### EFFECTOS DE VIOLACION DE HOMOCELASTICIDAD EN REGRESION LINEAL

En este capítulo, se contemplan los efectos de la violación de homocedasticidad, que se pudieron obtener de toda la bibliografía consultada, en el Modelo de Regresión Lineal y las formas en que ésta pudiera ocurrir.

Para tal fin, se consideró necesario dividir en cinco secciones el capítulo. Es importante hacer notar que todas las secciones se tratan a nivel introductorio excepto la quinta en donde se tratan los efectos de heterocedasticidad. Para respaldar todo lo que se escribe, se cita bibliografía en donde el lector interesado en profundizar podrá satisfacer su inquietud.

En la primera sección se enuncian las suposiciones básicas; en la segunda se trata la idea de estimador y para que sirve; en la tercera se da la definición de lo que es un residual y en la cuarta se mencionan, de manera somera, los métodos gráficos que entre otras cosas ayudan a saber si se cumplen las suposiciones. Finalmente en la última sección se ven los efectos de violación a la suposición de homocedasticidad como se mencionó al principio.

1) Muchas veces, tenemos que hacer abstracciones de la realidad. Por esta razón, nosotros nos encontramos en la necesidad de hacer representaciones por medio de algún modelo matemático que explique lo mejor posible el fenómeno en estudio.

Contemplemos las siguientes situaciones:

1.1) Supóngase que nos piden que demos la distancia a que se encuentra una partícula  $x$  al tiempo  $t$ , que sale de un punto dado a una velocidad constante  $v$ .

1.2) En que tiempo estará en el suelo un objeto que se tira verticalmente de una altura  $d$ .

1.3) Nos encomiendan estudiar el número de entradas de personas ( venta de boletos ) a un estadio.

Si nos ponemos a pensar; el número de entradas de personas ( venta de boletos ) vista como un fenómeno, se intuye que puede ser explicado por:

i) El número de gente que habita el lugar en donde se encuentra instalado el estadio.

ii) Que tan importantes son los equipos que se encuentran ( ya sean de beisbol, fútbol, fútbol americano, etc. y la importancia debida al número de partidos ganados en esa temporada. ).

iii) La cantidad de unidades de transporte.

iv) La época en que se llevan a cabo los encuentros, (en el sentido climatológico ), etc.

1.4) Supóngase que para fijar el impuesto de los negocios con ingresos mayores a \$ 2,000,000.00 , La Secretaría de Hacienda y Crédito Público, necesita hacer un estudio sobre las ventas de las tiendas en todo el país. Se piensa que los factores que influyen y por lo tanto se tomarán en cuenta son:

i) El área del local ( del negocio ).

ii) La cantidad de mercancía que cada tienda posee.

iii) La disponibilidad para atender al público, etc.

Los modelos que servirían para representar a cada uno de los fenómenos en estudio podrían ser de la forma siguiente:

1.-  $d = vt$  ; para el ejemplo 1.1

2.-  $t = 2 d/g$  ; para el caso 1.2, donde:  
 $d =$  altura y  
 $g = 9.8 \text{ m/seg}^2$  , la aceleración constante debida a la gravedad.

3.-  $Y = 3 \frac{X_1 X_2^2}{X_3}$  ; para el ejemplo 1.3

4.-  $Y = \left[ 8 \cdot 10^{20} X_1 \arctan(e^{3X_2}) \right] X_3^3 \cdot 10^{-16} \cdot \frac{1}{3C} R(X_1, X_2, X_3)$  ;

donde C es un factor de corrección y R una función desconocida de  $X_1$ ,  $X_2$  y  $X_3$ .

ó sencillamente como

5.-  $Y = X_1 \beta_1 + X_2 \beta_2 + X_3 \beta_3 + g(X_1, X_2, X_3)$

que en general puede escribirse como

6.-  $Y = X_1 \beta_1 + X_2 \beta_2 + \dots + X_{p/p} \beta_{p/p} + g(X_1, \dots, X_p)$

Podríamos de esta forma seguir enumerando modelos para describir tal ó cual fenómeno.

Por otra parte a los fenómenos se les puede clasificar en dos clases:

i) Determinísticos.

ii) No-Determinísticos.

Los no-deterministas ó aleatorios son aquellos en los que no se conoce de antemano lo que sucederá, situación que no sucede en los primeros.

Como una aclaración, debemos notar que en los fenómenos Determinísticos la relación entre el modelo y el fenómeno está completamente definida, tal es el caso de los modelos (1-3). Los No-determinísticos, contrario a lo anterior, están sujetos a cambios inesperados del fenómeno, es decir, la relación entre el modelo y el fenómeno no es exacta como en los modelos (4-6).

En este trabajo, se considerarán únicamente fenómenos no-determinísticos.

Dada la facilidad de cálculos y más importante aún, la facilidad para la interpretación de los resultados, (característica indispensable que debe poseer cualquier modelo), supondremos que la relación entre el fenómeno y las variables explicativas es de la forma (6), es decir lineal en las  $\beta$ 's.

Ahora, pudiera ocurrir que no tomáramos en cuenta todos los factores que influyen para explicar el fenómeno, ya sea por cuestión de costo, tiempo o bien por no conocerlos, entonces siempre existirá un término de discrepancia o fluctuación, esto puede verse en los ejemplos 1.3 y 1.4. En el ejemplo 1.3 un factor que pudiera afectar sería la transmisión por televisión de los encuentros o en el ejemplo 1.4 pudiera afectar el nivel de ingresos de los habitantes de la localidad donde se hiciera el estudio, etc. Dado esto, lo recomendable es incluir las variables explicatorias más relevantes para el estudio —denotadas por  $X_1, X_2, \dots, X_p$  en (6)— de la variable respuesta  $Y^1$ .

Por otro lado, es hasta cierto punto lógico, el pensar

---

1.- Generalmente a  $X_1, \dots, X_p$  se les llaman variables independientes y a  $Y$  variable dependiente. En este trabajo se usará esta convención.

que tanto las variables independientes como la dependiente estén sujetas a error, ya sea debido al vaciado de la información; que al tomar la muestra los interrogados no den la respuesta correcta; si se tiene que medir, que el observador no tome bien las medidas, o bien, los errores inherentes a las unidades de medida, etc.

Por lo tanto debe existir un término en el modelo que refleje estas discrepancias, al que denotaremos por  $u$ . El modelo general que trabajaremos será el siguiente:

$$7.- \quad Y = X_1\beta_1 + X_2\beta_2 + \dots + X_p\beta_p + u$$

donde  $X_1, X_2, \dots, X_p$  <sup>(1)</sup> son las variables independientes,  $\beta_1,$

$\beta_2, \dots, \beta_p$  son parámetros desconocidos los cuales debemos estimar y que funcionan como indicadores de la importancia de las variables independientes dentro del modelo;  $u$  es el término de discrepancia que no es observable por tanto es considerado como variable aleatoria, y  $Y$  es la variable dependiente la cual hereda la aleatoriedad del término  $u$ .

Si consideramos (7) como una observación de la variable aleatoria  $Y$ , entonces podemos tomar  $n$  observaciones repetidas de  $Y$  y  $X_1, \dots, X_p$ , es decir, tendríamos para la  $i$ -ésima observación

$$y_i = x_{1i}\beta_1 + x_{2i}\beta_2 + \dots + x_{pi}\beta_p + u_i$$

donde  $x_{ji}$  representa la  $i$ -ésima observación de la  $j$ -ésima variable independiente y  $u_i$  será la discrepancia introducida en la  $i$ -ésima observación. Esto nos permite escri -

---

(1) Generalmente el valor de  $X_1$  es igual a uno.

bir el modelo de forma más general, es decir:

$$8.- \quad \underline{Y} = \underline{X}\underline{\beta} + \underline{U}$$

donde  $\underline{Y}$  es un vector aleatorio de dimensión  $n \times 1$ ,

$\underline{X}$  es una matriz de variables independientes de dimensión  $n \times p$ ,

$\underline{\beta}$  es el vector de parámetros desconocidos de dimensión  $p \times 1$ ,

$\underline{U}$  es el vector de discrepancias de dimensión  $n \times 1$ .

En el desarrollo de este trabajo trataremos solamente el caso en que el rango de  $\underline{X}$  es  $p$ , es decir,  $\underline{X}$  es de rango completo. Para un tratamiento sencillo del modelo, es indispensable hacer ciertas suposiciones, puesto que sin ellas, tendríamos demasiadas dificultades para hacer inferencias sobre la población o fenómeno bajo estudio, en otras palabras, sin dichas suposiciones, no podríamos evaluar estadísticamente la importancia de cada una de las variables  $x_j$  ( $j= 1, \dots, p$ ) en la explicación del fenómeno en estudio.

Las suposiciones básicas son: (+)

- i) La relación funcional es lineal en los parámetros.
- ii) Los términos de discrepancia tienen media cero y varianza constante  $\sigma^2$ , i.e.  
$$E(u_i) = 0 \quad E(u_i^2) = \sigma^2 \quad \forall i$$
- iii) Son no correlacionados, i.e.  $E(u_i u_j) = 0 \quad \forall i \neq j$
- iv) Los términos de discrepancia se distribuyen normales con media cero y varianza  $\sigma^2$ , i.e.  
$$u_i \sim N(0, \sigma^2)$$

Debemos tener presente que las condiciones (iii) y (iv) implican independencia en los términos de error.

---

(+) Una excelente motivación de las suposiciones para el caso de análisis de varianza puede verse en C. Eisenhart (1947).

2) Hechas las suposiciones, estamos posibilitados para obtener los estimadores de los parámetros  $\underline{\beta}$  y  $\sigma^2$ . La idea para obtener estimadores de los parámetros  $\underline{\beta}$  y  $\sigma^2$  es minimizar la distancia entre las observaciones  $y_1$  y la función lineal de las  $x_j$  ( q.v. C. Eisenhart (1964) ), gráficamente se aprecia de la siguiente forma:

Si  $p = 2$  ,  $x_1 = 1$

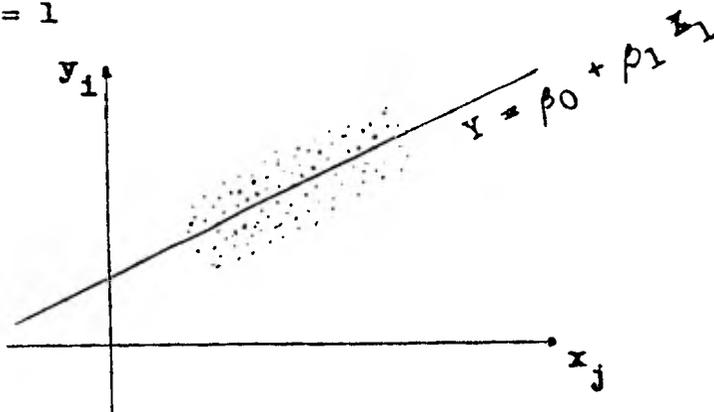


FIG. 1

Si  $p = 3$  ,  $x_1 = 1$

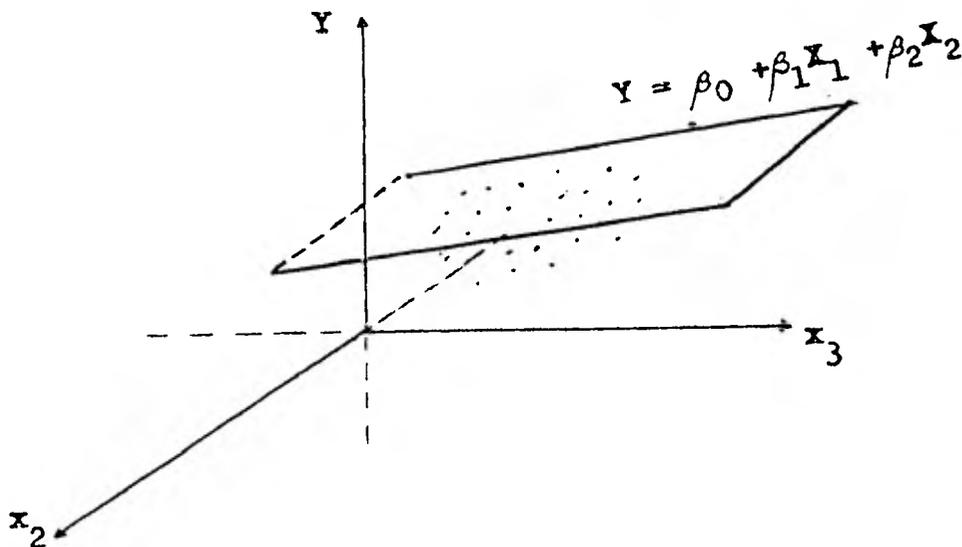


FIG. 2

En general, si tenemos un vector de observaciones de dimensión  $n$  y un conjunto de variables independientes  $X_1, \dots, X_p$  de las cuáles tenemos  $n$  observaciones, éstas últimas constituirán una matriz de dimensión  $n \times p$  cuyo rango será  $p$ . Entonces el modelo que tendremos será como (8), es decir:

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{U}$$

Veámoslo geoméricamente

Sea  $\underline{Y} \in V_n$  un espacio vectorial de dimensión  $n$ , las columnas de  $X$  puesto que son independientes entre sí, generan un espacio vectorial de dimensión  $p$  ( $V_p$ ) contenido en  $V_n$ . ( $V_p \subset V_n$ )

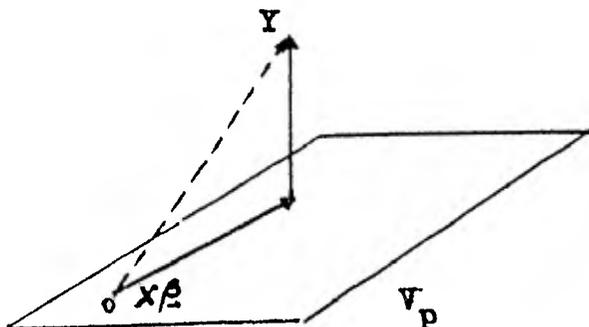


FIG. 3

Como queremos minimizar la distancia entre el punto  $\underline{Y}$  y  $V_p$ , tenemos que minimizar la siguiente norma al cuadrado:

$$9.- \quad \|\underline{Y} - \underline{X}\underline{\beta}\|^2$$

donde el mínimo se alcanza cuando  $\underline{X}\underline{\beta}$  es la proyección ortogonal de  $\underline{Y}$  sobre  $V_p$ . Si denotamos por  $\hat{\underline{\beta}}$  al punto que

cumple con esta propiedad, llegamos a que:

$$10.- \min \| Y - X \underline{\beta} \| ^2 = \| Y - X \hat{\underline{\beta}} \| ^2$$

donde el valor de  $\hat{\underline{\beta}}$  está dado por:

$$11.- \hat{\underline{\beta}} = (X'X)^{-1}X'Y$$

De aquí es evidente que el estimador de  $\underline{\beta}$  es lineal en  $\underline{Y}$  y puede mostrarse que es insesgado y de mínima varian<sup>1</sup>za. Por otro lado y sin salir de esta parte, la cantidad

$$\| Y - X \hat{\underline{\beta}} \| ^2$$

es llamada la SUMA DE CUADRADOS DEBIDA A LOS RESIDUALES y puesto que esta norma es mínima, provee un buen estimador de  $\sigma^2$ .

El estimador de  $\sigma^2$  obtenido por el método de mínimos cuadrados, ( es el método que tratamos grosso modo en la parte anterior inmediata), está dado por la siguiente expresión :

$$12.- \hat{\sigma}_{M.C.}^2 = \frac{\| Y - X \hat{\underline{\beta}} \| ^2}{n-p}$$

y éste es un estimador insesgado de la varianza. [ para un excelente tratamiento geométrico del método de Mínimos Cuadrados (obtención y propiedades de los estimadores) véase Scheffé (1959), págs. 8-23; Seber (1977), págs. 42-65 ] .

A diferencia del método de Mínimos Cuadrados, el método de Máxima Verosimilitud necesita además la suposición de normalidad, también provee estimadores con características muy deseables. Aunque el estimador de  $\underline{\beta}$  coincide con el estimador de  $\underline{\beta}$  por mínimos cuadrados; el estimador de  $\sigma^2$  no coincide con el estimador de  $\sigma^2$  por mínimos cua-

---

1.- Para esta demostración, bastan las suposiciones i-iii

drados, ya que el estimador máximo verosímil de  $\sigma^2$  es sesgado. Por cuestiones teóricas, es conveniente utilizar el estimador insesgado de  $\sigma^2$ .

Una vez obtenidos los estimadores (por cualquier método), utilizando la suposición de normalidad, podemos obtener la distribución de cada uno de ellos. En esta situación nos encontramos posibilitados para hacer entre otras las siguientes cosas:

- a) Obtener intervalos de confianza para cada  $\beta_j$ .
  - b) Obtener intervalos de confianza para  $\sigma^2$ .
  - c) Obtener intervalos de confianza para futuras predicciones de  $Y$ .
  - d) Obtener elipsoides de confianza para  $\beta$ .
  - e) Probar hipótesis sobre el vector  $\beta$ .
  - f) Medir estadísticamente la importancia de las variables independientes dentro del modelo.
- etc.

3) Recordando la definición de la suma de cuadrados debida a los residuales y denotando por  $\hat{Y}$  al vector estimado, tenemos:

$$\|Y - X\hat{\beta}\|^2 = \|Y - \hat{Y}\|^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

entonces a nivel de una observación, definimos a

$$13.- \quad Y_i - \hat{Y}_i = R_i$$

como el  $i$ -ésimo RESIDUAL. Esta diferencia mide numéricamente la parte no explicada por el modelo

4) Los residuales son una parte muy importante en el análisis de regresión puesto que en general ayudan, entre otras cosas, a revelar si se están cumpliendo ó no las su-

posiciones básicas bajo las cuáles se trabaja este modelo.

En este sentido existen métodos gráficos basados en residuales, que son muy útiles para saber de forma rápida que tan bien se están cumpliendo las suposiciones. (Seber, op. cit., págs. 162-172 ). Es conveniente aclarar que estos métodos serán más efectivos en tanto más experimentado esté el estadístico y más a fondo conozca el problema en el que trabaja.

Para hacer un estudio gráfico sistematizado de los residuales se recomienda proceder de la forma que a continuación se expone.

Antes que nada debemos tener presente las suposiciones que se refieren a los términos de discrepancia:

$$u_i \sim N(0, \sigma^2)$$
$$E(u_i) = 0$$
$$E(u_i u_j) = \begin{cases} \sigma^2 & \text{si } i=j \\ 0 & \text{si } i \neq j \end{cases}$$

aunque se ha visto que la correlación tiene poco efecto en los trazados gráficos de los residuales.

