3061
Zej.
2

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

Colegio de Ciencias y Humanidades

TRES METODOS DE ESTIMACION APLICADOS A UN MODELO DE CRECIMIENTO HUMANO

TESIS DE MAESTRIA EN CIENCIAS

Estad e Inves d operación

MARIO CORTINA BORJA

México, D. F. 1987

TRES METODOS DE ESTIMACION APLICADOS A UN MODELO DE CRECIMIENTO HUMANO

TESIS DE MAESTRIA EN CIENCIAS

Mario Cortina Borja

INDICE

0) INTRODUCCION

El propósito de este trabajo es comparar el funcionamiento de tres métodos de estimación ante el problema de estimar los parámetros del modelo de crecimiento humano Preece-Baines 1. Los tres métodos son el de mínimos cuadrados no lineales ordinarios [MCO] y dos casos del mínimos cuadrados no lineales generalizados: uno considerando a la matriz de varianza covarianza involucrada como una matriz circular simétrica [CIRC], y otro incorporando en tal matriz una estructura derivada de suponer que los residuales se comportan de acuerdo a un proceso autorregresivo de orden 1 [AR1].

En el primer capítulo de la tesis se exponen algunos modelos de crecimiento humano, haciendo énfasis en aquéllos que buscan describir el fenómeno desde el nacimiento hasta las edades adultas. En esta parte se exponen algunas ventajas y desventajas de los modelos considerados y se hace una breve discusión sobre trabajos en los que se han aplicado.

A continuación se presentan los métodos de estimación que se utilizaron en la estimación de PB1. En el tercer capítulo se muestran los resultados obtenidos con datos longitudinales. Estos resultados incluyen la estimación de 10 parámetros biológicos, así como una discusión acerca de la calidad de los ajustes, en los aspectos numéricos y estadísticos.

El capítulo cuarto es donde se presentan los estudios de simulación realizados para efectuar las comparaciones entre los métodos. Cabe señalar que las simulaciones se efectuaron utilizando dos clases de estructuras de error: suponiendo independencia y suponiendo un proceso autorregresivo de orden

i. Con base en las simulaciones, puede decirse que el método MCO es adecuado para estimar los parámetros del modelo, aún en casos en que la autocorrelación de los residuales sea significativa.

La tesis incluye dos apéndices: en el primero de ellos se expone la construcción de la familia de modelos de Preece y Baines; en el otro se discute muy brevemente la estimación para datos provenientes de estudios transversales.

La idea de realizar una tesis de regresión no lineal utilizando un modelo de crecimiento humano tiene dos orígenes: por un lado, el curso de Regresión y Análisis de Varianza impartido por Francisco Aranda y Silvia Ruiz Velasco en el semestre de verano de 1984; por otro, la información proporcionada por María Elena Sáenz acerca del modelo de Preece y Baines. Posteriormente, Francisco fungió como director de tesis, por lo que le estoy especialmente agradecido. En este mismo sentido, mucho agradezco los comentarios de Víctor Aguirre, Guillermo Daz, Luis Felipe Bazúa, Alfredo Bustos, Apolinar Calderón, Hortensia Góngora, Leticia Gracia-Medrano, Federico O'Reilly, Salvador Pérez Esteva y Enrique Villa. Por su parte, el Instituto de Investigaciones Antropológicas hizo posible este trabajo al proporcionar un ambiente (casi) óptimo para su realización. Desde luego, la mayor deuda la tengo con mi familia y mis amigos, que siempre se portaron más que a la altura. Sin ellos, seguiría diciendo "ora sí ya voy a acabar la tesis".

M.C.B.
Colonia Roma – Ciudad Universitaria,
8 de agosto de 1987 (en la mañana).

You can't always get what you want,
but if you try sometime,
you just might find you've got what you need.

Mick Jagger/Keith Richards

1) MODELOS DE CRECIMIENTO HUMANO

1.1) INTRODUCCION

Durante todo este trabajo, el término crecimiento se referirá al aumento de tamaño individual manifestado en el cuerpo a partir del nacimiento.

Este proceso, que en realidad abarca desde el momento de la concepción hasta que el individuo se convierte en adulto, presenta patrones generales en la forma en que ocurren los incrementos de tamaño respecto a la edad cronológica que son comunes a toda la especie humana (Faulhaber, 1976:7). En las gráficas 1A, 1B y 1C se ilustra lo anterior con los datos longitudinales de las estaturas de dos individuos (masculino y femenino), sus incrementos anuales y los incrementos anuales de los incrementos.

Dentro de la unidad en la forma en que ocurre el fenómeno hay variaciones entre poblaciones y entre individuos condicionadas esencialmente por factores hereditarios (influencias internas) y ambientales (influencias externas). Parece ser que los primeros tienen un mayor peso en la determinación de la forma en que ocurre el proceso, si bien resulta sumamente difícil separar la variación producida por ambos factores ya que se encuentran interactuando mutuamente (Susanne, 1980:236).

Aun así, dentro de esta variabilidad, es posible encontrar tendencias definidas como "tipos de formas particulares de las curvas individuales del crecimiento, las cuales son resultado de la interacción, a distintas edades cronológicas, de los factores antes mencionados" (Espinosa y Faulhaber, 1979:433).

En este trabajo únicamente se analizarán modelos que representen el cambio de variables antropométricas en función de la edad cronológica. La mayoría de los ejemplos que se presentan se refieren a la estatura. Puesto que la talla es una expresión integral del crecimiento, la construcción de modelos para esta medida es importante en áreas como ergonomía, genética, pediatría, endocrinología, nutrición, biología y antropología física. Por otra parte, al ser la estatura una medida que tiende hacia un valor final (al contario de lo que puede suceder con, por ejemplo, los pliegues cutáneos), los modelos construidos para ella pueden utilizarse en variables, como las longitudes de extremidades, que tengan esta propiedad.

Estos modelos son útiles en dos niveles de las investigaciones de crecimiento: por un lado permiten condensar y analizar información que sería difícil de obtener de otra forma y que puede ser útil en aspectos como construcción de estándares de crecimiento y detección de desnutrición; por otra parte, la construcción y el análisis de modelos matemáticos que intenten representar el proceso de crecimiento son, por sí mismos, una contribución al estudio del fenómeno.

Cabe mencionar rápidamente que hay otros métodos de análisis de datos longitudinales, expuestos en Roche *et al.* (1975) y en Roche (1980), que principalmente buscan predecir la talla final en función de variables como la estatura en la edad de la menarquia, la estatura final de los padres, la edad ósea, y otras características biológicas. Una comparación entre algunos de estos métodos aplicados a datos de niños mexicanos puede verse en el trabajo de Faulhaber (1982).

La construcción de modelos de crecimiento individual implica proponer alguna función que relacione a los cambios de tamaño en la estatura con la edad cronológica y una explicación para las variaciones que hay entre individuos y dentro del crecimiento particular que pueden no deberse a cambios en la edad. Estas variaciones pueden atribuirse a los errores de medición; a procesos fisiológicos particulares (Comas, 1966:218), a cambios estacionales en el tiempo en que se efectúan las mediciones (Lee, 1980); a las diferencias entre la fecha planeada para la observación y aquélla en la que realmente ésta se lleva a cabo y al hecho de que necesariamente se observa sólo parcialmente al proceso de crecimiento. La cantidad y complejidad de estas causas de variabilidad, así como las interacciones que se dan entre ellas señalan hacia la importancia de incorporar un componente estocástico en los modelos propuestos, si bien esta clase de supuestos no es incorporada en todos ellos, particularmente en los métodos que únicamente buscan suavizar los datos.

Generalmente se aceptan dos enfoques en la construcción de modelos de crecimiento: el no estructural y el estructural (Preece, 1978; Bock y Thissen, 1980; Preece y Heinrich, 1981).

En el primero, se intenta eliminar o disminuir los errores de medición y los efectos que puedan tener los factores mencionados anteriormente sobre las observaciones. Esto se hace con alguno de los siguientes métodos:

1) Suavizando las observaciones globalmente, con polinomios de Lagrange (Rice, 1965:§5.2). Joosens y Brems-Heyns (1975) y Zerbe (1979) presentan ajustes de esta clase para datos longitudinales de crecimiento humano.

2) Suavizando localmente, casi siempre con funciones spline cúbicas, que consisten en polinomios de grado tres ajustados para alguna partición del rango de edades con la restricción de que los valores de los polinomios deben coincidir en las fronteras de

intervalos adyacentes (Rice, 1985:§4.3). Algunos ejemplos de este tipo de ajustes con datos longitudinales pueden verse en Largo *et al.* (1978), Molinari *et al.* (1980) y Berkey *et al.* (1983a). Al parecer, este método produce mejores resultados que el anterior. En ambos casos no hay supuestos estocásticos.

3) Ajustando globalmente un polinomio con la teoría usual de la regresión lineal. Los parámetros se estiman generalmente por mínimos cuadrados suponiendo un componente estocástico de error aditivo para el modelo. Hay que hacer notar que para datos longitudinales lo más probable es que el supuesto de independencia para los errores no se satisfaga. Pomerance y Krall (1981) utilizan este método para datos longitudinales de crecimiento humano, señalando que, para la estatura, entre el 20% (para niños) y el 28% (para niñas) de los individuos mostró correlaciones significativas entre los residuales.

Los métodos 1 y 3 tienen la desventaja de que el grado del polinomio que se ajusta (es decir, el número de parámetros desconocidos), varía de individuo a individuo, lo que dificulta algunas comparaciones. Por otra parte, los modelos no estructurales no tienen forma de utilizar en su construcción cierta información conocida acerca del proceso de crecimiento, por ejemplo, su caracter asintótico, la existencia de un máximo local en la función de velocidad durante la adolescencia, etc. Si bien es posible reproducir estas características con ajustes polinomiales, ésto puede requerir de un número muy grande de parámetros por estimar. Janssens y Brems-Heyns (1975) y Zerbe (1979) utilizan polinomios de grados 18 y 17, respectivamente.

El suavizamiento por splines parece ser adecuado cuando se intenta describir el crecimiento individual durante periodos relativamente cortos (3 o 4 años); sin embargo, su aplicación en rangos de mayor tamaño tiene el inconveniente de que es necesario estimar un gran número de parámetros para obtener ajustes de buena calidad. Por otra parte, a menos que la investigación se centre en aspectos muy locales del crecimiento, lo usual es trabajar con modelos que cubran rangos de edades que abarquen a todo el proceso o bien, a etapas como la infancia o la adolescencia. La gráfica 1D muestra un ajuste con un polinomio de Lagrange en datos longitudinales de estatura; nótese que si bien la curva pasa exactamente por todos los puntos, el modelo no sirve para estimar parámetros de interés, como la estatura adulta. Además, el polinomio utilizado es de grado 8, lo que implica la necesidad de estimar 9 parámetros con 10 datos.

Bard (1974:1) llama al enfoque no estructural ajuste de curvas, mencionando que a los problemas anteriores debe sumarse la posible carencia de consideraciones estadísticas para evaluar la calidad de los ajustes. Este autor (1974:3) designa como ajuste de modelos llamado enfoque estructural. La principal característica de éste consiste en que a todos los individuos se les ajusta, estimando a los parámetros desconocidos con métodos estadísticos, una misma función. Esta es seleccionada como consecuencia de consideraciones teóricas respecto al fenómeno más que por motivos de conveniencia en cuanto a resumir la información. Por este motivo, los parámetros que involucra una función en un modelo estructural pueden interpretarse en términos del fenómeno, a diferencia de aquéllos presentes en los modelos no estructurales. De esta forma, los estimadores de los parámetros se consideran como resúmenes adecuados del proceso, al mismo tiempo que es posible hacer inferencias estadísticas y comparaciones sobre ellos. Otra ventaja de esta clase de modelos radica en el hecho de que es posible construir funciones de los parámetros que estimen a otros parámetros que no estén contenidos explícitamente en el modelo.

Recientemente, al exponer una metodología de estimación no paramétrica para curvas de crecimiento, Goldstein (1986) ha señalado algunas desventajas del enfoque estructural, llamado por él enfoque paramétrico. Estas son:

1) Al restringir el modelo a una sola función se tiene una forma que puede resultar demasiado rígida para permitir un rango razonable de variaciones individuales.

2) Al ajustar una misma curva sobre un rango de edades bastante amplio habrá variaciones locales (como la aceleración del crecimiento entre los 6 y 8 años conocida como *mid-growth spurt*) que pueden no ser observables con el modelo.

3) Los modelos paramétricos no modelan explícitamente a la variación debida a características que pueden variar con la edad factores estacionales, medio ambientales o bioquímicos, situación socioeconómica, etc. Como se ha visto, todas estas variables se incluyen en la estimación solamente a través del error. Sobre este último punto, Berkey y Laird (1986) presentan un método de estimación para modelos estructurales no lineales de crecimiento que puede incluir covariables como las mencionadas anteriormente. En ese trabajo, el procedimiento se ejemplifica con el modelo de Jenss-Bayley (sección 1.2.2).

Goldstein (1986) amplía el enfoque no paramétrico primeramente sugerido en el contexto de los estudios de crecimiento por Stützle *et al.* (1980) y por Gasser *et al.* (1984). Tal metodología no se analizará en este trabajo.

En el contexto de los modelos individuales de crecimiento, teniendo n observaciones para el i-ésimo individuo, el modelo se puede escribir como:

$$y_{ij} = f(x_{ij};\underline{\theta}) + \varepsilon_{ij}, \; (j = 1,\ldots,n) \qquad (1.1)$$

siendo y_{ij} la estatura en la edad x_{ij}, $f(.;.)$ la función de crecimiento, $\underline{\theta}$ un vector de p parámetros y ε_{ij} el error, el cual se supone que es no observable y que incluye tanto a los errores de medición como a los factores que no fueron considerados en la construcción de la función. Hay que hacer notar que la medición de la edad se supone libre de error.

Bard (1974:26) llama a la ecuación (1.1) modelo estándar reducido. En general, se le conoce como modelo de regresión, siendo x_{ij} (que puede ser un vector para cada pareja i,j) la variable(s) explicativa(s) y y_{ij} la variable dependiente.

Si f es de la forma

$$f(x_{ij};\underline{\theta}) = \sum_{k=1}^{p} \theta_k \, b_k(x_{ij}),$$

donde las b_k son funciones que sólo dependen de las variables explicativas, el modelo es lineal en $\underline{\theta}$. En cualquier otro caso, se dirá que el modelo es no lineal. Cabe hacer notar que en ciertos casos, mediante alguna transformación o reparametrización, un modelo no lineal puede llevarse a la forma anterior.

Como ejemplo de lo primero,

$$y = x^{\theta} \, exp(\varepsilon)$$

puede escribirse como

$$\ln(y) = \theta \ln(x) + \varepsilon.$$

Por otra parte, una reparametrización para llevar al modelo

$$y = exp(\theta) + \varepsilon$$

a una forma de regresión lineal consiste en hacer

$$\beta = exp(\theta).$$

Si no es posible linealizar al modelo en los parámetros, se dice que ésto es instrínsecamente no lineal (Ratkowsky, 1983:8-11).

1.2) ALGUNOS MODELOS DE CRECIMIENTO HUMANO

En esta sección se hace una revisión de la bibliografía consultada respecto a varios modelos de crecimiento. Su propósito es ilustrar algunos aspectos referentes a la construcción de las funciones mencionadas en la sección anterior.

1.2.1) MODELO DE JENSS-BAYLEY

Uno de los primeros intentos de aplicar un modelo estructural a datos longitudinales fue el de R.M. Jenss y N. Bayley, en 1937. La función por ellas propuesta puede escribirse como:

$$f(t;\underline{\theta}) = a + bt - exp(c+dt) \qquad (1.2)$$

Esta función ha sido utilizada para analizar el crecimiento en la talla y en el peso desde el nacimiento hasta los 7 años. La forma de la función se debe a que durante los primeros meses de vida el crecimiento puede considerarse como casi lineal con una pendiente relativamente grande (Berkey, 1982:222). Conforme la edad aumenta, se tiene una disminución en la velocidad de crecimiento:

$$d f/dt = b - d\, exp(c+dt),$$

representada por la diferencia entre el componente lineal y el exponencial.

Los parámetros a y $exp(c)$ están dados en unidades de longitud, mientras que b está en $(\text{tiempo})^{-1}$. El parámetro $exp(d)$ es adimensional, siendo $d < 0$.

Como:

$$\frac{d^2 f(t)/dt^2}{d^2 f(t+1)/dt^2} = exp(d),$$

siendo t y t+1 dos tiempos contiguos, $exp(d)$ se interpreta como una medida constante de la aceleración para cualquier momento en términos de la aceleración del anterior.

Deming y Washburn (1963) aplicaron esta función en un estudio longitudinal extendiendo el rango de edad hasta los 8 años, mientras que Manwani y Agarwal (1973) lo restringen al primer año de vida. A partir de la función (1.2), Berkey *et al.* (1983b) construyen estándares longitudinales de crecimiento para niños menores de 6 años. Berkey y Kent (1983) la utilizaron para ajustar datos de talla y peso en niños medidos entre los 3 meses y los 6 años. Schlaepfer (1987) presenta resultados de ajustes efectuados con este modelo para peso y talla de niños de una comunidad rural en el estado de Puebla. La finalidad del trabajo de esta investigadora es evaluar los efectos que diferentes clases de alimentación tienen sobre esas variables. Como se ve, a cincuenta años de haber sido propuesta por primera vez, la función de Jenss-Bayley es todavía ampliamente utilizada.

1.2.2) MODELO DE COUNT

En 1943, E.W. Count propuso explicar el proceso completo del crecimiento utilizando la unión de tres funciones; las dos primeras eran de la forma:

$$f(t;\underline{\theta}) = a + bt + c\ln(t); \qquad (1.3)$$

la tercera función, utilizada para modelar el crecimiento durante la adolescencia, fue la logística (subsección 1.2.3).

La función (1.3) ya había sido utilizada por Count con observaciones de medidas craneales en 1942. Las dos funciones de esta forma describían el crecimiento desde prácticamente el nacimiento ($t_{ij} > 0$) hasta los siete años y desde esta edad hasta el inicio de la adolescencia; durante la adolescencia, era una logística. Los problemas de trabajar con estas tres funciones están en el manejo de los 10 parámetros necesarios, así como en la dificultad de hacer coincidir los valores de las funciones en las fronteras de sus dominios, i.e., los puntos en los que se debía dejar de usar una función para empezar a considerar a otra.

El parámetro a se interpreta como la estatura en la primera edad observada; b es la asíntota inferior de la velocidad de crecimiento, además de ser la tasa de crecimiento para el componente lineal de la función, por lo que sus unidades son cm/tiempo; finalmente, c está dada en unidades de longitud y determina tanto la velocidad de crecimiento, ya que:

$$df/dt = b + ct,$$

como la aceleración, pues:

$$d^2f/dt^2 = -c/t^2.$$

Como se ve, la idea de Count era muy simple; la localización de los parámetros en la función está determinada por las dos ecuaciones anteriores.

Debido a que, al ser lineal en los parámetros, se tiene una gran facilidad para obtener y analizar estimaciones, el modelo propuesto por Count (ecuación 1.3) ha sido ampliamente utilizado con datos longitudinales de niños entre 0 y 7 años. Por ejemplo, Wingerd (1970) y Smith *et al.* (1983) lo ajustan para la talla y el peso de niños prematuros. Berkey (1982) presenta una comparación entre los modelos de Count y de Jenss-Bayley (subsección 1.2.1), concluyendo que el segundo proporciona ajustes más adecuados. Cabe mencionar que esta investigadora afirma no haber encontrado violaciones importantes a los supuestos usuales del método de mínimos cuadrados (lineales o no lineales). En un trabajo posterior (Berkey y Kent, 1983), se establece que este hecho parece deberse a que las edades en que se realizaron las observaciones están "ampliamente separadas en el tiempo" (p. 525). Tales separaciones son de entre 3 y 6 meses. Más adelante se discutirá sobre los métodos de estimación y sus problemas. Mientras, pueden verse, en la gráfica 1E a los modelos de Jenss-Bayley y Count, y en la gráfica 1F a sus funciones de velocidad.

1.2.3) MODELOS SIGMOIDALES

La ecuación diferencial

$$d f/dt = k f \cdot (A^n - f^n)/nA^n \tag{1.4}$$

donde A es la asíntota superior de f y k es una tasa constante de crecimiento, admite como solución a:

$$f(t;\underset{\sim}{\theta}) = A\cdot[1 \pm h\cdot exp(-kt)]^{-1/n} \tag{1.5}$$

siendo h una constante (Draper y Smith, 1981:§10.7; Preece y Heinrich, 1981:249). La ecuación (1.5) es conocida como la función de Richards (1959). En (1.4) se ve que la velocidad de crecimiento es proporcional a una función que representa la distancia entre la asíntota superior y otra función lineal de la talla (kf/nA^n), la cual está implícitamente relacionada con el tiempo, a través de f.

La ecuación (1.4) ha dado origen a varios modelos de crecimiento y de rendimiento. Ratkowsky (1983) presenta algunas de estas funciones.

Si n =1 en (1.5), se tiene la función logística, que puede escribirse (Marubini *et al.*, 1971) como:

$$f(t;\underset{\sim}{\theta}) = P + K\cdot[1 + exp(a-bt)]^{-1}. \tag{1.6}$$

Los parámetros de la ecuación anterior son los siguientes: P y P+K son las asíntotas inferior y superior de la curva, respectivamente; a es una constante adimensional de integración determinada de la posición del origen y b es una tasa de crecimiento constante, en unidades de $(\text{tiempo})^{-1}$. Con esta función se trata de representar al crecimiento desde el inicio de la adolescencia. Puede mostrarse que la curva de velocidad de la logística es simétrica alrededor de su punto de inflexión, localizado en $t^* = a/b$ (Draper y Smith, 1981:508). Por otra parte, como se tiene que:

$$f(t^*) = P + K/2,$$

entonces t^* resulta ser la edad en la que se alcanza la mitad del crecimiento obtenido desde la primera medición considerada.

Si $n \to 0$, (1.5) se transforma en:

$$f(t|\underline{\theta}) = P + K \cdot exp(-e^{a-bt}) \qquad (1.7)$$

que es la función de Gompertz, propuesta por el actuario inglés de ese apellido en 1825 para explicar la mortalidad. Los parámetros de esta función pueden interpretarse en los mismos términos usados para (1.6).

Marubini *et al.* (1971:239) muestran que la derivada de (1.7) es asimétrica, sesgada hacia la izquierda, con punto de inflexión en a/b. Para esta edad, en el modelo de Gompertz se ha alcanzado el 36.8% del crecimiento habido desde la primera medición. Aparentemente, esto podría ser una ventaja respecto a la función logística, en el sentido de que aproximaría mejor a la curva de velocidad para la adolescencia, en la que la aceleración es mayor antes del momento de máxima velocidad, disminuyendo a partir de esa edad (Marubini, 1978:129).

Las gráficas 16 y 11 muestran las curvas de distancia y de velocidad de las funciones logística y Gompertz, respectivamente.

Los ajustes producidos por estos dos modelos sigmoidales han sido comparados en varios estudios. Tanner *et al.* (1966) obtienen ajustes ligeramente mejores (en términos de varianza residual) con la función de Gompertz que con la logística; sin embargo, la conclusión de Marubini *et al.* (1971:251) es la contraria, al menos para mujeres, pues no contó con datos masculinos. Es interesante hacer notar que en ese estudio no se encontraron correlaciones significativas para los residuales en todos los casos examinados para ambas funciones. Además, los modelos fueron bastante robustos ante desviaciones de la periodicidad de medición. Estos resultados fueron confirmados en un estudio posterior realizado por Marubini y Tanner (Marubini *et al.* 1972).

Por su parte, Hauspie *et al.* (1980) comparan las dos funciones para ambos sexos durante la adolescencia. En este trabajo se reportan resultados similares a los de Marubini *et al.* (1971, 1972); respecto a la correlación de los residuales, ésta apareció en el 5% y el 3% de los casos para la logística y la de Gompertz, respectivamente.

Como se ve, no hay razones contundentes para preferir a un modelo sobre el otro. Sin embargo, el logístico ha tenido mayor aceptación tanto en estudios del crecimiento durante la adolescencia (Tanner *et al.* 1976) como en la construcción de nuevos modelos de crecimiento, como se verá más adelante.

1.2.4) MODELOS DOBLE Y TRIPLE LOGISTICO

Como su nombre lo indica, estos modelos están formados por la suma de dos y tres logísticas. En 1973 Bock *et al.* utilizaron dos logísticas para estudiar el rango completo de crecimiento: la primera representaba el crecimiento en la pre adolescencia y la suma de ambas ocurría a partir de la adolescencia. De esta forma buscaban disminuir la posibilidad de tener una división entre estos periodos. Este uso aditivo de logísticos tiene algunos problemas, señalados por El Lozy (1978) y Preece y Baines (1978); por ejemplo, la existencia de una asíntota inferior no hace de la logística un buen modelo para el crecimiento un poco antes de la adolescencia; además, como hace notar Preece (1978), "la suma de funciones no lineales destruye la estructura subyacente de las ecuaciones diferenciales originales, donde no hay una estructura similar". Por otra parte, el basarse únicamente en la curva de la estatura para describir el crecimiento parece ser menos adecuado que modelar utilizando desde el principio propiedades conocidas de la curva de velocidad (Stützle *et al.*, 1980:508).

El modelo doble logístico puede escribirse como:

$$f(t;\underline{\theta}) = a_1\cdot\{1+exp[-b_1(t-c_1)]\}^{-1} + (f-a_1)\cdot\{1+exp[-b_2(t-c_2)]\}^{-1} \tag{1.9}$$

siendo los parámetros como sigue (Bock y Thissen, 1980):a_1, b_1 y c_1 son, respectivamente, la ganancia en estatura,la velocidad máxima y la edad en la ésta se alcalnza, todo esto para la primera logística; f es la estatura final, por lo que $f-a_1$, b_2 y c_2 tienen el mismo significado para la adolescencia que el de los tres parámetros anteriores en la primera etapa del crecimiento.

Otra limitación de este modelo es el hecho de que, por la forma en que Bock *et al.* construyen la suma de logísticas, la estatura adulta (f) debe ser conocida en lugar de ser estimada como un parámetro más. Adicionalmente, el comportamiento de la autocorrelación de residuales deja mucho que desear respecto a los supuestos de la teoría usual de la estimación por mínimos cuadrados no lineales (Bock *et al.*, 1973). En fin, su derivada (que no fue estudiada en el trabajo original de Bock *et al.*) no ajusta suficientemente bien para casi todos los casos de las velocidades de crecimiento obtenidas por Tanner en 1966 (Marubini, 1978:131), particularmente entre los 0 y 5 años, además de no haber resultado robusto ante espaciamientos diferentes en las edades de las observaciones.

