



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
FACULTAD DE CIENCIAS

ANÁLISIS BAYESIANO DE LA SUPOSICIÓN DE  
NORMALIDAD EN EL MODELO DE REGRESIÓN LINEAL

TRABAJO DE TESIS QUE PARA OBTENER EL TÍTULO DE

ACTUARIO

PRESENTA

MARIA YOLANDA HERNÁNDEZ POLO

DIRECTOR DE TESIS :

M. EN C. RAUL RUEDA DÍAZ DEL CAMPO

MÉXICO, D.F.

NOVIEMBRE DE 1988



## **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 está 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.

# I N D I C E

	página
INTRODUCCION	1
1. CRITERIOS E INFERENCIAS ROBUSTOS	2
2. FAMILIA DE DISTRIBUCIONES EXPONENCIAL POTENCIA	9
3. DISTRIBUCIONES FINALES	13
3.1 DISTRIBUCIONES FINALES PARA EL CASO $\theta \in \mathbb{R}$	
3.1.1 $\beta$ CONOCIDO	
3.1.2 $\beta$ DESCONOCIDO	
3.2 DISTRIBUCIONES FINALES PARA EL MODELO DE REGRESION LINEAL	
4. IMPLEMENTACION	23
4.1 METODOS NUMERICOS EMPLEADOS	
4.2 PROGRAMAS GENERADOS	
5. EJEMPLOS	56
6. CONCLUSIONES GENERALES	98
APENDICE I	
APENDICE II	

## I. INTRODUCCION

Las inferencias estadísticas tienen dos bases fundamentales, una de ellas son las observaciones y la otra las suposiciones iniciales acerca del modelo generador de datos, ambas correspondientes a la población bajo estudio.

Cualquiera que sea el problema a estudiar, siempre debe considerarse la validez de las suposiciones iniciales, ya que, si se incurre en graves violaciones a las suposiciones asociadas al modelo con el que se pretende analizar un problema específico, puede producirse inestabilidad en las inferencias, es decir, para diferentes muestras de la misma población, pueden obtenerse conclusiones completamente diferentes.

El presente trabajo tiene como finalidad explorar, mediante el uso de la familia de distribuciones Exponencial Potencia (Box y Tiao, 1973), el efecto de la suposición de normalidad en las inferencias que se realicen acerca de los coeficientes de regresión en el modelo de Regresión Lineal, desde el punto de vista bayesiano.

En el capítulo 1 se incluye una breve discusión sobre los inconvenientes que pueden presentarse al asignar modelos de probabilidad que no son los adecuados al problema bajo estudio, con lo que surge el concepto de robustez, mismo que se analiza, mediante ejemplos, desde los enfoques clásico y bayesiano.

El capítulo 2 presenta a la familia de distribuciones Exponencial Potencia y algunas de sus propiedades que facilitan el análisis de la suposición de normalidad.

En el capítulo 3 se obtienen las distribuciones finales correspondientes a los coeficientes de regresión y también la distribución final para el parámetro de 'no normalidad'.

A fin de poder analizar el supuesto de normalidad, se elaboraron programas que realizan los cálculos numéricos necesarios. Estos, junto con los métodos numéricos empleados para tal fin, se presentan en el capítulo 4 y en el capítulo 5 se presentan algunos ejemplos ilustrativos. En el capítulo 6 se presentan las conclusiones generales del trabajo.

## 1. CRITERIOS E INFRENCIAS ROBUSTAS

Los métodos estadísticos son una herramienta útil para el desarrollo de la investigación científica. Estos métodos se basan en la construcción de modelos de probabilidad en base a la información que se obtiene de la observación de un fenómeno o suceso de interés.

Considérese un ejemplo tomado de Box y Tiao (1973) acerca de la vida útil de baterías producidas por una máquina. En condiciones normales puede suponerse que las vidas observadas de las baterías forman una muestra aleatoria de una distribución Normal con media  $\theta$  y varianza  $\sigma^2$ . La verosimilitud de una muestra de  $n$  observaciones  $\underline{Y} = y_1, \dots, y_n$  es entonces

$$P(\underline{Y}|\theta, \sigma^2) \propto \sigma^{-n} \exp \left[ -(2\sigma^2)^{-1} \sum_{i=1}^n (y_i - \theta)^2 \right], \quad y_i \in \mathbb{R}$$

sin embargo, puede haber situaciones en las cuales este modelo resulte inadecuado; puede suceder, por ejemplo, que durante el período de observación, una característica cualitativa  $x$  de un aditivo químico usado en la fabricación de las baterías se modifique y cause, vía una relación lineal, un cambio correspondiente en el tiempo medio de vida de las baterías. En este caso, un modelo más apropiado podría ser

$$P(\underline{Y}|x, \sigma^2, \theta_1, \theta_2) \propto \sigma^{-n} \exp \left[ -(2\sigma^2)^{-1} \sum_{i=1}^n (y_i - \theta_1 - \theta_2 x_i)^2 \right], \quad x_i, y_i \in \mathbb{R}$$

Alternativamente, pueden ocurrir otras situaciones en las que la primera batería de una corrida de producción fuera siempre defectuosa, o que las observaciones no se distribuyan independientemente pero sigan alguna serie de tiempo, o que su distribución difiera mucho de la Normal, etc.

En base a la naturaleza del problema, se forman las suposiciones básicas que conducen al modelo que se va a aplicar, y las conclusiones que se deriven del estudio se formularán en base a ese modelo. Es importante analizar cuan sensitivas son las conclusiones a la selección del modelo particular, esto es, es conveniente determinar en que medida las conclusiones generadas dependen de las suposiciones

iniciales sobre los datos. A este diagnóstico se le conoce con el nombre de análisis de robustez.

Se pueden distinguir dos tipos de robustez : robustez de criterio y robustez de inferencia.

Para mostrar la diferencia entre robustez de criterio y robustez de inferencia, se tomó el ejemplo de Box y Tiao (1973) referente al experimento de Darwin. El experimento consiste en analizar las diferencias en las alturas de diversos tipos de plantas, fertilizadas entre sí y en forma cruzada. Los datos del experimento son las alturas de 15 pares de plantas, cada par contiene una planta híbrida y otra no híbrida desarrolladas en la misma maceta y de la misma semilla. Se toman como observaciones las 15 diferencias de alturas correspondientes a los respectivos pares, y se encuentran en la siguiente tabla :

DIFERENCIAS DE ALTURAS (BOX Y TIAO, 1973)

49	23	24
-67	28	75
8	41	60
16	14	-48
6	56	29

Se supone que las diferencias son una muestra aleatoria de una población Normal  $N(\theta, \sigma^2)$  y que se requiere hacer inferencias sobre el parámetro  $\theta$ , probando la siguiente hipótesis

$$H_0: \theta = 0 \quad \text{vs} \quad H_a: \theta > 0$$

En la teoría muestral clásica, la hipótesis se prueba usando la estadística  $t$ . Bajo  $H_0$

$$t = \frac{n^{1/2} (\bar{Y} - \theta)}{\hat{\sigma}} \sim t_{(n-1)}$$

Si  $t_{(n-1)}^{(1-\alpha)}$  representa al cuantil  $1-\alpha$  de una distribución  $t$  de Student con  $n-1$  grados de libertad, usando este criterio, no se rechaza  $H_0$  con un nivel de significancia  $\alpha$  cuando  $t < t_{(n-1)}^{(1-\alpha)}$ . En

este caso  $n = 15$  y

$$\bar{Y} = 20.933 \quad \hat{\sigma}/\sqrt{n} = 9.746$$

$t = 2.1478$  y el nivel de significancia descriptivo es  $\alpha = 0.02485$ .

Box y Tiao supusieron, por otro lado, que la distribución de las observaciones era uniforme sobre algún rango desconocido  $\theta - \sqrt{3}\sigma$  a  $\theta + \sqrt{3}\sigma$  para mostrar el efecto de este grado extremo de no normalidad sobre el criterio  $t$ .

Mencionan que la distribución de la estadística  $t$  puede calcularse aproximadamente usando, por ejemplo, los trabajos de Geary(1936), Gayen(1949,1950), o de Box y Andersen(1955). Estos últimos autores, afirman que cuando la distribución de la población no es Normal, la distribución nula de  $t^2$  es aproximadamente una distribución F con  $\delta$  y  $\delta(n-1)$  grados de libertad, donde

$$\delta = 1 + E(b-3)/n \quad b = (n+2) \sum y_i^4 / [\sum y_i^2]^2$$

y

$$E(b-3) \approx \gamma_2 - n^{-1}(2\gamma_4 - 3\gamma_2^2 + 11\gamma_2) + n^{-2}(3\gamma_6 - 16\gamma_4\gamma_2 + 15\gamma_2^3 + 38\gamma_4 - 3\gamma_2^2 + 86\gamma_2)$$

donde

$$\gamma_{r-2} = K_r / K_2^{r/2} \quad r = 3, 4, \dots$$

y  $K_r$ 's los momentos de la distribución de las diferencias.

En el presente ejemplo  $\delta = 0.913$  por lo que  $t^2 \sim F_{(12, 78)}$  y el nivel de significancia descriptivo es  $\alpha = 0.02388$ .

Al comparar los niveles de significancia obtenidos, se observa que el criterio  $t$  es poco afectado al cambiar a la distribución Uniforme, por lo que hay indicios de que el criterio  $t$  es robusto bajo este caso extremo de no normalidad.

En general, un criterio es robusto, si los resultados obtenidos no cambian 'mucho' al modificar la forma de la distribución de probabilidad asumida.

Bajo el supuesto de que la distribución generadora de las observaciones es rectangular, Box y Tiao argumentan, que el criterio apropiado para realizar la prueba sobre  $\theta$  es

$$W = (n-1)(m-\theta) / h$$

donde

$$m = (y_{(n)} + y_{(1)}) / 2 \quad (\text{punto medio})$$

$$h = (y_{(n)} - y_{(1)}) / 2 \quad (\text{rango medio})$$

y  $|W| \sim F_{(2,2(n-1))}$  (Neyman y Pearson, 1928 ; Carlton, 1946).

Para el ejemplo se obtuvo  $m=4.0$ ,  $h=71$  , bajo  $H_0$   $W = 0.789$ . El nivel de significancia descriptivo es  $\alpha = 0.232$ .

Comparando los niveles de significancia descriptivos obtenidos al suponer normalidad y al suponer uniformidad, utilizando el criterio apropiado en cada caso, con un nivel de significancia  $\alpha' = 0.05$  , se observa que la inferencia obtenida en uno y otro es muy diferente, ya que, con la suposición de uniformidad no se rechaza  $H_0$ , mientras que suponiendo normalidad se rechaza, por lo que no hay robustez de inferencia aún cuando el criterio  $t$  es robusto. El problema surge entonces, cuando se tiene duda acerca de la distribución de las observaciones, ya que, dependiendo de ésta va a existir un criterio óptimo para producir inferencias acerca de los parámetros, y estas inferencias pueden cambiar aún cuando los criterios utilizados sean robustos.

La forma de analizar la robustez en el enfoque Bayesiano es un poco diferente de lo que se ha discutido. En el siguiente ejemplo se muestra la diferencia en la forma de análisis de los dos enfoques mencionados; el ejemplo fué tomado de Box y Tiao (1964).

Se desea comparar la dispersión en las observaciones realizadas por dos analistas, de una mezcla de carbón con polvo.

El analista A1 realizó 20 observaciones, mientras que el analista A2 efectuó 13. Las observaciones fueron obtenidas en forma independiente, y los resultados se muestran en la siguiente tabla

Analista A1	Analista A2
$X_1$	$X_2$
-10	- 8
16	- 3
- 8	20
9	22
5	3
- 5	5
5	10
11	14
25	-21
22	2
16	7

Analista A1

Analista A2

$X_1$	$X_2$
3	8
40	16
0	
-5	
16	
30	
14	
25	
28	

donde  $X_i = (\% \text{ de carbón} - 4.5) \cdot 100$   $i = 1, 2$ .

Se supone que ambas muestras provienen de variables aleatorias cuyas distribuciones son miembros de la clase de distribuciones Exponencial Potencia definida como

$$P(y|\theta, \sigma, \beta) = k \exp\left\{(-1/2) |(Y-\theta)/\sigma|^{2/(1+\beta)}\right\}$$

con

$$k = \left[ \Gamma[1+(1+\beta)/2] 2^{1+(1+\beta)/2} \sigma \right]^{-1} \quad Y, \theta \in \mathbb{R}, \sigma \in \mathbb{R}^+, \beta \in (-1, 1]$$

Se supone también que ambas distribuciones tienen un valor común de  $\beta$ , con  $\theta_1$  y  $\theta_2$  conocidas. Se encontró que para cualquier valor de  $\beta$  fijo,  $\beta = \beta_0$ , la prueba Uniformemente más Potente para verificar la hipótesis

$$H_0: \sigma_1^2 / \sigma_2^2 = 1 \quad \text{vs} \quad H_a: \sigma_1^2 / \sigma_2^2 > 1$$

se basa en el cociente

$$F(\beta_0) = \frac{n_2 \sum |y_{1i} - \theta_1|^{2/(1+\beta_0)}}{n_1 \sum |y_{2j} - \theta_2|^{2/(1+\beta_0)}}$$

en el cual, el numerador y el denominador son estadísticas suficientes para  $\sigma_1$  y  $\sigma_2$  respectivamente. En particular, cuando  $\beta = \beta_0$

$$F(\beta_0) \sim F_{(n_1(1+\beta_0), n_2(1+\beta_0))}, \text{ bajo } H_0.$$

Calcularon las probabilidades que se encuentran en la tabla 2.1 utilizando estadísticas equivalentes a la F y para los casos en que  $\beta = \beta_0$ , utilizaron una aproximación, obtenida por medio de la teoría de Permutaciones

Tabla 2.1 (Box y Tiao, 1964)

$\beta_0 \backslash \beta$	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6
-0.6	0.060	0.110	-----	-----	-----	-----	-----
-0.4	0.035	0.065	0.090	0.140	-----	-----	-----
-0.2	0.030	0.048	0.070	0.098	0.115	0.170	0.220
0.0	0.028	0.045	0.060	0.076	0.095	0.120	0.145
0.2	0.025	0.040	0.050	0.060	0.080	0.100	0.120
0.4	0.024	0.035	0.048	0.060	0.080	0.086	0.095
0.6	0.024	0.040	0.048	0.050	0.065	0.075	0.092

donde  $\beta$  es valor del parámetro en la distribución, y  $\beta_0$  el valor de  $\beta$  utilizado en la estadística de prueba. Las probabilidades reportadas corresponden a los niveles de significancia para varios valores de  $\beta$  en la distribución. Utilizaron el criterio Uniformemente más Potente para las distribuciones que cumplen con  $\beta = \beta_0$ , esto es, los elementos de la diagonal son los niveles de significancia para cada valor de  $\beta$  aplicando el criterio apropiado a cada distribución; en este caso, se observa que aún para valores extremos de  $\beta$  los valores no cambian mucho, por lo que se puede afirmar que existe robustez de inferencia. Ahora, si se fija un renglón, se tienen los niveles de significancia para un solo valor de  $\beta_0$ , esto es, aplicando un mismo criterio; por tanto, con los cambios en el nivel de significancia en cada renglón se observa si existe o no robustez de criterio. Puede observarse que los cambios ocurridos a lo largo de la diagonal son menores que aquéllos que ocurren horizontalmente, esto es, la inferencia apropiada acerca del cociente de varianzas es poco afectada por cambios en  $\beta$  en comparación con las probabilidades asociadas con un criterio particular.

Box y Tiao hacen referencia a un trabajo del mismo año, en el cual demuestran, desde un punto de vista bayesiano, que los elementos de la diagonal son las probabilidades finales de que  $\sigma_1^2 / \sigma_2^2 > 1$  para los correspondientes valores de  $\beta_0$ , cuando  $\log \sigma_1$  y  $\log \sigma_2$  se toman inicialmente con distribución localmente Uniforme.

En el enfoque Bayesiano, la robustez se determina como sigue : tómese como referencia la tabla 2.1, y considérese como distribuciones

iniciales las obtenidas al aplicar el modelo con  $\beta = \beta_0$ . Fijando un valor de  $\beta$  cualquiera, por ejemplo  $\beta = 0$ , y analizando el cambio en la distribución cuando  $\beta_0$  cambia se obtiene la robustez de inferencia. Como los cambios en las probabilidades son visibles, puede decirse que no hay robustez bajo no normalidad. Cabe aclarar que en este enfoque no existe robustez de criterio.

Por lo anterior, se tiene entonces que las inferencias dependen en gran medida del modelo inicial asignado a las observaciones cuando no se tiene robustez de inferencia, lo que se vuelve un problema, ya que en general no se puede garantizar la calidad del modelo.

El análisis se enfoca a las inferencias que se realizan acerca de los parámetros de localización en el modelo de Regresión Lineal, específicamente, es de interés analizar si estas inferencias son robustas bajo no normalidad en el contexto bayesiano, lo cual es de suma importancia si consideramos la discusión realizada en este capítulo, y en cualquier caso tener bases para asignar una distribución a las observaciones.

## 2. FAMILIA DE DISTRIBUCIONES EXPONENCIAL POTENCIA.

El modelo de Regresión Lineal es empleado frecuentemente en problemas de ajuste de curvas, por lo que es importante verificar a través de la información muestral la veracidad de las suposiciones del modelo, entre las que se encuentra la de normalidad de la variable de respuesta. Esta suposición es la que interesa analizar en el presente trabajo.

Como se sabe, la suposición de normalidad se puede justificar cuando las condiciones experimentales son tales que producen una tendencia que sugiera la aplicabilidad del teorema central de límite; sin embargo, aún en estas condiciones puede ocurrir que la distribución de las observaciones no sea precisamente Normal.

Sería más conveniente entonces, suponer que la distribución proviene de una clase de distribuciones simétricas y unimodales que incluye a la Normal. Diananda (1949), Box (1953) y Turner (1960) estudiaron una familia paramétrica de distribuciones que simplifica el problema de normalidad, estudiando el efecto de uno de sus parámetros sobre la forma de la distribución. Esta familia es la Exponencial Potencia (Box y Tiao, 1973), que surge de manera natural a partir de la distribución Normal.

La distribución Normal estándar puede escribirse como

$$P(x) = k \exp [-(1/2) |x|^q] \quad \text{con } q = 2$$

Permitiendo a  $q$  tomar valores diferentes de 2 se obtiene la **Familia de Distribuciones Exponencial Potencia**. Con  $q = 2/(1+\beta)$  y partiendo de la Normal no estándar, puede escribirse la forma general como

$$P(y|\theta, \phi, \beta) = k \phi^{-1} \exp \left\{ -(1/2) |(y-\theta)/\phi|^{2/(1+\beta)} \right\} \quad (2.1)$$

con  $\phi \in \mathbb{R}^+$ ,  $\beta \in (-1, 1]$ ;  $\theta, y \in \mathbb{R}$

donde

$$k^{-1} = \Gamma[1+(1+\beta)/2] 2^{(1+\beta)/2}$$

y los parámetros  $\theta$  y  $\phi$  corresponden a las medidas de localización y escala respectivamente

$$E(y) = \theta$$

$$\text{Var}(y) = \sigma^2 = 2^{(1+\beta)} \left[ \frac{\Gamma[(3/2)(1+\beta)]}{\Gamma[(1/2)(1+\beta)]} \right] \phi^2$$

esto puede comprobarse en el Apéndice I.

Utilizando esta última expresión, se puede reescribir la distribución como

$$P(y|\theta, \sigma, \beta) = w(\beta) \sigma^{-1} \exp \left[ -c(\beta) |(y - \theta)/\sigma|^{2/(1+\beta)} \right] \quad (2.2)$$

donde

$$c(\beta) = \left[ \frac{\Gamma[1.5(1+\beta)]}{\Gamma[0.5(1+\beta)]} \right]^{1/(1+\beta)}, \quad \beta \in (-1, 1), \quad \sigma \in \mathbb{R}^+$$

$$w(\beta) = \frac{\{\Gamma[1.5(1+\beta)]\}^{1/2}}{(1+\beta)\{\Gamma[0.5(1+\beta)]\}^{3/2}} \quad y, \theta \in \mathbb{R}$$

Es posible demostrar que cuando  $\beta \rightarrow -1$ , la familia de distribuciones tiende a la familia Rectangular, i.e.

$$\lim_{\beta \rightarrow -1} P(y|\beta, \theta, \sigma) = 1/(2\sqrt{3}\sigma) \quad \text{con } y \in (\theta - \sqrt{3}\sigma, \theta + \sqrt{3}\sigma)$$

cuando  $\beta = 1$  se tiene a la familia de distribuciones Doble Exponencial

$$P(y|\theta, \sigma) = \exp[-\sqrt{2}|(y-\theta)/\sigma|]/\sqrt{2}\sigma \quad y \in \mathbb{R}$$

y cuando  $\beta = 0$  se tiene a la familia de distribuciones Normales

$$P(y|\theta, \sigma) = \exp\{-0.5[(y-\theta)/\sigma]^2\}/\sqrt{2\pi}\sigma \quad y \in \mathbb{R}$$

El parámetro  $\beta$  juega entonces, un papel muy importante, puesto que, al fijarlo en cualquiera de los valores que puede tomar, se obtiene toda una familia de distribuciones. En especial, es de gran importancia el hecho de que se obtenga a la familia de distribuciones Normal al fijar  $\beta = 0$ .

Se sabe por otro lado, que existen medidas que describen algunos aspectos de las funciones de densidad, como es la asimetría, cuya medida es proporcionada por la expresión

$$\gamma_1 = E(y-\theta)^3 / \sigma^3$$

donde  $\theta$  es el valor esperado de  $y$  ( $E(y) = \theta$ ).

Si  $\gamma_1 < 0$  la distribución es asimétrica negativa (figura 1.1); si  $\gamma_1 = 0$  se dice que la distribución es simétrica (figura 1.2) y

finalmente si  $\gamma_1 > 0$ , la distribución es asimétrica positiva (figura 1.3).



$\gamma_1 < 0$   
FIGURA 1.1



$\gamma_1 = 0$   
FIGURA 1.2



$\gamma_1 > 0$   
FIGURA 1.3

Para la familia de distribuciones Exponencial Potencia,  $\gamma_1 = 0$  (Apéndice I), por lo que la familia de distribuciones es simétrica. La curtosis, mide el grado de acumulación de la densidad, es decir, que tan 'extendida' o 'picuda' es la curva. Su expresión está dada por

$$\gamma_2 = (E(y-\theta)^2 / \sigma^4) - 3$$

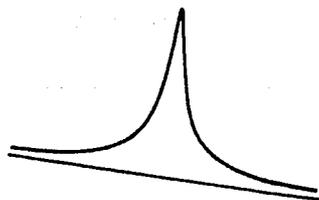
donde  $\theta$  es la media y  $\sigma^2$  la varianza de la distribución. Si  $\gamma_2 < 0$  se dice que la distribución es platicúrtica (figura 2.1); para la familia Normal,  $\gamma_2 = 0$  (figura 2.2) y si  $\gamma_2 > 0$ , se dice que la distribución es leptocúrtica (figura 2.3).



$\gamma_2 = 0$   
FIGURA 2.1



$\gamma_2 < 0$   
FIGURA 2.2



$\gamma_2 > 0$   
FIGURA 2.3

La expresión de  $\gamma_2$  para la familia de distribuciones Exponencial Potencia es (Apéndice I)

$$\gamma_2 = \frac{\Gamma[(1/2)(1+\beta)] \Gamma[(5/2)(1+\beta)]}{\{\Gamma[(3/2)(1+\beta)]\}^2} - 3$$

La expresión anterior involucra únicamente al parámetro  $\beta$ , de aquí, que a dicho parámetro se le considere como una medida de curtosis de la distribución; cuando  $\beta=0$ ,  $\gamma_2=0$ ; cuando  $\beta<0$ ,  $\gamma_2<0$ ; y cuando  $\beta>0$ ,  $\gamma_2>0$ ; por lo que  $\beta$  mide el grado de no normalidad de la distribución, y puesto que cuando  $\beta = 0$  se tiene curtosis 0, la familia de distribuciones Normal ocupa el lugar central de la familia de distribuciones Exponencial Potencia.

