UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

" INSTRUMENTOS ESTADISTICOS PARA EL ANALISIS ECONOMICO "

T E S I S

QUE PARA OBTENER EL TITULO DE

A C T U A R I O

P R E S E N T A :

ANA MARIA CRUZ RAMOS

MEXICO, D. F. 1986

INSTRUMENTOS ESTADISTICOS PARA EL ANALISIS ECONOMICO.

CONTENIDO

I N T R O D U C C I O N.

En estas páginas se busca dar una breve visión de métodos estadísticos de investigación en economía, presentando sus fundamentos y analizando sus distintas etapas. Con ello se persigue justificar la necesidad que tiene el Actuario de conocerlos.

Los métodos analizados aquí se dirigen a la construcción, desarrollo y análisis de modelos matemático-estadísticos para estudiar el comportamiento de fenómenos económicos. En esta labor se -- combinan dos fuentes de información:

a) La teoría, que conceptualiza y formula hipótesis acerca de los fenómenos económicos, y
b) Los datos, que representan mediciones de los hechos -- económicos.

La combinación de estas dos fuentes de forma adecuada culmina en un modelo teórico-empírico acerca de los fenómenos económicos en estudio.

El desarrollo del modelo provee una base para la contrastación empírica de hipótesis teóricas, y para el análisis de la estructura económica, así como una herramienta útil para fines prácticos como la predicción en materias de planificación económica (un ejemplo de ello puede verse en el Capítulo 5).

Los métodos desarrollados aquí llevan un orden lógico y, como se verá en su oportunidad, se retroalimentan para obtener distintos instrumentos.

En el Capítulo 1, se estudian métodos determinísticos, en los cuales, como es bien sabido, solamente tomamos en cuenta los datos tal y como se nos presentan en un momento dado.

En el Capítulo 2, se estudian métodos que involucran términos aleatorios (de ahí que sean conocidos como modelos estocásticos) y la forma de analizarlos.

El Capítulo 3 está estrechamente relacionado con el 2 y, su diferencia básica es que ahora no nos referiremos a un momento fijo en el tiempo, sino que serán una serie de momentos temporales sucesivos.

El Capítulo 4 se aboca a problemas bien específicos y tal vez los más conocidos de la teoría económica presentando algunas de las diferentes formas en que han sido investigados.

Finalmente, como ya se mencionó, el Capítulo 5 muestra - una forma de planeación sobre una actividad reconocidamente importante como lo es el " proceso de producción", haciéndose un ensayo de evaluación de proyectos para una empresa privada.

En el apéndice se han anexado algunas explicaciones puramente teóricas ya sea a manera de repaso de algunos conceptos o para que se tenga una forma distinta de visualizarlos.

CAPITULO 1.

AJUSTE PARA MODELOS DETERMINISTICOS.

1.0 MODELOS DETERMINISTICOS.

Un modelo es una representación abstracta de la realidad que pone de manifiesto lo más relevante de una cuestión particular dejando aparte todos los problemas restantes.

En la mayoría de las ciencias sociales, un modelo no es más que una idea del concepto que se tenga de la realidad; pero - esto varía en los modelos económicos debido a que pueden definirse por medio de variables y funciones perfectamente determinadas, y sobre las cuales puede tenerse información que servirá para hacer inferencias que a su vez mejoren, cambien o proyecten diferentes situaciones. Es por ello que los modelos económicos tienen su base en la economía pero sus métodos son completamente estadísticos.

El estudio estadístico de las relaciones entre diversas variables presenta los siguientes aspectos fundamentales a investigar: existencia de asociación entre las variables, dirección de la asociación, grado de asociación y, naturaleza y forma de la -- asociación (que es la base del desarrollo de este trabajo).

Los elementos fundamentales del análisis de regresión - son las variables y la ecuación de regresión.

Las variables pueden ser 2 o más, siendo una de ellas - dependiente del resto, que son independientes entre sí; por lo que la variable dependiente es aquella cuyo cambio se considera en función o dependencia del cambio o variación de la (s) variable (s) - independientes (s).

En este capítulo consideraremos que los valores que toman las variables son determinados (es decir, valores fijos que no presentan aleatoriedad alguna) por lo que los modelos reciben el nombre de determinísticos.

De las formas más usuales para encontrar la ecuación de regresión, se presentan a continuación:

i) Ajuste Gráfico.

ii) Regresión Lineal Simple.

iii) Regresión Ortogonal.

iv) Regresión Múltiple.

1.1 AJUSTE GRAFICO.

Cuando tenemos un conjunto de datos, y los representamos gráficamente, la simple observación visual de la nube de puntos - formada puede bastar para indicar genéricamente la existencia y el signo de la relación entre las variables.

Por supuesto, esta relación puede no ser obvia, pero la comparación con las gráficas del tipo de funciones más comúnmente utilizadas conducirá a conclusiones inmediatas sobre su comportamiento, existencia o no de alguna relación entre ellas, dirección de la relación (dada por el signo de los parámetros) y la forma genérica en la que estén asociadas.

Las funciones más consideradas para un ajuste de esta - naturaleza son:

lineal $y = \alpha + \beta x$

polinomial $y = \alpha + \beta_1 x + \beta_2 x^2$ (2° grado)

hipérbola rectangular $y = \beta / x$ ó $yx = \beta = cte$

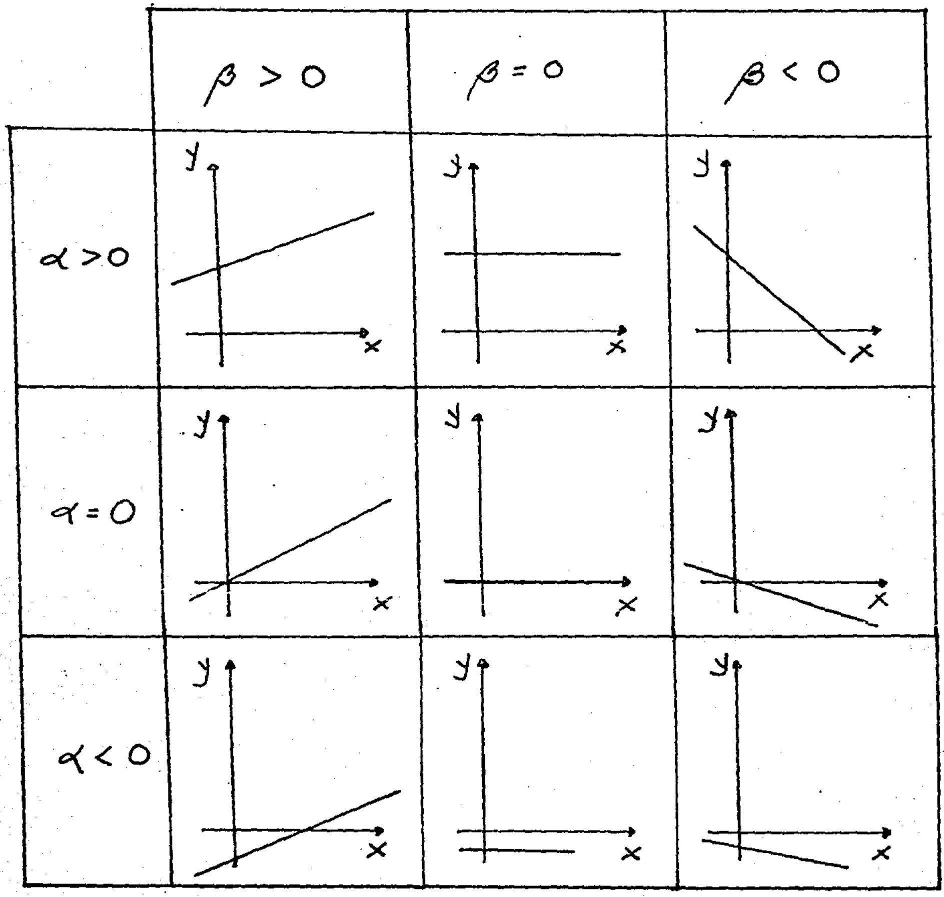
exponencial $y = \beta^x$ ó $\log y = x \log \beta$

logarítmica $y = \beta \log x$

cuyas gráficas se muestran en las siguientes páginas.

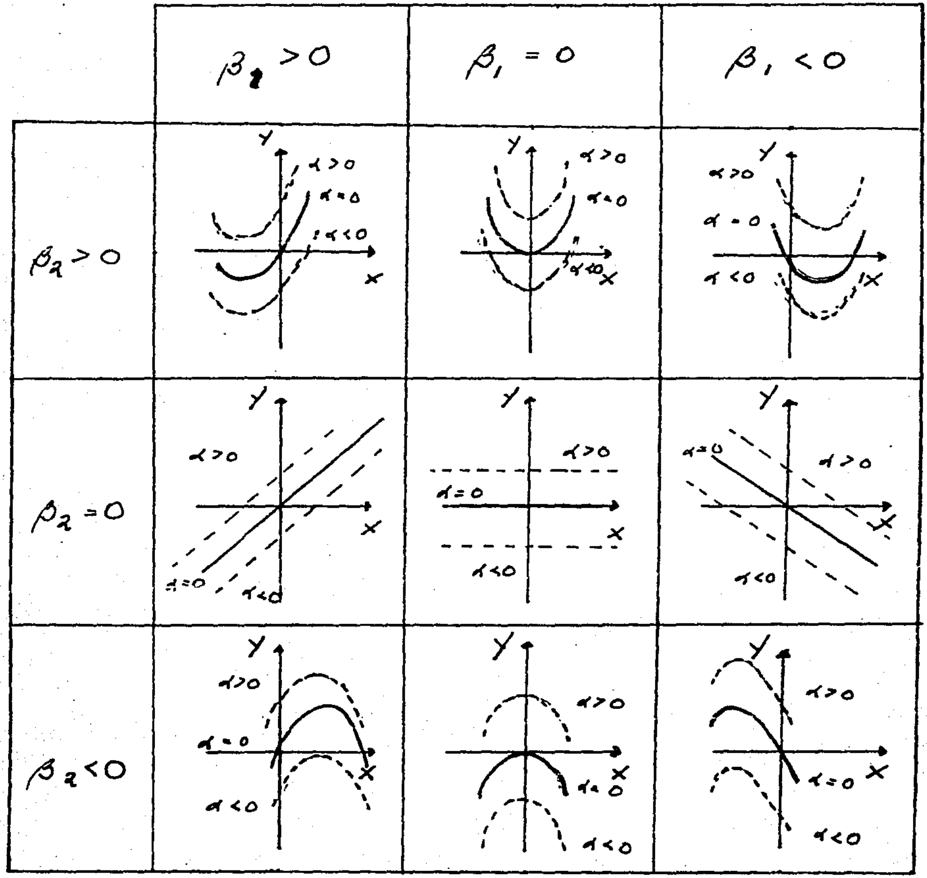
Modelo Lineal

$$y = \alpha + \beta x$$

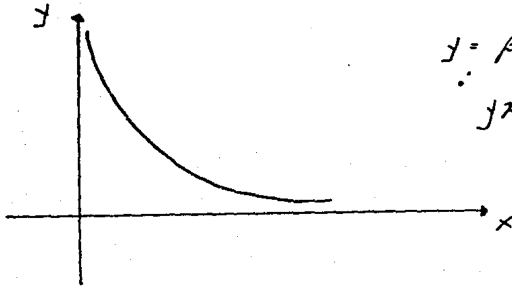


Modelo Polinomial (2º grado)

$$y = \alpha + \beta_1 x + \beta_2 x^2$$

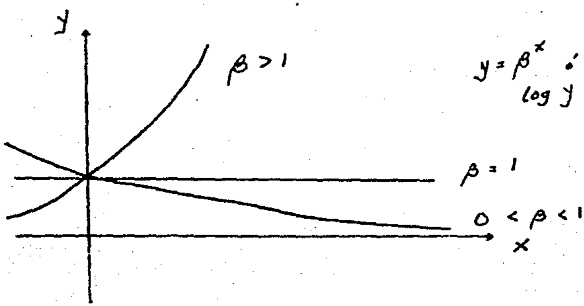


Modelo Hipérbola Rectangular.



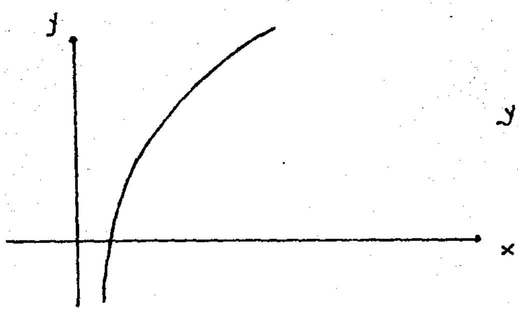
$$y = \beta / x \quad \therefore \quad yx = \beta = cte.$$

Modelo Exponencial.



$$y = \beta^x \quad \therefore \quad \log y = x \log \beta.$$

Modelo Logarítmico

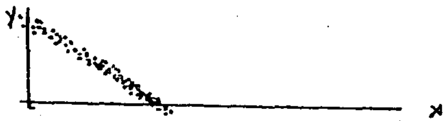


$$y = \beta \log x$$

1.2 REGRESION LINEAL SIMPLE

Este tipo de análisis es el que sirve para analizar la naturaleza y forma de asociación entre dos variables siempre y cuando dicha relación pueda ser expresada matemáticamente por la ecuación de una línea recta.

En el caso donde se nos presente un conjunto de datos en el cual, al graficar sus logaritmos se obtenga algo como:



la figura nos indica la existencia de algún tipo de dependencia entre los valores de x y de y.

Si nos concretáramos a saber lo que ocurriría tomando un simple ajuste de curvas, podríamos encontrar la relación entre las dos variables sin ninguna dificultad, solo que entonces nos enfrentaríamos al problema de deducir lo más exactamente posible la dependencia inminente entre ellas; y a la vez, poder generalizar, de alguna manera, todo lo referente a su comportamiento.

Para empezar, nunca debe perderse de vista que dicha relación sea expresada en la forma más sencilla y comprensible que se pueda.

En una situación como la planteada, podemos darnos cuenta de que es suficiente si buscamos una relación lineal, que se representará como una recta que pase a través del conjunto de puntos dado.

Por supuesto, lo que se espera es que la distancia entre los puntos y la recta sea lo más reducida posible; de ahí que se busque una recta tal que la suma de los cuadrados de las distancias entre ella y los puntos sea mínima.

Caso 1) Distancia a lo largo de una paralela al eje Y.

Una de las ecuaciones más conocidas para describirla es: $y = ax + b$. Solo que no puede considerarse aislada de las condiciones que le impusimos desde un principio.

Como tenemos dos magnitudes (x e y), con T observaciones, los puntos serán:

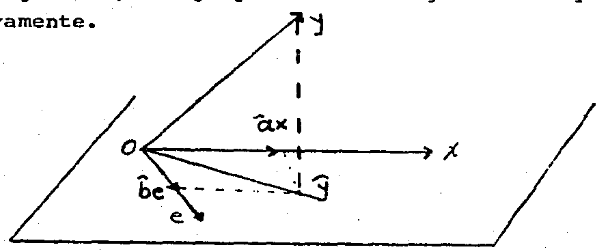
$$(x_1, y_1),\ (x_2, y_2),\ \ldots,\ (x_T, y_T)$$

i. e. .- $(x_t, y_t) \qquad t = 1, 2, \ldots, T$

necesitaremos entonces que: $(y_1 - ax_1 - b)^2 + (y_2 - ax_2 - b)^2 + \cdots + (y_T - ax_T - b)^2 = U$

sea mínima, en general: $U = \sum_{t=1}^{T} (y_t - ax_t - b)^2$

Geométricamente, sobre el espacio de las variables, $\hat{a}$ y $\hat{b}$ se interpretan respectivamente como la ordenada al orígen y la pendiente de la recta que hacen mínima la suma de los cuadrados de las distancias verticales de los puntos del orígen a los mismos; esta línea se llama "recta de regresión de y sobre x". Sobre el espacio de observaciones, $\hat{a}$ y $\hat{b}$ se interpretan como las componentes en x y e respectivamente de la proyección ortogonal de y sobre el subespacio de $\mathbb{R}^n$ engendrado por x y e; x y e no son colineales pero, en general, no son ortogonales, lo cual implica que $\hat{b}e$ y $\hat{a}x$ no sean, en general, las proyecciones ortogonales de y sobre x y e respectivamente.



Para obtener la solución al sistema $\sum_{t=1}^{T}(y_t - ax_t - b)^2$, tomemos en cuenta que f(x)= y = ax + b es una función derivable, por lo que bastará con encontrar los puntos para los cuáles el valor de la derivada sea 0.

Con objeto de derivar más fácilmente consideremos que para cada x , y:

$$(y_i - ax_i - b)^2 = y_i^2 - 2ax_i y_i - 2by_i + a^2x_i^2 + 2ax_i b + b^2$$

$$\frac{\partial f}{\partial a} = -2x_i y_i + 2ax_i^2 + 2x_i b = 0 = x_i y_i - ax_i^2 - x_i b$$

$$\frac{\partial f}{\partial b} = -2y_i + 2ax_i + 2b = 0 = y_i - ax_i - b$$

tomando ahora todas las restricciones será:

$$\sum_{t=1}^{T} x_t y_t - a\sum_{t=1}^{T} x_t^2 - b\sum_{t=1}^{T} x_t = 0 \quad \ldots\ldots\ldots\ldots \quad (1)$$

$$\sum_{t=1}^{T} y_t - a\sum_{t=1}^{T} x_t - \sum_{t=1}^{T} b = 0 \quad \ldots\ldots\ldots\ldots \quad (2)$$

(2) también puede verse como:

$$Tb = \sum_{t=1}^{T} y_t - a\sum_{t=1}^{T} x_t$$

y: $\bar{x} = \dfrac{\sum_{t=1}^{T} x_t}{T}$ $\qquad \bar{y} = \dfrac{\sum_{t=1}^{T} y_t}{T}$

por lo que, al despejar b se obtiene:

$$b = \bar{y} - a\bar{x}$$

Esto indica que la recta de ajuste debe pasar por el punto medio- $(\bar{x}, \bar{y})$. El punto de las medias será entonces fijo en nuestro sistema y los puntos (x_t, y_t) serán móviles. Un vector fijo que los una (diferente para cada uno) será de la forma li + mj, entonces, para cada (x_t, y_t):

$$(x_t - \bar{x})i + (y_t - \bar{y})j = t(li + mj)$$

$$(x_t - \bar{x}) = tl, \quad (y_t - \bar{y}) = tm$$

$$\frac{(x_t - \bar{x})}{l} = \frac{(y_t - \bar{y})}{m}$$

$$(y_t - \bar{y}) = \frac{m}{l}(x_t - \bar{x})$$

pero m/l define la pendiente de la recta lo mismo que a en la ecuación y= ax + b ; por lo tanto, a= m/l y podemos expresar nuestra ecuación como:

$$(y_t - \bar{y}) = a\,(x_t - \bar{x})$$

luego entonces, de (1):

$$\sum_{t=1}^{T} x_t y_t - a \sum_{t=1}^{T} (x_t)^2 - (\bar{y} - a\bar{x}) \sum_{t=1}^{T} x_t = 0$$

de ahí que:

$$a = \frac{\sum_{t=1}^{T} (y_t - \bar{y})(x_t - \bar{x})}{\sum_{t=1}^{T} (x_t - \bar{x})^2}$$

Tenemos ahora una manera de determinar la recta de ajuste partiendo de los valores de los datos. Como habrá puntos dispersos alrededor de la recta, si se substituyen los valores de a y b se tiene:

$$U = \sum_{t=1}^{T} (y_t - \bar{y})^2 - \frac{\left[\sum_{t=1}^{T} (y_t - \bar{y})(x_t - \bar{x})\right]^2}{\sum_{t=1}^{T} (x_t - \bar{x})^2}$$

aunque regularmente esta dispersión se indica como:

$$R^2 = 1 - \frac{U}{\sum_{t=1}^{T}(y_t-\bar{y})^2} \quad \text{ó} \quad R^2 = 1 - \frac{\sum_{t=1}^{T}(y_t-\bar{y})^2 - \frac{\left[\sum_{t=1}^{T}(y_t-\bar{y})(x_t-\bar{x})\right]^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}}{\sum_{t=1}^{T}(y_t-\bar{y})^2}$$

(donde $-1 \leq R^2 \leq 1$)

y se tiene entonces la manera de saber la existencia o no de variación conjunta entre las dos variables, según sea o no distinto de cero; la dirección de la ecuación, por su signo positivo o negativo y el grado de la covariación según sea el mayor o menor valor que alcance entre cero y más y menos uno.

En forma general, el parámetro b representa el punto en el que la recta de regresión corta al eje de las Y's e indica el valor que corresponde entonces a y cuando el valor de x es cero - según la ecuación. El parámetro a es la pendiente de la recta de regresión y señala la tasa de cambio en y por unidad de cambio en x.