Intentando corregir los malos ajustes que el modelo doble logístico presentaba para edades entre los 0 y los 5 años, Bock y Thissen (1976) presentaron un modelo de tres componentes aditivos logísticos: los dos primeros representan el crecimiento pre puberal y el tercero, el obtenido durante la adolescencia; es decir, se incluyó otra logística para modelar el crecimiento entre los 0 y los 5 años.

El modelo triple logístico puede estribirse como:

$$f(t;\underline{\theta}) = a_1\left\{q/(1+exp[-b_1(t-c_1)]) + p/(1+exp[-b_2(t-c_2)])\right\}$$

$$+ (f-a_1)/(1+exp[-b_3(t-c_3)]) \qquad (1.10)$$

De manera análoga a (1.9), los últimos dos componentes empiezan a actuar un poco antes de que el componente anterior alcance su asintota superior, aunque "sólo dos de ellos están activos simultáneamente" (Bock y Thissen, 1980:283). Este modelo involucra 9 parámetros y supone conocida a la estatura final con el fin de forzar a la curva a pasar por ese punto.

Como se esperaría, en virtud de la gran cantidad de parámetros, con este modelo se obtuvieron ajustes muy buenos en términos de varianza residual con observaciones longitudinales entre 1 y 20 años para las funciones de estatura y de velocidad (El Lozy, 1978), llegando incluso a mostrar la aceleración del crecimiento que ocasionalmente es posible observar entre los 6 y 8 años de edad. Este fenómeno es conocido como *mid-growth spurt* (Bock y Thissen, 1980, Molinari *et al.*, 1980, Gasser *et al.*, 1985) y, como se verá en la siguiente sección, no se observa con el modelo de Preece-Baines.

El Lozy (1978) ha señalado una objeción más a esta clase de modelos al afirmar que "mientras que matemáticamente es adecuado tener dos procesos de crecimiento simultáneamente, es difícil interpretar tal situación biológicamente". Otra comparación (desfavorable) para los modelos de Bock y Thissen puede verse en Preece y Baines (1978).

1.2.5) MODELOS DE PRECCE-BAINES

En 1978 M.A. Preece y M.J. Baines presentaron una familia de modelos con la que se tienen ajustes muy superiores a los del doble logístico y casi tan buenos como los del triple logístico pero con la mitad de sus parámetros y sin sus limitaciones y problemas estructurales. La construcción de esta familia proviene del sistema de ecuaciones diferenciales

$$df/dt = \sigma(t)\ (h_1 - f)$$

$$d\sigma/dt = \gamma(\sigma_1 - \sigma)\cdot(\sigma - \sigma_0),$$

siendo h_1 la estatura final. Nótese que la primera ecuación diferencial es parecida a la ecuación (1.4). Además, en virtud de una reparametrización, se puede mostrar que el modelo logístico satisface a la segunda ecuación diferencial, siendo σ_0 y σ_1 las asíntotas de tal logística. La construcción de la familia de Preece-Baines aparece en el apéndice 1.

El sistema anterior admite como soluciones a las siguientes funciones:

$$f_1(t;\underline{\theta}) = h_1 - 2(h_1-h_\theta)/[exp(\sigma_0(t-\theta)) + exp(\sigma_1(t-\theta))] \quad (1.11)$$

$$f_2(t;\underline{\theta}) = h_1 - (h_1-h_\theta)/(.5exp(\gamma\sigma_0(t-\theta)) + .5exp(\gamma\sigma_1(t-\theta)))^{1/\gamma} \quad (1.12)$$

$$f_3(t;\underline{\theta}) = h_1-4(h_1-h_\theta)/\Big[(exp(p_0(t-\theta))+exp(p_1(t-\theta)))\cdot(1+exp(q_1(t-\theta)))\Big] \quad (1.13)$$

donde h_1 es la estatura adulta; h_θ la estatura en la edad θ; σ_0, σ_1, p_0, p_1 y q_1 tasas constantes; θ es la edad en la cual se centra la logística $\sigma(t)$ en el sistema de ecuaciones diferenciales y γ una constante adimensional. Estas ecuaciones se conocen como modelos 1, 2 y 3 de Preece-Baines. Hay que hacer notar que, al contrario de lo que sucede con los modelos doble y triple logísticos, la estatura final es un parámetro por estimar. En las gráficas 11, 13 y 15 aparecen curvas de estatura, velocidad y aceleración para ambos sexos obtenidas con el modelo 1.

Hay que mencionar que este trabajo fue el único de los consultados en el que se reportaron con cierto detalle aspectos computacionales de los ajustes: convergencias, número de iteraciones, posibilidades de mínimos locales, etc. El modelo 1 fue el que menos problemas presentó en este rubro, mientras que el modelo 3 necesitó de más iteraciones para producir resultados satisfactorios y el modelo 2 no convergió en casi una cuarta parte de la muestra, mostrando, además, evidencia de presencia de soluciones no únicas. En cuanto a la varianza residual, no hubo diferencias significativas entre los tres modelos, si bien el modelo 3 (que tiene la reparametrización más pesada) tuvo los mejores resultados. Es interesante notar que esta estadística fue menor para los ajustes femeninos.

En cuanto al comportamiento distribucional de los residuales, Preece y Baines no encontraron evidencia de desviaciones para los supuestos de normalidad y homoscedasticidad. Esto lo evaluaron agrupando a los residuales en función del porcentaje de estatura adulta alcanzado y obteniendo estadísticas descriptivas para cada uno de estos grupos.

Al contrario de lo reportado en otros estudios, en el trabajo de Preece y Baines se da cuenta del comportamiento no independiente de los residuales. Estos autores aplicaron la estadística de Durbin-Watson (1971) y la prueba de rachas (Gibbons, 1971:63) a los residuales, encontrando evidencia de autocorrelación, siendo ésta ligeramente mayor para el primer modelo. Prácticamente todos los individuos tuvieron valores negativos en el valor normal estandarizado de la prueba de rachas, lo que indica un número de cambios de signo más bajo que el esperado bajo el supuesto de aleatoriedad. En general, esta prueba indicó una correlación más fuerte que la detectada con la estadística de Durbin-Watson.

Es importante señalar que Preece y Baines no toman en cuenta que ambas pruebas suponen que las observaciones están equiespaciadas, condición que, en general, no se cumple en sus datos (en casi todos los individuos, las mediciones fueron semestrales en la infancia, trimestrales en la pubertad y adolescencia y anuales de ahí en adelante). El hecho de que el número observado de cambios de signo para los residuales fuera consistentemente menor que el esperado bajo aleatoriedad indica la presencia de autocorrelación positiva de observación a observación. Puesto que en la adolescencia éstas se realizaron más frecuentemente que en cualquier otra etapa, lo que aumentaría la autocorrelación observada, es posible que las conclusiones que obtienen Preece y Baines respecto a la presencia de autocorrelación en los residuales sean más severas de lo que realmente hay en los modelos.

En resumen, estos autores concluyen que "si se está dispuesto a aceptar un cierto grado de autocorrelación, los tres modelos dan una representación aceptable de la curva de crecimiento" (Preece y Baines, 1978:18-19).

Es también importante la evidencia encontrada por estos autores en el sentido de que el modelo 1 subestima ligeramente la estatura desde el inicio de la curva hasta el punto en que se alcanza el 60-65% de la talla final. posteriormente, hay unatendencia a la sobreestimación (pequeña pero consistente) hasta la región del 70-75%, para finalmente no tener un patrón definido hacia la última parte de la curva. Estas tendencias no se observaron en los otros dos modelos.

En el trabajo de 1978, se propone la estimación de otros "parámetros biológicos" a partir de los estimadores obtenidos con los modelos originales. Estos aparecen, junto con las abreviaturas que se usarán en el resto del trabajo en la tabla 1.1. al final de este capítulo.

El análisis de estos nuevos parámetros proporciona la capacidad de profundizar en las relaciones entre diversos factores que intervienen en el proceso de crecimiento y da una justificación más al uso de esta clase de modelos como herramientas de primera importancia en estudios de crecimiento.

Más adelante se mencionarán varios trabajos en los que se aplicó el modelo 1; éste fue el único modelo de Preece-Baines del que se encontraron aplicaciones. Esta actitud se debe no sólo a que es el que tiene el menor número de parámetros, sino tal vez también al hecho de que es el más recomendado por Preece y Baines, además de ser el único miembro de la familia que mencionan Preece (1978) y Preece y Heinrich (1981) en sus trabajos de revisión sobre curvas de crecimiento. Todo ésto motivó que el modelo 1 (PB1) fuera elegido para trabajar en esta tesis.

Algunos trabajos en los que se aplicó el modelo PB1 se utilizó son los siguientes: Hauspie *et al.* (1980a, 1980b), Mirwald *et al.* (1981), Brown y Townsend (1982, 1983), Cameron *et al.* (1982), Billewickz y McGregor (1982), Tanner *et al.* (1982), Zacharias y Rand (1983, 1986), Shohoji y Sasaki (1984). Algunas ideas contenidas en estos trabajos fueron utilizadas en el análisis de los datos.

1.2.6) MODELO DE SHOHOJI-SASAKI

La proposición más reciente de un modelo paramétrico para el período completo del crecimiento encontrada en la literatura revisada es la de Shohoji y Sasaki (1984). La derivación del modelo no está expuesta en ese trabajo. El modelo consiste en la suma de dos funciones no lineales y puede escribirse como:

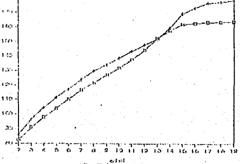
$$f(t;\underline{\theta}) = \mathcal{U}\cdot exp(-e^{a-bt}) + \rho(t)\cdot(1 - exp(-e^{a-bt}), \qquad (1.14)$$

siendo $\rho(t)$ la función de Count (ecuación 1.3), y $\mathcal{U}$, a y b parámetros, interpretándose a $\mathcal{U}$ como la estatura *"at the young adulthood"* (Shohoji y Sasaki, 1984:103). Este modelo tiene 6 parámetros. La justificación de esta ecuación está en términos de que el segundo sumando puede considerarse como la parte que determina a la estatura antes de la adolescencia decreciendo rápidamente conforme t crece, siendo entonces el primer componente (una función de Gompertz) el que produce principalmente el efecto de la aceleración durante la adolescencia.

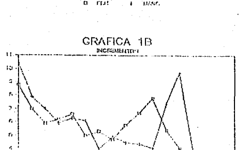
El trabajo de Shohoji y Sasaki es interesante más bien por ser el único en el que se encontraron estimaciones (utilizando el modelo 1.14) de algunas de medidas de no linealidad que se discutirán en los capítulos siguientes (en particular, el sesgo de Box) y porque presenta un estudio de simulación con el que genera varias edades y estaturas para un mismo individuo con el fin de analizar el comportamiento del sesgo del estimador al incrementar el tamaño de muestra.

Finalmente, cabe mencionar que la representación que se consigue de la curva de crecimiento no son muy buena, como puede apreciarse en las gráficas 1L y 1M. En ellas aparecen, calculados para ambos sexos, el modelo de Shohoji y Sasaki y su curva de velocidad mensual, respectivamente.

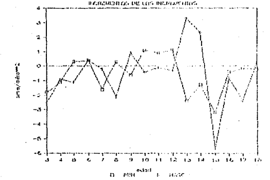
GRAFICA 1A



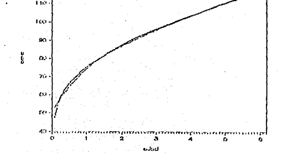
GRAFICA 1B



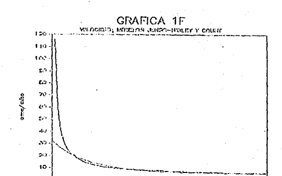
GRAFICA 10



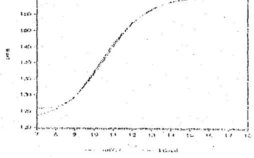
GRAFICA 1E



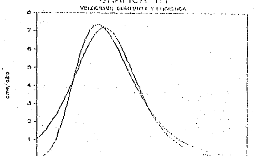
GRAFICA 1F



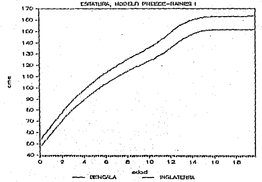
GRAFICA 1G



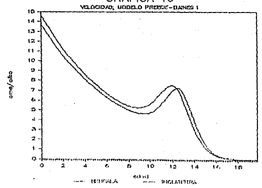
GRAFICA 1H



GRAFICA 1I

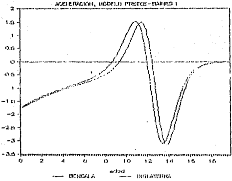


GRAFICA 1J

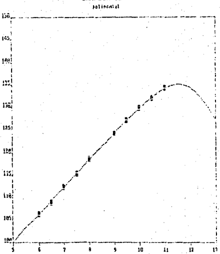


GRAFICA 1K

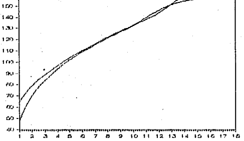
ACELERACION, MODELO PREECE - BAINES 1



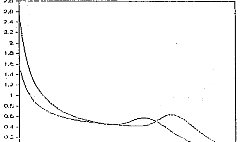
GRAFICA 1D


polinomial
150
145
140
135
130
125
120
115
110
105
100
5
6
7
8
9
10
11
12
13
edad

GRAFICA 3L



GRAFICA 3M



2) METODOS DE ESTIMACION Y PROCEDIMIENTOS DE COMPUTO

2.1) INTRODUCCION

Como se mencionó en el capitulo 0, el interés principal de este trabajo consiste en la comparación entre el uso del método de mínimos cuadrados ordinarios (MCO) y dos casos particulares del método de mínimos cuadrados generalizados (MCG), en la estimación de los 5 parámetros del modelo de crecimiento humano Preece-Baines 1 (PB1).

La diferencia entre ambos métodos radica en que la estimación MCO supone que matriz de varianza covarianza de los errores es la idéntica multiplicada por una constante, mientras que los métodos de mínimos cuadrados generalizados consideran que los errores pueden estar correlacionados. Dentro de la segunda clase de estimadores se trabajó con dos métodos de estimación: uno de ellos supone que la matriz de varianza covarianza de los errores es circular (Hannan, 1983:72); el otro impone a los errores propiedades de un proceso autorregresivo de algun orden determinado (Ballant y Grebel, 1976).

Hay varias dificultades para resolver el problema general planteado en el primer párrafo:

1) El modelo PB1 es no lineal en los parámetros, por lo que la estimación deberá realizarse por procedimientos iterativos; como consecuencia de ésto, será necesario controlar aspectos referentes a la calidad numérica de los ajustes (cotas de convergencia y tolerancia, valores iniciales, número de iteraciones, etc.). Por otra parte, los estimadores así obtenidos tendrán las propiedades de los estimadores correspondientes al caso lineal solamente en

forma asintótica (Malinvaud, 1966; Ratkowsky, 1983;§1.1; Gallant, 1987:§4), lo que hace necesario evaluar qué tan cerca se está de haber alcanzado con suficiente aproximación tales propiedades. Esto depende, en cierto grado, de "qué tan no lineal es un modelo no lineal", por lo que es importante establecer medidas de no linealidad para el modelo en cuestión (Beale, 1960, Box, 1971; Bates y Watts, 1980).

2) Ya que las observaciones longitudinales consisten en una serie de tiempo, los residuales podrían estar correlacionados, violando así uno de los supuestos usuales en regresión. Hay alternativas para incorporar a la estimación no lineal supuestos que hagan explícita la presencia de esta autocorrelación. (Gallant y Goebel, 1976; Glasbey, 1979; Gallant, 1987:§2.7); éstas son examinadas en este trabajo.

3) La única forma de estimar la variabilidad individual es a través de observaciones longitudinales; sin embargo, éstas son generalmente difíciles de obtener en un número suficiente para establecer en un modo satisfactorio las diferencias entre los métodos de estimación. Esto, aunado a los problemas que se mencionaron en el inciso 1), y al hecho de que las propiedades asintóticas de los estimadores para el caso no lineal funcionarán cuando los errores sean no correlacionados sugiere el uso de métodos Monte Carlo, tanto para la comparación de los métodos de estimación como para el estudio de las desviaciones que presentan los estimadores acerca de tales propiedades al incorporar una estructura de errores no independientes.

En este capítulo se revisan brevemente los principios que subyacen en los métodos de estimación de interés.

Antes de iniciar la exposición, es necesario señalar que en el transcurso del trabajo se intentó atacar el problema con el método de apoyo (*method of support*) presentado, fundamentalmente, en Edwards (1984) y en Sprott (1984). Este método consiste en, a partir del análisis directo de la función de verosimilitud, establecer niveles de plausibilidad sobre distintos posibles valores parámetro de interés sin hacer uso de supuestos de muestreo hipotético presentes, por ejemplo, en la construcción de regiones de confianza (Cox 1977). Así, se intentó construir regiones de plausibilidad para el parámetro a través de la función de log verosimilitud relativa (que era una función con 5 argumentos). Esto hubiera permitido evitar la construcción de regiones de confianza que eran aproximaciones probablemente insuficientes. El procedimiento para examinar esta función consistió en examinar la variación en cada componente de la log verosimilitud en una vecindad de su estimador de máxima verosimilitud (MV) determinada con ayuda de la inversa de la matriz de información esperada (Kalbfleisch, 1979:§11) fijando a los otros 4 componentes en sus estimadores MV. El problema con este método reside en que es muy difícil examinar una superfice de 5 dimensiones. Otro enfoque hubiera sido construir rejillas de dimensión 5 centradas en el estimador MV de todo el vector $\underline{\theta}$ y examinar la función de log verosimilitud relativa en cada celda para intentar tener una idea de su situación. Esta última aproximación parece ser aún más complicada de interpretar.

2.2) PRIMER METODO: ESTIMACION DE MINIMOS CUADRADOS ORDINARIOS

Es el método de estimación más ampliamente utilizado. De hecho, fuera del artículo de Dock y Thissen (1980), que utiliza al método bayesiano empírico para los modelos doble y triple logístico, no se encontró trabajo alguno en el que se utilizara otro método de estimación.

A partir de la ecuación (1.1), se define al error para la j-ésima observación como:

$$\varepsilon_j = y_j - f(x_j;\underline{\theta}^*),$$

siendo $\underline{\theta}^*$ el verdadero (y desconocido) valor del parámetro. Estas diferencias pueden calcularse sobre cualqier otro posible valor de $\underline{\theta}$, lo que define a las funciones

$$e_j(\underline{\theta}) = y_j - f(x_j;\underline{\theta}), \tag{2.1}$$

que son llamadas residuales. Claramente, $\varepsilon_j = e_j(\underline{\theta}^*)$.

El método de MCO consiste en encontrar los valores de $\underline{\theta}$ que minimizan a la función objetivo

$$\Phi(\underline{\theta}) = \sum_{j=1}^{n} (e_j(\underline{\theta}))^2 = \| \underline{e} \|^2 = \underline{e}'\underline{e} \tag{2.2}$$

donde $\underline{e}$ es el vector de residuales (de orden $n\times1$) y $\|\cdot\|$ es la norma de un vector.

Los supuestos básicos en la estimación MCO son los siguientes:

1) Los valores de las x's han sido observados sin error.

2) $\mathscr{E}(\varepsilon_j) = 0$.

3) $\mathscr{E}(\varepsilon_j^2) = \sigma^2$.

4) $\mathscr{E}(\varepsilon_i \cdot \varepsilon_j) = 0$, si $i \neq j$.

Las condiciones 1) y 3) implican que $\mathscr{E}(y_j) = f(x_j; \underline{\theta}^*)$ y que $var\ (y_j) = \sigma^2$; la condición 4) implica la no correlación entre residuales. Nótese que no se ha mencionado alguna necesidad de que la distribución de los errores sea especificada, si bien implícitamente se está suponiendo que ésta es simétrica alrededor de su esperanza (Randless y Wolfe, 1979:232), supuesto que resulta adecuado en situaciones en que parte del error se explica por errores de medición en los que no hay necesariamotne una mayor probabilidad de realizar observaciones equivocadas hacia valores mayores o menores que el verdadero.

En el caso de tener un modelo de regresión lineal en los parámetros, i.e., de la forma

$$\underline{y} = X\, \underline{\theta} + \underline{\varepsilon},$$

donde X es la matriz diseño, de orden nxp, el procedimiento para encontrar el valor del estimador MCO es derivar parcialmente a $S(\underline{\theta})$ respecto a cada componente de $\underline{\theta}$ y resolver el sistema de ecuaciones lineales homogéneas resultante. Estas ecuaciones se llaman ecuaciones normales. La solución será única cuando la matriz diseño sea de rango completo.

En el caso no lineal, las ecuaciones normales son:

$$\partial S/\partial\theta_k = -2 \sum_{j=1}^{n} e_j\ \partial f/\partial\theta_k = 0, \quad (k = 1,\dots p). \tag{2.3}$$

Las ecuaciones normales en el caso lineal eran funciones únicamente de las x's, puesto que al derivar con respecto al parámetro, éste se eliminaba, lo que permitía encontrar soluciones analíticas para el estimador. En el caso no lineal, las ecuaciones normales pueden ser no lineales en $\underline{\theta}$ y, en general, no admiten una solución analítica, por lo que el estimador deberá calcularse por medio de métodos numéricos que resuelvan aproximadamente a las ecuaciones normales, lo que puede ser difícil y computacionalmente costoso. En adelante, un estimador MCO se denotará como $\hat{\underline{\theta}}$.

El método más utilizado para encontrar un mínimo de la función objetivo (2.2) es el de Gauss-Newton. En este trabajo se utilizó la rutina NL2SOL (Calderón et al., 1983), basada en un algoritmo de tipo Gauss-Newton modificado que incorpora ciertas opciones para decidir cuándo utilizar una aproximación del operador Hessiano de la función objetivo. La idea general del método numérico de NL2SOL consiste en utilizar una variación del método de Gauss-Newton para encontrar la raíz del gradiente de la función objetivo $\Phi(\underline{\theta})$.

La base del método Gauss-Newton está en sustituir en la función objetivo una expansión de f en series de Taylor, despreciando los términos de orden mayor o igual a 2 (Gallant, 1976a:76), lo que puede expresarse como:

$$f(x_i;\underline{\theta}^*) \simeq f(x_i;\underline{\theta}_o) - \sum_{k=1}^{p} [f(x_i;\underline{\theta}_o)/\partial\theta_k] \cdot (\theta_k - \theta_{ko}) \qquad (2.4)$$

donde $\underline{\theta}_o$ es un punto que, se espera, está cerca de $\underline{\theta}^*$.

Si se hacen

$$f_i^o = f(x_i;\underline{\theta}_o),$$

$$\beta_k^o = \theta_k - \theta_{ko},$$

se tiene que f_o y b_o son los vectores de orden nx1 y px1 con componentes de las formas f_i^o y β_k^o , respectivamente, y

$$F_{ik}^o = \partial f(x_i;\underline{\theta}_o)/\partial\theta_k,$$

entonces la aproximación (2.4) implica que la función objetivo puede escribirse como:

$$S_T(\underline{\theta}) = \sum_{i=1}^{n} [y_i - f_i^o - \sum_{k=1}^{p} \beta_k^o F_{ik}^o]^2;$$

el subíndice T en S se debe a la aproximación por series de Taylor usada.

Esta suma de cuadrados es lineal en b_o, por lo que puede resolverse por métodos usuales como solución respecto a $b_o = \underline{\theta}-\underline{\theta}_o$. Sea ahora $b_1 = \underline{\theta}_1 - \underline{\theta}_o$. Entonces, el vector $\underline{\theta}_1$ puede considerarse como una nueva aproximación para la estimación de $\underline{\theta}^*$ al sustituirlo en lugar de $\underline{\theta}_o$ en la aproximación lineal anterior y obteniendo otra solución para la minimización de la suma de cuadrados correspondientes que ahora tendrá superíndices 1 en vez de o. Esto llevará a otra aproximación lineal, ahora en $\underline{\theta}_2$, y así sucesivamente.

Las iteraciones continúan hasta que se satisfaga algún criterio de convergencia. En este trabajo se examinaron tanto la convergencia en los valores de la función objetivo como la convergencia del gradiente a 0. El método puede converger lentamente, oscilar o aun divergir dependiendo de qué tan adecuada sea la aproximación lineal utilizada. Además, si la aproximación inicial θ_0 no está relativamente cerca de θ^*, el método puede converger a un mínimo local de la función objetivo, lo que podría proporcionar una estimación incorrecta.

Por otra parte, en el caso lineal, el carácter lineal del estimador MC en las observaciones ($\hat{\theta} = A\underline{y}$) es una condición para demostrar que los estimadores MC son insesgados; igualmente, el teorema de Gauss-Markov (Bard, 1974:319-321) asegura que $\hat{\theta}$ es el mejor estimador lineal insesgado (*meli*) para θ^*. Más aun, si se desconocen θ^* y σ^2, y se supone que el vector de errores se distribuye $N_n(0,\sigma^2\Sigma)$, siendo Σ la matriz de varianza covarianza de los errores, entonces el estimador

$$s^2 = \Phi(\hat{\theta})/(n-p)$$

es un *meli* para σ^2, $\hat{\theta}$ es de varianza mínima no sólo dentro de la clase de estimadores lineales insesgados, sino dentro de todos los estimadores insesgados y, finalmente, la distribución muestral de $\hat{\theta}$ será normal con media en θ^* y con matriz de varianza covarianza

$$V(\hat{\theta}) = (X'\Sigma^{-1}X)^{-1} .$$

En el caso no lineal, las soluciones del sistema de ecuaciones normales no son, en general, expresables como combinaciones lineales de los valores muestrales de la variable dependiente, por lo que $\hat{\theta}$ no será insesgado, ni de mínima varianza, ni distribuido normalmente. Así, en el caso no lineal, las propiedades del estimador MCO son "esencialmente desconocidas" (Ratkowsky, 1983:6). Sin embargo, si se supone que el vector de errores se distribuye normalmente, y bajo ciertas condiciones de regularidad (Gallant, 1987:§4.2), el estimador MCO converge casi seguramente a $\underline{\theta}^{*}$ y coincide con el estimador MV por lo que las propiedades asintóticas de esta última clase de estimadores, como insesgamiento, mínima varianza y distribución asintóticamente normal (Cox y Hinkley, 1979:§10) funcionan, aún con tamaños de muestra relativamente pequeños (Ratkowsky, 1983:7).