Las gráficas que pueden obtenerse son (las más usuales):

- 4.1 Graficación completa de residuales.
- 4.2 Graficación de residuales sobre papel normal.
- 4.3 Graficación ordenada de residuales con respecto al tiempo.
- 4.4 Graficación de los residuales con respecto a los valores ajustados  $\hat{y}_i$ .
- 4.5 Graficación de los residuales con respecto a las variables independientes  $X_j$ , ( $j = 1, 2, \dots, p$ )
- 4.6 Graficación a pares de las variables predictoras  $X_j$ , ( $j = 1, 2, \dots, p$ )

4.7 Graficación de residuales parciales contra cada  $X_j$ ,  
(  $j = 1, 2, \dots, p$  ).

4.1) Graficación Completa de los residuales.-

Se hace sin tomar en cuenta ningún factor influyente dentro del modelo, de esta manera se puede obtener una gráfica de frecuencias y en caso de que las suposiciones sean aproximadamente correctas, los residuales deben semejar a  $n$  observaciones con media cero, provenientes de una población normal (Fig. 4).

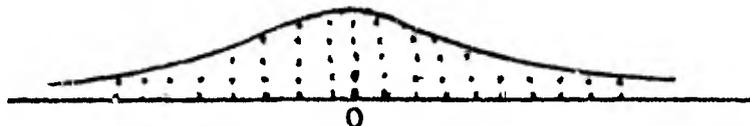


FIG. 4

4.2) Graficación de residuales sobre papel normal.-

Esta es una opción para tratar de detectar posibles violaciones en el modelo y las suposiciones, lo que se hace es graficar los residuales en papel normal, por tanto los puntos graficados deben localizarse cerca de la recta  $i = \text{Gau}_i^{-1} ( 1/n+1 )$  donde  $\text{Gau} (k) = P(x \geq k)$ .

Estos dos métodos ofrecen la posibilidad de detectar la existencia de:

- i) valores muy grandes en los residuales, (p.e. observaciones dudosas), aunque éstos dependerán del tamaño de  $n$ .
- ii) tendencia sistemática de los residuales, (los cuáles pudieran eliminarse posiblemente por una modificación en el modelo).
- iii) tendencia en los residuales que los haga 12 aparecer en forma asimétrica, etc.

4.3) Graficación ordenada de los residuales con respecto al tiempo.-

Esta graficación dependerá de si se tienen los residuales en orden cronológico, y en caso de respuesta afirmativa, estamos posibilitados para saber si afecta ó no el tiempo en el modelo. Podemos entonces enumerar los posibles casos a suceder:

a) Si el modelo resulta adecuado, se espera que la gráfica de los residuales con respecto al tiempo no rebasen cierta banda horizontal (Fig. 5).

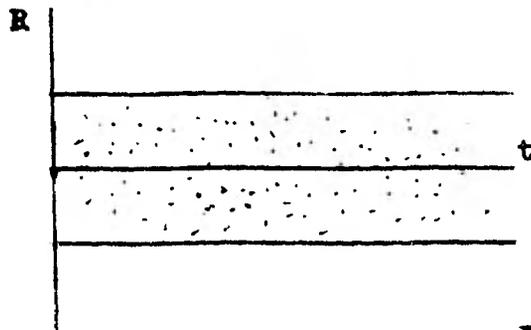


FIG. 5:

b) Puede suceder que haya un aumento (o disminución) sistemático ( en magnitud ) de los residuales con respecto al tiempo, ésta será representado como la fig. 6.(a,b).

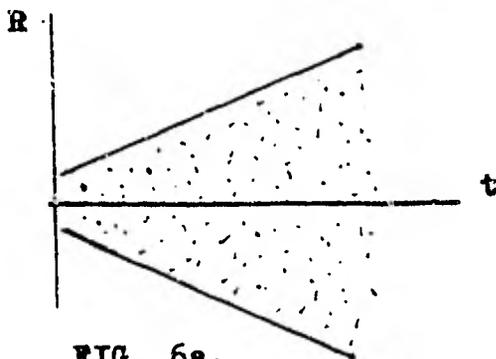


FIG. 6a.

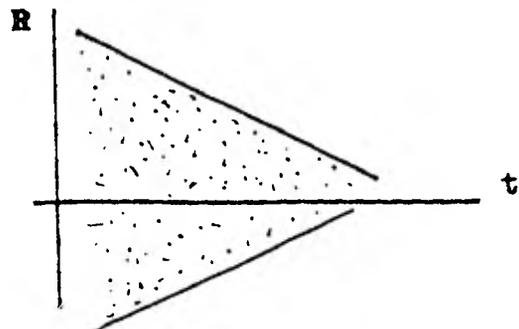


FIG. 6b.

Esto indicará que la varianza no es constante y depende del tiempo por lo que tendría que manejarse el modelo por mínimos cuadrados generalizados ó hacer una transformación en las observaciones.

c) En el caso de tener la siguiente gráfica (Fig. 7), indicaría que existe una tendencia con respecto al tiempo ó tal vez que los cálculos han sido incorrectos.

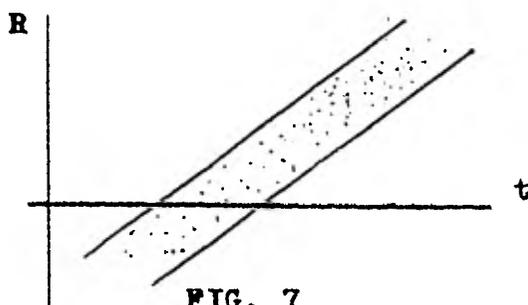


FIG. 7

d) En caso de necesitar la inclusión al modelo de un término cuadrático, ó un término no-lineal en el tiempo, la gráfica se comportaría de la siguiente forma (Fig. 8, a,b ).

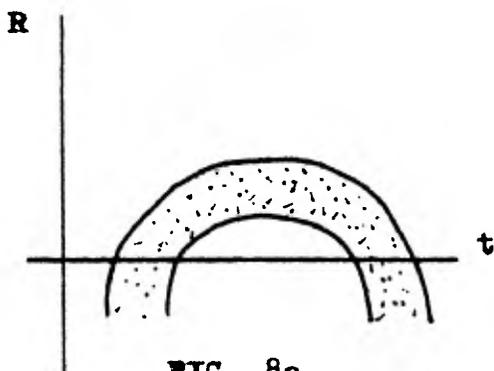


FIG. 8a.

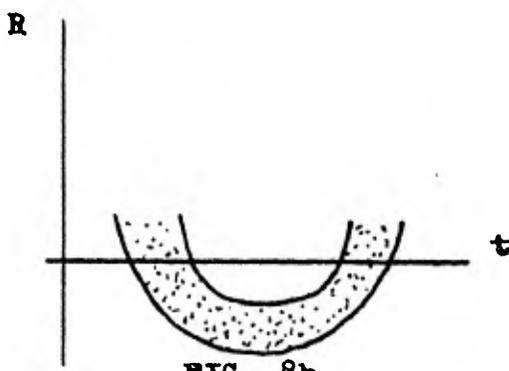


FIG. 8b.

e) Dentro de los posibles casos (usuales) puede también reflejarse la situación que el modelo necesite la inclusión de otra variable, cuyo efecto sea representado por ciertas

tendencias periódicas ya sean crecientes ó decrecientes. Este caso puede visualizarse en la fig. 9 (a,b).

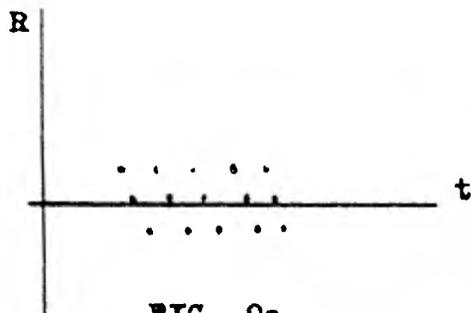


FIG. 9a.

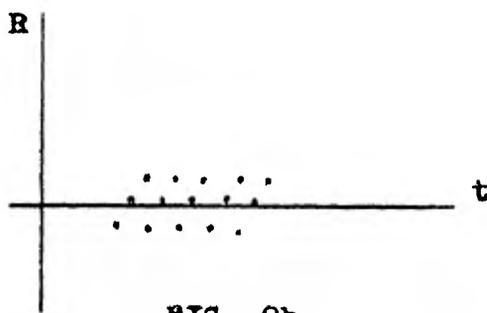


FIG. 9b.

#### 4.4) Graficación de los residuales con respecto a los valores ajustados $\hat{Y}_1$ .

La importancia de este trazado gráfico es que si se cumplen todas las suposiciones, se puede demostrar que  $\text{Cov}(\underline{R}, \hat{\underline{Y}}) = 0$ , y por lo tanto hay independencia por la normalidad. Este método es similar al anterior con la variante que en lugar de  $t$ , se sitúan los valores de  $\hat{Y}_1$ . El proceso de análisis es similar al anterior, es decir: si se tiene el caso de la figura 5, refleja que no se están violando las hipótesis; si se tiene el caso de la figura 6, nos está reflejando que la varianza está dependiendo de alguna variable predictora. El caso de la figura 7, indica errores causados por la omisión de un término constante  $b_0$  en el modelo propuesto. Cuando ocurre la situación de la figura 8, indica la necesidad de incluir términos cuadrados ó productos cruzados ó simplemente una transformación adecuada a las observaciones.

Con esta graficación se pueden detectar también observaciones aberrantes (outliers).

4.5) Graficación de residuales con respecto a cada variable  $X_j$ ;  $j = 1, 2, \dots, p$ .

Para este punto, se procede de la misma forma que los anteriores tanto en la graficación como en las interpretaciones con la ventaja de que aquí las interpretaciones son más específicas ya que se trata a cada una de las variables independientes.

4.6) Graficación de  $X_j$  contra  $X_k$ .

Si dos variables explicativas son altamente correlacionadas, lo más conveniente en éstos casos es no incluir a ninguna de las dos, ya que si se incluye una, la otra estará presente en la graficación de los residuales contra el tiempo.

4.7) Graficación de residuales parciales  $e'_i$  contra  $X_{ij}$ , ( $i = 1, 2, \dots, n$ ).

Con este trazado gráfico de los residuales parciales lo que se trata es de ver más de cerca la relación entre los residuales y la variable predictora  $X_j$ . Se define al  $i$ -ésimo residual como:

$$14.- \quad e'_i = Y_i - \hat{\beta}_0 - X_{i1} \hat{\beta}_1 - \dots - X_{i,j-1} \hat{\beta}_{j-1} - X_{i,j+1} \hat{\beta}_{j+1} - \dots - X_{i,p-1} \hat{\beta}_{p-1} = e_i + X_{ij} \hat{\beta}_j$$

Tal graficación lo que en esencia examina es la relación entre  $Y$  y  $X_j$  después de anular el efecto de las otras variables explicatorias en  $Y$ . Una propiedad interesante de esto es que si consideramos la regresión lineal a través del origen de  $e'_i$  en  $X_{ij}$ , entonces el estimador por mínimos cuadrados de la pendiente vuelve a ser  $\hat{\beta}_j$ . También, este método

do indica - en general - más precisamente como transformar  $Y$  para obtener linealidad. Finalmente es conveniente hacer notar que graficar  $y_i$  contra cada  $x_j$ , muy pocas veces es útil y puede ser completamente engañoso.

5) Es importante tratar siempre de que se cumplan lo mejor posible las suposiciones ya que cuando ésto no sucede, teóricamente, todo el análisis y en particular el conjunto de inferencias que de él se desprenden, resultan invalidadas.

Durante el desarrollo del trabajo se investigó -específicamente en Análisis de Regresión- los efectos que causa la violación de las suposiciones. En la búsqueda de éstas, nos encontramos con la siguiente afirmación, la cuál se pensó merecía la pena ser enfatizada:

Un estudio de esta clase [los efectos de la violación a las suposiciones] no puede ser exhaustivo, por una razón, pues suposiciones como éstas pueden ser violadas de muchas más formas que en las que pueden ser satisfechas. (Scheffé, op. cit., pág. 331).

Aunado a este problema se puntualiza que las violaciones (aparentemente) siempre se presentan en forma conjunta pero no de una manera pre-establecida. Por ejemplo, se puede presentar heterocedasticidad y anormalidad<sup>1</sup> ó no normalidad y correlación en los errores, etc.

Teniendo en cuenta todos los comentarios anteriores, mencionaremos en forma global los efectos que pudiera llevar consigo la violación de la suposición de homocedasticidad, que es la de interés en este trabajo. [Primeramente y sin perder generalidad, cabe decir que en los modelos de Aná

---

1.- Manejamos el término anormalidad para referirse a las curvas que "aparentemente" son normales pero no exactamente a la estándar, i.e.  $N(0, 1)$ ;  $\gamma = 0$ ;  $\gamma = \text{curtosis}$

lisis de Varianza balanceados, existe robustez<sup>1</sup> de las pruebas de igualdad de medias con respecto a desigualdad de varianzas.]

Ahora enfocándonos al modelo de Análisis de Regresión, el efecto que mayor énfasis hacen gran parte de los autores, es que el estimador de  $\beta$  obtenido por mínimos cuadrados ordinarios se vuelve ineficiente en comparación al estimador de  $\beta$  obtenido por mínimos cuadrados generalizados. El detalle es que para utilizar el método de mínimos cuadrados generalizados necesitaríamos conocer la forma de la varianza, es decir, saber la clase de heterogeneidad que se tiene en el problema. Este método es por sí mismo una forma de remediar la violación de homocedasticidad. Otra sería, tratar de obtener repeticiones de las observaciones ( en otras palabras, obtener varios muestreos ) de la misma población para poder comparar varianzas aunque esta posibilidad sería demasiado ambiciosa por el costo, pero no imposible.

Otra consecuencia que se tiene al existir heterocedasticidad es que en muchos casos, pruebas de hipótesis basadas en la distribución  $F$ , tienden a dar demasiados resultados significativos.

Dentro de la heterocedasticidad podemos tener dos grupos que son los siguientes:

Situaciones en las cuales existe una dependencia funcional de la varianza de una observación en la media del grupo al cual pertenece. Dependencia funcional de la varianza de una observación en su media ó valor esperado es una característica intrínseca de muchas distribuciones no-normales.

----- Puede ser una propiedad básica del fenómeno bajo

(1) Una prueba es robusta, si las inferencias [que de ella resultan] no son seriamente invalidadas por la violación de las suposiciones (en las que se basa) (Scheffé, op.cit. pág.360)

Investigación muy aparte de la forma de la distribución.

Inmediatamente, estos resultados, nos llevarían a pensar que antes de realizar el análisis, se valide la suposición de homocedasticidad (y en general todas las suposiciones que se hacen) mediante un procedimiento estadístico. En caso de que exista violación, buscar una transformación (en los datos) que sea apropiada, y así tratar de eliminarla ( vid. Guerrero y Hernández(1974); Lorda(1979)). Respecto a esto, Scheffé, señala que no es recomendable llevar a cabo una prueba preliminar de esta suposición - antes de aplicar el análisis de varianza, exponiendo las razones en el capítulo 10 de su libro(ibid., pág. 83.).

CAPITULO II  
PRUEBAS PARA LA DETECCION DE  
HETEROCEDASTICIDAD

Como se mencionó anteriormente, la suposición de igualdad de varianzas juega un papel muy importante en regresión. Este capítulo se dedicará a una revisión de la mayor parte de las pruebas existentes para detectar la violación de homocedasticidad.

Las primeras pruebas que se presentarán —debido a que son las más conocidas— serán la prueba propuesta por Bartlett(1937) y la prueba propuesta por Scheffé(1959). Después se citarán más pruebas en orden cronológico.

Recordemos que todas las pruebas fueron creadas con el propósito de verificar si la varianza en esa población es constante ( $0 \leq \sigma^2 < \infty$ ) Por lo tanto cada uno de los autores considerando diversas situaciones construyen su estadística de prueba y en caso de que éstas se distribuyan como alguna estadística estándar, señalan la región de rechazo de la hipótesis nula. En caso de que la estadística de prueba no se aproxime a ninguna de las usuales, proveen la región de rechazo de la hipótesis nula por medio de tablas especiales.

Como se mencionó en el capítulo anterior, la heterocedasticidad se puede presentar de dos formas:

- 1) Situaciones en las cuáles existe una dependencia funcional de la varianza en relación a alguna observación.
- 2) Situaciones en las cuáles puede ser una propiedad básica del fenómeno en investigación.

Dentro de estas dos formas existen una infinidad de

subformas de heterogeneidad, principalmente de la clase (1), señalada en el párrafo anterior.

Goldfeld & Quandt(1965), establecen la diferencia entre las pruebas que persiguen detectar heterocedasticidad la cual se deba a alguna variable incluida en el modelo o sencillamente detectar si existe heterocedasticidad, donde no interesa el tipo de relación funcional que existe en la varianza, llamándolas pruebas CONSTRUCTIVAS y pruebas NO-CONSTRUCTIVAS respectivamente.

Adelantándonos un poco, podemos citar como pruebas constructivas las pruebas propuestas por Goldfeld & Quandt(1965), Glejser(1969) y Zroeter(1978). Entre las no-constructivas tenemos, las pruebas propuestas por Bartlett(1937), Scheffé (1959), Hartley(1950), Cadwell(1953), Levene(1960), Miller (1968), Leyard(1973), etc.

A continuación, se tratarán todas estas pruebas con detalles, rayando en el nivel informativo, ya que como se dijo antes, no se pretende en ningún momento hacer una presentación muy formal de cada una de las pruebas aquí citadas.

Recordando, sabemos que en el modelo de regresión se hacen varias suposiciones básicas las cuales validan el estudio de los datos y los resultados que de éste se desprenden, particularmente la suposición de homocedasticidad. Por lo tanto, la hipótesis de interés primordial es

$$2.1) \quad H_0 : \sigma_i^2 = \sigma_j^2 \quad \forall i \neq j$$

Las hipótesis alternativas son numerosas. Básicamente, la diferencia la impondrá la prueba que se vaya a utilizar. En

otras palabras, las hipótesis alternativas serán de acuerdo a si se utiliza una prueba constructiva o no-constructiva.

PRUEBA DE BARTLETT(1937).-

Lo que se desea probar con esta estadística es si existe igualdad de varianza de muestra a muestra. Para este caso, Bartlett considera  $k$  muestras independientes, donde cada una contiene  $n_i$  ( $i = 1, 2, \dots, k$ ) observaciones las cuales son independientes e idénticamente distribuidas. Cuando la distribución es normal (y el coeficiente de curtosis es cero), demuestra que bajo  $H_0$ :

$$B = -\frac{1}{\alpha} \sum_{i=1}^k (n_i - 1) \log(s_i^2 / s^2)$$

se distribuye aproximadamente como una  $\chi^2_{(k-1)}$  Bartlett(1937)

donde:

$$c = 1 + \frac{1}{3(k+1)} \left[ \sum_{i=1}^k \left( \frac{k}{n_i - 1} \right) - \frac{1}{\sum_{i=1}^k (n_i - 1)} \right]$$

es un factor de corrección para muestras muy pequeñas ( $n_i=3$ )

y

$$s_i^2 = \frac{\sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2}{(n_i - 1)}, \quad s^2 = \frac{\sum_{i=1}^k (n_i - 1) s_i^2}{\sum_{i=1}^k (n_i - 1)}$$

Intuitivamente, lo que Bartlett hizo es comparar por medio de razones la variación de cada grupo(o muestra) con respec

to a la variación total ( la variación en el total de todas las muestras).