Caso 2) Distancia a lo largo de una paralela al eje X. Podemos representar la ecuación por x= ay + b.
Con los puntos:

$$(x_t, y_t) \qquad t = 1, 2, \ldots, T$$

y sujeta a:

$$\sum_{t=1}^{T} (x_t - ay_t) - b)^2 \quad \text{sea mínima.}$$

entonces:

$$(x_t - ay_t - b)^2 = x_t^2 - 2ax_ty_t - 2bx_t + a^2y_t^2 + 2ay_tb + b^2$$

$$\frac{\partial f}{\partial a} = x_ty_t - ay_t^2 - y_tb = 0 \qquad \frac{\partial f}{\partial b} = x_t - ay_t - b = 0$$

al considerar todas las restricciones:

$$\sum_{t=1}^{T} x_t y_t - a \sum_{t=1}^{T} y_t^2 - b \sum_{t=1}^{T} y_t = 0 \quad \ldots\ldots \quad (3)$$

$$\sum_{t=1}^{T} x_t - a \sum_{t=1}^{T} y_t - \sum_{t=1}^{T} b = 0 \quad \ldots\ldots \quad (4)$$

de (4): $$Tb = \sum_{t=1}^{T} x_t - a \sum_{t=1}^{T} y_t$$

por lo que: $$b = \bar{x} - a\bar{y}$$

igual que la anterior, esta recta también pasa por el punto de las medias $(\bar{x}, \bar{y})$.

Podemos escribir la ecuación como:

$$(x_t - \bar{x}) = a\,(y_t - \bar{y})$$

y, de (3)

$$a = \frac{\sum_{t=1}^{T} (x_t - \bar{x})(y_t - \bar{y})}{\sum_{t=1}^{T} (y_t - \bar{y})^2}$$

la pendiente es ahora diferente a la anterior, además de:

$$U = \sum_{t=1}^{T} (x_t - \bar{x})^2 - \frac{\left[\sum_{t=1}^{T} (x_t - \bar{x})(y_t - \bar{y})\right]^2}{\sum_{t=1}^{T} (y_t - \bar{y})}$$

La relación entre las pendientes de las dos regresiones simples es igual al cuadrado del coeficiente de correlación lineal R^2.

1.3 REGRESION ORTOGONAL.

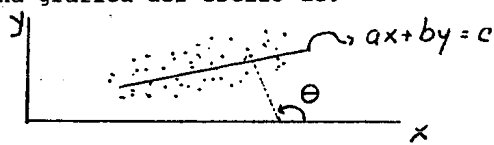
Este tipo de regresión consiste en tomar en cuenta la - distancia que existe entre la recta de ajuste y los puntos de la gráfica.

Otra vez, lo que se busca es una expresión presentada lo más sencillamente posible que refleje esta situación de una manera general.

Consideremos ahora la ecuación de una recta en 2 dimensiones de la forma:

$$ax + by = c$$

y tendremos una gráfica del estilo de:



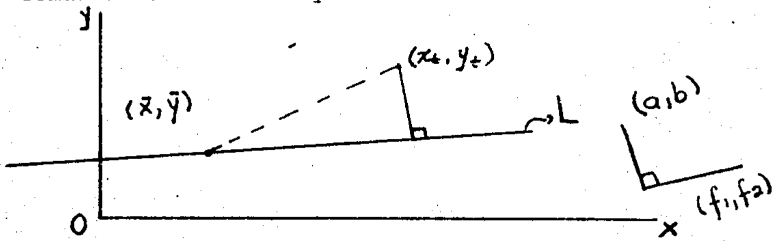
el ángulo que forma el vector ortogonal a la recta ax + by = c con el eje X nos dá los valores:

$$a = \cos\theta \qquad\qquad b = \operatorname{sen}\theta$$

$$c = \text{distancia de la recta al orígen.}$$

y, como $\cos^2\theta + \operatorname{sen}^2\theta = 1$, además de restringir a distancias -- mínimas, tendremos que $a^2 + b^2 = 1$.

Tomando en cuenta lo que está sucediendo será:



la componente principal (como conjunto de dos magnitudes) estará dada por:

$$v_t = f_1 (x_t - \bar{x}) + f_2 (y_t - \bar{y}) \qquad \textit{para cada } t.$$

en donde se cumple que $f_1^2 + f_2^2 = 1$ (la que tiene la mayor media cuadrática del conjunto de observaciones).

Para determinar la componente principal del sistema, necesitamos encontrar una recta L que pase por el punto de las medias $(\bar{x}, \bar{y})$, y hace máxima la suma v de los cuadrados de las proyecciones v_t sobre L de los vectores de componentes $(x_t - \bar{x})$ y $(y_t - \bar{y})$. (i. e. $V = \sum_{t=1}^{T} v_t = \sum_{t=1}^{T} f_1 (x_t - \bar{x}) + f_2 (y_t - \bar{y})$).

Sea μ_t la distancia de (x_t, y_t) a la recta L, al sumar $v_t^2 + \mu_t^2$ y sustituír los valores se tendrá:

$$(x_t - \bar{x})^2 + (y_t - \bar{y})^2$$

que es independiente de la posición de L. Esto indica que la recta L que hace máxima V es la misma que hace mínima la suma de U de los cuadrados de las distancias μ_t: por lo que esta es la regresión ortogonal y los coeficientes de la recta buscada deben ser $f_1 = b$ y $f_2 = -a$.

Todo esto quiere decir que:

Como $f(x, y)$ = ax + by - c , tenemos una función derivable y, de acuerdo con la teoría de los multiplicadores de Lagrange:

$$U = \sum_{t=1}^{T} (a x_t + b y_t - c)^2 - \lambda (a^2 + b^2 - 1)$$

$$\frac{\partial(U)}{\partial(a)} = 2 \sum_{t=1}^{T} (a x_t + b y_t - c) x_t - 2 \lambda a = 0 \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; (1)$$

$$\frac{\partial(U)}{\partial(b)} = 2 \sum_{t=1}^{T} (a x_t + b y_t - c) - 2 \lambda b = 0 \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; (2)$$

$$\frac{\partial(U)}{\partial(c)} = 2 \sum_{t=1}^{T} (a x_t + b y_t - c) = 0 \; . \; . \; . \; . \; . \; . \; . \; (3)$$

$$\text{De } (3): \quad \sum_{t=1}^{T} a x_t + \sum_{t=1}^{T} b y_t = Tc \quad \therefore \quad a\bar{x} + b\bar{y} = c$$

por lo que podemos escribir la ecuación de la recta como:

$$a\,(x_t - \bar{x}) + b\,(y_t - \bar{y}) = 0 \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Si consideramos ahora las ecuaciones (1) y (2), al substituír el valor de c se tiene:

$$a\left[\sum_{t=1}^{T}(x_t - \bar{x})^2 - \lambda\right] + b\sum_{t=1}^{T}(x_t - \bar{x})(y_t - \bar{y}) = 0$$

$$a\sum_{t=1}^{T}(x_t - \bar{x})(y_t - \bar{y}) + b\left[\sum_{t=1}^{T}(y_t - \bar{y})^2 - \lambda\right] = 0$$

si se dividen por el número total de puntos, T:

$$a\left[\frac{\sum_{t=1}^{T}(x_t - \bar{x})^2}{T} - \frac{\lambda}{T}\right] + b\,\frac{\sum_{t=1}^{T}(x_t - \bar{x})(y_t - \bar{y})}{T} = 0$$

$$a\,\frac{\sum_{t=1}^{T}(x_t - \bar{x})(y_t - \bar{y})}{T} + b\left[\frac{\sum_{t=1}^{T}(y_t - \bar{y})}{T} - \frac{\lambda}{T}\right] = 0$$

se obtienen los momentos centrales de 2° orden de x_t e y_t, denota dos por:

$$m_{11} = \frac{\sum_{t=1}^{T}(x_t - \bar{x})^2}{T} \qquad\qquad m_{22} = \frac{\sum_{t=1}^{T}(y_t - \bar{y})^2}{T}$$

$$m_{12} = \frac{\sum_{t=1}^{T}(x_t - \bar{x})(y_t - \bar{y})}{T}$$

entonces:

$$a\left(m_{11} - \frac{\lambda}{T}\right) + b\,m_{12} = 0$$

$$a\,m_{12} + b\left(m_{22} - \frac{\lambda}{T}\right) = 0$$

es un sistema de ecuaciones sin solución distinta de cero, a menos que su determinante se anule. Debido a las condiciones que impusimos al principio, tenemos que resolverlo junto con $a^2 + b^2 = 1$.

Si tomamos el determinante:

$$\begin{vmatrix} m_{11} - \frac{\lambda}{T} & m_{12} \\ m_{12} & m_{22} - \frac{\lambda}{T} \end{vmatrix} = 0$$

se tendrán dos raíces de la ecuación, λ_1 y λ_2, donde

$$\lambda_1 + \lambda_2 = \sum_{t=1}^{T} (x_t - \bar{x})^2 + \sum_{t=1}^{T} (y_t - \bar{y})^2 = U + V$$

y el valor λ_1 de la menor raíz mide la dispersión de los puntos observados alrededor de la regresión ortogonal*.

* Ver página 117.

1.4 REGRESION MULTIPLE.

El análisis de regresión múltiple consiste en la extensión del análisis de regresión lineal simple a cualquier número de variables de las cuales una es dependiente y las otras son independientes.

En este tipo de estudio se trata de buscar la ecuación que pueda servir mejor para expresar matemáticamente la relación de x_1 (variable dependiente) con x_2, x_3, x_4, ... (variables independientes).

Cuando nos referimos a sólo dos magnitudes, por razones de simplicidad mantuvimos una relación funcional lineal, si tomamos en cuenta los mismos principios y los adaptamos al nuevo conjunto de observaciones, podremos representar la tendencia general como:

$$x_1 = a_{12} x_2 + a_{13} x_3 + \cdots + a_{1n} x_n + b_1$$

en un espacio de n dimensiones, donde los puntos se ven descritos por $(x_1, x_2, \ldots, x_n)$ y que, junto con el vector $(a_{12}, \cdots, a_{1n})$ nos describen a su vez un plano P de n-1 dimensiones y distancia b_1 al orígen.

El valor de los coeficientes a_{1i} y b_1 será determinado de modo que el cuadrado de la distancia media entre los puntos y P sea mínima. Si tomamos distancias paralelas al eje X_1:

$$U = \frac{1}{T} \sum_{t=1}^{T} \left[x_{1t} - a_{12} x_{2t} - a_{13} x_{3t} - \cdots - a_{1n} x_{nt} - b_1 \right]^2$$

Al igual que en los casos anteriores, la función es derivable, por lo que existen:

$$\frac{\partial U}{\partial a_{12}}, \frac{\partial U}{\partial a_{13}}, \frac{\partial U}{\partial a_{14}}, \ldots, \frac{\partial U}{\partial a_{1n}}, \frac{\partial U}{\partial b_1}$$

y, de

$$\left[x_{1t} - a_{12}x_{2t} - a_{13}x_{3t} - \dots - a_{1n}x_{nt} - b_1\right]^2$$

en forma directa:

$$\frac{\partial U}{\partial b_1} = \frac{1}{T}\sum_{t=1}^{T}\left[x_{1t} - a_{12}x_{2t} - a_{13}x_{3t} - \dots - a_{1n}x_{nt}\right] = 0$$

entonces: $\bar{x}_1 = a_{12}\bar{x}_2 + a_{13}\bar{x}_3 + a_{14}\bar{x}_4 + \dots + a_{1n}\bar{x}_n + b$

que demuestra que P pasa por el punto medio $\bar{x} = (\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_n)$.

Considerando ahora como punto fijo a $\bar{x}$, $(a_{12}, a_{13}, \dots, a_{1n})$ como vector fijo y $x = (x_1, x_2, \dots, x_n)$ como punto móvil, podemos escribir:

$$(x_1 - \bar{x}_1) = a_{12}(x_2 - \bar{x}_2) + a_{13}(x_3 - \bar{x}_3) + \dots + a_{1n}(x_n - \bar{x}_n)$$

y U se puede expresar como:

$$U = \frac{1}{T}\sum_{t=1}^{T}\left[(x_{1t} - \bar{x}_1) - a_{12}(x_{2t} - \bar{x}_2) - \dots - a_{1n}(x_{nt} - \bar{x}_n)\right]^2$$

de ahí que:

$$\frac{\partial U}{\partial a_{12}} = \sum_{t=1}^{T}\frac{(x_{2t} - \bar{x}_2)(x_{1t} - \bar{x}_1)}{T} - a_{12}\sum_{t=1}^{T}\frac{(x_{2t} - \bar{x}_2)^2}{T} - \dots - a_{1n}\sum_{t=1}^{T}\frac{(x_{2t} - \bar{x}_2)(x_{nt} - \bar{x}_n)}{T} = 0$$

y, por analogía:

$$\frac{\partial U}{\partial a_{13}} = \sum_{t=1}^{T}\frac{(x_{3t} - \bar{x}_3)(x_{1t} - \bar{x}_1)}{T} - a_{12}\sum_{t=1}^{T}\frac{(x_{3t} - \bar{x}_3)(x_{2t} - \bar{x}_2)}{T} - \dots - a_{1n}\sum_{t=1}^{T}\frac{(x_{3t} - \bar{x}_3)(x_{nt} - \bar{x}_n)}{T} = 0$$

$$\vdots$$

$$\frac{\partial U}{\partial a_{1n}} = \sum_{t=1}^{T}\frac{(x_{nt} - \bar{x}_n)(x_{1t} - \bar{x}_1)}{T} - a_{12}\sum_{t=1}^{T}\frac{(x_{nt} - \bar{x}_n)(x_{2t} - \bar{x}_2)}{T} - \dots - a_{1n}\sum_{t=1}^{T}\frac{(x_{nt} - \bar{x}_n)^2}{T} = 0$$

Para manejar más fácilmente estos resultados, sea:

$$m_{ij} = \frac{1}{T} \sum_{t=1}^{T} (x_{it}-\bar{x}_i)(x_{jt}-\bar{x}_j) \qquad \begin{array}{l} i=1,2,\ldots,n \\ j=1,2,\ldots,n \end{array}$$

momento central de 2° orden, entonces:

$$\begin{array}{l} m_{11} - m_{12}a_{12} - a_{13}m_{13} - \cdots - m_{1n}a_{1n} = 0 \\ m_{21} - m_{22}a_{12} - m_{23}a_{13} - \cdots - m_{2n}a_{1n} = 0 \\ \vdots \\ m_{n1} - m_{n2}a_{12} - m_{n3}a_{13} - \cdots - m_{nn}a_{1n} = 0 \end{array}$$

y podemos resolverla usando una matríz de n x n:

$$M = \frac{1}{T} \sum_{t=1}^{T} (x_t-\bar{x})(x_t-\bar{x})'$$

i.e.
$$\begin{vmatrix} m_{11} & -m_{12} & -m_{13} & - \cdots - m_{1n} \\ m_{21} & -m_{22} & -m_{23} & - \cdots - m_{2n} \\ m_{31} & -m_{32} & -m_{33} & - \cdots - m_{3n} \\ \vdots & \vdots & \vdots & \vdots \quad \vdots \\ m_{n1} & -m_{n2} & -m_{n3} & - \cdots - m_{nn} \end{vmatrix}$$

Sean:

$a' = (a_{12}, a_{13}, \ldots, a_{1n})$ vector de componentes a_{ij}, $j = 2, 3, \ldots, n$

$x' = (x_2, x_3, \ldots, x_n)$ vector de componentes x_i, $i = 2, 3, \ldots, n$

m' = vector de componentes m_{1i} $\quad i = 2, 3, \ldots, n$

M' = matríz $(n-1) \times (n-1)$

m'' = transpuesta de m'

entonces
$$M = \begin{vmatrix} m_{11} & m'' \\ m' & M' \end{vmatrix}$$

de donde la ecuación del plano P resulta:

$$(x_1 - \bar{x}_1) = a^{1)}(x' - \bar{x}')$$

y el conjunto de ecuaciones normales:

$$m' - M'a' = 0$$

el sistema que forman, tiene solución específica si el determinante que se asocia a los a_{1i}, $i=2,3,\ldots,n$ es distinto de cero, esto es, si M' es no singular. Por lo cual, si $(M')^{-1}$ denota la matríz inversa de M' :

$$a' = (M')^{-1}m'$$

el valor que hace mínimo a U se obtiene utilizando este resultado de manera que:

$$U = m_{11} - 2a^{1)}m' + a^{1)}M'a'$$

de donde:

$$U = m_{11} - m^{1)}(M')^{-1}m'$$

y la dispersión de los puntos alrededor del plano estará dada por:

$$R^2 = 1 - \frac{U}{m^{11}} \qquad R^2 = \frac{m^{1)}(M')^{-1}m'}{m_{11}}$$

donde R^2 se conoce como " coeficiente de correlación múltiple" de x_1 con respecto a x_2, x_3, ..., x_n.

1.5 COMPARACION ENTRE ESTOS MODELOS.

Las regresiones simple y múltiple son las más utilizadas en la práctica. La mejor ventaja que se le puede atribuír a estos métodos es la sencillez para realizar los cálculos y lo preciso de sus conceptos.

Las regresiones ortogonales presentan un poco de dificultad para muchas personas por todos los medios de que se vale para poder realizarse, aunque, como sus valores se encuentran entre los de los dos casos de la regresión lineal simple, pueden considerarse de alguna manera más exactos.

Ambos métodos representan una forma muy directa de obtener resultados rápidos sobre la idea que se tenga del comportamiento de las variables pero, al hacer inferencias en los datos tal y como se nos presentan, estamos dejando de ver muchos posibles errores cometidos desde el momento en que se decidió tomarlos, hasta su captación y gráfica de comportamiento.

Se hace necesario entonces, conocer qué alcance general podemos atribuír tanto a los datos como a las regresiones calculadas a partir de ellos, sin tomar o considerar las posibles fuentes de error. Para esto se desarrollará el Capítulo 2, en donde se estudian modelos considerando los posibles errores.

1.6 EJEMPLOS.

REGRESION LINEAL SIMPLE Y MULTIPLE.

Las operaciones denominadas oferta y demanda se encuentran estrechamente vinculadas al concepto de precio, lo que justifica que se les defina en los siguientes términos: " la demanda de cualquier cosa, dado el precio, es la cantidad que se compra a ese precio por unidad de tiempo"*. Como podemos observar, se incluyen en esta definición los conceptos de precio y de unidad de tiempo, ya que la demanda carece de significado si no se estipula un precio y, por ser necesario, relacionarla a un determinado período de tiempo.

En cuanto a la oferta, podemos considerarla como " la cantidad que se pone a la venta a un precio determinado por unidad de tiempo"*, y es la operación lógica que complementa la idea de la demanda.

La oferta y la demanda presentan cuatro modalidades, ya que a menudo los costos de diferentes artículos están relacionados porque algunas cosas se producen simultáneamente (oferta conjunta), otros sirven para propósitos similares (oferta compuesta), otros sólo son útiles juntos (demanda conjunta) y la mayoría de ellos pueden utilizarse para varios fines diferentes (demanda compuesta).

El ejemplo elegido será presentado de una manera sencilla. Si tomamos la teoría más simplificada de la demanda (por período de tiempo)de un bien, podemos considerarla solamente relacionada con su precio. Con base en esto, el comportamiento del consumidor es tal que una cantidad menor (mayor) del bien es demandada cuando su precio aumenta (disminuye).

Concretaremos el análisis a un mercado competitivo, es decir, trataremos solamente en un mercado donde el número de com-

* DOMINGUEZ VARGAS, Sergio."Teoría Económica", Ed. Porrúa, S. A.

pradores y vendedores es tan grande que ningún comprador o vendedor puede controlar por sí solo o influír en forma significativa sobre el precio en el mercado del bien que se está tratando.

Si:

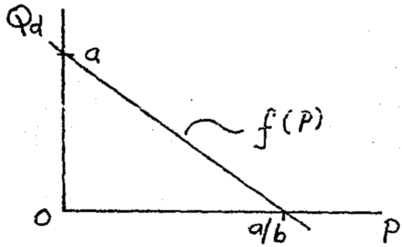
Q_d = cantidad demandada del bien
P = precio del bien

entonces:

$$Q_d = f(P)$$

Como el precio de cada bien estará determinado externamente en este modelo, podemos considerar a P como una variable exógena, y debido a que Q_d está determinada dentro de él, será una variable endógena (como tal, Q_d es dependiente y sus valores se hallan dentro del modelo).

Así, Q_d y P presentan una variación directa. En forma general, puede decirse entonces que la gráfica que representa eate fenómeno se parece a:



$$Q_d = a - bP$$

para $P \geq 0$

de tal modo que

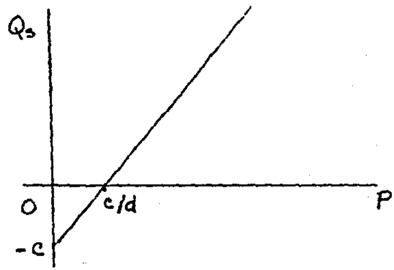
a) cuando $P = 0$, $Q_d = a$
b) cuando $P = a/b$, $Q_d = 0$
c) una unidad de aumento (disminución) en P tiene como resultado una disminución (aumento) de b unidades en Q_d .