Otra forma de analizar la calidad de las estimaciones en el caso no lineal consiste en examinar el grado de no linealidad del modelo respecto a los parámetros. Esto tiene relevancia no sólo por la parte numérica de la estimación, que requiere de una aproximación lineal; también está la cuestión acerca de la pertinencia del uso de la teoría asintótica, basada en estimadores lineales, por lo que, como se ha visto, si el modelo no fuera muy no lineal, los estimadores serían más parecidos a estimadores lineales, lo que daría apoyo al uso de esta teoría con tamaños de muestra pequeños.

Respecto al primer punto, Hartley (1961) afirma que el procedimiento de estimación no lineal basado en métodos de Gauss-Newton casi siempre converge satisfactoriamente, dependiendo de la precisión de la máquina, de la aproximación inicial $\underline{\theta}_{o}$ y del modelo. Puesto que en este caso se cuenta con una computadora muy precisa (la Burroughs B7800 de la Dirección de Cómputo Académico de la UNAM) y una rutina altamente robusta, además de que la experiencia con el modelo de Froese-Baines indica que se tiene una convergencia rápida (Froese y Baines reportan un promedio de 7 iteraciones con una tolerancia de 10^{-8} para la diferencia absoluta en cada iteración) y se dispone de aproximaciones iniciales de buena calidad, no se insistirá mucho en la parte numérica de la estimación. Por lo anterior, la atención se concentrará en la no linealidad del modelo.

Como señala Ratkowsky (1983:13), "cuando se incrementa el tamaño de muestra, el comportamiento de los estimadores MCO no lineales se hace cada vez más lineal", ésto como consecuencia de la linealización realizada vía teorema de Taylor, lo que permite extender asintóticamente las propiedades de la regresión lineal (Malinvaud, 1970:957). Jennrich (1969) demuestra que bajo ciertas condiciones de regularidad los estimadores MC no lineales son consistentes. Este resultado es generalizado por Hannan (1971) para el caso en que los errores no se distribuyen independientemente sino que se suponen generados por un proceso estocástico estacionario. Gallant y Goebel (1976) y Glasbey (1979) discuten más ampliamente el problema de la estimación MCO no lineal con residuales autocorrelacionados. Más adelante se regresará a este punto.

El primer trabajo que consideró medidas de no linelidad fue el de Beale (1960), quien sugiere unas medidas de curvatura para el espacio de soluciones de las ecuaciones normales, ya que en el caso lineal este espacio es un hiperplano. Bates y Watts (1980) dividen la no linealidad de un modelo en dos componentes: uno llamado de "no linealidad intrínseca", debido a la curvatura del espacio de soluciones y otro de "no linealidad causada por efectos paramétrales", que mide el espaciamiento y la curvatura de las proyecciones de los parámetros sobre el espacio de las soluciones;

en el caso lineal, tales proyecciones son equiespaciadas y paralelas.

Guttman y Meeter (1965) demostraron que las medidas de curvatura de Beale tienden a subestimar la no linealidad. Bates y Watts (1980) proponene correcciones a tales medidas.

Por su parte, Gillis y Ratkowsky (1978) encuentran, usando métodos Monte Carlo, que el sesgo del estimador calculado por el método de Box (1971) es un buen indicador para la no linealidad, principalmente en aspectos asociados a los efectos del parámetro.

Bates y Watts (1980) demuestran que este hecho se debe principalmente a que el sesgo de Box está relacionado con la medida propuesta por ellos para medir la no linealidad por efectos parametrales. En todos los casos estudiados por estos autores, esta medida resultó ser mayor que la propuesta para la no linealidad intrínseca, lo que "ayuda a explicar porqué Gillis y Ratkowsky encontraron que el sesgo de Box ... estaba fuertemente relacionado con las propiedades del estimador MCO" (Ratkowsky, 1983:14).

Según Ratkowsky (1983:186), la medida de no linealidad intrínseca (que, en vista de los estudios de Bates y Watts podría suponerse despreciable si es que se tiene una medida de no linealidad por efectos parametrales suficientemente pequeña) está asociada a las diferencias en las predicciones para los valores de la variable respuesta, mientras que el otro componente de no linealidad tiene que ver más bien con las propiedades estadísticas de los estimadores MCO (Ratkowsky, 1983:190).

El uso del sesgo de Box como única medida de no linealidad en este trabajo obedeció a que tiene una interpretación más directa que las curvaturas de Beale y Bates y Watts; además permite estimar el sesgo particualr de cada uno de los componentes de $\hat{\underline{\theta}}$.

Si $\underline{\phi} = \hat{\underline{\theta}} - \underline{\theta}^{*}$ es el sesgo del estimador MCO, se puede demostrar que, suponiendo normalidad y homoscedasticidad,

$$\mathcal{B}(\underline{\phi}) = -.5\sigma^{2} \, [\sum_{i=1}^{n} f_i' f_i]^{-1} \sum_{j=1}^{n} f_j' \; tr \; ([\sum_{k=1}^{n} f_k' f_k]^{-1} H_j) \qquad (2.5)$$

donde

$$f_i = \nabla f(x_i ; \underline{\theta}^{*}),$$

tr (.) indica la traza de una matriz y H_j es la matriz hessiana (de orden pxp) evaluada en la j-ésima observación y en $\underline{\theta}^{*}$.

En el caso lineal, las matrices H_j son cero, por anularse las segundas derivadas parciales de un modelo lineal, con lo que $\mathcal{B}(\underline{\phi})$=0. Además, si la varianza se puede reducir, el sesgo también disminuirá. Para calcular estimaciones del sesgo, se sustituyen los estimadores s^2 y $\hat{\underline{\theta}}$ en (2.5). En todo caso el sesgo es menor que el error estándar de $\hat{\underline{\theta}}$ (Box, 1971:177).

2.3) SEGUNDO Y TERCER METODOS: ESTIMACION MCG

Puede demostrarse que es equivalente minimizar $\mathcal{B}(\underline{\theta})$ a maximizar a la función de verosimilitud en el caso no lineal con errores distribuidos normalmente (Bard, 1974:63). Esta propiedad es cierta aun si los errores están autocorrelacionados; en tal caso, el método de máxima verosimilitud corresponde al método de MCG, *i.e.*, se debe encontrar el estimador $\hat{\underline{\theta}}$ que minimice la suma de cuadrados expresada como:

$$\underline{e}\, \Sigma_e^{-1}\, \underline{e}\,, \tag{2.6}$$

siendo Σ_e la matriz de varianza covarianza de los errores.

Si se supone que la distribución conjunta de los errores es normal multivariada de orden n (siendo n el número de observaciones longitudinales para cada individuo) con media $\underline{0}$ y matriz de varianza covarianza Σ_e, la función de log verosimilitud puede escribirse como:

$$\lambda(\underline{\theta}) = C - .5 \sum_{j=1}^{n} \sum_{k=j}^{n} \sigma^{jk}\, e_j(\underline{\theta})\, e_k(\underline{\theta}) \tag{2.7}$$

donde C es un término independiente de $\underline{\theta}$, e_j es el j-ésimo residual y σ^{jk} es el elemento (j-k) de Σ_e^{-1} (Bock y Thissen, 1980:270).

En las ecuaciones (2.8) y (2.9) Σ_e se supone conocida. De no ser así, se debe encontrar un estimador y sustituirlo en esas fórmulas.

Gallant y Goebel (1976) presentan dos formas de estimar Σ_e. Ambas serán seguidas en este trabajo.

La primera es la desarrollada inicialmente por Hannan (1971) y utiliza una matriz circular simétrica de Toeplitz (Hannan, 1983:2; Fuller, 1976:9; véase su expresión 2.), *i.e.*:

$$\hat{\Sigma}_e = \begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \dots & \hat{\gamma}(q-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & \dots & \hat{\gamma}(q-2) \\ \vdots & \vdots & & \vdots \\ \hat{\gamma}(q-1) & \hat{\gamma}(q-2) & \dots & \hat{\gamma}(0) \end{bmatrix}, \tag{2.8}$$

estimándose las autocovarianzas de orden h (γ(h)), bajo el supuesto de que los errores se comportan de acuerdo a un proceso estacionario de segundo orden, con:

$$\gamma_h^a = (1/(n-h))\sum_{t=1}^{n-h} e_t \cdot e_{t+h} \qquad (2.9)$$

o con:

$$\gamma_h^b = (1/n)\sum_{t=1}^{n-h} e_t \cdot e_{t+h} \,. \qquad (2.10)$$

Ambos estimadores de γ_h son asintóticamente insesgados (Fuller, *op. cit.* §6.2); además, si bien γ_h^a puede tener un error cuadrático medio menor que γ_h^b para ciertas series de tiempo, éste garantiza que la matriz Toeplitz construida sea positiva definida.

Aunque γ_h^a puede resultar intuitivamente más atractivo, ya que pondera diferencialmente a las autocovarianzas según el número de términos involucrados en su estimación, Gallant y Goebel utilizan a γ_h^b, pues tienen n = 254 y hacen h = 1, 2,..., 60, con lo que las autocovarianzas son estimadas siempre con un número de términos bastante grande. Aquí se utilizará a γ_h^a.

El método de estimación que minimiza a la ecuación (2.6) estimando circularmente a Σ_e^{-1} se denominará MCO-CIRC.

En el trabajo de Gallant y Goebel se muestra (mediante un estudio de simulación) que, para el modelo de crecimiento exponencial en el caso de que los errores sean estacionarios, es posible imponer restricciones más fuertes que las implicadas por la estructura de una matriz banda circular simétrica; por ejemplo, se puede suponer que los errores forman un proceso autorregresivo de orden q, (AR(q)) es decir, si $\{e_t\}$ son los errores y $\{\omega_t\}$ es una sucesión de variables aleatorias *iid* $N(0,\sigma^2)$, entonces:

$$e_t + a_1 e_{t-1} + a_2 e_{t-2} + \cdots + a_q e_{t-q} = u_t, \quad (t = 0, \pm 1, \pm 2, \ldots),$$

siendo las raíces del polinomio $z^q + a_1 z^{q-1} + \cdots + a_q$ menores que 1 en valor absoluto.

El uso de este supuesto en la estimación de Σ_e, si q es pequeña, parece incrementar la eficiencia para los estimadores MC generalizados con muestras pequeñas, respecto al uso del estimador circular.

El procedimiento de estimación utilizado bajo el supuesto autorregresivo aparece en Gallant y Goebel (1976:62). Primeramente se obtienen los residuales a partir de la estimación por MC ordinarios $(\hat{e}_t)_1^n$ para calcular las autocovarianzas hasta de orden q, mediante la ecuación,

$$\hat{\gamma}_h = (1/n)\sum_{t=1}^{n-|h|} \hat{e}_t \hat{e}_{t+|h|}, \quad (h = 0, 1, \ldots, q). \quad (2.11)$$

Sean

$$\hat{\Gamma}_q = \begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \cdots & \hat{\gamma}(q-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & \cdots & \hat{\gamma}(q-2) \\ \vdots & \vdots & & \vdots \\ \hat{\gamma}(q-1) & \hat{\gamma}(q-2) & \cdots & \hat{\gamma}(0) \end{bmatrix}, \quad (2.12)$$

$$\hat{\gamma}_q = [\hat{\gamma}(1), \hat{\gamma}(2), \ldots, \hat{\gamma}(q)]', \quad (2.13)$$

y $F(\theta)$ la matriz de $n \times p$ cuyo t-ésimo renglón es el vector gradiente de f evaluado en la t-ésima observación.

Entonces, usando las ecuaciones de Yule-Walker (Chatfield, 1983;§4.2.1) se tiene que:

$$\hat{\underline{a}} = -\hat{\Gamma}_q^{-1}\,\hat{\gamma}_q \qquad (2.14)$$

y

$$\hat{\sigma}^2 = \hat{\gamma}(0) + \hat{\underline{a}}'\,\hat{\gamma}_q\,. \qquad (2.15)$$

Haciendo

$$\hat{\Gamma}_q^{-1} = \hat{P}_q'\,\hat{P}_q\,,$$

con una descomposición de Cholesky (Kennedy y Gentle, 1980;§7.4; Calderón, 1985), y

$$\hat{P} = \left(\begin{array}{cccccccc|c} & \sqrt{\hat{\sigma}^2} & \hat{P}_q & & & & & & 0 \\ \hline \hat{a}_q & \hat{a}_{q-1} & \cdots & \hat{a}_1 & 1 & & & & \\ & \hat{a}_q & \hat{a}_{q-1} & \cdots & \hat{a}_1 & 1 & & & \\ & & \ddots & \ddots & & \ddots & \ddots & & \\ & & & \hat{a}_q & \hat{a}_{q-1} & & \hat{a}_1 & 1 & \end{array}\right) \begin{array}{l} \} \; q \text{ ren.} \\ \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} (n-q) \text{ ren.} \end{array} \qquad (2.16)\,,$$

se obtiene al estimador MCG- AR1 (así llamado por el supuesto autorregresivo para los errores; se denotará como $\tilde{\underline{\theta}}$) minimizando a la forma cuadrática:

$$Q_n(\underline{\theta}) = (1/n)[\hat{P}\underline{y} - \hat{P}f(\underline{\theta})]'[\hat{P}\underline{y} - \hat{P}f(\underline{\theta})], \qquad (2.17)$$

considerando a $\hat{\underline{\theta}}$ como valor inicial para las iteraciones.

Del estimador anterior se pueden calcular

$$\tilde{\sigma}^2 = [\hat{P}y - \hat{P}f(\tilde{\theta})]' [\hat{P}y - \hat{P}f(\tilde{\theta})]/(n-p)$$

y

$$\tilde{C} = [F'(\tilde{\theta})\ \hat{P}'\hat{P}\ F(\tilde{\theta})]^{-1}.$$

Gallant y Goebel demuestran que

$$\sqrt{n}(\tilde{\theta} - \theta^*)$$

se distribuye asintóticamente como una normal multivariada de orden p, siendo $\tilde{\sigma}^2 n\tilde{C}$ un estimador consistente para la matriz de varianza covarianza del estimador. Intuitivamente se esperaría que si el procedimiento de estimación se repitiera utilizando como estimacion inicial a $\tilde{\theta}$, los estimadores serían más eficientes para muestras chicas. Sin embargo, si bien estos autores recomiendan la repetición del procedimiento para por lo menos dos pasos, concluyen que la ganancia esperada en eficiencia sólo será efectiva si los errores efectivamente siguen un proceso autorregresivo; en otros casos es posible que la estimación sea ligeramente menos precisa que cuando sólo se efectúa en un paso.

Una consecuencia de las propiedades asintóticas de $\tilde{\theta}$ que es importante analizar es la cuestión de si la distribución de la estadística

$$t_k = (\tilde{\theta}_k - \theta_k)/\sqrt{\tilde{\sigma}^2\ \tilde{c}_{kk}} \quad (k = 1,\ldots,p) \tag{2.18}$$

donde $\tilde{c}_{kk}$ es el k-ésimo elemento de la diagonal de $\tilde{C}$, para muestras pequeñas puede aproximarse con suficiente precisión por la distribución t de Student con (n-p) grados de libertad, con el fin de establecer intervalos de confianza para $\tilde{\underline{\theta}}$. En adelante se llamará "t-cociente".

Gallant y Goebel encontraron (con métodos Monte Carlo) que, para el modelo de crecimiento exponencial

$$f(x;\underline{\theta}) = \theta_1 exp(\theta_2 x) + \varepsilon,$$

si los errores están autocorrelacionados, los intervalos de confianza construidos a partir de la suposición de que la estadística anterior se distribuye t(n-p) pueden ser demasiado estrechos, *i.e.*, la distribución de (2.10) bajo diferentes supuestos para la estructura aleatoria de los errores (por ejemplo, AR(1), AR(2), ARMA (2,2)) tiene colas considerablemente más pesadas que las de una t de Student. Su idea será retomada en este trabajo para producir tablas con los cuantiles usuales aproximados para la distribución del t-cociente para el modelo PB1 suponiendo diferentes estructuras de error. Estas tablas podrán utilizarse para construir intervalos de confianza aproximados para cada componente del parámetro. Esto último se haría a partir de la expresión

$$\tilde{\theta}_k \pm t_k \, (\tilde{\sigma}^2 \, \tilde{C}_{kk})^{1/2} \, .$$

Por otra parte, una forma de aproximarse al orden del proceso autorregresivo consiste en utilizar alguna estadística como

$$T_h = \tilde{\rho}_h \sqrt{(n+1)} \; / \; [1 + 2 \sum_{j=1}^{h-1} \tilde{\rho}_j^2 \;]^{1/2} \qquad (2.19)$$

que, asintóticamente tiene una distribución normal (Bowerman y O'Connell, 1979:342). En caso de que la hipótesis H_o: $T_h = 0$ se acepte para toda $h \geq q+1$, es razonable pensar que las autocorrelaciones tienden a 0 para rezagos mayores a q. Otra posibilidad aparece en Chatfield (1983:§4.2.2) y consiste en el análisis de las autocorrelaciones parciales muestrales. Estas se denotan por $\tilde{\pi}_h$ y miden el exceso de correlación para rezagos de orden h que no es explicado por un modelo AR(h-1). Los valores de $\tilde{\pi}_h$ tales que

$$|\tilde{\pi}_h| > z_\alpha /\sqrt{n},$$

siendo α el nivel de significancia y z_α el cuantil $(1-\alpha/2)$ de una normal (0,1), se consideran como significativamente distintos de 0. Puede demostrarse que la función de autocorrelación de un modelo AR(q) se "corta" para rezagos mayores que q; entonces, se esperara que el valor de q corresponda al orden de autocorrelación parcial a partir del cual las $\tilde{\pi}_h$ no difieren significativamente de 0. Tal criterio tambien se aplicó en este trabajo.

Antes de concluir con la presentación de los dos métodos MCG utilizdos es necesario mencionar un caso particular, adecuado en situaciones en que es posible suponer independencia entre los residuales al mismo tiempo que varianzas desiguales respecto a la edad. Esto sucede, por ejemplo, en estudios transversales en los que se ajusta PB1 a la sucesión de medias obtenidas en cada grupo de edad (Tanner *et al*, 1982; Jordán *et al.*, 1979). Puesto que, por un lado, la varianza de la estatura cambia con la edad, y, por otro, se tienen distintos tamaños de muestra en cada grupo de edad, estas medias tendrán diferentes varianzas, expresadas como:

$$var\ (\bar{y}_k) = \sigma_k^2/n_k \quad , \tag{2.20}$$

siendo $\bar{y}_k$, σ_k^2 y n_k la media, la varianza y el tamaño de muestra, respectivamente, para para el k-ésimo grupo de edad.

En este caso, la matriz Σ_e es diagonal, con $var\ (\bar{y}_k)$. Puesto que el interés principal del trabajo estuvo puesto en los ajustes longitudinales, la presentación de ejemplos que utilicen a la ecuación (2.20) se reservará para el apéndice 2.

3) ANALISIS DE RESULTADOS PRELIMINARES

3.1) DATOS INICIALES

Las observaciones utilizadas en este capitulo provienen de estudios realizados en el Departamento de Crecimiento y Desarrollo del Instituto de Pediatría de la Universidad de Londres. Pude disponer de ellos gracias a una cinta grabada en esa Universidad que me fue amablemente proporcionada por la maestra María Elena Sáenz, investigadora del Instituto de Investigaciones Antropológicas de la UNAM, quien estudió bajo la supervisión del profesor J. M. Tanner entre 1981 y 1982.

Básicamente, la cinta contiene cuatro archivos (llamados ICC masculino, ICC femenino, NCH masculino y NCH femenino) con datos longitudinales para varias medidas antropométricas. En una nota aparecida recientemente, Tanner y Cox (1985) informan acerca del banco de datos longitudinales con que cuenta la Universidad de Londres; éste incluye, entre otros, a los archivos mencionados.

Por otra parte, todas las observaciones fueron realizadas por R.H. Whitehouse, lo que ciertamente reduce los errores de medición. Como se mencionó en el capítulo 1, en este trabajo se analizaron principalmente datos referentes a la estatura total. En los trabajos de Cameron *et al.* (1982) y de Tanner *et al.* (1982) pueden verse ajustes del modelo PB1 con otras variables antropométricas. En el primero de ellos se expone una interesante metodología, útil para investigar el orden en el que los diferentes segmentos corporales alcanzan su velocidad máxima de crecimiento. En el segundo se ajusta el modelo PB1 a las medias de datos transversales provenientes de Japón con el fin de analizar el aumento secular tanto en la talla como en el segmento inferior.

Un análisis de datos longitudinales (que incluye a los datos reportados en el trabajo de Tanner, Hayashi, Preece y Cameron (1982)) se presenta en el apéndice 2.

Regresando a los datos aquí analizados, se observó que el espaciamiento temporal en las mediciones variaba tanto entre archivos diferentes como al interior de cada archivo. En los archivos ICC, la mayor parte de las observaciones se efectuaron desde el nacimiento hasta los 19 años en intervalos anuales; sin embargo, hubo una cantidad considerable de individuos para los cuales sólo se contaba con observaciones previas a la adolescencia. Para la mayoría de los casos contenidos en los archivos NCH se tienen mediciones realizadas desde los 2 o 3 años hasta los 19 o 20 años; las observaciones son semestrales hasta antes de los 11 años y entre los 14 y 15 años, trimestrales entre los 11 y los 14 años y anuales a partir de los 15 años.

Originalmente se utilizó todo el material correspondiente a estaturas que venía en la cinta. No se tuvieron ajustes aceptables para todo el conjunto de datos. Para estas primeras estimaciones, lo mismo que para las siguientes, se utilizó la rutina NL2SOL (Calderón *et al*, 1983) con una tolerancia de 10^{-6} para la convergencia y con 50 iteraciones como máximo admisible.

Generalmente, los casos en los que no se logró un ajuste aceptable fueron individuos con 8 o menos observaciones, o bien, tales que no se disponía de observaciones para edades posteriores a la adolescencia. Desde luego, hubo contraejemplos para estas afirmaciones. Si bien hubiera sido interesante estudiar el comportamiento de las estimaciones respecto al tamaño de muestra, en este trabajo no se hizo mucho énfasis en este punto. En la siguiente sección de este capítulo se presentarán estos resultados preliminares.

Al mencionar en el párrafo anterior a situaciones en las que el ajuste no fue aceptable, me refiero a aquellos casos en los que se obtuvieron estimaciones de los parámetros del modelo totalmente equivocadas; el ejemplo más común consistió en la estimación del parámetro de estatura final (h_1) con valores superiores a los 2 metros asociada a estimaciones con edades superiores a los 17 años para el parámetro θ. Esto sucedió generalmente en individuos para los que no se disponía de mediciones en la post adolescencia, lo que indica la influencia de esta clase de observaciones en la correcta estimación del modelo.

3.2) ESTIMACIONES PRELIMINARES

3.2.1) ANALISIS DE LOS ESTIMADORES

Un primer paso consistió en obtener estimadores con el método ordinario de mínimos cuadrados no lineales (MCO) para todos los individuos disponibles. Esto se hizo con los siguientes objetivos: a) probar el comportamiento de la subrutina NL2SOL frente a los datos y el modelo para tener idea de, por ejemplo, el número de iteraciones promedio, las normas de los gradientes de la función objetivo al ser evaluados en los estimadores, etc.; b) detectar casos en los que la estimación fue inadecuada; c) tener resultados que pudieran servir como fuente para generar estimaciones con diferentes métodos y d) explorar el comportamiento de los residuales; ésto último sólo en una forma aproximada, ya que las pruebas utilizadas (rachas y Durbin-Watson) para contrastar la hipótesis de independencia para los residuales suponen que las observaciones están equiespaciadas en el tiempo, lo que no se tenía exactamente para estos datos.

A lo largo del trabajo, se buscó analizar series individuales que hubieran completado su ciclo de crecimiento según el criterio dado por Preece y Baines (1978) de considerar únicamente casos cuyas dos últimas mediciones disponibles tuvieran una diferencia menor que 1 cm. Igualmente, se intentó seleccionar individuos que tuvieran las fechas de medición cercanas a las fechas correspondientes para cada edad exacta.

Para los casos masculinos, el archivo ICC contó con 61 individuos iniciales, de los cuales 47 tuvieron ajustes aceptables (77%), mientras que en el archivo NCH, de 32 individuos, 23 fueron bien ajustado (72%). En el archivo ICC femenino, estos números fueron 58 y 35 (60%) y para el NCH femenino, 28 y 24 (86%).

A partir de esos resultados preliminares y de las consideraciones acerca de los individuos con crecimiento completo y con mediciones efectuadas en las fechas exactas, se decidió eliminar, primeramente, a los datos del archivo ICC femenino, que, dentro de los aceptados, contaba con pocos casos que realmente hubieran completado su ciclo de crecimiento y que satisfacieran los requerimientos acerca del equiespaciamiento y la exactitud de las fechas de medición, así fuera en forma aproximada.

Los valores iniciales utilizados para todas las estimaciones fueron las medias (para cada sexo) de los estimadores reportadas por Preece y Baines (1978) con los datos de los archivos NCH.

Los resultados obtenidos para los archivos ICC masculino, NCH masculino y NCH femenino sin modificación alguna en las edades de medición, aparecen en las tablas 3.1, 3.2 y 3.3, respectivamente. Sólo se presentan los individuos cuyo ajuste fue satisfactorio.

A continuación se explica el significado de las columnas incluidas en estas tablas; posteriormente se discutirán los resultados obtenidos con cada archivo.

Las primeras tres columnas de estas tablas son el número de serie del individuo, el número de observaciones longitudinales disponibles y la forma en que terminó la rutina de mínimos cuadrados no lineales (por convergencia del gradiente o por convergencia en x). Posteriormente están los estimadores para cada componente del parámetro.