Cuando los tamaños de cada muestra son iguales, esta prueba coincide con la prueba desarrollada por Neyman & Pearson(1931) la cuál está basada en el principio de cociente de verosimilitud y es dada por:

$$M = K(n-1) \left\{ \log_e \left( \sum_{i=1}^k \frac{S_i^2}{k} \right) - \sum_{i=1}^k \left( \log_e \frac{S_i^2}{k} \right) \right\}$$

(de hecho la prueba de Bartlett es una modificación de la prueba M).

Cuando la distribución es arbitraria (diferente de la distribución normal), Box(1953) demuestra que bajo  $H_0$  converge a  $\left[ 1 + \left( \frac{\gamma}{2} \right) \right] \chi^2_{(k-1)}$  donde  $\gamma$  es la curtosis<sup>(1)</sup> de

la distribución dada por  $\gamma = E(x - \mu)^4 / \sigma^4 - 3$

Observa además que:

- 1) Si  $\gamma > 0$ , entonces suponiendo que B es aproximadamente  $\chi^2$ , la estadística B dá muchos resultados significativos.
- ii) Si  $\gamma < 0$ , entonces suponiendo que B es aproximadamente  $\chi^2$ , se obtienen muy pocos resultados significativos.

Por tanto, resumiendo el proceso de prueba obtenemos lo siguiente:

Supóngase que tenemos n residuales, al ajustar un modelo de regresión, independientes, donde además se supone que provienen de una población normal. Entonces, podemos dividir el conjunto de residuales de tamaño n en k sub -

---

(1) Recuérdese que la curtosis es una medida que nos indica que tan "chata" o tan pronunciada está la curva.

muestras con la restricción de que el número de elementos en cada submuestra deben ser mínimo  $n_i \geq 2$  para  $i=1,2,\dots,k$ ; las  $n_i$ 's no necesariamente tienen que ser iguales.

Constrúyase B donde

$$S_i^2 = \sum_{j=1}^{n_i} (e_{ij} - \bar{e}_i)^2 / (n_i - 1) \quad i=1,2,\dots,k$$

$$S^2 = \sum_{i=1}^k (n_i - 1) S_i^2 / \sum_{i=1}^k (n_i - 1)$$

Por tanto, bajo  $H_0$ , B se distribuye como una estadística  $\chi^2$  con  $(k-1)$  grados de libertad si B rebasa el valor de  $\chi^2_{(k-1, 1-\alpha)}$ , rechazamos la hipótesis nula al  $(1-\alpha)100\%$  de significancia.

Nota. Es importante aclarar que si  $k=2$  ( # de muestras ) y/o  $n_i = 2$ , la prueba no funciona. (Ver Cap. III).

#### PRUEBA DE SCHEFFÉ(1959).-

Esta prueba es una variante de la prueba S para igualdad de medias. Al igual que la prueba de Bartlett, lo que se desea es verificar la igualdad de varianzas en  $k$  muestras (o grupos).

Esta prueba está basada en el análisis de varianza de los logaritmos de las varianzas muestrales, la cual es totalmente insensitiva a la forma de las distribuciones de las medias estimadas.

La idea para probar la suposición de homocedasticidad es la siguiente:

Supóngase que se tiene un conjunto de residuales  $\{e_i\}$  de tamaño  $n$  es decir  $\{e_i\}_{i \in J_n}$ ;  $n \in \mathbb{N}$ , si se considera

una división del conjunto de residuales de tal forma que cada división tenga al menos dos elementos; entonces podemos construir grupos con estas divisiones con la característica que dentro de cada grupo estamos dispuestos a suponer igualdad de varianza. Por tanto, si hay  $J_i$  divisiones en el  $i$ -ésimo grupo, podemos denotar por  $S_{ij}^2$  (1) a la varianza muestral de la  $j$ -ésima división en el  $i$ -ésimo grupo y por consiguiente bajo las suposiciones básicas

$$E(S_{ij}^2) = \sigma_i^2 \quad j = 1, 2, \dots, J_i$$

Sea  $y_{ij} = \log_e S_{ij}^2$ ,

es demostrable (2) que

$$E(y_{ij}) \sim \log_e \sigma_i^2 = \sigma_i^{2*}$$

y

$$\text{Var}(y_{ij}) \sim \frac{2}{n_{ij}-1} + \frac{\gamma_{2,ij}}{n_{ij}}$$

donde  $n_{ij}$  es el tamaño de la muestra de la que fué calculada  $S_{ij}^2$  y  $\gamma_{2,ij}$  es la medida de curtosis de la misma muestra (o grupo) la cual es dada por

$$\gamma_2 = \sigma^{-4} \mu_4 - 3$$

con  $\mu_4$  igual al cuarto momento central de la población.

Por otro lado, bajo las suposiciones básicas, es lógico pensar que todos los grupos tendrán el mismo grado de curtosis, por esta razón se adiciona esta suposición y por consiguiente es equivalente a probar la hipótesis

$$H^* : \sigma_1^{2*} = \sigma_2^{2*} = \dots = \sigma_I^{2*}$$

---

(1) donde  $S_{ij}^2$  es dado por  $\sum_{k=1}^{n_{ij}} (e_{ijk} - \bar{e}_{ij})^2 / n_{ij} - 1$

(2) Schaffé(1959) pág. 84.

Si a ésto se le aplica el análisis de varianza, obtenemos una estadística de prueba de la forma siguiente -siempre que los grupos sean del mismo tamaño-

$$F_c = \frac{Y_e}{I - 1} \frac{\sum_i J_i (y_{i.} - \bar{y})^2}{\sum_i \sum_j (y_{ij} - y_{i.})^2}$$

con

$$\bar{y} = \sum_i J_i y_{i.} / \sum_i J_i \quad Y_e = \sum_i (J_i - 1).$$

Esta estadística bajo la hipótesis nula se distribuye como una F con I-1 y  $Y_e$  grados de libertad.

Para el caso en que los tamaños de grupos sean diferentes (al menos uno) la estadística de prueba varía, obteniéndose una estadística de la forma

$$F'_c = \frac{Y_e}{I - 1} \frac{\sum_i v_i (\hat{n}_i - \hat{n})^2}{\sum_i \sum_j v_{ij} (y_{ij} - \hat{n}_i)^2}$$

donde las  $v_i$ ,  $v_{ij}$  son ponderaciones para compensar la desigualdad entre grupos y están dadas por:

$$v_{ij} = n_{ij} - 1, \quad v_i = \sum_j v_{ij}, \quad v = \sum_i v_i \quad y$$

$$Y_e = \sum_i (J_i - 1), \quad \hat{n}_i = \sum_j v_{ij} y_{ij} / v_i \quad y$$

$$\hat{n} = \sum_i v_i \hat{n}_i / v.$$

Entonces bajo la hipótesis nula, la  $F'_c$  se distribuye como una F con I - 1,  $Y_e$  grados de libertad.

Podemos ver claramente de  $F_c$  y  $F'_c$  que lo que se tra-

ta de medir es la variación de grupo a grupo con respecto a la variación dentro de los grupos.

La inconveniencia de esta prueba es que si todas las  $J_i$  son igual a 1 no sirve y no es sensitiva si  $\sum_i (J_i - 1)$  es muy pequeño. En caso de tener muestras muy pequeñas, se debe subdividir de tal forma que  $\sum_i (J_i - 1)$  no sea muy pequeña.

#### PRUEBA DE HARTLEY(1950B).-

Hartley, basado en la prueba de Bartlett y con las demostraciones de Bartlett y Kendall(1946) que el  $\log_e S^2$  es aproximadamente distribuido como una variable normal, propuso la siguiente prueba: Toma  $K$  muestras de igual tamaño provenientes de poblaciones normales y calcula los cuadrados medios de cada una, después divide el mayor cuadrado medio entre el menor de ellos.

Pero como mencionamos antes, Bartlett y Kendall mostraron que  $\log_e S^2$  se distribuye aproximadamente normal con varianza  $\text{Var}(\log_e S^2) \cong 2/(n-1)$ , (con  $n$  igual al número de observaciones en cada muestra), por lo que se pueden calcular puntos porcentuales para el cociente por medio de

$$F_{\max}(\alpha) = \exp \left\{ w_k(\alpha) \sqrt{[2/(n-1)]} \right\}$$

donde  $w_k(\alpha)$  es el punto  $100\alpha\%$  del "rango"  $w$ , en muestras normales de tamaño  $k$ . Con el fin de checar la precisión de la aproximación, Hartley hizo dos comparaciones con la distribución nula exacta. Los casos que consideró fueron:

- 1) donde tiene dos muestras y diferentes números de observaciones.

ii) donde tiene dos observaciones y diferentes números de muestras.

Estas consideraciones fueron a nivel  $\alpha = 0.05$  para el primer caso y a los niveles  $\alpha = 0.01$  y  $\alpha = 0.05$  para el segundo caso.

David, H.A.(1952) observó que la aproximación en general tiende a subestimar los puntos porcentuales y presenta unas tablas corregidas a los niveles 5 y 1%. Estas tablas las obtuvo por cuadratura exacta.

Resumiendo, la prueba es:

Obténganse  $k$  muestras (o grupos) de residuales del mismo tamaño y calcúlense los cuadrados medios de cada una, elíjase el mayor y el menor y divídase, es decir, si denotamos por  $S_{\min}^2$  a la menor varianza muestral y por  $S_{\max}^2$  a la mayor varianza muestral, entonces defínase

$$F_{\max} = \frac{S_{\max}^2}{S_{\min}^2}$$

obténgase para los valores de  $k$  y  $v$ , los puntos porcentuales de  $F_{\max}$  y compárese. Si  $F_{\max}$  es mayor que el valor de tablas, se rechaza la hipótesis nula  $H_0$ . TABLA II

#### PRUEBA DE CADWELL(1953).

La prueba propuesta por Cadwell(1953) se basa en el cociente entre el rango máximo y el rango mínimo de  $k$  muestras las cuales se supone que se extrajeron independientemente de poblaciones normales. Esta prueba, similar a la propuesta por Hartley(1950) es tomada del criterio en el cual se basa la prueba de Bartlett; con la diferencia que en las

primeras dos se pierde una considerable potencia. Cadwell aproxima la distribución del rango a una potencia apropiada de la estadística  $\chi^2$ . Las tablas para tal prueba, obtenidas por Cadwell, resultan de una transformación de las tablas de David(1952). Harter(1963), considerando el caso para  $k = 2$ , dió puntos porcentuales exactos para esta prueba. A partir de ésto, Leslie & Brown(1966) vieron la importancia de extender las tablas obtenidas por Harter a valores grandes de  $k$  y al mismo tiempo dieron una versión exacta de las tablas obtenidas por Cadwell en su artículo original. La prueba es de la siguiente forma:

Supóngase que se tienen  $k$  grupos de residuales, donde cada grupo contiene el mismo número de elementos. Supóngase además que son independientes y que provienen de poblaciones normales.

Una hipótesis que nos interesaría probar es:

$$H_0 : \sigma_1 = \sigma_2 = \dots = \sigma_k = \sigma$$

lo cual es equivalente a probar si la dispersión en las muestras son iguales. Entonces, podemos calcular el rango de cada una de las muestras o grupos, de donde escogemos el máximo rango y el mínimo rango.

Defínase

$$w(k,n) = \frac{w_{\text{máx}}}{w_{\text{mín}}}$$

donde  $w$  denota el rango de una muestra, el resultado obtenido lo comparamos con los valores de tablas obtenidos por Leslie & Brown (Tabla III) al buscar el número de muestras  $k$  y el número de elementos por muestra  $n$ . Si el valor de  $w(k,n)$

rebasa el valor de tablas, rechazamos la hipótesis nula  $H_0$ .

### PRUEBA DE LEVENE(1960).-

Levene propuso una prueba para examinar la igualdad de varianza en varios grupos, basada en los valores absolutos de los residuales. Además propuso varios estimadores centrales como variantes a dicha prueba.

Draper & Hunter(1969) enfatizan al respecto:

Levens(1960) también mencionó "La posibilidad de usar un análisis de varianza similar en los valores absolutos de los residuales de otras regresiones, a fin de estudiar la varianza de los residuales. No hay razón de porque esto no pudiera trabajarse también en general como en el caso especial aquí considerado". Abajo extendemos [Draper & Hunter] la prueba de Levene un paso más y la aplicamos a residuales los cuales provienen de grupos que pueden ser desiguales. Hasta donde sabemos no ha sido hecho ningún estudio en situaciones sobre grupos desiguales, pero hay razones para creer que, en esta situación también, la prueba es robusta.

#### **METODO:**

Supóngase que tenemos  $p$  grupos de residuales

$e_{ij}$  como sigue:

Grupo 1:  $e_{11}, e_{12}, \dots, e_{1n_1}$  promedio  $\bar{e}_1, V(e_{1j}) = \sigma_1^2$   
Grupo 2:  $e_{21}, e_{22}, \dots, e_{2n_2}$ , promedio  $\bar{e}_2, V(e_{2j}) = \sigma_2^2$   
:  
:  
:  
Grupo  $p$ :  $e_{p1}, e_{p2}, \dots, e_{pn_p}$ , promedio  $\bar{e}_p, V(e_{pj}) = \sigma_p^2$

Construir de estas observaciones :

$$z_{ij} = |e_{ij} - \bar{e}_i| \quad j = 1, 2, \dots, n_i; \quad i = 1, 2, \dots, p$$

para obtener:

Grupo 1:  $z_{11}, z_{12}, \dots, z_{1n_1}$  , total  $z_1$  .  
Grupo 2:  $z_{21}, z_{22}, \dots, z_{2n_2}$  , total  $z_2$  .  
:  
:

⋮

Grupo p:  $z_{p1}, z_{p2}, \dots, z_{pn_p}$ , total  $z_p$ .

Hacer un análisis de varianza estándar en las  $z_{ij}$  como si sigue:

Fuente	Suma de cuadrados	Grados de libertad	Cuadrados medios
Entre grupos	$\sum_{i=1}^p \frac{z_i^2}{n_i} - \frac{G^2}{\sum_i n_i}$	$p - 1$	$S_1^2 F_1 = \frac{S_1^2}{S_2}$
Dentro de grupos	por diferencia	$\sum_{i=1}^p (n_i - 1)$	$S^2$
Media	$G^2 / \sum_{i=1}^p n_i$	1	
Total	$\sum_{i=1}^p \sum_{j=1}^{n_i} z_{ij}^2$	$\sum_{i=1}^p n_i$	

donde  $G = z_1 + z_2 + \dots + z_p = \text{total grande.}$

Si  $F_1 > F \left\{ (p - 1), \sum (n_i - 1), 1 - \alpha \right\}$  es significativa, y decimos que hay evidencias de que existen diferencias entre  $\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2$ . Si  $F_1$  no es significativa, no rechazamos la hipótesis  $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_p^2$ .

Razonamiento heurístico:

El análisis de varianza es sabido que es robusto para probar diferencias entre medias de grupos. Las  $z_{ij}$ , con  $j = 1, 2, \dots, n_i$ , son todos estimadores de  $\sigma_i^2$ , para cada  $i = 1, 2, \dots, p$ . Probamos, robustamente, si existen diferencias entre las medias de los  $p$  conjuntos de estimadores. Si no entonces concluimos que no hay evidencia contra la hipótesis que  $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_p^2$ . (Draper & Hunter(1969), op.cit., pág. 39).

Las variables alternativas para la prueba que Levene propuso fueron utilizar  $\sqrt{z_{ij}}$  o  $\log z_{ij}$ . Brown y Forsythe (1974), utilizaron estimadores diferentes de  $\bar{e}_i$  al calcular las  $z_{ij}$ , estos fueron los siguientes, sustitución de la media por la mediana. Ellos denotan el cociente como  $W_{50}$ . Sustitución de la media por la "Media reducida"; ésta, es obtenida de eliminar un porcentaje<sup>1</sup> igual tanto de las observaciones más pequeñas como de las observaciones más grandes denotando el cociente como  $W_{10}$ . Lo que Brown & Forsythe argumentaron es que estos estimadores son robustos para el caso cuando las poblaciones básicas están sesgadas.

#### PRUEBA PARAMÉTRICA DE GOLDFELD & QUANDT(1965).-

Esta es una prueba constructiva propuesta por Goldfeld & Quandt. El fin de esta estadística es verificar si existe alguna relación funcional de la varianza con alguna variable independiente incluida en el modelo. El inconveniente en este tipo de pruebas es que si la relación funcional probada no es la verdadera, se tienen que volver a hacer todos los cálculos, sin embargo, recordemos que los métodos gráficos pueden dar una idea al respecto. Los autores consideran el modelo de regresión :

$$(1) \quad \underline{Y} = X\underline{\beta} + \underline{U} \quad ; \quad u_i \sim N(0, \phi_i^2 \sigma^2) \quad , \quad E(u_i u_j) = 0$$

donde se supone que  $\phi_i$  es conocida pero  $\sigma^2$  no lo es. Además las  $x_j$  se suponen independientes.

Si sucede esta relación, el modelo puede ser trabajado como:

---

1.- El porcentaje que ellos eliminaron fué el 10% pero esta eliminación puede hacerse con diferentes porcentajes. En otras palabras, es arbitraria, según la conveniencia del estadístico.

$$(2) \quad \frac{1}{\beta_1} \underline{Y} = \frac{1}{\beta_1} \underline{X} \underline{\beta} + \underline{V} \quad \text{con } E(\underline{v}_1) = 0, \\ E(\underline{v}_1 \underline{v}_j) = \begin{cases} 0 & \text{si } i \neq j \\ \sigma^2 & \text{si } i=j \end{cases}$$

de esta manera, se pueden obtener los estimadores de  $\underline{\beta}$ . En particular si  $\beta_1 = 1$ . Los estimadores de (1) y (2) coinciden. Pero cuando  $\beta_1 = \chi_m$  (una variable incluida en el modelo) el modelo es similar a (2), es decir,

$$(3) \quad \underline{Y} / \chi_m = \underline{X} / \chi_m \underline{\beta} + \underline{V}$$

Por lo tanto, si existe una relación de este tipo, es válido pensar que si colocamos las observaciones de acuerdo a un orden creciente (o decreciente) en  $\chi_m$  podremos de alguna forma saber, mediante una prueba de significancia, si existe realmente efecto de la variable que se sospecha, esto es, de  $\chi_m$ .