### 3. DISTRIBUCIONES FINALES

En el enfoque Bayesiano de la Estadística, es necesario obtener las distribuciones finales de los parámetros de interés, para realizar inferencias acerca de la población bajo estudio. En el presente trabajo, es de gran interés obtener las distribuciones finales correspondientes a los coeficientes del modelo de Regresión Lineal, esto con el fin de analizar el efecto de la suposición de normalidad de la variable de respuesta, sobre las inferencias que se realicen para estos coeficientes en el Análisis de Regresión, y también es de interés obtener la distribución final de  $\beta$ , para analizar la información que contienen las observaciones sobre la misma.

Por simplicidad, se obtendrán primeramente las distribuciones finales para el caso  $\theta \in \mathbb{R}$ , esto es de una muestra aleatoria de observaciones con media  $\theta$  y varianza  $\sigma^2$ , y posteriormente, se generalizarán los resultados al caso del modelo de Regresión Lineal.

#### 3.1 DISTRIBUCIONES FINALES PARA EL CASO $\theta \in \mathbb{R}$ .

Considérese una muestra de  $n$  observaciones independientes e idénticamente distribuidas de un miembro de la familia de distribuciones Exponencial Potencia,  $y_1, \dots, y_n$ ; para la cual, la función de verosimilitud está dada por

$$P(\underline{Y}|\theta, \sigma, \beta) = \prod_{i=1}^n P(y_i|\theta, \sigma, \beta) \quad , \quad \theta, y_i \in \mathbb{R} ; \sigma \in \mathbb{R}^+ ; \beta \in (-1, 1]$$

entonces

$$P(\underline{Y}|\theta, \sigma, \beta) = w(\beta)^n \sigma^{-n} \exp \left[ -c(\beta) \sum_{i=1}^n |(y_i - \theta)/\sigma|^{2/(1+\beta)} \right] \quad (3.1.1)$$

con  $w(\beta)$  y  $c(\beta)$  definidas en (2.2).

Para facilitar el desarrollo posterior de las distribuciones finales se considerará primero el caso para  $\beta$  fijo y conocido, y así analizar el comportamiento de dichas distribuciones en este caso particular.

##### 3.1.1 $\beta$ CONOCIDO.

La familia de distribuciones Exponencial Potencia pertenece a la

familia de distribuciones de Localización - Escala. Una distribución inicial de referencia para  $\theta$  y  $\sigma$  con  $\beta$  conocido es la obtenida al aplicar la regla de Jeffreys a la familia de Localización - Escala (Box y Tiao, 1973)

$$P(\theta, \sigma) \propto \sigma^{-1} \quad (3.1.1.1)$$

Utilizando esta distribución inicial de referencia, la distribución final de  $\theta$  y  $\sigma$  es

$$P(\theta, \sigma | \underline{Y}) \propto P(\underline{Y} | \theta, \sigma) P(\theta, \sigma)$$

la función de verosimilitud (para  $\beta$  conocido) es

$$P(\underline{Y} | \theta, \sigma) \propto \sigma^{-n} \exp \left[ -c(\beta) \sum_{i=1}^n |(y_i - \theta) / \sigma|^{2/(1+\beta)} \right] \quad (3.1.1.2)$$

de (3.1.1.1) y (3.1.1.2) se tiene que

$$P(\theta, \sigma | \underline{Y}) \propto \sigma^{-(n+1)} \exp \left[ -c(\beta) \sum_{i=1}^n |(y_i - \theta) / \sigma|^{2/(1+\beta)} \right] \quad (3.1.1.3)$$

Para obtener la distribución final de  $\theta$ , se integra (3.1.1.3) con respecto a  $\sigma$  y se divide por la constante de proporcionalidad, esto es

$$P(\theta | \underline{Y}) = \frac{\int P(y | \theta, \sigma) P(\theta, \sigma) d\sigma}{\iint P(y | \theta, \sigma) P(\theta, \sigma) d\theta d\sigma}$$

Así, se tiene que

$$\int_0^{\infty} P(\theta, \sigma | \underline{Y}) d\sigma \propto \int_0^{\infty} \sigma^{-(n+1)} \exp \left[ -c(\beta) \sum_{i=1}^n |y_i - \theta|^{2/(1+\beta)} \sigma^{-2/(1+\beta)} \right] d\sigma$$

En el Apéndice (A 2.1.6) de Box y Tiao (1973) se da la siguiente expresión para  $a > 0$ ,  $p > 0$  y  $\alpha > 0$

$$\int_0^{\infty} x^{-(p+1)} \exp\{-ax^{-\alpha}\} dx = (1/\alpha) a^{-(p/\alpha)} \Gamma(p/\alpha) \quad (3.1.1.4)$$

en este caso, sean  $p = n$ ,  $a = c(\beta) \sum_{i=1}^n |y_i - \theta|^{2/(1+\beta)}$ ,  $\alpha = 2/(1+\beta)$ , entonces

$$\int_0^{\infty} P(\theta, \sigma | \underline{Y}) d\sigma \propto (1/\alpha) a^{-(p/\alpha)} \Gamma(p/\alpha)$$

$$\propto (1+\beta)/2 \{c(\beta)M(\theta)\}^{-n(1+\beta)/2} \Gamma[(n/2)(1+\beta)] \quad (3.1.1.5)$$

donde 
$$M(\theta) = \sum_{i=1}^n |y_i - \theta|^{2/(1+\beta)} \quad (3.1.1.6)$$

Por otro lado

$$\int_0^{\infty} \int_{-\infty}^{\infty} P(\theta, \sigma | \underline{Y}) d\theta d\sigma = \int_{-\infty}^{\infty} \int_0^{\infty} P(\theta, \sigma | \underline{Y}) d\sigma d\theta$$

empleando (3.1.1.5) se obtiene

$$\int_{-\infty}^{\infty} \int_0^{\infty} P(\theta, \sigma | \underline{Y}) d\sigma d\theta = \left[ (1-\beta)/2 [c(\beta)]^{-n(1+\beta)/2} \Gamma[(n/2)(1+\beta)] \int_{-\infty}^{\infty} [M(\theta)]^{-n(1+\beta)/2} d\theta \right]$$

entonces la distribución final de referencia para  $\theta$  usando la distribución inicial (3.1.1.1), es

$$P(\theta | \underline{Y}) = \frac{[M(\theta)]^{-n(1+\beta)/2}}{\int_{-\infty}^{\infty} [M(\theta)]^{-n(1+\beta)/2} d\theta} = [M(\theta)]^{-n(1+\beta)/2} J(\beta)^{-1} \quad (3.1.1.7)$$

donde 
$$J(\beta) = \int_{-\infty}^{\infty} [M(\theta)]^{-n(1+\beta)/2} d\theta, \quad \beta \text{ fijo} \quad (3.1.1.8)$$

puesto que  $\beta$  es conocido, se puede analizar lo que pasa con  $P(\theta | \underline{Y})$  para valores extremos de  $\beta$  o para  $\beta=0$ ; esta distribución es una función monótona de  $\theta$ . Utilizando este hecho, pueden derivarse algunas de las propiedades de la distribución final de  $\theta$

i)  $P(\theta | \underline{Y})$  es continua, y para  $\beta \in (-1, 1)$  es diferenciable y unimodal, aunque no necesariamente simétrica; la moda se localiza en el intervalo  $[y_{(n)}, y_{(n)}]$ . El valor modal es el estimador máximo verosímil de  $\theta$  (ver Apéndice II).

ii)  $P(\theta | \underline{Y})$  cuando  $\beta = 0$ .

En este caso

$$M(\theta) = \sum_{i=1}^n (y_i - \theta)^2 = \sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - \theta)^2 = (n-1) S^2 + n(\bar{y} - \theta)^2$$

entonces

$$P(\theta|Y) \propto [M(\theta)]^{-n/2} = [(n-1)S^2 + S^2[\sqrt{n}(\bar{y}-\theta)/S]^2]^{-n/2}$$

sea  $u = \sqrt{n}(\bar{y} - \theta)/S$ , entonces

$$P(\theta|Y) \propto [(n-1)S^2 + S^2u^2]^{-n/2} \propto [1 + u^2/(n-1)]^{-n/2} \sim \text{STU}(\theta|n-1, 0, 1)$$

donde  $\text{STU}(x|n, \mu, \nu)$  denota a una distribución Student con  $n$  grados de libertad, parámetro de localización  $\mu$  y parámetro de escala  $\nu$

y

$$J(\beta) \propto \int_{-\infty}^{\infty} [M(\theta)]^{-n/2} d\theta = \int_{-\infty}^{\infty} \text{STU}(\theta|n-1, 0, 1) d\theta = 1$$

por tanto

$$t = \sqrt{n}(\bar{y}-\theta)/S \sim t_{(n-1)} \quad \text{i.e. } P(\theta|Y) = \text{STU}(\theta|n-1, \bar{y}, S^2/n)$$

esto es, cuando  $\beta = 0$ , la distribución final de referencia de  $\theta$  es una distribución  $t$ .

iii)  $P(\theta|Y)$  cuando  $\beta \rightarrow -1$ .

En el Apéndice II se demuestra que

$$\lim_{q \rightarrow \infty} [M(\theta)]^{-1/q} = h + |m - \theta|$$

donde

$$m = (1/2)(y_{(1)} + y_{(n)}) \quad \text{es el punto medio}$$

y

$$h = (1/2)(y_{(n)} - y_{(1)}) \quad \text{es el rango medio}$$

usando este resultado se tiene entonces

$$\lim_{\rho \rightarrow -1} [M(\theta)]^{(\rho+1)/2} = h + |m - \theta|$$

cuando  $\beta \rightarrow -1$

$$J(\beta) = \int_{-\infty}^{\infty} (h + |m - \theta|)^{-n} d\theta$$

pero

$$|m-\theta| = \begin{cases} m-\theta & \text{si } m-\theta > 0, \theta < m \\ -(m-\theta) & \text{si } m-\theta < 0, \theta > m \end{cases}$$

entonces

$$\begin{aligned} J(\beta) &= \int_{-\infty}^m [h + (m-\theta)]^{-n} d\theta + \int_m^{\infty} [h - (m-\theta)]^{-n} d\theta \\ &= 2 [h/(n-1)] [1/h^n] \end{aligned}$$

entonces

$$[J(\beta)]^{-1} = (1/2) [h/(n-1)]^{-1} [h^n]$$

y por tanto

$$\begin{aligned} \lim_{\beta \rightarrow -1} P(\theta|Y) &= \lim_{\beta \rightarrow -1} [J(\beta)]^{-1} [M(\theta)]^{(1+\beta)/2} \\ &= (1/2) [h/(n-1)]^{-1} \left\{ 1 + [1/(n-1)] |(n-1)(\theta-m)/h| \right\}^{-n} \end{aligned}$$

sea  $w = (n-1)(m-\theta)/h$ , entonces

$$\lim_{\beta \rightarrow -1} P(\theta|Y) \propto [(n-1)/2h] \left\{ 1 + |w|/(n-1) \right\}^{-n}$$

que es el núcleo de una densidad  $F_{(2, 2(n-1))}$

iv) Cuando  $\beta \rightarrow 1$ ,  $P(\theta|Y)$  no puede expresarse en términos de funciones simples de las observaciones. Sin embargo en el límite, la moda es la mediana de las observaciones si  $n$  es impar, y es algún valor único entre los valores de las observaciones medias si  $n$  es par. Cuando  $\beta = 1$  y  $n$  es par, la densidad es constante para valores de  $\theta$  entre las observaciones medias.

v) En ciertos casos, es posible expresar  $P(\theta|Y)$  en términos de un número fijo de funciones de las observaciones. Sin embargo, en general, la distribución final de  $\theta$  no puede expresarse en términos de funciones simples de las observaciones.

Las diferencias entre las distintas distribuciones finales de  $\theta$  para diversos valores de  $\beta$  son evidentes; y debido a que el problema es precisamente que  $\beta$  es desconocido, estas distribuciones no son muy útiles para el presente análisis, además, como se mencionó en v), no

siempre es posible expresar a  $P(\theta|Y)$  en términos de funciones simples de las observaciones.

Como  $\beta$  es desconocido debe expresarse un estado de incertidumbre, respecto al valor de  $\beta$ , asignándole una función de distribución inicial  $p(\beta)$ ; esta incertidumbre deberá reflejarse en la distribución final de  $\theta$ .  $P(\beta)$  será primeramente considerada una distribución no informativa y posteriormente tratará de incorporarse información sobre el valor de  $\beta$  a través de ella.

### 3.1.2 $\beta$ DESCONOCIDA.

Supóngase por el momento, que ya se cuenta con una distribución inicial para  $\beta$ ,  $P(\beta)$ ; debe encontrarse entonces,  $P(\beta|y)$  y después  $P(\theta|y)$ . Supóngase que inicialmente  $\beta$  tiene una distribución independiente de la media  $\theta$  y de la desviación estándar  $\sigma$ , tal que

$$P(\theta, \sigma, \beta) = P(\beta)P(\theta, \sigma) \quad (3.1.2.1)$$

y como antes, se considera la distribución inicial de referencia para  $(\theta, \sigma)$

$$P(\theta, \sigma) \propto \sigma^{-1} \quad (3.1.2.2)$$

Debe tenerse cuidado con las suposiciones de independencia de los parámetros, Box y Tiao comentan este problema diciendo que aunque parezca razonable suponer, inicialmente, que  $\theta$  es independiente de  $\sigma$  y  $\beta$ , la independencia de  $\beta$  y  $\sigma$  no es tan evidente. Considérese la forma (2.1) de la familia de distribuciones Exponencial Potencia, usando a  $\phi$  como parámetro de escala

$$\phi = f(\beta) \sigma$$

donde  $f(\beta)$  es una función de  $\beta$ . De (3.1.2.1) se tiene que  $\log \sigma$  y  $\beta$  son independientes. Suponiendo que

$$P(\log \sigma, \beta) = P(\log \sigma)P(\beta)$$

pero  $P(\log \sigma) \propto c$ , y puesto que  $\log \phi = \log f(\beta) + \log \sigma$ , se tiene localmente que para  $\beta$  dado,  $P(\log \phi) \propto c$ ; y si  $\log \sigma$  es localmente uniforme e independiente de  $\beta$ , entonces  $\log \phi$  y  $\beta$  son aproximadamente independientes (ver Apéndice II).

Empleando los resultados anteriores, la distribución final conjunta de  $(\theta, \sigma, \beta)$  es

$$P(\theta, \sigma, \beta | \underline{Y}) \propto P(\theta, \sigma, \beta) P(\underline{Y} | \theta, \sigma, \beta) = \sigma^{-1} P(\beta) P(\underline{Y} | \theta, \sigma, \beta)$$

donde  $P(\underline{Y} | \theta, \sigma, \beta)$  es la función de verosimilitud (3.1.1). Entonces

$$P(\theta, \sigma, \beta | \underline{Y}) \propto \sigma^{-1} P(\beta) w(\beta) \sigma^{-n} \exp \left\{ -c(\beta) \sum_{i=1}^n |(y_i - \theta) / \sigma|^{2/(1+\beta)} \right\}$$

donde  $c(\beta)$  y  $w(\beta)$  son las definidas en (2.2). Integrando  $P(\theta, \sigma, \beta | \underline{Y})$  sobre  $\sigma$ , se obtiene la distribución final conjunta de  $\theta$  y  $\beta$

$$\begin{aligned} P(\theta, \beta | \underline{Y}) &\propto \int_0^{\infty} P(\theta, \sigma, \beta | \underline{Y}) d\sigma \\ &= p(\beta) w(\beta)^n \int_0^{\infty} \sigma^{-(n+1)} \exp \left[ -c(\beta) \sum_{i=1}^n |y_i - \theta|^{2/(1+\beta)} \sigma^{-2/(1+\beta)} \right] d\sigma \end{aligned}$$

empleando la expresión (3.1.1.4), con

$$p = n, \quad a = c(\beta) \sum_{i=1}^n |y_i - \theta|^{2/(1+\beta)}, \quad \alpha = 2/(1+\beta)$$

se tiene

$$\begin{aligned} P(\theta, \beta | \underline{Y}) &\propto p(\beta) w(\beta)^n [(1+\beta)/2] \left[ c(\beta) \sum_{i=1}^n |y_i - \theta|^{2/(1+\beta)} \right]^{-n(1+\beta)/2} \Gamma [n(1+\beta)/2] \\ &\propto P(\beta) [M(\theta)]^{-n(1+\beta)/2} \Gamma [1+(n/2)(1+\beta)] \{ \Gamma [1+(1/2)(1+\beta)] \}^{-n} \end{aligned} \quad (3.1.2.3)$$

esta distribución es útil para la obtención de las probabilidades finales para  $\theta$  y  $\beta$  respectivamente.

Integrando (3.1.2.3) sobre  $\theta$ , se obtiene la distribución final para  $\beta$

$$\begin{aligned} P(\beta | \underline{Y}) &= \int_{-\infty}^{\infty} P(\theta, \beta | \underline{Y}) d\theta \\ &\propto P(\beta) \Gamma [1+(n/2)(1+\beta)] \{ \Gamma [1+(1/2)(1+\beta)] \}^{-n} J(\beta) \end{aligned} \quad (3.1.2.4)$$

con  $J(\beta)$  definida en (3.1.1.8).

Similarmente, la distribución final de  $\theta$  se obtiene integrando sobre  $\beta$ , la expresión (3.1.2.3), esto es

$$P(\theta | \underline{Y}) = \int_{-1}^1 P(\theta, \beta | \underline{Y}) d\beta, \quad \theta \in \mathbb{R} \quad (3.1.2.5)$$

Todo lo anterior se ha desarrollado para cualquier  $P(\beta)$ . En particular, si no se desea incorporar información inicial acerca de  $\beta$ , puede utilizarse una distribución inicial de referencia.

Una distribución inicial de referencia factible es la Uniforme, puesto que su incorporación se reduce a eliminar el término  $P(\beta)$  en la expresión (3.1.2.5).

Si se estuviera interesado en incorporar información inicial, como por ejemplo, la normalidad de la distribución de ciertas observaciones, podría entonces escogerse una distribución inicial para  $\beta$ , de tal forma que sea unimodal y centrada en  $\beta = 0$ . Una distribución con las características mencionadas que se puede aplicar es la distribución Beta simétrica con media 0 y extendida de -1 a 1 con un parámetro  $a$  ajustable. Específicamente, se tiene

$$P(\beta) = w(1+\beta^2)^{a-1}, \quad \beta \in (-1,1]$$

donde  $w = \Gamma(2a) [\Gamma(a)]^{-2} 2^{-(2a-1)} \quad a \geq 1$

cuando  $a = 1$  la distribución es uniforme; con  $a > 1$  se tiene una distribución simétrica con moda en  $\beta = 0$ , y conforme se incrementa  $a$  la distribución se concentra cada vez más alrededor de  $\beta = 0$ ; cuando  $a$  tiende a infinito,  $p(\beta)$  se aproxima a la función  $\delta$

$$\delta_{\beta}(0) = \begin{cases} 1 & \text{si } \beta = 0 \\ 0 & \text{en otro caso} \end{cases}$$

representando una suposición de normalidad exacta.

### 3.2 DISTRIBUCIONES FINALES PARA EL MODELO DE REGRESION LINEAL.

El modelo de Regresión Lineal es

$$E(\underline{Y}|X) = X\underline{\theta}$$

donde  $\underline{Y}$  es un vector de observaciones de  $n \times 1$ ,  $X$  una matriz de elementos fijos y de rango completo  $n \times k$  y  $\underline{\theta}$  un vector de coeficientes desconocidos.

Ahora, en lugar de suponer que las observaciones provienen de una distribución Normal, supongamos que la distribución de  $\underline{Y}$  pertenece a la familia de distribuciones Exponencial Potencia y que las observaciones son independientes e idénticamente distribuidas, entonces

$$P(\underline{Y} | X, \underline{\theta}, \sigma, \beta) = w(\beta) \sigma^{-1} \exp \left\{ -c(\beta) \sum_{i=1}^n |(y_i - x_i' \underline{\theta}) / \sigma|^{2/(1+\beta)} \right\} \quad (3.2.2)$$

donde  $x_i'$  es el  $i$ -ésimo renglón de  $X$ . La función de verosimilitud de  $(\underline{\theta}, \sigma, \beta)$  es entonces

$$L(\underline{\theta}, \sigma, \beta | \underline{Y}) \propto [w(\beta)]^n \sigma^{-n} \exp \left\{ -c(\beta) \sum_{i=1}^n |(y_i - x_i' \underline{\theta}) / \sigma|^{2/(1+\beta)} \right\} \quad (3.2.3)$$

siguiendo argumentos similares a los del caso  $\theta \in \mathbb{R}$ , supóngase inicialmente que

$$P(\underline{\theta}, \sigma, \beta) = P(\beta) P(\underline{\theta}, \sigma)$$

integrando sobre  $\sigma$  se obtiene la distribución final conjunta de  $(\underline{\theta}, \beta)$

$$P(\underline{\theta}, \beta | \underline{Y}) = P(\underline{\theta} | \beta, \underline{Y}) P(\beta | \underline{Y})$$

que es la distribución final condicional de  $\underline{\theta}$  y  $\beta$ . Entonces

$$P(\underline{\theta} | \beta, \underline{Y}) = J(\beta)^{-1} [M(\underline{\theta})]^{-n(1+\beta)/2} \quad \theta_j \in \mathbb{R} \quad j=1, \dots, k \quad (3.2.3)$$

donde

$$M(\underline{\theta}) = \sum |(y_i - x_i' \underline{\theta})|^{2/(1+\beta)} \quad (3.2.4)$$

y

$$J(\beta) = \int_{\mathbb{R}^k} [M(\underline{\theta})]^{-n(1+\beta)/2} d\underline{\theta} \quad (3.2.5)$$

además

$$P(\theta_j | \beta, \underline{Y}) = \int_{\mathbb{R}^{k-1}} P(\underline{\theta} | \beta, \underline{Y}) d\theta_1 \dots d\theta_{j-1} d\theta_{j+1} \dots d\theta_k, \quad j=1, 2, \dots, k \quad (3.2.6)$$

y

$$P(\beta | \underline{Y}) \propto P_u(\beta | \underline{Y}) P(\beta) \quad \beta \in (-1, 1] \quad (3.2.7)$$

donde

$$P_u(\beta | \underline{Y}) \propto \Gamma [1+n(1+\beta)/2] \{ \Gamma [1+(1+\beta)/2] \}^{-n} J(\beta)$$

y como en la sección anterior

$$P(\underline{\theta} | \underline{Y}) = \int_{-1}^1 P(\underline{\theta} | \beta, \underline{Y}) P(\beta | \underline{Y}) d\beta \quad (3.2.8)$$

$$P(\theta_j | \underline{Y}) = \int_{\mathbb{R}^{k-1}} P(\underline{\theta} | \underline{Y}) d\theta_1 \dots d\theta_{j-1} d\theta_{j+1} \dots d\theta_k, \quad j=1, \dots, k \quad (3.2.9)$$

Las distribuciones que nos interesan y que además son las que sirven para realizar el análisis de Normalidad y Robustez son  $P(\beta|Y)$ ,  $P(\theta_j|\beta, Y)$ ,  $P(\theta_j|Y)$ ,  $j = 1, 2, \dots, k$ .

La distribución  $P(\beta|Y)$  es útil para realizar análisis de robustez de la distribución del parámetro de no normalidad,  $\beta$ ; dicho análisis puede realizarse gráficamente; por ejemplo, pueden obtenerse curvas para varios valores de  $a$  ( $a=1, 3, 6, \dots$ ) y si las curvas no difieren 'mucho', podemos decir que las inferencias que se hagan sobre  $\beta$  serán robustas bajo diferentes distribuciones iniciales.

Con  $P(\theta_j|\beta, Y)$  ( $j = 1, 2, \dots, k$ ), puede observarse si existe robustez de inferencia con respecto a  $\theta_j$ . El procedimiento para hacer el análisis gráficamente es el siguiente :

se grafican algunas de las distribuciones de  $\theta_j$  para valores de  $\beta$  igualmente espaciados a lo largo del intervalo  $(-1, 1]$ . Si las curvas son muy parecidas entre sí, puede decirse que las inferencias que se realicen para el parámetro  $\theta_j$  serán robustas, es decir, que dichas inferencias no dependen de la suposición distribucional inicial que se asignó a las observaciones. Lo anterior también implica que  $P(\theta_j|Y)$  también será insensitiva a cambios en  $\beta$ , esto es, las curvas para diferentes valores de  $a$  serán muy parecidas.