Siguen los números de residuales negativos y positivos, el número de rachas que hay en ellos y el resultado de la prueba de rachas. Este aparece indicado a la izquierda de la columna de rachas: los signos "+" y "-" se refieren a que el número de rachas fue mayor o menor, respectivamente, que el esperado bajo la hipótesis de aleatoriedad. Si no aparece signo alguno en esta columna, se entiende que se aceptó la hipótesis de aleatoriedad. Para los archivos NCH, donde se tenía, generalmente, más de 20 observaciones por individuo, se utilizó la estadística z que tiene una distribución normal estándar bajo la hipótesis nula (Gibbons, 1971). El nivel de significancia considerado fue, tanto en esta prueba como en todas las siguientes, .05.

A continuación aparecen el valor de la estadística de Durbin-Watson para los residuales y el resultado de la prueba de autocorrelación efectuada con la estadística anterior. Un signo "+" o uno "-" en esta columna se refieren, respectivamente, a que se aceptó la hipótesis de autocorrelación de primer orden positiva o negativa; el símbolo "\#" indica que el valor de la estadística DW cayó en la región de indecisión de la prueba. La ausencia de signo alguno señala la aceptación de la aleatoriedad.

En las siguientes tres columnas aparecen la varianza de los residuales y las estadísticas $\sqrt{b_1}$ y b_2 de D'Agostino y Pearson

(1973) para probar normalidad. La notación "sk" o "ku" a la izquierda de la columna $\sqrt{b_1}$ se debe a que la hipótesis de normalidad se rechazó debido a un sesgo o una kurtosis excesivos; el símbolo "*" indica que ambos indicaron desviaciones de la normal. Es importante señalar que no se incluye en estas tablas la media de los residuales debido a que en todos los casos fue, en valor absoluto, menor que 10^{-8}.

La columna marcada como "sesgos" se refiere a cuáles componentes tuvieron un sesgo que excedía al 1% del estimador, señalándose entre paréntesis de cuánto es este porcentaje; la aparición del símbolo "*" indica que los sesgos excedieron al 10% del valor de los componentes correspondientes. El criterio del 1%, propuesto por Ratkowsky (1983:20-21), al parecer con bases puramente empíricas, es un indicador de conducta no lineal en la combinación de datos y modelo considerada. Como afirma Ratkowsky (1983:13), con modelos de regresión no lineal "es más adecuado hablar de combinaciones de modelos con conjuntos de observaciones que únicamente de modelos, puesto que un conjunto de datos observados en conjunción con un modelo determina la conducta de este último".

La última columna de la tabla 3.3 aparece marcada como "det(vcv)". Esto se refiere al determinante de la matriz de varianza covarianza de los residuales calculada a partir de las expresiones (2.8) y (2.9). Más adelante se regresará a este punto.

Finalmente, es necesario hacer notar que en estas tres primeras tablas no se reportan los resultados obtenidos con la estadística t presentada en Ratkowsky, (1983:62.8):

$$t = \hat{\theta} \,/\, var^{1/2}(\hat{\theta})$$

para contrastar si la estimación de cada componente estaba bien determinada: en todos los casos, los estimadores de los cinco componentes del parámetro llevaron a aceptar tal hipótesis. Esta estadística se usó en el sentido mencionado por este autor, es decir, comparándola contra una distribución t de Student con (n-p) grados de libertad, que son los asociados con la varianza residual.

Los aspectos numéricos del ajuste resultaron satisfactorios en la mayor parte de los casos: las normas del gradiente tuvieron una media (para ambos sexos) del orden de 10^{-5}; el promedio de iteraciones fue 7.6, que se compara satisfactoriamente con las 7 iteraciones obtenidas en promedio por Preece y Baines. Incluso, es posible pensar que la compración es favorable a las estimaciones aquí presentadas, pues estos autores no mencionan sus niveles de tolerancia ni las normas de los gradientes que obtuvieron.

Con el fin de disminuir cualquier perturbación proveniente de inexactitudes respecto a la edad de medición y al equiespaciamiento de las observaciones, se ajustaron los datos, utilizando interpolación lineal. Esto se hizo únicamente cuando se tenían mediciones que distaran en a lo más 6 meses y sin realizar más de dos de estos ajustes por individuo.

De esta forma se dispuso de 29 casos masculinos ajustados, con 16 observaciones anuales, entre los 4 y los 19 años y de 16 casos femeninos ajustados, con 23 observaciones semestrales entre los 5 y los 16 años. Cabe señalar que inicialmente se pensó trabajar con las edades entre 5 y 19 años para los datos femeninos; no se hizo así por consideraciones concernientes a los residuales, como se verá en la siguiente subsección.

Las tablas 3.4 y 3.5 muestran los ajustes del modelo PB1 para estos individuos; la notación es la usada en las tres primeras tablas de este capítulo.

Para cada uno de estos casos se estimaron 10 parámetros biológicos. La selección de estos últimos se basó en los trabajos de Preece y Baines (1978), Billewicz y McGregor (1982) y Zacharias y Rand (1983). Las edades de "despegue" de la velocidad de crecimiento en la pre adolescencia (abreviada TO, por *take off*) y de máxima velocidad durante la adolescencia (PV, por *peak velocity*), que corresponden a puntos estacionarios de la curva de velocidad, se estimaron encontrando, con un algoritmo de bisección (Rice, 1983:§8.2) modificado, las raíces de la función de aceleración. Puesto que esta función tiene otras raíces además de las que aparecen en la adolescencia, la modificación hecha al algoritmo de bisección consistió en una rutina que garantizara que las raíces que se encontraran fueran las que interesaba obtener. A partir de estas edades se generó el resto de los parámetros biológicos considerados; éstos fueron: las estaturas en TO (HTO, por *height at take off*) y en PV (HPV), las velocidades en TO (VTO) y en PV (VPV), los incrementos en estatura habidos entre TO y PV (INC TO-PV) y entre PV y la última edad disponible (INC PV-AD) y los porcentajes de la estatura final alcanzados en las edades TO y PV, abreviados %AD-TO y %AD-PV, respectivamente.

En las tablas 3.6 y 3.8 aparecen, separados por sexo, los 15 parámetros considerados para cada individuo. Las tablas 3.7 y 3.9 son las matrices de correlación de estos parámetros para cada sexo. Los valores críticos para el coeficiente de correlación al nivel de significancia .05 fueron .3494 para los hombres y .5197 para las mujeres.

Es posible extraer una gran cantidad de información referente a los procesos de maduración de esta clase de tablas. Los trabajos ya mencionados de Dillewickz y McGregor y de Zacharias y Rand son ejemplos de su uso para examinar problemas de biología humana. En el primero de ellos se analizan las diferencias estacionales (época de lluvias y época de secas) en la velocidad de crecimiento para el peso y la talla en varias comunidades rurales de Gambia. Por su parte, Zacharias y Rand estudian la relación entre crecimiento y edad de la menarquia a través del análisis de la aceleración de crecimiento durante la adolescencia.

Aun sin pretender realizar interpretaciones biológicas completas, una inspección de las matrices de correlación resalta ciertas relaciones interesantes. Para efectuar esa clase de interpretaciones podrían efectuarse análisis estadísticos más completos sobre esta matriz: por ejemplo, componentes principales o correlación canónica para identificar grupos de parámetros biológicos. De cualquier forma, debe tomarse en cuenta lo señalado por Preece y Baines (1978:15) en el sentido de que varios de estos parámetros biológicos son combinaciones lineales de otros de ellos, lo que llevará a multicolinealidades en la matriz de correlación.

A continuación se describen algunas de tales relaciones; primeramente se comentan algunas parejas de parámetros con correlaciones significativas coincidentes en ambos sexos.

h_1 y HTO (+): evidencia, simplemente que los individuos que llegan altos al momento del "despegue", tienen las mayores estaturas finales.

s_0 y [illegible] (+); s_0 y %AD PV (+); v_0 y v_1 (+): puesto que, como [illegible] en Preece y Baines (1978), los porcentajes de altura adulta que se tienen en las edades de "despegue" y de velocidad máxima pueden verse como un indicador (si bien un tanto grueso) del grado de maduración, es posible interpretar este resultado en el sentido de que ϕ_0 y ϕ_1 funcionan como reflejo de tal característica: los individuos de maduración más temprana tendrían valores altos en estos parámetros.

θ y TO (+); *θ* y VTO (-); *θ* y VPV (-): la primera resulta obvia a partir de la construcción del modelo; las otras dos resultan interesantes, pues muestran que los individuos que despegan más tardíamente lo hacen con menor velocidad y alcanzan menor velocidad máxima que aquéllos con despegues tempranos.

HTO y h_1 (+); HTO y %AD-TO (+); HTO y INC TO-PV (-): la primera ya fue mencionada; las dos últimas indican que mientras con más altura se llegue al "despegue", se acumula menos estatura entre esta edad y la edad de velocidad máxima.

VPV y TO (-); VPV e INC TO-PV (+): implican que mientras más tardíamente se empiece a acelerar el crecimiento en la adolescencia, menor velocidad máxima se alcanza, si bien se tiene una mayor acumulación de talla después del despegue que la que se observa en individuos con despegue temprano.

INC PV-AD y %AD-PV (-): refleja el hecho de que mientras se llegue con menor grado de maduración al momento de la máxima velocidad, se tendrá una ganancia de estatura mayor en las edades siguientes a tal momento; tal vez una especie de compensación.

Es interesante notar que PV y HPV no tuvieron coincidencia alguna en sus correlaciones para diferentes sexos.

Respecto a la matriz de correlación femenina, además de las correlaciones significativas mencionadas en los párrafos anteriores, se tuvieron, entre otras, positivas, entre HPV y h_1, HPV y HTO, HPV y %AD-TO, HPV y %AD-PV y PV y TO, y negativa entre PV y VTO. Por su parte, para la matriz de correlación masculina las correlaciones significativas para HPV fueron negativas: con PV y TO; PV se asoció, también negativamente, con INC PV-AD y positivamente con %AD-PV. Estas últimas correlaciones indican, simplemente que mientras más diferido se encuentre el incremento de velocidad en la adolescencia, menos significativo resultará éste, situación no encontrada en las mujeres.

En esta misma línea, la tabla 3.14 presenta tanto las estadísticas descriptivas más importantes como los resultados de los contrastes de igualdad de medias efectuados entre ambos sexos para los 15 parámetros biológicos. En los casos en que no fue posible aceptar previamente la hipótesis de igualdad de varianzas (utilizando una distribución F para el cociente de varianzas), se calculó la estadística presentada en, por ejemplo, Lindgren (1968) junto con sus grados de libertad aproximados. Excepto en IN TD-PV, e IN PV-AD se rechazó la hipótesis de igualdad de medias. Esto se indica con el símbolo "*" a la izquierda de la estadística.

Los ajustes realizados por este método no presentaron problemas numéricos serios. Siguiendo a Gallant y Goebel (1976), se calcularon estimadores MCG-CIRC utilizando como valores iniciales aquéllos obtenidos en la estimación MCO y a los residuales ahí calculados para estimar a la matriz Σ_{*}. Los estimadores así obtenidos se incorporaron como valores iniciales en una segunda estimación circular que, desde luego, utilizó los residuales generados a partir de estos primeros estimadores circulares para calcular la $\hat{\Sigma}_{*}$ correspondiente. Esto se hace con el fin de observar si los estimadores son cada vez más estables. En este caso, se repitió este procedimiento durante 10 pasos, siendo el paso 0 la estimación MCO. Las gráficas 3A y 3B muestran los promedios de los determinantes de las matrices $\hat{\Sigma}_{\bullet}$ obtenidas para cada individuo y del número de iteraciones necesario para obtener soluciones satisfactorias. Como se ve, es posible hablar de una situación invariante a partir del cuarto paso, si bien las magnitudes absolutas de los promedios de determinantes y de iteraciones no cambian significativamente.

Los valores obtenidos para el cuarto paso fueron considerados como los estimadores máximo verosímiles.

La tabla 3.10 es la tabla análoga a las tablas 3.6 y 3.8 para la estimación MCG-CIRC (o MV) mencionada en el capítulo anterior. Como se ve, las diferencias son muy pequeñas. La ventaja que tendrían los estimadores presentados en la tabla 3.10 es que, al ser máximo verosímiles los estimadores de $\underline{\theta}_i$, se garantiza, por el principio de invarianza (Fisher, 1973), que los 10 estimadores biológicos calculados a partir de ellos heredan esta propiedad.

La tabla 3.12 muestra las estadísticas descriptivas para los métodos MCO y MCG-CIRC (o MV). Como se aprecia inmediatamente, no hay diferencias significativas entre los estimadores calculados con ambos métodos.

3.2.2) ANALISIS DE LOS RESIDUALES

Como se mencionó en la subsección anterior, se trabajó con observaciones anuales entre los 4 y los 19 años (16 observaciones) para 29 series masculinas y con observaciones semestrales entre los 5 y los 17 años (23 observaciones) para 16 series femeninas.

Las gráficas 3C, 3D, 3E y 3F muestran el comportamiento de las medias y varianzas de los residuales contra la edad en cada sexo. En las dos últimas gráficas aparecen los resultados para datos femeninos obtenidos al ajustar el modelo con datos semestrales entre 5 y 19 años.

La tabla 3.11 muestra el número de residuales que fueron mayores que 1 cm, la media y la varianza de los residuales y los parámetros que tuvieron sesgos mayores que el 1% del estimador para las estimaciones MCO y MV. Como se ve, la comparación es ligeramente favorable al método de MCO.

El hecho notable de que para este método en todos los casos la suma de los residuales fue 0 puede interpretarse como indicador de un comportamiento casi lineal de los estimadores MCO. Esta interpretación se refuerza con los bajos valores relativos del sesgo de Box (hubo muy pocos individuos con sesgos mayores que el 10% del estimador) y con el hecho de que la matriz de varianza covarianza de los residuales fue, en todos los casos casi singular (la media de los determinantes fue del orden de 10^{-10}). Estas tres características se cumplen en regresión lineal (Weisberg, 1982:§5.1).

Este comportamiento casi lineal de los estimadores es una ventaja del modelo en dos sentidos: por un lado, da elementos para considerar que las propiedades asintóticas de los estimadores MCO (insesgamiento, varianza mínima, máxima verosimilitud en el caso normal y distribución normal asintótica) se satisfacen para muestras relativamente pequeñas (Ratkowsky, 1983:30); por otro, hace innecesario el uso de de ciertas clases de residuales generalizados (Cox y Snell, 1971; Cook y Tsiao, 1985) que resultan adecuados para modelos altamente no lineales.

A partir de las gráficas 3C y 3D se estableció que el comportamiento de los residuales podría considerarse como semejante a un proceso estocástico estacionario de segundo orden

(Hannan, 1983). Para los residuales femeninos, aunque también hubiera sido posible considerarlos como generados por un proceso estacionario de segundo orden, en vista del comportamiento de las varianzas para edades entre los 16 y los 19 años, se decidió trabajar sólo con edades entre los 5 y los 16 años, como ya se ha mencionado, repitiéndose los ajustes necesarios.

Las gráficas 3C y 3E muestran un comportamiento que, desde luego, se observó en la mayor parte de los ajustes individuales: el modelo sobreestima para edades entre los 6 y los 9 años, subestima entre los 10 a los 12 o 13, vuelve a sobreestimar entre los 12 y los 14 y no tiene un patrón de comportamiento claramente definido en las edades posteriores. Las gráficas 3K y 3L muestran las medias y varianzas de los residuales obtenidos con MCO y con MCG-CIRC contra la edad. Las tendencias descritas anteriormente aparecen con bastante claridad.

Las gráficas 3G, 3H, 3I, y 3J muestran, para cada sexo, las autocorrelaciones y las autocorrelaciones parciales junto con los puntos hasta los cuales se podrían considerar como iguales a cero, a un nivel de significancia de 5%. Estos estimadores así como los puntos para contrastar la hipótesis anterior fueron calcualdos con los métodos descritos en el capitulo 2.

Como se ve, los residuales masculinos no presentan autocorrelaciones significativas; para los femeninos, sólo la autocorrelación de primer orden lo fue. Esto puede ser consecuencia de la menor separación existente entre cada observación femenina respecto a las observaciones masculinas.

Otra forma de detectar comportamientos diferenciales del ajuste, sugerida por [illegible] y [illegible], consiste en agrupar los residuales no en términos de la edad, sino del porcentaje de estatura adulta alcanzado por el individuo. Los resultados aparecen en la tabla 3.13. Aparece en la primera columna el rango de porcentaje de estatura final considerado, el número de residuales que entraron en tal rango, su media, desviación estándar y error estándar así como el coeficiente de asimetría $\sqrt{b_1}$. Llaman la atención la presencia, detectada a través de las medias, de un ciclo de sobre y subestimación semejante al descrito anteriormente, el comportamiento de las desviaciones estándares, que no está relacionado con el porcentaje de estatura adulta, y los cambios en el coeficiente de asimetría que resultó ser no significativo en ningún caso. La idea central de esta tabla consiste en observar la existencia de patrones de subestimación o sobreestimación del modelo con respecto a un parámetro que, aunque en forma gruesa, da indicación acerca de la velocidad de maduración del individuo.

3.3) A MANERA DE CONCLUSION

respecto al comportamiento del modelo y de los dos métodos de estimación utilizados en estos análisis preliminares pueden establecerse las siguientes afirmaciones:

1) Los ajustes fueron de muy buena calidad tanto en términos numéricos como en términos estadísticos cuando el número de observaciones era mayor que 10.

2) Aunque sí hay presencia de autocorrelación en los residuales, ésta disminuye al espaciarse más las observaciones. Véanse los resultados obtenidos con la estadística Durbin-Watson y con la prueba de rachas en las tablas 3.1 a 3.6: son consistentes con el enunciado anterior.

3) Los valores obtenidos con los métodos MCO y MCO-CIRC son muy similares, siendo mucho menos costoso y complicado obtener estimadores con el primero. Por otra parte, los resultados obtenidos con las medidas de no linealidad y con los residuales son evidencia en favor del buen comportamiento del método de MCO.
4) El modelo presenta claramente un ciclo de sobreestimación - subestimación - sobreestimación respecto a la edad y también respecto al porcentaje de estatura adulta alcanzado.

TABLA 3.1

ARCHIVO ICC MASCULINO 47 CASOS

# SERIE	n	fin	x1	x2	x3	x4	x5	neg	pos	rachas	dw	var	b1	b2	sesgos
4	15	g	181.2448	169.7242	0.120542	1.525418	14.04203	7	8	9	1.99414	0.41050	0.00310	4.04038	no
15	16	g	165.0000	153.5603	0.115030	1.246345	13.38286	8	8	9	1.62661	0.47796	0.45945	3.18976	no
24	18	g	167.2344	153.5520	0.103511	1.346032	12.79451	8	10	11	2.28719	0.71161	-0.24955	2.57615	no
35	18	g	187.4340	174.9351	0.108307	1.069490	15.11641	9	9	10	2.14710	0.80536	-0.70085	4.29642	s1(2)
43	16	n	177.7242	166.4642	0.103096	1.401579	15.90440	9	7	6 +	1.04299	0.45352	-0.80284	4.90871	s1(1)
46	14	g	155.1736	146.1365	0.105688	1.769069	15.90898	6	8	8	2.52171	0.12758	0.04807	1.96165	no
47	14	g	195.8002	182.0900	0.104611	0.635694	15.45765	6	8	6	1.60800	0.81421	-0.25734	1.72725	todos
51	15	g	166.7841	153.4186	0.110679	1.184895	12.69953	7	8	7	1.68942	0.30435	0.12452	2.15964	no
58	16	g	177.8546	165.7696	0.098673	0.964853	15.51961	10	6	7	1.42723	1.14003	0.58351	2.49483	todos
59	16	g	174.0100	161.9950	0.106720	1.289070	13.92786	8	8	7	2.05735	1.81273	-0.30783	2.87006	s0(1),s1(3)
67	13	g	175.4137	163.1142	0.111758	1.131960	13.87250	7	6	6 *	1.30982	0.32772	0.01419	1.89593	s1(1)
69	11	n	171.6374	161.6372	0.117951	1.034864	14.34030	4	7	6 *	2.20074	0.79172	-0.53120	2.16456	todos
78	15	g	191.5126	178.7750	0.110268	1.033829	14.94393	7	8	6	1.46577	1.15687	0.40929	3.23847	s0(2),s1(5)
90	13	g	181.7264	168.9410	0.099562	0.814808	15.29527	6	7	8 *	2.01836	0.38216	-0.38533	2.32193	todos*
91	14	n	176.9365	165.8292	0.117656	1.079904	16.34060	8	6	6 *	2.35520	0.51860	-0.20192	3.03294	s1(2)
97	17	g	178.0856	162.8379	0.086928	1.023152	13.12734	8	9	8	1.60559	0.44936	-0.44048	2.15200	no
100	19	g	181.0322	167.6665	0.101385	1.364876	14.24741	9	10	9	2.26460	0.22566	-0.49355	2.83357	no
104	17	g	185.2714	170.3415	0.107206	1.148031	13.16195	8	9	11	1.66481	-0.64013	-0.51420	2.50130	no
105	16	g	176.9103	164.2645	0.095946	1.103954	16.04331	8	8	6	1.45972	0.70985	0.15793	1.83603	s0(1),s1(2)
114	17	g	155.4364	142.6940	0.088781	1.123276	13.87267	7	10	7	1.84211	0.24403	ku-0.17822	1.63313	no
116	17	n	172.8643	158.3799	0.093500	1.104479	14.49223	8	9	7 *	1.08412	0.83766	ku 0.12328	1.71381	s1(1)
117	17	g	196.5067	173.8150	0.103309	1.043801	15.12074	7	10	6 +	0.92519	0.57090	-0.57439	2.39005	s1(1)
118	15	g	171.2273	160.0639	0.110328	1.307120	13.94240	7	8	9 +	2.92351	0.24283	0.25769	2.12929	no
124	15	g	173.2673	160.9722	0.103901	1.459654	14.64764	5	10	7	1.90028	0.16141	sk-0.85294	3.74706	no
129	17	g	168.8600	156.9560	0.103810	1.457723	14.20692	9	8	9 *	1.16626	0.34952	0.30474	2.61000	no
137	15	g	181.1016	170.0123	0.113951	1.186541	14.71463	6	9	10	2.18264	0.49889	-0.39388	2.15729	s0(1),s1(2)
144	20	n	175.2127	162.3042	0.090662	1.208467	15.12045	8	12	13 -	2.89479	0.32226	-0.25267	2.56759	no
149	14	g	170.4612	163.3301	0.103444	1.160885	13.00660	8	6	6 *	1.58339	0.71674	0.20136	2.20931	s1(2)
153	18	g	182.9412	171.1713	0.099567	0.985270	16.31051	9	9	- 6 +	1.14453	0.77605	-0.14897	2.10420	s1(2)
155	15	n	180.2089	168.9675	0.117521	1.142204	12.64473	10	5	8	2.18171	0.18538	0.64016	1.89747	no
157	15	g	185.9720	173.2371	0.113455	1.387247	13.54485	8	7	8	2.45095	0.63045	-0.36737	2.89765	s1(1)
162	11	n	168.6645	155.9598	0.109457	0.996303	13.40002	5	6	6 *	2.24381	1.97208	-0.42939	2.11141	todos
170	18	g	192.4372	177.0964	0.090940	0.106453	15.38454	8	10	- 6 +	1.06508	0.72993	-0.17140	2.92617	s1(1)
179	17	g	188.4821	174.9755	0.107117	1.174851	14.41440	7	10	7 *	1.25427	0.05684	-0.10951	2.63517	s1(1)
185	16	g	161.7970	149.7905	0.090158	1.303620	14.32265	8	8	+ 12 -	3.10869	0.23101	0.10349	2.13783	no
191	17	n	169.6557	155.0064	0.100030	1.269835	13.27644	9	8	7	1.73707	1.06006	0.37931	2.19406	no
192	15	g	172.6083	158.3956	0.101094	1.143078	13.03769	9	6	8	1.39000	0.73255	0.25469	2.02943	s1(1)
197	17	g	179.9203	167.9388	0.106055	0.711827	14.59872	9	8	6 +	1.12559	1.72448	0.25063	2.25745	todos*
202	15	n	184.6370	170.2227	0.104460	1.080264	13.90170	9	6	8	2.05693	0.44549	0.30179	2.29947	s1(1)
208	18	g	177.8066	162.9064	0.107040	1.243030	12.16024	10	8	+ 13	2.44917	0.38099	0.21344	3.12622	no
210	20	g	175.2589	162.4484	0.098976	1.191467	14.36033	10	10	9	1.54034	0.24923	-0.59923	2.97421	no
211	17	n	177.5665	161.7963	0.087150	0.948203	14.27870	7	10	6	1.71178	0.52626	-0.50932	2.08649	s1(1)
214	18	g	164.1032	151.0891	0.107263	1.301571	13.27712	8	10	7	1.69098	0.75704	sk-0.91813	3.87224	no
217	18	g	180.6160	166.9505	0.092601	0.971690	15.32750	7	11	10	1.77532	0.37306	-0.30039	2.15012	s1(1)
218	18	n	182.4300	167.7230	0.100314	1.140194	13.60415	9	9	8	1.43595	0.83504	-0.40491	2.46621	no
219	17	g	177.3363	164.5347	0.109594	1.074102	13.61811	9	8	6	1.35716	0.34020	-0.33926	2.03010	no
222	17	[illegible]	159.5008	157.0704	0.106595	1.220347	13.94482	8	9	10	1.56804	0.49066	0.53740	3.20158	no
	16.1		176.6687	163.7907	0.104866	1.141041	14.25030	7.80	8.29	7.106					
	1.94		8.692546	8.302977	0.009029	0.250548	0.999276	1.34	1.52	2.707					