Basados en esta idea, Goldfeld & Quandt propusieron la siguiente prueba.

Se ordenan las  $n$  observaciones de tal forma que los valores de  $\chi_m$  queden desde el valor más chico hasta el más grande (no perdiendo de vista el orden original). Habiendo hecho tal ordenación, se eliminan  $k^{(L)}$  observaciones centrales de tal forma que  $\frac{n-k}{2} \gg p$  (donde  $p$  es el número de parámetros en el modelo).

Se obtiene la recta de regresión para cada uno de los subconjuntos constituidos por  $\frac{n-k}{2}$  observaciones y se calcula la suma de cuadrados debido a los residuales de cada una de las regresiones. Denótese por  $S_1$  y  $S_2$  la suma de cuadrados de las  $\frac{n-k}{2}$  primeras observaciones relativas a  $\chi_m$

1-En las simulaciones hechas por los autores, observaron que el valor de  $k$  para el cual se obtiene mayor potencia de la prueba es  $k=8$  si  $n=30$  y  $k=16$  si  $n=60$ .

y  $\frac{n-k}{2}$  últimas observaciones relativas a  $\chi_m$  respectivamente.

Constrúyase el cociente  $R = S_2 / S_1$ .

Entonces, bajo  $H_0$ ,  $R$  se distribuye como una estadística  $F$  con  $(n-k/2, n-k/2)$  grados de libertad<sup>1</sup>. Por lo tanto, si  $R > F_{(n-k/2, n-k/2)}$  se rechaza  $H_0$  al nivel de significancia  $\alpha$ .

Intuitivamente, esta prueba lo que hace es medir el efecto de la variable  $\chi_m$  por medio del cociente entre las sumas de cuadrados de los residuales obtenidos de cada una de las regresiones. Puesto que si es cierta la hipótesis alternativa  $S_2$  será mucho mayor que  $S_1$  y por consiguiente  $R$  será grande.

Los autores hicieron las siguientes observaciones:

a) Puesto que la suma de cuadrados de los residuales puede ser expresada como una forma cuadrática de los errores verdaderos, el cociente  $R$  es homogéneo de grado cero en los términos de errores verdaderos; por tanto el cociente es independiente de  $\sigma_u$  bajo la hipótesis  $H_0$  y  $\sigma_v$  si la hipótesis alternativa es verdadera.

b) El cociente  $R$  es independiente de los coeficientes de regresión.

c) Normalmente no podemos usar la prueba estándar para la homogeneidad de varias varianzas ya que generalmente no tenemos observaciones repetidas para las  $\chi$ 's dadas.

d) La potencia de esta prueba claramente dependerá del valor de  $k$ , el número de observaciones omitidas; para valores grandes de  $k$  la potencia será pequeña pero no es obvio que la potencia aumente monótonamente cuando  $k$  tiende a cero.

e) La potencia de la prueba claramente dependerá de la naturaleza de la muestra de valores para la variable la cual afecta. Así, si la varianza de  $\chi_m$  es relativamente pequeña a la media de  $\chi_m$ , la potencia puede esperarse pequeña y recíprocamente. (Goldfeld & Quandt, op.cit., pág. 541.).

Harvey & Phillips por su parte puntualizan lo siguiente:

Nótese que omitiendo algunas observaciones centrales lleva a una reducción en los grados de libertad. Sin embargo, es la demostración de esto puede verse en GRAYBILL, F.A. (1961) pág. 31.

te efecto es compensado ya que las varianzas de los términos de error asociados a las observaciones centrales son unitarios, y por tanto sacar estas  $\chi$ 's significa que R tenderá a ser más grande bajo la hipótesis alternativa. (Harvey & Phillips (1974)).

### PRUEBA JACKKNIFE (MILLER(1968)).-

Para el lector no familiarizado con el jackknife se incluye una breve descripción: Sea  $\theta$  un parámetro desconocido, y sea  $(X_1, \dots, X_N)$  una muestra de  $N$  observaciones independientes e idénticamente distribuidas con función de distribución acumulativa  $F_X$ , la cual depende de  $\theta$ . Supóngase que es obtenible un método (sesgado o insesgado) para estimar  $\theta$ . Además supóngase que los datos son divididos en  $n$  grupos de tamaño  $k$  ( $N=nk$ ); i.e.,  $(X_1, \dots, X_k)$ ,  $(X_{k+1}, \dots, X_{2k})$ , ...,  $(X_{(n-1)k+1}, \dots, X_{nk})$ . Esta división puede ser determinada por la estructura del experimento o arbitrariamente impuesto por el estadístico. Sea  $\hat{\theta}$  el estimador de  $\theta$  basado en las  $N=nk$  observaciones, y sea  $\hat{\theta}_{-i}$ ,  $i=1, 2, \dots, n$ , el estimador de  $\theta$  obtenido de las  $(n-1)k$  observaciones resultantes de eliminar el  $i$ -ésimo grupo. Estos forman los nuevos estimadores (Tukey(1958) les da el nombre de "PSEUDOVALORES"):

$$(1) \quad \tilde{\theta}_i = n\hat{\theta} - (n-1)\hat{\theta}_{-i}, \quad i = 1, 2, \dots, n.$$

El estimador jackknife de  $\theta$  es el promedio de las  $\tilde{\theta}_i$ , i.e.,

$$(2) \quad \tilde{\theta}_\cdot = n^{-1} \sum_{i=1}^{n-1} \tilde{\theta}_i = n\hat{\theta} - (n-1)\hat{\theta}_{-\cdot},$$

donde

$$\hat{\theta}_{-\cdot} = \left( \sum_{i=1}^{n-1} \hat{\theta}_{-i} \right) / n.$$

El jackknife elimina exactamente un término de sesgo  $n^{-1}$ . Digamos, si

$$(3) \quad E(\hat{\theta}) = \theta + cN^{-1} + O(N^{-2})$$

entonces

$$(4) \quad E(\tilde{\theta}_\cdot) = \theta + O(n^{-2})$$

Quenouille [1949] concibió el jackknife para llevar a cabo esta reducción en sesgo. Tukey fué un paso más y propuso que en muchas instancias las  $\hat{\theta}_{-i}$  son aproximadamente independientes, e idénticamente distribuidas. Si esta proposición es

correcta, entonces

$$(5) \quad (n(n-1))^{-1} \sum_{i=1}^{n-1} (\tilde{\theta}_i - \tilde{\theta}_\cdot)^2$$

sería un estimador (aproximado) de la  $\text{Var}(\tilde{\theta}_\cdot)$ , y

$$(6) \quad (\tilde{\theta}_\cdot - \theta) \left[ (n(n-1))^{-1} \sum_{i=1}^{n-1} (\tilde{\theta}_i - \tilde{\theta}_\cdot)^2 \right]^{-1/2}$$

sería aproximadamente distribuido como una variable  $t$  con  $n-1$  grados de libertad. El cociente (6) podría por consiguiente ser usado para construir un intervalo de confianza para  $\theta$  aproximado o desarrollar una prueba de significancia en  $\theta$  aproximada.

Basándose en la teoría anteriormente expuesta, Miller (1968) propuso la siguiente prueba. Supóngase que se tienen  $N$  residuales  $e_1, e_2, \dots, e_N$  los cuales son independientes e idénticamente distribuidos de acuerdo a  $F((e-\mu_e)\sigma_e^{-1})$ .

Para verificar si existe varianza constante, se puede construir un intervalo de confianza para el parámetro, es decir,  $\sigma^2 = E(\xi - E(\xi))^2$ , o equivalentemente probar la hipótesis  $H_0: \sigma^2 = \sigma_0^2$ , donde  $\sigma_0^2$  es una constante específica.

Cuando  $\sigma^2$  es desconocida, sabemos muy bien que un buen estimador clásico de  $\sigma^2$  está dado por  $S^2 = \sum_{i=1}^n (e_i - \bar{e})^2 / (n-1)$ .

Por lo tanto, si denotamos por  $\theta$  a  $\sigma^2$  y  $\hat{\theta}$  a  $S^2$  entonces el jackknife para grupos de tamaño 1 (i.e.  $k=1$ ) da:

$$\tilde{\theta}_1 = S_e^2 + n(n-2)^{-1} \left[ (e_1 - \bar{e})^2 - n^{-1} \sum_{j=1}^n (e_j - \bar{e})^2 \right],$$

donde

$$(7) \quad \tilde{\theta}_\cdot = S_e^2$$

y por consiguiente

$$\sum_{i=1}^{n-1} (\tilde{\theta}_i - \tilde{\theta}_\cdot)^2 = n^2 (n-2)^{-2} \sum_{i=1}^n ((e_i - \bar{e})^2 - n^{-1} \sum_{j=1}^n (e_j - \bar{e})^2)^2$$

Ahora, aplicando (6), bajo la hipótesis nula, tenemos una variable  $t$  con  $n-1$  grados de libertad.

Por otro lado, si  $k > 1$ , las expresiones no son nada fáciles para calcular, de hecho es indispensable un programa de computadora para agilizar los cálculos ya que manualmente es un enredo. Para el lector interesado en este caso puedo referirse a Miller, op. cit., pág. 569.

Miller, demuestra en su artículo que la transformación logaritmo de  $S^2$  frecuentemente produce una distribución más cercana a la normal y estabilización de la varianza. Por esta razón es mucho más conveniente utilizar el  $\log S^2$  que  $S^2$  y particularmente en la teoría jackknife. Por tanto si denotamos por

$$\begin{aligned}
 \theta &= \log \sigma^2 \\
 \hat{\theta} &= \log S^2 \\
 (8) \quad \tilde{\theta}_i &= n \log S^2 - (n-1) \log S_{-i}^2 \\
 \tilde{\theta}_\cdot &= n \log S^2 - (n-1) n^{-1} \sum_{i=1}^n \log S_i^2
 \end{aligned}$$

y por ende

$$(n-1)^{-1} \sum_{i=1}^n (\tilde{\theta}_i - \tilde{\theta}_\cdot)^2 = (n-1) \sum_{i=1}^n (\log S_{-i}^2 - n^{-1} \sum_{j=1}^n \log S_{-j}^2)^2$$

donde sabemos que  $S_{-i}^2$  es la varianza muestral basada en  $(n-1)k$  observaciones las cuales son obtenidas de eliminar el  $i$ -ésimo grupo con  $k$  observaciones.

Entonces, si regresamos al problema de verificar si existe varianza constante cuando tenemos una sola muestra, se podría utilizar la expresión (6) con la variante de que en lugar de usar las  $\tilde{\theta}_i$  y  $\tilde{\theta}_\cdot$  originales, se debe sustituir por la que da la expresión (8) y de esta forma obtenemos una prueba para  $H_0 : \log \sigma^2 = \log \sigma_0^2$  o sencillamente construir un intervalo de confianza para  $\log \sigma^2$ .

Supóngase ahora que tenemos 2 muestras, es decir, sean  $e_1, e_2, \dots, e_N$  independientes e idénticamente distribuidas de acuerdo a  $F((e - \mu_e) \sigma_e^{-1})$  y  $r_1, r_2, \dots, r_M$  independientes e idénticamente distribuidas de acuerdo a  $F((r - \mu_r) \sigma_r^{-1})$ . Supóngase además que tanto las  $e_i$ 's como las  $r_j$ 's están regidas por la misma distribución excepto posiblemente por una localización y escala diferente de los parámetros. De nuevo, tanto la localización y escala de los parámetros como la función de distribución son desconocidas, excepto que se supone que la distribución posee cuartos momentos finitos.

Entonces lo que nos interesaría probar sería la hipótesis  $H_0: \sigma_e^2 = \sigma_r^2$  o algo equivalente a esto sería construir un intervalo de confianza al cociente o diferencia de las varianzas. Dados todos los resultados anteriores y procediendo de forma similar, se obtiene la estadística:

$$\left[ (\bar{e}_{\tilde{\theta}} - \theta_e) - (\bar{r}_{\tilde{\theta}} - \theta_r) \right] \left[ (n(n-1))^{-1} \sum_{i=1}^n (\tilde{\theta}_{e_i} - \bar{e}_{\tilde{\theta}})^2 + (m(m-1))^{-1} \sum_{i=1}^m (\tilde{\theta}_{r_i} - \bar{r}_{\tilde{\theta}})^2 \right]^{-1/2}$$

la cual es usada como pivote para construir un intervalo de confianza o llevar a cabo la prueba. De nuevo,  $\theta_e = \log \sigma_e^2$ ,  $\theta_r = \log \sigma_r^2$  y  $\tilde{\theta}_{e_i}, \tilde{\theta}_{r_i}$  están basados en  $\hat{\theta}_e = \log S_e^2$  y  $\hat{\theta}_r = \log S_r^2$  respectivamente y esta estadística se distribuye como una variable  $t$ .

Lo que restaría ver es cuantos grados de libertad corresponden a la  $t$ , naturalmente, esto dependerá del tamaño de cada muestra. Se tienen dos casos. Si  $N=M$  y las divisiones son iguales en cada una de las muestras, entonces los grados de libertad de la estadística  $t$  serán  $n+m-2$ ; pero si  $N \neq M$  los gra

dos de libertad apropiados para la estadística  $t$  es un tanto ambiguo como en la prueba de diferencia de medias con varianzas desiguales.

Finalmente es importante hacer notar que la mejor elección para el tamaño de  $k$  es 1, aunque cuando los tamaños de muestra sean muy grandes con el fin de economizar tiempo en cálculos no habrá de otra que elegir el valor de  $k$  mayor que uno.

PRUEBA DE LAYARD(LAYARD(1973))  
(EXTENSIÓN DE LA PRUEBA DE JACKKNIFE).-

Puesto que Miller trabaja la prueba para el caso de dos muestras, Layard<sup>1</sup> propone una extensión directa a la prueba, considerando  $k$  muestras de la forma siguiente:

$$\text{Sea } \tilde{\theta}_{ij} = n_i \log S_i^2 - (n_i - 1) \log S_{i(j)}^2$$

con

$i = 1, 2, \dots, k$ , donde  $i$  denota las muestras

$j = 1, 2, \dots, n_i$ , donde  $j$  denota el tamaño de la muestra.

$$S_{i(j)}^2 = \left( \frac{1}{n_i - 2} \right) \sum_{h \neq j} (e_{ih} - \bar{e}_{i(j)})^2$$

$$\bar{e}_{i(j)} = \left( \frac{1}{n_i - 1} \right) \sum_{h \neq j} e_{ih}$$

después calcula la estadística  $F$  de un análisis de varianza con un criterio de clasificación en las  $\tilde{\theta}_{ij}$ , digamos

$$F_{\text{calc}} = \frac{\sum_{i=1}^k n_i (\bar{\theta}_{i.} - \bar{\theta}_{..})^2 / (k-1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (\theta_{ij} - \bar{\theta}_{i.})^2 / \sum_{i=1}^k (n_i - 1)}$$

la cual bajo la hipótesis nula la distribución de  $F_c$  es aproxi

ximadamente una  $F$  con  $(k-1, \sum (n_i - 1))$  grados de libertad.

PRUEBA DE GLEJSER (1969).-

El método propuesto por Glejser es el siguiente. Supóngase que tenemos un conjunto de residuales  $\{u_i\}_{i \in J_n}$ ;  $n \in \mathbb{N}$  obtenidos por mínimos cuadrados ordinarios del modelo de regresión lineal

(1)  $\underline{Y} = \underline{X} \underline{\beta} + \underline{u}$  ; donde  $u_i$  se distribuyen independientemente normales con

$$E(u_i) = 0, E(u_i u_j) = \begin{cases} \sigma_u^2 & \text{si } i=j \\ 0 & \text{si } i \neq j \end{cases}$$

Supóngase además que al término  $u$  se le puede descomponer en series de potencia de la forma

$$(2) \quad u = v P_g(z) = v \left\{ m_0 + m_1 f(z) + m_2 [f(z)]^2 + \dots + m_g [f(z)]^g \right\}$$

donde  $v$  es una variable aleatoria con  $E(v) = 0$  y  $E(v^2) = \sigma^2$ ,  $P_g(z)$  es un polinomio de grado  $g$  en función de una variable matemática  $z$ . Esta variable  $z$ , usualmente será una de las  $x_j$  de  $x$  la cual se presume afecta a la varianza.

Glejser supone que en la práctica,  $f(z)$  podría ser de la forma  $x_j^{1/2}$ ,  $x_j^{-1/2}$  o  $\log x_j$  y  $g$  de orden no mayor a 2, es mas, considera que en práctica es suficiente  $g=1$ . Aunque el valor de las  $m$ 's son desconocidos, Glejser pide que los términos del polinomio sean todos no negativos. En caso de estar en lo correcto, la relación de la varianza de  $u$  sería de la forma:

$$(3) \quad \sigma_u^2 = \sigma^2 [P_g(z)]^2$$

Por otro lado, tomando en cuenta (2), pueden suceder 2 cosas:

a)  $m_0 \neq 0$  y  $m_i = 0 \quad \forall i \neq 0$ , esto representaría que sí se está cumpliendo la suposición de homocedasticidad.

b)  $m_i \neq 0$  para alguna  $i \neq 0$ , en este caso, se piensa que existe heterocedasticidad y sabríamos, en cierta medida, la relación funcional específica de la varianza.

Ahora, si obtenemos la esperanza de las  $u$ 's, tendríamos lo siguiente:

$$(4) \quad E(|u|) = E(|v|)m_0 + E(|v|)m_1 f(z) + \dots + \\ + E(|v|)m_i [f(z)]^i + \dots + E(|v|)m_g [f(z)]^g$$

Glejser, propone estimar  $E(|v|)m_i$  por medio de una recta de regresión de los valores absolutos de los residuales  $|\hat{u}_j|$  en las  $[f(z)]^i$ . Aunque enfatiza que los términos de discrepancia en esta regresión ( $u^a$ , digamos) generalmente no tienen esperanza cero, él piensa que este efecto de sesgo, en general, no será importante en comparación con la ordenada al origen. Por tal motivo no los toma en consideración.

Entonces, sin apartarse de lo anterior, dado que se hacen estimaciones, se pueden hacer pruebas de hipótesis sobre las  $m$ 's y aunque Glejser no lo menciona claramente, las hipótesis a probar serían

$$(5) \quad H_0(I): m_0 = 0 \quad \text{vs.} \quad H_A(I): m_0 > 0 \\ H_0(II): m_1 = 0 \quad \text{vs.} \quad H_A(II): m_1 > 0$$

De este proceso pueden suceder dos casos los cuales Glejser considera relevantes:

a) El estimador de  $E(|v|)m_0$  (la ordenada al origen) no difiera significativamente de cero, mientras que el estimador de  $E(|v|)m_1$  (la pendiente), sea positiva y significativa. En

este caso, no se rechazaría  $H_0(I)$  pero si  $H_0(II)$  y se concluiría existencia de heterocedasticidad específica (a la que él llama heterocedasticidad "Clásica") la cual se elimina utilizando Mínimos Cuadrados Generalizados.

b) Los estimadores de  $E(|v|)_{m_0}$  y  $E(|v|)_{m_1}$  son ambos positivos y significativamente distintos de cero sobre un rango relevante. En este caso se concluiría la existencia de heterocedasticidad mixta.