Al mismo tiempo que eso ocurre, los comerciantes quieren ofrecer más mercancía con precios altos para aumentar sus ingresos. Si nuevamente consideramos una relación simple, la oferta que se presente al público de artículos de precio más alto tenderá a ser mayor en beneficio de sus utilidades. Nuevamente tendremos una relación lineal.

Si:

Q_s = cantidad ofrecida de un bien

entonces:



$$Q_s = -c + dP$$

esto implica que:

a) $Q_s \gtrless 0$ solamente cuando $P \gtrless c/d$

(es decir, la oferta no se presentará a menos que P sea mayor o cuando menos igual a c/d).

b) una unidad de aumento (disminución) en P, tiene como resultado una disminución (aumento) de d unidades en Q_s .

Parece en cierta forma natural esperar que exista una relación más evidente entre estos fenómenos.

Supongamos entonces, que el bien en discusión es comprado y vendido en un mercado aislado, con cambios en este mercado que tienen efectos despreciables en otras partes de la economía y viceversa. Esto significa que el nuestro, será un análisis parcial.

Como primer paso, recordemos que las funciones de demanda y oferta representan el comportamiento de consumidores y produc tores. Vistos de ese modo, las funciones representan los propósitos o planes de compradores y vendedores de adquirir o vender ciertas cantidades del bien en respuesta a ciertos precios. En el caso en que esos propósitos coinciden, es decir, la cantidad demandada es exactamente igual a la cantidad ofrecida, nos encontraremos frente a un estado de balance o equilibrio. Esto implica que es necesario especificar una condición de equilibrio, que será entonces de la forma:

$$Q_d = Q_s$$

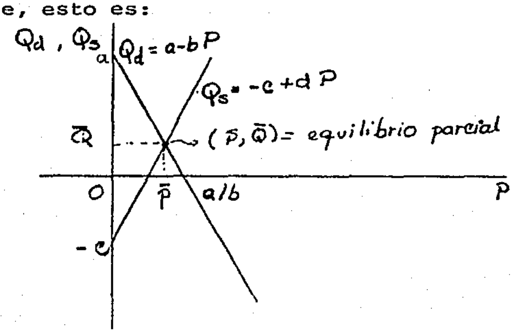
que nos dice que el exceso de demanda ($Q_d - Q_s$) debe ser cero. A esto se le conoce como mercado " clasificado", y el equilibrio alcanzado será un equilibrio parcial, al cual, como ha sido definido en un estado de balance (sin tendencia o cambio), puede dársele el nombre de equilibrio estático.

Tendremos ahora un modelo:

$$\begin{aligned} Q_d &= Q_s \\ Q_d &= a - bP \qquad (a,b > 0) \\ Q_s &= -c + dP \qquad (c,d > 0) \end{aligned}$$

donde las incógnitas $\bar{Q}_d$, $\bar{Q}_s$ y $\bar{P}$ serán los "valores de equilibrio" de las variables Q_d , Q_s y P respectivamente (ahora las tres variables Q_d , Q_s y P son endógenas, ya que sus valores de equilibrio serán determinados dentro del modelo de mercado de -- equilibrio parcial). Como $Q_d = Q_s$, entonces $\bar{Q}_d = \bar{Q}_s = \bar{Q}$.

Gráficamente, esto es:



y: $Q_d = Q_s$

$$a - b\bar{P} = -c + d\bar{P}, \quad a + c = d\bar{P} + b\bar{P} \Rightarrow (a+c) = (b+d)\bar{P}$$

$$\therefore \quad \bar{P} = \frac{a+c}{b+d}$$

ahora, como $\bar{Q} = \bar{Q}_s = \bar{Q}_d$:

$$\bar{Q} = a - b\bar{P} = a - \frac{b\,(a+c)}{b+d} = \frac{a\,(b+d) - b\,(a+c)}{b+d} = \frac{ad - bc}{b+d}$$

$\bar{P}$ y $\bar{Q}$ dependerán entonces de los parámetros a, b, c y d, esto quiere decir que son valores determinados, pero aún así, es necesario que sean positivos para que tengan interpretación económica. Lo único que necesitamos es pedir que $ad > bc$.

Este mismo modelo se puede ver como regresión múltiple. Si consideramos a, b, c y d positivos:

$$1\,Q_d - 1\,Q_s + 0\,P = 0$$
$$1\,Q_d + 0\,Q_s + b\,P = a$$
$$0\,Q_d + 1\,Q_s - d\,P = -c$$

la matríz será   vector   vector

$$\underset{M}{\begin{vmatrix} 1 & -1 & 0 \\ 1 & 0 & b \\ 0 & 1 & -d \end{vmatrix}} \quad \underset{x}{\begin{vmatrix} Q_d \\ Q_s \\ P \end{vmatrix}} = \underset{e}{\begin{vmatrix} 0 \\ a \\ -c \end{vmatrix}}$$

debemos encontrar el valor de $|M|$ para comprobar que existe una solución única para $Mx = e$.

$$|M| = \begin{vmatrix} 1 & -1 & 0 \\ 1 & 0 & b \\ 0 & 1 & -d \end{vmatrix} = 1 \begin{vmatrix} 0 & b \\ 1 & -d \end{vmatrix} - (-1) \begin{vmatrix} 1 & b \\ 0 & -d \end{vmatrix}$$

$$= -b-d = -(d+b) \neq 0$$

Como $|M| \neq 0 \Rightarrow \exists\, M^{-1}$

Para encontrar M^{-1} necesitaremos la matríz de cofactores C

$$C = \begin{vmatrix} \begin{vmatrix} 0 & b \\ 1 & -d \end{vmatrix} & -\begin{vmatrix} 1 & b \\ 0 & -d \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \\ -\begin{vmatrix} -1 & 0 \\ 1 & -d \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 0 & -d \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} \\ \begin{vmatrix} -1 & 0 \\ 0 & b \end{vmatrix} & -\begin{vmatrix} 1 & 0 \\ 1 & b \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} \end{vmatrix} = \begin{vmatrix} -b & d & 1 \\ -d & -d & -1 \\ -b & -b & 1 \end{vmatrix}$$

entonces $C' = \text{adj } M$

$$C' = \begin{vmatrix} -b & -d & -b \\ d & -d & -b \\ 1 & -1 & 1 \end{vmatrix}$$

y M^{-1} estará dada por:

$$M^{-1} = \frac{\text{adj } M}{|M|} = -\frac{1}{(d+b)} \begin{vmatrix} -b & -d & -b \\ d & -d & -b \\ 1 & -1 & 1 \end{vmatrix}$$

que debe satisfacer $M^{-1}M = I$.

la solución será:

$$M^{-1}e = \frac{-1}{(d+b)}\begin{bmatrix} -b & -d & -b \\ d & -d & -b \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ a \\ -c \end{bmatrix}$$

$$= -\frac{1}{(d+b)}\begin{bmatrix} -(ad-bc) \\ -(ad-bc) \\ -(a+c) \end{bmatrix} = \begin{bmatrix} \bar{Q}_d \\ \bar{Q}_s \\ \bar{P} \end{bmatrix}$$

i.e: $\bar{Q} = \bar{Q}_d = \bar{Q}_s = \frac{ad-bc}{d+b}$

$$\bar{P} = (a+c)/(b+d)$$

que es igual a la encontrada de la manera anterior.

CAPITULO 2

METODOS ESTADISTICOS PARA MODELOS ESTOCASTICOS.

2.0 MODELOS ESTOCASTICOS.

El tratamiento de los procesos estocásticos constituye una parte importante de la teoría de probabilidades, es decir; los modelos que se desarrollarán en este capítulo involucran términos aleatorios no considerados anteriormente.

El término aleatorio se conoce como " error " en la ecuación. Hay dos tipos de errores:

a) Error de medida.- Existen varias razones por las cuales Y puede medirse incorrectamente. Por ejemplo, al tomar los datos si se trata de una encuesta, las personas entrevistadas pueden mentir sobre su situación real.

b) Error estocástico.- Ocurre por la irreproducibilidad inherente de fenómenos sociales. Aún cuando no se tenga un error de medida, siempre hay diferencias no predecibles que son llamadas aleatorias o estocásticas.

Tomando esto en cuenta, se desarrollarán los modelos:

a) Regresión Lineal Simple.

b) Regresión Lineal Múltiple.

c) Modelos Lineales con errores en las variables.

2.1 MODELO DE REGRESION LINEAL SIMPLE.

El significado principal y tal vez, más "natural" de -- linealidad es que la esperanza condicional de Y es una función - lineal de las X_t, lo cual indica, geométricamente, que la curva de regresión es una línea recta.

El segundo sentido de linealidad es que la espectativa condicional de Y, $E(Y/X_t)$, es una función lineal de los parámetros, pudiendo ser o no lineal en la variable X.

Nosotros consideraremos el segundo sentido ya que el modelo será considerado así en forma más amplia.

Para poder representar un fenómeno por medio de un modelo matemático se necesitan ciertas consideraciones, que en este - caso serán: el modelo debe ser el más sencillo que pueda encontrarse, ya que se refiere a una variable endógena que depende de una sola variable exógena y una perturbación aleatoria no observable. Es decir, una magnitud Y se considera determinada por medio de - otra magnitud X. La dependencia que supusimos permite relacionarlas como:

$$Y = aX + b + \xi$$

donde a y b son desconocidos y ξ es el término aleatorio no observable.

Hay que recordar que la magnitud observable X y la no observable ξ son consideradas como las causas que determinan el valor que toma la magnitud Y; ya que el término ξ representa el efecto de todos los factores que no podemos o no sabemos identificar. El problema estadístico consiste en la estimación de los parámetros a y b.

Uno de los métodos más eficaces y populares del análisis de regresión es el de mínimos cuadrados ordinarios, que se debe a Carl Friederich Gauss.

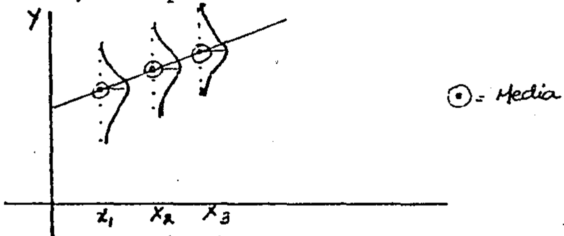
En este apartado, tomaremos el enfoque gaussiano para - dos variables:

$$Y_t = aX_t + b + \xi_t$$

donde Gauss hace los siguientes supuestos:

1.-* $E(\xi_t / X_t) = 0 \quad \forall t.$

es decir, el valor esperado condicional de ξ_t al X_t dado, es cero. Geométricamente, esto puede verse:



Aquí, las variables X_t, Y_t se consideran como observadas sin error, ya que aunque los datos económicos estén indiscutiblemente viciados por errores de medida más o menos grandes, se estima frecuentemente que estos errores son despreciables en comparación con los elementos no observables que influyen en la determinación de las variables endógenas.

Por otra parte, si no se tuviera el supuesto de que la esperanza matemática de las ξ es nula para cualesquiera valores - tomados por X, los parámetros a y b incluídos en el modelo no serían identificables. Esto es, esta condición significa que el término aditivo ξ del modelo toma valores tanto positivos como negativos para todo valor de la variable exógena, lo cual no se satisface si el efecto de los factores no observables o no identificables se correlaciona con el efecto de la magnitud explicativa . Matemáticamente podemos expresarlo como:

$$E(Y_t / X_1, X_2, \ldots, X_i) = aX_t + b$$

donde los parámetros a y b serían entonces características de la

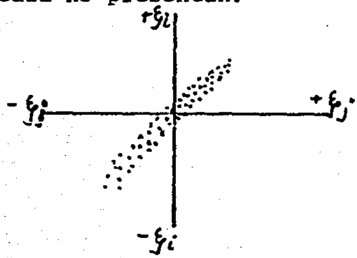
* Se considera que los errores son independientes.

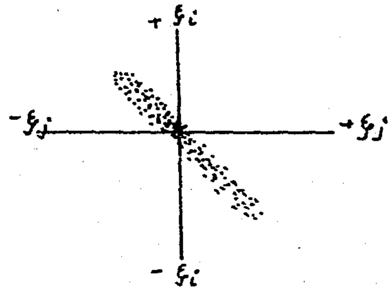
ley de probabilidad condicional de las Y. Si conocemos los valores de los parámetros a y b y de los X_t, podremos obtener la esperanza de Y directamente por medio de $aX_t + b$.

2.- $$\begin{aligned} cov(\xi_i, \xi_j) &= E(\xi_i - E(\xi_i)][\xi_j - E(\xi_j)] \\ &= E(\xi_i, \xi_j) \quad \text{por el supuesto 1} \\ &= 0 \qquad i \neq j \end{aligned}$$

donde i y j son dos observaciones diferentes y donde cov significa covarianza.

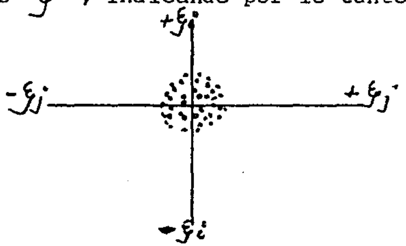
Esto quiere decir que las perturbaciones ξ_i y ξ_j no están correlacionadas. Técnicamente, este supuesto se conoce como de " no correlación serial" o de " no autocorrelación". Lo que dice que: dada X_t, las desviaciones de dos cualesquiera valores de Y de su media no presentan:





En la primera figura, la correlación positiva es cuando tenemos ξ positivo seguido de ξ positivo, o ξ negativo seguido de ξ negativo. En la segunda figura, la correlación negativa es cuando tenemos ξ positivo seguido de ξ negativo y viceversa.

Si las perturbaciones siguen patrones de tipo sistemático, como los anteriores, se dice que existe autocorrelación o correlación serial, y lo que requiere este supuesto es precisamente que no la exista. La siguiente figura no muestra un patrón sistemático para los ξ , indicando por lo tanto, cero correlación:

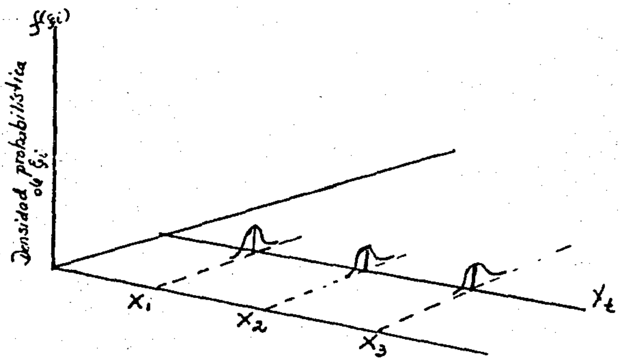


3.-

$$
\begin{aligned}
var\,(\xi_t / X_t) &= E\,[\xi_t - E(\xi_t)]^2 \\
&= E(\xi^2) \quad \text{por el supuesto 1} \\
&= \sigma^2
\end{aligned}
$$

donde var = varianza.

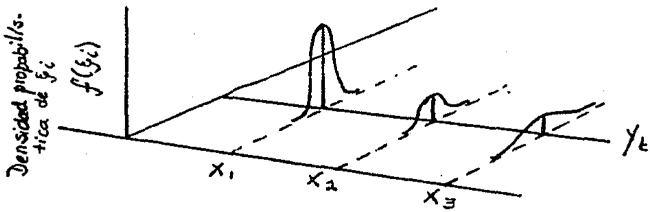
Expresa que la varianza de ξ_t para cada X_t (i. e., la varianza condicional de ξ_t) es un número positivo constante e - igual a σ^2 . Técnicamente lo que se representa aquí es el supuesto de " homoscedasticidad", igual (homos) dispersión (cedasticidad) o igual varianza. Podemos representarlo geométricamente por:



El caso opuesto se conoce como heteroscedasticidad, desviación desigual o varianza desigual. Simbólicamente puede escribirse:

$$Var\,(\xi_t\,/X_t) = \sigma_t^2$$

Como puede verse, en esta ecuación aparece un subíndice, que quiere decir que la varianza de la población ya no es constante. Podemos representar esto como:



4.- $$cov\,(\xi_t\,, X_t) = E[\xi_t - E(\xi_t)]\,[X_t - E(X_t)]$$
$$= 0$$

es decir, la perturbación ξ y la variable explicatoria X no están correlacionadas. En cualquier caso de correlación entre X y ξ es difícil aislar la influencia individual de X y de ξ sobre Y.

Aparte de la muestra en estudio, no se dispone de ninguna información con respecto a los coeficientes numéricos a y b, los cuales pueden tomar a priori valores positivos, negativos o nulos cualesquiera.

AJUSTE POR MINIMOS CUADRADOS.

Tal y como se ha considerado el modelo, las mejores estimaciones para a y b son las facilitadas por los coeficientes de la regresión simple de Y en relación con X .

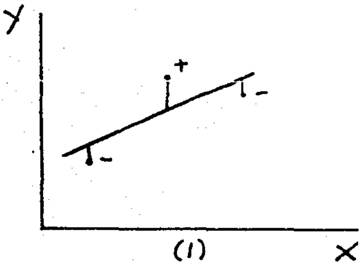
Para llevarla a cabo es importante considerar que se tomará como un buen ajuste aquel que " minimice el error total". Los criterios tentativos lógicos a seguir serían:

∘ Considerar una línea de ajuste que minimice la suma de los errores:

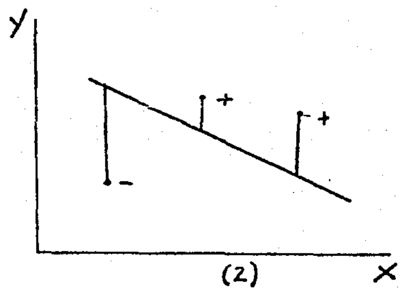
$$\sum_{t=1}^{T} (Y_t - \hat{Y}_t)$$

donde Y_t = valor real
$\hat{Y}_t$ = valor ajustado.

Desafortunadamente, esto no funciona. Al usar este criterio las dos líneas de las figuras siguientes se ajustan igualmente aunque la primera sea intuitivamente mejor que la segunda, y la suma de los errores sea cero.



(1)



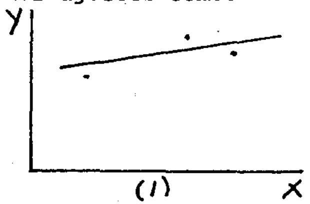
(2)

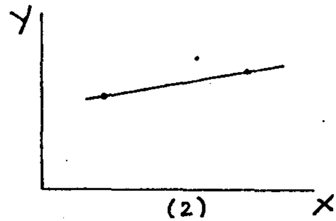
El problema es el signo. Este criterio debe rechazarse porque no proporciona distinción alguna entre ajustes buenos y malos.

°° Hay dos formas de evitar los problemas de signo. El primero es minimizar la suma de los valores absolutos de los errores:

$$\sum_{t=1}^{T} | Y_t - \hat{Y}_t |$$

Pero, como a los grandes errores positivos no se les permite interponerse con los grandes errores negativos, este criterio daría malos ajustes como:





La figura 2 satisface el criterio mejor que la 1, aunque, de hecho, se puede ver que la 2° gráfica podría indicar que el unir los dos puntos extremos se ajusta mejor que cualquier otra recta.

Esta tal vez no sea la mejor solución al problema porque no presta ninguna atención al punto medio cualquiera que sea su valor, por lo que pudiera ser preferible elegir la 1° gráfica ya que considera todos los puntos.

°°° Como una segunda opción para evitar el problema del signo, se minimiza la suma de los cuadrados de los errores:

$$\sum_{t=1}^{T} (Y_t - \hat{Y}_t)^2$$

y ahora sí tendremos el criterio conocido por " mínimos cuadrados" o " mínimos cuadrados ordinarios".

Entre sus ventajas están:

a) Al evitar los problemas de signo elevando al cuadrado los errores, los mínimos cuadrados son sencillos de manejar algebráicamente.

b) Hay dos justificaciones teóricas para los mínimos cuadrados: El teorema de Gauss-Markov, y el criterio de máxima verosimilitud para un modelo de regresión normal.

El teorema de Gauss-Markov dice que " dentro de la clase de estimadores lineales insesgados de b , el estimador mínimo cuadrático $\hat{b}$ tiene varianza mínima. De la misma manera, $\hat{a}$ es el estimador de varianza mínima de a ".