TABLA 3.2

ARCHIVO NCH MASCULINO   23 CASOS

# SERIE	n	fin	m1	m2	m3	m4	m5	neg	pos	rachas	z		dw	var	b1	b2	sesgos
128	35	g	171.3111	158.6533	0.102863	1.170820	14.62451	10	17	9	-3.26	*	0.68350	0.32268	0.11975	2.26232	no
129	34	g	182.3009	169.3382	0.100956	1.200376	14.39173	14	20	11	-2.33	*	1.00735	0.37094	0.01530	2.67423	no
222	30	g	168.0030	157.9606	0.111529	1.314470	14.55065	15	15	9	-2.60	*	0.93904	0.26372	0.48731	3.31604	no
259	26	q	173.3925	161.7778	0.112803	0.917827	14.20204	13	13	8	-2.40	*	0.89428	0.59048	-0.03455	1.96273	#0,#1
251	33	n	175.1712	161.9381	0.105142	1.290624	14.18724	17	16	7	-3.71	*	0.48829	0.64410	0.79333	3.78038	#1
305	32	n	173.8562	161.4142	0.107931	1.209706	14.01603	18	14	12	-1.74	*	0.70399	0.27497	0.85215	2.97148	no
334	32	g	183.1044	170.0907	0.111701	1.096078	14.64367	17	15	13	-1.42	*	1.05336	0.29677	0.23194	1.96703	no
336	34	g	163.9767	152.8949	0.111532	1.257314	14.90821	17	17	13	-1.74	*	1.26143	0.21865	-0.18113	2.41929	no
340	34	n	170.0694	160.1471	0.108377	1.022414	16.28620	17	17	10	-2.79	*	1.25580	0.32927	0.12989	2.58341	#0,#1
402	20	g	195.3409	179.1161	0.078156	0.764075	14.26386	9	11	9	-0.88	*	1.12461	0.54882	0.32203	2.36210	#0,#1
422	34	g	167.2795	156.6292	0.124460	1.484261	13.65303	13	21	10	-2.61	*	1.09774	0.29926	-0.31429	2.34777	no
426	33	g	174.3735	163.4171	0.110164	0.817682	14.46745	16	17	8	-3.36	*	1.02563	0.24854	-0.23557	2.21275	#1
501	34	g	179.0412	166.1991	0.105071	1.349150	14.62905	17	17	9	-3.13	*	0.77135	0.23423	-0.11033	1.80180	no
514	34	q	173.8787	162.6707	0.106740	1.137793	15.39299	19	15	10	-2.74	*	0.93167	0.57836	0.40186	2.40394	#1
525	30	g	179.4677	167.5138	0.119067	1.225554	13.01906	17	13	10	-2.17	*	0.85161	0.51941	0.14365	2.23843	no
534	35	g	180.9106	169.0452	0.112975	1.426025	14.60967	17	18	6	-4.29	*	0.65209	0.34763	-0.16750	1.99204	no
631	34	g	180.1300	164.6668	0.093436	0.839221	14.12999	16	18	8	-3.48	*	0.50283	0.92024	-0.86701	2.95378	#0,#1
634	27	g	173.0719	161.2962	0.112310	1.470723	14.13682	15	12	13	-0.53		2.11711	0.16722	-0.19305	2.27905	no
639	41	n	177.1133	165.2345	0.106345	1.345299	14.66246	21	20	14	-2.37	*	1.11529	0.32530	-0.10253	2.84370	no
642	33	g	178.7384	164.5187	0.117613	1.184191	12.95663	18	15	10	-2.63	*	0.75035	0.36898	0.42369	2.48750	no
647	40	g	181.1164	168.8216	0.102975	1.129857	14.72219	21	19	10	-3.52	*	0.76772	0.42723	-0.16079	2.14149	no
649	37	g	188.9631	176.5705	0.111194	1.270047	14.30992	19	18	10	-3.17	*	0.70762	0.24744	-0.56747	3.45688	no
649	33	g	175.8760	163.7053	0.099212	1.263276	15.54592	10	15	12	-1.91	*	1.01505	0.25601	0.69911	3.81910	#1
	32.8		177.1771	164.8528	0.108719	1.185859	14.44426	16.6	16.2	10.04	-2.95						
	4.25		6.601800	5.856205	0.006768	0.196071	0.703769	2.59	2.55	2.010	0.890						

TABLA 3.3

ARCHIVO MCM FEMENINO 24 CASOS

# SERIE	n	fin	m1	m2	m3	m4	m5	neg	pos	rachas	z	dw	var	b1	b2	sesgos	det (ucv)
113	36	g	159.0001	147.2142	0.113355	1.023903	12.91299	17	19	12	-2.36 •	0.67705	0.45839	-0.66843	3.30720	no	1.3E-2
114	31	q	159.2198	147.1697	0.103608	1.109675	13.60711	16	15	12	-1.64 •	0.53479	0.28600	-0.01651	3.18401	no	1.7E-2
120	40	м	163.2906	150.7967	0.108640	1.073729	12.92042	19	21	12	-2.87 •	0.99914	0.20441	-0.37714	2.47024	no	1.1E-2
121	29	g	161.2502	152.1006	0.160054	1.215202	11.06010	17	12	17	0.73	2.33195	0.08064	0.33509	2.74057	no	9.0E-3
209	31	g	163.6211	153.7422	0.130035	1.360759	11.49642	12	19	14	-0.66 ●	1.44834	0.15400	-0.03031	3.09308	no	2.0E-2
321	19	g	156.0980	145.9935	0.117839	0.775146	12.75269	11	8	7	-1.58 •	0.99751	0.14435	-0.42869	2.02096	s0,s1,th	5.2E-3
323	24	g	167.1610	156.3332	0.131768	1.095126	11.98454	13	11	10	-1.23 •	1.06320	0.32783	ku 0.26247	1.60801	s0,s1,th	1.3E-3
325	36	g	161.9500	151.5639	0.127887	1.106439	12.32299	16	20	18	-1.29	1.61829	0.13929	-0.27839	2.08741	no	8.3E-2
326	36	g	163.9737	153.4510	0.124001	1.231447	12.60161	18	18	14	-1.69 •	0.80238	0.23853	-0.36288	2.06117	no	1.7E-2
327	39	g	167.9000	156.3334	0.115032	0.991385	13.89442	19	20	12	-2.76 •	0.80366	0.30567	-0.02901	2.57992	no	4.2E-3
329	37	м	172.8319	160.2081	0.113403	0.917540	12.82091	18	19	15	-1.50	1.61183	0.15073	0.06702	2.40986	s0,s1	9.5E-4
409	30	м	160.7570	157.7574	0.126902	0.951774	12.46400	14	16	14	-0.72 •	1.32715	0.37517	-0.65763	4.15442	s1	1.9E-2
428	39	м	162.1422	151.8442	0.119138	0.906111	13.47348	19	20	10	-3.41 •	0.69180	0.36155	0.22144	2.15617	s1	5.4E-3
501	33	g	163.2759	151.9815	0.129783	1.207234	11.37663	15	18	11	-2.27 •	1.37468	0.18543	0.37062	3.13971	no	2.7E-2
502	23	м	161.0553	152.2096	0.134629	0.914004	12.88056	10	13	8	-1.87 •	1.00760	0.22283	-0.20009	2.26347	todos	1.1E-3
503	33	g	162.7929	151.8412	0.141376	1.239138	11.26217	16	17	6	-4.07 •	0.69139	0.25078	-0.40428	2.33829	no	2.7E-2
514	18	м	156.2717	144.5587	0.123167	0.981957	12.40316	10	8	8	-0.93	2.12490	0.15969	0.47297	3.76346	s0,s1	4.7E-3
518	23	g	153.8651	142.3842	0.124348	1.012268	13.03222	16	12	10	-1.06	1.78793	0.11957	0.29671	2.06989	no	1.9E-2
537	27	g	160.5364	128.9786	0.137073	1.017674	12.97953	15	17	15	-1.50 •	1.35091	0.06906	0.02603	2.38334	no	8.9E-3
603	21	g	166.1878	149.1673	0.068769	0.503773	12.46406	10	11	12	0.24	1.88287	0.24325	0.14376	2.48869	todos*	5.4E-4
618	37	м	167.0427	156.0589	0.125795	1.183382	11.70792	15	22	16	-0.98	1.97965	0.24427	-0.48587	2.08307	no	1.7E-2
624	24	g	161.9049	152.8186	0.162388	1.563460	11.07771	11	13	9	-1.65 ●	1.42705	0.18244	-0.17499	4.06086	s1	1.1E-2
640	36	g	166.6507	156.9533	0.129623	1.001927	14.02327	19	17	12	-2.36 •	0.71053	0.40003	0.31723	2.46842	s0,s1,th	4.0E-3
643	24	м	156.9994	147.1621	0.140197	1.640143	10.96181	11	11	10	-0.87 ●	1.28960	0.13753	-0.04794	2.07819	s0,s1	2.7E-2
	30.8		163.1304	151.9966	0.125979	1.091312	12.4279	15	15.7	11.70	-1.62						
	6.68		4.610608	4.562750	0.019035	0.230710	0.883124	3.16	4.12	2.064	1.044						

TABLA 3.4

MASCULINO; 29 CASOS AJUSTADOS n = 16
ANUAL: 4 - 19

# SERIE	fin	w1	w2	w3	w4	w5	neg	pos	rachas	dw	var	b1	b2	sesgos
24	q	167.3199	153.8501	0.104480	1.356892	12.90044	6	10	7	2.29413	0.54801	-0.29544	2.25697	no
25	y	187.9137	174.8472	0.104637	0.925295	15.11912	8	8	6 •	1.05094	0.30560	-0.04401	2.01016	no
43	g	177.8583	166.2597	0.102889	1.324100	15.86737	9	7	6 *	1.09906	0.42652	-0.92658	3.52744	no
51	g	166.9086	153.8068	0.108906	1.178019	12.81666	7	9	7	1.45944	0.17449	-0.07410	2.12307	no
59	q	173.6047	161.8431	0.110653	1.378562	13.85962	8	8	7	2.00690	0.08673	0.17828	2.25853	no
97	g	170.8455	162.3685	0.096262	1.001381	13.09206	9	7	8	1.02234	0.24554	-0.22860	1.79389	no
104	g	185.2331	170.8006	0.110425	1.209033	13.29492	7	9	7	1.90238	0.46938	-0.22132	2.16720	no
114	w	158.4905	142.7203	0.088786	0.110760	13.92190	8	8	7	1.75914	0.48926 ku	0.17802	1.71808	no
116	g	172.3458	158.7347	0.097863	1.157305	14.57867	8	8	7 *	1.24529	0.59079	0.18728	1.79322	no
117	g	186.1673	173.7069	0.106274	1.150174	15.00062	10	6	7	1.39541	0.45793	0.20175	2.67351	no
124	g	173.6512	161.3289	0.102572	1.243234	14.80716	9	7	8	1.51990	0.12929	-0.44056	3.53091	no
129	g	168.7435	156.0542	0.105404	1.411080	14.14177	7	9	7	2.40841	0.29763	0.17972	2.66082	no
144	w	175.1206	162.4508	0.097920	1.194324	15.12220	8	8	11	2.31190	0.32267	0.13990	1.79372	no
153	g	182.9130	171.7562	0.104236	1.080622	16.39639	8	8	6	1.88502	0.58161	0.22377	2.33493	s0,s1
163	w	167.7492	157.8062	0.174258	1.226459	12.99366	7	9	7	2.14563	0.29241	-0.25809	2.26843	no
170	g	191.7978	177.0799	0.095375	1.160405	15.36881	8	8	9	2.07619	0.19343	0.19157	2.80273	no
191	g	169.5267	155.7976	0.101360	1.202288	13.22363	9	7	6 *	1.35467	0.74052	0.38286	1.82187	no
209	w	177.8521	162.7074	0.106579	1.107965	12.14334	7	9	7	2.34052	0.22086	0.01992	3.67984	no
211	g	177.6523	162.2524	0.091175	0.959292	14.49035	7	9	6 •	1.03005	0.55045	0.07142	1.90632	no
214	g	163.8791	152.2967	0.114833	1.421171	13.29702	9	7	11	2.79556	0.47920	-0.58668	2.75953	no
217	w	180.8966	167.2101	0.094588	0.956106	15.42159	8	8	6	1.86700	0.31371	-0.42198	2.48979	s1(1)
218	w	182.8073	168.5828	0.106233	1.200054	13.78817	8	8	11	2.20837	0.27572	0.25892	2.15615	no
219	g	177.1905	164.9749	0.113467	1.136405	13.72400	8	8	7	1.42363	0.31496	0.10478	1.94294	no
222	g	169.6030	157.6581	0.110031	1.288575	14.07565	6	10	9	1.93069	0.12246	-1.02078	3.73936	no
128	g	171.1064	158.5897	0.103476	1.197205	14.51793	9	7	6	1.47035	0.62793	-0.04464	2.35066	s1
631	w	179.8354	164.4442	0.094027	0.832101	14.11680	8	8	6 *	1.76445	0.97653	-1.24780	4.08309	s0,s1
639	q	177.2398	165.3229	0.101810	1.254560	14.70709	9	7	7	2.11966	0.38413	0.45959	2.19338	no
647	g	180.7394	168.9901	0.107044	1.235025	14.73839	8	0	7	1.15400	0.22941	-0.04963	2.23574	no
649	w	188.8353	176.5600	0.112446	1.303293	14.30709	9	7	7	1.36997	0.23148	0.25025	2.54629	no
		176.1449	163.1431	0.104264	1.150530	14.20176	8	8	7.344					
		8.067575	7.811269	0.009126	0.242553	0.900409	0.94	0.94	1.460					

TABLA 3.5

FEMENINO; 16 CASOS AJUSTADOS n = 23
SEMESTRAL 5 - 16

# SERIE	fin	m1	m2	m3	m4	m5	neg	pos	rachas	z		dm	var	b1	b2	sesgos
113	g	159.157	147.842	0.1191	1.0162	13.0146	11	12	0	-1.49	*	0.92291	0.59824	0.01570	2.72597	s0,s1
114	g	161.326	148.118	0.0944	0.9373	13.0447	13	10	10	-0.65	#	1.27432	0.33459	-0.40069	2.98774	s0,s1,th
120	g	162.888	151.006	0.1202	1.1944	13.0166	12	11	13	0.64		2.21436	0.11967	0.10669	2.10618	no
121	g	161.144	152.281	0.1637	1.2471	11.0939	10	13	15	0.65		2.62013	0.05736	am-1.42152	6.14411	no
209	m	165.038	153.141	0.1199	1.2384	11.4151	11	12	9	-1.07	#	1.40648	0.12561	-0.35636	2.38472	no
325	g	162.114	151.540	0.1267	1.1374	12.3062	11	12	11	-0.21		1.96399	0.11024	sk-0.73508	2.83417	no
326	g	163.494	153.233	0.1274	1.2854	12.5399	10	13	9	-1.09		1.81022	0.23399	-0.40477	3.56479	s1
327	g	160.325	157.301	0.1209	1.0402	14.0452	12	11	12	0.21		1.45241	0.12739	0.43314	2.84397	s0,s1
409	g	168.677	150.679	0.1384	1.0373	12.6265	10	13	12	0.22		1.92676	0.25430	sk-1.20515	5.06447	s0,s1
428	m	162.648	152.596	0.1220	0.9769	13.6377	10	13	6	-2.39	*	1.06330	0.33150	ku-0.16418	1.79590	s0,s1,th
501	g	163.116	152.149	0.1329	1.2680	11.4295	11	12	10	-0.64		1.87938	0.20608	0.75406	4.35154	no
503	g	162.330	151.478	0.1464	1.2944	11.2501	10	13	7	-1.96	*	1.02472	0.22224	-0.39318	1.98821	no
518	m	154.320	142.717	0.1198	0.9413	13.8051	12	11	9	-1.07		2.30271	0.09856	0.63714	5.11364	s1
539	g	168.666	150.600	0.1311	0.9391	12.8524	14	9	10	-0.22		1.58254	0.06747	0.54822	2.64284	s1
616	g	167.713	156.825	0.1374	1.2560	11.8187	10	13	15	1.52	#	2.66720	0.13933	0.04533	2.19699	no
640	g	169.436	150.363	0.1153	0.7867	14.3491	10	13	12	0.22	*	1.17199	0.30467	-0.05393	2.17408	todos*
media		163.826	152.098	0.1272	1.0990	12.6903										
d.s.		3.956	4.255	0.0148	0.1544	1.0329										

TABLA 3.6

MASCULINO; DATOS AJUSTADOS; PARAMETROS BIOLOGICOS

# SERIE	h1	h0	s0	s1	theta	to	pv	hto	hpv	vto	vpv	incto-pv	incpv-ad	%ad-to	%ad-pv
24	167.3199	153.0501	0.104480	1.358892	12.90044	9.106	12.612	127.615	164.236	4.572	7.104	36.622	3.084	76.27	98.16
25	187.9137	174.8472	0.104637	0.925295	15.11912	10.566	12.802	146.610	159.431	5.086	6.195	12.621	28.402	78.13	84.04
43	177.5553	166.2597	0.102009	1.324100	15.86737	8.915	12.620	143.655	163.918	3.960	7.510	20.630	13.673	80.91	92.32
51	166.9086	153.5065	0.100906	1.170019	12.81666	8.765	13.604	125.865	159.996	5.055	6.049	34.131	6.913	75.41	95.86
59	173.6047	161.8431	0.110653	1.378562	12.05962	10.190	13.431	130.632	158.005	4.289	8.995	19.374	15.600	79.85	91.01
97	178.8455	167.3665	0.086262	1.001381	13.09266	8.218	13.450	129.243	165.531	4.798	8.252	36.287	13.319	72.27	92.56
104	185.2331	170.8006	0.110425	1.209833	13.29492	9.360	13.796	141.242	175.247	5.409	8.067	34.005	9.986	76.25	94.61
114	155.4905	142.7283	0.088786	1.107683	13.92190	9.352	13.379	117.553	130.455	3.732	7.834	20.902	17.035	75.50	89.04
116	172.3458	158.7347	0.097563	1.157305	14.37567	10.301	12.904	131.478	144.958	4.449	6.890	13.480	27.308	76.29	84.11
117	186.1673	173.7069	0.106274	1.150174	15.08062	10.967	12.662	148.103	156.330	4.580	5.479	8.226	29.830	79.55	83.97
124	173.6612	161.3239	0.102572	1.243284	14.00716	10.708	12.669	136.795	145.414	4.205	5.481	8.619	20.237	78.78	83.74
129	168.7435	156.0542	0.105484	1.411080	14.14177	10.444	13.067	133.898	147.366	4.037	7.762	13.469	21.378	79.35	87.33
144	175.1206	162.4505	0.099928	1.194324	15.12228	10.965	12.640	137.134	144.659	4.231	5.110	7.523	30.463	78.31	82.60
153	182.9138	171.7562	0.104236	1.080632	16.39639	11.910	15.821	148.102	169.063	4.086	6.913	20.691	13.847	80.91	92.41
163	167.7452	157.6662	0.134258	1.226459	12.99366	9.452	13.636	136.031	161.630	4.967	5.290	25.599	6.119	81.09	96.35
170	191.7978	177.0799	0.095375	1.160408	15.36081	11.056	12.725	147.827	156.060	4.663	5.558	8.233	35.738	77.07	81.37
191	169.8267	155.7976	0.101360	1.202780	13.22363	9.751	14.061	128.829	162.608	4.592	6.500	33.859	6.837	75.99	95.97
208	177.8571	162.7074	0.106579	1.187965	12.14334	8.151	11.760	131.310	158.749	5.510	10.201	27.349	19.031	73.81	89.26
211	177.6523	162.2524	0.091175	0.959292	14.43035	9.570	13.116	130.374	151.335	4.906	7.930	20.961	26.317	73.39	85.19
214	163.8791	152.2967	0.114033	1.421171	13.29702	9.748	14.191	129.393	158.919	4.373	5.511	29.526	4.960	78.96	96.97
217	180.5964	167.2101	0.094508	0.956106	15.42159	10.649	13.033	139.224	150.040	4.498	5.719	11.616	29.757	77.09	83.52
218	182.5073	168.5828	0.106253	1.200854	13.78817	7.762	13.441	140.310	165.352	5.039	9.451	25.042	17.156	76.88	90.60
219	177.1906	164.9747	0.113467	1.136485	13.72400	9.729	13.277	139.381	162.251	4.929	7.994	22.870	14.940	78.66	91.57
222	169.8930	157.6501	0.110031	1.208575	14.07565	10.260	13.170	135.246	149.832	4.497	8.219	16.586	19.971	78.47	88.24
128	171.1864	158.5097	0.103474	1.197795	14.51793	10.434	12.875	133.101	145.572	4.405	6.635	12.391	25.614	77.88	85.04
631	179.0354	164.4442	0.094827	0.832101	14.11688	9.051	13.031	131.422	156.300	5.302	7.592	24.878	23.536	73.08	86.91
639	177.2339	165.3229	0.104010	1.254560	14.70709	10.748	12.698	141.522	150.471	4.172	5.582	8.948	26.763	79.85	84.90
647	180.7394	168.9501	0.107044	1.235025	14.75839	10.799	12.600	145.101	153.908	4.312	5.607	8.872	26.460	80.28	85.19
649	188.0353	176.5680	0.112446	1.303273	14.30789	10.553	12.923	151.837	164.705	4.659	7.320	12.949	24.050	80.41	87.26
media	176.1449	163.1431	0.104264	1.104906	14.20176	9.968310	13.18434	136.7200	156.5990	4.601827	6.994137	19.87120	20.56855	77.61068	88.99655
ds	8.047575	7.811259	0.009126	0.142947	0.900409	0.891826	0.705622	7.884546	8.418073	0.446352	1.328273	9.290028	11.42620	2.426772	4.816777
cv	0.045800	0.047879	0.087529	0.120640	0.069034	0.089465	0.053519	0.057665	0.053755	0.096994	0.189912	0.467512	0.555556	0.031268	0.054123
max	191.7978	177.0799	0.134258	1.421171	16.39639	11.91	15.821	151.837	175.247	5.51	10.201	36.622	60.463	81.09	98.16
min	155.4905	142.7283	0.086261	0.832100	12.14334	8.151	11.76	117.553	135.455	3.732	5.11	7.523	3.084	72.27	81.37
b1	-0.70534	-0.64193	0.067062	-0.21604	0.074272	-0.15887	1.351955	-0.24147	-0.31271	0.212765	0.513784	1.053427	-0.91894	-0.78983	0.727087
b2	2.908201	2.996252	5.262331	2.935095	2.530734	2.445929	7.644075	2.886471	2.482626	2.481654	2.429784	1.881228	0.767402	2.380503	1.876218

TABLA 3.7

MASCULINO; MATRIZ DE CORRELACION

	h1	h0	s0	s1	theta	to	pv	hto	hpv	vto	vpv	incto-pv	incpv-sd	%sd-to	%sd-pv
h1	1.000	0.994	-0.082	-0.361	0.442	0.347	-0.129	0.839	-0.042	0.378	-0.008	-0.327	0.381	0.003	-0.426
h0	0.984	1.000	0.038	-0.272	0.512	0.446	-0.092	0.921	-0.096	0.285	-0.080	-0.402	0.377	0.257	-0.417
s0	-0.082	0.039	1.000	0.820	-0.282	0.061	0.154	0.269	0.101	0.101	-0.129	0.087	-0.358	0.624	0.424
s1	-0.361	-0.272	0.820	1.000	-0.231	0.044	-0.022	0.003	0.158	-0.400	-0.001	0.044	-0.285	0.533	0.359
theta	0.442	0.812	-0.282	-0.231	1.000	0.764	0.066	0.610	-0.509	-0.815	-0.405	-0.731	0.871	0.466	-0.815
to	0.347	0.446	0.061	0.044	0.764	1.000	0.151	0.623	-0.513	-0.472	-0.515	-0.796	0.561	0.631	-0.584
pv	-0.129	-0.092	0.154	-0.022	0.066	0.151	1.000	-0.019	-0.618	-0.106	-0.033	0.381	-0.460	0.149	0.811
hto	0.839	0.921	0.269	0.003	0.610	0.623	-0.019	1.000	-0.102	0.014	-0.224	-0.830	0.337	0.612	-0.361
hpv	-0.042	-0.096	0.101	0.158	-0.809	-0.513	-0.618	-0.182	1.000	0.374	0.087	0.158	-0.047	-0.260	0.044
vto	0.378	0.285	0.101	-0.400	-0.815	-0.472	-0.106	0.014	0.374	1.000	0.317	0.447	-0.183	-0.814	0.108
vpv	-0.008	-0.080	-0.129	-0.001	-0.405	-0.815	-0.033	-0.224	0.087	0.317	1.000	0.425	-0.324	-0.400	0.242
incto-pv	-0.327	-0.402	0.087	0.044	-0.731	-0.796	0.381	-0.830	0.158	0.447	0.425	1.000	-0.798	-0.493	0.884
incpv-sd	0.381	0.377	-0.358	-0.285	0.571	0.561	-0.460	0.337	-0.047	-0.183	-0.324	-0.798	1.000	0.064	-0.809
%sd-to	0.003	0.257	0.624	0.533	0.466	0.631	0.149	0.612	-0.260	-0.514	-0.400	-0.493	0.064	1.000	-0.041
%sd-pv	-0.426	-0.417	0.424	0.359	-0.615	-0.504	0.611	-0.361	0.044	0.108	0.242	0.884	-0.889	-0.041	1.000