Para los casos en que el estimador de  $E(|v|)_{m_0} = 0$ ,  $E(|v|)_{m_1} = 0$  y  $E(|v|)_{m_0} \neq 0$ ,  $E(|v|)_{m_1} = 0$ , se concluye Homocedasticidad.

Por su parte Hickman & Long (1980) hicieron una crítica detallada a la prueba de Glejser de la cual se desprenden los siguientes comentarios:

" Glejser dice, los términos en la suma de (1.2) [lo que equivale a (2)] son todos no negativos. Esto implica que  $f(x)$  tiene siempre ser ya sea no-positivo o no-negativo en un caso dado".

"...No es claro del artículo de Glejser si o no la hipótesis alternativa podría ser  $\alpha > 0$  o  $\alpha \neq 0$  [Las  $\alpha$ 's que trabajan los autores son las  $m$ 's]. Su desarrollo parece requerir la primera, pero su simulación aparentemente usa la última alternativa".

Glejser deshecha este problema [el problema del sesgo en los términos  $u$ 's] como si generalmente no fuese importante dado que él piensa que  $E(\xi_1)$  [ $E(u_1^*)$ ] será pequeña relativa a  $\alpha_0$ . Pero el problema no podría ser deshechado tan a la ligera. La habilidad de esta prueba para distinguir entre heterocedasticidad pura y mixta depende de la aceptación o rechazo de  $H_0(I)$ . Para tamaños de muestra finitos, si  $H_0(I)$  es re-

chazada, no hay forma de decir si el rechazo es debido a  $M_0 > 0$  [ $m_0 > 0$ ] o  $E(\xi_i) > 0$ . De hecho, cuando  $H_0(I)$  es aceptada, puede ser que  $M_0 > 0$  y  $E(\xi_i) < 0$  de tal forma que  $\hat{\alpha}_0$  es cercano a cero. Estos dos efectos son confundidos y no hay forma aparente de separarlos. También, la aceptación de  $H_0(I)$  pudiera venir sencillamente a causa de una gran varianza de  $\hat{\alpha}_0$ .

Así, en fundamentos teóricos no esperamos [Hickman & Long] que la prueba de Glejser permita correctamente distinguir entre heterocedasticidad pura y mixta.

#### PRUEBA DE LAYARD(1973).-

Layard sugiere una prueba basada en la estadística  $\chi^2$ . Este procedimiento utiliza la distribución de los logaritmos de las varianzas muestrales.

Supóngase que se tienen  $k$  muestras independientes :

$e_{i1}, e_{i2}, \dots, e_{in_i}, i = 1, 2, \dots, k$  y además  $e_{ij}, j=1, 2, \dots, n_i$  son independientes e idénticamente distribuidas con f.d.

$F((e_i - \mu_i)/\sigma_i)$  con cuartos momentos finitos pero tanto  $F$  como la localización y la escala de los parámetros son desconocidos.

Entonces lo que se desea probar es, si existe igualdad de varianza en las  $k$  muestras.

El procedimiento consiste en una prueba  $\chi^2$  la cuál está en función de la curtosis, es decir

$$\text{si } S_i^2 = \frac{\sum_{j=1}^{n_i} (e_{ij} - \bar{e}_i)^2}{n_i - 1}$$

es la varianza muestral de la  $i$ -ésima muestra correspondiente.

Entonces, para  $n_i$  grande, se puede mostrar<sup>1</sup> que:

$$(n_i - 1) \log S_i^2 \sim N(\log \sigma_i^2, r_i^2) \text{ donde}$$

---

1.- cf. Miller (1968) págs. 570-71.

$$r_i^2 = 2 + \left\{ 1 - (1/n_i) \right\} \gamma$$

y

$$\gamma = E(\varepsilon - \mu)^4 / \sigma^4 - 3$$

es el coeficiente de curtosis de la distribución.

Sea

$$r^2 = 2 + \left\{ 1 + (1/\bar{n}) \right\} \gamma \quad \text{con} \quad \bar{n} = \sum n_i / k$$

por tanto si  $n_i$  es grande (razonablemente), las  $r_i^2$  son aproximadamente igual a  $r^2$ , un término constante.

Ahora, haciendo una transformación al vector<sup>1</sup>:

$$\left[ (n_1-1)^{1/2} \log S_1^2, \dots, (n_k-1)^{1/2} \log S_k^2 \right]$$

Layard, muestra que bajo  $H_0$ ,

$$S' = \sum (n_i-1) \left( \log S_i^2 - \frac{\sum (n_i-1) \log S_i^2}{\sum (n_i-1)} \right)^2 / r^2$$

se distribuye asintóticamente como una  $\chi^2$  con  $k-1$  grados de libertad. Pero como en práctica  $\gamma$  es desconocida, no es posible basarse en  $S'$  para probar  $H_0$ . Por otro lado si se utiliza un estimador consistente de  $\gamma$ , el cual es dado por

$$\hat{\gamma} = \frac{\left( \sum_i n_i \right) \sum_i \sum_j (e_{ij} - \bar{e}_i)^4}{\left\{ \sum_i \sum_j (e_{ij} - \bar{e}_i)^2 \right\}^2} - 3$$

la prueba se basaría en:

$$S = \frac{r^2}{r^{\wedge 2}} S' \quad \text{con} \quad r^{\wedge 2} = 2 + \left[ 1 - (1/\bar{n}) \right] \hat{\gamma}, \text{ la cual}$$

se distribuye también como  $\chi^2$  con  $k-1$  grados de libertad.

1.- Layard op.cit., pág. 196

Scheffé puntualizó que para poder obtener un estimador razonablemente bueno de  $\gamma$  es indispensable tener muestras muy grandes.

Por tanto, si tenemos que bajo  $H_0$ ,  $S$  se distribuye como una estadística  $\chi^2$ , rechazamos la hipótesis nula si  $S$  rebasa el valor de  $\chi^2$  con  $k-1$  grados de libertad al  $(1-\alpha)100\%$  de significancia.

PRUEBA DE HARVEY & PHILLIPS (1973-74)  
(PRUEBA POR RESIDUALES RECURSIVOS) .-

Lo que motiva esta prueba es lo siguiente. Supóngase que se tiene el modelo de regresión lineal

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{u}$$

con las suposiciones básicas, en particular, hacen énfasis en que los términos  $u$ 's son independientes entre sí, con esperanza cero y que se distribuyen normalmente, Sobre la varianza no se menciona nada ya que la prueba es precisamente para detectar la forma de ella.

Entonces, la forma de la varianza del término  $\underline{u}$  será:

$$E(u u') = \sigma^2 V$$

donde  $V$  es una matriz diagonal y que bajo la hipótesis nula, es decir, de acuerdo a las suposiciones básicas  $V = I_n$ .

Por otra parte, similarmente a los residuales por mínimos cuadrados ordinarios los residuales recursivos (que se definen mas adelante) son lineales en la variable dependiente  $Y$  e insesgados en el sentido que tienen esperanza cero. Pero conforme a las suposiciones básicas, los residuales por mínimos cuadrados ordinarios en general, son correlacionados y heterocedásticos. En otras palabras, tenemos que en la hipóte-

sis nula  $E(uu') = \sigma^2 I_n$ , si definimos  $\hat{u}$  como el vector de residuales por mínimos cuadrados, entonces  $E(\hat{u}\hat{u}') = \sigma^2 (I_n - X(X'X)^{-1}X')$ .

Esto no ocurre con los residuales recursivos ya que la forma de construcción hace que bajo la hipótesis nula, sean independientes e idénticamente distribuidos. De esta forma podemos tener que  $E(\tilde{u}\tilde{u}') = \sigma^2 I_{n-k}$ , donde  $\tilde{u}$  denota el vector de residuales recursivos de dimensión  $(n-k) \times 1$ .

La explicación viene del hecho de que es posible definir a lo mas  $n-k$  residuales si estos son restringidos a tener una matriz de covarianza escalar<sup>1</sup>.

La idea para obtener los residuales recursivos es calcular  $n-p+1$  estimadores de  $\beta$ , donde el primer estimador  $\hat{\beta}_p$  es tá basado en las primeras  $p$  observaciones, el segundo en las  $p+1$  observaciones, etc.

Apoyándose en esa idea, dichos estimadores pueden ser estimados recursivamente utilizando la siguiente expresión:

$$\hat{\beta}_k = b_{k-1} + \frac{(X'_{k-1}X_{k-1})^{-1} \chi_k (y_k - \chi_k' \hat{\beta}_{k-1})}{1 + \chi_k' (X'_{k-1}X_{k-1})^{-1} \chi_k}$$

donde  $X_{k-1}$  es la matriz generada por los primeros  $k-1$  renglones de  $X$ ,  $\chi_k'$  es el  $k$ -ésimo renglón de  $X$  y  $y_k$  es el  $k$ -ésimo valor del vector  $y$ .

Por otro lado, sólo se necesita realizar una sola inversión matricial, es decir, sólo se necesita invertir  $(X'_{k-1}X_{k-1})$  ya que para obtener la inversa de  $X'_k X_k$  se puede utilizar la fórmula recursiva:

$$(X'_k X_k)^{-1} = (X'_{k-1} X_{k-1})^{-1} - \frac{(X'_{k-1} X_{k-1})^{-1} \chi_k \chi_k' (X'_{k-1} X_{k-1})^{-1}}{1 + \chi_k' (X'_{k-1} X_{k-1})^{-1} \chi_k}$$

1.- para la demostración de esto ver Theil(1971, pg.203-204) 46

y por tanto, se pueden definir los residuales recursivos como las  $n-p$  estadísticas:

$$\tilde{u}_k = \frac{y_k - x_k' \hat{\beta}_{k-1}}{\left\{ 1 + x_k' (X_{k-1}' X_{k-1})^{-1} x_k \right\}^{1/2}} ; k = p+1, \dots, n$$

Se ha demostrado<sup>1</sup> que bajo la hipótesis nula estos residuales son insesgados con valor esperado igual a cero y que tiene una matriz de covarianzas escalar.

Por tanto, ya que los residuales recursivos poseen estas características, las cuales son deseables, se construye una prueba de significancia bastante fácil. La prueba es como sigue:

Si se pueden obtener  $n-p$  residuales recursivos, a este conjunto, por su parte se le puede dividir en dos grupos con  $m$  elementos, donde  $m \leq \frac{n-k}{2}$ . Entonces, si denotamos por  $\tilde{u}_1$  al vector de los primeros  $m$  residuales recursivos y por  $\tilde{u}_2$  al vector de los últimos  $m$  residuales recursivos, no es muy difícil demostrar<sup>2</sup> que bajo la hipótesis nula:

$$\frac{\tilde{u}_2' \tilde{u}_2}{\tilde{u}_1' \tilde{u}_1} \sim F_{(m,m)}.$$

Intuitivamente lo que hace la prueba es medir la diferencia entre la suma de cuadrados debida a los residuales del segundo grupo con respecto a la suma de cuadrados de los residuales del primero.

En caso de que se viole la hipótesis, el cociente sería bastante grande y su valor numérico en consecuencia rebasaría el valor de  $F_{(m,m)}$ . De hecho es una prueba similar a la pro -

1.- por ejemplo ver Harvey & Phillips (1973)

2.- cf. Graybill(1961), pág.31

puesta por Goldfeld & Quandt.

PRUEBA DE ZROETER(1978)  
(PRUEBA PARA MODELOS CON PREDICTORES NO ALEATORIOS).-

La idea de esta prueba es comparar la magnitud de cada residual con respecto a la magnitud del vector de residuales.

Para construir la estadística de prueba, se multiplica cada uno de los cocientes por un valor, para ponderarlos y se suman todos los términos. Se compara con un valor basado en cuantiles de las tablas del cociente Von-Newmann y cuantiles de la prueba Durbin-Watson.

Formalmente, el proceso de prueba es el siguiente. Supóngase que tenemos el modelo de regresión lineal donde las variables independientes no son aleatorias y el término de discrepancia  $u_i$ , es independiente de cualquier otro, con  $E(u_i)=0$  y  $E(u_i^2) = \sigma_i^2$ . Que aunque la forma de  $\sigma_i^2$  no puede ser conocida, se supone que las observaciones han sido ordenadas de tal forma que  $\sigma_{i-1}^2 < \sigma_i^2$  ( $i=2, \dots, n$ ). (el ordenamiento de las observaciones pudiera ser de acuerdo al criterio de Goldfeld & Quandt(1965)).

Entonces la hipótesis de interés es  $H_0$  contra la hipótesis alternativa  $H_1: \sigma_i^2 > \sigma_{i-1}^2$ .

Para hacer tal prueba, considérese lo siguiente:

Sea  $A$  un subconjunto no-vacío de  $\{1, 2, \dots, n\}$  y sea  $\{u_i; i \in A\}$  el cual denota un conjunto de residuales que en general se aproximen a los términos de discrepancia  $u$  en  $\{u_i; i \in A\}$ . Entonces para nuestro propósito se supone que  $A$  es un subconjunto de  $\{1, 2, \dots, n\}$  construido de  $(n-p)$  elementos y que

los residuales  $\tilde{u}_i (i \in A)$  bajo la hipótesis nula tienen esperanza cero y matriz de covarianza escalar. En particular estos residuales pueden ser formados por un conjunto de residuales recursivos (cf. Harvey & Phillips(1974)).

Sea  $\{h_{in}; i \in A\}$  algún conjunto de  $n_A$  escalares no aleatorios teniendo la propiedad  $h_{in} \leq h_{jn}$  si  $i < j$ . Sea además  $\mathcal{Z}(A,i)$ ,  $i \in A$  una función tal que  $i$  es el  $\mathcal{Z}(A,i)$ -ésimo elemento mas pequeño de  $A$ . Entonces, para cada uno de los valores de  $t$  tomados de 1 a 6, la ecuación siguiente tiene  $h_{in}$  no decreciente en  $i$ .

$$h_{in} = (-1)^{t+1} 2(1 - \cos \psi_t(i)) , \quad i \in A$$

donde

$$\begin{aligned} \psi_t(i) = \pi \left\{ (\delta_{t2} + \delta_{t4} + \delta_{t6})(n-p+1) + (\delta_{t5} + \delta_{t6})(J-1) + \right. \\ \left. + (-1)^{t+1} \mathcal{Z}(A,i) \right\} / \left\{ n-p+(1-\delta_{t1}-\delta_{t2})J + \right. \\ \left. + \delta_{t1} + \delta_{t2} \right\} \end{aligned}$$

con  $J$  igual a un entero positivo mayor o igual a 2, el cual es elegido arbitrariamente y  $\delta_{ij}$  la  $\delta$  de Kronecker definida como :

$$\delta_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j. \end{cases}$$

Consiguientemente, para cada uno de los valores de  $t$  se produce una prueba exacta al  $100\alpha$  por ciento de significancia (Zroeter(1978), págs. 1312-1315) en la que se rechaza  $H_0$  si y solo si

$$\begin{aligned} \tilde{h}_n &= \sum_{i \in A} \tilde{w}_{in} h_{in} \\ \text{con } \tilde{w}_{in} &= \tilde{u}_i^2 / \left( \sum_{j \in A} \tilde{u}_j^2 \right) ; \quad j \in A \end{aligned}$$

es mayor que

$$\begin{aligned} & \left\{ (n-p+1)^{-1} (n-p) ( \delta_{t1} v_U(n-p+1; \alpha) - \delta_{t2} v_L(n-p+1; \alpha) ) + \right. \\ & + 4( \delta_{t3} + \delta_{t5} ) - ( \delta_{t3} + \delta_{t6} ) d_U(n-p+ J, J; \alpha) - \\ & \left. - ( \delta_{t4} + \delta_{t5} ) d_L(n-p+ J, J; \alpha) \right\} . \end{aligned}$$

donde  $v_U, v_L; d_U, d_L$  se obtienen en las tablas del cociente Von-Newmann y Durbin-Watson (Tabla VI y Tabla VII) respectivamente.

## CAPITULO III

### SIMULACIONES

Este capítulo, se incluye con la finalidad de dar una idea, al lector, del comportamiento de cada una de las pruebas, para la detección de heterocedasticidad, vistas en el capítulo anterior.

Toda la información fué extraída de artículos (de los cuales se dá la referencia exacta) publicados por diferentes personas quienes han realizado simulaciones experimentales, ya sea para calcular la potencia de alguna prueba propuesta por ellos o simplemente para comparar la potencia de pruebas que ellos consideran de interés. Se trató de dar en forma clara y breve la manera en que se llevaron a cabo cada uno de los estudios y los resultados que se obtuvieron (algunos autores hacen recomendaciones, las cuales también se incluyen); persiguiendo con esto, que cada quien pueda formarse un criterio de como proceder en determinadas circunstancias.

Para finalizar esta pequeña introducción, se aclara que el procedimiento para la presentación de la información fué en orden cronológico para ser coherentes respecto a la presentación de las pruebas.

#### BISHOP Y NAIR(1939).-

Realizaron un estudio sobre la prueba propuesta por Neyman y Pearson(1931) y la prueba Bartlett(1937). Este análisis consistió en revisar cada uno de los métodos de aproximación

sugeridos por diversas personas tales como Nayer(1936), Batlett(1937) y Nair(1939), en base a estas aproximaciones, calcularon la potencia de cada prueba y establecieron las diferencias.

Notaron que cuando los grados de libertad (en los que esta basada  $S_i^2$ ,  $i=1,2,\dots,k$ ), así como para cuando se tenían dos grupos; la prueba de Bartlett da una subestimación de significancia de violación de homogeneidad. Notaron además, que si cada uno de los  $k$  estimadores de la varianza  $S_i^2$  es están basados en los mismos grados de libertad, no hay diferencia entre el uso de la prueba de Bartlett y la de Neyman y Pearson

#### Puntualizan:

"Debería ser recordado que Bartlett nunca esperó que su método pudiera ser muy preciso cuando los estimadores estuvieran basados en grados de libertad tan pequeños como 1". (Bishop y Nair, op. cit., pág. 96).

aunque ellos observaron que cuando los grados de libertad en los que se basan los estimadores fueran tan pequeños como la unidad, la prueba podría detectar solamente en el caso de que la violación de homogeneidad fuera muy grande.