La regresión de Y con respecto a X se defina como $\hat{Y} = \hat{a}X+\hat{b}$, donde a y b deben ser tales que:

$$U = \sum_{t=1}^{T} (y_t - aX_t - b)^2 \text{ sea mínima}$$

El desarrollo entonces debe ser el mismo que en el capítulo anterior, así que podemos decir que:

$$\hat{a} = \frac{\sum_{t=1}^{T} (Y_t - \bar{Y})(X_t - \bar{X})}{\sum_{t=1}^{T} (X_t - \bar{X})^2}$$

$$\hat{b} = \bar{Y} - \hat{a}\bar{X}$$

donde $\bar{X} = \frac{\sum_{t=1}^{T} X_t}{T}$ $\qquad \bar{Y} = \frac{\sum_{t=1}^{T} Y_t}{T}$

Ya que se tiene un ajuste, tal vez lo más natural ahora sería preguntarse qué tan bueno es; para eso, si usamos desviaciones de la media, es decir $x_t = (X_t - \bar{X})$, $y_t = (Y_t - \bar{Y})$:

$$U = \sum_{t=1}^{T} (y_t - \hat{a}x_t - \hat{b})^2$$

$$\frac{\partial U}{\partial b} = -2 \sum_{t=1}^{T} (y_t - \hat{a}x_t - \hat{b}) = 0$$

$$= \sum_{t=1}^{T} (y_t - \hat{a}x_t - \hat{b}) = 0$$

$$\sum_{t=1}^{T} y_t - \hat{a}\sum_{t=1}^{T} x_t - T\hat{b} = 0$$

$$\sum_{t=1}^{T} y_t - \hat{a}\sum_{t=1}^{T} x_t = T\hat{b}$$

pero $\sum_{t=1}^{T} x_t = 0$, entonces: $\sum_{t=1}^{T} y_t = T\hat{b}$

$$\hat{b} = \frac{\sum_{t=1}^{T} y_t}{T} = \bar{y}$$

Al hacer lo mismo para $\hat{a}$:

$$\frac{\partial U}{\partial a} = \sum_{t=1}^{T} (2)(-x_t)(y_t - \hat{a}x_t - \hat{b}) = 0$$

$$= \sum_{t=1}^{T} (x_t)(y_t - \hat{a}x_t - \hat{b}) = 0$$

$$= \sum_{t=1}^{T} x_t y_t - \hat{a}\sum_{t=1}^{T} x_t^2 - \hat{b}\sum_{t=1}^{T} x_t = 0$$

otra vez, como $\sum_{t=1}^{T} x_t = 0$; $\hat{b}\sum_{t=1}^{T} x_t = 0$

$$\sum_{t=1}^{T} x_t y_t - \hat{a}\sum_{t=1}^{T} x_t^2 = 0 \quad \therefore \quad \hat{a} = \frac{\sum_{t=1}^{T} x_t y_t}{\sum_{t=1}^{T} x_t^2}$$

Estos estimadores se conocen como " estimadores mínimo cuadrados" por derivarse de este principio. Sus características son:

1.- Están expresados únicamente en términos de cantidades observadas.
2.- Son estimadores puntuales, es decir, dada la muestra, cada estimador proporcionará un solo valor (punto) del parámetro relevante

Una vez que se han obtenido los estimadores de los mínimos cuadrados, la línea de regresión ajustada tendrá las siguentes propiedades:

1.- Pasa a través de la media muestral (esto quedó bien determina do durante el desarrollo).

2.- El valor medio del Y estimado ($\hat{Y}_t$) es igual al valor medio del Y observado debido a que:

$$\begin{aligned} \hat{y}_t &= \hat{a} x_t + b \\ &= \hat{a} x_t + (\bar{y}_t - \hat{a} \bar{x}_t) \\ &= \bar{y}_t + \hat{a} (x_t - \bar{x}_t) \end{aligned}$$

sumando a ambos lados sobre los valores muestrales y dividiendo por T, se tiene:

$$\bar{\hat{y}}_t = \bar{y}$$

donde utilizamos que $\sum_{t=1}^{T} (x_t - \bar{x}) = 0$, porque:

si definimos

$$x = (X_t - \bar{X})$$

$$\begin{aligned} \sum_{t=1}^{T} x_t &= \sum_{t=1}^{T} (X_t - \bar{X}) \\ &= \sum_{t=1}^{T} X_t - T\bar{X} \end{aligned}$$

y, como $\quad \bar{X} = \sum_{t=1}^{T} X_t / T$

entonces $\quad T\bar{X} = \sum_{t=1}^{T} X_t$

de ahí que $\quad \sum_{i=1}^{T} X_t = T\bar{X} - T\bar{X} = 0$

3.- El valor medio de los residuos es cero. De la ecuación:

$$-2\sum_{t=1}^{T}(y_t-\hat{a}x_t-\hat{b})=0$$

como $\xi_t = y_t - \hat{a}x_t - \hat{b}$

entonces

$$-2\sum_{t=1}^{T}\xi_t = 0\ ,\quad donde\ \bar{\xi}=0$$

4.- Los residuos ξ_t no están correlacionados con el valor predicho de y_t, lo cual se puede verificar

$$\begin{aligned}\sum_{t=1}^{T}\hat{y}_t\xi_t &= \hat{a}\sum_{t=1}^{T}x_t\xi_t\\ &= \hat{a}\sum_{t=1}^{T}x_t(y_t-\hat{a}x_t)\\ &= \hat{a}\sum_{t=1}^{T}x_t y_t-\hat{a}^2\sum_{t=1}^{T}x_t^2\\ &= \hat{a}^2\sum_{t=1}^{T}x_t^2-\hat{a}^2\sum_{t=1}^{T}x_t^2\\ &= 0\end{aligned}$$

donde $\hat{a} = \dfrac{\sum_{t=1}^{T}x_t y_t}{\sum_{t=1}^{T}x_t^2}$

5.- Los residuos ξ_t no están correlacionados con X_t, esto es $\sum_{t=1}^{T}\xi_t x_t = 0$.

Debido a que los estimadores son función de los datos, pero dado que los datos pueden fácilmente cambiar de una muestra a otra y que los estimadores por lo tanto cambiarían inmediatamente, lo que necesitaremos será alguna medida de la confiabilidad de los estimadores $\hat{a}$ y $\hat{b}$. En estdística, la precisión de un estimador se mide por su error standar, por lo que necesitaremos el cálculo de sus esperanzas y varianzas.

Si escribimos: $$\hat{a} = \sum_{t=1}^{T} \left\{ \frac{x_t}{K} \right\} Y_t \qquad \sum_{t=1}^{T} x_t^2 = K$$

entonces
$$\hat{a} = \sum_{t=1}^{T} \omega_t Y_t = \omega_1 Y_1 + \omega_2 Y_2 + \cdots + \omega_T Y_T .$$
donde
$$\omega_t = x_t / K$$

lo que implica, que $\hat{a}$ es una combinación lineal de las variables - aleatorias Y_t .

Como la esperanza es un operador lineal, entonces
$$E(\hat{a}) = \omega_1 E(Y_1) + \omega_2 E(Y_2) + \cdots + \omega_T E(Y_T)$$
$$= \sum_{t=1}^{T} \omega_t E(Y_t)$$

y, al haber asumido que las variables Y_i son independientes
$$Var(\hat{a}) = \omega_1^2 Var Y_1 + \cdots + \omega_T^2 Var Y_n$$
$$= \sum_{t=1}^{T} \omega_t^2 Var(Y_t)$$

y, $E(Y_t)$ cae sobre una línea recta (la verdadera línea de regresión)
$$E(\hat{a}) = \sum_{t=1}^{T} \omega_t (a x_t + b)$$
$$= a \sum_{t=1}^{T} \omega_t x_t + b \sum_{t=1}^{T} \omega_t$$

si volvemos a tomar $\omega_t = x_t / K$
$$E(\hat{a}) = \frac{a}{K} \sum_{t=1}^{T} (x_t)(x_t) + \frac{b}{K} \sum_{t=1}^{T} x_t$$

y, dado que $\sum_{t=1}^{T} x_t = 0$, $E(\hat{a}) = 0 + \frac{a}{K} \sum_{t=1}^{T} x_t^2$

Nota: $Cov(\hat{a}, \hat{b}) = E(\hat{a} - a)(\hat{b} - b) = E\left\{ \frac{1}{n} \sum_t \xi_t - \frac{\bar{X} \sum_t (X_t - \bar{X}) \xi_t}{\sum_t (X_t - \bar{X})^2} \right\}$
$$\left\{ \frac{\sum_j (X_j - \bar{X}) \xi_j}{\sum_j (X_j - \bar{X})^2} \right\} = E\left\{ \frac{\sum_t (X_t - \bar{X}) \xi_t^2}{n \sum_i (X_t - \bar{X})^2} - \frac{\bar{X} \sum (X_i - \bar{X}) \xi_i^2}{[\sum (X_i - \bar{X})^2]^2} = - \frac{\bar{X} \sigma^2}{\sum_i (X_i - \bar{X})^2} \right.$$

y como habíamos definido $K = \sum_{t=1}^{T} x_t^2$

esto es

$$E(\hat{a}) = \frac{a}{\sum_{t=1}^{T} x_t^2} \sum_{t=1}^{T} x_t^2 \qquad \therefore \quad E(\hat{a}) = a$$

es decir, $\hat{a}$ es un estimador insesgado para a .

Para calcular su varianza:

al sustituír $\sigma^2 = Var\, Y_t$

$$Var(\hat{a}) = \omega_1^2\, var\, Y_1 + \ldots + \omega_T^2\, var\, Y_t$$

$$Var(\hat{a}) = \sum_{t=1}^{T} \omega_t^2 \sigma^2$$

$$Var(\hat{a}) = \sum_{t=1}^{T} \frac{x_t^2}{K^2} \sigma^2$$

$$= \frac{\sigma^2}{K^2} \sum_{t=1}^{T} x_t^2$$

y, volviendo al valor de $K = \sum_{t=1}^{T} x_t^2$

tendremos que $Var(\hat{a}) = \dfrac{\sigma^2}{\sum_{t=1}^{T} x_t^2}$

Con respecto a $\hat{b}$:

Sea $e_t = x_t \Big/ \sum_{t=1}^{T} x_t^2$

entonces $\hat{b} = \bar{Y} - \hat{a}\,\bar{X}$

$$\hat{b} = \frac{\sum_{t=1}^{T} Y_t}{T} - \bar{X} \sum_{t=1}^{T} c_t Y_t$$

$$\hat{b} = \sum_{t=1}^{T} \left(\frac{1}{T} - \bar{X} c_t \right) Y_t$$

y $$E(\hat{b}) = \sum_{t=1}^{T} \left(\frac{1}{T} - \bar{X} c_t \right) E(Y_t)$$

$$= \sum_{t=1}^{T} \left(\frac{1}{T} - \bar{X} c_t \right) (a X_t + b)$$

al multiplicar cruzado

$$E(\hat{b}) = \sum_{t=1}^{T} \left(\frac{b}{T} - \bar{X} c_t b + \frac{a X_t}{T} - \bar{X} c_t a X_t \right)$$

$$= b + a\bar{X} - a\bar{X}$$

$$\therefore \quad E(\hat{b}) = b$$

y su varianza: $$\hat{b} = \sum_{t=1}^{T} \left(\frac{1}{T} - \bar{X} c_t \right) Y_t$$

$$Var(\hat{b}) = Var \left[\sum_{t=1}^{T} \left(\frac{1}{T} - \bar{X} c_t \right) Y_t \right]$$

$$Var(\hat{b}) = \sum_{t=1}^{T} \left(\frac{1}{T} - \bar{X} c_t \right)^2 Var\, Y_t$$

$$= \sigma^2 \sum_{t=1}^{T} \left(\frac{1}{T} - \bar{X} c_t \right)^2$$

$$= \sigma^2 \sum_{t=1}^{T} \left(\frac{1}{T^2} - \frac{2\bar{X} c_t}{T} + \bar{X}^2 c_t^2 \right)$$

como $\sum_{t=1}^{T} c_t = 0$ y $\sum_{t=1}^{T} c_t^2 = \dfrac{1}{\sum_{t=1}^{T} x_t^2}$

$$= \sigma^2 \left(\frac{1}{T} + \frac{\bar{X}^2}{\sum_{t=1}^{T} x_t^2} \right) = \sigma^2 \; \frac{\sum_{t=1}^{T} X_t^2 + T\bar{X}^2}{T \sum_{t=1}^{T} X_t^2}$$

$$y \quad \sum_{t=1}^{T} x_t^2 = \sum_{t=1}^{T} X_t^2 - T\bar{X}^2$$

esto implica que

$$Var(\hat{b}) = \sigma^2 \frac{\sum_{t=1}^{T} X_t^2 - T\bar{X}^2 + T\bar{X}^2}{T \sum_{t=1}^{T} x_t^2}$$

$$\therefore \qquad Var(\hat{b}) = \sigma^2 \frac{\sum_{t=1}^{T} X_t^2}{T \sum_{t=1}^{T} x_t^2}$$

Todas las cantidades pueden estimarse a partir de los datos con excepción de σ^2, que es estimada mediante:

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{T} \xi^2}{T-2}$$

donde $$\sum_{t=1}^{T} \xi_t^2 = \sum_{t=1}^{T} y_t^2 - \hat{a}^2 \sum_{t=1}^{T} x_t^2$$

Solamente queda por ver en qué medida se ajusta la línea de regresión muestral a los datos; porque, si todas las observaciones coinciden con la línea de regresión, obtendríamos un " ajuste perfecto", lo cual raras veces ocurre. Generalmente tiende a haber algunos ξ_t positivos y otros negativos, con la esperanza de que los residuos localizados alrededor de la línea de regresión sean lo más pequeños posible. El coeficiente de determinación r^2 (en el caso de dos variables) o R^2 (regresión múltiple) es una medida resumen que nos dice qué tan exactamente la línea de regresión muestral se ajusta a los datos.

Para calcular r^2 recordemos que $Y_t = \hat{Y}_t + \xi_t$

o, en forma de desviación $y_t = \hat{y}_t + \xi_t$

elevando al cuadrado

$$\sum_{t=1}^{T} y_t^2 = \sum_{t=1}^{T} \hat{y}_t^2 + \sum_{t=1}^{T} \xi_t^2 + 2 \sum_{t=1}^{T} \hat{y}_t \xi_t$$

$$\sum_{t=1}^{T} y_t^2 = \sum_{t=1}^{T} \hat{y}_t^2 + \sum_{t=1}^{T} \xi_t^2 \qquad \text{porque } \sum_{t=1}^{T} \hat{y}_t \xi_t = 0$$

y, como $\hat{y}_t = \hat{a} x_t$

$$\sum_{t=1}^{T} y_t^2 = \hat{a}^2 \sum_{t=1}^{T} x_t^2 + \sum_{t=1}^{T} \xi_t^2$$

Las sumas de los cuadrados de esta expresión puede escribirse:

$$\sum_{t=1}^{T} y_t^2 = \sum_{t=1}^{T} (Y_t - \bar{Y})^2$$

variación* total de Y con respecto a su media muestral, conocido como suma total de cuadrados (STC).

$$\sum_{t=1}^{T} (\hat{Y}_t - \bar{Y})^2 = \hat{a}^2 \sum_{t=1}^{T} x_t^2$$

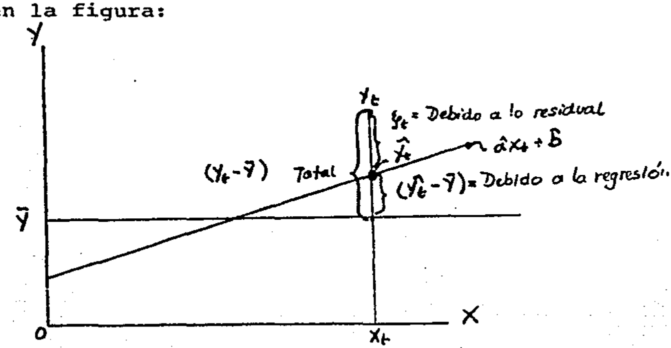
variación de los valores de Y con respecto a su media, ya que, $\bar{\hat{Y}} = \bar{Y}$, y que se conoce también como suma de cuadrados debi

* Los términos variación y varianza son diferentes. Variación es la suma de las desviaciones de una variable con respecto a su media, elevadas al cuadrado. Varianza es igual a la variación dividida por el número de grados de libertad. El término " número de grados de libertad" es igual al número total de observaciones en la muestra menos el número de restricciones (lineales) impuesto en ellas.

da a la regresión, suma explicada de cuadrados (SEC).

$\sum_{t=1}^{T} \xi_t^2$ es la desviación residual o no explicada de los valores de Y con respecto a la línea de regresión, o simplemente la suma residual de cuadrados (SRC).

Por consiguiente, STC = SEC + SRC nos muestra que la variación de los valores observados de Y alrededor de su media pueden ser partidos en dos, uno atribuíble a la línea de regresión y otro a las fuerzas aleatorias en razón de que no todas las observaciones reales de Y caen sobre la línea de regresión. Podemos ver esto en la figura:



Dividiendo ambos lados por STC:

$$1 = \frac{\text{SEC}}{\text{STC}} + \frac{\text{SRC}}{\text{STC}}$$

$$= \frac{\sum_{t=1}^{T} (\hat{Y}_t - \bar{Y})}{\sum_{t=1}^{T} (Y_t - \bar{Y})^2} + \frac{\sum_{t=1}^{T} (\xi_t)^2}{\sum_{t=1}^{T} (Y_t - \bar{Y})^2}$$

que definimos como

$$r^2 = \frac{\sum_{t=1}^{T} (\hat{Y}_t - \bar{Y})^2}{\sum_{t=1}^{T} (Y_t - \bar{Y})^2} = \frac{SEC}{STC}$$

La cantidad definida como r^2 se conoce por " coeficiente de determinación"(muestral). r^2 mide la proporción o porcentaje de la variación total en Y explicada por el modelo de regresión. Sus propiedades importantes son:

1.- Es una cantidad no negativa.
2.- Sus límites son $0 \leq r^2 \leq 1$

Una cantidad muy relacionada con r^2 pero conceptualmente diferente es el " coeficiente de correlación" que es una medida del grado de asociación entre dos variables. Puede calcularse como:

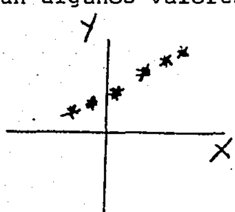
$$r = \pm \sqrt{r^2}$$

o a partir de su definición

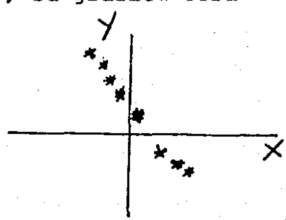
$$r = \frac{\sum_{t=1} x_t y_t}{\sqrt{(\sum_{t=1} x_t^2)(\sum_{t=1} y_t^2)}}$$

$$r = \frac{T \sum_{t=1}^{T} X_t Y_t - (\sum_{t=1} X_t)(\sum_{t=1} Y_t)}{\sqrt{[T \sum_{t=1} X_t^2 - (\sum_{t=1} X_t)^2][T \sum_{t=1} Y_t^2 - (\sum_{t=1} Y_t)^2]}}$$

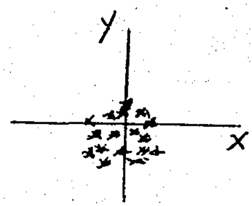
según algunos valores de r, su gráfica será



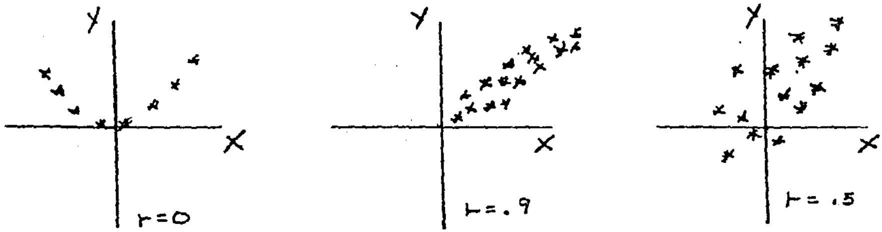
$r = +1$



$r = -1$



$r = 0$



Algunas de las propiedades de r son:

1.- Puede ser positivo o negativo, el signo dependerá del numerador, que mide la covariación de las dos variables.

2.- Estará entre los límites de -1 y 1, esto es, $-1 \leq r \leq 1$.

3.- Es de naturaleza simétrica; es decir, el coeficiente de correlación entre X e Y (r_{xy}) es igual al coeficiente de correlación de Y y X (r_{yx}).

4.- Es independiente del orígen y la escala; es decir, si definimos $X_t^* = aX_t + c$ y $Y_t^* = bY_t + d$, donde $a > 0$, $b > 0$, c y d constantes, el r entre X^* e Y^* es igual al r entre X e Y.

5.- Si X e Y son estadísticamente independientes (esto es, si y sólo si $f(x,y) = f(x)f(y)$; la función de densidad de probabilidad conjunta pude expresarse como producto de las funciones de densidad de probabilidad marginales), el coeficiente de correlación entre ellas es cero; pero si $r = 0$ esto no quiere decir que las dos variables sean independientes. En otras palabras, correlación cero no implica necesariamente independencia.

6.- Es una medida de asociación lineal o dependencia lineal únicamente y no tiene sentido utilizarlo para describir relaciones no lineales.

7.- Aunque mide la asociación lineal entre dos variables, no necesariamente implica una relación causa-efecto.

ESTIMACION DE INTERVALOS.

En estadística la confiabilidad de un estimador puntual se mide por su error estándar o por su varianza. En lugar de confiar en la estimación puntual únicamente, veremos cuál es la probabilidad de que el estimador puntual se encuentre dentro de cierto rango o intervalo alrededor del verdadero parámetro.