En el caso de que para algunos valores de  $\beta$  las distribuciones de  $P(\theta_j|\beta, Y)$  difieran del resto, o que todas las distribuciones graficadas tengan medias muy diferentes, se observaría el comportamiento de  $P(\theta_j|Y)$ ; ya que puede ocurrir que en promedio, ponderado sobre  $\beta$ , la distribución final de  $\theta_j$  no cambie bajo diferentes 'grados' de suposición de normalidad, esto indicaría que finalmente las inferencias serían robustas a pesar de que no se detectó la robustez con las distribuciones marginales.

En el caso extremo de que las distribuciones finales de  $\theta_j$  no proporcionen evidencia de que exista robustez de inferencia, lo que se haría es, mediante  $P(\beta|Y)$  con  $a = 1$ , averiguar cuál es la probabilidad de que  $\beta$  tome los valores correspondientes a los casos problema, y dependiendo de ésta, se puede estar en condiciones de optar por una posible distribución de  $\beta$ , quedándose con aquella que acumule mayor probabilidad.

#### 4. IMPLEMENTACION.

Como ya se ha discutido, con la ayuda de las distribuciones finales que se obtuvieron en el capítulo anterior, puede efectuarse el análisis de normalidad. Dichas distribuciones, no tienen expresión analítica definida - generalmente dependen de una integral-, por lo que para realizar el análisis en problemas concretos, se generaron programas en FORTRAN con el auxilio de algunos métodos numéricos.

##### 4.1 METODOS NUMERICOS EMPLEADOS.

Básicamente se requirió lo siguiente

i) Generar variables aleatorias Exponencial Potencia a fin de simular experimentos con valores de los parámetros fijos y conocidos. Esto con el afán de ejemplificar los resultados que pueden obtenerse en diferentes situaciones; básicamente se está interesado en generar valores que provengan de diferentes distribuciones y así analizar el cambio en las distribuciones finales de  $\theta$  y  $\beta$ .

ii) Obtener las distribuciones finales para  $\theta$  y  $\beta$ .

iii) Obtener probabilidades para la distribución final de  $\beta$ .

Para lograr lo anterior se emplearon los siguientes métodos:

- Generación de variables aleatorias Uniformes (0,1).

Se utilizó el método congruencial, Rubinstein (1981), que consiste en lo siguiente: a partir de una semilla inicial  $x_0$  se obtienen las variables de la forma  $x_{i+1} \equiv (ax_i + c) \pmod{m}$   $i = 1, 2, \dots$  donde  $a, c$  y  $m$  son enteros no negativos.

Rubinstein demuestra que esto es suficiente para que  $x_{i+1} \in [0, m)$  Este generador tiene período completo si y sólo si  $c$  es primo relativo de  $m$ , es decir,  $a \equiv 1 \pmod{g}$  para cada factor primo  $g$  de  $m$  y también si  $m$  es múltiplo de 4, es decir,  $a \equiv 1 \pmod{4}$ ; al cumplirse ambas

generado  $m$  variables. Rubinstein asegura que esta propiedad se cumple si  $m = 2^{25}$ ,  $a = 2^7 + 1$  y  $c = 1$ . Con esto se genera  $x_i \sim U(0, m)$ ; para obtener variables aleatorias  $U(0, 1)$

$$U_i = x_i / m \sim U(0, 1)$$

- Cálculo de la función Gamma ( $\Gamma(x)$ ).

Se utilizó la aproximación polinomial (Abramowitz & Stegun, 1965):

$\Gamma(x+1) = x! = 1 + b_1 x + b_2 x^2 + \dots + b_8 x^8 + e(x)$ , con  $x \in [0, 1]$ ,  $|e(x)| \leq 3(10^{-7})$  donde

$b_1 = -0.577191652$	$b_2 = 0.988205891$
$b_3 = -0.897056937$	$b_4 = 0.918206857$
$b_5 = -0.756704078$	$b_6 = 0.482199394$
$b_7 = -0.193527818$	$b_8 = 0.035868343$

- Obtención de las variables aleatorias Exponencial Potencia.

Forsythe (1972) dió una modificación al método de rechazo para obtener muestras de una distribución continua, cuya densidad puede representarse como:

$$f(x) = c e^{-B(x)}$$

donde  $B(x)$  es una función creciente de  $x$  sobre  $(0, \infty)$ . El método es una modificación de una técnica usada por Von-Neumann (1951) para generar variables aleatorias exponenciales. Requiere de tres tablas,  $\{q_k\}, \{r_k\}, \{d_k\}$ , para  $k = 0, 1, \dots, K$  que se obtienen como sigue:

Sea  $q_0 = 0$ , para cada  $k = 1, 2, \dots$  se elige  $q_k$  tan grande como sea posible y sujeta a las siguientes restricciones

$$(1) \quad q_k - q_{k-1} \leq 1$$

$$(2) \quad B(q_k) - B(q_{k-1}) \leq 1$$

enseguida se calcula

$$r_k = \int_0^{q_k} f(x) dx, \quad k = 0, 1, \dots, k$$

aquí,  $K$  se escoge como el índice de  $r$  de tal forma que  $r_k$  es el número más grande posible menor que 1, el método de integración usado fue el

de Gauss - Aitken (Byte, 1986), finalmente se calcula

$$d_k = q_k - q_{k-1}$$

por simplicidad se definen las funciones

$$G_k(x) = E(q_{k-1} + x) - E(q_{k-1}) \quad , \quad k = 1, 2, \dots, K$$

En el algoritmo, que a continuación se presenta, los pasos 1 al 3 determinan el intervalo  $[q_{k-1}, q_k)$  al que pertenece la variable generada ( $y$ ), los pasos restantes determinan el valor de la variable dentro del intervalo elegido.

1.  $k \leftarrow 1$ . Generar una variable aleatoria uniforme  $U(0,1)$ .
2. Si  $U \leq r_k$ , ir a 4.
3. Si  $U > r_k$ ,  $k \leftarrow k+1$  y regresar a 2.
4. Generar otra variable aleatoria uniforme  $U(0,1)$ , y  $w \leftarrow U d_k$ .
5.  $T \leftarrow G_k(w)$ .
6. Generar otra variable aleatoria uniforme  $U^*(0,1)$ .
7. Si  $U^* \geq T$ , ir a 11.
8. Si  $U^* < T$  generar otra variable aleatoria uniforme  $U(0,1)$ .
9. Si  $U < U^*$ ,  $T \leftarrow U$  y regresar a 6.
10. Si  $U \geq U^*$ , regresar a 4.
11.  $y \leftarrow q_{k-1} + w$

En el caso tratado en este trabajo, la función de densidad es

$$P(y|\theta, \sigma, \beta) = w(\beta) \sigma^{-1} \exp \left\{ -c(\beta) \left| (y-\theta)/\sigma \right|^{2/(1+\beta)} \right\} \quad , \quad \text{con}$$

$$\theta, y \in \mathbb{R} \quad , \quad \sigma \in \mathbb{R}^+ \quad , \quad \beta \in [-1, 1)$$

y  $c(\beta)$ ,  $w(\beta)$  dadas en (2.2)

Para facilitar los cálculos, se estandariza la densidad como sigue sea  $x = (y-\theta)/\sigma \rightarrow dx/dy = 1/\sigma$  y  $y = \sigma x + \theta$  entonces

$$\begin{aligned} P(x|\theta, \sigma, \beta) &= P(y(x)|\theta, \sigma, \beta) / |J| \\ &= P(\theta + x\sigma|\theta, \sigma, \beta) \sigma \\ &= w(\beta) \exp \left\{ -c(\beta) \left| (\theta + x\sigma - \theta)/\sigma \right|^{2/(1+\beta)} \right\} \sigma \end{aligned}$$

por tanto

$$P(x|\theta, \sigma, \beta) = w(\beta) \exp\{-c(\beta)|x|^{2/(1+\beta)}\} = P(x|\theta=0, \sigma=1, \beta) \quad \beta \in (-1, 1]$$

que es la densidad estandarizada. Esta función es simétrica con respecto al 0, por lo que

$$P(x|\theta=0, \sigma=1, \beta) = 2w(\beta) \exp\{-c(\beta)|x|^{2/(1+\beta)}\}, \quad \beta \in (-1, 1] \\ x \in \mathbb{R} \\ x \in [0, \infty)$$

El algoritmo se aplica directamente, adicionando un paso para obtener el signo de la variable:

12. Generar una variable aleatoria  $U(0,1)$ . Si  $U < 0.5$ ,  $y \leftarrow -y$ .

- Integración numérica Gauss - Aitken.

El algoritmo utilizado es el presentado por David M. Smith (Byte, 1986), en el cual, se exhibe un método para aproximar integrales con un número mínimo de evaluaciones de la función a integrar ( $f(x)$ ).

Para usar la extrapolación de Aitken, se necesita un conjunto de aproximaciones a la integral de  $f(x)$  desde  $a$  hasta  $b$ , donde  $a$  y  $b$  son los límites de integración. Para aproximar el área bajo  $f(x)$  se usó la fórmula de cuadratura de Gauss; una ventaja del método de Gauss sobre otros métodos de integración, es que la función no se evalúa en ninguno de los puntos extremos del intervalo de integración, esto es una ventaja para funciones con singularidades en los extremos.

Una vez que se ha obtenido una tabla de valores mediante Gauss, se tiene una secuencia de aproximaciones  $T_1, T_2, T_3, \dots$ , las cuales convergen al valor correcto de la integral. En teoría el análisis numérico indica que tal sucesión es linealmente convergente; la extrapolación de Aitken usa esta propiedad de convergencia lineal para obtener mayor exactitud a partir de los resultados obtenidos por el método de Gauss, sin necesidad de evaluaciones adicionales de la función.

Si  $T$  representa el valor exacto de la integral, y  $e_i$  el error de la aproximación de Gauss en el renglón  $i$ -ésimo de la tabla, entonces  $T_i - T = e_i$ . Por la definición de convergencia lineal, existe una constante  $c$  para la cual  $e_i = ce_{i-1}$  para cada  $i$ . Suponiendo esto, resultan tres ecuaciones

$$t_i - t = e_i$$

$$t_{i+1} - t = ce_i$$

$$t_{i+2} - t = c^2 e_i$$

resolviendo para  $t$  obtenemos

$$t = t_{i+2} - (t_{i+2} - t_{i+1})^2 / (t_{i+2} - 2t_{i+1} + t_i)$$

esta es la expresión para la extrapolación de Aitken. En la práctica, la convergencia no es precisamente lineal, así que, el cálculo de  $t$  no es aún el resultado exacto, pero es frecuentemente más exacto que cualquiera de los tres valores usados para generarlo.

- Obtención de probabilidades finales.

Las probabilidades finales involucran el cálculo de  $J(\beta)$  (3.1.5). Esta función de  $\beta$  se desestabiliza a medida que  $\beta$  se acerca a  $-1$ , lo cual provoca que la forma de las distribuciones finales, tanto para  $\beta$  como para  $\theta$ , no puedan ser obtenidas fácilmente con métodos de integración con ponderaciones sobre las abscisas, ya que caen en severas dificultades, especialmente cuando se trata de integración múltiple.

Por lo anterior, resulta más conveniente calcular estas integrales por el método de Monte-Carlo, el cual, aunque es menos exacto que los métodos de cuadratura convencionales, presenta menos problemas cuando se realizan las aproximaciones. El método se describe a continuación:

Sea  $F$  función de distribución sobre algún espacio de probabilidad  $\Omega$ , y sea  $g$  una función real valuada sobre el mismo espacio, entonces, el valor esperado de  $g$  con respecto a  $F$  es

$$\mu(g, F) = \int_{\Omega} g(x) dF(x)$$

por la ley fuerte de los grandes números, si el valor esperado de  $g$  existe, entonces, con probabilidad 1, la media de  $g$  para una muestra  $z_1, z_2, \dots, z_M$  de  $F$

$$m(g, F, M) = \left[ \sum_{i=1}^M g(z_i) \right] / M$$

tiende a  $\mu(g, F)$  cuando  $M \rightarrow \infty$ ,  $z_i$   $i = 1, \dots, M$  independientes.

Para una integral definida con límites  $(a, b)$ , la media es

$$m(g, F, M) = [(b-a)/M] \sum_{i=1}^M g(z_i)$$

por simplicidad, sólo se considerará el caso de Regresión Lineal Simple, esto es,  $\underline{\theta} = (\theta_1, \theta_2)$ , entonces las distribuciones quedan como

$$J(\beta) = [(b_2 - a_2)/n_2] \sum_{i=1}^{n_2} [(b_1 - a_1)/n_1] \sum_{j=1}^{n_1} \left[ \sum_{k=1}^n |y_k - \theta_{1j} - x_k \theta_{2i}|^{2/(1+\beta)} \right]^{-n(1+\beta)/2}$$

$$P(\beta|y) = \frac{\Gamma[1+(n/2)(1+\beta)] \{ \Gamma[1+(1/2)(1+\beta)] \}^{-n} J(\beta) P(\beta)}{[(b_3 - a_3)/n_3] \sum_{i=1}^{n_3} \left[ \Gamma[1+(n/2)(1+\beta_i)] \{ \Gamma[1+(1/2)(1+\beta_i)] \}^{-n} J(\beta_i) P(\beta_i) \right]}$$

$$P(\theta_1 | \beta, y) = J(\beta)^{-1} [(b_1 - a_1)/n_1] \sum_{j=1}^{n_1} \left[ \sum_{i=1}^n |y_i - \theta_{1j} - x_i \theta_{2j}|^{2/(1+\beta)} \right]^{-n(1+\beta)/2}$$

$$P(\theta_2 | \beta, y) = J(\beta)^{-1} [(b_2 - a_2)/n_2] \sum_{i=1}^{n_2} \left[ \sum_{j=1}^n |y_j - \theta_{1j} - x_j \theta_{2i}|^{2/(1+\beta)} \right]^{-n(1+\beta)/2}$$

$$P(\theta_1 | y) = [(b_3 - a_3)/n_3] \sum_{i=1}^{n_3} P(\theta_1 | \beta_i, y)$$

$$P(\theta_2 | y) = [(b_3 - a_3)/n_3] \sum_{i=1}^{n_3} P(\theta_2 | \beta_i, y)$$

con  $P(\beta)$  definida en (3.2.3.1)

$(a_1, b_1)$  y  $n_1$  son los límites de integración y el número de particiones para  $\theta_1$  respectivamente.

$(a_2, b_2)$  y  $n_2$  son los límites de integración y el número de particiones para  $\theta_2$  respectivamente.

$(a_3, b_3)$  y  $n_3$  son los límites de integración y el número de particiones para  $\beta$  respectivamente.

Por otro lado, para obtener las distribuciones finales de  $\underline{\theta}$  y  $\beta$ , es necesario conocer el rango de valores de  $\underline{\theta}$  sobre el cual se va a integrar. Estos valores se obtuvieron a partir de los estimadores máximo verosímiles de las medias respectivas, las cuales coinciden con las modas (Apéndice II), y un rango aproximado de valores que pueden

tomar con mayor probabilidad.

Para calcular los estimadores, se empleó el método de Newton para resolver sistemas de ecuaciones en  $n$  dimensiones. Geométricamente el método de Newton ajusta tangentes (planos o hiperplanos) a cada superficie en el espacio de dimensión  $n+1$ , entonces resuelve para estas intersecciones, en vez de resolver para las intersecciones originales. Si los planos tangentes son buenos representantes de las superficies entre las intersecciones actuales y las raíces que se buscan, el método tendrá éxito.

Se desea resolver el sistema

$$\begin{aligned} f_1(x_1, \dots, x_n) &= 0 \\ f_2(x_1, \dots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, \dots, x_n) &= 0 \end{aligned}$$

teniendo un punto  $a$  inicial, expandiendo cada función alrededor de  $a$  en el espacio de dimensión  $n$  por medio de series de Taylor, conservando sólo los términos lineales e igualando estas expansiones a 0 se obtiene

$$0 = f_1(a) + \partial f_1 / \partial x_1 (\delta_1) + \dots + \partial f_1 / \partial x_n (\delta_n)$$

$$0 = f_2(a) + \partial f_2 / \partial x_1 (\delta_1) + \dots + \partial f_2 / \partial x_n (\delta_n)$$

$\vdots$

$$0 = f_n(a) + \partial f_n / \partial x_1 (\delta_1) + \dots + \partial f_n / \partial x_n (\delta_n)$$

donde  $\delta_i$  representa desviaciones de  $a$ . Después de resolver este conjunto de ecuaciones lineales para  $\{\delta_i\}$  se calcula una nueva aproximación a la raíz con

$$x_i = a_i + \delta_i, \quad i=1, \dots, n$$

El valor inicial para  $a$  fue el estimador máximo verosímil de  $\theta = (\theta_1, \theta_2)$  para una distribución Normal

$$\theta_1 = \bar{y} - \bar{\theta}_2 \bar{x} \quad ; \quad \theta_2 = \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Las funciones que intervienen en el sistema a resolver son

$$f(\theta) = \sum_i |y_i - \theta_1 - \theta_2 x_i|^{2/(1+\beta)} \quad i = 1, 2, \dots, n$$

$$f(\theta) = \begin{cases} \sum_i (y_i - \theta_1 - \theta_2 x_i)^{2/(1+\beta)} \\ -\sum_i (y_i - \theta_1 - \theta_2 x_i)^{2/(1+\beta)} \end{cases} \quad (a)$$

(b)

$$\frac{\partial f_1(\theta)}{\partial \theta_1} = \begin{cases} -[2/(1+\beta)] \sum_i (y_i - \theta_1 - \theta_2 x_i)^{[2/(1+\beta)]-1} \\ [2/(1+\beta)] \sum_i (y_i - \theta_1 - \theta_2 x_i)^{[2/(1+\beta)]-1} \end{cases} \quad (a)$$

(b)

$$\frac{\partial f_1(\theta)}{\partial \theta_2} = \begin{cases} [2/(1+\beta)] \{ [2/(1+\beta)] - 1 \} \sum_i x_i (y_i - \theta_1 - \theta_2 x_i)^{[2/(1+\beta)]-1} \\ -[2/(1+\beta)] \{ [2/(1+\beta)] - 1 \} \sum_i x_i (y_i - \theta_1 - \theta_2 x_i)^{[2/(1+\beta)]-1} \end{cases} \quad (a)$$

(b)

$$= \frac{\partial f_2(\theta)}{\partial \theta_2}$$

$$\frac{\partial f_2(\theta)}{\partial \theta_1} = \begin{cases} [2/(1+\beta)] \{ [2/(1+\beta)] - 1 \} \sum_i x_i^2 (y_i - \theta_1 - \theta_2 x_i)^{[2/(1+\beta)]-1} \\ -[2/(1+\beta)] \{ [2/(1+\beta)] - 1 \} \sum_i x_i^2 (y_i - \theta_1 - \theta_2 x_i)^{[2/(1+\beta)]-1} \end{cases} \quad (a)$$

(b)

(a) si  $y_i - \theta_1 - \theta_2 x_i > 0$

(b) si  $y_i - \theta_1 - \theta_2 x_i < 0$

$$i = 1, 2, \dots, n$$

El rango de valores para  $\theta$  se obtuvo a partir del cálculo del intervalo de confianza para una distribución Normal, tomando como centro del intervalo la moda correspondiente en cada caso para asegurar que se está cubriendo el intervalo con mayor densidad

intervalo para  $\theta_1$   $\hat{\theta}_1 \pm t_{(n-2)}^* \hat{\sigma} \left\{ \sum x_i^2 / (n-1) \sum (x_i - \bar{x})^2 \right\}^{1/2}$

intervalo para  $\theta_2$   $\hat{\theta}_2 \pm t_{(n-2)}^* \hat{\sigma} \left\{ n / (n-2) \sum (x_i - \bar{x})^2 \right\}^{1/2}$

con  $\hat{\sigma} = (\hat{\theta}_1, \hat{\theta}_2)$  la moda correspondiente y

$$\hat{\sigma}^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n-2)$$

donde  $\hat{y}_i = \hat{\theta}_1 + \hat{\theta}_2 x_i$ ;  $\hat{\theta}$  es el estimador máximo verosímil para la Normal.

- Obtención de probabilidades para la distribución final de  $\beta$ .

Se utilizó el Spline cúbico para interpolación de valores. Se recurrió a este método para obtener curvas lo más suaves posible.

#### 4.2 PROGRAMAS GENERADOS.

Se generaron programas que aplican los métodos expuestos en 4.1; además, se requirió graficar las distribuciones finales de  $\theta$  y  $\beta$ .

Los programas se diseñaron para ser procesados por una minicomputadora VAX.

Para la generación de variables aleatorias Exponencial Potencia, se empleó la adaptación al método sugerido por Forsythe (1972); para calcular los cuantiles  $\{q_k\}$  y las probabilidades acumuladas  $\{r_k\}$ , se cuenta con el programa GEP1, en el cual las integrales requeridas se calculan con las subrutinas llamadas GAUSS y AITKEN. Estas integrales se determinan por el método de Gauss y posteriormente se refinan usando la extrapolación de Aitken. Las probabilidades se obtienen para las distribuciones pertenecientes a la familia de distribuciones Exponencial Potencia cuyo parámetro  $\beta$  especifique el usuario.

Una vez obtenidas las tablas anteriores, se procede a generar las variables aleatorias Exponencial Potencia para el valor específico de  $\beta$  proporcionado en GEP1, para lo cual se cuenta con el programa GEP2. En este programa se utiliza la subrutina UN para la generación de variables aleatorias Uniformes  $U(0,1)$ , empleando el método congruencial. Las variables aleatorias se obtienen para un modelo de Regresión Lineal  $E(Y|X) = X\theta$ , y la dimensión máxima que se maneja para  $\theta$  es de 14.

El programa MON1 se diseñó para la obtención de las distribuciones finales de los parámetros de interés  $(\theta_1, \theta_2, \beta)$ , para lo cual se requiere obtener los estimadores máximo verosímiles correspondientes a  $\theta_1$  y  $\theta_2$ ; estos estimadores son calculados en la subrutina RAIZ, mediante el método de Newton para resolver sistemas de ecuaciones en

$\mathbb{R}^n$ , posteriormente se obtienen rangos de valores más probables para  $\theta_1$  y  $\theta_2$ , estos rangos de valores son obtenidos por la subrutina ETHETA.

Para el cálculo de la función Gamma mediante aproximación polinomial, se empleó la subrutina GAMMA.

Finalmente, para integrar numéricamente las expresiones correspondientes a las distribuciones finales de cada parámetro, se empleó el método de integración Monte Carlo. El programa MONI puede modificarse fácilmente para adaptarse al caso del Modelo Lineal General.

Una vez obtenidos algunos de los valores para la distribución final de  $\beta$ , se procede a interpolarlos utilizando el método de Spline Cúbico y así obtener probabilidades para diferentes valores de  $\beta$ . Para realizar los cálculos anteriores se generó el programa PROBA3.

Mediante el paquete ENERGRAPHICS se graficaron los valores obtenidos en el programa MONI para las distribuciones finales de  $\beta$ ,  $\theta_1$  y  $\theta_2$ , y los obtenidos para las distribuciones marginales de  $\theta_1$  y  $\theta_2$ .

A continuación se presentan los listados de los programas generados.

G E P 1.

C Cálculo de probabilidades (RK) para diferentes cuantiles (QK)  
C para la generación de variables aleatorias Exponencial Potencia  
C con un valor de  $\beta$  fijo.  
C Algoritmo tomado de :  
C FORSYTHE, GEORGE E., MATHEMATICS OF COMPUTATION, 1972.