Para cuando los estimadores están basados en diferentes grados de libertad, Bartlett argumentó que su estadística era un perfeccionamiento del método Neyman y Pearson, puesto que si se eliminan varios parámetros de regresión, de muestras desiguales en tamaño, pudiera ser posible que las fluctuaciones de una varianza reducida a uno o dos grados de libertad, aparenten la discrepancia real en muestras más estables.

Por otro lado, los autores del artículo reafirman lo anterior por medio de una simulación, concluyendo que la prueba Neyman--Pearson es muy sesgada, respecto al riesgo de error tipo I nominal, al menos para cuando las diferencias entre el número de observaciones por grupo y los grados de libertad (en los que se basan los estimadores) es muy grande.

HARTLEY(1940), por su parte demuestra que la prueba de Bartlett no es muy precisa si el número de muestras ( $k$ ) es muy grande. ( $k=30$  prueban).

GOLDFELD & QUANDT(1965), con el fin de calcular la potencia de la prueba propuesta por ellos, llevaron a cabo una simulación de un modelo de regresión lineal simple con una variable independiente. Variaron el tamaño de la muestra ( $n=30,60$ ), el número de observaciones centrales omitidas ( $k=0,4,8,12,16$ ), la media ( $\mu_x$ ) y la desviación estándar de los valores de  $x$  ( $\sigma_x$ ). Los valores de  $x$  fueron generados de una distribución uniforme para cada combinación de los valores de  $\mu_x$  y  $\sigma_x$ ; restringiendo el uso de los valores de  $x$  a aquellos que cumplieran con  $x > 1$ .

Para cada combinación de parámetros, calcularon 100 muestras de cada tamaño. Los valores del término de discrepancia  $u$ , los generaron de una población  $N(0,1)$ . Específicamente, el modelo utilizado fué:

$$(1) \quad Y_i = \beta_0 + \beta_1 X_i + u_i$$

por tanto las hipótesis probadas fueron:

$$(2) \quad \hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

vs.

$$(3) \quad \hat{Y}_i / X_i = \hat{\beta}_0 1 / X_i + \hat{\beta}_1$$

Las estadísticas obtenidas se compararon con los cuantiles correspondientes de una distribución  $F$  al nivel de significancia 5%. Entonces, para conocer la potencia de la prueba, se obtuvieron las frecuencias relativas (en los 100 ensayos) de

i) Los casos en que se rechaza la hipótesis falsa (3)

ii) Los casos en que se rechaza la hipótesis correcta (2)

Como vieron que los dos casos eran equivalentes, se limitaron al caso (2). A partir de los resultados a los que llegaron, hicieron las siguientes conclusiones:

a) La máxima potencia de la prueba, para cuando los tamaños de muestra son 30 o 60, la obtuvieron con el número de observaciones centrales omitidas ( $k$ ) igual a 8 y 16 respectivamente.

b) Cuando existen incrementos en la desviación estándar de la variable  $x$  ( $\sigma_x$ ) relativa a su media; la potencia de la prueba mejora.

c) En tanto el tamaño de la muestra aumente, la potencia de la prueba mejora.

PEARSON (1966), comparó las pruebas Bartlett, Hartley y Cadwell (para lo cual necesitó tamaños iguales de muestras).

El estudio fué como sigue: consideró 1000 muestras aleatorias, extraídas de poblaciones normales  $N(0,1)$ ; cada una consistiendo de  $n=5$  observaciones. Estas 1000 muestras, las dividió en 200 grupos de  $k=5$  muestras cada uno. Multiplicó, la varianza, la desviación estándar o rango (según se requiriera) por factores apropiados tales que las 5 muestras se pudieran considerar como muestras extraídas de poblaciones normales con

diferente varianza, desviación estándar o rango ( $\sigma_t, t=1,2,3,4,5$ ) y examinó la potencia de cada prueba para detectar heterocedasticidad.

Estas pruebas de hipótesis las realizó a 5% y 1% niveles de significancia, de lo cual concluye lo siguiente:

a) En general, la prueba de Hartley es ligeramente más potente que la prueba propuesta por Cadwell. Aunque a cualquier proporción de los valores elegidos de  $k$  y  $n$ , la diferencia entre las dos pruebas es muy pequeña.

b) La potencia de la prueba de Bartlett, es considerablemente mayor que cualquiera de las dos pruebas. Aunque en algunas hipótesis alternativas que él consideró, la diferencia no es mucha. Esto le llevó a pensar que la prueba de Hartley y la de Cadwell se podría esperar fueran sensitivas en situaciones donde hay una varianza poblacional muy grande y una muy pequeña.

c) Si se necesita un análisis rápido sobre presencia de heterocedasticidad podría ser muy útil la prueba de Cadwell. Pero si ya se poseen por separado los estimadores muestrales de la varianza y además se dispone de tiempo, recomienda la prueba de Bartlett. Sin embargo, enfatiza que no debe olvidarse que ninguna de las pruebas es robusta a violación de la su posición de normalidad.

Seis años después GARTSIDE(1972) hizo un estudio similar de estas tres pruebas y comenta lo siguiente:

a) Si se desconoce la hipótesis alternativa pero se cree que existe normalidad lo recomendable es utilizar la prueba de Bartlett (prueba que encontró más potente).

b) Si se necesita una prueba rápida y se cree que exis-

te normalidad, él, aconseja utilizar cualesquiera de las dos pruebas Hatley o Cadwell.

MILLER(1968).-

Comparó su prueba con la prueba propuesta por Box-Scheffé y la prueba propuesta por Levene a 5% y 1% niveles de significancia. En su estudio, Miller, consideró dos variantes que Levene sugiere para su prueba<sup>(1)</sup>. Su experimento consistió en lo siguiente, generó dos veces 1000 pares de muestras aleatorias de una distribución Uniforme, el tamaño de cada muestra fué de  $N=25$ . A cada una de las muestras las sometió a una transformación apropiada para obtener muestras de distribuciones como: Normal, doble exponencial y doble exponencial sesgada.

En un primer estudio, expuso lo siguiente:

a) La prueba Jackknife con  $k=1$ , es generalmente mas potente, comparada con la prueba S de Levene. La prueba Jackknife aumenta en potencia cuando se trabaja al nivel de significancia 1%.

b) La potencia de la prueba Jackknife no es tan grande cuando  $k=5$  como cuando  $k=1$ . Sin embargo tiene un mejor control en el riesgo de error tipo I respecto al nivel de significancia nominal.

c) La prueba S de Levene<sup>(2)</sup> es robusta, pero no es tan potente como la prueba de Box-Scheffé o la prueba Jackknife en distribuciones con colas muy largas. En particular al nivel de significancia 1% se obtiene la mas baja potencia.

---

1.- Las variantes son a) obtener  $u_1 = (\sum x_i - \bar{x})^2$  y b) obtener  $z_1 = |x_i - \bar{x}|$  las cuales denota por la prueba S y la prueba Z respectivamente, cada una de ellas se distribuye como una estadística t-Student. NOTA. Draper & Hunter(1969) en su artículo, hacen una extensión de la prueba Z.

2.- En un estudio que Levene realizó, encontró que la prueba Z es preferible (en términos de potencia) a la prueba S. Sin embar-

d) La prueba de Box con  $k=5$  es robusta y muy similar (en potencia) a la prueba Jackknife con  $k=5$ . Pero en este caso, por facilidad de cálculos, aconseja la prueba de Box -- Scheffé.

En un segundo estudio que realizó, Miller, corroboró los resultados anteriores. Finalmente puntualiza que para muestras grandes, la prueba Jackknife (con  $k=1$ ) es muy potente en comparación con la prueba S de Levene y la de Box-Scheffé. Sin embargo, en tanto se incrementa el tamaño de la muestra, la cantidad de cálculos que se necesitan para dicha prueba es demasiado. Cuando se utiliza  $k > 1$ , usar la prueba Jackknife es equivalente a utilizar la prueba S de Levene ( en cuestión de potencia).

GLEJSER(1969), con el fin de medir la efectividad de su prueba, la comparó con la prueba propuesta por Goldfeld & Quandt.

El estudio lo realizó utilizando un modelo de regresión lineal simple

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

donde varió la desviación estándar de las  $u$ . Las  $u$ 's fueron generadas de una población normal con media igual a 50. Tomó varios tamaños de muestras y consideró ocho tipos de relación funcional de la varianza (Ver Glejser(1969) pág. 318). Para cada caso, generó 1000 muestras. Los valores de  $v$  los generó de una población normal con media cero y varianza unitaria.

El criterio para evaluar la potencia de las pruebas, fué tomar en cuenta la frecuencia relativa del error tipo I (i.e. detectar heterocedasticidad cuando no existe). De los resultados obtenidos después del estudio, Glejser concluye lo siguiente

---

Miller en este artículo, demuestra que la prueba Z no es asintóticamente distribución libre. Por tal motivo no la consideró competidora para su prueba

te:

a) La prueba de Glejser es favorable comparativamente a la de Goldfeld & Quandt<sup>(1)</sup>. Aunque para desviaciones estándar pequeñas de  $\alpha$  y tamaños de muestra reducidos, la potencia de ambas es muy baja.

b) La prueba de Glejser tiene ventaja con respecto a la de Goldfeld & Quandt ya que la primera puede detectar heterocedasticidad mixta.

Glejser, dió las condiciones para los que se obtienen óptimas respuestas de su prueba, i.e. la prueba ayuda bastante en los siguientes casos:

i) Cuando la muestra consta al menos de 30 observaciones.

ii) Cuando la desviación estándar de  $\alpha$  es grande (p.e.  $\sigma_x=30$  y media igual a 50).

iii) Cuando la potencia de la  $\alpha$  que aparece en la forma funcional de  $u$ , en valores absolutos no es menor que 1.

Por su parte, Hickman y Long(1980), hicieron una crítica a la prueba propuesta por Glejser. En este estudio exhaustivo (vía un estudio similar al realizado por Glejser) llegaron a las siguientes conclusiones:

1) La prueba de Glejser identificará homocedasticidad correctamente con muestras tan pequeñas como 30 observaciones sólo si la varianza de la variable independiente  $\alpha$  es muy grande.

2) Para el caso de heterocedasticidad para ( $u=vx$ ), el método de Glejser detecta bien cuando el tamaño de la muestra es mayor que 30 y los valores de la variable  $\alpha$  tienen una desviación estándar grande. Sin embargo, esto no sucede cuando la varianza del término de error  $u$  decrece en tanto que los valo-

---

1.- La prueba de Goldfeld & Quandt es mas potente solo cuando se tiene heterocedasticidad pura y la desviación estándar de  $\alpha$  es muy grande.

res de  $\alpha$  se incrementan (producen resultados no satisfactorios).

3) En el caso de heterocedasticidad mixta, similar al punto anterior, los resultados obtenidos de aplicar la prueba de Glejser son bastante insatisfactorios.

La razon argumentada del por qué de los resultados es, que al no tomar en cuenta el sesgo de los estimadores de los términos de disorepancia, en el ajuste por regresión de los residuales con respecto a los valores de  $f(z)$ , no hay forma de saber si la aceptación o rechazo de la hipótesis es debida a que  $m_0 > 0$  y  $E(u_i^2) > 0$ . Por ejemplo, cuando la hipótesis nula I se acepta es porque  $\hat{m}_0 > 0$  y  $E(u_i^2) < 0$ . De hecho, en la forma funcional de la varianza que ellos (Hickman y Long) consideran,  $E(u_i^2) < 0$  mientras que  $\hat{m}_0 > 0$ . Esto les lleva a aceptar  $H_0(I)$ :  $m_0 = 0$  cuando realmente debería ser rechazada.

GAMES, WINKLER Y PROBERT (1972). Hicieron un estudio sobre la robustez de las pruebas Hartley, Bartlett, Levene<sup>(1)</sup> y Scheffé<sup>(2)</sup>. El estudio consistió en dos simulaciones de muestreo, donde en el primero se obtuvieron 1000 grupos de 3 muestras de 6 observaciones cada una. En este estudio se compararon las pruebas Bartlett, Hartley, L-A y L-x<sup>2</sup>.

El segundo estudio consistió de 1000 grupos de 3 muestras, cada una consistiendo de 18 observaciones, en éste, se compararon las pruebas Bartlett y Scheffé<sup>(3)</sup>.

1.- Los autores, al igual que en el estudio de Miller (1968) tra tan las variantes de la prueba de Levene  $|x_i - \bar{x}|$  y  $(x_i - \bar{x})^2$  las cuales denotaron como L-A y L-x<sup>2</sup> respectivamente.

2.- Games, Winkler y Probert, se refieren a la prueba de Scheffé como la prueba de Bartlett-Kendall pero como en este trabajo siempre se ha tratado como la prueba de Scheffé, se seguirá trabajando con este mismo nombre.

3.- La prueba de Scheffé fué trabajada de dos formas, la muestra

Las observaciones las generaron de una distribución normal, donde variaron los parámetros  $\mu$  y  $\sigma^2$ . El total de variaciones fueron 6 obteniéndose de esta manera distintas formas poblacionales, éstas son:

- 1.- Población normal (0,1)
- 2.- Población normal ligeramente sesgada
- 3.- Población normal moderadamente sesgada
- 4.- Población normal extremadamente sesgada (muy leptocurtica).
- 5.- Simétrica leptocurtica
- 6.- Rectangular

Cada una de las estadísticas consideradas en el estudio fué probada a los niveles de significancia 5%, 2.5%, 1% y 0.1% (cuando fué posible). Para obtener la potencia de cada prueba, calcularon la frecuencia relativa del rechazo de  $H_0$  cuando era verdadera.

Los resultados que obtuvieron fueron, entre otros, los que a continuación se describen:

En el primer estudio encontraron que

a) En general, la estadística de Bartlett y la de Hartley se comportan de la siguiente forma: En poblaciones PLATICURTI - CAS producen pruebas conservativas, mientras que cuando existe LEPTOCURTOSIS la probabilidad de error tipo I es inflado. En particular la prueba de Bartlett, en el caso de las poblaciones 4, 5 y 6 mostró que en tanto que  $n$  crece, la probabilidad de error tipo I se incrementa aun mas. Esto, ya había sido comentado por Box(1953).

b) Cuando las varianzas difieren, la estadística de Bart-

---

(la cual constaba de 18 observaciones) se subdividió en 9 submuestras de 2 observaciones y en 6 submuestras de 3 observaciones cada una. A estos casos, los autores les llamaron LEV2 y LEV3 respectivamente

Lett y la de Harley resultaron mas potentes que la L-A y  $L-x^2$ , de estas dos últimas la mas potente fué L-A. Aunque esta potencia siempre estuvo acompañada con un riesgo de error tipo I muy inflado.

c) Sobre la falta de potencia, tanto en L-A como en  $L-x^2$ , los autores creen que esto, probablemente se debe a que los valores tanto de  $|x_{ik} - \bar{x}|$  como de  $(x_{ik} - \bar{x})^2$  no son independientes en una muestra y esta dependencia aumente cuando el valor de  $n$  es pequeño.

En el segundo estudio, encontraron que:

a) Para las poblaciones 1,2 y 3 la prueba de Bartlett fué mas potente, la prueba LEV3 mostró el mejor control de error tipo I en las tres poblaciones, pero siempre estuvo acompañada de una baja potencia.

b) La prueba LEV3 mostró ligeras desviaciones significantes del nivel 0.5 y esta junto con la LEV2 fueron ligeramente conservativas en todas las poblaciones.

c) Para las poblaciones 4 y 5, encontraron que la prueba fué mas potente, sin embargo el riesgo de error tipo I llegó a 37%. La prueba LEV3 mostró una potencia razonable y de nuevo la prueba LEV2 fué conservativa y de menor potencia.

Respecto a las pruebas conservativas, los autores puntualizan:

"La probabilidad de error tipo I inflada es un problema mucho mas serio que una probabilidad de error tipo I conservativa, puesto que es posible para una prueba ser conservativa y tenga una potencia mas alta que otras alternativas". (Games, et.al., op.cit., pág.893).

De estos resultados, los autores dan las siguientes recomendaciones generales:

1) Si se tiene razón para creer que las poblaciones muestreadas son platicurticas o mesocurticas, lo aconsejable sería utilizar la prueba de Bartlett o la prueba de Hartley. Aunque si bien es cierto que producen pruebas conservativas en poblaciones platicurticas, tienen mas potencia que cualquier alternativa.

2) Si se sospecha que las poblaciones muestreadas son leptocurticas, o no se tiene claro cual es la forma de la población y se desea una probabilidad de error tipo I igual a 0.05; lo recomendable es utilizar el método LEV3 y aunque tiene poca potencia, esta puede ser compensada con un incremento en n. Si se quiere hacer inferencias sobre varianzas como fin prioritario, lo que recomiendan es utilizar un tamaño de muestra mínimo de 20 observaciones.

En la prueba de Scheffé, los autores experimentaron sobre la partición óptima de la muestra. Para esto, utilizaron tamaños iguales de submuestras con 2,3,4,5 y 6 observaciones y una población con alto grado de curtosis. De esto obtuvieron los siguientes resultados.

Si se tiene muestras grandes ( $n=36$ , a este tamaño probaron los autores), el utilizar tamaños de submuestras igual a 4,5 o 6 produce una diferencia muy pequeña en potencia que cuando se usan tamaños de submuestras igual a 3. En caso de tener muestras pequeñas ( $n=18$ ), la prueba refleja una ligera mejoría en potencia utilizando submuestras de tamaño 4,5 o 6 que en submuestras de tamaño 3. En cambio, cuando utilizaron submuestras de tamaño igual a 2, obtuvieron una potencia inferior. Es

por esta razón que los autores recomiendan utilizar submuestras de tamaño igual a 3. Hacen la aclaración de que puede suceder que la muestra no pudiera ser subdividida en submuestras de tamaño igual a 3, a lo que ellos recomiendan, obtener el mayor número de submuestras de 3 observaciones y la última que fuera de 4 o 5, la cual debe ignorarse el tamaño, ya que el efecto no es significativo.

LAYARD(1973). Con el fin de probar la eficiencia de su prueba, llevó a cabo dos simulaciones de muestreo generadas de distribuciones Uniforme, Normal y doble exponencial. En estos dos estudios, comparó las pruebas propuestas por Bartlett, Layard ( $\chi^2$ ), Box-Scheffé y Layard(extensión de la prueba de Miller).

La primera simulación constó de 500 conjuntos (para cada población) de 4 muestras donde cada muestra contenía 25 observaciones. Después de aplicar cada una de las pruebas se cuantificó el número de rechazos de la hipótesis nula con el propósito de calcular la potencia. La segunda simulación, fué similar a la primera con la variante de que cada muestra tenía 10 observaciones. Todo el proceso se realizó a los niveles de significancia 5% y 1%.

Para aplicar la prueba de Scheffé, cada muestra fué subdividida en submuestras de tamaño 5.