TABLA 3.8

FEMENINO; DATOS AJUSTADOS; PARAMETROS BIOLOGICOS

# SERIE	h1	h0	se	s1	theta	to	pv	hto	hpv	vto	vpv	incto-pv	incpv-ad	%adto	%adpv
113	159.157	147.842	0.1191	1.0162	13.0146	8.968	12.282	123.457	142.892	5.081	6.918	19.436	16.246	77.57	89.78
114	161.326	148.118	0.0944	0.9373	13.0447	9.025	13.219	120.387	143.699	4.450	7.710	23.312	17.627	74.62	89.07
120	162.888	151.506	0.1202	1.1944	13.0166	9.232	12.526	127.615	147.705	4.880	7.879	20.089	15.183	78.35	90.60
121	161.144	152.281	0.1637	1.2471	11.0939	6.073	10.366	133.144	147.414	5.690	6.895	14.270	13.730	82.62	91.48
209	165.038	153.141	0.1199	1.2384	11.4151	7.687	10.968	126.733	149.161	5.355	9.042	22.428	16.677	76.42	89.94
325	162.114	151.548	0.1257	1.1374	12.3062	8.548	11.711	128.928	147.388	4.906	7.124	18.460	14.727	79.53	90.92
326	163.494	153.235	0.1274	1.2554	12.8399	8.993	12.094	131.781	149.897	4.633	7.618	18.116	13.597	80.60	91.68
327	168.325	157.301	0.1209	1.0402	14.0452	10.067	13.340	133.866	152.560	5.005	6.882	18.994	15.765	79.35	90.63
409	168.697	158.679	0.1384	1.0373	12.6265	9.048	11.723	137.086	152.973	5.470	6.520	15.887	15.724	81.26	90.68
420	162.648	152.896	0.1220	0.9768	13.6377	9.620	12.788	130.854	147.619	4.728	6.022	16.765	15.028	80.45	90.76
801	163.116	152.149	0.1329	1.2600	11.4275	7.947	10.938	128.927	148.226	5.275	8.132	19.300	14.890	79.04	90.87
803	162.330	151.478	0.1464	1.2944	11.2881	7.998	10.708	128.193	147.036	5.912	8.369	18.843	15.293	78.97	90.80
518	154.320	142.717	0.1198	0.9413	13.8051	8.978	12.173	117.623	136.746	5.395	6.740	19.123	17.573	76.22	88.61
539	168.666	158.000	0.1211	0.9391	12.8524	9.069	11.762	137.095	152.243	5.283	6.039	15.147	16.423	81.28	90.26
618	167.713	156.825	0.1374	1.2568	11.8187	8.393	11.289	133.584	152.633	5.498	8.003	19.049	15.080	79.65	91.01
640	169.426	158.363	0.1153	0.7867	14.3491	10.074	12.903	135.322	150.524	5.147	5.641	15.201	18.822	79.91	88.89
media	163.826	152.898	0.1272	1.0996	12.6903	8.857	11.924	129.643	148.045	5.169	7.226	18.401	15.774	79.12	90.37
d.s.	3.956	4.255	0.0148	0.1544	1.0329	0.693	0.876	5.437	4.104	0.381	0.916	2.418	1.366	2.02	0.86
c.v.	0.024	0.028	0.1160	0.1404	0.0814	0.078	0.073	0.042	0.028	0.074	0.127	0.131	0.087	0.03	0.01
max	169.436	4.031	0.0146	0.1529	1.0295	10.074	13.340	137.095	152.973	5.912	9.042	23.312	18.822	82.62	91.68
min	154.320	142.717	0.0944	0.7867	11.0939	7.687	10.366	117.623	136.746	4.450	5.641	14.270	13.597	74.62	88.61
b1	-0.88440	-1.02904	0.046615	-0.11251	-0.05525	0.075702	-0.07184	-1.45381	-2.30404	-0.01782	0.103117	0.228682	0.569268	-0.74333	-0.60548
b2	2.655699	2.006640	4.070678	0.797210	1.322053	1.638673	1.065562	2.877440	4.216492	2.388473	1.641513	2.388082	1.179148	2.753277	1.966065

TABLA 3.9

FEMENINO; MATRIZ DE CORRELACION

	h1	h0	s0	s1	theta	to	pv	hto	hpv	vto	vpv	incto-pv	incpv-ad	%adto	%adpv
h1	1.000	0.972	0.082	-0.001	0.024	0.256	0.069	0.843	0.942	0.071	-0.145	-0.303	0.068	0.451	0.227
h0	0.972	1.000	0.263	-0.035	-0.042	0.240	-0.018	0.945	0.966	0.144	-0.248	-0.494	-0.093	0.646	0.367
s0	0.082	0.263	1.000	0.565	-0.724	-0.404	-0.795	0.487	0.293	0.759	0.051	-0.603	-0.638	0.733	0.647
s1	-0.001	-0.035	0.565	1.000	-0.855	-0.742	-0.703	0.045	0.183	0.321	0.780	0.209	-0.775	0.163	0.733
theta	0.024	-0.042	-0.724	-0.855	1.000	0.919	0.949	-0.139	-0.183	-0.597	-0.633	0.002	0.612	-0.262	-0.593
to	0.256	0.249	-0.404	-0.742	0.919	1.000	0.893	0.204	0.119	-0.521	-0.729	-0.260	0.370	0.009	-0.296
pv	0.069	-0.018	-0.795	-0.703	0.949	0.893	1.000	-0.146	-0.090	-0.756	-0.460	0.175	0.468	-0.313	-0.439
hto	0.843	0.945	0.487	0.045	-0.139	0.204	-0.146	1.000	0.910	0.220	-0.356	-0.712	-0.291	0.860	0.534
hpv	0.942	0.966	0.293	0.183	-0.183	0.119	-0.090	0.910	1.000	0.093	-0.046	-0.350	-0.270	0.620	0.540
vto	0.071	0.144	0.759	0.321	-0.597	-0.521	-0.756	0.220	0.093	1.000	0.154	-0.340	-0.073	0.297	0.089
vpv	-0.145	-0.248	0.051	0.780	-0.633	-0.729	-0.460	-0.356	-0.046	0.154	1.000	0.727	-0.277	-0.443	0.226
incto-pv	-0.303	-0.494	-0.603	0.209	0.002	-0.260	0.175	-0.712	-0.350	-0.340	0.727	1.000	0.200	-0.809	-0.208
incpv-ad	0.060	-0.083	-0.638	-0.775	0.612	0.370	0.468	-0.291	-0.270	-0.073	-0.277	0.200	1.000	-0.556	-0.956
%ad-to	0.451	0.646	0.733	0.163	-0.262	0.089	-0.313	0.860	0.620	0.297	-0.443	-0.809	-0.556	1.000	0.679
%adto	0.227	0.367	0.647	0.733	-0.593	-0.296	-0.439	0.534	0.540	0.089	0.226	-0.208	-0.956	0.679	1.000

TABLA IV

FEMENINO; DATOS AJUSTADOS; PARAMETROS BIOLOGICOS

ESTIMACION MAXIMO VEROSIMIL

serie	h1	h0	s0	s1	theta	to	pv	hto	hpv	vto	vpv	incto-pv	incpv-ad	%adto	%adpv
113	159.653	147.533	0.1118	0.8927	12.9620	8.580	12.034	121.315	141.504	5.227	6.641	20.189	18.149	75.99	88.63
114	161.647	148.073	0.0926	0.0859	13.8479	8.854	13.148	119.560	143.279	4.509	6.975	23.719	18.268	74.01	88.64
120	162.466	151.529	0.1198	1.1051	13.0260	9.219	12.531	127.519	147.704	4.370	7.857	20.186	15.262	[illegible]	90.64
121	161.130	152.394	0.1658	1.2686	11.1149	8.134	10.405	133.523	147.641	5.672	6.902	14.118	13.489	82.87	91.63
209	165.676	153.361	0.1246	1.3073	11.4381	7.878	11.025	127.852	149.600	5.367	9.233	21.748	16.077	77.17	90.30
325	162.287	151.636	0.1257	1.0947	12.3116	8.461	11.671	128.472	147.146	4.945	7.009	18.674	15.141	79.16	90.67
326	163.540	153.268	0.1272	1.2707	12.5406	8.991	12.091	131.735	149.868	4.651	7.557	18.134	13.672	80.55	91.64
327	168.136	157.103	0.1209	1.0156	14.0155	10.011	13.263	133.291	152.131	5.057	6.772	18.841	16.005	74.24	90.49
409	168.737	159.501	0.1361	1.0019	12.5907	8.946	11.623	136.500	152.436	5.557	6.420	15.937	16.300	80.90	90.34
428	162.639	152.405	0.1204	0.9192	13.6082	9.503	12.624	130.303	146.007	4.931	5.864	16.504	15.932	80.12	90.27
501	163.208	152.340	0.1340	1.2579	11.4754	7.998	10.968	123.255	148.346	5.219	8.024	19.091	14.865	79.20	90.84
503	162.474	151.514	0.1445	1.2338	11.2773	7.943	10.675	127.897	146.720	5.968	8.133	18.831	15.705	78.72	[illegible]
518	154.416	142.712	0.1187	0.9332	13.0862	8.940	12.168	117.396	136.701	5.308	6.739	19.305	17.715	76.03	90.53
539	168.598	158.540	0.1314	0.9385	12.8367	9.063	11.739	137.701	152.139	5.298	6.038	15.068	16.459	81.50	90.21
610	167.919	156.791	0.1338	1.1918	11.8163	8.264	11.234	132.812	152.243	5.545	7.914	19.431	15.676	79.09	90.60
640	169.489	153.239	0.1128	0.7347	14.3568	10.067	12.621	135.349	149.064	5.230	5.627	13.715	20.425	79.86	87.96

TABLA 3.11

FEMENINO

# SERIE	MINIMOS CUADRADOS resi> 1 cm	media	var	sesgos	det(vcs)	MAXIMA VEROSIMILITUD resi> 1 cm	media	var	sesgos
113	4	0.00000	0.59824	s0,s1	1.3E-02	3	-0.00463	0.63781	todos
114	1	0.00000	0.33459	s0,s1,th	1.7E-02	2	0.02045	0.35171	todos
120	0	0.00000	0.11967	no	1.1E-02	0	-0.01236	0.12237	no
121	0	0.00000	0.05736	no	9.0E-03	0	0.00506	0.05891	no
209	0	0.00000	0.12561	no	2.0E-02	0	0.02415	0.14383	no
325	0	0.00000	0.11024	no	3.3E-02	0	-0.00765	0.01143	no
326	0	0.00000	0.23394	s1	1.7E-02	2	0.00118	0.23469	s1
327	0	0.00000	0.12379	s0,s1	4.2E-03	0	0.02153	0.13959	s0,s1
409	0	0.00000	0.25430	s0,s1	1.9E-02	1	0.00509	0.25965	s0,s1
428	0	0.00000	0.33131	s0,s1,th	5.4E-03	0	0.02692	0.33452	todos
501	1	0.00000	0.20608	no	2.7E-02	1	0.00971	0.21343	no
503	0	0.00000	0.22224	no	2.0E-02	0	0.00735	0.23523	s0,s1
518	0	0.00000	0.09856	s1	1.9E-02	0	-0.01152	0.10067	s0
539	0	0.00000	0.06717	s1	8.9E-03	0	0.00678	0.06781	s1
618	0	0.00000	0.13933	no	1.7E-02	0	-0.02045	0.15131	no
640	1	0.00000	0.38467	todos	4.0E-03	1	0.02146	0.40512	todos1

TABLA 3.12
FEMENINO

ESTADISTICAS DESCRIPTIVAS; 16 INDIVIDUOS

		MEDIA	D.S.	C.V.	MAX.	MIN.	SQRT(B1)	B2
HI								
	M.C.	163.8199	4.0681	0.0249	169.3459	154.3195	-0.4509	2.9168
	M.V.	163.9009	4.0237	0.0246	169.4889	154.4163	-0.4617	3.0176
HTHETA								
	M.C.	152.8981	4.3944	0.0287	158.6785	142.7172	-0.4928	2.9765
	M.V.	152.8625	4.3732	0.0286	158.5397	142.7121	-0.5447	2.9931
S0								
	M.C.	0.1271	0.0159	0.1198	0.1637	0.0944	0.3839	4.1171
	M.V.	0.1261	0.0159	0.1268	0.1658	0.0926	0.4499	4.2987
S1								
	M.C.	1.0998	0.1595	0.1450	1.2944	0.7867	-0.2863	1.8540
	M.V.	1.0707	0.1772	0.1655	1.3078	0.7347	-0.1925	1.7811
THETA								
	M.C.	12.6453	1.0312	0.0815	14.3491	11.0939	-0.0080	1.8815
	M.V.	12.6448	1.0167	0.0804	14.3568	11.1149	0.0200	1.9019

TABLA 3.13
FEMENINO
RESIDUALES AGRUPADOS POR % DE ESTATURA ADULTA

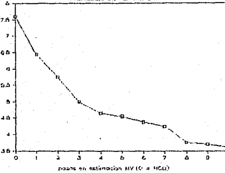
% EST. AD.	n	MEDIA	D.S.	E.S.	C.A.
60.00 - 64.99	12	0.38	0.51	0.15	0.52
65.00 - 69.99	34	-0.17	0.43	0.07	-0.86
70.00 - 74.99	47	-0.08	0.35	0.05	-0.79
75.00 - 79.99	47	0.03	0.39	0.06	-0.02
80.00 - 84.99	48	0.16	0.39	0.06	-0.36
85.00 - 89.99	40	-0.10	0.44	0.07	0.14
90.00 - 94.99	38	-0.04	0.41	0.07	-0.51
95.00 - 100.00	162	0.01	0.35	0.03	-0.19

TABLA 3.14
RCO

PARAMETROS		MEDIA	t	GL	D.S.	F (16,29)	C.V.	MAX.	MIN.	SQRT(B1)	B2
h1											
	FEM	163.820	-3.4348 *	42.56	4.089	0.2371 *	0.025	169.346	154.320	-0.451	2.911
	MASC	175.765			8.396		0.049	191.798	155.491	-0.339	2.432
h3											
	FEM	152.698	-2.8555 *	43.00	4.394	0.2989 *	0.029	158.679	142.717	-0.499	2.977
	MASC	162.709			8.039		0.049	177.080	142.728	-0.311	2.585
s0											
	FEM	0.127	1204.1297 *	21.21	0.016	2.7301 *	0.120	0.164	0.094	0.384	4.117
	MASC	0.104			0.010		0.092	0.134	0.086	0.060	4.264
s1											
	FEM	1.100	-1.7757 *	43.00	0.160	1.1773	0.145	1.294	0.787	-0.286	1.854
	MASC	1.184			0.147		0.124	1.421	0.832	-0.158	2.567
theta											
	FEM	12.645	-5.4431 *	43.00	1.031	1.7364	0.082	14.349	11.094	-0.008	1.832
	MASC	14.132			0.783		0.055	15.422	12.817	-0.070	1.674 *
to											
	FEM	8.857	-5.0187 *	43.00	0.693	0.8434	0.078	10.074	7.607	0.076	1.639 *
	MASC	10.004			0.755		0.076	11.056	8.218	-0.361	2.024
pv											
	FEM	11.924	-23.1935 *	20.28	0.876	3.2097 *	0.073	13.340	10.366	-0.072	1.866
	MASC	13.228			0.489		0.037	14.390	12.640	0.467	2.382
hto											
	FEM	129.613	-2.9061 *	43.00	5.457	0.4971	0.042	137.075	117.673	-1.454 *	2.577
	MASC	136.233			7.819		0.057	151.837	117.553	-0.101	2.459
hpv											
	FEM	148.045	-2.2719 *	42.61	4.104	0.2402 *	0.028	152.973	135.746	-2.354 *	4.216
	MASC	155.756			8.374		0.054	175.247	138.455	0.072	2.302
vto											
	FEM	5.169	37.8251 *	37.77	0.381	0.8819 *	0.074	5.912	4.450	-0.018	2.388
	MASC	4.411			0.405		0.098	5.487	3.732	0.116	2.452
vpv											
	FEM	7.226	4.2692 *	41.39	0.916	0.4388 *	0.127	9.042	5.641	0.103	1.642 *
	MASC	6.721			1.383		0.206	9.451	3.673	0.025	2.076
into-pv											
	FEM	18.401	-0.3112	33.87	2.418	0.0622 *	0.131	23.312	14.270	0.227	2.398
	MASC	19.523			9.692		0.496	36.622	7.523	1.199 *	1.624
Incpv-ad											
	FEM	15.774	-1.0753	29.32	1.366	0.0132 *	0.087	18.822	13.597	0.569	1.179 *
	MASC	21.154			11.901		0.563	60.483	3.034	3.172 *	4.789
Iad-to											
	FEM	79.115	2.3645 *	43.00	2.024	0.7934	0.026	82.620	74.620	-0.743	2.753
	MASC	77.503			2.272		0.029	81.090	72.270	-0.834 *	2.405
Iad-pv											
	FEM	90.365	1.9074 *	30.94	0.862	0.0298 *	0.010	91.660	88.610	-0.605	1.966
	MASC	89.727			4.993		0.056	98.160	81.370	0.942 *	1.606 *

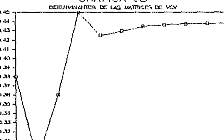
GRAFICA 3A

PROMEDIOS DE ITERACIONES



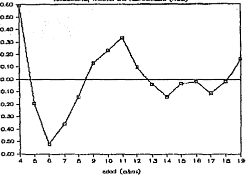
GRAFICA 3B

DETERMINANTES DE LAS MATRICES DE VCV



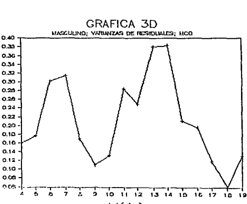
GRAFICA 3C

MASCULINO; MEDIAS DE RESIDUALES (MCO)

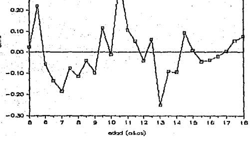


GRAFICA 3D

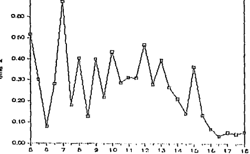
MASCULINO; VARIANZAS DE RESIDUALES; MCO



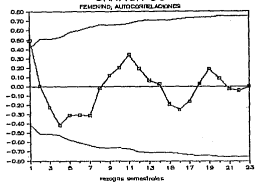
GRAFICA 3E



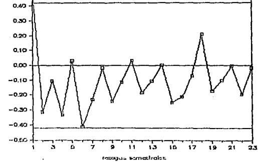
GRAFICA 3F



GRAFICA 3G

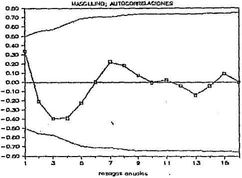


GRAFICA 3H



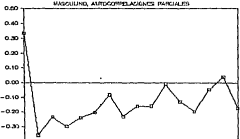
GRAFICA 3I

MASCULINO; AUTOCORRELACIONES

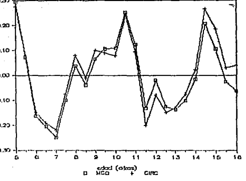


GRAFICA 3J

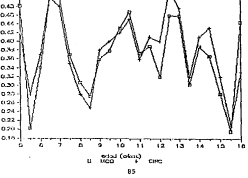
MASCULINO, AUTOCORRELACIONES PARCIALES



GRAFICA 3K



GRAFICA 3L



4) COMPARACION DE TRES METODOS DE ESTIMACION BAJO DIFERENTES ESTRUCTURAS DE ERROR

4.1) INTRODUCCION

Como se vió en el capítulo anterior, el número de individuos con los que se trabajó es relativamente reducido, lo que hace cuestionable cualquier intento de comparación entre métodos de estimación. Como un intento de profundizar en el conocimiento del modelo PB1, se presentan en este capítulo los resultados de un estudio de simulación Monte Carlo efectuado para ambos sexos considerando tres métodos de estimación: MCO, MC generalizado utilizando una matriz circular (que en adelante se llamará CIRC, y que no es otro que el método llamado MV en el capítulo anterior) y autorregresivo de primer orden (AR1). Los principios de éstos métodos de estimación fueron expuestos en el capítulo 2. Las ideas fundamentales de los procedimientos realizados en las simulaciones de este capítulo están tomadas de los trabajos de Gallant y Goebel (1976) y de Gallant (1987).

4.2) CONSTRUCCION DE LAS SIMULACIONES

A grandes rasgos, estos procedimientos consisten en lo siguiente:

Supóngase que se simulan, para cada sexo, R individuos, cada uno de ellos con observaciones en los tiempos

$$t = (t_1, \ldots, t_n).$$

Para cada repetición, se tiene que:

$$\tilde{y}(t_j) = f(\underline{\theta}^*;\ t_j) + u_j, \qquad j = 1, \ldots, n \qquad (4.1)$$

donde f es el modelo PDI, $\underline{\theta}^*$ es un valor paramétral fijo y u_j es un error aleatorio.

Los errores se generan de las siguientes formas:

$$\begin{array}{ll} nlid & u_j = e_j \\ ar1 & u_j = \alpha_1 u_{j-1} + e_j \end{array}$$

siendo e_j una variable aleatoria con distribución normal con media 0 y varianza σ^2. La generación de errores se realizó con la rutina FTGEN del paquete IMSL (IMSL Corporation, 1977).

Los pasos seguidos en cada simulación fueron los siguientes:

a) Generar $f(\underline{\theta}^*;\ t_j) + u_j$, $j = 1, \ldots, n$.

b) Para cada una de las estructuras de error consideradas obtener los estimadores MCO para a partir de ellos y de los residuales calculados, obtener los estimadores CIRC y AR1.

El hecho de que el valor del parámetro utilizado sea el mismo ($\underline{\theta}^*$) para todos los individuos se debe a que un punto de interés de las simulaciones está en obtener estimaciones del error cuadrático medio (ECM) que produce cada método bajo diferentes estructuras de error. La naturaleza no lineal del modelo dificulta la posibilidad de considerar distintos valores paramétricos en la estimación del ECM de cada método.

Para cada caso, $\underline{\theta}^*$ fue el valor del estimador de un individuo cuyo ajuste se consideró como adecuado. Estos valores fueron los correspondientes al individuo 120 en las mujeres y al individuo 639 en los hombres y son los siguientes:

FEMENINO:

h_1 = 162.888; h_θ = 151.506;

σ_0 = .1202; σ_1 = 1.1944; θ = 13.0166.

MASCULINO:

h_1 = 177.2398; h_θ = 165.3229;

σ_0 = .10481; σ_1 = 1.25456; θ = 14.70789.

Por otra parte, los valores utilizados como σ^2 y α_1 en las simulaciones fueron .02164 y .066047 para las simulaciones femeninas y masculinas, respectivamente y .496 y .393 para el coeficiente autorregresivo. Estos valores se obtuvieron como la varianza de la serie de medias de los residuales para cada sexo. Los coeficientes α_1 correspondientes se estimaron por máxima verosimilitud con la rutina FTMXL del paquete IMSL.

En todas las combinaciones de estructura de error con método de estimación se calculó el estimador del k-ésimo componente del ECM como:

$$\widehat{ECM}_k = \sum_{i=1}^{R} (\hat{\theta}_{ik} - \theta_k^*)^2 / R, \qquad (4.2)$$

siendo $\hat{\theta}_{ik}$ el k-ésimo componente del estimador en la i-ésima repetición, θ_k^* el k-ésimo componente del valor paramétrico y R el número de repeticiones. El estimador de la varianza del estimador $\widehat{ECM}_k$ se obtuvo con la expresión:

$$\hat{\sigma}^2_k = \sum_{i=1}^{R} ((\hat{\theta}_{ik}-\theta^*_k)^2 - \widehat{ECM}_k)^2/(R-1)$$

4.3) RESULTADOS

Otras preguntas que se buscaron responder con el estudio de simulación fueron referentes a dos clases de distribuciones marginales: por un lado la de los estimadores; por otro de la estadística t_k (ecuación 2.13) para las posibles combinaciones de método de estimación y estructura de error. Respecto a la primera, es importante verificar que la distribución del parámetro es normal para muestras pequeñas, pues si así sucede, se tendrá un elemento más para afirmar la baja no linealidad de los estimadores. El interés de examinar la segunda clase de distribuciones radica en analizar en qué condiciones puede aproximarse satisfactoriamente por una distribución t de Student con (n-1) grados de libertad para muestras pequeñas.

Con este fin se construyeron, para cada componente del estimador, tablas de tres tipos:

1) Las tablas 4.1 y 4.2 muestran, para las 6 combinaciones y para cada componente del estimador, su media y varianza, su $\widehat{ECM}$ y la varianza de éste así como la eficiencia respecto al estimador MCO correspondiente. Como se ve, todas las combinaciones produjeron estimaciones muy cercanas a $\underline{\theta}^*$. Es igualmente notable que la eficiencia de los estimadores CIRC resultó, para ambos errores, ser menor que 1, acentuándose un poco cuando el error es *ari*, mientras que la correspondiente a los estimadores CRI fue prácticamente 1. Aunque los resultados no se reportan en tabla alguna, se constrastó la hipótesis de

normalidad para los componentes de los estimadores tanto con las estadísticas $\sqrt{b_1}$ y b_2 de D'Agostino y Pearson como con la prueba de Kolmogorov-Smirnov, utilizando el paquete SPSS9 (Hadlai y Nie, 1981). Esta hipótesis fue aceptada, si bien hubo algunas excepciones, localizadas todas ellas en los parámetros σ_0 y σ_1.

2) Una primera aproximación a la distribución empírica de los t-cocientes aparece en las tablas 4.3 y 4.4. Se tienen ahí los cuantiles .05 y .95 para cada componente en las 6 combinaciones método-error comparándolos con los valores de tablas de la distribución t de Student con los grados de libertad correspondiente (n-p). Unicamente la combinación MCO-$niid$ produjo, para los cinco componentes, valores que, en vista del error estándar pueden considerarse como equivalentes a los de una t de Student. Para las combinaciones restantes, se puede observar un cierto grado de asimetría y el hecho de que la distribución empírica tiene colas más pesadas que una t de Student. El error estándar que aparece mencionado en estas tablas se calculó con la fórmula

$$[\ \alpha\ (1-\alpha)\ /\ R\ ]^{1/2}\ /\ [\ f(t_\alpha)\ ](4.4) \qquad (4.3)$$

(Schaefer, 1974; Gallant y Goebel, 1976), siendo $f(t_\alpha)$ la función de densidad de una t de Student y t_α el α-ésimo cuantil de esa distribución. En este caso, α = .05. El cálculo de este error estándar se basa en suponer que si $\tilde{t}_k$ se distribuye según una t de Student, entonces el estimador Monte Carlo de t_α se distribuye asintóticamente normal, con error estándar dado en (4.3).

3) Las tablas 4.5 y 4.6 presentan una visión más completa de las distribuciones empíricas de los t-cocientes. Otra vez se observa un gran parecido entre éstas y la t de Student correspondiente cuando se estima MCO con *niid*, encontrándose asimetrías y colas más pesadas en las otras combinaciones. Los errores estándares que aparecen en estas tablas se obtuvieron debido a que, bajo las mismas consideraciones hechas en el párrafo anterior, el estimador Monte Carlo de P $[\ \tilde{t}_k \leq t_\alpha\ ]$ tiene un error estándar

$$[\ \alpha\ (1-\alpha)\ /\ R\ ]^{1/2} \tag{4.4}$$

La expresión anterior fue utilizada previamente a las simulaciones con el fin de explorar las consecuencias que tendría la adopción de diferentes números de repeticiones en la estimación de las probabilidades mencionadas en este párrafo.

Las gráficas 4A a 4D presentan algunos ejemplos de las situaciones arriba descritas. En 4A aparecen las distribuciones para h_1 bajo errores normales con los tres métodos; la gráfica 4B tiene las distribuciones del mismo estimador pero con errores *ari*. En 4C se tiene un ejemplo en que la distribución del estimador no fue normal: se trata de o_0 para CIRC-*niid*. Finalmente, la gráfica 4D ilustra un resultado que se repitió en todos los t-cocientes. La distribución de MCO-*ari*, al igual que la correspondiente a MCO-*niid*, está centrada en 0, aunque tiene colas más pesadas, así como un cierto grado de asimetría.