Para ser más específicos, supongamos que queremos saber qué tan cerca está $\hat{a}$ de a. Con tal fín, trataremos de encontrar dos números positivos, α y β, que están dentro de 0 y 1, de manera que la probabilidad de que el intervalo ($\hat{a} - \beta$, $\hat{a} + \beta$) contenga el valor verdadero de a sea $1-\alpha$. Simbólicamente:

$$Pr\,(\hat{a} - \beta \leq a \leq \hat{a} + \beta) = 1 - \alpha \quad \cdots I$$

Tal intervalo se conoce como " intervalo de confianza"; $1-\alpha$ se conoce como " coeficiente de confianza" y α ($0 < \alpha < 1$) se conoce como"nivel de significancia". Los límites del intervalo de confianza se conocen como " límites de confianza" o " valores críticos". En la práctica, α y $1-\alpha$ se expresan en forma de porcentaje como 100 α y 100 ($1-\alpha$) por ciento.

Es importante darnos cuenta de que:

1.- La ecuación I no expresa que la probabilidad de que a esté entre los límites dados sea $1-\alpha$. Dado que a, aunque desconocido, se supone un número fijo, tiene que estar dentro o fuera del intervalo. Lo que I afirma es que la probabilidad de construír un intervalo que contenga a a es $1-\alpha$.

2.- El intervalo I es aleatorio, es decir, varía de una muestra a otra porque se basa en $\hat{a}$ que es aleatorio.

3.- Como el intervalo de confianza es aleatorio, las afirmaciones probabilísticas relacionadas con él se entenderán en un sentido de largo plazo, es decir, en muestras repetidas. Más específicamente, I quiere decir que si en muestras repetidas se construyen intervalos como el anterior con base en una probabilidad $1-\alpha$, a largo plazo, en promedio, tales intervalos contendrán en $1-\alpha$ de los casos el verdadero valor del parámetro.

4.- El intervalo I es aleatorio mientras $\hat{a}$ sea desconocido, pero cuando tenemos una muestra específica y hemos obtenido un valor numérico específico de $\hat{a}$, el intervalo I deja de ser aleatorio; está fijo. En este caso ya no podemos hacer la afirmación probabilística I, es decir, no podemos afirmar que la probabilidad de que un intervalo fijo incluya el verdadero valor a sea $1-\alpha$. En una situación como esta, a se encuentra ya sea dentro del intervalo fijo o fuera de él, entonces la probabilidad es 1 ó 0.

CONSTRUCCION DE INTERVALOS DE CONFIANZA.

La regresión lineal normal supone que cada ξ_t está distribuído normalmente con:

$$\text{Media: } E(\xi_t) = 0$$

$$\text{Var: } E(\xi_t^2) = \sigma^2$$

$$\text{Cov}(\xi_i, \xi_j): E(\xi_i \xi_j) = 0 \quad i \neq j$$

que puede expresarse como:

$$\xi_t \sim N(0, \sigma^2)$$

De acuerdo con este supuesto, los estimadores mínimos cuadrados ordinarios tienen las propiedades:

1.- Son insesgados.

2.- Tienen varianza mínima. (Junto con (1), esto quiere decir que son insesgados con varianza mínima, estimadores "eficientes").

3.- Consistentes. Esto es, a medida que el tamaño de la muestra---- aumenta indefinidamente, los estimadores convergen al valor poblacional verdadero.

4.- Están normalmente distribuídos con:

$$E(\hat{b}) = b$$

$$Var(\hat{b}) = \sigma_b^2 = \frac{\sum_{t=1}^{T} X_t^2}{T \sum_{t=1}^{T} x_t^2} \sigma^2$$

$$E(\hat{a}) = a$$

$$Var(\hat{a}) : \sigma_a^2 = \frac{\sigma^2}{\sum_{t=1}^{T} x_t^2}$$

5.- $(T-2)\,\hat{\sigma}^2/\sigma^2 \sim \chi^2$ con T-2 grados de libertad.

6.- $(\hat{a}, \hat{b})$ están distribuídos independientemente de

7.- $(\hat{a}, \hat{b})$ tienen varianza mínima para todas las clases de es-- estimadores insesgados.

De esta manera, por ejemplo, la variable:

$$Z = \frac{\hat{a} - a}{\text{desv. est. } (\hat{a})} = \frac{(\hat{a} - a)\sqrt{\sum_{t=1}^{T} x_t^2}}{\sigma}$$

es una variable normal estandarizada. Parece, entonces, que pode-- mos utilizar la distribución normal para hacer afirmaciones probabilísticas acerca de $\hat{a}$ siempre que la varianza poblacional verdadera se conozca. Sin embargo, σ^2 se conoce muy rara vez, y en la práctica se determina con el estimador insesgado $\hat{\sigma}^2$, al sustituír:

$$t = \frac{(\hat{a} - a)\sqrt{\sum_{t=1}^{T} x_t^2}}{\hat{\sigma}}$$

refiriéndose ahora al error estándar estimado.

Si tenemos:

$$Z_1 = \frac{\hat{a} - a}{desv.est.(\hat{a})} = \frac{(\hat{a} - a)\sqrt{\sum_{t=1}^{T} x_t^2}}{\sigma}$$

$$y \ Z_2 = (T-2)\,\hat{\sigma}^2 / \sigma^2$$

$Z_1 \sim N(0,1)$ para poder conocer σ

$Z_2 \sim \chi^2$ con T-2 grados de libertad

entonces:

" Si Z_1 es una variable normal estandarizada [$Z_1 \sim N(0,1)$]y Z_2 es otra variable que sigue la distribución χ^2 con K grados de libertad y es independiente de Z_1 ,la variable definida por:

$$t = \frac{Z_1}{\sqrt{Z_2}/\sqrt{K}} = \frac{Z_1\sqrt{K}}{\sqrt{Z_2}}$$

sigue la distribución "t de Student" con K grados de libertad.

Por lo tanto, en lugar de usar una distribución normal podemos usar la distribución t para establecer el intervalo de confianza para a de la siguiente manera:

$$P_r\left(-t_{\alpha/2} \leq t \leq t_{\alpha/2}\right) = 1 - \alpha$$

donde el valor de t en medio de esta doble desigualdad es el valor de t dado por:

$$t = \frac{(\hat{a} - a)\sqrt{\sum_{t=1}^{T} x_t^2}}{\hat{\sigma}}$$

y donde $t_{\alpha/2}$ es el valor de t obtenido de la distribución t para el nivel de significancia $\alpha/2$ y T-2 grados de libertad.

Sustituyendo:

$$Pr\left[-t_{\alpha/2} \leq \frac{(\hat{a}-a)}{desv.\ est(\hat{a})} \leq t_{\alpha/2}\right] = 1-\alpha$$

reordenando:

$$Pr\left[\hat{a} - t_{\alpha/2}\, desv.\ est(\hat{a}) \leq a \leq \hat{a} + t_{\alpha/2}\, desv.\ est.(\hat{a})\right] = 1-\alpha$$

y esto nos proporciona un intervalo de confianza al 100($1-\alpha$) por ciento para a . Este mismo orden de ideas nos sirve para b.

INTERVALO DE CONFIANZA PARA σ^2.

Bajo el supuesto de normalidad:

$$\chi^2 = (T-2)\frac{\hat{\sigma}^2}{\sigma^2}$$

sigue la distribución χ^2 con T-2 grados de libertad, por ello podemos utilizar la distribución χ^2 para establecer el intervalo de confianza de σ^2 :

$$Pr\left(\chi^2_{1-\alpha/2} \leq \chi^2 \leq \chi^2_{\alpha/2}\right) = 1-\alpha$$

reemplazando χ^2 :

$$Pr\left[(T-2)\frac{\hat{\sigma}^2}{\chi^2_{\alpha/2}} \leq \sigma^2 \leq (T-2)\frac{\hat{\sigma}^2}{\chi^2_{1-\alpha/2}}\right] = 1-\alpha$$

da el intervalo de confianza para σ^2 de 100 ($1-\alpha$) %.

PRUEBA DE HIPOTESIS.

El problema de la prueba de hipótesis estadística puede expresarse en los siguientes términos: ¿ es un hecho u observación compatible con alguna hipótesis, o no lo es? La palabra " compatible" en este contexto quiere decir " suficientemente" cerca del valor hipotético, de suerte que nos lleve a aceptar la hipótesis que queremos analizar. De esta manera, si una teoría o una ex periencia a priori nos lleva a creer que el verdadero coeficiente es uno, se preguntará si el $\hat{a}$ observado es consistente con la hipótesis en juicio. Si es así, la hipótesis puede ser aceptada, si no, se rechaza.

En el lenguaje de la estadística, la hipótesis propuesta, o hipótesis sometida a análisis se conoce como hipótesis nula y se designa con el símbolo H_0 . La hipótesis nula suele probarse contra la hipótesis alterna que se designa con H_1 . La teoría de pruebas de hipótesis se ocupa de diseñar reglas o procedimientos que permitan decidir cuándo aceptar o rechazar una hipótesis nula y:

H_0 puede rechazarse si queda excluída del intervalo de confianza correspondiente a a .

2.2 MODELO DE REGRESION LINEAL MULTIPLE.

La regresión lineal múltiple es la extensión de la regresión lineal simple para tomar en cuenta a más de una variable independiente X. Es obviamente la técnica apropiada cuando queremos investigar los efectos de diferentes variables que actúan sobre Y al mismo tiempo. Aún cuando únicamente estemos interesados en los efectos de una sola variable sobre otra, es bueno incluír las otras variables que influyen sobre Y en un análisis de regresión múltiple por dos razones:

1.- Para reducir el error estocástico, y con ello, reducir la varianza residual. Esto hace que los intervalos de confianza sean más precisos.

2.- Para eliminar el sesgo que puede darse si simplemente ignoramos una variable que afecta a Y sustancialmente.

La manera más sencilla de expresar la relación que buscamos será:

$$E(Y_t) = a + bX_{1t} + cX_{2t}$$

donde los regresores X_{1t} y X_{2t} están medidos nuevamente como desviaciones de sus medias.

La diferencia entre el valor observado y esperado de Y_t es el término estocástico o de error ξ_t. De ahí que cualquier valor observado Y_t puede expresarse como su valor esperado más el término de disturbancia:

$$y_t = a + bx_{1t} + cx_{2t} + \xi_t$$

en donde ξ_t es considerado en la misma forma que en el apartado anterior.

De acuerdo a lo desarrollado en el capítulo anterior, tendremos que minimizar:

$$\sum_{t=1}^{T}(y_t - a - bx_{1t} - cx_{2t})^2$$

donde

$$\sum_{t=1}^{T}\varphi_t^2 = \sum_{t=1}^{T}(y_t - \hat{y}_t)^2 = \sum_{t=1}^{T}(y_t - \hat{a} - \hat{b}x_{1t} - \hat{c}x_{2t})^2$$

y las ecuaciones normales serán

$$\sum_{t=1}^{T}Y_t = T\hat{a} + \hat{b}\sum_{t=1}^{T}X_{1t} + \hat{c}\sum_{t=1}^{T}X_{2t}$$

$$\sum_{t=1}^{T}X_{1t}Y_t = \hat{a}\sum_{t=1}^{T}X_{1t} + \hat{b}\sum_{t=1}^{T}X_{1t}^2 + \hat{c}\sum_{t=1}^{T}X_{1t}X_{2t}$$

$$\sum_{t=1}^{T}X_{2t}Y_t = \hat{a}\sum_{t=1}^{T}X_{2t} + \hat{b}\sum_{t=1}^{T}X_{1t}X_{2t} + \hat{c}\sum_{t=1}^{T}X_{2t}^2$$

y si las consideramos en forma de desviaciones de sus medias

$$\hat{a} = \bar{Y} - \hat{b}\bar{X}_1 - \hat{c}\bar{X}_2$$

$$\hat{c} = \frac{\left(\sum_{t=1}^{T}x_{2t}y_t\right)\left(\sum_{t=1}^{T}x_{1t}^2\right) - \left(\sum_{t=1}^{T}x_{1t}y_t\right)\left(\sum_{t=1}^{T}x_{1t}x_{2t}\right)}{\left(\sum_{t=1}^{T}x_{1t}^2\right)\left(\sum_{t=1}^{T}x_{2t}^2\right) - \left(\sum_{t=1}^{T}x_{1t}x_{2t}\right)^2}$$

$$\hat{b} = \frac{\left(\sum_{t=1}^{T}x_{1t}y_t\right)\left(\sum_{t=1}^{T}x_{2t}^2\right) - \left(\sum_{t=1}^{T}x_{2t}y_t\right)\left(\sum_{t=1}^{T}x_{1t}x_{2t}\right)}{\left(\sum_{t=1}^{T}x_{1t}^2\right)\left(\sum_{t=1}^{T}x_{2t}^2\right) - \left(\sum_{t=1}^{T}x_{1t}x_{2t}\right)^2}$$

El estimador $\hat{b}$ mide el cambio en Y por variaciones unitarias en X_1 mientras se mantiene constante X_2. De la misma manera, $\hat{c}$ mide el cambio en Y por variaciones unitarias en X_2 mientras se mantiene constante X_1. $\hat{b}$ y $\hat{c}$ son coeficientes de regresión parcial. $\hat{a}$, $\hat{b}$ y $\hat{c}$ son estimadores insesgados por ser estimadores de mínimos cuadrados ordinarios (como quedó demostrado en el apartado anterior)

Ahora:

$$Var(\hat{b}) = \frac{\sum_{t=1}^{T} x_{2_t}^2}{\left(\sum_{t=1}^{T} x_{1_t}^2\right)\left(\sum_{t=1}^{T} x_{2_t}^2\right) - \left(\sum_{t=1}^{T} x_{1_t} x_{2_t}\right)^2}$$

$$Var\ \hat{c} = \frac{\sum_{t=1}^{T} x_{1_t}^2}{\left(\sum_{t=1}^{T} x_{1_t}^2\right)\left(\sum_{t=1}^{T} x_{2_t}^2\right) - \left(\sum_{t=1}^{T} x_{1_t} x_{2_t}\right)^2}$$

y el estimador para σ^2 será

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{T} \hat{y}_t^2}{T-3} \quad con: \sum_{t=1}^{T} \hat{y}_t^2 = \sum_{t=1}^{T} y_t^2 - \hat{b} \sum_{t=1}^{T} y_t x_{1_t} - \hat{c} \sum_{t=1}^{T} y_t x_{2_t}$$

Sus características son:

1.- Pasan por las medias $\bar{Y}$, $\bar{X}_1$, $\bar{X}_2$.

2.- El valor medio de Y_t estimado ($\hat{Y}_t$) es igual al valor medio de los valores observados Y_t, lo cual puede verse:

$$\hat{Y}_t = \hat{a} + \hat{b} X_{1_t} + \hat{c} X_{2_t}$$

$$\hat{Y}_t = (\bar{Y} - \hat{b}\bar{X}_1 - \hat{c}\bar{X}_2) + \hat{b}X_{1t} + \hat{c}X_{2t}$$

$$\hat{Y}_t = \bar{Y} + \hat{b}(X_{1t} - \bar{X}_1) + \hat{c}\,(X_{2t} - \bar{X}_2)$$

como desviaciones de sus medias

$$\hat{Y}_t = \bar{Y} + \hat{b}\,x_{1t} + \hat{c}\,x_{2t}$$

sumando sobre todos los valores y dividiendo por T

$$\bar{\hat{Y}} = \bar{Y}$$

porque $\sum_{t=1}^{T} x_{1t} = \sum_{t=1}^{T} x_{2t} = 0$

3.- $\sum_{t=1}^{T} \xi_t = 0$

4.- Los residuos ξ_t no están correlacionados con $\hat{Y}_t$, es decir, $\sum_{t=1}^{T} \xi_t \hat{Y}_t = 0$.

5.- Los residuos ξ_t no están correlacionados con X_{1t} ni con X_{2t} esto es $\sum_{t=1}^{T} \xi_t X_{1t} = \sum_{t=1}^{T} \xi_t X_{2t} = 0$.

6.- Son estimadores insesgados.

R^2 estará dado por:

$$R^2 = \frac{\hat{b} \sum_{t=1}^{T} y_t x_{1t} + \hat{c} \sum_{t=1}^{T} y_t x_{2t}}{\sum_{t=1}^{T} y_t^2}$$

$R^2_{\bar{x}}$ al igual que r^2 está comprendido entre 0 y 1. Si es 1 significa que la línea de regresión ajustada explica el 100% de la variación de Y. Por otro lado, si es cero, el modelo no explica nada de las variaciones en Y. Regularmente R^2 estará entre los dos extremos y se dice que el ajuste es "mejor" mientras más cerca de 1 esté R^2.

De la misma manera que en el caso de dos variables, existe el " coeficiente de correlación" que será múltiple, R , pero siempre será positivo. R es una medida de asociación entre Y y todas las variables explicatorias conjuntamente. En la práctica, se considera más significativo R^2 y a R se le dá poca importancia.

2.3 MODELOS LINEALES CON ERRORES EN LAS VARIABLES.

En los apartados anteriores, hemos admitido que las variables x_t eran observadas sin error, pero; aunque esta hipótesis quede justificada frecuentemente en la realidad, es importante ver qué ocurre cuando no sucede así, ya que las variables de un modelo no siempre son observables directamente y frecuentemente se sustituyen por otras muy relacionadas con ellas (aunque no sea en forma rigurosa).

En estas circunstancias, es conveniente modificar las hipótesis básicas del modelo que habíamos considerado.

Tomemos el modelo de regresión y supongamos que existe una relación lineal entre las variables exógenas x_t y las endógenas y_t , entonces:

$$y_t = A x_t + \varepsilon_t$$

Pero las variables x_{it} , y_{it} no son observadas exactamente, sino que solamente se tienen las evaluaciones aproximadas x_{it}^* , y_{it}^* , donde:

$$y_t^* = y_t + \xi_t \qquad\qquad x_t^* = x_t + \zeta_t$$

en el cual ξ_t y ζ_t son los vectores de los errores en las variables (que se suponen aleatorios al igual que en los casos anteriores).

Entonces, al substituír x_t por x_t^* , y_t por y_t^*, se obtiene:

$$y_t^* = A x_t^* + (\varepsilon_t - \xi_t + A\zeta_t)$$

es decir, los errores ahora serán:

$$(\xi_t - \xi_t + A\,\zeta_t)$$

si hubiera independencia en estos errores podríamos trabajar con mínimos cuadrados*. A este respecto, se admite generalmente que los errores tienen una esperanza matemática nula y que son independientes de los valores tomados por x_t e y_t ; lo cual implica que también sean independientes de los errores en las ecuaciones, ξ_t . En estas condiciones, $\xi_t - \xi_t + A\,\zeta_t$ son independientes de x_t pero no de x_t^* . A un valor elevado de x_t^* corresponde frecuentemente un valor elevado de $\xi_t - \xi_t + A\,\zeta_t$. Esto indica que es indispensable la independencia de los errores con respecto a x_t^* para poder utilizar las teorías expuestas anteriormente sobre los modelos de regresión.

Si las variables exógenas se miden sin error, tendremos que $\zeta_t = 0$ y $\xi_t^* - \xi_t$ entonces $x_t^* = x_t$ y los errores serán independientes, por lo que las propiedades estudiadas anteriormente permanecen válidas aunque existan errores de medida en las variables endógenas, a condición de que no existan tales errores en las variables exógenas, y que los de las variables endógenas obedezcan a las mismas hipótesis que los errores en las ecuaciones.

Como desafortunadamente las variables exógenas a veces se ven afectadas por errores importantes de medida, necesitaremos algunos cambios para poder medirlas:
Supondremos que ξ_{it} y ζ_{it} tienen esperanzas matemáticas nulas independientemente de los valores tomados por x_t e y_t . Sin embargo, en la práctica, puede suceder que las observaciones se vean afectadas por errores sistemáticos.

* Malinvaud, Edmond. " Instrumentos Estadísticos Para el Análisis Econométrico". Ed. Ariel., pág. 366.

Para poder allanar esta dificultad, tomemos el modelo:

$$y_t = a x_t + b + \xi_t$$

en donde:

$$E(\xi_t / x_t ; \xi_t) = u_1 y_t + u_0$$

$$E(\zeta_t / x_t ; \xi_t) = w_1 x_t + w_0$$

de esa manera, los errores sistemáticos son funciones lineales de las variables a las que afectan.

Podremos entonces definir las nuevas variables $\hat{x}_t$, $\hat{y}_t$ y los nuevos errores $\hat{\xi}_t$, $\hat{\zeta}_t$ donde:

$$y_t^* = \hat{y}_t + \hat{\xi}_t \qquad x_t^* = \hat{x}_t + \hat{\zeta}_t$$

con

$$E(\hat{\xi}_t / \hat{y}_t ; \xi_t) = 0 \qquad E(\hat{\zeta}_t / \hat{x}_t ; \xi_t) = 0$$

será suficiente entonces la transformación:

$$\hat{y}_t = (1 + u_1) y_t + u_0$$

$$\hat{x}_t = (1 + w_1) x_t + w_0$$

Esto es, las nuevas variables serán la suma de las variables verdaderas y de los errores sistemáticos en observaciones correspondientes, y siguen un modelo de regresión simple

$$\hat{y}_t = \hat{a}\,\hat{x}_t + \hat{b} + \xi_t$$

donde

$$\hat{a} = a \, \frac{1+u_1}{1+\omega_1}$$

$$\hat{b} = b(1+u_1) + u_0 - w_0 \, \frac{a\,(1+u_1)}{1+\omega_1}$$

2.4 TRANSFORMACION DE VARIABLES.

En los apartados anteriores hemos tomado como base que el modelo que describen los datos es lineal en los parámetros. Debido a que algunas veces no ocurre así, para poder llevar a cabo los desarrollos precedentes, el análisis tiene que realizarse entonces sobre variables transformadas. La necesidad de transformar los datos surge porque la variable original, o el modelo en términos de la variable original, viola uno o más de los supuestos considerados. Las hipótesis que más frecuentemente se violan son las referentes a la linealidad en el modelo y la constancia de la varianza en los errores.