C Entrada :  $\beta$   
C Salida : valores para RK y QK que se guardan en el archivo  
C DAT1.DAT  
C

```
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
DIMENSION Q(100),R(100),TABLA(30)
COMMON WB,CB,BE
COMMON /TABLA/ TABLA
WRITE(*,300)
OPEN(S,FILE='DAT1.DAT',STATUS='NEW')
WRITE(*,*)' BETA : '
READ(*,*)B
WRITE(S,*)B
WRITE(*,*)
Q(1)=ODO
R(1)=ODO
WRITE(S,*)Q(1),R(1)
BE=1DO+B
A=1.5DO*BE
C=0.5DO*BE
GA=GAMA(A)
GC=GAMA(C)
CB=(GA/GC)**(1DO/BE)
WB=(GA**0.5DO)/(BE*(GC**1.5DO))
V=0.1DO/CB
DO 100 J=2,100
L=J-1
RK=ODO
Q(J)=QK(BE,V,Q(L))
```

```

KM=5
CALL DINTE(KM,ODO,Q(J),RK)
R(J)=RK
C1=1D0-R(J)
IF(C1.LE.0.0001D0)GOTO 200
WRITE(5,*)Q(J),R(J)
100 CONTINUE
200 WRITE(5,*)Q(J),R(J)
300 FORMAT(/,17X,'FAMILIA DE DISTRIBUCIONES EXPONENCIAL POTENCIA',
-,23X,'GENERACION DE VARIABLES ALEATORIAS '///)
END

```

C  
C  
C  
C  
C

Cálculo de QK

```

FUNCTION QK(BE,V,QA)
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
W=(V+QA**(2/BE))**(BE/2D0)
QK=W
Z=0.1D0+QA
IF(Z.LE.W)QK=Z
RETURN
END

```

C  
C  
C  
C  
C  
C

Integración GAUSS-AITKEN  
para la obtención de las probabilidades RK

```

SUBROUTINE DINTE(KM,A1,B1,TABLAB)
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
DIMENSION TABLA(30)
COMMON WB,CB,BE
COMMON /TABLA/ TABLA
KL=KM

```

```

CALL GAUSS(KM,A1,B1)
DO 400 J=1,20
CALL AITKEN(KL,TABLA2,RELEER)
KL=KL-2
IF(KL .LT. 3)GO TO 500
400 CONTINUE
500 TABLAB=TABLA2
RETURN
END

```

```

C
C
C Cálculo de la integral de GAUSS
C

```

```

SUBROUTINE GAUSS(KM,A1,B1)
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
DIMENSION TABLA(30)
EXTERNAL F2
COMMON /TABLA/ TABLA
NLines=KM
NSUBS=1
DO 600 JLINE=2,NLines+1
XH=(B1-A1)/NSUBS
XH2=XH/2D0
XR=XH2/(3D0**0.5D0)
START1=A1-XH2-XR
START2=A1-XH2+XR
SUM=0D0
DO 650 K=1,NSUBS
SUM=SUM+F2(START1+K*XH)+F2(START2+K*XH)
650 CONTINUE
SUM=SUM*XH2
NSUBS=2*NSUBS
TABLA(JLINE)=SUM
600 CONTINUE
RETURN
END

```



F2=2D0\*WB\*DEXP(-CB\*(DABS(X)\*\*(2D0/BE)))

RETURN

END

C

C Función GAMA

C

FUNCTION GAMA(S)

IMPLICIT DOUBLE PRECISION(A-H,O-Z)

DIMENSION H(8)

H(1)=-0.577191652D0

H(2)=0.988205891D0

H(3)=-0.897056937D0

H(4)=0.918206857D0

H(5)=-0.756704078D0

H(6)=0.482199394D0

H(7)=-0.193527818D0

H(8)=0.035868343D0

EM=1D0

GAM=0D0

G=S

8 IF(G.LE.2D0)GOTO 9

G=C-1D0

EM=EM\*G

GOTO 8

9 IF(G.LE.1D0)G=G+1D0

F=G-1D0

DO 10 IK=1,8

GAM=H(IK)\*(F\*\*IK)+GAM

10 CONTINUE

IF(S.LT.1D0)GOTO 11

GAMA=(1D0+GAM)\*EM

GO TO 12

11 GAMA=(1D0+GAM)/F

12 RETURN

END

G E P 2.

C  
C  
C  
C  
C  
C  
C  
C  
C  
C  
C

Obtención de muestras de variables aleatorias EXPONENCIAL POTENCIA  
para un valor de  $\beta$  fijo

Entrada : DAT1.DAT que contiene BETA, RK Y QK

Salida : muestras aleatorias con los valores de THETA, SIGMA  
y X (variable independiente) dadas por terminal

```
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
DIMENSION Q(150),R(150),Y(60),TH(14)
EXTERNAL GAMA
COMMON X1,A,C,EM
COMMON /THETA/ TH
OPEN(5,FILE='DAT1.DAT',STATUS='OLD')
READ(5,*)B
I=1
100 READ(5,*,END=200)Q(I),R(I)
I=I+1
GO TO 100
200 WRITE(*,*)' TAMAÑO DE MUESTRA : '
READ(*,*)N
WRITE(*,*)
WRITE(*,*)' NUMERO DE VARIABLES INDEPENDIENTES : '
READ(*,*)M
WRITE(*,*)
WRITE(*,*)' VARIANZA DESEADA : '
READ(*,*)SIGMA
WRITE(*,*)
L=M+1
DO 50 J=1,L
WRITE(*,800)J
```

```

READ(*,*)TH(J)
50 CONTINUE
WRITE(*,*)
WRITE(*,*)' UNA SEMILLA INICIAL ;'
READ(*,*)X1
A=(2D0**7)+1D0
C=1D0
EM=2D0**35
DO 700 I=1,N
K=2
CALL UN(U1)
U1=U1*2D0
IF(U1 .LT. 1D0)GO TO 300
S=-1D0
U1=U1-1D0
GO TO 350
300 S=1D0
350 IF(U1 .LE. R(K))GO TO 400
K=K+1
GO TO 350
400 CALL UN(U2)
W=U2*(Q(K)-Q(K-1))
Z=Q(K-1)+W
T=BE(Z,B)-BE(Q(K-1),B)
500 CALL UN(U3)
IF(U3 .GE. T)GO TO 600
CALL UN(U4)
IF(U4 .GE. U3)GO TO 400
T=U4
GO TO 500
600 Y(I)=S*(Q(K-1)+W)
CALL TRANSF(Y(I),M,SIGMA)
700 CONTINUE
800 FORMAT(/,' THETA',I2,' :')
END

```

C

C  
C EXPONENCIAL POTENCIA  
C  
C

```
FUNCTION BE(X,B)
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
A1=1D0+B
A2=X**(2D0/A1)
A3=1.5D0*A1
A4=0.5D0*A1
B3=GAMA(A3)
B4=GAMA(A4)
BE=((B3/B4)**(1D0/A1))*A2
RETURN
END
```

C  
C  
C Transformación de la muestra generada mediante los valores  
C  
C de entrada THETA, SIGMA y X  
C  
C

```
FUNCTION TRANSF(YE,M,SIGMA)
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
DIMENSION XI(15),TH(14)
COMMON /THETA/ TH
RP=TH(1)
DO 1 J=1,M
WRITE(*,10)J
READ(*,*)XI(J)
RP=RP+XI(J)*TH(J+1)
1 CONTINUE
W1=(YE*(SIGMA**0.5D0))+RP
WRITE(*,20)W1
10 FORMAT(/,' DAR X',I2,' :')
20 FORMAT(//,12X,'Y = ',F10.3,/)
RETURN
END
```

C  
C Función GAMMA  
C

```
FUNCTION GAMA(S)  
IMPLICIT DOUBLE PRECISION(A-H,O-Z)  
DIMENSION H(8)  
H(1)=-0.577191652D0  
H(2)=0.988205891D0  
H(3)=-0.897056937D0  
H(4)=0.918206857D0  
H(5)=-0.756704078D0  
H(6)=0.482199394D0  
H(7)=-0.193527818D0  
H(8)=0.035868343D0  
EM=1D0  
GAM=0D0  
G=S  
8 IF(G.LE.2D0)GOTO 9  
G=G-1D0  
EM=EM*G  
GOTO 8  
9 IF(G.LE.1D0)G=G+1D0  
F=G-1D0  
DO 10 IK=1,8  
GAM=H(IK)*(F**IK)+GAM  
10 CONTINUE  
IF(S.LT.1D0)GOTO 11  
GAMA=(1D0+GAM)*EM  
GO TO 12  
11 GAMA=(1D0+GAM)/F  
12 RETURN  
END
```

C  
C  
C SUBROUTINE UN(UNI)

```
C      Generador de v.a. Uniformes en el intervalo (0,1)
C
C
C      Entrada: X1 semilla inicial
C      Salida : UNI número aleatorio en (0,1)
C
C
```

```
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
COMMON X1,A,C,EM
X2=A*X1+C
X2=DMOD(X2,EM)
UNI=X2/EM
X1=X2
RETURN
END
```

MON1.

```
C
C
C   Obtención de probabilidades finales de la distribución
C   EXPONENCIAL POTENCIA utilizando el método de integración
C   de MONTECARLO.
C
C   Se tiene como entrada únicamente las observaciones
C   correspondientes a un modelo de REGRESION LINEAL SIMPLE
C   (X), los cuales se encuentran en el archivo DATOE.DAT
C
C   IMPLICIT DOUBLE PRECISION(A-H,O-Z)
C   DIMENSION X(50,2),THETA(15,2),PPTB(7,15,15),AJBETA(7),PU(7),
C   -PPT(4,15,15),PPT1(4,15),PPT2(4,15),PPTB1(7,15),PPTB2(7,15),
C   -PB(4,7),SUM1(4),S(4,7),H(8)
C   COMMON X,N
C   EXTERNAL GAMA
C   OPEN(1,FILE='DATOE.DAT',STATUS='OLD')
C   OPEN(2,FILE='DATOS.DAT',STATUS='NEW')
C   OPEN(3,FILE='DATOA.DAT',STATUS='NEW')
C   DO 50 I=1,50
C   READ(1,*,END=60)(X(I,J),J=1,2)
50 CONTINUE
60 N=I-1
   WRITE(*,500)
   CALL ETHETA(T1,T2,AT1,AT2)
   WRITE(*,*)
   WRITE(*,*)'VALOR MINIMO DE BETA PARA CALCULAR EL INTERVALO'
   WRITE(*,*)'DE INTEGRACION :'
   READ(*,*)BETA
   CALL RAIZ(BETA,T1,T2,R1,R2)
   WRITE(*,*)
   WRITE(*,*)'THETA1= ',R1,' THETA2= ',R2
   WRITE(*,*)
   WRITE(*,*)'AMPLIACION HACIA LA IZQUIERDA PARA THETA1'
   READ(*,*)IC
```

```

WRITE(*,*)'AMPLIACION HACIA LA IZQUIERDA PARA THETA2'
READ(*,*)ID
THETA(1,1)=R1-AT1-IC
THETA(1,2)=R2-AT2-ID
WRITE(*,*)
WRITE(*,*)'VALOR MAXIMO DE BETA PARA CALCULAR EL INTERVALO'
WRITE(*,*)'DE INTEGRACION :'
READ(*,*)BETA
CALL RAIZ(BETA,T1,T2,R1,R2)
WRITE(*,*)'THETA1= ',R1,' THETA2= ',R2
WRITE(*,*)'AMPLIACION HACIA LA DERECHA PARA THETA1'
READ(*,*)IE
WRITE(*,*)'AMPLIACION HACIA LA DERECHA PARA THETA2'
READ(*,*)IF
WRITE(*,510)
THETA(15,1)=R1+AT1+IE
THETA(15,2)=R2+AT2+IF
P1=(THETA(15,1)-THETA(1,1))/14DO
P2=(THETA(15,2)-THETA(1,2))/14DO
DO 100 I=2,14
J=I-1
THETA(I,1)=THETA(J,1)+P1
THETA(I,2)=THETA(J,2)+P2
100 CONTINUE
CTE1=(THETA(15,1)-THETA(1,1))/15DO
CTE2=(THETA(15,2)-THETA(1,2))/15DO
BETA=-0.9DO
DO 180 I=1,7
PU(I)=0DO
BE=1DO+BETA
P=DELE(N)
C=(P/2DO)*BE
SUMA=0DO
DO 170 J=1,15
DO 160 K=1,15
S1=0DO

```

```

DO 150 L=1,N
S2=X(L,2)-THETA(J,1)-(THETA(K,2)*X(L,1))
S2=(DABS(S2))**(2DO/BE)
S1=S1+S2
150 CONTINUE
PPTB(I,J,K)=S1**(-C)
SUMA=CTE1*CTE2*PPTB(I,J,K)+SUMA
160 CONTINUE
170 CONTINUE
AJBETA(I)=SUMA
WRITE(*,*)'AJBETA = ',AJBETA(I),' ',I
G1=1DO+C
G2=1DO+(0.5DO*BE)
PU(I)=GAMA(G1)
PU(I)=PU(I)*(GAMA(G2)**(-N))
BETA=BETA+0.3DO
180 CONTINUE
A=1DO
DO 184 I=1,4
W=GAMA(2DO*A)*(GAMA(A)**(-2))*(2DO**((-2DO*A)+1DO))
BETA=-0.9DO
SUM1(I)=ODO
DO 182 J=1,7
S(I,J)=PU(J)*W*((1DO-BETA**2)**(A-1DO))
PB(I,J)=AJBETA(J)*S(I,J)
SUM1(I)=SUM1(I)+PB(I,J)
BETA=BETA+0.3DO
182 CONTINUE
IF (I.EQ.1)A=A-1DO
A=A+3DO
184 CONTINUE
DO 220 I=1,4
DO 210 J=1,15
DO 200 K=1,15
SUM=ODO
DO 190 L=1,7

```

```

SUM=SUM+PPTB(L,J,K)*S(I,L)
190 CONTINUE
PPT(I,J,K)=SUM/SUM1(I)
200 CONTINUE
210 CONTINUE
220 CONTINUE
DO 240 I=1,4
DO 230 J=1,7
PB(I,J)=PB(I,J)/(SUM1(I)*0.2571428D0)
230 CONTINUE
240 CONTINUE
DO 246 I=1,7
DO 244 J=1,15
DO 242 K=1,15
PPTB(I,J,K)=PPTB(I,J,K)/AJBETA(I)
242 CONTINUE
244 CONTINUE
246 CONTINUE
DO 270 I=1,7
DO 260 J=1,15
DO 250 K=1,15
PPTB1(I,J)=PPTB1(I,J)+CTE2*PPTB(I,J,K)
PPTB2(I,J)=PPTB2(I,J)+CTE1*PPTB(I,K,J)
250 CONTINUE
260 CONTINUE
270 CONTINUE
DO 300 I=1,4
DO 290 J=1,15
DO 280 K=1,15
PPT1(I,J)=PPT1(I,J)+CTE2*PPT(I,J,K)
PPT2(I,J)=PPT2(I,J)+CTE1*PPT(I,K,J)
280 CONTINUE
290 CONTINUE
300 CONTINUE
WRITE(2,410)
BETA=-0.9D0

```

```

DO 320 I=1,7
WRITE(2,420)BETA,(PB(J,I),J=1,4)
WRITE(3,*)BETA,(PB(J,I),J=1,4)
BETA=BETA+0.3D0
320 CONTINUE
WRITE(2,430)
DO 350 I=1,2
WRITE(2,440)I,I,I
DO 340 J=1,15
IF(I.EQ.2)GOTO 330
WRITE(2,450)THETA(J,I),(PPTB1(K,J),K=1,7)
GO TO 340
330 WRITE(2,450)THETA(J,I),(PPTB2(K,J),K=1,7)
340 CONTINUE
WRITE(2,460)
350 CONTINUE
DO 380 I=1,2
WRITE(2,470)I,I
DO 370 J=1,15
IF(I.EQ.2)GOTO 360
WRITE(2,480)THETA(J,I),(PPT1(K,J),K=1,4)
GO TO 370
360 WRITE(2,480)THETA(J,I),(PPT2(K,J),K=1,4)
370 CONTINUE
WRITE(2,490)
380 CONTINUE
410 FORMAT(1H1,5(/),6X,'DISTRIBUCION FINAL DE BETA P(BETA/Y)',
-/,4X,45(' '),//,5X,'BETA',7X,'A=1',5X,'A=3',5X,'A=6',5X,'A=9',
-//,4X,45(' '),//)
420 FORMAT(/,5X,F4.1,3X,4(F7.2,1X))
430 FORMAT(/,3X,45(' '))
440 FORMAT(1H1,5(/),8X,'DISTRIBUCION MARGINAL DE THETA',11,' DADO',
-' BETA P(THETA',11,' /BETA,Y)',/,3X,74(' '),//,40X,'B E T A',
-//,3X,'THETA',11,7X,'-0.9',5X,'-0.6',5X,'-0.3',6X,'0.0',6X,'0.3',
-6X,'0.6',6X,'0.9',/,3X,74(' '),//)
450 FORMAT(/,2X,F10.4,7(1X,F8.3))

```

```

460 FORMAT(5X,74(' '))
470 FORMAT(1H1,5(/),14X,'DISTRIBUCION FINAL DE THETA',I1,/,3X,60(' '),
-//,5X,'THETA',I1,11X,'A=1',8X,'A=3',8X,'A=6',8X,'A=9',//,3X,
-60(' '),//)
480 FORMAT(/,5X,F10.4,3X,4(F8.4,3X))
490 FORMAT(3X,60(' '))
500 FORMAT(11X,'OBTENCION DE DISTRIBUCIONES FINALES DE REFERENCIA')
510 FORMAT(15X,//,' - - - ESPERE UN MOMENTO POR FAVOR - - -')
END

```

C

C Cálculo de la moda de THETA1 y THETA2

C para  $\beta$  dado

C

```

SUBROUTINE RAIZ(B,T1,T2,R1,R2)
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
DIMENSION X(50,2)
COMMON X,N
E=2D0/(1D0+B)
E1=E-1D0
1 F1=0D0
F2=0D0
P1=0D0
P2=0D0
P3=0D0
DO 2 I=1,N
A=X(I,2)-T1-(T2*X(I,1))
M=-1
IF(A.LT.0D0)M=1
E2=DABS(A)**(E-2D0)
F1=F1+(M*E*(DABS(A)**E1))
F2=F2+(M*E*X(I,1)*(DABS(A)**E1))
P1=P1+(E*E1*E2)
P2=P2+(E*E1*X(I,1)*E2)
P3=P3+(E*E1*E2*(X(I,1)**2))
2 CONTINUE
D=(P1*P3)-(P2**2)

```

```

D1=(P3*(-F1)+F2*P2)/D
D2=(P1*(-F2)+F1*P2)/D
R1=T1+D1
R2=T2+D2
Q1=DABS(R1-T1)
Q2=DABS(R2-T2)
DL1=0.0001D0
IF(Q1.LT.DL1.AND.Q2.LT.DL1)GOTO 3
T1=R1
T2=R2
GO TO 1
3 RETURN
END

```

C

C Cálculo de los estimadores máximo verosímiles  
C para  $\theta_1$  y  $\theta_2$  cuando  $\beta = 0$  así como  
C intervalos de confianza para determinar el  
C área a integrar  
C

```

SUBROUTINE ETHETA(T1,T2,AT1,AT2)
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
DIMENSION X(50,2)
COMMON X,N
P=DBLE(N)
XP=0D0
YP=0D0
DO 4 I=1,N
XP=XP+X(I,1)
YP=YP+X(I,2)
4 CONTINUE
XP=XP/P
YP=YP/P
AN=0D0
AD=0D0
DO 5 I=1,N
CN=X(I,1)-XP

```

```

AN=AN+(CN*X(I,2))
AD=AD+(CN**2)
5 CONTINUE
T2=AN/AD
T1=YP-(T2*XP)
SE=ODO
SCX=ODO
DO 6 I=1,N
YE=T1+T2*X(I,1)
SE=SE+((X(I,2)-YE)**2)
SCX=SCX+(X(I,1)**2)
6 CONTINUE
RSE=(SE/P)**(0.5D0)
AT1=1.5D0*RSE*((SCX/((P-1D0)*AD))**0.5D0)
AT2=1.5D0*RSE*((P/((P-2D0)*AD))**0.5D0)
RETURN
END

```

C

C

Función GAMMA

C

```

FUNCTION GAMA(S)
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
DIMENSION H(8)
H(1)=-0.577191652D0
H(2)=0.988205891D0
H(3)=-0.897056937D0
H(4)=0.918206857D0
H(5)=-0.756704078D0
H(6)=0.482199394D0
H(7)=-0.193527818D0
H(8)=0.035868343D0
EM=1D0
GAM=ODO
1 G=S
8 IF(G.LE.2D0)GO TO 9
G=G-1D0

```

```
EM=EM*G
GO TO 8
9 IF(G.LE.1D0)G=G+1D0
  F=G-1D0
  DO 10 IK=1,8
    GAM=H(IK)*(F**IK)+GAM
10 CONTINUE
  IF(S.LT.1D0)GO TO 11
  GAMA=(1D0+GAM)*EM
  GO TO 12
11 GAMA=(1D0+GAM)/F
12 RETURN
  END
```

### PROBA 3.

```
C
C
C      Obtención de probabilidades correspondientes a P(BETA/Y)
C
C      usando la aproximación por SPLINE CUBICO usando los
C
C      datos obtenidos por medio de MONTECARLO.
C
C      Los datos se obtienen del archivo INF.DAT
C
C
C      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
C      DIMENSION X(7),Y(4,7),B(7),C(7),D(7)
C      CHARACTER R*1,ARCH*7
C      OPEN(1,FILE='DATOA.DAT',STATUS='OLD')
C      OPEN(2,FILE='INF1.DAT',STATUS='NEW')
C      DO 15 I=1,7
C      READ(1,*)X(I),(Y(J,I),J=1,4)
15  CONTINUE
C      WRITE(*,*) ' '
C      WRITE(*,*) 'NOMBRE DEL ARCHIVO'
C      READ(*,21)ARCH
C      WRITE(2,22)ARCH
17  CALL SPLINE(X,Y,B,C,D,J)
C      WRITE(*,*) ' '
C      WRITE(*,*) 'TECLEA EL LIMITE INF Y EL LIMITE SUP DE INTEGRACION'
C      READ(*,*)RLI,RLS
C      WRITE(2,*) 'LIM INF ',RLI,' LIM SUP ',RLS
C      CALL EVALUA(X,Y,B,C,D,J,RLI,RLS,PROB)
C      WRITE(*,*) ' '
C      WRITE(*,*)PROB
C      WRITE(2,*) 'PROBABILIDAD = ',PROB
C      WRITE(*,*) ' '
C      WRITE(*,*) ' OTRO VALOR DE A?'
C      READ(*,20)R
```

```

IF(R .EQ. 'S')GO TO 17
20 FORMAT(A1)
21 FORMAT(A7)
22 FORMAT(/, 'NOMBRE DEL ARCHIVO : ',A7)
STOP
END

```

C  
C

```

SUBROUTINE SPLINE(X,Y,B,C,D,J)
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
DIMENSION X(7),Y(4,7),B(7),C(7),D(7)
WRITE(*,*)'TECLEA EL VALOR DE A'
READ(*,*)J1
WRITE(2,*)'A = ',J1
J=1
IF(J1 .EQ. 3)J=2
IF(J1 .EQ. 6)J=3
IF(J1 .EQ. 9)J=4
D(1)=X(2)-X(1)
C(2)=(Y(J,2)-Y(J,1))/D(1)
DO 20 I=2,6
D(I)=X(I+1)-X(I)
B(I)=2D0*(D(I-1)+D(I))
C(I+1)=(Y(J,I+1)-Y(J,I))/D(I)
C(I)=C(I+1)-C(I)
20 CONTINUE
B(1)=-D(1)
B(7)=-D(6)
C(1)=0D0
C(7)=0D0
C(1)=C(3)/(X(4)-X(2))-C(2)/(X(3)-X(1))
C(7)=C(6)/(X(7)-X(5))-C(5)/(X(6)-X(4))
C(1)=C(1)*D(1)**2/(X(4)-X(1))
C(7)=-C(7)*D(6)**2/(X(7)-X(4))
DO 25 I=2,7
T=D(I-1)/B(I-1)

```

```

      B(I)=B(I)-T*D(I-1)
      C(I)=C(I)-T*C(I-1)
25  CONTINUE
      C(7)=C(7)/B(7)
      DO 30 IB=1,6
      I=7-IB
      C(I)=(C(I)-D(I)*C(I+1))/B(I)
30  CONTINUE
      B(7)=(Y(J,7)-Y(J,6))/D(6)+D(6)*(C(6)+2D0*C(7))
      DO 35 I=1,6
      B(I)=(Y(J,I+1)-Y(J,I))/D(I)-D(I)*(C(I+1)+2D0*C(I))
      D(I)=(C(I+1)-C(I))/D(I)
      C(I)=3D0*C(I)
35  CONTINUE
      C(7)=3*C(7)
      D(7)=D(6)
      RETURN
      END

```

```

SUBROUTINE EVALUA(X,Y,B,C,D,J,RLI,RLS,PROB)
IMPLICIT DOUBLE PRECISION (A-H,O-Z)
DIMENSION X(7),Y(4,7),B(7),C(7),D(7)
PROB=ODO
IF(RLI.GE.-1D0.AND.RLI.LE.-.6D0)M=1
IF(RLI.GE.-.6D0.AND.RLI.LE.-.3D0)M=2
IF(RLI.GE.-.3D0.AND.RLI.LE.ODO)M=3
IF(RLI.GE.ODO.AND.RLI.LE.O.3D0)M=4
IF(RLI.GE.O.3D0.AND.RLI.LE.O.6D0)M=5
IF(RLI.GE.O.6D0.AND.RLI.LE.1D0)M=6
IF(RLS.GE.-1D0.AND.RLS.LE.-.6D0)K=1
IF(RLS.GT.-.6D0.AND.RLS.LE.-.3D0)K=2
IF(RLS.GT.-.3D0.AND.RLS.LE.ODO)K=3
IF(RLS.GT.ODO.AND.RLS.LE.O.3D0)K=4
IF(RLS.GT.O.3D0.AND.RLS.LE.O.6D0)K=5

```

```
IF(RLS.GT.0.6D0.AND.RLS.LE.1D0)K=6
DO 60 I=M,K
Z=(X(I+1)-X(I))
PROB=Y(J,I)*Z+B(I)/2*Z**2+C(I)/3*Z**3+D(I)/4*Z**4+PROB
60 CONTINUE
RETURN
END
```

## 5. EJEMPLOS

Los ejemplos presentados en este trabajo, tienen como finalidad mostrar la utilidad de las distribuciones obtenidas en el análisis de normalidad utilizando los programas expuestos en el capítulo anterior.