TABLA 4.1

MCO
FEMENINO

ERROR	media	var	ecm	var(ecm)	eff
			h1		
MIID	162.8909	0.016461	0.016550	0.000499	1
AR1	162.8097	0.035013	0.034948	0.002525	1
			h0		
MIID	151.5160	0.009176	0.009258	0.000177	1
AR1	151.5081	0.025260	0.025223	0.001275	1
			s0		
MIID	0.120194	0.000001	0.000001	0.000000	1
AR1	0.120220	0.000002	0.000002	0.000000	1
			s1		
MIID	1.194791	0.000431	0.000431	0.000000	1
AR1	1.196873	0.001101	0.001105	0.000003	1
			theta		
MIID	13.01844	0.000358	0.000361	0.000000	1
AR1	13.01632	0.000961	0.000959	0.000002	1

CIRC
FEMENINO

ERROR	media	var	ecm	var(ecm)	eff
			h1		
MIID	162.0945	0.023009	0.023083	0.001323	0.692966
AR1	162.0833	0.046340	0.046269	0.005069	0.755320
			h0		
MIID	151.5144	0.010630	0.010603	0.000239	0.866661
AR1	151.5064	0.026403	0.026431	0.001592	0.954290
			s0		
MIID	0.120227	0.000002	0.000002	0.000000	0.621179
AR1	0.120201	0.000003	0.000003	0.000000	0.855612
			s1		
MIID	1.195722	0.000738	0.007304	0.000001	0.583692
AR1	1.190205	0.001542	0.001554	0.000005	0.710956
			theta		
MIID	13.01813	0.000406	0.000408	0.000000	0.986736
AR1	13.01583	0.001032	0.001030	0.000002	0.931032

AR1
FEMENINO

ERROR	media	var	ecm	var(ecm)	eff
			h1		
MIID	162.8987	0.016340	0.016423	0.000482	1.007749
AR1	162.8894	0.035705	0.035634	0.002695	0.980744
			h0		
MIID	151.5159	0.009203	0.009281	0.000181	0.997508
AR1	151.5001	0.024857	0.024812	0.001257	1.016558
			s0		
MIID	0.120185	0.000001	0.000001	0.000000	0.991871
AR1	0.120231	0.000002	0.000002	0.000000	1.032124
			s1		
MIID	1.194818	0.000426	0.000426	0.000000	1.011863
AR1	1.196884	0.001105	0.001109	0.000003	1.032124
			theta		
MIID	13.01846	0.000360	0.000362	0.000000	0.997045
AR1	13.01626	0.000955	0.000953	0.000002	1.006315

TABLA 4.2

ERROR	MCO MASCULINO media	var	ecm	var(ecm)	eff	CIRC MASCULINO media	var	ecm	var(ecm)	eff	RR1 MASCULINO media	var	ecm	var(ecm)	eff
			h1					h1					h1		
NIID	177.2829	0.031058	0.031041	0.001847	1	177.2352	0.039613	0.039555	0.003616	0.784775	177.2334	0.031656	0.031634	0.001974	0.981255
AR1	177.2394	0.001211	0.049496	0.004360	1	177.2365	0.057929	0.057822	0.006543	0.056000	177.2388	0.049865	0.049748	0.004483	0.994931
			h0					h0					h0		
NIID	165.3205	0.024119	0.024080	0.001080	1	165.3219	0.027949	0.027895	0.001504	0.863222	165.3200	0.024383	0.024312	0.001085	0.990432
AR1	165.3221	0.038354	0.038280	0.002790	1	165.3215	0.040128	0.040047	0.003117	0.985081	165.3223	0.038937	0.038861	0.002913	0.985051
			s0					s0					s0		
NIID	0.104826	0.000001	0.000001	0.000000	1	0.104821	0.000001	0.000001	0.000000	0.674022	0.104817	0.000001	0.000001	0.000000	0.930168
AR1	0.104702	0.000002	0.000002	0.000000	1	0.104804	0.000002	0.000002	0.000000	0.793595	0.104791	0.000002	0.000002	0.000000	0.932661
			s1					s1					s1		
NIID	1.287421	0.001804	0.001809	0.000006	1	1.257586	0.002606	0.002610	0.000013	0.693134	1.257427	0.001872	0.001876	0.000006	0.964223
AR1	1.253698	0.002466	0.002163	0.000012	1	1.255906	0.003468	0.003465	0.000034	0.624541	1.259032	0.002242	0.002237	0.000012	0.966659
			theta					theta					theta		
NIID	14.70929	0.001211	0.001210	0.000003	1	14.70968	0.001359	0.001359	0.000003	0.890428	14.70924	0.001229	0.001228	0.000003	0.985681
AR1	14.70915	0.001522	0.001920	0.000007	1	14.70900	0.002019	0.002016	0.000009	0.952337	14.70914	0.001936	0.001933	0.000008	0.992921

TABLA 4.3

ERROR	MCO MASCULINO 5%	MCO MASCULINO 95%	CIRC MASCULINO 5%	CIRC MASCULINO 95%	AR1 MASCULINO 5%	AR1 MASCULINO 95%
		h1		h1		h1
NIID	-1.771	1.660	-3.241	3.291	-2.298	2.300
AR1	-2.290	2.544	-3.684	3.562	-2.947	3.207
		h0		h0		h0
NIID	-1.803	1.894	-4.143	3.769	-2.468	2.422
AR1	-2.487	2.567	-4.095	4.625	-2.992	3.108
		s0		s0		s0
NIID	-1.809	1.759	-3.554	3.831	-2.630	2.494
AR1	-2.557	2.466	-4.255	4.299	-3.277	2.795
		s1		s1		s1
NIID	-1.897	1.949	-3.587	3.008	-2.586	2.582
AR1	-2.519	2.204	-3.624	3.019	-2.907	2.469
		theta		theta		theta
NIID	-1.753	1.699	-3.403	3.574	-2.471	2.514
AR1	-2.467	2.426	-3.677	4.211	-3.091	2.932
		t		t		t
gl = 11	-1.796	1.796	-1.796	1.796	-1.796	1.796
error std.	0.219	0.219	0.219	0.219	0.219	0.219

TABLA 4.4

ERROR	MCO FEMENINO 5%	MCO FEMENINO 95%	CIRC FEMENINO 5%	CIRC FEMENINO 95%	AR1 FEMENINO 5%	AR1 FEMENINO 95%
		h1		h1		h1
NIID	-1.809	1.779	-3.381	3.174	-2.221	2.257
AR1	-3.191	2.556	-4.002	3.860	-2.817	2.693
		h0		h0		h0
NIID	-1.540	1.825	-3.407	3.500	-2.038	2.200
AR1	-3.171	2.912	-4.616	4.530	-2.983	2.896
		s0		s0		s0
NIID	-1.662	1.769	-3.479	3.697	-2.059	2.248
AR1	-2.701	2.945	-3.099	4.558	-2.783	2.769
		s1		s1		s1
NIID	-1.759	1.745	-3.619	3.144	-2.168	2.231
AR1	-2.936	3.033	-4.352	4.474	-3.019	2.869
		theta		theta		theta
NIID	-1.485	1.634	-3.453	3.524	-2.084	2.081
AR1	-3.039	2.756	-4.711	4.309	-2.963	2.671
		t		t		t
gl = 18	-1.734	1.734	-1.734	1.734	-1.734	1.734
error std.	0.199	0.199	0.199	0.199	0.199	0.199

TABLA 4.5
DISTRIBUCION EMPIRICA; T-COCIENTE

VALORES TABULARES
gl = 18

c	P [T<=c]	MIID MCO FEMENINO h1 P [t<=c]	h0 P [t<=c]	s0 P [t<=c]	s1 P [t<=c]	theta P [t<=c]	AR1 MCO FEMENINO h1 P [t<=c]	h0 P [t<=c]	s0 P [t<=c]	s1 P [t<=c]	theta P [t<=c]	error std
-2.6520	0.0100	0.010	0.010	0.008	0.008	0.010	0.070	0.084	0.062	0.066	0.080	0.0044
-2.1010	0.0250	0.026	0.016	0.022	0.022	0.010	0.096	0.118	0.102	0.106	0.120	0.0070
-1.7340	0.0500	0.050	0.036	0.040	0.052	0.030	0.146	0.160	0.148	0.150	0.154	0.0097
-1.3300	0.1000	0.096	0.072	0.086	0.088	0.072	0.214	0.222	0.216	0.206	0.216	0.0134
-0.6880	0.2500	0.224	0.210	0.250	0.220	0.208	0.326	0.360	0.354	0.312	0.372	0.0194
0.0000	0.5000	0.456	0.442	0.524	0.500	0.456	0.492	0.474	0.476	0.492	0.514	0.0224
0.6880	0.7500	0.736	0.738	0.744	0.742	0.724	0.676	0.634	0.650	0.644	0.662	0.0194
1.3300	0.9000	0.882	0.898	0.900	0.886	0.920	0.790	0.700	0.778	0.770	0.780	0.0134
1.7340	0.9500	0.948	0.942	0.946	0.940	0.960	0.856	0.832	0.828	0.822	0.836	0.0097
2.1010	0.9750	0.972	0.978	0.976	0.976	0.982	0.908	0.876	0.876	0.874	0.890	0.0070
2.6520	0.9900	0.994	0.988	0.986	0.994	0.990	0.948	0.910	0.900	0.914	0.930	0.0044

VALORES TABULARES
gl = 18

c	P [T<=c]	MIID CIRC FEMENINO h1 P [t<=c]	h0 P [t<=c]	s0 P [t<=c]	s1 P [t<=c]	theta P [t<=c]	AR1 CIRC FEMENINO h1 P [t<=c]	h0 P [t<=c]	s0 P [t<=c]	s1 P [t<=c]	theta P [t<=c]	error std
-2.5520	0.0100	0.086	0.092	0.094	0.096	0.086	0.120	0.164	0.154	0.152	0.154	0.0044
-2.1010	0.0250	0.110	0.114	0.150	0.150	0.118	0.160	0.202	0.186	0.176	0.204	0.0070
-1.7340	0.0500	0.144	0.144	0.190	0.184	0.140	0.200	0.238	0.226	0.210	0.252	0.0097
-1.3300	0.1000	0.210	0.190	0.256	0.236	0.190	0.252	0.304	0.270	0.266	0.310	0.0134
-0.6000	0.2500	0.290	0.306	0.356	0.344	0.314	0.348	0.390	0.370	0.352	0.410	0.0194
0.0000	0.5000	0.468	0.448	0.404	0.460	0.478	0.518	0.492	0.478	0.476	0.518	0.0224
0.6880	0.7500	0.666	0.620	0.632	0.606	0.622	0.662	0.608	0.600	0.576	0.630	0.0194
1.3300	0.9000	0.782	0.764	0.742	0.750	0.758	0.762	0.722	0.710	0.720	0.728	0.0134
1.7340	0.9500	0.840	0.828	0.810	0.828	0.826	0.828	0.776	0.764	0.770	0.764	0.0097
2.1010	0.9750	0.870	0.868	0.858	0.880	0.864	0.858	0.804	0.812	0.818	0.816	0.0070
2.5520	0.9900	0.916	0.892	0.900	0.914	0.894	0.886	0.838	0.874	0.854	0.862	0.0044

VALORES TABULARES
gl = 10

c	P [T<=c]	MIID AR1 FEMENINO h1 P [t<=c]	h0 P [t<=c]	s0 P [t<=c]	s1 P [t<=c]	theta P [t<=c]	AR1 AR1 FEMENINO h1 P [t<=c]	h0 P [t<=c]	s0 P [t<=c]	s1 P [t<=c]	theta P [t<=c]	error std
-2.5520	0.0100	0.030	0.022	0.020	0.028	0.026	0.072	0.074	0.066	0.068	0.080	0.0044
-2.1010	0.0250	0.058	0.045	0.044	0.060	0.048	0.094	0.122	0.100	0.106	0.108	0.0070
-1.7340	0.0500	0.090	0.060	0.076	0.078	0.066	0.140	0.160	0.138	0.154	0.152	0.0097
-1.3300	0.1000	0.126	0.114	0.140	0.110	0.116	0.202	0.222	0.214	0.196	0.200	0.0134
-0.6000	0.2500	0.260	0.246	0.204	0.204	0.220	0.316	0.332	0.336	0.304	0.368	0.0194
0.0000	0.5000	0.448	0.452	0.540	0.510	0.456	0.506	0.472	0.470	0.474	0.512	0.0224
0.6880	0.7500	0.704	0.692	0.720	0.710	0.690	0.680	0.650	0.642	0.638	0.684	0.0194
1.3300	0.9000	0.858	0.852	0.874	0.842	0.844	0.802	0.794	0.790	0.782	0.796	0.0134
1.7340	0.9500	0.906	0.904	0.914	0.906	0.920	0.872	0.840	0.848	0.850	0.852	0.0097
2.1010	0.9750	0.934	0.936	0.942	0.944	0.950	0.896	0.878	0.886	0.882	0.876	0.0070
2.5520	0.9900	0.970	0.968	0.966	0.962	0.972	0.934	0.928	0.926	0.928	0.942	0.0044

TABLA 4.6
DISTRIBUCION EMPIRICA; T-COCIENTE

VALORES TABULARES

gl = 11

c	P [T<=c]	MIIO MCO MASCULINO h1 P [t<=c]	h0 P [t<=c]	s0 P [t<=c]	s1 P [t<=c]	theta P [t<=c]	AR1 MCO MASCULINO h1 P [t<=c]	h0 P [t<=c]	s0 P [t<=c]	s1 P [t<=c]	theta P [t<=c]	error std
-2.7100	0.0100	0.006	0.002	0.008	0.016	0.008	0.030	0.038	0.040	0.036	0.034	0.0044
-2.2010	0.0250	0.024	0.022	0.022	0.038	0.022	0.056	0.074	0.068	0.066	0.068	0.0070
-1.7960	0.0500	0.046	0.054	0.052	0.054	0.046	0.110	0.114	0.106	0.102	0.098	0.0097
-1.3630	0.1000	0.108	0.092	0.094	0.118	0.098	0.174	0.178	0.152	0.162	0.146	0.0134
-0.6970	0.2500	0.264	0.268	0.240	0.262	0.210	0.302	0.302	0.310	0.320	0.286	0.0194
0.0000	0.5000	0.524	0.516	0.480	0.500	0.406	0.510	0.498	0.510	0.518	0.498	0.0224
0.6970	0.7500	0.778	0.750	0.746	0.722	0.742	0.692	0.702	0.684	0.738	0.690	0.0194
1.3630	0.9000	0.914	0.914	0.090	0.892	0.900	0.840	0.840	0.828	0.874	0.818	0.0134
1.7960	0.9500	0.966	0.944	0.950	0.930	0.958	0.888	0.890	0.904	0.912	0.894	0.0097
2.2010	0.9750	0.986	0.966	0.970	0.960	0.974	0.928	0.926	0.934	0.948	0.926	0.0070
2.7100	0.9900	0.988	0.986	0.994	0.980	0.990	0.950	0.952	0.968	0.980	0.964	0.0044

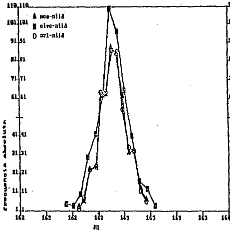
VALORES TABULARES

c	P [T<=c]	MIIO CIRC MASCULINO h1 P [t<=c]	h0 P [t<=c]	s0 P [t<=c]	s1 P [t<=c]	theta P [t<=c]	AR1 CIRC MASCULINO h1 P [t<=c]	h0 P [t<=c]	s0 P [t<=c]	s1 P [t<=c]	theta P [t<=c]	error std
-2.7180	0.0100	0.086	0.102	0.096	0.082	0.076	0.096	0.098	0.112	0.084	0.088	0.0044
-2.2010	0.0250	0.122	0.132	0.152	0.112	0.102	0.136	0.134	0.136	0.122	0.110	0.0070
-1.7960	0.0500	0.166	0.178	0.198	0.150	0.146	0.182	0.180	0.196	0.170	0.154	0.0097
-1.3630	0.1000	0.240	0.252	0.244	0.220	0.192	0.254	0.248	0.250	0.248	0.212	0.0134
-0.6970	0.2500	0.340	0.364	0.376	0.342	0.324	0.360	0.338	0.360	0.374	0.318	0.0194
0.0000	0.5000	0.508	0.510	0.478	0.458	0.490	0.508	0.514	0.600	0.518	0.502	0.0224
0.6970	0.7500	0.604	0.642	0.582	0.622	0.662	0.658	0.654	0.648	0.666	0.644	0.0194
1.3630	0.9000	0.806	0.776	0.724	0.718	0.770	0.768	0.756	0.736	0.788	0.752	0.0134
1.7960	0.9500	0.864	0.820	0.770	0.872	0.828	0.828	0.806	0.796	0.844	0.808	0.0097
2.2010	0.9750	0.896	0.864	0.842	0.882	0.880	0.876	0.850	0.842	0.880	0.856	0.0070
2.7100	0.9900	0.930	0.900	0.904	0.928	0.922	0.896	0.894	0.870	0.922	0.886	0.0044

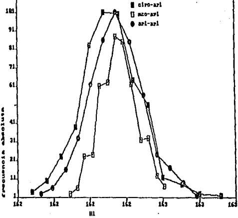
VALORES TABULARES

c	P [T<=c]	MIIO AR1 MASCULINO h1 P [t<=c]	h0 P [t<=c]	s0 P [t<=c]	s1 P [t<=c]	theta P [t<=c]	AR1 AR1 MASCULINO h1 P [t<=c]	h0 P [t<=c]	s0 P [t<=c]	s1 P [t<=c]	theta P [t<=c]	error std
-2.7180	0.0100	0.024	0.030	0.044	0.040	0.032	0.060	0.068	0.070	0.064	0.060	0.0044
-2.2010	0.0250	0.050	0.072	0.076	0.072	0.066	0.090	0.100	0.102	0.100	0.082	0.0070
-1.7960	0.0500	0.098	0.122	0.110	0.104	0.108	0.144	0.142	0.144	0.136	0.130	0.0097
-1.3630	0.1000	0.160	0.160	0.156	0.174	0.152	0.212	0.192	0.204	0.190	0.176	0.0134
-0.6970	0.2500	0.330	0.322	0.300	0.292	0.284	0.328	0.316	0.330	0.346	0.314	0.0194
0.0000	0.5000	0.520	0.510	0.466	0.504	0.490	0.500	0.496	0.512	0.512	0.484	0.0224
0.6970	0.7500	0.726	0.706	0.698	0.672	0.698	0.684	0.660	0.642	0.700	0.670	0.0194
1.3630	0.9000	0.854	0.830	0.836	0.822	0.834	0.818	0.812	0.796	0.840	0.778	0.0134
1.7960	0.9500	0.910	0.896	0.890	0.884	0.939	0.870	0.850	0.850	0.892	0.852	0.0097
2.2010	0.9750	0.946	0.936	0.928	0.914	0.934	0.890	0.869	0.909	0.932	0.900	0.0070
2.7180	0.9900	0.972	0.960	0.964	0.954	0.968	0.930	0.930	0.944	0.964	0.944	0.0044

GRAFICA 4A

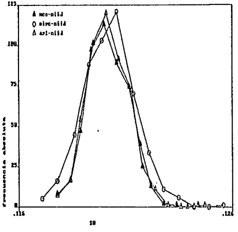


GRAFICA 4B


FEMENINO; H1
circ-ari
arc-ari
ari-ari
frecuencia absoluta
110
101
91
81
71
61
41
31
21
11
1
162
162
162
162
163
163
163
H1

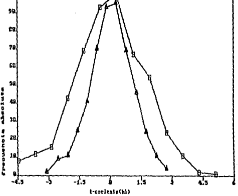
GRAFICA 4C

PENDULO; S0; ERRORES NORMALES



GRAFICA 4D

MASCULINO; MCO; T-COCIENTE (HI) CON DOS ERRORES



REFERENCIAS

Bard, Y. (1974). *Nonlinear parameter estimation*. Academic Press, Nueva York.

Bates, D.M. y D.G. Watts. (1980). "Relative curvature measure of nonlinearity (with discussion)" *JRSS-B*, 42, 1-25.

Beale, E.M.L. (1960). "Confidence regions in nonlinear estimation", *JRSS-B* 22, 41-76.

Berkey, C.S. (1982). "Comparison of two longitudinal growth models for preschool children". *Biometrics*, 38, 221-234.

Berkey, C.S., R.B. Reed e I. Valadian. (1983a). "Midgrowth spurt in height of Boston children", *Ann.Hum.Biol.* 10, 25-30.

Berkey, C.S., R.B. Reed e I. Valadian. (1983b). "Longitudinal growth standards for preschool children", *Ann.Hum.Biol.* 10, 57-67.

Berkey, C.S. y R.L. Kent. (1983). "Longitudinal principal components and non-linear regression models of early childhood growth", *Ann.Hum.Biol.* 10, 523-536.

Berkey, C.S. y N.M. Laird. (1986). "Nonlinear growth curve analysis: estimating the population parameters", *Ann.Hum.Biol.* 13, 111-128.

Billewickz, W.Z. e I.A. McGregor. (1982). "A birth-to-maturity longitudinal study of heights and weights in two West African (Gambian) villages, 1951-1975", *Ann.Hum.Biol.* 9, 309-320.

Bock, R.D., H. Wainer, A. Petersen, D. Thissen, J. Murray y A.F. Roche. (1973). "A parametrization for individual growth curves", *Hum.Biol.* 45,

Bock, R.D. y D. Thissen. (1976). "Fitting multi-component models for growth in stature", *Proceedings of the 9th International Biometric conference*. The Biometric Society, Boston.

Bowerman, B.L. y R.T. O'Connell. (1979). *Forecasting & time series*, Duxbury, Belmont.

Box, M.J. (1971). "Bias in nonlinear estimation". *JRSS-B* 33, 171-201.

Brillinger, D.R. (1975). *Time series. Data Analysis and theory*. Holt, Rinehart and Winston, Nueva York.

Brown, T. y G.C. Townsend. (1982). "Adolescent growth in height of australian aboriginals analysed by the Preece-Baines function: a longitudinal study". *Ann.Hum.Biol.* 9, 495-506.

Calderón, A. (1985). *Guía para el uso de la biblioteca básica de programas de análisis numérico. Parte I.* Comunicaciones Técnicas, serie azul, 50. IIMAS, UNAM, México.

Calderón, A., S. Gómez, J. Alonso, J. Capistrán, P. Guerrero, J.L. Morales y M. Sánchez. (1983). *NL2SOL: una subrutina altamente robusta para resolver problemas de mínimos cuadrados no lineales.* Comunicaciones técnicas, serie azul, 89, IIMAS, UNAM, México.

Cameron, N., J.M. Tanner y R.H. Whitehouse. (1982). "A longitudinal analysis of the growth of limb segment in adolescence". *Ann.Hum.Biol.* 9, 211-220.

Comas, J. (1966). *Manual de Antropología Física*, UNAM, México.

Cook, R.D. y C.L. Tsai. (1985). "Residuals in nonlinear regression", *Biometrika*, 72, 23-29.

Count, E.W. (1942). "A quantitative analysis of growth in certain human skull dimensions". *Hum.Biol.* 14, 143-166.

Count, E.W. (1943). "Growth patterns of human physique: an approach to kinetic anthropometry". *Hum.Biol.* 15, 1-32.

Cox, D.R. (1978). "Foundations of statistical inference: a case for eclecticism". *Australian Journal of Statistics*, 20, 43-57.

Cox, D.R. y D.V. Hinkley. (1979). *Theoretical Statistics.* Chapman and Hall, Londres.

D'Agostino, R. y E..S. Pearson, (1973). "Tests for departure from normality: Empirical distributions of b2 and √b1", *Biometrika*, 60, 613-622.

Deming,J. (1957). "Application of the Gompertz curve to the observed pattern of growth in lenght of 48 individuals boys and girls during the adolescence", *Hum.Biol.* 39, 83-122.

Deming, J. y A.H. Washburn. (1963). "Application of the Jenss curve to the observed pattern of growth during the first eight years of life in forty boys and forty girls", *Hum.Biol.* 41.

Draper, N.R. y H. Smith. (1981). *Applied regression analysis*, 2nd. ed., Wiley, Nueva York.

Durbin, J. y G.S. Watson. (1971). "Testing for serial correlation in least squares regression", *Biometrika*, 58, 1-19.

Edwards, A.W.F. (1984). *Likelihood*, Cambridge University Press, Cambridge.

El-Lozy, M. (1978)."A critical analysis of the double and triple logistic growth curves". *Ann.Hum.Biol.*, 5, 389-394.

Espinosa, G. y J.Faulhaber. (1979). "El análisis de cúmulos aplicado a datos longitudinales de crecimiento". *Anales de Antropología*, **XVI**, 433-447.

Faulhaber, J. (1976). *Estudio longitudinal del crecimiento*, INAH, México.

Faulhaber, J. (1982). "La predicción de la estatura adulta según varios métodos en niños mexicanos", *Anales de Antropología*, **XIX**, 93-120.

Fearn, T. (1975). "A bayesian approach to growth curves". *Biometrika*, **64**, 89-100.

Fuller, W.A. (1976). *Introduction to statistical time series*, Wiley, Nueva York.

Gallant, A.R. (1975a). "Non linear regression", *Am.Statistician*, **29**,73-81.

Gallant, A.R. (1975b). "Testing a subset of the regression parameters of a non linear regression model", *JASA*, **70**, 927-932.

Gallant, A.R. (1976). *Confidence regions for the parameters of a non linear regression model*, Institute of Satistics mim. series, **1077**, North Carolina State University, Raleigh.

Gallant, A.R. y J.J. Goebbel (1976). "Non linear regression with autocorrelated errors", *JASA*, **71**, 961-967.

Gallant, A.R., (1987), *Nonlinear statistical models*, Wiley, Nueva York.

Gasser, T., H.G. Muller, W. Kohler, A. Prader, R. Largo y L. Molinari, (1985). "An analysis of the mid-growth and adolescent spurts of height based on acceleration", *Ann.Hum.Biol.*, **12**, 129-148.

Gibbons, J.D. (1971). *Non parametric statistical inference*, McGraw-Hill, Nueva York.

Gillis, P.R. y D.A. Ratkowsky, (1978) "The behaviour of estimators of the parameters of various yield-density relationships", *Biometrics*, **34**, 191-198.

Glasbey, C.A. (1979). "Correlated residuals in non linear regression applied to growth data". *Appl.Stat.*, **28**, 251-259.