Para satisfacer las suposiciones del modelo de regresión que hemos venido considerando, en vez de trabajar con las variables originales, algunas veces es necesaria una transformación en ellas. Esta necesidad puede darse por diferentes motivos:

1) La teoría puede ser muy específica al considerar una relación no lineal entre las variables. Una transformación puede hacer que la relación cambie a lineal. Por ejemplo, la "función de producción" de Cobb-Douglas puede escribirse como:

$$Y = A X_1^{\alpha} X_2^{\beta}$$

donde A, α y β son constantes positivas y

Y = producto

X_1 = trabajo

X_2 = capital

visto de esta manera, no podemos utilizar nuestros modelos de regresión lineal directamente; pero, si tomamos logaritmos:

$$\log Y = \log A + \alpha \log X_1 + \beta \log X_2$$

que ahora nos describe una función lineal y podremos usar los métodos que ya habíamos considerado.

2) La variable dependiente y analizada puede tener una distribución de probabilidad cuya varianza está relacionada con la media, es decir, si la media está relacionada con el valor de la variable

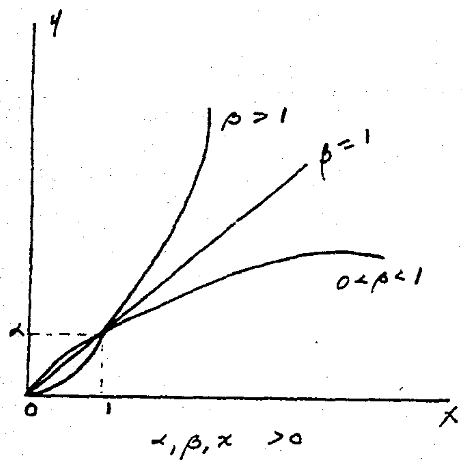
independiente X , la varianza de Y cambiará con X , por lo cual la varianza no será constante y, regularmente la distribución de Y será no normal bajo estas condiciones. La no normalidad invalida las pruebas de significancia. La varianza no constante de los errores producirá estimadores que son insesgados pero que no serán los mejores en el sentido de precisión. Con todo esto, regularmente se transforman los datos para asegurar normalidad y constancia de la varianza en los errores.

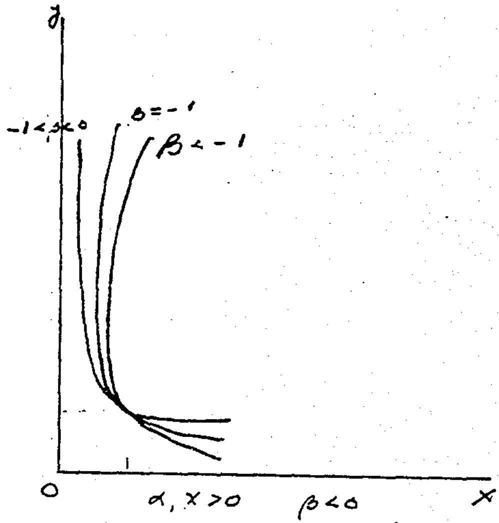
En la práctica, se eligen las transformaciones para asegurar la constancia de la varianza (" transformaciones que estabilizan la varianza"). Afortunadamente, es una coincidencia que las transformaciones que estabilizan la varianza también son buenas transformaciones normalizadoras.

Ejemplos de estas transformaciones son:

Función	Transformación	Forma lineal
$Y = \alpha X^{\beta}$	$Y' = \ln Y$, $X' = \ln X$	$Y' = \ln \alpha + \beta X'$

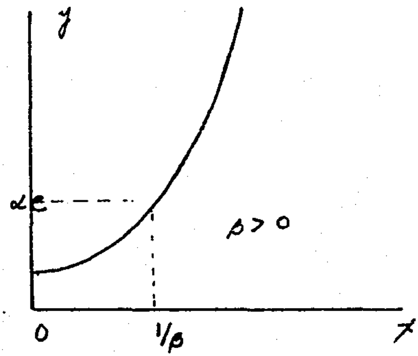
Gráficas:

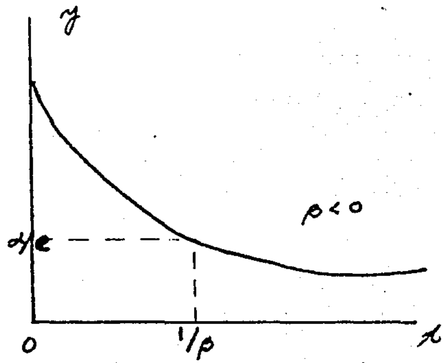




Función	Transformación	Forma Lineal
$y = \alpha e^{\beta x}$	$y' = \ln y$	$y' = \ln \alpha + \beta X$

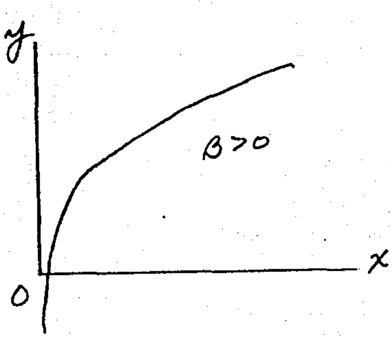
Gráficas:

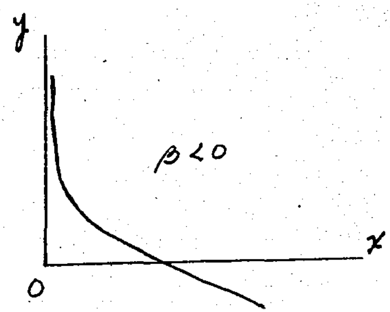




Función	Transformación	Forma Lineal
$y = \alpha + \beta \ln X$	$x' = \ln X$	$y = \alpha + \beta X'$

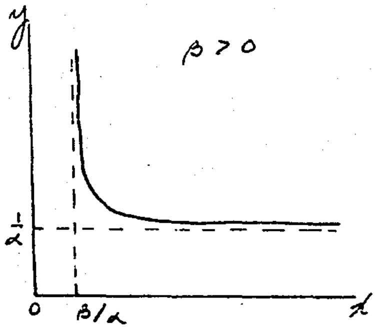
Gráficas:

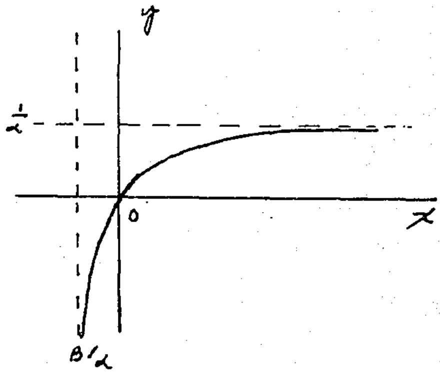




Función	Transformación	Forma Lineal
$Y = \frac{X}{\alpha X - \beta}$	$Y' = \frac{1}{Y}$, $X' = \frac{1}{X}$	$Y' = \alpha - \beta X'$

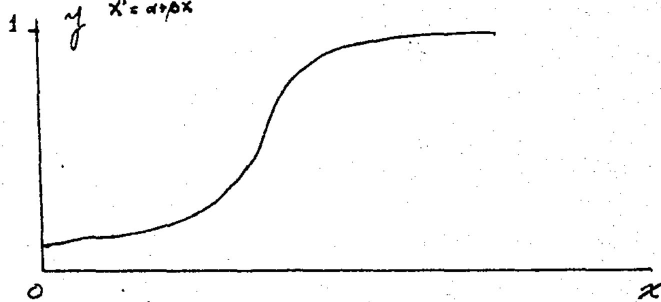
Gráficas:





Función	Transformación	Forma Lineal
$Y = \frac{e^{\alpha+\beta X}}{1+e^{\alpha+\beta X}}$	$Y' = \ln\left(\frac{Y}{1-Y}\right)$ $X' = \alpha+\beta X$	$Y' = \alpha + \beta X'$

Gráfica:



2.5 COMPARACION ENTRE ESTOS MODELOS.

Como fue mencionado en su oportunidad, la regresión - múltiple es considerada sencillamente como una extensión de la regresión lineal simple pero tiene ventajas importantes, que son: reducir los errores estocásticos y eliminar el sesgo que resulte de no considerar variables que pudieran afectar el comportamiento de la variable en estudio.

Debemos estar conscientes de que al utilizar un modelo de regresión de la forma

$$y = ax + b$$

estamos suponiendo constantes todas las otras variables que podrían verse involucradas de alguna manera.

El número de regresores (variables independientes) tomados en cuenta para la ecuación del modelo de regresión correspondiente dependerá directamente del fenómeno en estudio.

CAPITULO 3

METODOS ESTADISTICOS PARA MODELOS AUTORREGRESIVOS.

3.0 EL PAPEL DE LOS REZAGOS EN LA ECONOMIA.

En el análisis económico, la dependencia de la variable y con respecto a otras variables x no es instantánea generalmente. Es frecuente ver que y responde a x después de cierto tiempo; a ese período se le da el nombre de rezago o retraso.

Cuando en el análisis de regresión (en series de tiempo)* el modelo incluye no solo los valores comunes de las variables explicatorias, sino también sus valores rezagados, se dice que tenemos un modelo de rezagos distribuídos; pero, si el modelo incluye uno o más valores rezagados de la variable dependiente (entre las variables explicatorias) se dice que tenemos un modelo autorregresivo.

Existen tres razones principales por las cuales ocurren estos rezagos:

1.- Psicológicas.- Debido a la fuerza de costumbre, ante una baja en el precio o un aumento en el ingreso, la gente no cambia inmediatamente sus hábitos de consumo; además del hábito y de que el proceso de cambio puede implicar algunas pérdidas (dependiendo de que el cambio se considere como permanente o transitorio).

2.- Tecnológicas.- Si suponemos que el precio del capital con relación al del trabajo decrece, volviéndose económicamente posible la sustitución de trabajo por capital; las empresas no se apresurarán a sustituír trabajo por capital, especialmente si creen que después de la caída temporal, el precio del capital supera sus niveles anteriores.

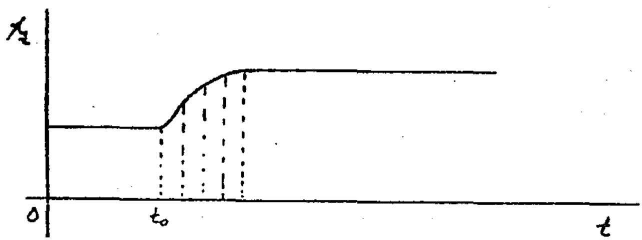
3.- Institucionales.- Las obligaciones contractuales pueden impedir a las empresas cambiar de un tipo de trabajo a otro o de una materia prima a otra.

Por estas razones, los rezagos ocupan un lugar central en la economía, que se reflejan en la metodología económica de corto y largo plazo.

* Una serie temporal está formada por un conjunto de valores obtenidos de observaciones referentes al mismo fenómeno, realizados en una sucesión de momentos de tiempo, normalmente a intervalos iguales.

3.1 ESTIMACION DIRECTA POR REGRESIONES LINEALES.

Las leyes de comportamiento que intervienen en los modelos económicos describen cómo una o más variables endógenas se determinan a partir de los valores tomados por ciertas variables exógenas. Es poco frecuente que la dependencia entre variables exógenas y endógenas sea instantánea; sería entonces inexacto suponer que la adaptación de los comportamientos se efectúa totalmente en un instante, después de cierto plazo fijo. Si las variables exógenas permanecieran constantes durante un largo período, y luego, sufrieran una fuerte modificación (momento t_o en la gráfica), para no variar después, no se observará una variación drástica de cada variable endógena en el momento $t_o + \theta$, sino más bien un paso progresivo de un valor de equilibrio a otro; esto puede representarse como:



Los modelos realistas para las relaciones económicas requieren frecuentemente de la consideración de variables con retardos en las variables explicatorias.

Si suponemos un modelo de la forma:

$$y_t = \alpha + \beta x_{t-\theta} + \xi_t$$

se asume que el efecto de $x_{t-\theta}$ aparece solamente en el periodo t y se completa dentro de ese período. Pero es más frecuente conside_ rar que el efecto de una variable está distribuído sobre diferentes períodos. Entonces podemos representar este modelo como:

$$y_t = \alpha_0 x_t + \alpha_1 x_{t-1} + \cdots + \alpha_h x_{t-h} + \beta + \xi_t$$

siendo notorio de esta manera que queda comprendido en la categoría general de los modelos de regresión que fueron desarrollados en el capítulo anterior.

3.2 METODO DE MINIMOS CUADRADOS.

Supongamos que existe una sola variable endógena y_t y una sola variable exógena x_t, tomando el valor x° antes de t_o y el valor x' a partir de t_o. Supongamos, además, que la relación entre x_t e y_t sea lineal y exacta, de tal forma que el valor de equi-librio de y_t sea $\alpha x^o + \beta$ antes de t_o y $\alpha x' + \beta$ cuando se ha finalizado el ajuste al nuevo valor de x_t. En el intervalo, podemos admitir entonces que el valor de y_t depende linealmente de x° y de x' con coeficientes de función solamente de la diferencia de $t-t_o$. Podemos escribir, entonces:

$$y_{t_o+\tau} = \alpha d_\tau x' + \alpha (1-d_\tau) x^o + b \qquad \tau \geq 0$$

y, haciendo

$$c_o = d_o \qquad c_\tau = d_\tau - d_{\tau-1} \qquad \tau > 0$$

resulta:

$$y_{t_o+\tau} = \alpha (c_o + c_1 + \ldots + c_\tau) x' + \alpha (c_{\tau+1} + c_{\tau+2} + \ldots + c_h) x^o + \beta$$

si admitimos que d_τ es igual a 1 a partir de h, por lo cual $c_\tau = 0$ para $\tau > h$.

En general, si la variable exógena x_t sigue una evolución cualquiera, el valor de la variable endógena y_t puede considerarse como una función lineal de los valores tomados anteriormente por x_t, o sea:

$$y_t = \alpha \sum_{\tau=0}^{h} c_\tau x_{t-\tau} + \beta$$

con

$$\sum_{\tau=0}^{h} c_\tau = 1$$

de donde

$$y_t = \alpha_o x_1 + \alpha_1 x_{t-1} + \ldots + \alpha_h x_{t-h} + \beta$$

sin condición alguna sobre los coeficientes α_τ.

Al considerar los errores aleatorios:

$$y_t = \alpha_o x_1 + \alpha_1 x_{t-1} + \ldots + \alpha_h x_{t-h} + \beta + \xi_t$$

Regularmente se tiene a priori algún conocimiento sobre los coeficientes $\alpha_0, \alpha_1, \ldots, \alpha_h$. La influencia del valor tomado por una variable exógena durante el período t_0 debe ser especialmente importante para los valores de la variable endógena en el curso de los períodos que siguen inmediatamente a t_0. En muchos casos, α_0 podrá ser mayor que todos, o se podrá admitir que $\alpha_\tau < \alpha_{\tau-1}$ para todo τ entre 1 y h.

Como casi siempre los datos no permiten una determinación muy precisa de los coeficientes del modelo, y dado que nuestros conocimientos a priori sobre la sucesión de los coeficientes son muy reducidos, la elección de las hipótesis sobre α_τ debe inspirarse en consideraciones de simplicidad.

Las hipótesis más comúnmente tomadas son:

- La " distribución exponencial"

$$c_\tau = (1-c)\,c^\tau$$

que representa un decrecimiento geométrico y elimina el problema de la elección de retardo máximo h. Tiene el inconveniente de no poder aplicarse al conjunto de la sucesión más que si las α_τ decrecen desde los primeros términos.

Cuando no hay más que una variable exógena, esta distribución conduce a un modelo en el que la forma autorregresiva es particularmente simple:

$$y_t = d \sum_{\tau=0}^{\infty} c^\tau x_{t-\tau} + b + \xi_t$$

que al hacer

$$e = b(1-c) \qquad \eta_t = \xi_t - c\,\xi_{t-1}$$

puede expresarse como:

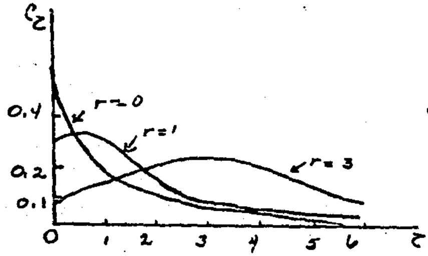
$$y_t = c\,y_{t-1} + d\,x_t + e + \eta_t$$

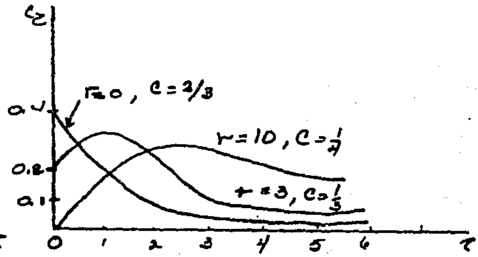
- La"distribución de Pascal".

Aquí se define a c_τ de la forma:

$$c_\tau = (1-c)^{r+1} C^{r}_{r+\tau} c^{\tau}$$

donde se hace intervenir solamente a los parámetros c y r, y nos conduce nuevamente a la distribución exponencial cuando r=0 y da, para los demás valores de r, una gran variedad de evoluciones posibles*:





Esto nos conduce a que, en vez de considerar la forma directa del modelo con rezagos escalonados, se puede pensar en estimar los coeficientes desconocidos sobre la forma autorregresiva.

Podemos entonces escribir la forma autorregresiva como:

$$y_t = c\, y_{t-1} + d\, x_t + e + \delta_t$$

si las α_τ obedecen a una distribución exponencial $\alpha_\tau = d c^\tau$; y

$$y_t = 2c\, y_{t-1} - c^2 y_{t-2} + d\, x_t + e + \delta_t$$

si las α_τ obedecen a una distribución de Pascal de parámetro r=1.

Para la distribución exponencial, la estimación por ajuste mínimo cuadrática puede llevarse a cabo aplicando las formas de regresión desarrolladas en el capítulo anterior. Para la distribución de Pascal, es necesario tener un desarrollo distinto.

*MALINVAUD, Edmond. "Instrumentos Estadísticos de la Econometría" pág. 540.

Tendremos entonces que minimizar:

$$\sum_{t=1}^{T}\left(y_t - 2c\,y_{t-1} + c^2 y_{t-2} - d\,x_t - e\right)^2$$

y al derivarla resultará:

$$\sum_{t=1}^{T}\left(y_{t-1} - c\,y_{t-2}\right)\left(y_t - 2c\,y_{t-1} + c^2 y_{t-2} - d\,x_t - e\right) = 0 \quad \ldots\ldots (1)$$

$$\sum_{t=1}^{T} x_t\left(y_t - 2c\,y_{t-1} + c^2 y_{t-2} - d\,x_t - e\right) = 0 \quad \ldots\ldots (2)$$

$$\sum_{t=1}^{T}\left(y_t - 2c\,y_{t-1} + c^2 y_{t-2} - d\,x_t - e\right) = 0 \quad \ldots\ldots (3)$$

y si $\bar{y}_{-1} = \frac{1}{T}\sum_{t=1}^{T} y_{t-1}$ $\qquad \bar{y}_{-2} = \frac{1}{T}\sum_{t=1}^{T} y_{t-2}$

$$-c^3[x^2_{-2}] + 3[x_{-1}x_{-2}] + cd[x_{-2}z] - c([x\,x_{-2}] + 2[x^2_{-1}]) - d[x_{-1}z] + [x\,x_{-1}] = 0$$

$$c^2[x_{-2}z] - 2c[x_{-1}z] - d[z^2] + [xz] = 0$$

$$c^2\bar{x}_{-2} - 2c\bar{x}_{-1} - d\bar{z} - e + \bar{x} = 0$$

una vez determinados c y d, la tercera ecuación dará el valor de e. La segunda da el valor de d, cuando se ha logrado determinar c; y c se obtiene eliminando d de las dos primeras ecuaciones y resolviendo la ecuación de tercer grado resultante.

Solamente importan las raíces inferiores a 1, ya que "Se sabe que existe una raíz que tiende en probabilidad hacia el valor real c, que tiene un módulo inferior a 1. Pero no es seguro, a priori, que en una muestra cualquiera, la ecuación de tercer grado tenga una solución comprendida entre cero y uno. Si no es así, el sistema no sirve" *

*MALINVAUD, Edmond. Op. Cit. Pág 540.