Los pasos a seguir para la obtención de las distribuciones son los siguientes:

Se puede realizar, como ya se especificó antes, la simulación de variables aleatorias Exponencial Potencia. Para esto, se requiere ejecutar los programas GEP1 y GEP2. La información que se debe proporcionar por medio de la pantalla es la siguiente :

Al ejecutar GEP1

→ El valor de  $\beta$  ( $\in (-1,1]$ ) correspondiente al miembro de la familia que se desea simular.

Al ejecutar GEP2

→ El tamaño de muestra requerido ( $n$ ).

→ El número de variables independientes,  $x$ .

→ La varianza ( $\sigma^2$ ).

→ Una semilla inicial (para comenzar la simulación). Esta puede ser cualquier número Natural.

→ Los valores de  $x_i$ . Se recomienda que estos valores sean obtenidos aleatoriamente para generar así una muestra aleatoria.

La información se obtiene por pantalla, y corresponde a la variable  $y_i$ ,  $i=1,2,\dots,n$ .

El programa MON1 solo obtiene probabilidades finales para el caso de Regresión Lineal Simple, y se puede ejecutar cuando ya se cuenta con las observaciones (o ya se generó una muestra aleatoria). La información que se debe proporcionar es la siguiente :

Por medio de archivo

→ Las observaciones ( $x,y$ ) en formato libre.

Por medio de pantalla

→ El valor mínimo y el máximo de  $\beta$  para calcular el intervalo de integración. Esto a fin de tener flexibilidad en la obtención de las probabilidades, ya que dependiendo del valor que tenga  $\beta$ , la probabilidad va a estar acumulada en uno u otro lado del intervalo

$(-1, 1]$ .

→ La ampliación del intervalo de integración de  $\underline{\theta}$  con el mismo fin que el punto anterior.

Se obtienen cinco tablas de probabilidades :  $P(\beta|\underline{Y})$ ,  $P(\theta_1|\beta, \underline{Y})$ ,  $P(\theta_2|\beta, \underline{Y})$ ,  $P(\theta_1|\underline{Y})$  y  $P(\theta_2|\underline{Y})$ .

Finalmente, se cuenta con un programa, PROBA3, que calcula probabilidades acumuladas para  $\beta$ . La información que se proporciona de entrada es la siguiente :

- El valor de el parámetro  $a$  correspondiente a la densidad a integrar.
- El límite inferior y superior de integración. Esto para obtener  $P(a < \beta < b)$ .

EJEMPLO 1.

Se simuló una muestra aleatoria de tamaño 25, con  $\beta = 0$ ,  $\sigma^2 = 3$ ,  $\theta_1 = 2.5$ , y  $\theta_2 = 1.4$ . Los valores generados fueron :

X	Y
-9.9	-11.365
0.6	5.404
-6.9	-5.669
6.8	12.858
2.7	5.404
-5.0	-6.293
7.4	11.886
-4.9	-2.992
-8.9	-10.378
7.4	14.492
-8.4	-9.242
-8.3	-9.897
-4.7	-6.364
8.1	13.158
-7.5	-7.603
-9.6	-12.272
-8.2	-8.151
7.7	11.747
6.1	11.504
7.6	14.436
2.6	6.797
4.6	10.040
7.3	15.621
9.3	16.159
4.0	5.893

Los resultados obtenidos fueron :

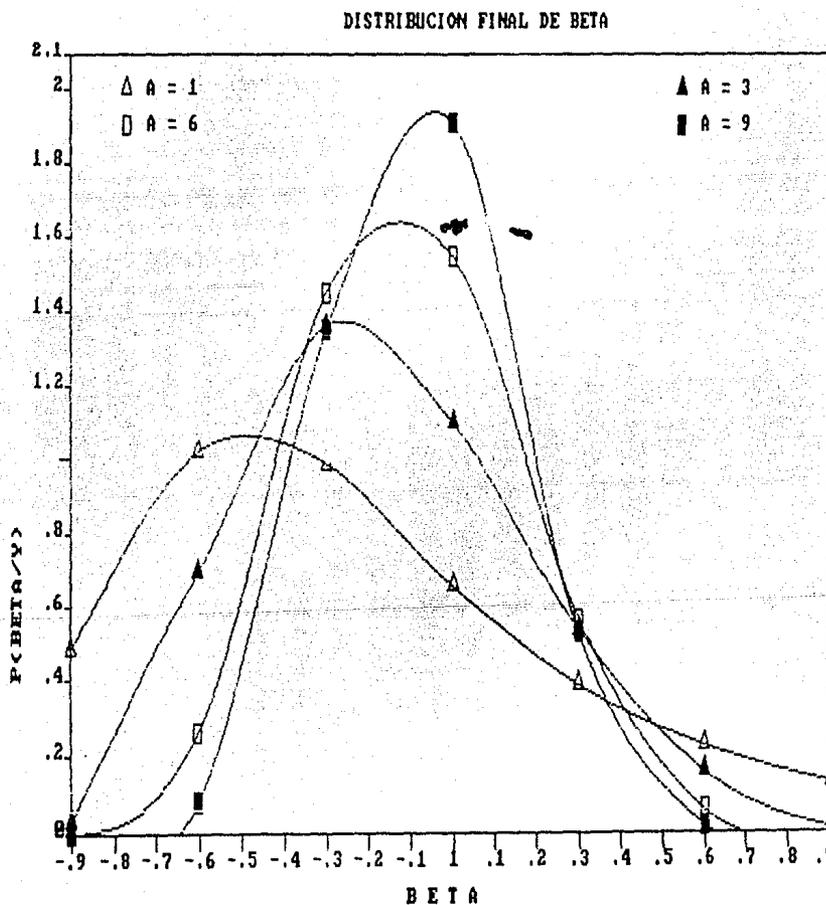
DISTRIBUCION FINAL DE BETA P(BETA|Y)

BETA	A=1	A=3	A=6	A=9
-0.9	0.48	0.03	0.00	0.00
-0.6	1.02	0.70	0.26	0.08
-0.3	0.99	1.36	1.45	1.35
0.0	0.66	1.10	1.55	1.91
0.3	0.39	0.54	0.57	0.53
0.6	0.23	0.16	0.06	0.02
0.9	0.13	0.01	0.00	0.00

NOTA : En los encabezados de las tablas y las gráficas se denota a las densidades como distribuciones.

$P(\beta|Y)$ .

Se observa que para  $A = 1$ , esto es, cuando se está utilizando distribución inicial no informativa para  $\beta$ , la muestra no proporciona información acerca de una tendencia central de límite, puesto que acumula mayor probabilidad en valores menores que cero. Además,  $P(\beta \leq -0.6) = 0.2418$  y  $P(\beta \geq 0.06) = 0.0527$ , esto indica que existe una gran posibilidad de que  $\beta \leq -0.6$ , y por tanto, se debe observar el comportamiento del parámetro de interés,  $\theta$ , para valores de  $\beta \leq -0.6$ . Esto nos hace pensar que no existe robustez en la inferencia que se realice acerca de  $\beta$ .

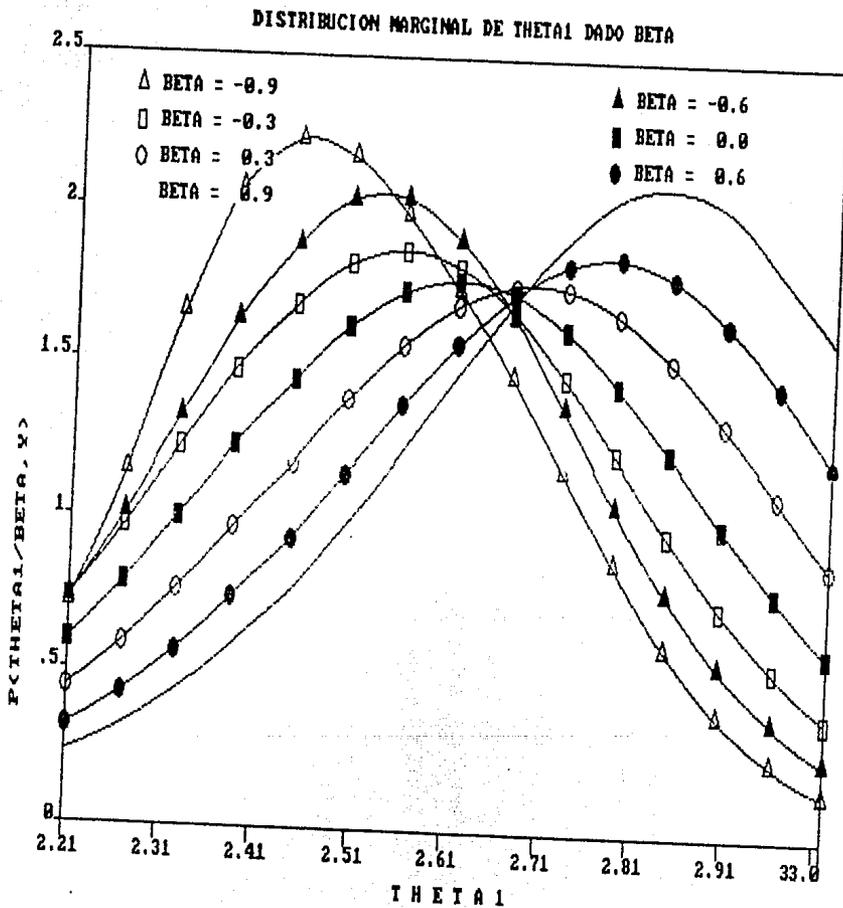


DISTRIBUCION MARGINAL DE THETA1 DADO BETA P(THETA1 |BETA,Y)

THETA1	B E T A						
	-0.9	-0.6	-0.3	0.0	0.3	0.6	0.9
2.2117	0.716	0.727	0.731	0.594	0.440	0.315	0.232
2.2694	1.145	1.006	0.968	0.786	0.590	0.428	0.319
2.3271	1.652	1.319	1.224	1.003	0.767	0.569	0.434
2.3848	2.057	1.629	1.473	1.227	0.967	0.740	0.582
2.4425	2.220	1.884	1.683	1.440	1.177	0.935	0.762
2.5001	2.166	2.030	1.823	1.616	1.381	1.148	0.975
2.5578	1.984	2.037	1.870	1.732	1.560	1.365	1.215
2.6155	1.739	1.904	1.816	1.773	1.692	1.565	1.466
2.6732	1.462	1.663	1.672	1.732	1.759	1.728	1.706
2.7309	1.167	1.364	1.462	1.616	1.752	1.831	1.915
2.7886	0.872	1.056	1.217	1.440	1.669	1.857	2.057
2.8463	0.604	0.778	0.968	1.227	1.519	1.792	2.087
2.9040	0.392	0.550	0.738	1.003	1.322	1.650	2.017
2.9617	0.244	0.374	0.542	0.786	1.100	1.446	1.851
3.0194	0.149	0.248	0.348	0.594	0.875	1.201	1.595

$$P(\theta_1 | \beta, Y).$$

La distribución de  $\theta_1$  varía mucho dependiendo del valor que tome  $\beta$ , por lo que, en principio, no se puede hacer suposición alguna acerca de la distribución de las observaciones, esto es, no hay robustez.

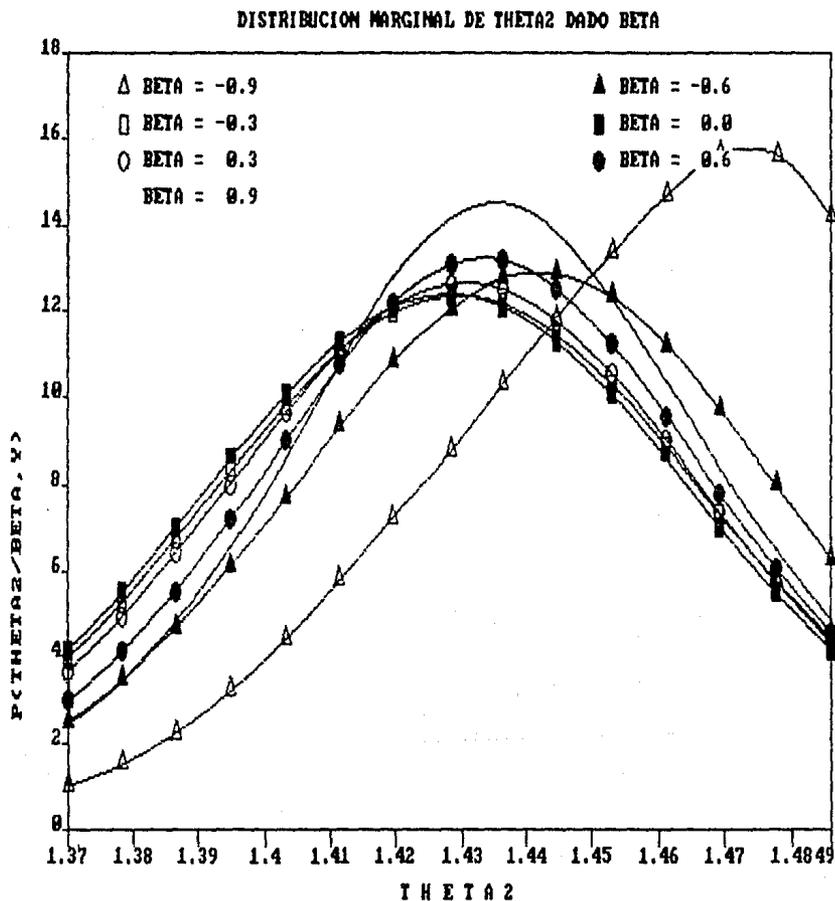


DISTRIBUCION MARGINAL DE THETA2 DADO BETA P(THETA2 |BETA,Y)

THETA2	B E T A						
	-0.9	-0.6	-0.3	0.0	0.3	0.6	0.9
1.3793	1.016	2.512	3.965	4.145	3.665	2.997	2.449
1.3876	1.528	3.472	5.265	5.489	4.931	4.129	3.464
1.3959	2.250	4.671	6.740	6.998	6.407	5.530	4.808
1.4041	3.210	6.092	8.294	8.567	8.016	7.175	6.519
1.4124	4.397	7.676	9.790	10.050	9.627	8.977	8.586
1.4207	5.758	9.308	11.065	11.279	11.056	10.731	10.768
1.4289	7.222	10.825	11.957	12.093	12.109	12.189	12.787
1.4372	8.735	12.033	12.341	12.378	12.619	13.064	14.132
1.4455	10.274	12.751	12.161	12.093	12.498	13.179	14.490
1.4537	11.822	12.853	11.441	11.279	11.764	12.517	13.822
1.4620	13.339	12.309	10.286	10.050	10.534	11.213	12.257
1.4703	14.698	11.196	8.847	8.567	8.994	9.551	10.333
1.4785	15.612	9.681	7.292	6.998	7.342	7.775	8.366
1.4868	15.589	7.971	5.774	5.489	5.742	6.060	6.490
1.4951	14.170	6.268	4.403	4.145	4.315	4.534	4.839

$$P(\theta_2 | \beta, Y)$$

La distribución de  $\theta_2$  varía menos que la anterior con respecto a  $\beta$ , pero para valores extremos de  $\beta$  ( $\leq -0.6$  ó  $\geq 0.6$ ) las distribuciones difieren mucho del resto. En este caso tampoco se tiene evidencia de la distribución de las observaciones, por lo que tampoco se detecta robustez.

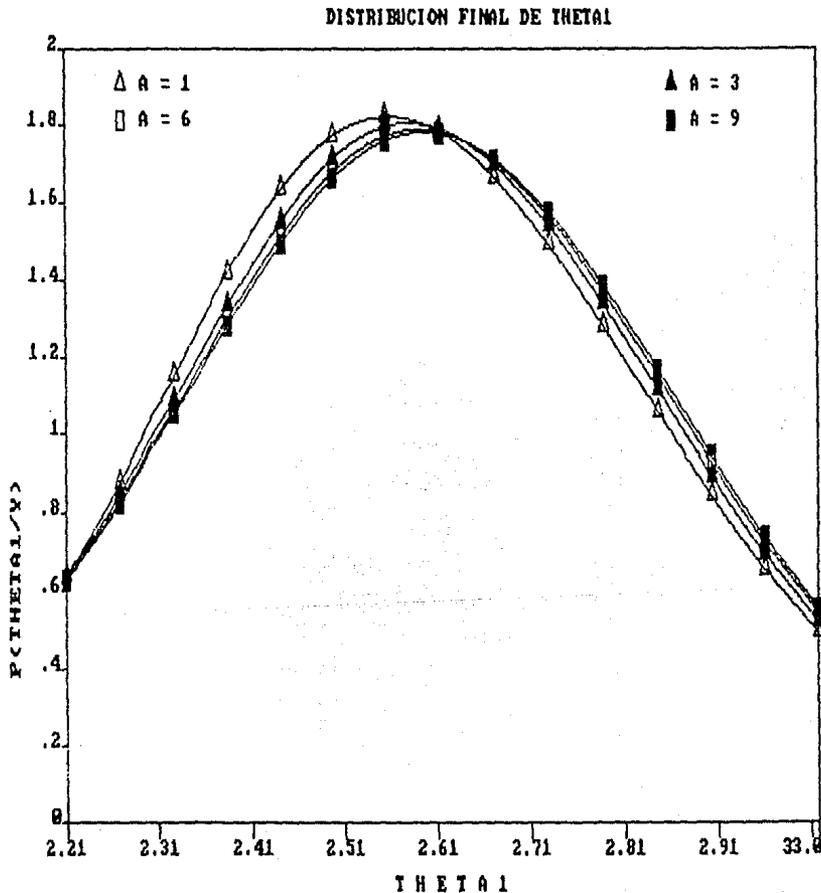


DISTRIBUCION FINAL DE THETA1

THETA1	A=1	A=3	A=6	A=9
2.2117	0.6348	0.6336	0.6272	0.6219
2.2694	0.8775	0.8494	0.8344	0.8254
2.3271	1.1533	1.0904	1.0649	1.0516
2.3848	1.4203	1.3344	1.2999	1.2829
2.4425	1.6348	1.5522	1.5137	1.4951
2.5001	1.7720	1.7140	1.6790	1.6620
2.5578	1.8217	1.7972	1.7729	1.7611
2.6155	1.7834	1.7911	1.7828	1.7787
2.6732	1.6672	1.7003	1.7092	1.7137
2.7309	1.4920	1.5416	1.5650	1.5768
2.7886	1.2813	1.3387	1.3713	1.3879
2.8463	1.0588	1.1163	1.1523	1.1706
2.9040	0.8460	0.8966	0.9306	0.9482
2.9617	0.6558	0.6950	0.7239	0.7390
3.0194	0.4930	0.5206	0.5433	0.5553

$P(\theta_1 | \underline{Y})$  y  $P(\theta_2 | \underline{Y})$

A pesar de que las distribuciones de  $\theta_1$  y  $\theta_2$  difieren mucho al variar  $\beta$ , cuando se hace el promedio sobre las distribuciones marginales de  $\theta_j$ ,  $j=1,2$  ponderadas por  $P(\beta | \underline{Y})$ , y comparar las distribuciones finales usando una distribución inicial no informativa para  $\beta$  con distribuciones que incluyen distribuciones iniciales para  $\beta$  con fuerte evidencia de normalidad en la distribución de la muestra ( $\underline{Y}$ ), éstas se aproximan considerablemente. La forma de la distribución final de  $\beta$  influye grandemente en estos resultados, puesto que asigna



## DISTRIBUCION FINAL DE THETA2

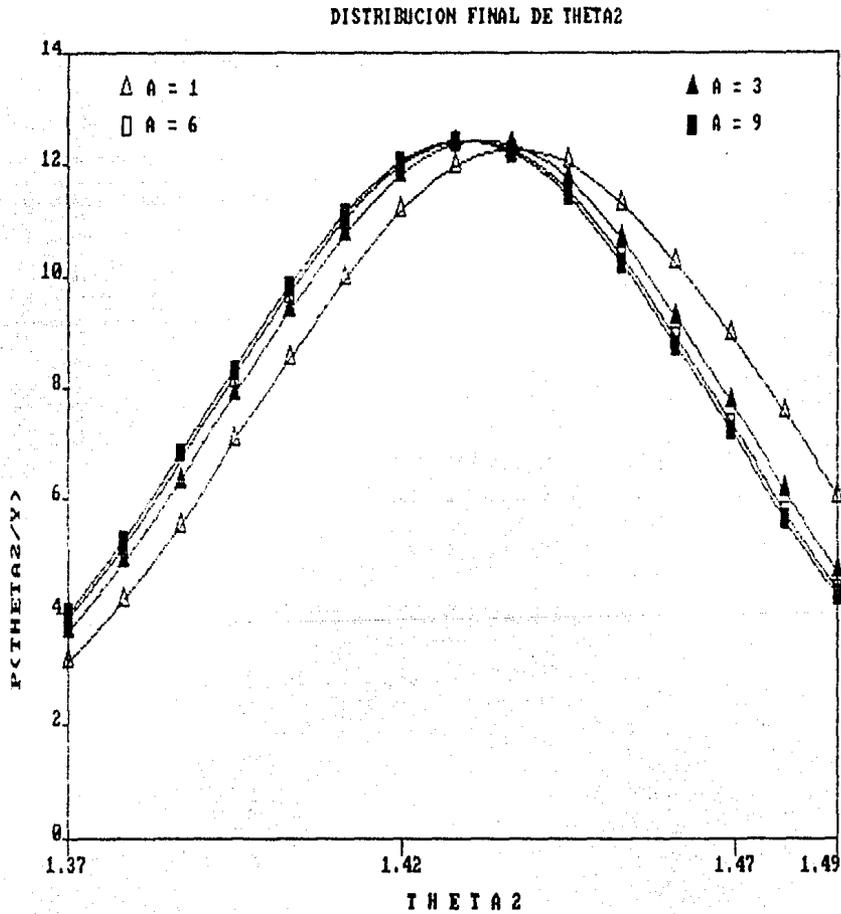
---

THETA2	A=1	A=3	A=6	A=9
1.3793	3.1145	3.6503	3.8814	3.9762
1.3876	4.2136	4.8846	5.1694	5.2857
1.3959	5.5206	6.3108	6.6384	6.7710
1.4041	6.9855	7.8524	8.1990	8.3374
1.4124	8.5130	9.3877	9.7173	9.8463
1.4207	9.9582	10.7566	11.0273	11.1295
1.4289	11.1579	11.7898	11.9618	12.0214
1.4372	11.9537	12.3410	12.3870	12.3943
1.4455	12.2434	12.3255	12.2375	12.1911
1.4537	12.0055	11.7435	11.5336	11.4410
1.4620	11.2964	10.6794	10.3764	10.2518
1.4703	10.2401	9.2832	8.9235	8.7833
1.4785	8.9604	7.7266	7.3483	7.2082
1.4868	7.5459	6.1679	5.8063	5.6783
1.4951	6.0623	4.7299	4.4131	4.3045

---

mayor probabilidad a  $\beta$  en valores menores pero cercanos a cero , no así a los valores que causan problemas.