Goldstein, H. (1979). *The design and analysis of longitudinal studies*, Academic Press, Londres.

Guttman, I. y D.A. Meeter (1965). "On Beale's measures of nonlinearity". *Technometrics*, **7**, 623-637.

Hannan, E.J. (1971) "Non linear time series regression", *J.Appl.Prob.* **8**, 767-780.

Hannan, E.J. (1903) *Time series analysis*, Chapman and Hall, Londres.

Hartley, H.O. (1964). "Exact confidence regions for parameters in non linear regression laws", *Biometrika* 51, 347-353.

Hauspie, R.C., S.R. Das, M.A. Preece y J.M. Tanner (1980), "A longitudinal study of the growth in height of boys and girls of West Bengal (India) aged six months to 20 years" *Ann. Hum. Biol.* 7, 429

Hauspie, R.C., A. Wachholder, G. Baron, F. Cantraine, C. Susanne y M. Graffar (1980), "A comparative study of the fit of four different functions to longitudinal data of growth in height in belgian girls", *Ann. Hum. Biol.*, 7, 347

Hauspie, R.C., S.R. Das, M.A. Preece y J.M. Tanner (1982), "Degree of resemblance of the pattern of growth among sibs in families of West Bengal (India)", *Ann. Hum. Biol.* 9, 171-174.

Hull, C.H. y N.H. Nie (1981) *SPSS Update 7-9*. McGraw- Hill, Nueva York

Jennrich, R.I. (1969). "Asymptotic properties of non linear least squares estimator", *Ann. Math. Stat.*, 40, 633-643.

Jenss, R.M. y N. Bailey (1937). "A mathematical model for studying the growth of child", *Hum. Biol.* 9 556-563.

Johnston, F.E. (1980). "Research design and sample selection in studies of growth and development", en Johnston, F.E., A.F. Roche y C. Susanne (eds.): *Human physical growth and maturation: methodologies and factors*, pp. 5-20, Plenum, Nueva York.

Joosens, R.M. y E. Brems-Heyns (1975) "High power polynomial regression for the study of distance, velocity and acceleration of growth", *Growth*, 39, 535-

Jordán, R. (1979). *Desarrollo humano en Cuba*, Editorial Científica y Técnica, Ministerio de Cultura, La Habana.

IMSL Corporation (1977) *IMSL Reference Manual*, Houston.

Kalbfleisch, J.G. (1979) *Probability and Statistical Inference*, Springer Verlag, Berlin.

Kennedy, W.J. y J.E. Gentle (1980) *Statistical Computing*, Dekker, Nueva York.

Lee, P.A. (1980) "Independence of seasonal variation of growth from temperature change", *Growth*, 44, 54-57.

Largo, R.H., T. Gasser, A. Prader, W. Stützle y P.J. Huber (1978), "Analysis of the adolescent growth spurt using smoothing spline functions", *Ann. Hum. Biol.* 5, 421-

Malinvaud, E. (1970). "The consistency of nonlinear regressions", *Ann. Math. Stat.*, **41**, 956-969.

Manwani, A.H. y K.N. Agarwal (1973). "The growth pattern of indian infants during the first year of life", *Hum. Biol.*, **45**, 341-349.

Marubini, E. (1978). "The fitting of longitudinal growth data of man", en Gedda, L. y P. Parisi (eds.) *Auxology: Human growth in health and disorder*, pp. 121-131, Academic Press, Londres.

Marubini, E., L.F. Resele y G. Barghini (1971), "A comparative fitting of the Gompertz and the Logistic functions to longitudinal height data during adolescence in girls". *Hum. Biol.* **44**, 237-252.

Marubini, E. L.F. Resele, J.M. Tanner y R.H. Whitehouse (1972) "The fitting of Gompertz and Logistic curves to longitudinal data during adolescence on height sitting, height and biacromial diameter in boys and girls of the Harpenden growht study", *Hum. Biol.* **45**, 511-524.

Mirwald, R.C., D.A. Bailey, N. Cameron y R.L. Rasmuncen (1981), "Longitudinal comparison of aerobic power in active and inactive boys aged 7.0 to 17.0", *Ann. Hum. Biol.*, **8**, 405-414.

Molinari, L., R.H. Largo y A. Prader (1980), "Analysis of the growth spurt at age seven (mid-growth spurt)", *Helvetica Pediatrica Acta* **35**, 325-334.

Nelder, J.A. (1961), "The fitting of a generalization of the logistic curve", *Biometrics*, **17**, 89-110.

Nelder, J.A. (1962), "An alternative form of a generalized logistic equation", *Biometrics*, **18**, 614-616.

Preece, M.A. (1978), "Analysis of the Human growth curve", *Postgraduate Medical Journal*, *suppl.* 1, **54**, 77-86.

Preece, M.A. y M.J. Baines (1978), "A new family of mathematical models describing the human growth curve", *Ann. Hum. Biol.*, **5**, 1-24.

Preece, M.A. e I. Heinrich (1981), "Mathematical modelling of individual growth curves", *British Medical Bulletin*, **37**, 247-253.

Randless, R.H. y D.A. Wolfe (1979), *Introduction to the theory of non parametric statistics*, Wiley, Nueva York.

Ratkowsky, D.A. (1983), *Non linear regression modelling*, Marcel Dekker, Nueva York.

Richards, F.J. (1959), "A flexible growth function for empirical use", *Journal of Experimental Biology*, **10**, 290-300.

Roche, A.F., H. Wainer y D. Thissen (1975), *Predicting adult stature for individuals*, *Monographs in Pediatrics 3*, Kargen, Basilea.

Schafer, R.E. (1974), "On assesing the precision of Simulations", *Journal of Statistical Computation and Simulation*, 3, 67-70.

Schlaepfer, L.V. (1986), "Consideraciones de tipo metodológico en relación al análisis de un estudio longitudinal de crecimiento de niños", *Memorias del 1er foro de Estadística Aplicada*, UNAM, México.

Silvey, S.D. (1975), *Statistical Inference*, Chapman and Hall, Londres.

Smith, E.O., R.J. Schandler , C. Garza y B.L. Nichols (1983), "Modelling the growth pattern of premature infants", *Growth*, 47, 340-347.

Sprott, D.A. (1984) "Verosimilitud y estimación máximo verosímil", versión en español de un artículo publicado en *Mathematical Reports of the Academy of Science*, 6, 1-15. Royal Society of Canada.

Stützle, W., Th. Gasser, L. Molinari, R.H. Largo, A. Prader y P.J. Huber (1980), "Shape-invariant modeling of human growth", *Ann. Hum. Biol.*, 7, 507-528.

Susanne, C. (1980), "Developmental genetics of man", en Johnston, F.E., A.F. Roche y C. Susanne (eds.): *Human Physical Growth and Maturation: methodologies and factors*, 221-242, Plenum, Nueva York.

Tanner, J.M. (1978), "Human growth standards: construction and use", en *Auxology in health and disorder*, 109-121.

Tanner, J.M. y N. Cox (1986), "The London studies of Human Growth", *Ann. Hum. Biol.*, 13, 510-512.

Tanner, J.M., R.H. Whitehouse, E. Marubini y L.F. Resele (1976), "The adolescent growth spurt of boys and girls of the Harpenden growth study", *Ann. Hum. Biol.*, 3, 109.

Tanner, J.M., T. Hayashi, M.A. Preece y N. Cameron (1982), "Increase in lengh of leg relative to trunk in Japanese children and adults from 1957 to 1977: comparison with the British and with Japanese Americans", *Ann. Hum. Biol.*, 9, 411-423.

Thissen, D., R.D. Bock, H. Wainer y A.F. Roche (1976), "Individual growth in stature: a comparison of four growth studies in the U.S.A.", *Ann. Hum. Biol.*, 3, 529.

Zacharias, L. y W.M. Rand (1983), "Adolescent growth in height and its relation to menarche in contemporary american girls", *Ann. Hum. Biol.*, 10, 209-222.

Zacharias, L. y W.M. Rand (1986), "Adolescent growth in height and its relation to menarche in contemporary american girls", *Ann. Hum. Biol.*, **13**, 369-386.

Zerbe, G.O. (1979), "A new non parametric technique for constructing the percentiles and normal ranges for growth curves determined from longitudinal data", *Growth*, **43**, 263.

APENDICE 1) CONSTRUCCION DE LOS MODELOS DE PREECE-BAINES

Como se ha mostrado en varios estudios (Marubini et al. 1971, Marubini 1978, Preece 1978, Hauspie et al. 1980, Preece y Heinrich 1981, Tanner et al., 1982), la función logística proporciona, para la adolescencia y post adolescencia, mejores ajustes en términos de varianza residual que muchas otras funciones. Sin embargo, esta función no es un buen modelo para la curva de crecimiento completa principalmente debido a que su derivada es simétrica, (sección 1.2) lo que la hace ser inadecuada para modelar la velocidad de crecimiento en la infancia y pre adolescencia, ya que en estos periodos la velocidad debe ser mayor que en la adolescencia y post adolescencia.

En vista de lo anterior y de que un problema central para modelar la curva completa de crecimiento está localizado en la adolescencia, Preece y Baines parten de la función logística para construir su familia de modelos.

La forma usual de escribir una logística es:

$$h\ (t) = P + K\ (1 + exp(a - bt))^{-1} \tag{A1}$$

donde P y P + K son las asíntotas superior e inferior, respectivamente, y a/b es el instante en que la velocidad h' es máxima, esto es, el eje de simetría de h. La ecuación (A1) puede expresarse como:

$$h\ (t) = \frac{h_0\ exp\ (\gamma h_0\ (t-\theta)) + h_1\ exp\ (\gamma h_1\ (t-\theta))}{exp\ (\gamma h_0\ (t-\theta)) + exp\ (\gamma h_1\ (t-\theta))} \tag{A2}$$

siendo $h_o = P$, $h_1 = P + K$, $\gamma = b/K$ y $\theta = a/b$. La interpretación de esta reparametrización es obvia en todos los componentes, excepto para γ, que puede verse como una constante proporcional a $h'(\theta) = \gamma (h_1 - h_o)^2/4$. Preece y Baines basan su desarrollo en el hecho de que la ecuación (A2) es solución de la ecuación diferencial

$$dh/dt = \gamma (h - h_o) (h_1 - h) \tag{A3}$$

siendo h_o y h_1 las asintotas de h en esta ecuación. Si t está cerca de 0, entonces, por la simetría de la derivada de una logística, h'(t) toma valores parecidos a los que tiene cuando t es grande.

Esto lleva a pensar en modelar dh/dt sustituyendo $\gamma (h - h_o)$ por alguna función que incremente la parte derecha de (A3) en los valores bajos de t.

La ecuación (A3) puede escribirse como:

$$dh/dt = s(h) (h_1 - h),$$

y, entonces,

$$s(h) = \gamma(h - h_o) = dh/dt (h_1 - h)^{-1} \tag{A4}$$

es decir, s(h) es una logística con asintotas 0 y b.

Esto implica que dh/dt = 0 cuando t = 0, lo que es inadmisible, pues la mayor velocidad de crecimiento ocurre en los primeros meses de vida.

Con el fin de resolver tal problema, Preece y Baines trabajaron con (A4) usando datos reales para evaluar numéricamente dh/dt mediante incrementos observados y conociendo la estatura final h_1. De esta manera encontraron que s(h) se comportaba como una curva sigmoidal con asíntota inferior mayor que 0, como se esperaría. Este resultado sugirió que s podría modelarse como una (otra) logística y así, considerando a s en función del tiempo (en vez de la estatura), y utilizando a (A3), se tiene el siguiente sistema de ecuaciones diferenciales:

$$ds/dt = \gamma\ (s_1 - s)\ (s - s_0) \qquad (A5)$$

$$dh/dt = s(t)\ (h_1 - h) \qquad (A6)$$

Para resolver el sistema, se encuentra una solución de la primera ecuación y posteriormente se sustituye en (A6) con el fin de obtener h.

La ecuación (A5) puede expresarse como:

$$\int ds/(s_1 - s)\ (s - s_0) = -\ (t + C)\ \gamma$$

siendo C una constante de integración.

Entonces:

$$(t + C)\ \gamma = (s_1 - s_0)^{-1}\ ln\ |(s - s_1)(s_1 - s)|$$

$$\Rightarrow ln\ |(s - s_0)/(s_1 - s)| = ln\ (s - s_1)(s_1 - s),$$

debido a que la logística satisface $s \geq s_0$ y $s_1 \geq s_1$.

Así:

$$s = \frac{s_1 + s_0 \exp[-\gamma (t - C)(s_1 - s_0)]}{1 + \exp[-\gamma (t - C)(s_1 - s_0)]} .$$

Sea $C = -\theta$; entonces,

$$s(t) = \frac{s_1 \exp(\gamma(t-\theta)) + s_0 \exp(\gamma(t-\theta))}{\exp(\gamma(t-\theta)s_1) + \exp(\gamma(t-\theta)s_0)} \qquad (A7)$$

satisface (A5).

Por otra parte, de la ecuación (A6) se tiene que:

$$\int (h_1 - h)^{-1} dh = \int s(t)\, dt .$$

Integrando ambos lados:

$$-\ln (h_1 - h) = \ln ([\exp(\gamma(t-\theta)s_1) + \exp(\gamma(t-\theta)s_0)]^{1/\gamma} + \ln C_1,$$

siendo C_1 una constante de integración..

Si se hacen $t = 0$ y $h(\theta) = h_0$, se ve que:

$$-\ln(h_1 - h_0) = \ln(2^{1/\gamma}) + \ln C_1$$

$$\Rightarrow \qquad C_1 = [2^{1/\gamma} (h_1 - h_\theta)]^{-1},$$

$$-\ln(h_1 - h) - \ln C_1 = \ln ([\exp(\gamma(t-\theta)s_0) + \exp(\gamma(t-\theta)s_1)]^{1/\gamma}),$$

y, finalmente,

$$h = h_1 - \frac{2^{1/\gamma}(h_1 - h_\theta)}{[exp(\gamma(t-\theta)s_0) + exp(\gamma(t-\theta)s_1)]^{1/\gamma}} \quad (A8)$$

La ecuación (A8) es conocida como modelo 2; los parámetros s_0 y s_1 son las asíntotas de la logística s y están medidas en unidades de $(\text{tiempo})^{-1}$; θ es un parámetro en el tiempo y representa el momento en que s alcanza su máxima velocidad; h_θ es el valor de la estatura en el momento θ; h_1 es la talla final y γ es una constante adimensional.

Si en (A7) se hace $\gamma = 1$, se obtiene el modelo 1:

$$h = h_1 - \frac{2\,(h_1 - h_\theta)}{exp((t-\theta)s_0) + exp((t-\theta)s_1)} \quad (A9)$$

Preece y Baines (1978:22) hacen notar que si γ toma valores cercanos a 1 en el modelo 2, h se ajusta bien para valores grandes de t, y "haciendo $\gamma = 1$, de hecho se tiene un muy buen modelo, independientemente de su conexión con el modelo 2". Además, el modelo 1 tiene sólo 5 parámetros, uno menos que el modelo 2.

La velocidad de crecimiento para h en la ecuación (A7) es:

$$dh/dt = s(t)\,(h_1 - h) = \frac{2^{1/\gamma}(h_1 - h_\theta)\,(s_0e_0+s_1e_1)}{(e_0+e_1)\,(e_0+e_1)^{1/\gamma}} \quad (A10)$$

donde $e_0 = exp(\gamma(t-\theta)s_0)$ y $e_1 = exp(\gamma(t-\theta)s_1)$; la aceleración se obtiene con la regla del producto:

$$d^2h/dt^2 = (h_1 - h)\ ds/dt - s\ dh/dt \tag{A11}$$
$$= (h_1 - h)\ [s^{-2}\ (1 + \gamma) + s\ (s_1 + s_0) - s_1 s_0].$$

Las velocidades máxima y mínima se obtendrían analíticamente igualando la ecuación (A11) a 0 y resolviendo para s (eliminando la solución trivial $h_1 = h$); las soluciones de (A11) son de la forma:

$$\frac{s_1 + s_0 \pm \sqrt{(s_1+s_0)^2 - 4(1-\gamma)s_1 s_0}}{2\ (1 + \gamma)} \tag{A12}$$

Para obtener analíticamente las edades de velocidades máxima y mínima bastaría sustiuir las soluciones de la forma (A12) en la ecuación (A7). A partir de estas soluciones se estimarían otros parámetros biológicos, como los que fueron mencionados en el capítulo 1.

Una modificación a los modelos 1 y 2 consiste en utilizar una función con primera derivada no simétrica en el lugar de la logística s que aparece en la ecuación (A7). La posibilidad que consideraron Preece y Baines es la producida con la suma de dos logísticas. La idea de trabajar con tal clase de funciones está expuesta en Bock (1973) y Bock y Thissen (1980). Estos autores utilizan sumas de dos y tres logísticas considerando que cada una de ellas modela una parte de la curva completa de crecimiento; sin embargo, cabe hacer notar que la idea de Preece y Baines no es la de asumir dos o tres componentes aditivas que probablemente no tienen un sentido biológico claro. En mi opinión, sólo intentan dar una forma asimétrica a la función s buscando mejorar el ajuste: así como eligieron una suma de logísticas, bien podrían haber considerado otra función.

Todo esto reforzaría analíticamente la parte más débil de la construcción del modelo, misma que debió sostenerse únicamente con la evidencia empírica, lo que, por otra parte, es totalmente válido, pues el patrón de comportamiento observado por Preece y Baines para los incrementos de la talla (gráfica A2) puede considerarse universal. De cualquier forma, a continuación se expone la construcción de otros dos modelos de Preece-Baines que involucran una s no simétrica.

Si se hace $\gamma = 1$ y s (en la ecuación A7) se considera como la suma de dos logísticas con diferentes parámetros ($s = p + q$), se tiene que:

$$dp/dt = (p_1 - p)(p - p_0) \qquad \text{y} \qquad dq/dt = (q_1 - q)(q - q_0).$$

Un desarrollo análogo al anterior conduce a las siguientes soluciones:

$$p(t) = \frac{p_0 \, exp((t-\theta_1)p_0) + p_1 \, exp((t-\theta_1)p_1)}{exp((t-\theta_1)p_0) + exp((t-\theta_1)p_1)}$$

y

$$q(t) = \frac{q_0 \, exp((t-\theta_2)q_0) + q_1 \, exp((t-\theta_2)q_1)}{exp((t-\theta_1)p_0) + exp((t-\theta_1)p_1)}.$$

Ahora, es necesario resolver

$$dh/dt = (p + q)(h_1 - h)$$

para h; es decir:

$$\int dh/(h_1 - h) = \int [p(t) + q(t)]\, dt \qquad (A13)$$

$$\rightarrow - \ln (h_1 - h) = \ln [exp((t-\theta_1)p_0) + exp((t-\theta_1)p_1)] + \ln [exp((t-\theta_2)q_0) + exp((t-\theta_2)q_1)] + \ln C, \qquad (A14)$$

donde C es una constante de integración.

El siguiente paso sería, de manera análoga al procedimiento seguido a partir de (A7), imponer una condición para la ecuación (A14): por ejemplo, $h(\theta_1) = h(\theta_2)$ y encontrar un valor adecuado para C. Como esto no es posible, Preece y Baines suponen, advirtiendo que ésto no tiene estrictamente que ser así, que

$$\theta_1 = \theta_2 = \theta$$

desde antes de la ecuación (A13), con lo que tienen una sola constante de integración, y, resolviendo la ecuación análoga a (A14), obtienen el modelo 3:

$$h = h_1 - \frac{4\,(h_1 - h_\theta)}{[exp((t-\theta)p_0)+exp(t-\theta)p_1)]\,[exp((t-\theta)q_0)+exp((t-\theta)q_1)]} \qquad (A15)$$

Al llegar a este punto, Preece y Baines hacen notar que uno de los dos nuevos parámetros (por ejemplo q_0) podría hacerse 0 "sin afectar la generalidad del modelo" (p.23). Esto lleva a la siguiente forma del modelo 3:

$$h = h_1 - \frac{4\,(h_1 - h_\theta)}{[exp((t-\theta)p_0) + exp(t-\theta)p_1)]\,[1 + exp((t-\theta)q_1)]} .$$

Nótese que si $p_0 = p_1$ se recupera el modelo 1. Por otra parte, con el modelo 3 no es posible obtener las soluciones para las velocidades máxima y mínima analíticamente, puesto que las raíces de

$$\frac{p_0 exp(p_0(t-\theta)) + p_1 exp(p_1(t-\theta))}{exp(p_0(t-\theta)) + exp(p_1(t-\theta))} + \frac{q_0 exp(q_0(t-\theta)) + q_1 exp(q_1(t-\theta))}{exp(q_0(t-\theta)) + exp(q_1(t-\theta))}$$

no admiten una expresión analítica.

Como se mencionó en el capítulo 1, el modelo 1 presenta ventajas sobre los otros miembros de la familia, por lo que fue el único con el que se realizaron los ajustes en este trabajo. Después de haber revisado la literatura disponible acerca de modelos de crecimiento, creo que no es posible más que estar de acuerdo con Preece y Baines (1978: 10) cuando afirman que el propósito de su trabajo fue "desarrollar modelos robustos que fueran fácilmente ajustados, más que *primas donnas* que requieren de grandes cuidados para cada ajuste individual".

APENDICE 2) ESTIMACIONES CON DATOS TRANSVERSALES

Aunque el modelo PB1 fue desarrollado originalmente para ajustar observaciones longitudinales, ha sido utilizado para analizar datos provenientes de estudios transversales. Ejemplos de lo anterior aparecen en los trabajos de Jordán (1979), aunque sólo aparecen las curvas ajustadas, sin hacer referencia alguna a los parámetros y de Tanner *et al.* (1982), en el que se ajusta el modelo a las medias para grupos de edad anuales obtenidas en las mediciones de escolares llevadas a cabo en Japón en 1957, 1967 y 1977. El objetivo de este último trabajo es analizar el aumento secular en medidas tales como la longitud del segmento superior y la estatura total. Estos autores encuentran que el modelo es una valiosa herramienta para el análisis de estudios transversales, pues permite estimar la estatura adulta promedio que alcanzaría la población, así como la edad en la que el incremento en estatura promedio de la población es máximo. Es interesante hacer notar que el equivalente de este parámetro biológico en el caso de estudios longitudinales para un solo individuo es la edad en la que la velocidad de crecimiento es máxima.

En el segundo trabajo mencionado en el párrafo anterior, los ajustes se efectúan por MCO, sin que haya referencia alguna a la calidad de la estimación (por ejemplo, varianza residual, sesgos, determinación de los estimadores, etc.) ni a la posibilidad de incorporar en la estimación la información contenida en el conocimiento de los tamaños de muestra y las desviaciones estándares para cada grupo de edad.

Es posible distinguir dos clases de suavizamiento efectuado sobre los datos cuando se ajusta un modelo utilizando las medias de los grupos de edad : uno producido al redondear edades que no tienen la magnitud del centro del grupo y otro por concentrar toda la información contenida en los datos únicamente en la estimación de la media.

El primero es relativamente inevitable. En realidad, cualquier edad que se considere en el análisis será una aproximación más o menos exacta a la edad real del individuo en el momento de la medición. Al parecer, el redondear en grupos de edades anuales o semestrales es suficientemente preciso; la gran mayoría de los trabajos consultados así lo hacen tanto en el análisis propiamente dicho como en la presentación de resultados.

En cuanto al segundo punto, es claro que la estimación de la media del grupo de edad estará sujeta a errores de muestreo; de aquí la importancia de incluir tales errores en la estimación de los parámetros de PB1. La forma más simple de realizar ésto consiste en utilizar los errores estándares de la media de cada grupo de edad. Estas estadísticas son una medida de la precisión obtenida en la estimación de la media e involucran a la desviación estándar de la las estaturas y el tamaño de muestra conseguido. Se espera que, aproximadamente, una de cada 20 estimaciones de la media distaría del verdadero valor del parámetro en más de dos veces el error estándar, o bien, que una de cada 100 lo haría en más de tres veces el error estándar.

En cuanto a la posibilidad de involucrar la información acerca de la variabilidad de las estaturas dentro del grupo de edad y de los tamaños de muestra, es fácil ver que una posible solución consiste en estimar con mínimos cuadrados generalizados utilizando como matriz de varianza covarianza a una matriz diagonal con los errores estándar de las medias de cada grupo de edad. El uso de esta matriz es plausible, puesto que se espera que los errores sean independientes en un estudio transversal y que las ponderaciones necesarias para los residuales varíen según los grupos de edad.

Bajo este supuesto, es fácil obtener los estimadores correspondientes con los métodos presentados en las secciones 2.2 y 2.3.

En el transcurso del trabajo se analizaron los datos provenientes de los trabajos de Jordán (1979) y Tanner *et al.* (1982) estimando los parámetros de PB1 con MCO y con MCG utilizando la matriz diagonal ya mencionada.

Todos los ajustes terminaron en términos numéricos muy buenos; por ejemplo, la norma del gradiente de la función objetivo fue orden de 10^{-4} en promedio y el ajuste se obtuvo en 7 iteraciones, en promedio. Los resultados de las pruebas de rachas y de Durbin Watson indicaron aleatoriedad para los residuales, las varianzas fueron similares a las obtenidas en los ajustes longitudinales, casi todos los sesgos representaron menos del 1% del respectivo componente del estimador y fue posible obtener estimadores para todos los parámetros biológicos derivados de la estimación del modelo.

En cuanto a las diferencias entre las dos clases de ajuste, no es posible emitir un juicio decisivo. Al igual que en los casos longitudinales analizados, los valores de los estimadores obtenidos por ambos métodos fueron muy similares. Sin embargo, no es posible generalizar lo aquí observado. Para tal efecto, una alternativa está en realizar estudios con métodos Monte Carlo, considerando diferentes errores estándares y diferentes grupos de edades.

Por último, cabe señalar que cuando se ajustó el modelo para edades menores que 2 años, como se hizo con los datos de Jordán, disminuyó la calidad de la estimación. Para los primeros meses de vida se tuvo una sobreestimación notable, fenómeno que se repitió entre los 3 y los 8 años, para ambos sexos. Al utilizar datos referentes a edades mayores que 3 años se tuvo una diferencia significativa en los valores de los parámetros respecto a las estimaciones anteriores; sin embargo, el ajuste fue mucho mejor, en términos de varianza residual, sesgos y aleatoriedad de los residuales.