3.3 AJUSTE DE MINIMOS CUADRADOS EN MUESTRAS PEQUEÑAS.

En Econometría se utilizan frecuentemente series estadísticas cortas, por eso es interesante conocer más exactamente las distribuciones de las estimaciones obtenidas por el método de mínimos cuadrados para las muestras pequeñas. Desafortunadamente, estas distribuciones son muy complicadas y no ha sido posible establecerlas más que para casos muy particulares.

Consideremos el más sencillo de los modelos:

$$x_t = b\,x_{t-1} + \xi_t \qquad |b| < 1$$

utilizando mínimos cuadrados:

$$U = \sum_{t=2}^{T} \left(x_t - b\,x_{t-1} \right)^2 \quad \text{debe ser mínima.}$$

$$\frac{\partial U}{\partial b} = (2)\left(-\sum_{t=2}^{T} x_{t-1}\right) \sum_{t=2}^{T} \left(x_t - b\,x_{t-1}\right) = 0$$

$$= \sum_{t=2}^{T} x_{t-1} \left(\sum_{t=2}^{T} x_t - \sum_{t=2}^{T} b\,x_{t-1} \right) = 0$$

$$= \sum_{t=2}^{T} x_{t-1}\,x_t - \sum_{t=2}^{T} b\,x_{t-1}^2 = 0$$

$$\therefore \quad \hat{b} = \frac{\sum_{t=2}^{T} x_t\,x_{t-1}}{\sum_{t=2}^{T} x_{t-1}^2}$$

Su esperanza matemática de 1/T, viene dada por:

$$E(\hat{b}) \approx b\left(1 - \frac{2}{T}\right)$$

Cuando T crece indefinidamente, el error sistemático $-\frac{2b}{T}$ se hace infinitamente pequeño en relación con la desviación de $\hat{b}$, que es:

$$\sqrt{\frac{1-b^2}{T}}$$

CAPITULO 4

METODOS ESTADISTICOS PARA MODELOS DE ECUACIONES SIMULTANEAS.

4.0 PROBLEMAS DE IDENTIFICACION DE OFERTA Y DEMANDA.

Desde la 1° guerra mundial se intentaron determinar empíricamente las leyes de la demanda relacionando la cantidad pedida con el precio del producto. Para ello, se utilizaron las series estadísticas que arrojaban las cantidades consumidas y los precios en un mismo país y durante diferentes períodos. Algunos de los resultados obtenidos parecían sorprendentes.

En 1927 se demostró que los ajustes realizados no hacían aparecer necesariamente las leyes de la demanda. Si volvemos al ejemplo del capítulo 1, el precio y la cantidad consumida de un producto se determinan simultáneamente después de una confrontación entre la demanda de los consumidores y la oferta de los productores.

Indudablemente, este esquema teórico corresponde sólo aproximadamente a los hechos. Además del precio, pueden actuar otros factores sobre las ofertas y las demandas, pero si el efecto de estos factores es débil de alguna manera, el conjunto de datos observados para una sucesión de períodos deberá situarse cerca del punto de equilibrio, formando una nube que no puede darnos información acerca de la forma de la curva de la demanda.

Tomando un modelo aleatorio lineal con ecuaciones simultáneas:

$$\begin{cases} q_t = b_1 p_t + f_1 + \eta_{1t} \\ q_t = b_2 p_t + f_2 + \eta_{2t} \end{cases}$$

donde p_t = precio, q_t = cantidades.
b_1, b_2, f_1, f_2 son coeficientes constantes desconocidos y η_{1t}, η_{2t} son los términos aleatorios de medias nulas (que traducen el efecto de factores no identificables).

si: $b_1 \neq b_2$

$$q_{20} = \frac{f_1 - f_2}{b_2 - b_1} \qquad a_{10} = \frac{b_2 f_1 - b_1 f_2}{b_2 - b_1}$$

$$\xi_{1t} = \frac{b_2 \eta_{1t} - b_1 \eta_{2t}}{b_2 - b_1} \qquad \xi_{2t} = \frac{\eta_{1t} - \eta_{2t}}{b_2 - b_1}$$

tendremos un sistema de la forma:

$$\begin{cases} q_t = a_{10} + \xi_{1t} \\ p_t = a_{20} + \xi_{2t} \end{cases}$$

Estas dos ecuaciones nos indican claramente que la cantidad y el precio deben tomar valores distribuídos al azar alrededor de q_{10} y q_{20}, que a su vez definen las coordenadas del punto de intersección de la oferta media y la demanda media.

Una muestra de observaciones sobre p_t y q_t (t=1,2,...T) permite una determinación más o menos precisa de q_{10} y a_{20}, y de la ley de distribución del par (ξ_{1t}, ξ_{2t}).

Pero esto no basta para determinar los coeficientes b_1, f_1, b_2 y f_2, y la ley de distribución de (η_{1t}, η_{2t}), ya que las relaciones q_{10} y q_{20} no pueden resolverse de manera única con respecto a b_1, f_1, b_2 y f_2.

De esa manera, las relaciones estrucrurales oferta y demanda no son identificables.

Puede también ocurrir que la oferta y la demanda dependan ambos de factores distintos que el precio. Sobre una gráfica de dos dimensiones, la nube de puntos no tiene ya ninguna forma determinada; pero este hecho, no impide una identificación de las relaciones de oferta y demanda. Algunos de los factores que actúan sobre la oferta pueden dejar la demanda invariable, mientras que ciertos factores localizados que actúan sobre la demanda pueden -

dejar la oferta invariable; con esto, se pueden identificar las dos leyes.

Los obstáculos más frecuentes para la comprensión del problema son:

Para construír una tabla de demanda hacemos variar el precio del bien y observamos lo que ocurriría con la cantidad comprada en un periodo cualquiera de tiempo, dentro del cual no varía ningún otro dato que pudiera enturbiar nuestro experimento: Esto es, al hacer variar el precio no podemos variar al mismo tiempo los ingresos de las familias, ni el precio de un producto sustitutivo, ni cualquier otra cosa que pudiera desplazar la curva de demanda.

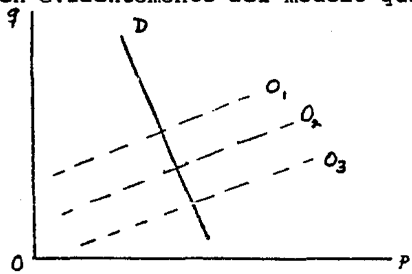
La segunda dificultad es más sutil y, aunque no es tan probable que se presente, es más difícil de resolver que la anterior pues" Si la cantidad vendida por una persona es exactamente la cantidad comprada por otra; ¿ Cómo es que la igualdad entre la oferta y la demanda determina el precio de equilibrio?"

La cantidad comprada ha de ser siempre igual a la cantidad vendida, cualquiera que sea el precio. Independientemente de que el mercado esté o no en equilibrio, el estadístico que se dedique a reunir los datos de las cantidades compradas y vendidas encontrará siempre la misma cifra por ambos lados, pues no son sino aspectos distintos de un mismo fenómeno.

Lo que en realidad interesa es que " el precio al cual la cantidad de consumidores están dispuestos a seguir comprando" sea igual exactamente a " la cantidad que los productores estén dispuestos a seguir vendiendo". A este precio no habrá tendencia a subir ni a bajar.

4.1 ESTIMACION DE LAS LEYES DE DEMANDA.

Si consideramos un caso sencillo que permite la identificación de la demanda, aquel cuya nube de puntos se distribuye alrededor de una curva de demanda estable; por ejemplo, para gran número de productos agrícolas se observa efectivamente una relación inversa entre precio y cantidad consumida. Dicha relación es debida a los consumidores, ya que la cosecha, que varía de un año a otro, debe venderse a un precio tal que los consumidores están dispuestos a absorberla en su totalidad. El mercado comporta entonces una demanda estable (curva D) y una oferta que se desplazará hacia arriba o hacia abajo según la abundancia de la cosecha (curvas O_1, O_2, O_3). La nube de puntos observados en una sucesión de años se reparte alrededor de la curva de demanda e informa perfectamente acerca de su posición. Las propiedades de los diferentes estimadores posibles dependen evidentemente del modelo que se juzgue adecuado.



Considerando que las variaciones de la oferta reflejan las de las condiciones meteorológicas (ya que la cosecha se vé afectada directamente por ellas), podemos admitir un índice θ_t que mida el estado de estas condiciones en el año t. Esto queda representado por:

$$q_t = b_1 p_t + c_1 \theta_t + f_1 + \eta_{1t}$$
$$q_t = b_2 p_t + f_2 + \eta_{2t}$$

donde:

p_t = precio en el año t

q_t = cantidad en el año t

η_{1t}, η_{2t} = términos aleatorios.

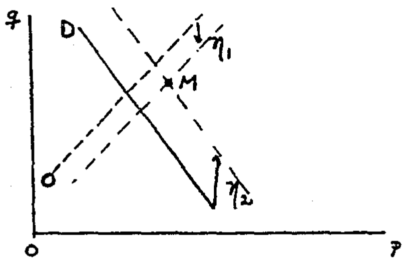
en el cual las dos relaciones representan respectivamente la oferta y la demanda. Para determinar el par (p_t , q_t) en un precio para el que θ_t tiene un valor dado, trazaremos la función de oferta media (la esperanza)

$$q_t = b_1 p_t + c_1 \theta_t + f_1$$

y la función de demanda media:

$$q_t = b_2 p_t + f_2 .$$

en un diagrama con p en abscisas y q en ordenadas:



En ausencia de factores no identificados, la cantidad q_t y el precio p_t tomarían los valores definidos por el punto de intersección de O y D (como en el ejemplo del capítulo 1), pero, debido a η_{1t} y η_{2t}, los valores observados corresponden al punto de intersección M de las dos curvas deducidas respectivamente de O y D mediante desplazamientos verticales de longitudes η_{1t} y η_{2t} . Si θ_t tomara el mismo valor para todas las observaciones, los puntos observados quedarían repartidos alrededor de la intersección de O y D, correspondiendo - una perturbación η_2 positiva a valores elevados de p y de q; y una perturbación positiva η_1 a un valor elevado de q, pero pequeño de p.

4.2 ESTIMACIONES DE FUNCIONES DE PRODUCCION.

Básicamente, la función de producción es, por una parte, una relación técnica entre cantidades de varios factores de producción o insumos, y por la otra, la cantidad de producto o producción que generan. Estos términos sugieren que la producción es un proceso físico mediante el cual los insumos se transforman en productos; sin embargo, la función de producción no se limita a las variaciones posibles dentro de un proceso técnico de producción sino que cubre toda la gama de métodos de producción concebibles. La teoría económica se interesa por factores de producción muy abstractos como trabajo y capital.

La función de producción resume las restricciones técnicas que delimitan el comportamiento económico en la variación de las proporciones relativas de los distintos insumos. El principal interés de las teorías económicas se centra en la substitución de un factor de producción por otro.

La función de producción de una empresa representa las limitaciones técnicas que se imponen a esta empresa. Determina las cantidades de producto que pueden obtenerse a partir de cualquier combinación de factores puestos en juego.

Si consideramos la función de producción:

$$X = f(V_1, V_2, \ldots, V_T) \qquad \ldots\ldots\ldots \qquad (1)$$

en donde X es el volúmen de producción y V_j, $j=1,2,\ldots T$, denota cantidades o volúmenes de varios factores productivos, en esta etapa no es necesario precisar la forma de la función de producción ni introducir todavía un término de perturbación. Ignoramos también todos los problemas de definición y medición dando por hecho que tenemos suficientes observaciones de X y de V_j así como los precios correspondientes que denotaremos por P, P_j.

Lo primero que es necesario establecer es que sería un error leer (1) como si V_j fuera una variable independiente exógena determinante de X. La producción no es espontánea y ciertamente no consiste en juntar algunos insumos dados arbitrariamen te y en esperar pacientemente el resultado. Por el contrario, la determinación conjunta de V_j, y por lo tanto de X, está gobernada por consideraciones económicas que debemos incorporar al modelo. El enfoque común consiste en suponer que el productor determina V_j de modo tal que maximice los beneficios W donde:

$$W = PX - \sum_{j=1}^{T} P_j V_j$$

teniendo P y P_j como variables predeterminadas.

Si existe un máximo definido de W, el productor establecerá el valor de los insumos V_j a un nivel que satisfaga:

$$\frac{\partial W}{\partial V_j} = 0 \qquad j = 1, 2, \ldots, T$$

y estos valores están determinados por la solución de este sistema de T ecuaciones simultáneas, que se convierten en:

$$^{*}P \frac{\partial f}{\partial V_j} = P_j \qquad j = 1, 2, \ldots, T$$

y en esta forma se conocen con el nombre de " condiciones de productividad marginal".

Estas condiciones de productividad marginal se utilizan para determinar las cantidades del insumo como función de precios dados. En este caso, las ecuaciones no pueden tomarse una por una, lo que dificulta su utilización.

Si nos referimos al caso en donde exista solamente un producto Q, y dos factores, cuyas cantidades serán N y K (en la mayor parte de las veces, N representa el factor trabajo y K el capital) la función de producción se escribe entonces:

$$Q = f(N, K)$$

que fuera hasta hace poco, la base exclusiva del estudio de funciones de producción, y que forman entonces una función de producción Cobb-Douglas:

$$Q = A N^{\alpha} K^{\beta}$$

que, como vimos en el capítulo 2, puede ser sometida a una transformación para convertirla en una relación lineal.

*En el sentido económico.

Por supuesto, estos no son los únicos criterios para --- analizar funciones de producción. También en la Organización de las Naciones Unidas (O. N. U.) se ha desarrollado una investigación - para ello aplicándolo a industrias y basándose en funciones de producción que se expresen en forma polinómica cuya estructura matemática es de un polinomio de grado menor o igual a tres. Para P insumos la forma de escribirlos es:

$$E(y_i) = \beta_0 + \sum_{j=1}^{P} \beta_j X_{ij} + \sum_{j=1}^{P} \beta_{jj} X_{ij}^2 + \sum_{j}^{P} \sum_{<l}^{P} \beta_{jl} X_{ij} X_{il}$$

$$i = 1, 2, \ldots, n.$$

donde las β son los parámetros desconocidos.

Análogamente, se puede hacer el mismo tipo de estudios considerando que las rectas son, en especial, polinomios de 1er -- grado; de ahí que pueda verse fácilmente como expansión por series de Taylor. Se define después una manera para calcular esparenza y varianza de las variables para regresión simple y múltiplc. Este es el caso específico que se presenta en el capítulo 5.

4.3 MODELOS RECURSIVOS.

Sea:

$$B x_t + C z_t + \eta_t = 0$$

la forma estructural de un modelo de ecuaciones múltiples. El modelo recibe el nombre de recursivo si existe un orden de las variables endógenas tal que la matriz B sea triangular y si:

$$b_{ij} = 0 \qquad j > i$$

$$E(\eta_{it}, \eta_{jt}) = 0 \qquad j \neq i$$

la i-ésima ecuación estructural puede considerarse entonces como representando la determinación causal de la i-ésima variable endógena x_i a partir del error ξ_i y de las demás variables que intervienen en dicha ecuación. No hay ninguna relación de dependencia de x_i hacia $x_1, x_2, \ldots, x_{i-1}$. Una regresión múltiple de x_i respecto a las demás variables que intervienen en la i-ésima ecuación estructural proporciona un estimador convergente* de esta ecuación.

Esta definición y esta propiedad se generalizan a los modelos en los que intervienen variables endógenas retardadas. Consideremos un modelo propuesto para describir el mercado de ciertos artículos de orígen agrícola (la carne de cerdo en particular).

La oferta es función del precio del año anterior, mientras que la demanda lo es del año considerado. La decisión de producción debe tomarse con mucha anticipación sobre la base de los precios observados en la campaña precedente. Este modelo puede escribirse:

$$q_t = a_1 p_{t-1} + f_1 + \eta_{1t}$$

$$p_t = \frac{1}{b} q_t + f_2 + \eta_{2t}$$

*En el sentido de función.

podemos representar las dependencias mediante el esquema:

$$\begin{array}{ccccccc} \longrightarrow p_{t-1} & \longrightarrow & q_t & \longrightarrow & p_t & \longrightarrow & q_{t+1} \\ \eta_{2,t-1}\nearrow & & \eta_{1,t}\nearrow & & \eta_{2,t}\nearrow & & \eta_{1,t+1}\nearrow \end{array}$$

pero el modelo no constituye una cadena causal pura a menos que los errores satisfagan las condiciones restrictivas:

$$E(\eta_{2t}\,\eta_{1,\,t-\tau}) = 0 \qquad \forall\ \tau \geq 0$$

$$E(\eta_{2t}\,\eta_{2,\,t-\tau}) = 0 \qquad \forall\ \tau > 0$$

es decir, para que no exista relación entre η_{2t} y q_t, es necesario que η_{2t} no esté relacionada con η_{1t}, ni con $\eta_{2,t-1}$, $\eta_{1,t-1}$, etc. Los errores η_{2t} relativos a diferentes años no deben presentar ninguna correlación y tampoco deben estar relacionados con los valores anteriores de η_{1t}.

El sistema puede considerarse como un modelo autorregresivo sobre el vector de componentes (p_t , q_t).

La ausencia de relación entre los errores debe estipularse para que la regresión mínimo cuadrática proporcione un estimador convergente de un modelo autorregresivo. (De no ser así, los errores también estarían relacionados con las variables predeterminadas, con respecto a las cuales se ha calculado la regresión).

En general, si consideramos el modelo:

$$B x_t + C z_t + D_1 x_{t-1} + \cdots + D_h x_{t-h} + \eta_t = 0$$

en donde las diferentes ecuaciones constituyen las relaciones estructurales, el modelo se llama recursivo si los errores no están relacionados entre sí, y si existe un orden en las variables endógenas (i=1,2,...,n) tal que la matríz B sea triangular, es decir, si:

$$b_{ij} = 0 \qquad \forall\ j > i$$

$$E(\eta_{it}\,\eta_{j,\,t-\tau}) = 0 \qquad \forall\ i,j \ \text{y}\ \forall\ \tau > 0$$

$$E(\eta_{it}\,\eta_{jt}) = 0 \qquad \forall\ j \neq i$$

En un modelo recursivo una regresión múltiple de x_i con relación a las demás variables que intervienen en la i-ésima ecuación estructural proporciona un estimador convergente para esta ecuación. Los modelos recursivos presentan la ventaja de poder estimarlos mediante procedimientos sencillos, debido a que cada ecuación es independiente de las demás; es decir, se tendrán varias regresiones. Debido a esto, el tratar por separado cada ecuación queda comprendido dentro de los capítulos anteriores.

CAPITULO 5

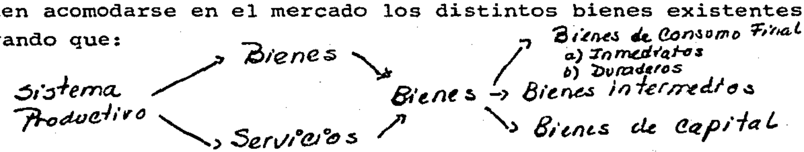
APLICACIONES DE LOS METODOS ESTADISTICOS.

EVALUACION DE PROYECTOS.

Una de las partes más importantes del análisis económico es la de evaluación de proyectos para empresas públicas y/o privadas. Hay dos tipos de evaluación aunque en realidad se trate de una forma con dos puntos de vista diferentes. Cuando se toma en cuenta el punto de vista de la empresa, se tendrá una evaluación financiera, y, cuando se toma en cuenta el punto de vista de la comunidad, tendremos una evaluación económica.

Si se toma en cuenta una evaluación financiera, al mismo tiempo será económica ya que entonces se verá involucrado alguno de los tres sectores en los que se ha dividido la economía mexicana.

La evaluación de proyectos se refiere a la forma en la que pueden acomodarse en el mercado los distintos bienes existentes considerando que:



En el sistema productivo un bien tiene su valor de uso de la transformación de materia prima o materia bruta. Los bienes se consideran comprendidos en alguna de las siguientes categorías:

Inmediatos.- Como los alimentos y el vestido, ya que se consumen por todos los miembros de la sociedad en todo momento.

Duraderos.- Son los bienes más sofisticados, como los automóviles, refrigeradores, etc

Intermedios.-Son los que componen a un bien final (que puede ser inmediato o duradero), como la madera, turbinas de un avión, motores, etc.

De Capital.- Son los que Marx llama capital constante, como el capital de las grandes industrias, barcos, etc.

La mayor producción de bienes casi siempre está dirigida a los de consumo final, en particular, los inmediatos. Los bienes más rentables son los de capital porque en ellos se involucra perfectamente un proceso de producción. En los bienes de capital no se requiere de mucha gente, hay incluso algunos que no son producidos en ciertos países por la mano de obra , las nuevas inversiones requeridas, etc., lo que influye es:

i) La experiencia

ii) El mercado existente.

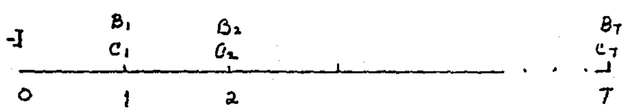
iii) La recuperación de la inversión.