Como conclusión se tiene que, a pesar de que existen grandes discrepancias en la forma de la distribución marginal de  $\theta_1$  y de  $\theta_2$ , esto no afecta a la distribución final, por lo que las inferencias que se puedan realizar acerca de  $\theta$  son insensitivas a cambios bajo no normalidad y por tanto la suposición de normalidad es aceptable. Esto es, a fin de cuentas sí existe robustez en las inferencias que se realicen acerca de  $\theta$ .



EJEMPLO 2.

Se simuló una muestra aleatoria de tamaño 15 con  $\beta = -0.8$ ,  $\sigma^2 = 0.35$ ,  $\theta_1 = 2.5$ ,  $\theta_2 = 4.5$ . Los valores generados fueron :

X	Y
1.5	9.25
-9.9	-42.042
7.0	32.918
7.1	33.598
4.4	21.517
-9.1	-39.557
-5.4	-21.612
-6.3	-26.636
-4.6	-18.2
0.5	4.755
2.8	14.654
-0.8	-0.52
2.9	16.165
6.1	29.432
-8.7	-37.139

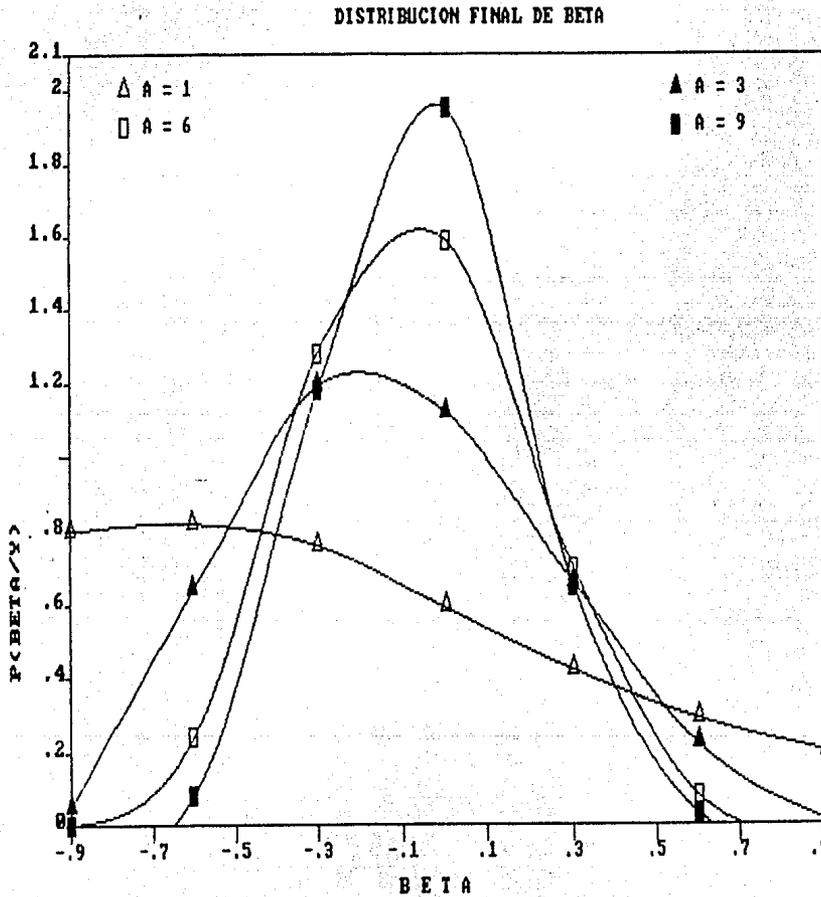
Los resultados obtenidos fueron :

DISTRIBUCION FINAL DE BETA P(BETA/Y)

BETA	A=1	A=3	A=6	A=9
-0.9	0.80	0.05	0.00	0.00
-0.6	0.82	0.64	0.24	0.08
-0.3	0.76	1.19	1.28	1.18
0.0	0.59	1.12	1.59	1.95
0.3	0.42	0.66	0.70	0.65
0.6	0.29	0.22	0.08	0.03
0.9	0.20	0.01	0.00	0.00

$P(\beta|Y)$

Para el caso de interés ( $A = 1$ ), la probabilidad de que  $\beta$  tome valores pequeños ( $\beta \leq -0.6$ ) es grande ( $\approx 0.25$ ), no es de extrañar este hecho si consideramos que la muestra proviene de una distribución con  $\beta = -0.8$ .

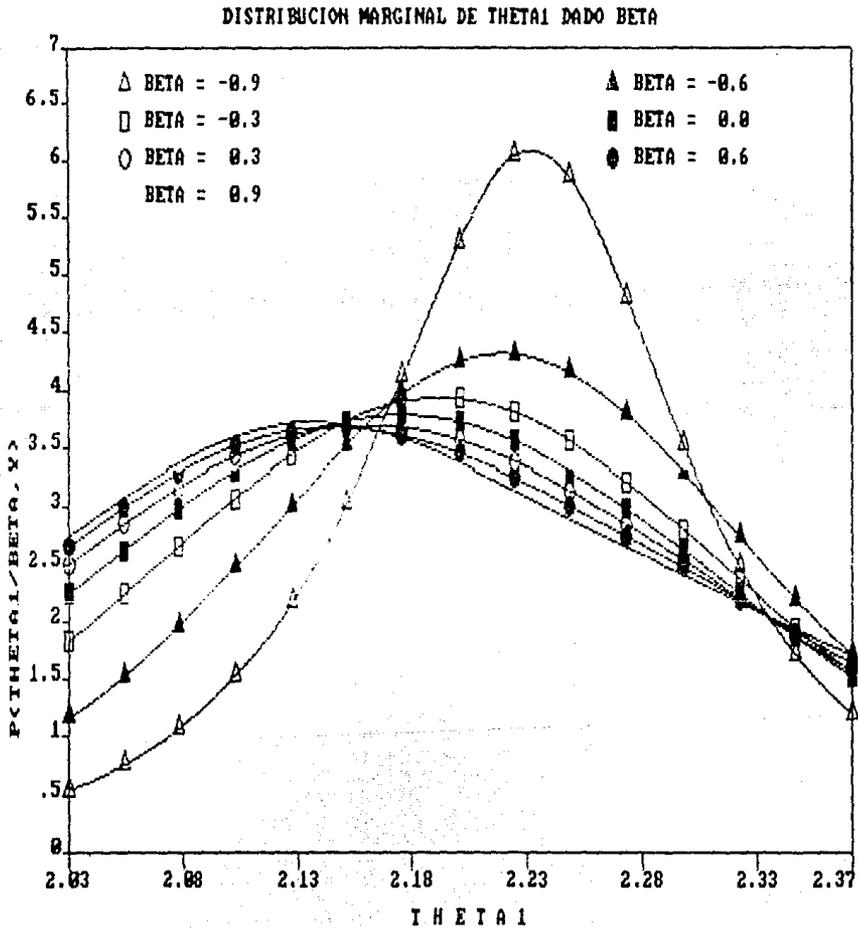


DISTRIBUCION MARGINAL DE THETA1 DADO BETA  $P(\text{THETA1} | \text{BETA}, Y)$

THETA1	B E T A						
	-0.9	-0.6	-0.3	0.0	0.3	0.6	0.9
2.0344	0.542	1.166	1.829	2.240	2.500	2.658	2.747
2.0588	0.761	1.527	2.234	2.620	2.852	2.996	3.081
2.0831	1.075	1.965	2.658	2.984	3.170	3.292	3.373
2.1075	1.524	2.466	3.074	3.308	3.429	3.518	3.594
2.1318	2.158	3.004	3.444	3.563	3.606	3.651	3.714
2.1561	3.022	3.526	3.730	3.726	3.688	3.674	3.698
2.1805	4.113	3.963	3.898	3.782	3.675	3.605	3.584
2.2048	5.271	4.243	3.925	3.724	3.569	3.454	3.377
2.2291	6.034	4.310	3.808	3.560	3.385	3.249	3.138
2.2535	5.849	4.149	3.560	3.304	3.139	3.007	2.892
2.2778	4.787	3.788	3.210	2.980	2.847	2.742	2.645
2.3021	3.514	3.292	2.799	2.615	2.528	2.462	2.397
2.3265	2.468	2.737	2.365	2.236	2.199	2.181	2.158
2.3508	1.716	2.192	1.942	1.866	1.875	1.904	1.930
2.3751	1.196	1.702	1.555	1.523	1.568	1.635	1.701

$$P(\theta_1 | \beta, Y)$$

La distribución de  $\theta_1$  es muy variable dependiendo del valor de  $\beta$ , en particular, se observan grandes diferencias en las distribuciones cuando  $\beta = -0.9$  y  $\beta = 0.6$ , además alrededor de  $\beta = -0.9$  se acumula la mayor probabilidad (gráfica anterior), por lo que no hay evidencia de que la suposición de normalidad sea adecuada, aún más, no existe evidencia de robustez en las inferencias que se realicen con  $\theta_1$ .

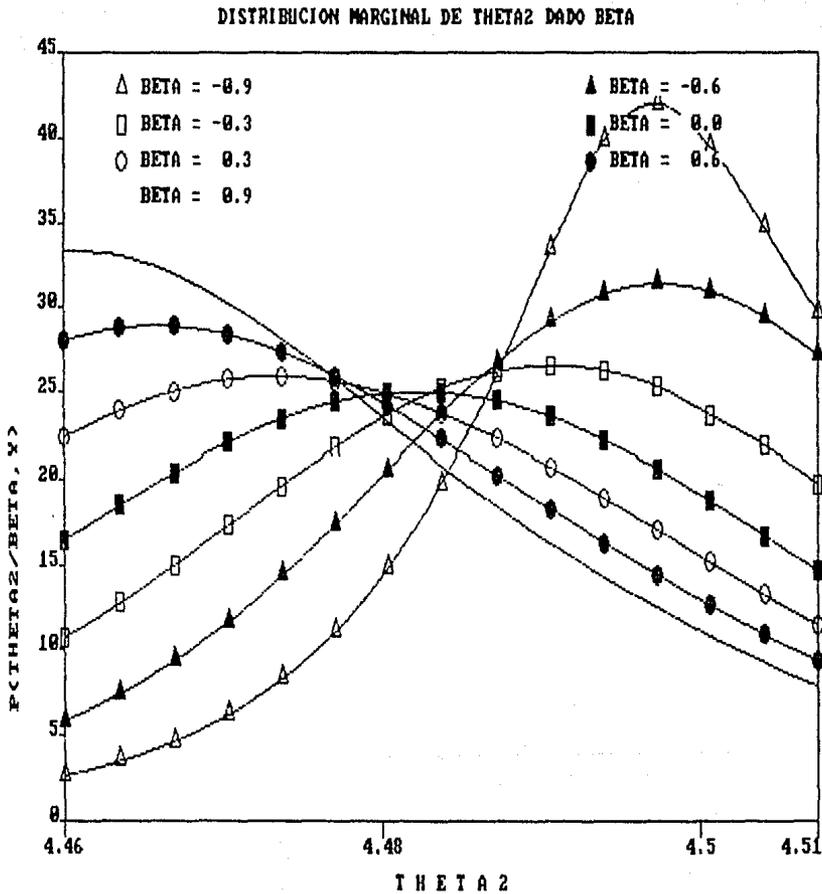


DISTRIBUCION MARGINAL DE THETA2 DADO BETA P(THETA2 /BETA,Y)

THETA2	B E T A						
	-0.9	-0.6	-0.3	0.0	0.3	0.6	0.9
4.4666	2.651	5.875	10.731	16.447	22.372	28.066	33.346
4.4700	3.518	7.456	12.778	18.475	23.925	28.820	33.086
4.4734	4.677	9.385	14.993	20.388	25.074	28.931	32.018
4.4768	6.227	11.688	17.306	22.085	25.758	28.435	30.319
4.4802	8.305	14.363	19.619	23.472	25.953	27.426	28.217
4.4836	11.093	17.360	21.810	24.461	25.667	25.996	25.847
4.4869	14.827	20.565	23.740	24.991	24.933	24.249	23.334
4.4903	19.782	23.787	25.266	25.025	23.815	22.308	20.820
4.4937	26.143	26.770	26.264	24.562	22.393	20.289	18.435
4.4971	33.514	29.220	26.642	23.633	20.748	18.287	16.275
4.5005	39.867	30.859	26.360	22.298	18.956	16.343	14.326
4.5039	41.998	31.487	25.438	20.640	17.080	14.472	12.547
4.5072	39.484	31.030	23.948	18.755	15.176	12.672	10.871
4.5106	34.792	29.557	22.008	16.741	13.303	10.970	9.312
4.5140	29.786	27.262	19.762	14.691	11.511	9.399	7.911

$$P(\theta_2 | \beta, \underline{y})$$

Las distribuciones tienen grandes diferencias con respecto al valor que toma  $\beta$ , para toda  $\beta$ , por lo que no existe robustez para  $\theta_2$ .



## DISTRIBUCION FINAL DE THETA1

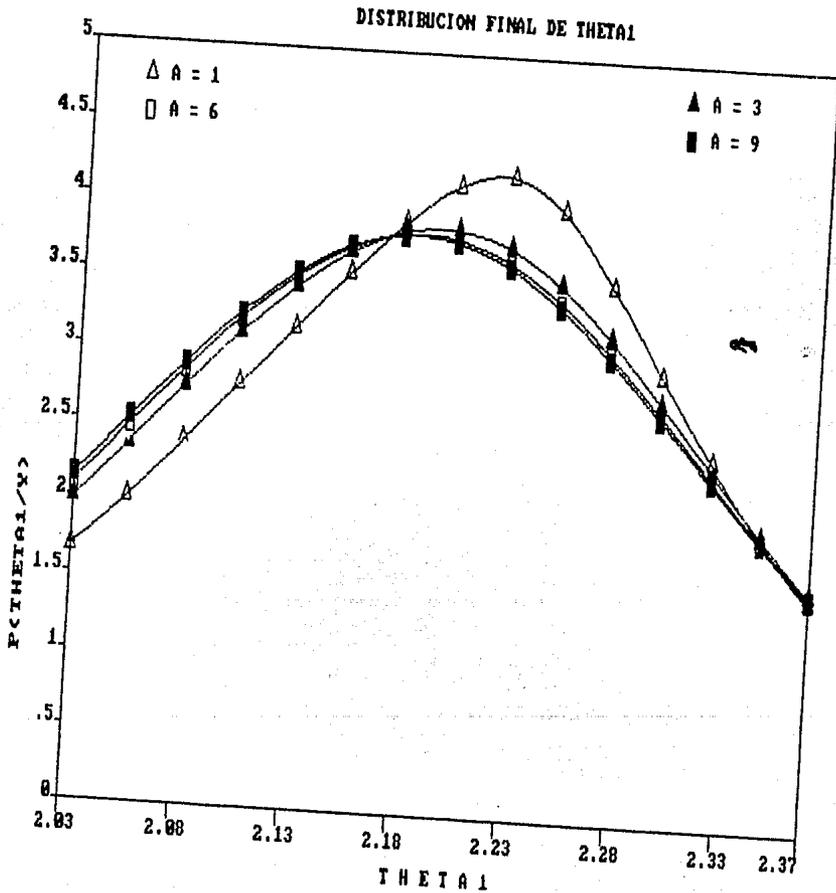
---

THETA1	A=1	A=3	A=6	A=9
2.0344	1.6678	1.9846	2.0954	2.1405
2.0588	2.0070	2.3592	2.4760	2.5221
2.0831	2.3745	2.7412	2.8547	2.8981
2.1075	2.7602	3.1069	3.2055	3.2414
2.1318	3.1512	3.4281	3.4990	3.5233
2.1561	3.5302	3.6752	3.7070	3.7165
2.1805	3.8762	3.8226	3.8077	3.8014
2.2048	4.1369	3.8493	3.7884	3.7679
2.2291	4.2124	3.7442	3.6491	3.6186
2.2535	3.9952	3.5095	3.4034	3.3685
2.2778	3.5185	3.1705	3.0759	3.0419
2.3021	2.9473	2.7703	2.6978	2.6686
2.3265	2.4031	2.3509	2.3010	2.2785
2.3508	1.9263	1.9445	1.9134	1.8974
2.3751	1.5223	1.5721	1.5548	1.5444

---

$$P(\theta_1 | Y) \text{ y } P(\theta_2 | Y)$$

Observando las gráficas puede verse, que las distribuciones finales de  $\theta$  son muy diferentes cuando  $A = 1$  y cuando  $A = 9$ , esto quiere decir que las inferencias que se realicen acerca de  $\theta$  serán diferentes dependiendo del valor que se suponga para  $\beta$ . Se ve que en general, no podemos suponer normalidad de los datos, puesto que si la

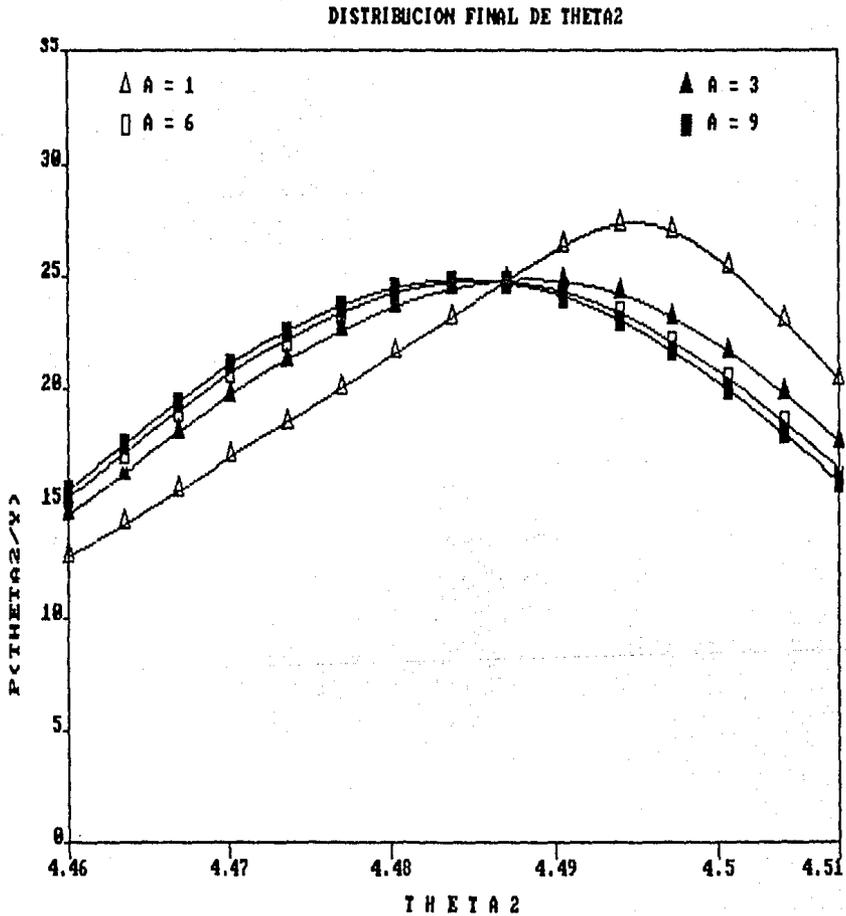


DISTRIBUCION FINAL DE THETA2

THETA2	A=1	A=3	A=6	A=9
4.4666	12.6165	14.4980	15.2410	15.5701
4.4700	14.0520	16.2816	17.1349	17.5070
4.4734	15.5033	18.0364	18.9709	19.3715
4.4768	16.9736	19.7112	20.6775	21.0837
4.4802	18.4722	21.2519	22.1808	22.5619
4.4836	20.0034	22.5975	23.4057	23.7267
4.4869	21.5722	23.6841	24.2820	24.5081
4.4903	23.1847	24.4492	24.7513	24.8530
4.4937	24.8239	24.8348	24.7728	24.7312
4.4971	26.3535	24.7880	24.3282	24.1394
4.5005	27.3095	24.2551	23.4262	23.1029
4.5039	27.0133	23.1996	22.1059	21.6758
4.5072	25.3944	21.6626	20.4361	19.9366
4.5106	23.0212	19.7659	18.5115	17.9817
4.5140	20.3702	17.6479	16.4389	15.9141

distribución final de  $(\theta_1, \theta_2)$  cambia dependiendo del valor que toma  $\beta$ , las inferencias que se realicen acerca de dicho parámetro serán sensitivas a cambios en  $\beta$ . No hay robustez, y por tanto, se puede llegar a conclusiones equivocadas.

En este caso, la muestra proviene de una distribución exponencial potencia con  $\beta = -0.8$ , y si supusiéramos normalidad en las observaciones, se llegaría a resultados falsos.



EJEMPLO 3.

El presente ejemplo y el siguiente fueron tomados de Draper (1981) .

Se realizó un estudio sobre el efecto en el producto de un proceso químico, obteniéndose los siguientes resultados (en forma codificada)

X	Y
13.3	3.5
16.9	5.1
19.9	4.8
23.2	6.7
26.3	6.0
30.1	9.5
42.6	8.1

Los resultados obtenidos fueron :

DISTRIBUCION FINAL DE BETA P(BETA/Y)

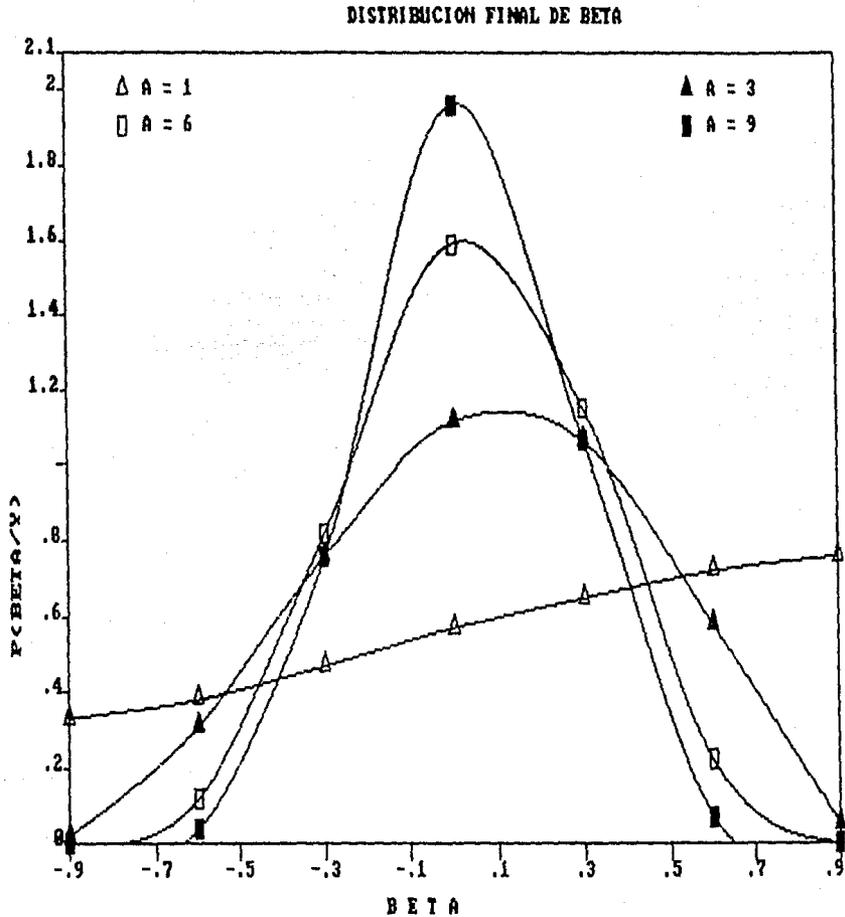
---

BETA	A=1	A=3	A=6	A=9
-0.9	0.33	0.02	0.00	0.00
-0.6	0.38	0.31	0.12	0.04
-0.3	0.47	0.76	0.82	0.76
0.0	0.57	1.11	1.59	1.96
0.3	0.65	1.06	1.15	1.06
0.6	0.72	0.58	0.22	0.07
0.9	0.76	0.05	0.00	0.00

---

$P(\beta|Y)$

Para  $A=1$  se observa que la distribución final de  $\beta$  tiene tendencia casi uniforme, y se acumula más probabilidad para valores grandes de  $\beta$ , aunque esto no es muy marcado.