Los resultados que el autor obtuvo fueron los siguientes:

a) La prueba de Bartlett presentó muy poca potencia para la distribución uniforme (la cual tiene  $\delta = -1.2$ ) y mucha potencia para la distribución normal y doble exponencial.

b) La prueba  $\chi^2$  y Jackknife presentaron muy buena potencia a los dos niveles de significancia, aunque la prueba Jackknife se mostró mas sensitiva a la forma de la curva.

c) En el caso de una curva normal, la prueba de Bartlett, rebasa a la prueba  $\chi^2$  y Jackknife.

d) La prueba de Scheffé, en general, presentó poca potencia pero fué la mas robusta a la forma poblacional.

De este análisis, el autor, hace el siguiente comentario.

Las pruebas  $\chi^2$  y Jackknife son aconsejables si se quiere un procedimiento razonablemente robusto pero bastante potencia. De estas dos, la prueba  $\chi^2$  es mucho mas fácil de calcular. Finalmente, hizo la aclaración de que las pruebas trabajaron mejor con muestras de tamaño 25 que con las de tamaño 10.

HARVEY Y PHILLIPS(1974). Con el fin de saber la efectividad de su prueba, la compararon con la prueba de Goldfeld & Quandt, para poder hacerlo hicieron una simulación de muestras, el procedimiento fué similar al estudio realizado por Goldfeld & Quandt, de los resultados que obtuvieron sacaron la siguiente conclusión.

La prueba por Residuales Recursivos es preferible globalmente a la prueba de Goldfeld & Quandt ya que es mas potente y mas flexible en la ordenación de las observaciones.

BROWN Y FORSYTHE(1974). Llevaron a cabo un experimento para comparar la potencia de cada una de las siguientes pruebas Levene<sup>(1)</sup>, Layard ( $\chi^2$ ) y Layard (Jackknife). El experimento consistió en generar pares de muestras aleatorias<sup>(2)</sup> de poblaciones Gaussianas, Ji-cuadrada (con 4 grados de libertad),

---

1.- Los autores consideran la prueba de Levene (denotada por W) en la siguiente forma  $z_{ij} = |x_{ij} - \bar{x}_i|$  y las variantes  $\sqrt{z_{ij}}$ ,  $\log(z_{ij})$  donde las últimas mostraron muy poca potencia respecto a  $z_{ij}$  (a la estadística W formada por esta transformación la denotaron  $W_0$ ). Además utilizaron estimadores centrales mas robustos que la media aritmética (como la mediana y la media reducida. A las estadísticas generadas por estas transformaciones las denotaron por

t-student (con 4 grados de libertad) y Cauchy. La distribución Ji-cuadrada, la sesgaron a la derecha en tanto que las distribuciones t-student y Cauchy fueron simétricas pero con colas largas. En la segunda muestra de cada par se transformaron los datos con el fin de que las varianzas fueran diferentes. Finalmente cada par de muestras (de cada población y para cada estadística) lo repitieron 1000 veces en bloques de 100 y cuantificaron el número de rechazos de  $H_0$  cuando ésta era verdadera.

Los resultados obtenidos en este estudio fueron los que a continuación se citan.

a) En caso de muestras pequeñas, la estadística  $W_{50}$  se mostró conservativa cuando existe igualdad de varianzas en poblaciones Gaussianas.

b) En caso de muestras desiguales, la prueba Jackknife, da resultados muy inflados en distribuciones con colas grandes.

c) La prueba  $\chi^2$  se desvía del nivel de significancia nominal con muestras muy pequeñas.

d) De los tres tipos de la prueba de Levene,  $W_{10}$  es aparentemente mas robusta y se acerca mas al 5% que  $W_{50}$ , mientras que  $W_0$  es más sensible a la forma poblacional.

A partir de esto, los autores aconsejan lo siguiente: Si se tiene poblaciones con colas grandes y se quiere probar igualdad de varianzas, lo recomendable es utilizar una estadística de la forma  $W_{10}$  ya que la pérdida en potencia que se tiene ( $W_{50}$  y  $W_{10}$  respectivamente).

2.- Los tamaños de muestra de cada par son variables, es decir, hay casos que dentro de cada par, los tamaños de muestra son diferentes.

cuando se utiliza  $W_{10}$  en lugar de utilizar  $W_0$  es pequeña relativa al incremento de probabilidad de error tipo I debida a no-normalidad.

Si se cree que existe simetría, lo recomendable es una estadística de la forma  $W_{50}$ .

BESELMAN, GAMES Y CLINCH(1979). Realizaron una simulación de muestreo para probar la potencia de las pruebas Jackknife,  $W_{10}$ ,  $W_{50}$  y Scheffé. Este estudio consistió en generar muestras aleatorias de las poblaciones Normal, Ji-cuadrada ( con 2 grados de libertad ) y t-student ( con cinco grados de libertad ), donde las últimas dos poblaciones con un grado de curtosis ( $\hat{\nu}$ ) igual a 6. Tomaron dos tamaños de muestra las cuales fueron divididas en cuatro grupos de diferentes tamaños; la muestra mas pequeña fué de  $N_1=180$  observaciones ( las submuestras fueron de tamaños  $n_k = 18, 36, 54, 72$  ).

La segunda muestra fué de tamaño  $N_2=420$  y las submuestras de tamaños  $n_k=36, 72, 108, 144$  <sup>(1)</sup>. Las varianzas fueron igual a 8, 7, 5 y 4 las cuales aparearon directa e inversamente, de acuerdo a los tamaños de las submuestras. El nivel de significancia utilizado fué de 5%. Para aplicar la prueba de Scheffé, tuvieron que dividir cada submuestra en subgrupos. Esta división fué de acuerdo a la sugerencia de Martin y Games (1977) basada en la raíz cuadrada positiva de la menor submuestra las cuales fueron 5 y 6 para  $N_1$  y  $N_2$  respectivamente.

Los resultados que obtuvieron de este análisis fueron:

a) La prueba Jackknife reflejó tendencia a inflar el riesgo de error tipo I en poblaciones Leptocurticas. Cuando las muestras son de tamaños desiguales esta tendencia se incrementa

---

1.- En particular, los autores consideraron submuestras de tamaño  $n_k = 45 \forall k$ .

ta y cuando los tamaños de muestra son iguales la prueba da mayor potencia sin embargo el riesgo de error tipo I sigue inflado.

b) La prueba  $W_{10}$  tiene un comportamiento razonablemente bueno en distribuciones simétricas. Pero presenta un riesgo de error tipo I inflado en poblaciones asimétricas.

c) Las pruebas  $W_{50}$  y Scheffé, en general, garantizan un buen control de error tipo I. Aunque las dos presentan un ligero sesgo positivo.

Para concluir, resumiremos los resultados que se expusieron en el transcurso de este capítulo. Con este propósito, se presentan los resultados obtenidos de cada prueba (i.e., el comportamiento de cada prueba en diferentes situaciones). De nuevo, el orden en que se citan las pruebas es análogo al ya descrito.

#### PRUEBA DE BARTLETT.-

- Da una subestimación de significancia de violación de homocedasticidad cuando los grados de libertad en que se basan los  $S_i$  son tan pequeños como la unidad y solo la detecta cuando la varianza es muy grande.
- Cuando el número de muestras es muy grande, la prueba no es muy precisa para detectar heterocedasticidad.
- No es robusta a no-normalidad.
- Cuando se cree que existe normalidad y se desconoce la hipótesis alternativa, esta prueba es la más recomendable.
- Comparada con la prueba de Hartley y la de Cadwell, esta prueba es muy potente.
- Si se tienen poblaciones normales platicúrticas o mesocúrticas, es una prueba conservativa, sin embargo tiene buena potencia.
- En caso de poblaciones leptocúrticas, tiene buena potencia pero el riesgo del error tipo I es inflado. Esto se incrementa si  $n$  crece.
- Cuando las varianzas difieren, es mejor que la prueba de Levene ( $L-A$  y  $L-x^2$ ).
- En el caso de poblaciones normales platicúrticas o mesocúrticas, es mejor que la prueba de Layard ( $\chi^2$  o

Jackknife).

#### PRUEBA DE SCHEFFÉ.-

- Es robusta a la forma poblacional pero la potencia no es muy alta.
- En efectividad la prueba es similar a Jackknife, aunque por cuestión de cálculos, ésta es preferible.
- Cuando se utilizan submuestras de tamaño 3, la prueba provee una potencia alta.
- En cuestión de control de error tipo I es similar a la prueba de Levene( $W_{50}$ ).

#### PRUEBA DE HARTLEY.-

- Se necesitan tamaños de muestra iguales.
- Es mas potente que la prueba de Cadwell aunque la diferencia es muy pequeña.
- Es muy sensible a no-normalidad.
- Si existe normalidad y ya sea platicurtosis o mesocurtosis la prueba es conservativa.
- Si la población es normal leptocurtica, la probabilidad de error tipo I es inflado.

#### PRUEBA DE CADWELL.-

- Es muy sensible a no-normalidad.
- Es preferible a las pruebas Bartlett y Hartley si se quiere rapidéz en la detección.
- Los tamaños de muestra deben ser iguales.

#### PRUEBA DE LEVENE.-

En caso de dos muestras:

- La prueba Z es muy sensible a la forma poblacional.

- La prueba S es sensible a poblaciones con colas muy grandes.

En caso de varias muestras:

- $W_{50}$  es recomendable si la población es simétrica y en caso de que el tamaño de la muestra sea pequeño, la prueba es conservativa.
- $W_{10}$  es aparentemente mas robusta a la forma de la población.
- $W_0$  es la mas sensible, de las tres ( $W_0$ ,  $W_{50}$ ,  $W_{10}$ ), a la forma poblacional. En poblaciones simétricas, la probabilidad de error tipo I es bastante inflada.
- $W_{50}$  al igual que la prueba de Scheffé, presenta buen control de error tipo I, aunque la potencia es baja.

#### PRUEBA PARAMETRICA DE GOLDFELD Y QUANDT.-

- Cuando, se utilizan muestras de tamaño 30 o 60, la prueba obtiene máxima potencia cuando  $k$  toma el valor de 8 y 16 respectivamente.
- La potencia de la prueba aumenta si se incrementa la desviación estándar y/o incrementa el tamaño de la muestra.
- La prueba es menos potente que la de Residuales Recursivos.

#### PRUEBA DE MILLER.-

Dos muestras ( $k$  es el número de observaciones por grupo):

- Es mas potente que la prueba S de Levene si se tra -  
baja con  $k=1$  pero no funciona igual si  $k=1$  y se vuelve un enredo al calcularla.
- Si  $k=1$  la prueba es similar (en potencia) a la prueba S de Levene.

Varias muestras ( $k$  es el número de muestras):

- En poblaciones con colas muy grandes, la prueba da resultados muy inflados (comparados con los que deberían ser) cuando  $n_i \neq n_j$ .
- En poblaciones leptocurticas, la prueba tiende a inflar la probabilidad de error tipo I. Si  $n_i \neq n_j$  esta tendencia se incrementa.
- La prueba trabaja mejor con muestras grandes ( $n \geq 25$ )

#### PRUEBA DE GLEJSER.-

- La prueba detecta bien heterocedasticidad pura con muestras de tamaño mínimo igual a 30, sólo si la varianza de  $u$  crece y los valores de  $\chi_m$  son muy grandes.
- La prueba no funciona cuando la varianza de  $u$  crece y los valores de  $\chi_m$  decrecen.
- La prueba da resultados insatisfactorios cuando se trata de detectar heterocedasticidad mixta.

#### PRUEBA DE LAYARD.-

- La prueba es robusta a la forma poblacional y es bastante potente.
- Es preferible a la prueba Jackknife dado que es más fácil de calcular.
- La prueba, en general, se desvía del nivel de significancia nominal con muestras pequeñas.

#### PRUEBA DE HARVEY Y PHILLIPS.-

- La prueba es más potente que la de Goldfeld & Quandt y es más flexible en el ordenamiento.

PRUEBA DE ZROETER.-

- Esta prueba es demasiado sofisticada. Esta involucra una serie de cálculos lo cual la hace inaccesible.

## LISTA DE TABLAS

TABLA	Pág.
I. Percentiles de la distribución $\chi^2$ .	74
II. David(1952)-Hartley(1950). Puntos porcentua-	75
III. Leslie & Brown(1966).	76
IV. Percentiles de la distribución t-Student.	78
V. Distribución F.	79
VI. Valores críticos del cociente Von-Newmann.	81
VII. Estadística Durbin-Watson (d).	83

TABLA I. Percentiles de la distribución  $\chi^2$ . (1)

df	Per cent									
	0.5	1	2.5	5	10	90	95	97.5	99	99.5
1	0.000039	0.00016	0.00098	0.0039	0.0158	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.1026	0.2107	4.61	5.99	7.38	9.21	10.60
3	0.0717	0.115	0.216	0.352	0.584	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.15	1.61	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27
18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81	37.16
20	7.43	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57	40.00
24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
120	83.85	86.92	91.58	95.70	100.62	140.23	146.57	152.21	158.95	163.64

Para valores grandes de los grados de libertad la siguiente fórmula aproximada puede ser usada:

$$\chi^2_n = n \left( 1 - \frac{2}{9n} + z_a \sqrt{\frac{2}{9n}} \right)^3$$

donde  $z_a$  es la desviación normal y  $n$  es el número de grados de libertad. Por ejemplo  $\chi^2_{0.99} = 60 \left( 1 - \frac{2}{9 \cdot 60} + 2.326 \sqrt{\frac{2}{9 \cdot 60}} \right)^3 = 60 \left( 1.1379 \right)^3 = 88.4$  para el 99avo. percentil con 60 grados de libertad.

TABLA II. David(1952)-Hartley(1950). Puntos porcentuales superiores del cociente  $S_{\max}^2 / S_{\min}^2$  en un conjunto de  $k$  medias cuadradas, cada una basada sobre  $v$  grados de libertad. (Se supone una variación normal).

(a) Puntos al 5%

$v \backslash k$	2	3	4	5	6	7	8	9	10	11	12
2	39.0	87.5	142	202	266	333	403	475	550	626	704
3	15.4	27.8	39.2	50.7	62.0	72.9	83.5	93.9	104	114	124
4	9.60	15.5	20.6	25.2	29.5	33.6	37.5	41.1	44.6	48.0	51.4
5	7.15	10.8	13.7	16.3	18.7	20.8	22.9	24.7	26.6	28.2	29.9
6	5.82	8.38	10.4	12.1	13.7	15.0	16.3	17.5	18.6	19.7	20.7
7	4.99	6.94	8.44	9.70	10.8	11.8	12.7	13.5	14.3	15.1	15.8
8	4.43	6.00	7.18	8.12	9.03	9.78	10.5	11.1	11.7	12.2	12.7
9	4.03	5.34	6.31	7.11	7.80	8.41	8.95	9.45	9.91	10.3	10.7
10	3.72	4.85	5.67	6.34	6.92	7.42	7.87	8.28	8.66	9.01	9.34
12	3.28	4.16	4.79	5.30	5.72	6.09	6.42	6.72	7.00	7.25	7.48
15	2.86	3.54	4.01	4.37	4.68	4.95	5.19	5.40	5.59	5.77	5.93
20	2.48	2.95	3.29	3.54	3.76	3.94	4.10	4.24	4.37	4.49	4.59
30	2.07	2.40	2.61	2.78	2.91	3.02	3.12	3.21	3.29	3.36	3.39
60	1.67	1.85	1.96	2.04	2.11	2.17	2.22	2.26	2.30	2.33	2.36
$\infty$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

(b) Puntos al 1%

$v \backslash k$	2	3	4	5	6	7	8	9	10	11	12
2	199	448	729	1036	1362	1705	2063	2432	2813	3204	3605
3	47.5	85	120	151	184	21(6)	24(9)	28(1)	31(0)	33(7)	36(1)
4	23.2	37	49	59	69	79	89	97	106	113	120
5	14.9	22	28	33	38	42	46	50	54	57	60
6	11.1	15.5	19.1	22	25	27	30	32	34	36	37
7	8.89	12.1	14.5	16.5	18.4	20	22	23	24	26	27
8	7.50	9.9	11.7	13.2	14.5	15.8	16.9	17.9	18.9	19.8	21
9	6.54	8.5	9.9	11.1	12.1	13.1	13.9	14.7	15.3	16.0	16.6
10	5.85	7.4	8.6	9.6	10.4	11.1	11.8	12.4	12.9	13.4	13.9
12	4.91	6.1	6.9	7.6	8.2	8.7	9.1	9.5	9.9	10.2	10.6
15	4.07	4.9	5.5	6.0	6.4	6.7	7.1	7.3	7.5	7.8	8.0
20	3.32	3.8	4.3	4.6	4.9	5.1	5.3	5.5	5.6	5.8	5.9
30	2.63	3.0	3.3	3.4	3.6	3.7	3.8	3.9	4.0	4.1	4.2
60	1.96	2.2	2.3	2.4	2.4	2.5	2.5	2.6	2.6	2.7	2.7
$\infty$	1.00	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Los valores en la columna  $k=2$  y en las filas  $v=2$  e  $\infty$  son exactos. En todos los demás el tercer dígito puede tener un error de pocas unidades para los puntos al 5% y algunas más para los puntos al 1%. Las figuras del tercer dígito entre paréntesis para  $v=3$  son las más inseguras.

TABLA III. Leslie & Brown(1966). Puntos porcentuales del cociente  $w_{\max}/w_{\min}$ .

$n \backslash k$     2    3    4    5    6    7    8    9    10    11    12

Puntos superiores del 5%.