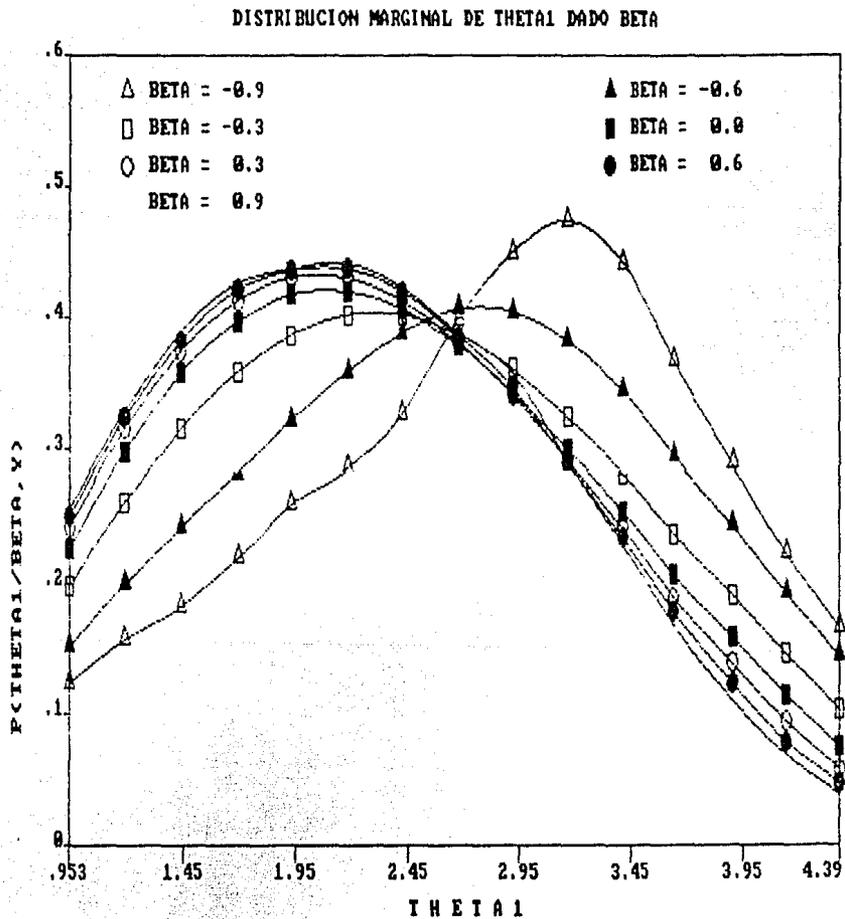


DISTRIBUCION MARGINAL DE THETA 1 DADO BETA  $P(\text{THETA1}|\text{BETA},Y)$

THETA1	B E T A						
	-0.9	-0.6	-0.3	0.0	0.3	0.6	0.9
0.953	0.123	0.151	0.197	0.226	0.241	0.250	0.256
1.199	0.157	0.199	0.261	0.299	0.316	0.324	0.329
1.444	0.182	0.242	0.316	0.358	0.375	0.383	0.390
1.690	0.218	0.283	0.358	0.397	0.413	0.421	0.426
1.936	0.259	0.322	0.387	0.418	0.431	0.436	0.437
2.182	0.286	0.359	0.403	0.420	0.430	0.437	0.441
2.428	0.327	0.389	0.403	0.407	0.413	0.420	0.423
2.674	0.399	0.407	0.388	0.380	0.382	0.387	0.390
2.920	0.449	0.405	0.361	0.343	0.341	0.347	0.358
3.165	0.473	0.383	0.324	0.300	0.292	0.290	0.291
3.411	0.442	0.344	0.282	0.254	0.241	0.234	0.229
3.657	0.367	0.295	0.237	0.206	0.189	0.178	0.168
3.903	0.289	0.243	0.191	0.159	0.139	0.124	0.111
4.149	0.222	0.192	0.146	0.115	0.095	0.080	0.070
4.395	0.166	0.144	0.104	0.076	0.059	0.048	0.040

$$P(\theta_1 | \beta, Y)$$

Se observa que para valores grandes de  $\beta$  ( $\beta \geq 0$ ) las distribuciones al parecer no difieren mucho entre sí, pero para valores de  $\beta \leq -0.3$ , las distribuciones se dispersan más. Esto quiere decir que no existe robustez en las inferencias que se realicen con respecto a  $\theta_1$ , puesto que las inferencias que se pueden realizar dependen del valor que se suponga para  $\theta_1$ .



DISTRIBUCION MARGINAL DE THETA2 DADO BETA P(THETA2 | BETA, Y)

B E T A

THETA2	-0.9	-0.6	-0.3	0.0	0.3	0.6	0.9
0.084	0.905	1.360	1.964	2.424	2.662	2.734	2.709
0.093	1.832	2.268	2.931	3.494	3.855	4.029	4.067
0.103	3.056	3.474	4.076	4.666	5.146	5.473	5.657
0.112	4.420	4.874	5.320	5.883	6.482	7.012	7.431
0.122	6.058	6.414	6.615	7.109	7.798	8.530	9.234
0.131	8.104	8.074	7.926	8.299	8.976	9.730	10.418
0.141	10.674	9.772	9.194	9.382	9.914	10.529	11.092
0.150	13.737	11.283	10.307	10.260	10.527	10.851	11.151
0.160	15.329	12.252	11.115	10.830	10.783	10.787	10.798
0.169	13.388	12.381	11.469	10.993	10.634	10.338	10.122
0.179	10.775	11.643	11.258	10.672	10.051	9.497	9.056
0.188	8.895	10.269	10.424	9.803	9.032	8.354	7.826
0.198	7.045	8.497	8.946	8.351	7.519	6.761	6.148
0.207	5.332	6.370	6.849	6.393	5.640	4.891	4.238
0.217	3.270	3.888	4.426	4.263	3.800	3.305	2.873

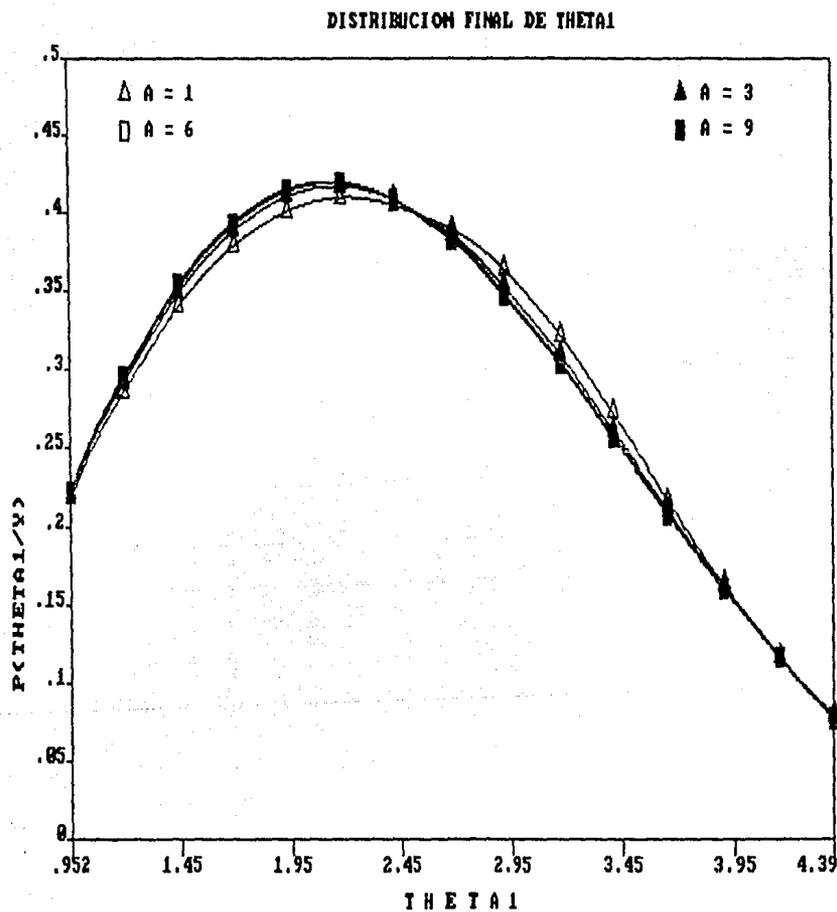


## DISTRIBUCION FINAL DE THETA1

THETA1	A=1	A=3	A=6	A=9
0.9528	0.2189	0.2217	0.2232	0.2239
1.1986	0.2857	0.2917	0.2946	0.2959
1.4445	0.3401	0.3484	0.3521	0.3537
1.6903	0.3784	0.3877	0.3916	0.3933
1.9361	0.4004	0.4099	0.4134	0.4148
2.1820	0.4095	0.4168	0.4187	0.4194
2.4278	0.4044	0.4081	0.4081	0.4079
2.6737	0.3889	0.3856	0.3835	0.3825
2.9195	0.3638	0.3524	0.3483	0.3467
3.1654	0.3211	0.3086	0.3046	0.3031
3.4112	0.2714	0.2606	0.2573	0.2561
3.6571	0.2169	0.2107	0.2087	0.2079
3.9029	0.1630	0.1615	0.1606	0.1602
4.1488	0.1169	0.1166	0.1160	0.1157
4.3946	0.0787	0.0781	0.0774	0.0771

$$P(\theta_1 | Y)$$

Las distribuciones finales de  $\theta_1$  son muy parecidas entre sí, por por tanto, aunque marginalmente algunas distribuciones tienen tendencia diferente ( $\beta \leq -0.3$ ) en promedio tienen comportamiento similar al caso  $\beta = 0$ , por lo que la suposición de normalidad no afectaría las inferencias que se realicen sobre  $\theta_1$ .



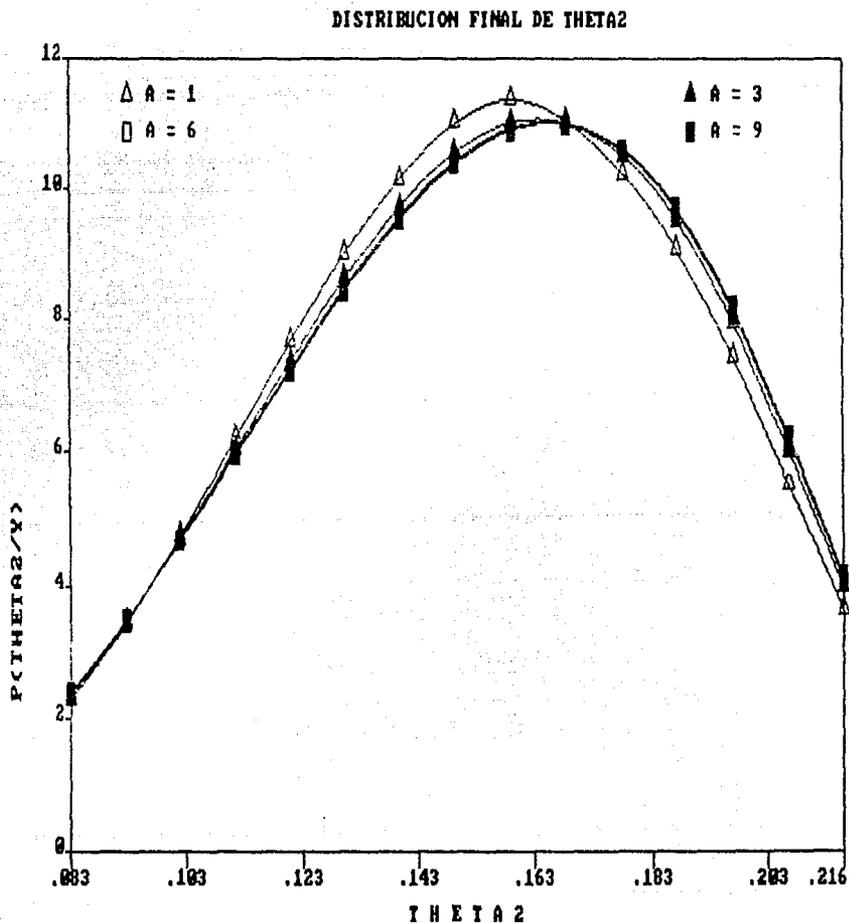
DISTRIBUCION FINAL DE THETA2

THETA2	A=1	A=3	A=6	A=9
0.0837	2.2861	2.3555	2.3827	2.3944
0.0932	3.4342	3.4627	3.4750	3.4805
0.1027	4.7629	4.7109	4.6925	4.6848
0.1122	6.2027	6.0365	5.9738	5.9472
0.1217	7.6857	7.3790	7.2663	7.2193
0.1312	9.0089	8.6332	8.4931	8.4346
0.1407	10.1460	9.7229	9.5749	9.5149
0.1502	10.9948	10.5441	10.4124	10.3628
0.1597	11.3700	11.0060	10.9164	10.8857
0.1692	11.0411	11.0030	10.9922	10.9893
0.1787	10.2087	10.4980	10.5758	10.6050
0.1882	9.0605	9.5034	9.6399	9.6925
0.1977	7.4588	7.9778	8.1468	8.2129
0.2072	5.5274	6.0158	6.1823	6.2489
0.2167	3.6327	3.9719	4.0964	4.1477

$$P(\theta_2 | Y)$$

La distribución para  $A = 1$  es ligeramente diferente a las demás, es decir, a las distribuciones con un mayor grado de suposición de normalidad, pero sus medias están muy cercanas entre sí.

Por lo que anterior, se puede decir que las inferencias que se realicen acerca de  $\theta_2$  serán robustas; por lo que la suposición de normalidad es aplicable, puesto que no afecta a las inferencias que se realicen para  $(\theta_1, \theta_2)$ , con la reserva de que si  $\theta_2$  es un parámetro muy importante para el estudio, debe analizarse más a fondo la naturaleza de las observaciones para no llegar a resultados erróneos.



EJEMPLO 4.

Se piensa que el número de latas dañadas en un cargamento de cajas que contienen a dichas latas es una función de la velocidad de la caja al momento del impacto. 13 cajas fueron seleccionadas al azar para examinar si esta suposición es cierta. Los datos obtenidos fueron los siguientes:

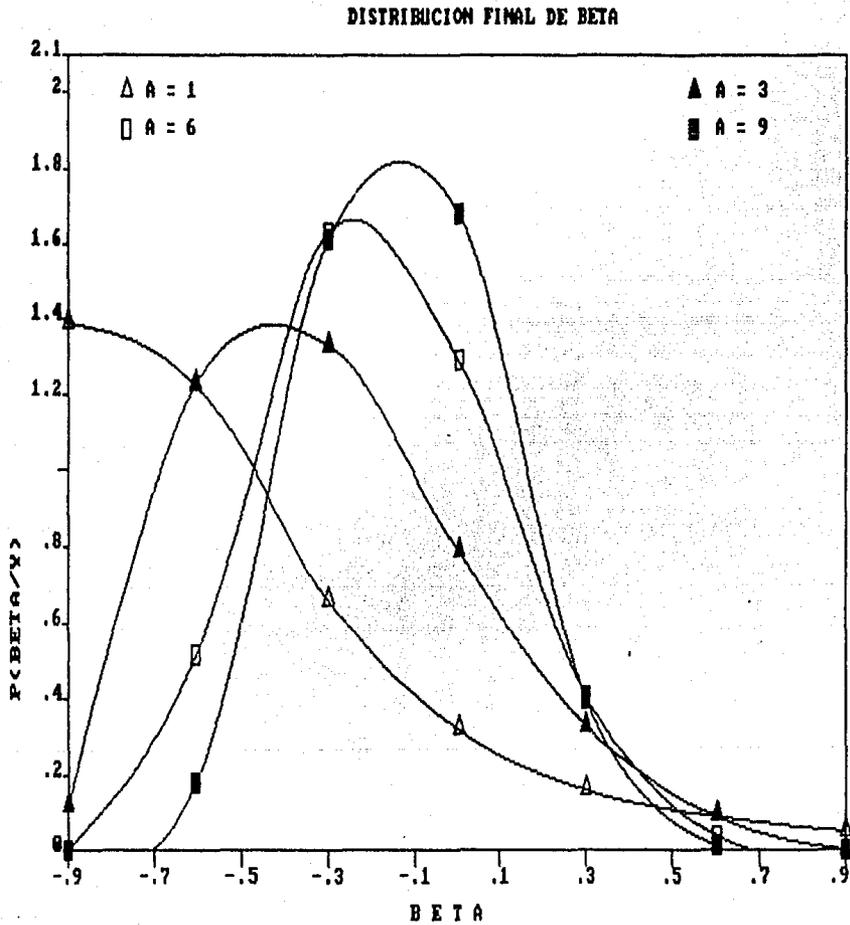
X (VELOCIDAD DE LAS LATAS AL IMPACTO)	Y (NUMERO DE LATAS DAÑADAS )
460	0.3
450	0.3
440	0.4
430	0.4
420	0.6
410	0.5
450	0.5
440	0.6
430	0.6
420	0.6
410	0.7
400	0.6
420	0.6
410	0.6
400	0.6

Los resultados obtenidos fueron :

DISTRIBUCION FINAL DE BETA P(BETA/Y)				
BETA	A=1	A=3	A=6	A=9
-0.9	1.39	0.12	0.00	0.00
-0.6	1.22	1.23	0.52	0.18
-0.3	0.66	1.33	1.63	1.61
0.0	0.32	0.79	1.28	1.68
0.3	0.16	0.33	0.41	0.40
0.6	0.09	0.09	0.04	0.01
0.9	0.05	0.00	0.00	0.00

$P(\beta|Y)$

Para  $A = 1$ , la probabilidad de que  $\beta$  tome valores pequeños ( $\beta \leq -0.6$ ) es muy grande,  $P(\beta \leq -0.6) = 0.41$ , por lo que se debe tener cuidado con la forma de las distribuciones de  $\theta_j$  ( $j=1,2$ ) en estos casos.

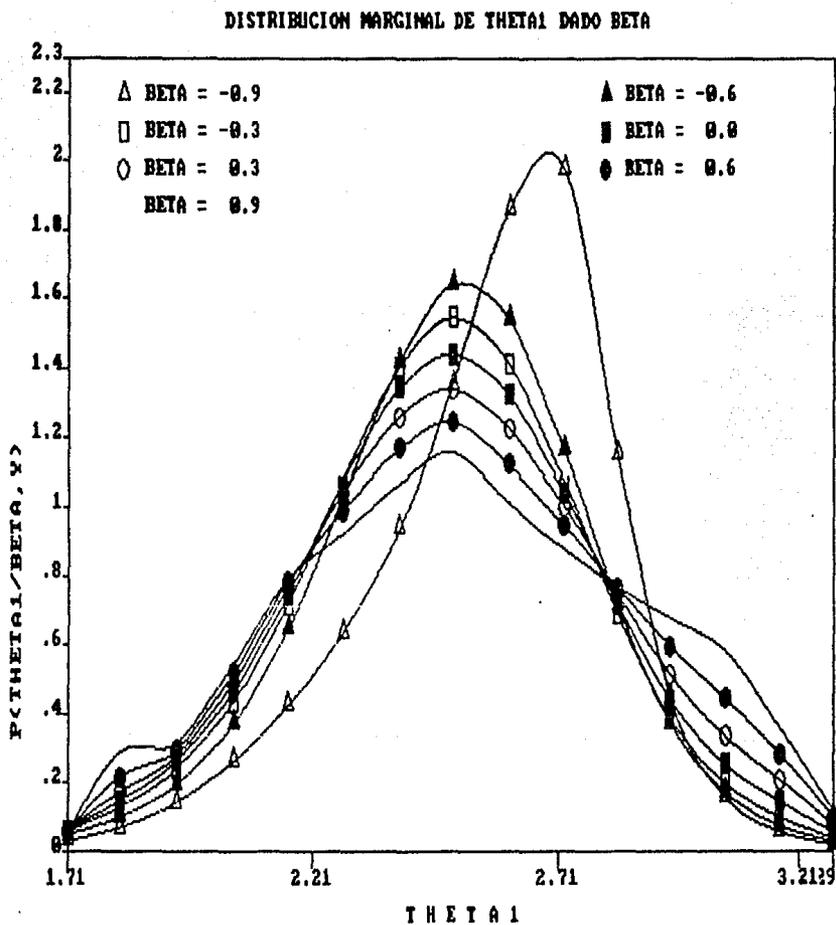


DISTRIBUCION MARGINAL DE THETA1 DADO BETA  $P(\text{THETA1} | \text{BETA}, Y)$

THETA1	B E T A						
	-0.9	-0.6	-0.3	0.0	0.3	0.6	0.9
1.7183	0.031	0.050	0.062	0.064	0.062	0.057	0.051
1.8306	0.071	0.102	0.131	0.150	0.175	0.219	0.289
1.9429	0.145	0.198	0.244	0.266	0.281	0.297	0.316
2.0553	0.262	0.370	0.434	0.464	0.490	0.518	0.544
2.1676	0.423	0.648	0.720	0.747	0.768	0.784	0.790
2.2799	0.637	1.027	1.079	1.074	1.042	0.990	0.923
2.3923	0.937	1.419	1.406	1.343	1.261	1.167	1.068
2.5046	1.356	1.642	1.547	1.443	1.340	1.245	1.157
2.6169	1.860	1.542	1.415	1.323	1.225	1.120	1.012
2.7293	1.975	1.163	1.074	1.042	1.003	0.947	0.878
2.8416	1.150	0.717	0.690	0.717	0.747	0.764	0.766
2.9539	0.446	0.375	0.389	0.442	0.515	0.596	0.677
3.0663	0.162	0.176	0.201	0.255	0.336	0.446	0.583
3.1786	0.061	0.078	0.103	0.149	0.212	0.286	0.361
3.2909	0.023	0.031	0.042	0.060	0.081	0.103	0.122

$$P(\theta_1 | \beta, Y) \text{ y } P(\theta_2 | \beta, Y)$$

El efecto de  $\beta$  sobre estas distribuciones es básicamente en las varianzas, ya que para los diferentes valores de  $\beta$  las



DISTRIBUCION MARGINAL DE THETA2 DADO BETA P(THETA2 /BETA,Y)

THETA2	B E T A						
	-0.9	-0.6	-0.3	0.0	0.3	0.6	0.9
-0.0064	10.998	15.609	23.941	39.021	59.541	82.275	103.702
-0.0061	27.546	35.089	45.143	63.394	89.217	120.860	155.244
-0.0059	73.411	79.546	90.423	113.130	147.327	194.915	257.280
-0.0056	202.671	170.256	175.879	199.056	229.842	262.672	293.470
-0.0054	522.843	325.578	313.206	324.595	336.721	342.058	339.178
-0.0051	897.620	528.644	488.091	472.944	454.255	427.590	394.998
-0.0049	845.539	700.791	643.042	600.841	556.007	507.658	457.944
-0.0046	616.219	746.143	703.031	655.397	608.465	564.891	524.133
-0.0044	425.785	644.921	638.868	610.334	572.531	529.572	484.000
-0.0041	289.725	466.826	490.272	487.456	472.797	448.322	416.826
-0.0039	192.028	294.414	327.068	338.667	347.324	352.994	353.890
-0.0037	119.180	168.157	196.953	209.366	219.452	229.114	237.304
-0.0034	65.866	89.973	110.300	118.915	122.413	123.526	123.359
-0.0032	32.073	46.399	58.790	65.224	70.614	78.050	88.356
-0.0029	14.497	23.655	30.994	37.662	49.494	71.502	106.316

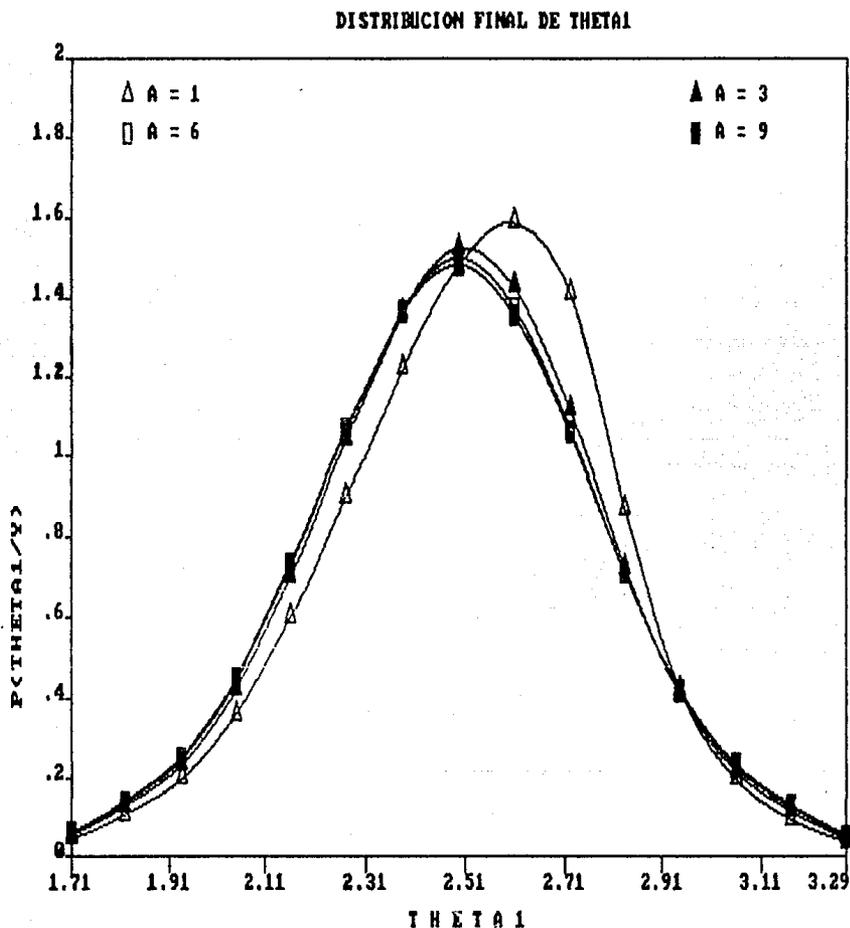


DISTRIBUCION FINAL DE THETA1

THETA1	A=1	A=3	A=6	A=9
1.7183	0.0471	0.0576	0.0612	0.0625
1.8306	0.1079	0.1298	0.1388	0.1426
1.9429	0.1999	0.2355	0.2496	0.2555
2.0553	0.3609	0.4214	0.4419	0.4501
2.1676	0.5979	0.6991	0.7249	0.7336
2.2799	0.8992	1.0423	1.0654	1.0702
2.3923	1.2216	1.3644	1.3693	1.3637
2.5046	1.4792	1.5246	1.5005	1.4837
2.6169	1.5860	1.4270	1.3790	1.3604
2.7293	1.4122	1.1149	1.0672	1.0566
2.8416	0.8697	0.7249	0.7092	0.7089
2.9539	0.4230	0.4128	0.4197	0.4249
3.0663	0.1998	0.2205	0.2323	0.2383
3.1786	0.0959	0.1167	0.1278	0.1334
3.2909	0.0376	0.0466	0.0513	0.0537

$P(\theta_1 | \underline{Y})$  y  $P(\theta_2 | \underline{Y})$

Se observa que las distribuciones para  $A = 1$  son diferentes de las distribuciones con fuerte suposición de normalidad, y como existe una gran probabilidad de que el verdadero valor de



## DISTRIBUCION FINAL DE THETA2

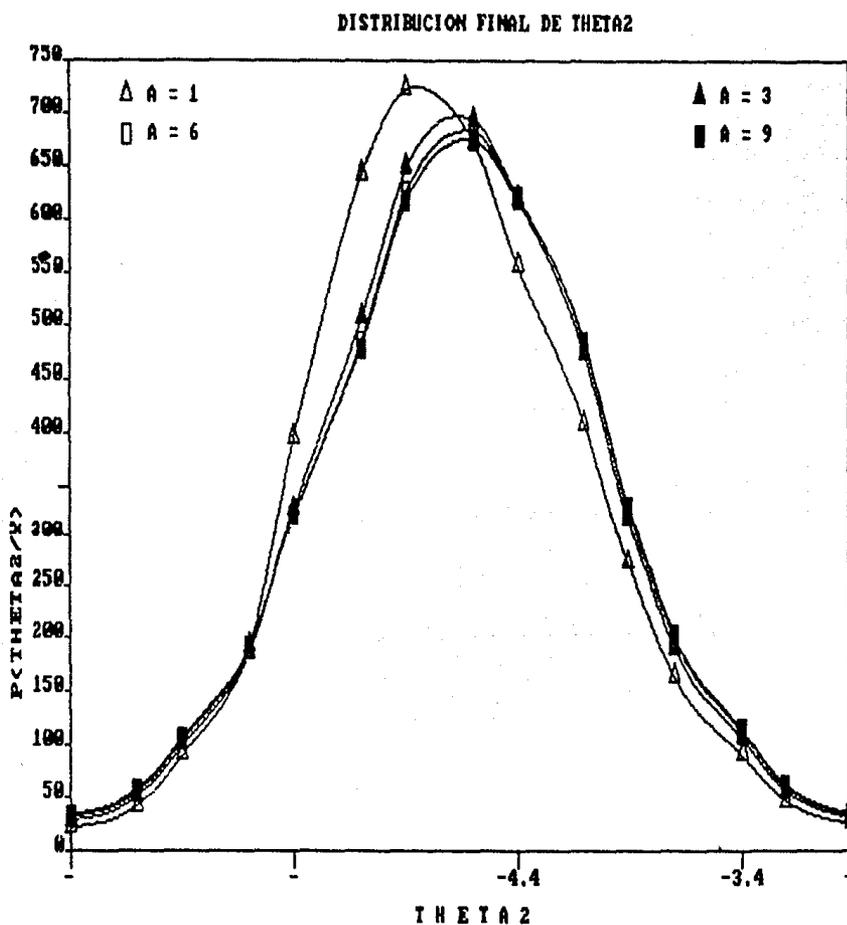
---

THETA2	A=1	A=3	A=6	A=9
-0.0064	21.8162	28.4125	32.0915	33.9520
-0.0061	42.2263	50.7191	55.1636	57.3791
-0.0059	89.7584	98.4736	103.4290	105.9703
-0.0056	191.3541	186.3507	189.2830	191.5099
-0.0054	394.7658	328.7063	321.4531	321.2316
-0.0051	641.5938	506.3428	484.5647	479.7181
-0.0049	720.7585	648.3712	626.5233	618.0146
-0.0046	672.2456	692.8345	681.8275	674.1881
-0.0044	555.1447	619.9683	622.1713	619.5853
-0.0041	408.5176	473.4654	483.8720	486.0228
-0.0039	271.4029	317.2155	328.8240	332.7534
-0.0037	163.3996	190.6346	199.8140	203.4219
-0.0034	89.7592	105.5854	111.7878	114.3818
-0.0032	47.2209	56.8249	60.6617	62.2859
-0.0029	26.0372	32.0958	34.5341	35.5857

---

$\beta$  esté cercano a  $-0.9$ , si se supusiera normalidad y realmente el valor esperado del parámetro estuviera cercano a  $-0.9$ , llegaríamos a conclusiones erróneas, por lo que no se puede considerar que las inferencias sean robustas.

En este caso, se sugiere tomar más observaciones de la variable, si esto es posible, para tratar de asignar una distribución al parámetro  $\beta$ , en todo caso, la decisión depende de las conclusiones que se esperan obtener del estudio, esto es, si se trata de un estudio exploratorio, se puede suponer normalidad en las observaciones y obtener un valor aproximado de los parámetros.



## CONCLUSIONES GENERALES.

Cuando se realiza un análisis estadístico, se recomienda hacer un estudio de robustez, principalmente de inferencia, puesto que, como ya se ha discutido, las suposiciones distribucionales que se asignan inicialmente a las observaciones pueden afectar las inferencias que se realicen sobre la población bajo estudio.

La familia de distribuciones Exponencial Potencia, es muy útil para analizar el efecto de la suposición de normalidad de las observaciones, ya que, por medio de las distribuciones marginales de  $\underline{\theta}$  y finales para  $\underline{\theta}$  y  $\beta$ , además de realizarse el análisis de normalidad, se obtiene la información sobre el modelo que sugieren las observaciones mismas. También, por medio de dichas distribuciones puede realizarse el análisis de robustez, y dependiendo del resultado que se obtenga se puede determinar la acción a seguir.

Con  $P(\beta|\underline{Y})$  y  $P(\theta_j|\beta, \underline{Y})$ ,  $j=1,2,\dots,n$ , se puede determinar si existe robustez observando los cambios ocurridos en estas distribuciones al variar  $\beta$ . Si no se detecta robustez con las distribuciones anteriores, se puede analizar la robustez con  $P(\theta_j|\underline{Y})$ ,  $j=1,2,\dots,n$ . En caso de que las observaciones no proporcionen indicios de la existencia de robustez, se puede determinar el valor de  $\beta$  más probable y así tener una idea de la distribución de  $\beta$  que sugieren las observaciones. En cualquier caso, con  $P(\theta_j|\underline{Y})$  ya se cuenta con los estimadores de  $\theta_j$ , lo cual es una ventaja más sobre el análisis estadístico clásico, puesto que podemos estimar los parámetros de interés sin tener una distribución determinada de las observaciones.

Todo lo anterior es muy importante, en especial, para el caso de Regresión Lineal, puesto que la metodología desarrollada para el Análisis de Regresión, tiene como base fundamental la suposición distribucional de la variable de respuesta.

## A P E N D I C E S

## APENDICE I.

- Obtención de  $E(y)$  y  $\text{Var}(y)$ .

Sea  $y$  que se distribuye como Exponencial Potencia (2.2), se demostrará que  $E(y) = \theta$ , para lo cual se hallará primeramente  $E(y-\theta)$

$$E(y-\theta) = \int_{-\infty}^{\infty} (y-\theta) w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |y-\theta|^{2/(1+\beta)}\right\} dy$$

sea  $u=y-\theta$   $du=dy$ , entonces

$$\begin{aligned} E(u) &= \int_{-\infty}^{\infty} u w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |u|^{2/(1+\beta)}\right\} du \\ &= \int_{-\infty}^0 u w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |u|^{2/(1+\beta)}\right\} du + \\ &= \int_0^{\infty} u w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |u|^{2/(1+\beta)}\right\} du = 0 \end{aligned}$$

Entonces  $E(u) = E(y-\theta) = 0$  Por tanto  $E(y) = \theta$

Ahora se demostrará que  $\text{Var}(y) = \sigma^2$

Se sabe que  $\text{Var}(y) = E(y^2) - [E(y)]^2$

sea  $u = y - \theta$  entonces  $du = dy$  haciendo el cambio de variable se tiene

$$\begin{aligned} E(u^2) &= \int_{-\infty}^{\infty} u^2 w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |u|^{2/(1+\beta)}\right\} du \\ &= \int_{-\infty}^0 u^2 w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |u|^{2/(1+\beta)}\right\} du + \\ &= \int_0^{\infty} u^2 w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |u|^{2/(1+\beta)}\right\} du \end{aligned}$$

como  $u^2$  y es no negativa para toda  $u$ , se tiene que

$$E(u^2) = 2 \int_0^{\infty} u^2 w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |u|^{2/(1+\beta)}\right\} du$$

sean  $p = 3$   $a = c(\beta)/\sigma^{2/(1+\beta)}$   $\alpha = 2/(1+\beta)$  entonces

empleando la expresión (3.1.1.4) se tiene que

$$E(u^2) = (2 w(\beta)/\sigma) \{ [(1+\beta)/2] [-c(\beta)/\sigma^{2/(1+\beta)}] \}^{-3(1+\beta)/2} \Gamma[(3/2)(1+\beta)]$$

sustituyendo los valores de  $w(\beta)$  y  $c(\beta)$  y simplificando, se obtiene

$$E(u^2) = \sigma^2 \Gamma[(3/2)(1+\beta)]^{-1} \Gamma[(3/2)(1+\beta)] = \sigma^2$$

por otro lado

$$E(u^2) = E[(y-\theta)^2] = E(y^2) - \theta^2 \quad \text{entonces} \quad E(y^2) = \sigma^2 + \theta^2$$

$$\text{por tanto} \quad \text{Var}(y) = E(y^2) - [E(y)]^2 = \sigma^2$$

- Obtención de  $\gamma_1$  y  $\gamma_2$  para la familia de distribuciones Exponencial Potencia.

i) Se sabe que  $\gamma_1 = E(y-\theta)^3/\sigma^3$  donde  $\theta$  es la media y  $\sigma^2$  es la varianza de la distribución, entonces

$$E(y-\theta)^3 = \int_{-\infty}^{\infty} (y-\theta)^3 w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |y-\theta|^{2/(1+\beta)}\right\} dy$$

sea  $u = y-\theta$  entonces  $du = dy$  haciendo el cambio de variable se tiene que

$$\begin{aligned} E(u^3) &= \int_{-\infty}^{\infty} u^3 w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |u|^{2/(1+\beta)}\right\} du \\ &= \int_{-\infty}^0 u^3 w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |u|^{2/(1+\beta)}\right\} du + \\ &= \int_0^{\infty} u^3 w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |u|^{2/(1+\beta)}\right\} du = 0 \end{aligned}$$

por tanto  $\gamma_1 = 0$

ii)  $\gamma_2$  tiene la expresión  $\gamma_2 = [E(y-\theta)^4/\sigma^4] - 3$

para la familia de distribuciones Exponencial Potencia

$$E(y-\theta)^4 = \int_{-\infty}^{\infty} (y-\theta)^4 w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |y-\theta|^{2/(1+\beta)}\right\} dy$$

sea  $u = y-\theta$ ,  $dy=du$  se tiene que

$$\int_{-\infty}^{\infty} u^4 w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |u|^{2/(1+\beta)}\right\} du$$

como  $u^4$  es no negativa para toda  $u$  se tiene que

$$E(u^4) = 2 \int_0^{\infty} u^4 w(\beta) \sigma^{-1} \exp\left\{-[c(\beta)/\sigma^{2/(1+\beta)}] |u|^{2/(1+\beta)}\right\} du$$

empleando (3.1.1.4) y haciendo

$$p = 5 \quad a = c(\beta)/\sigma^{2/(1+\beta)} \quad \alpha = 2/(1+\beta) \quad \text{se tiene}$$

$$E(u^4) = 2 w(\beta) \sigma^{-1} [(1+\beta)/2] [c(\beta)/\sigma^{2/(1+\beta)}]^{-5(1+\beta)/2} \Gamma[(5/2)(1+\beta)]$$

sustituyendo  $w(\beta)$  y  $c(\beta)$  por sus expresiones (2.2), y simplificando, obtenemos

$$E(u^4) = \sigma^4 \left\{ \Gamma[(1/2)(1+\beta)] \Gamma[(5/2)(1+\beta)] / \{\Gamma[(3/2)(1+\beta)]\}^2 \right\} = E(y-\theta)^4$$

$$\text{por tanto } \gamma_2 = \left\{ \Gamma[(1/2)(1+\beta)] \Gamma[(5/2)(1+\beta)] / \{\Gamma[(3/2)(1+\beta)]\}^2 \right\} - 3$$

APENDICE II.

t) Se desea demostrar que  $\lim_{q \rightarrow \infty} [M(\theta)]^{1/q} = h + |m-\theta|$   
 consideremos una serie monótona creciente finita de números positivos  $\{a_n\}$  y un número  $s$ , tal que

$$S = \left[ \sum_{i=1}^n a_i^q \right]^{1/q} = a_n \left[ \sum_{i=1}^n (a_i/a_n)^q \right]^{1/q}$$

como  $\{a_n\}$  es creciente,  $(a_i/a_n) \leq 1 \quad \forall i \in \{1, \dots, n\}$  entonces (a)

$$(S / a_n) = \left[ \sum_{i=1}^n (a_i/a_n)^q \right]^{1/q}$$

y

$$\log (S / a_n) = (1/q) \log \left[ \sum_{i=1}^n (a_i/a_n)^q \right]$$

como  $0 \leq (a_i/a_n) \leq 1$  entonces  $0 \leq (a_i/a_n)^q \leq 1$  con  $q > 1$   
 entonces se tiene que

$$\lim_{q \rightarrow \infty} \log \left[ \sum_{i=1}^n (a_i/a_n)^q \right] = \log \left\{ \lim_{q \rightarrow \infty} \left[ \sum_{i=1}^n (a_i/a_n)^q \right] \right\} = \log r$$

$r \in [1, n]$

por tanto

$$\lim_{q \rightarrow \infty} \{ \log (S / a_n) \} = \lim_{q \rightarrow \infty} \{ (1/q) \log r \} = 0$$

$$\lim_{q \rightarrow \infty} (S / a_n) = 1, \quad \lim_{q \rightarrow \infty} S = a_n$$

por construcción  $a_n$  es el mayor de la serie, por tanto

$$\begin{aligned} \lim_{q \rightarrow \infty} [M(\theta)]^{1/q} &= \max_i |y_i - \theta| \\ &= \max \left\{ |\theta - y_{(s)}|, |\theta - y_{(m)}| \right\} \\ &= \begin{cases} h + (m-\theta) & \theta < m \\ h & \theta = m \\ h + (\theta-m) & \theta > m \end{cases} \end{aligned}$$

donde

$$m = (1/2) (y_{(n)} + y_{(1)}) \quad \text{es el punto medio}$$

$$h = (1/2) (y_{(n)} - y_{(1)}) \quad \text{es el rango medio}$$

por tanto

$$\lim_{q \rightarrow \infty} [M(\theta)]^{1/q} = h + |m - \theta|$$

(i)  $\hat{\theta}$  coincide con el estimador Máximo Verosímil.

$$P(z | \theta, \alpha, \beta) = [w(\beta)]^n \sigma^{-n} \exp \left[ [c(\beta) / \sigma^{2/(1+\beta)}] \sum_{i=1}^n |y_i - \theta|^{2/(1+\beta)} \right]$$

$$\text{sea } M(\theta) = \sum_{i=1}^n |y_i - \theta|^{2/(1+\beta)}$$

sea  $\hat{\theta} \in \mathbb{R}$  tal que  $M(\hat{\theta}) \leq M(\theta) \quad \forall \theta \in \mathbb{R}$

como  $c(\beta) \in \mathbb{R}^+ \quad \forall \beta \in (-1, 1)$  y  $\sigma^{2/(1+\beta)} \in \mathbb{R}^+ \quad \forall \sigma \in \mathbb{R}^+$ , entonces

$$-[c(\beta) / \sigma^{2/(1+\beta)}] M(\hat{\theta}) \geq -[c(\beta) / \sigma^{2/(1+\beta)}] M(\theta) \quad \forall \theta \in \mathbb{R}, \sigma \in \mathbb{R}^+ \\ \text{y } \beta \in (-1, 1)$$

entonces

$$P(z | \hat{\theta}, \sigma, \beta) \geq P(z | \theta, \sigma, \beta) \quad \forall \theta \in \mathbb{R}, \sigma \in \mathbb{R}^+ \text{ y } \beta \in (-1, 1) \quad (1)$$

sea  $(\hat{\theta}, \hat{\sigma}, \hat{\beta}) \in \mathbb{R} \times \mathbb{R}^+ \times (-1, 1)$  el estimador máximo verosímil de  $(\theta, \sigma, \beta)$  si existe, entonces

$$-w(\hat{\beta}) \hat{\sigma}^{-n} \exp \left\{ -[c(\hat{\beta}) / \hat{\sigma}^{2/(1+\hat{\beta})}] M(\hat{\theta}) \right\} \\ \geq [w(\hat{\beta})] \hat{\sigma}^{-n} \exp \left\{ -[c(\hat{\beta}) / \hat{\sigma}^{2/(1+\hat{\beta})}] M(\hat{\theta}) \right\}$$

es decir

$$P(z | \hat{\theta}, \hat{\sigma}, \hat{\beta}) \geq P(z | \theta, \sigma, \beta) \quad \forall (\theta, \sigma, \beta) \in \mathbb{R} \times \mathbb{R}^+ \times (-1, 1)$$

entonces

$$P(z | \hat{\theta}, \hat{\sigma}, \hat{\beta}) \geq P(z | \tilde{\theta}, \hat{\sigma}, \hat{\beta}) \geq P(z | \theta, \sigma, \beta) \quad \forall (\theta, \sigma, \beta) \in \mathbb{R} \times \mathbb{R}^+ \times (-1, 1) \\ \text{y } \tilde{\theta} \in \mathbb{R} \text{ tal que } M(\tilde{\theta}) \leq M(\hat{\theta}) \dots (2)$$

aun más  $P(z | \hat{\theta}, \hat{\sigma}, \hat{\beta}) \geq P(z | \tilde{\theta}, \hat{\sigma}, \hat{\beta})$ , entonces

$$[w(\hat{\beta})]^n \hat{\sigma}^{-n} \exp \left\{ -[c(\hat{\beta}) / \hat{\sigma}^{2/(1+\rho)}] M(\hat{\theta}) \right\} \geq$$

$$[w(\hat{\beta})]^n \hat{\sigma}^{-n} \exp \left\{ -[c(\hat{\beta}) / \hat{\sigma}^{2/(1+\rho)}] M(\tilde{\theta}) \right\}$$

entonces  $M(\hat{\theta}) \leq M(\tilde{\theta})$  y por (2)  $M(\tilde{\theta}) \leq M(\hat{\theta})$   
 por tanto  $M(\hat{\theta}) \leq M(\tilde{\theta})$ , es decir, el estimador máximo verosímil de  $\theta$   
 es el valor que minimiza a  $M(\theta)$ .

## B I B L I O G R A F I A

- Abramowitz, Milton & Stegun, Irene. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Dover Publicatinos, Inc., New York, 1965.
- Acton, Forman S. *Numerical Methods that work*, Harper & Row, 1970.
- Box G.E.P., (1953) 'A note on Regions For Test of Kurtosis, *Biometrika* 40, p. 465.
- Box, G.E.P. & Andersen, S.L. (1955). *Permutation Theory in the Derivation of Robust Criteria and the Study of Departures from Assumptions*, *J. Roy. Statist. Soc., Series B* 17,1.
- Box, G.E.P. & Tiao G.C., (1962). *A further look at Robustness via Bayes Theorem*, *Biometrika*, 49 3 and 4, p. 419.
- Box G.E.P. & Tiao G.C. (1964), *A note on Criterion Robustness and Inference Robustness*, *Biometrika* 51, 1 and 2, p.169.
- Box, G.E.P & Tiao, G.E.C. *Bayesian Inference in Statistical Analysis*, Addison-Wesley, 1973.
- Carlton, G.A. (1946) *Estimating the Parameters of a Rectangular Distribution*, *Am. Math. Statist.* 17, p. 355.
- Diananda, P.H. (1949) *Note on Some Properties of Maximum Likelihood Estimates*, *Proc. Camb. Phil. Soc.* 45, p.536.
- Draper, Norman, Richard *Applied Regression Analysis*, New York :Willey, 1981.
- Forsythe, George E. (1972). *Von Neumann's comparison methods of Random Sampling from the Normal and other Distributions*, *Mathematics of Computation*, V26, No. 120, p.817.

Forsythe, George E. *Computer Methods for Mathematical Computations*, Prentice-Hall, 1977.

Geary, R.C. (1936). *The Distributions of Student's Ratio for Non-Normal Samples*, J. Royal Statist. Soc., Supplement 3, p.178.

Gayen, A. K. (1949) *The Distribution of Student's t in Random Sample of any Size Drawn from Non-Normal Universe*, Biometrika 36, p. 353.

Gayen, A.K. (1950) *The distribution of Variance Ratio in Random Sample of any Size Drawn From a Non Normal Universe*, Biometrika 37, p. 236.

Neyman, J. & Pearson, E.S. (1928) *On the Use and Interpretation of Certain Criteria for Purposes of Statistical Inference, Part. I.*, Biometrika 20A, p.175.

Reuben & Rubinstein. *Simulation and the Monte Carlo Methods*, John Wiley & Sons, New York, 1981.

Smith, David M. *A program for approximating Integrals*, Byte, December, 1986, p.113.

Turner, M.C. (1960) *On Heuristic Estimation Methods*, Biometrics 16, p. 299.