3	6.267	9.392	11.09	14.30	16.40	18.35	20.19	21.93	23.59	25.18	26.71
4	3.071	5.335	6.371	7.237	7.992	8.669	9.285	9.854	10.38	10.88	11.35
5	3.157	4.018	4.643	5.149	5.580	5.958	6.298	6.607	6.892	7.150	7.404
6	2.744	3.381	3.831	4.187	4.487	4.747	4.978	5.187	5.377	5.553	5.717
7	2.494	3.007	3.362	3.640	3.871	4.070	4.245	4.403	4.546	4.678	4.800
8	2.325	2.760	3.056	3.286	3.476	3.638	3.781	3.909	4.024	4.130	4.227
9	2.203	2.584	2.841	3.039	3.201	3.340	3.461	3.568	3.660	3.755	3.837
10	2.110	2.452	2.681	2.855	2.998	3.120	3.226	3.320	3.404	3.482	3.553
11	2.037	2.349	2.556	2.714	2.842	2.951	3.046	3.129	3.205	3.274	3.337
13	1.923	2.198	2.375	2.508	2.617	2.708	2.787	2.857	2.920	2.977	3.029
16	1.820	2.049	2.198	2.309	2.399	2.475	2.540	2.597	2.648	2.695	2.737
21	1.709	1.900	2.022	2.113	2.186	2.246	2.299	2.344	2.385	2.422	2.456
31	1.592	1.745	1.841	1.912	1.969	2.016	2.056	2.091	2.123	2.151	2.176
61	1.459	1.571	1.641	1.692	1.732	1.765	1.794	1.818	1.840	1.860	1.867
$\infty$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Puntos superiores del 2.5%

3	8.920	13.38	17.08	20.37	23.36	26.15	28.76	31.24	33.60	35.87	38.05
4	5.078	6.805	8.116	9.212	10.17	11.02	11.80	12.52	13.19	13.82	14.42
5	3.837	4.863	5.607	6.209	6.722	7.173	7.577	7.945	8.284	8.599	8.893
6	3.238	3.970	4.485	4.894	5.238	5.537	5.802	6.041	6.260	6.461	6.649
7	2.886	3.461	3.858	4.169	4.428	4.650	4.847	5.023	5.184	5.331	5.467
8	2.654	3.133	3.459	3.711	3.920	4.098	4.255	4.395	4.522	4.639	4.746
9	2.489	2.903	3.182	3.396	3.572	3.722	3.853	3.970	4.076	4.173	4.261
10	2.365	2.733	2.978	3.166	3.319	3.449	3.563	3.664	3.755	3.838	3.914
11	2.269	2.601	2.822	2.989	3.126	3.242	3.342	3.432	3.512	3.585	3.652
13	2.127	2.411	2.596	2.737	2.851	2.946	3.029	3.103	3.169	3.229	3.283
16	1.988	2.226	2.380	2.495	2.589	2.667	2.734	2.794	2.847	2.895	2.939
21	1.848	2.043	2.168	2.261	2.335	2.398	2.451	2.498	2.540	2.577	2.612
31	1.703	1.857	1.954	2.026	2.083	2.130	2.171	2.206	2.238	2.266	2.292
61	1.540	1.652	1.721	1.772	1.812	1.845	1.874	1.898	1.920	1.940	1.958
$\infty$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLA III. Leslie & Brown(1966). Puntos porcentuales del cociente  $w_{\max}/w_{\min}$  continuación.

$k/n$	2	3	4	5	6	7	8	9	10	11	12
Puntos superiores del 1%											
3	14.16	21.24	27.12	32.33	37.09	41.53	45.73	49.74	53.55	57.12	60.41
4	6.970	9.316	11.10	12.59	13.89	15.05	16.11	17.08	17.99	18.85	19.66
5	4.914	6.202	7.137	7.893	8.537	9.104	9.612	10.07	10.50	10.90	11.27
6	3.985	4.860	5.477	5.967	6.378	6.736	7.053	7.340	7.602	7.843	8.068
7	3.461	4.127	4.587	4.947	5.247	5.505	5.733	5.938	6.124	6.295	6.452
8	3.126	3.667	4.035	4.321	4.558	4.760	4.938	5.096	5.240	5.372	5.492
9	2.892	3.352	3.661	3.900	4.096	4.263	4.409	4.539	4.657	4.764	4.863
10	2.720	3.122	3.391	3.597	3.765	3.908	4.032	4.143	4.243	4.334	4.418
11	2.588	2.948	3.186	3.368	3.516	3.642	3.751	3.848	3.935	4.014	4.087
13	2.396	2.698	2.896	3.045	3.166	3.269	3.357	3.436	3.506	3.570	3.628
16	2.211	2.460	2.621	2.742	2.840	2.922	2.993	3.055	3.111	3.161	3.206
21	2.029	2.230	2.358	2.454	2.531	2.595	2.650	2.698	2.741	2.780	2.815
31	1.845	2.000	2.099	2.171	2.229	2.277	2.318	2.354	2.386	2.414	2.440
61	1.642	1.752	1.821	1.874	1.914	1.947	1.976	2.001	2.023	2.043	2.054
$\infty$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Puntos superiores del 0.5%											
3	20.05	30.09	38.43	45.82	52.56	58.82	64.70	70.28	75.60	80.71	85.02
4	8.825	11.78	14.03	15.91	17.65	19.01	20.35	21.68	22.73	23.81	24.83
5	5.897	7.427	8.537	9.436	10.20	10.88	11.48	12.03	12.54	13.01	13.46
6	4.636	5.638	6.345	6.907	7.379	7.789	8.153	8.482	8.783	9.060	9.319
7	3.947	4.691	5.205	5.608	5.943	6.232	6.488	6.717	6.925	7.117	7.294
8	3.516	4.109	4.514	4.828	5.088	5.310	5.506	5.681	5.839	5.984	6.118
9	3.220	3.717	4.053	4.311	4.523	4.705	4.863	5.005	5.133	5.249	5.357
10	3.005	3.435	3.723	3.944	4.124	4.278	4.412	4.531	4.638	4.736	4.826
11	2.840	3.223	3.476	3.660	3.827	3.961	4.077	4.180	4.273	4.358	4.436
13	2.600	2.922	3.129	3.286	3.414	3.521	3.614	3.697	3.771	3.838	3.900
16	2.383	2.640	2.807	2.932	3.033	3.118	3.192	3.257	3.315	3.367	3.415
21	2.167	2.371	2.502	2.600	2.678	2.744	2.800	2.850	2.894	2.934	2.971
31	1.951	2.107	2.206	2.279	2.337	2.386	2.427	2.463	2.496	2.525	2.551
61	1.719	1.830	1.899	1.950	1.990	2.024	2.053	2.078	2.101	2.121	2.140
$\infty$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLA IV. Percentiles de la distribución t-student.<sup>1</sup>

df	t <sub>0.00</sub>	t <sub>0.10</sub>	t <sub>0.20</sub>	t <sub>0.30</sub>	t <sub>0.40</sub>	t <sub>0.50</sub>	t <sub>0.60</sub>	t <sub>0.70</sub>
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552	2.878
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528	2.845
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518	2.831
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500	2.807
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492	2.797
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485	2.787
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467	2.763
29	0.256	0.530	0.854	1.311	1.699	2.045	2.462	2.756
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457	2.750
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704
60	0.254	0.527	0.848	1.296	1.671	2.000	2.390	2.660
120	0.254	0.526	0.845	1.289	1.658	1.980	2.358	2.617
∞	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576
df	-t <sub>0.40</sub>	-t <sub>0.30</sub>	-t <sub>0.20</sub>	-t <sub>0.10</sub>	-t <sub>0.05</sub>	-t <sub>0.025</sub>	-t <sub>0.01</sub>	-t <sub>0.005</sub>

Cuando la tabla es leída mediante las entradas inferiores, se debe anteponer un signo negativo a los valores tabulados.

1.- Parcialmente tomados de la tabla III de R. A. Fisher y Frank Yates, Statistical Tables for Biological, Agricultural and Medical Research, 3a. edición, Oliver & Boyd, Ltd., Edinburgh, 1948. Con permiso de los autores y editores.

Grados de libertad para el numerador.

Grados de libertad para el denominador.

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	$\infty$
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
$\infty$	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Las interpolaciones deberán llevarse a cabo usando los recí-  
procos de los grados de libertad.

1.- Reproducida de M. Merrington, C.M. Thompson, "Tables of  
percentage points of the inverted beta(F) distribution,"  
Biometrika, vol.33, p.73,1943. Con permiso de los autores y  
editores.

TABLA V. Distribución F, quantos superiores del 5% ( $F_{0.95}$ ).<sup>1</sup>

Grados de libertad para el numerador.

Grados de libertad para el denominador.

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	$\infty$
1	4,052	5,000	5,403	5,625	5,764	5,859	5,928	5,982	6,023	6,056	6,106	6,157	6,209	6,235	6,261	6,287	6,313	6,339	6,366
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.0	13.9	13.8	13.7	13.7	13.6	13.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.53	2.45	2.36	2.27	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
$\infty$	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Las interpolaciones deberán llevarse a cabo usando los recí-  
procos de los grados de libertad.

1.- Reproducida de M.Merrington, C.M. Thompson, "Tables of  
percentage points of the inverted beta(F) distribution,"  
Biometrika, vol.33, p.73, 1943. Con permiso de los autores y  
editores.

TABLA V.i. Distribución F, puntos superiores del 1%(F 0.99).<sup>1</sup>

TABLA VI. Valores críticos para el cociente de Von-Neumann.<sup>1</sup>

$$P\left(\frac{\delta^2}{s^2} < k\right) = \int_0^k w\left(\frac{\delta^2}{s^2}\right) d\left(\frac{\delta^2}{s^2}\right)$$

$k \backslash n$	4	5	6	7	8	9	10	11	12
0.25				0.00001	0.00001	0.00001	0.00001		
0.30				0.00007	0.00007	0.00005	0.00004	0.00002	0.00001
0.35			0.00006	0.00027	0.00021	0.00014	0.00009	0.00005	0.00003
0.40			0.00047	0.00065	0.00047	0.00031	0.00019	0.00012	0.00007
0.45			0.00126	0.00126	0.00088	0.00059	0.00038	0.00025	0.00016
0.50		0.00038	0.00246	0.00214	0.00150	0.00103	0.00069	0.00046	0.00031
0.55		0.00223	0.00409	0.00333	0.00237	0.00168	0.00116	0.00080	0.00055
0.60		0.00493	0.00615	0.00486	0.00355	0.00259	0.00185	0.00132	0.00094
0.65		0.00830	0.00865	0.00678	0.00511	0.00382	0.00282	0.00208	0.00152
0.70		0.01225	0.01161	0.00913	0.00710	0.00544	0.00414	0.00313	0.00235
0.75		0.01673	0.01505	0.01197	0.00958	0.00753	0.00587	0.00455	0.00351
0.80	0.00356	0.02171	0.01900	0.01534	0.01263	0.01015	0.00809	0.00642	0.00508
0.85	0.01302	0.02717	0.02348	0.01932	0.01631	0.01338	0.01089	0.00883	0.00714
0.90	0.02257	0.03310	0.02851	0.02403	0.02068	0.01729	0.01436	0.01188	0.00980
0.95	0.03223	0.03949	0.03412	0.02957	0.02579	0.02196	0.01858	0.01565	0.01316
1.00	0.04199	0.04634	0.04035	0.03598	0.03171	0.02745	0.02363	0.02025	0.01733
1.05	0.05186	0.05364	0.04728	0.04325	0.03849	0.03384	0.02959	0.02578	0.02241
1.10	0.06184	0.06140	0.05500	0.05137	0.04618	0.04120	0.03655	0.03232	0.02852
1.15	0.07194	0.06963	0.06361	0.06036	0.05482	0.04957	0.04458	0.03997	0.03577
1.20			0.07323	0.07020	0.06445	0.05901	0.05375	0.04882	0.04425
1.25						0.06956	0.06412	0.05894	0.05407
1.30								0.07040	0.06531

1.- Reproducida de "The Annals of Mathematical Statistics", Vol. 13, p.213, 1942, con el permiso de los editores.

TABLA VI. Valores críticos para el cociente de Von-Neumann-continuación.

$k \backslash n$	15	20	25	30	40	50	60
0.35	0.00001						
0.40	0.00002						
0.45	0.00004						
0.50	0.00009	0.00001					
0.55	0.00018	0.00002					
0.60	0.00033	0.00005	0.00001				
0.65	0.00059	0.00012	0.00002				
0.70	0.00100	0.00024	0.00005	0.00001			
0.75	0.00161	0.00044	0.00011	0.00003			
0.80	0.00250	0.00076	0.00023	0.00007	0.00001		
0.85	0.00375	0.00127	0.00044	0.00015	0.00002		
0.90	0.00547	0.00206	0.00079	0.00030	0.00004	0.00001	
0.95	0.00778	0.00323	0.00135	0.00057	0.00010	0.00002	
1.00	0.01079	0.00489	0.00222	0.00102	0.00022	0.00005	0.00001
1.05	0.01465	0.00720	0.00355	0.00176	0.00044	0.00012	0.00003
1.10	0.01950	0.01033	0.00550	0.00294	0.00085	0.00026	0.00008
1.15	0.02550	0.01448	0.00826	0.00474	0.00158	0.00054	0.00019
1.20	0.03280	0.01986	0.01208	0.00738	0.00280	0.00108	0.00043
1.25	0.04155	0.02670	0.01723	0.01117	0.00476	0.00206	0.00092
1.30	0.05189	0.03524	0.02402	0.01644	0.00780	0.00376	0.00185
1.35	0.06396	0.04571	0.03276	0.02357	0.01235	0.00656	0.00355
1.40	0.07787	0.05834	0.04379	0.03298	0.01892	0.01098	0.00649
1.45		0.07333	0.05743	0.04511	0.02810	0.01769	0.01133
1.50			0.07398	0.06038	0.04055	0.02750	0.01893
1.55				0.07920	0.05696	0.04131	0.03034
1.60					0.07797	0.06006	0.04675
1.65						0.08465	0.06942
1.70							0.09949

Valores de  $k$  para los cuales  $P\left(\frac{\sigma^2}{s^2} < k\right) = 0$

$n$	$k$	$n$	$k$
4	0.7811	15	0.0468
5	0.4775	20	0.0259
6	0.3215	25	0.0164
7	0.2311	30	0.0113
8	0.1740	40	0.0063
9	0.1357	50	0.0040
10	0.1088	60	0.0028
11	0.0891		
12	0.0743		

TABLA VII.<sup>1</sup> Estadística Durbin-Watson (d). Puntos de significancia de  $d_L$  y  $d_U$  : 1%.

n	k' = 1		k' = 2		k' = 3		k' = 4		k' = 5	
	$d_L$	$d_U$								
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = número de observaciones.

k' = número de variables explicativas.

1.- Reproducida de Biometrika, vol. 41, p. 175, 1951, con el permiso de los editores.

TABLA VII.i.<sup>1</sup> Estadística Durbin-Watson (d). Puntos de significancia de  $d_L$  y  $d_U$  : 5%.

n	k' = 1		k' = 2		k' = 3		k' = 4		k' = 5	
	$d_L$	$d_U$								
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = número de observaciones.

k' = número de variables explicativas.

1.- Reproducida de Biometrika, vol. 41, p. 173, 1951, con el permiso de los editores.

## BIBLIOGRAFIA

- Bartlett, M.S. and Kendall, D.G.(1946), The Statistical Analysis of Variance-Heterogeneity and Logarithmic Transformation, Supplement to The Journal of the Royal Statist. Soc., 8, No.1, 128-138.
- Bishop, D.J. and Nair, U.S.(1939), A note on Certain Methods of Testing for the Homogeneity of a Set of Estimate Variances, Journal of the Royal Statist.Soc., 6, No.1, 89-99.
- Box, G.E.P.(1953), Non-Normality and Tests on Variances, Biometrika, Vol.40, No. 3-4, 318-335.
- Brown, M.B. and Forsythe, A.B.(1974), Robust Tests for the Equality of Variance, Journal of American Statistical Association, Vol. 69, 364-67.
- Cadwell, J.H.(1953), Approximating to the Distributions of Measures of Dispersion by a Power of  $\chi^2$ , Biometrika, Vol. 40, 336-346.
- David, H.A.(1952), Upper 5% and 1% Points of the Maximum F-Ratio, Biometrika, Vol. 39, 422-424.
- Draper, N.R. and Hunter, W.G.(1969), Transformations: Some Examples Revisited, Technometrics, Vol. 11, No.1, 23-40.
- Eisenhart, C.(1947), The Assumptions Underlying the Analysis of Variance, Biometrics, Vol. 3, 1-21.
- (1964), The Meaning of Least in Least Squares, Journal of Washington Academic Sciences, Vol. 54, 24-33.
- Games, P.A, Winkler, H.B. and Probert, D.A.(1972), Robust Tests for Homogeneity of Variance, Educ. Psychol. Meas., Vol. 32, 887-909.
- Gartside, P.S.(1972), A Study of Methods for Comparing Several Variances, Journal of American Statistical Association, Vol.67, No.338, 342-346.

- Glejser, H.(1969), A New Test for Heteroscedasticity, Journal of American Statistical Association, Vol. 64, 316-23.
- Goldfeld, S.M. and Quandt, R.B.(1965), Some Tests for Homoscedasticity, Journal of American Statistical Association, Vol.60, 539-47.
- Graybill, F.A.(1961), An Introduction to Linear Statistical Models, Vol.I, McGraw-Hill.
- Guerrero, G.V.M. y Hernández, A.F.M.(1974), Uso de Transformaciones en los Modelos Lineales, Tesis de Licenciatura, Fac. Ciencias, UNAM.
- Harter, H.L.(1963), Percentage Points of the Ratio of Two Ranges and Power of the Associated Test, Biometrika, Vol.50, 187-94.
- Hartley, H.O.(1940), Testing the Homogeneity of a Set of Variances, Biometrika, Vol.31, 249-55.
- (1950), The Maximum F-Ratio as a Short-cut Test for Homogeneity of Variances, Biometrika, Vol.37, 308-312.
- Harvey, A.C. and Phillips, G.D.A.(1974), A Comparison of the Power of Some Tests for Heteroscedasticity in the General Linear Model, Journal of Econometrics, Vol. 2, 307-16.
- Hickman, E.P. and Long, C.S.(1980), A Note on Glejser's Test for Heteroscedasticity, Communications in Statistics. Theor. and Method., Vol 49, No.12, 1209-20.
- Johnston, J.(1963), Econometric Methods, McGraw-Hill Book Co., New York.
- Keselman, H.J., Games, P.A. and Cunch, J.J.(1979), Tests for Homogeneity of Variance, Commun. Statist.-Simula. Computa., ( B812 ), 113-129.
- Layard, M.W.J.(1973), Robust Large-Sample Tests for Homogeneity of Variances, Journal of American Statistical Association, Vol. 68, No. 341.

- Leslie, R.T. and Brown, B.M.(1966), Use of Range in Testing Heterogeneity of Variance, Biometrika, Vol. 53, No. 1-2, 221-27.
- Levene, H. (1960), Robust Tests for Equality of Variances, in I. Olkin, Ed. Contributions to Probability and Statistics, Palo Alto, California: Stanford University Press, 278-92.
- Lorda, A.R.A(1979), La Transformación de Variables, Tesis Profesional, Fac. Ciencias, UNAM.
- Miller, R.G.(1968), Jackknifing Variances, Annals of Mathematical Statistics, Vol. 29, No. 2, 567-82.
- Pearson, E.S.(1966), Alternative Tests for Homogeneity of Variance, some Monte-Carlo Results, Biometrika, Vol. 53, No. 1-2, 229-34.
- Quenouille, M.H.(1949), Approximate Tests of Correlation in Time-Series, Journal of The Royal Statistical Society, Serie B, Vol. 11, 68-84.
- Scheffé, H.(1959), The Analysis of Variance, Wiley:New York.
- Seber, G.A.F.(1977), Linear Regression Analysis, Wiley: New York.
- Tukey, J.W.(1958), Bias and Confidence in Not-Quite Large Samples, Annals of Mathematical Statistics, Vol. 29, 614.