Al estar relacionada la actividad económica en forma directa con el sistema productivo, se proporcionan bienes y servicios para los cuales, dado que existan o no, se tienen los métodos de planeación y de operación. A partir de ellos se implanta un mejor diseño y/o mejores medidas correctivas. En este ejemplo, nos referiremos al método de planeación.

En la evaluación de proyectos trataremos de medir la rentabilidad de un sistema productivo no existente. Esto es, vamos a considerar si un proyecto es o no rentable.

En proyectos de la iniciativa pública se requieren los máximos beneficios para la sociedad ya que dicho proyecto afecta a toda la comunidad. En proyectos de la iniciativa privada se requieren las máximas utilidades para la empresa.

Como a futuro los precios no van a mantenerse constantes, por lo general la tasa para el cálculo será la del interés bancario del mercado en donde piensan colocarse los bienes. Los inversionistas hacen un horizonte de planeación en donde:



I = Inversión

B_t = Beneficios o Ingresos

C_t = Costos

y la rentabilidad se mide trayendo a valor presente, de modo que:

$$R = -I + \sum_{t=1}^{T} \frac{B_t - C_t}{(1+i)^t}$$

donde i será fijada por el interés bancario.

Por ello, el criterio que usaremos será:

$$VPN = -I + \sum_{t=1}^{T} \frac{(B_t - C_t)}{(1+i)^t}$$

y, si $VPN > 0$ el sistema productivo es rentable

$VPN < 0$ el sistema productivo no es rentable

porque la rentabilidad se mide con respecto a la inversión y el tiempo.

La recuperación de activos, los intereses y la inversión no necesariamente son al principio del horizonte de planeación.

Debido a que en México se considera como mercado fuerte los bienes de consumo final, este estudio se referirá a la producción de un bien duradero.

El análisis debe empezar por conocer el entorno del sistema, formado por: competidores, proveedores, clientes, usuarios, etc., y de ahí deberán buscarse diferentes combinaciones hasta lograr los objetivos. Una vez que se haya elegido el mejor de los proyectos, este deberá implantarse y controlarse en su desarrollo para evitar desviaciones inconvenientes.

Los datos necesarios para la evaluación serán:

Utilidad Bruta = Ingresos - Costos

Utilidad de Operación = Utilidad Bruta - Intereses

Utilidad Neta = Utilidad de Operación - (Reparto de Utilidades + Impuestos sobre la renta)

Hasta aquí para efectos fiscales (Estado de pérdidas y ganancias)

Depreciación
Intereses
- Inversión de Activos y Capital (que se usa porque si no, estaríamos considerando dos veces la depreciación).

El período de recuperación estará dado por:

$$PR = \frac{\text{Monto de la Inversión}}{\text{Beneficio Bruto Medio Anual*}}$$

que es la media aritmética de la diferencia entre ingresos y gastos del proyecto durante su vida.*

El análisis nos dirá cuál es la capacidad que tiene la empresa para hacer frente al proyecto.

Supondremos que los beneficios y los costos se comportan como variables aleatorias, de modo que lo que queremos encontrar es el valor esperado del valor presente neto, $E[VPN]$. Para ello supondremos que VPN se comporta como una normal y buscaremos que la varianza (promedio de variación de los valores de su media) tienda a cero.

Como las proyecciones serán con base en datos históricos, lo que se estará haciendo no es más que análisis de regresión, pero utilizando métodos distintos a los desarrollados en los dos primeros capítulos de este trabajo.

Usando el llamado " Enfoque de 1er orden", cuando se trate de regresión lineal simple tendremos:

x = v. a. $\quad x = u + \xi \quad u$ = determinístico
h = función lineal. $\quad \xi$ = estocástico

$$y = a + bx$$

$$E(y) = h(E(x)) = a + bE(x)$$

$$Var(y) = \left[\left. \frac{d(a+bx)}{dx} \right|_{E(x)} \right]^2 \sigma_x^2 = b^2 \sigma_x^2$$

Cuando se trate de dos variables:

$$y^{o} = (x_1^{o}, x_2^{o})$$

utilizando desarrollo en series de Taylor:

$$y = h\left(x_1^{o}, x_2^{o}\right) \left.\frac{dy}{dx_1}\right|_{x_1^{o}} (x_1 - x_1^{o}) + \left.\frac{dy}{dx_2}\right|_{x_2^{o}} (x_2 - x_2^{o}) \Big)$$

al tomar valores esperados:

$$y^{o} = [E(x_1), E(x_2)]$$

$$y = h\left[E(x_1), E(x_2) + \left.\frac{dy}{dx_1}\right|_{E(x_1)} (x_1 - E(x_1)) + \left.\frac{dy}{dx_2}\right|_{E(x_2)} (x_2 - E(x_2))\right]$$

entonces

$$E(Y) = h\left(E(x_1), E(x_2)\right)$$

$$Var(Y) = Var\left(\left.\frac{\partial h}{\partial x_1}\right|_{E(x_1)} (x_1 - E(x_1)) + \left.\frac{\partial h}{\partial x_2}\right|_{E(x_2)} (x_2 - E(x_2))\right)$$

$$Var(x+y) = Var(x) + Var(y) + 2\, cov(x, y)$$

$$Var(K_1 x + K_2 y) = K_1^2\, var(x) + K_2^2\, var(y) + 2 K_1 K_2\, cov(x, y)$$

$$Var[Y] = \left(\left.\frac{\partial h}{\partial x_1}\right|_{E(x_1)}\right)^2 \sigma^2_{x_1} + \left(\left.\frac{\partial h}{\partial x_2}\right|_{E(x_2)}\right)^2 \sigma^2_{x_2}$$

$$+ 2\left(\left.\frac{\partial h}{\partial x_1}\right|_{E(x_1)}\right)\left(\left.\frac{\partial h}{\partial x_2}\right|_{E(x_2)}\right)\left(cov(x_1, x_2)\right)$$

y, como

$$\rho_{x_1 x_2} = \frac{cov(x_1, x_2)}{\sigma_{x_1}\, \sigma_{x_2}}$$

$$Var\ [Y] = \left(\frac{\partial h}{\partial x_1} \bigg|_{E(x_1)} \right)^2 \sigma_{x_1}^2 + \left(\frac{\partial h}{\partial x_2} \bigg|_{E(x_2)} \right)^2 \sigma_{x_2}^2$$

$$+ 2 \left(\frac{\partial h}{\partial x_1} \bigg|_{E(x_1)} \right) \left(\frac{\partial h}{\partial x_2} \bigg|_{E(x_2)} \right) \left(\rho_{x_1, x_2}\, \sigma_{x_1}\, \sigma_{x_2} \right)$$

finalmente, para regresión lineal múltiple:

$$Y = G\ (X_1, X_2, \ldots, X_T)$$

$$Y^o = (\mu_{x_1}, \mu_{x_2}, \ldots, \mu_{x_T})$$

$$Y = G\,(\mu_{x_1}, \mu_{x_2}, \ldots, \mu_{x_T}) + \sum_{t=1}^{T} \frac{\partial G}{\partial x_t}\bigg|_{\mu_{x_t}} (x_t - \mu_{x_t})$$

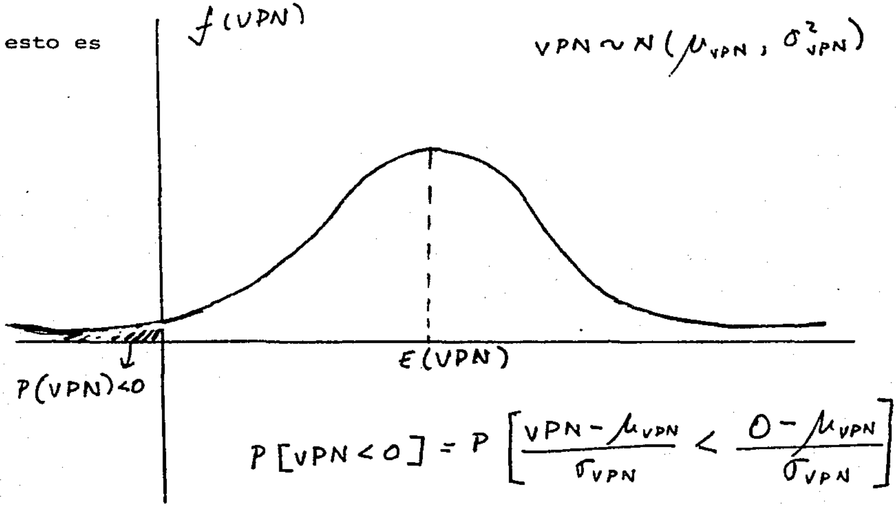
$$E\ (Y) = G\ (\mu_{x_1}, \mu_{x_2}, \ldots, \mu_{x_T})$$

$$Var\ (Y) = \left(\sum_{t=1}^{T} \frac{\partial G}{\partial X_t}\bigg|_{\mu_{x_t}} \right)^2 \sigma_{X_t}^2$$

$$+ 2 \sum_{t=1}^{T} \sum_{j=t+1}^{T} \left(\frac{\partial G}{\partial X_t}\bigg|_{\mu_{x_t}} \right) \left(\frac{\partial G}{\partial X_j}\bigg|_{\mu_{x_j}} \right) \rho_{x_t x_j} \cdot \sigma_{x_t}\, \sigma_{x_j}$$

Como consideramos que el comportamiento del *VPN* es normal, lo que nos interesa a fín de cuentas, es cuándo la probabilidad será mayor que cero.

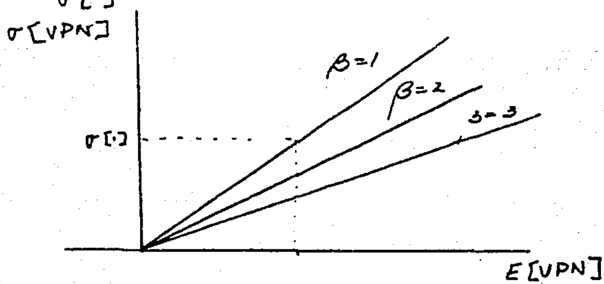
esto es



$$P[VPN < 0] = P\left[\frac{VPN - \mu_{VPN}}{\sigma_{VPN}} < \frac{0 - \mu_{VPN}}{\sigma_{VPN}}\right]$$

según el teorema del límite central. Y, por Chebyshev:

$$Si\ \beta = \frac{E[\cdot]}{\sigma[\cdot]} = \frac{1}{\sqrt{[\cdot]}}$$



a medida que β aumenta, la probabilidad de los valores reales se encontrará cercana a los que estamos estudiando, y mientras que $\sigma[\cdot] \rightarrow 0$, entonces $\beta \rightarrow \infty$.

Veremos ahora el tratamiento que se le dará a una alternativa de una planta productora de cemento. Los datos fueron proporcionados por la División de Estudios de Posgrado de la Facultad de Ingeniería, U.N.A.M.

El análisis será sobre la instalación de una planta con capacidad de producir 200 000 toneladas de cemento al año. Los costos iniciales y de operación son:

Concepto	Costo por ton. producida de cemento.	Costo Anual.
Gastos Gen. (C_1)	1.50	300 000
Mat. Primas (C_2)	1.40	280 000
Yeso (C_3)	0.36	72 000
Antracita (C_4)	0.84	168 000
Carbón (C_5)	1.92	384 000
Electricidad (C_6)	1.20	240 000
Mano de Obra (C_7)	0.80	160 000
Manten. (C_8)	0.60	120 000
Rep. May. (C_9)	0.75	150 000
Empaque (C_{10})	1.40	280 000
Transporte (C_{11})	4.00	800 000
TOTAL	14.77	2 954 000

Los precios están dados en dólares porque el dólar es un buen indicador de precios y la planta estará diseñada para producir cemento de exportación.

El equipo podrá ser depreciado en función de su vida útil, los edificios en 20 años, las obras civiles en 10 y la investigación y ayuda técnica en 5 años. Los impuestos (después de las depreciaciones) constituyen el 20% sobre las ganancias netas.

La construcción de la planta durará dos años, al término de los cuales se iniciará la producción de 200 000 toneladas cuya venta se tiene asegurada (por convenio) a un precio de 23 dólares por tonelada de cemento.

Para el análisis, todos los costos de operación constituyen variables aleatorias, las cuales son probabilísticamente independientes; asimismo, los costos de investigación y el capital inicial presentan un carácter estocástico (independiente del resto) el cual se mide en un coeficiente de variación del 10%. Los costos de los equipos adquiridos se tratan también como variables aleatorias considerando que su variabilidad aumenta con el tiempo (= 10% a los 5 años, 20% a los 10 años y 30% a los 15 años), y que su comportamiento probabilístico está perfectamente correlacionado entre ellos mismos pero es independiente de los demás. Los coeficientes de variación de los costos anuales de operación son:

C_1	0.10
C_2	0.10
C_3	0.05
C_4	0.05
C_5	0.10
C_6	0.01
C_7	0.05
C_8	0.05
C_9	0.05
C_{10}	0.01
C_{11}	0.20

Los costos iniciales estarán dados por:

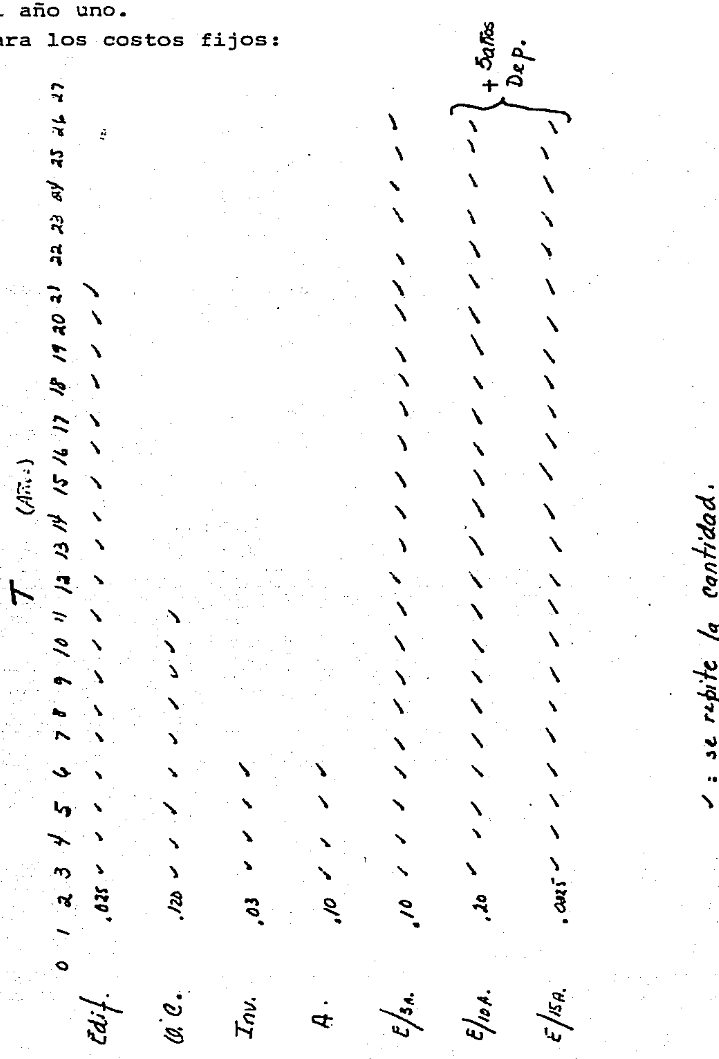
CONCEPTO		\$
Capital Inicial	(C_0)	300 000
Terreno	(T)	250 000
Edificios	(E)	500 000
Obras Civiles	(O)	1 200 000
Investigación	(C_{12})	150 000
Ayuda Técnica	(A)	500 000
Equipo Renovable:		
Cada 5 años	(C_{13})	500 000
Cada 10 años	(C_{14})	2 000 000
Cada 15 años	(C_{15})	1 300 000
TOTAL		6 700 000

La matríz de correlación es:

	C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}
C_0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C_1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C_2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
C_3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
C_4	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
C_5	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
C_6	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
C_7	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
C_8	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
C_9	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
C_{10}	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
C_{11}	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
C_{12}	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
C_{13}	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
C_{14}	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
C_{15}	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1

y lo que hay que calcular ahora es $E[VPN]$. Los costos iniciales serán pagados la mitad al comenzar (año cero) y la otra mitad en el año uno.

Para los costos fijos:



	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
T (Años)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Ingresos			86	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
− Costos Fijos			.66	✓	✓	✓	✓	.53	✓	✓	✓	✓	.41	✓	✓	✓	✓	✓	✓	✓	✓	✓	.38	✓	✓	✓	✓	✓
− Costos Var.			2.9	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
= Utilidad Bruta			1.04	✓	✓	✓	✓	1.17	✓	✓	✓	✓	1.24	✓	✓	✓	✓	✓	✓	✓	✓	✓	1.32	✓	✓	✓	✓	
− Intereses			–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	
= Utilidades de Op.			1.04	✓	✓	✓	✓	1.17	✓	✓	✓	✓	1.29	✓	✓	✓	✓	✓	✓	✓	✓	✓	1.32	✓	✓	✓	✓	
− (Rep. de Ut.			–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	
+ I.S.R.			.20	✓	✓	✓	✓	.23	✓	✓	✓	✓	.25	✓	✓	✓	✓	✓	✓	✓	✓	✓	.26	✓	✓	✓	✓	
= Utilidad Neta			.84	✓	✓	✓	✓	.94	✓	✓	✓	✓	1.04	✓	✓	✓	✓	✓	✓	✓	✓	✓	1.06	✓	✓	✓	✓	
+ (Depreciación			.66	✓	✓	✓	✓	.53	✓	✓	✓	✓	.41	✓	✓	✓	✓	✓	✓	✓	✓	✓	.38	✓	✓	✓	✓	
+ Intereses)			–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
− Inv. de Act y Costos	3.35	3.35	–	–	–	–	.5	–	–	–	–	2.5	–	–	–	–	1.8	–	–	–	–	2.5	–	–	–	–	–	

= Flujo Neto

$$= -3.35 - 3.35\,V_{.10}^{1} - .5\,V_{.10}^{6} - 2.5\,V_{.10}^{11} - 1.8\,V_{.10}^{16} - 2.5\,V_{.10}^{21} + 1.5\left(a_{\overline{6}|} - a_{\overline{1}|}\right)$$

$$+ 1.47\left(a_{\overline{11}|} - a_{\overline{6}|}\right) + 1.45\left(a_{\overline{21}|} - a_{\overline{11}|}\right) + 1.44\left(a_{\overline{26}|} - a_{\overline{21}|}\right) + 1.43\,V_{.10}^{27}$$

$$= E[VPN] = 3.99\ \bar{M}$$

Como $$V\,PN = \sum_{t=1}^{T} \frac{B_t - C_t}{(1+i)^t} - I$$

ahora B = constante = $4.6\ \bar{M}$ $\qquad$ $\bar{M}$= millones de dólares.

y necesitamos saber $VAR\,[V.P.N.]$

$$= \left(\frac{\partial VPN}{\partial C_t}\bigg|_t\right)^2 \sigma_{C_t}^2 + 2 \sum_{t=1}^{T} \sum_{j=t+1}^{T-1} \left(\frac{\partial VPN}{\partial C_t}\bigg|_t\right)\left(\frac{\partial VPN}{\partial C_j}\bigg|_t\right)$$

donde la tasa de interés fijada será del 10%.

Dado que

$$VPN = -\begin{matrix}\text{Costos}\\ \text{Iniciales}\end{matrix} + \sum_{t=2}^{24}\left(B - \sum_{\ell=1}^{11} C_\ell\right)\left(\frac{1}{(1+i)^t}\right) - \sum_{t=2}^{26}\left(\frac{C_{13}^t + C_{14}^t + C_{15}^t}{(1+i)^t}\right)$$

donde

C_{13}^t solo para t = 6, 11, 16 y 21

C_{14}^t solo para t = 11, 21

C_{15}^t solo para t = 16

$$VPN = -\frac{1}{2}\left(C_0 + T + E + O + C_{12} + A + C_{13} + C_{14} + C_{15}\right)$$

$$-\frac{1}{2}\left(C_0 + T + E + O + C_{12} + A + C_{13} + C_{14} + C_{15}\right)(1+i)^{-1}$$

$$+ \sum_{t=2}^{26}\left(B - \sum_{\ell=1}^{11} C_\ell\right)\left(\frac{1}{(1+i)^t}\right) \quad \text{Beneficios menos costos al valor presente.}$$

$$- \sum_{t=2}^{24}\left(\frac{C_{13}^t + C_{14}^t + C_{15}^t}{(1+i)^t}\right) \quad \text{costos.}$$

lo que necesitaremos ahora serán los valores parciales

$$\frac{\partial VPN}{\partial C_0} = -\frac{1}{2} - \frac{1}{2(1+i)} = -\frac{1}{2}\left(1 + \frac{1}{(1+i)}\right)$$

que será la misma para C_{12}

En los casos de $\ell = 1$ hasta $\ell = 11$

$$\frac{\partial VPN}{\partial C_\ell} = \frac{\partial \sum_{t=1}^{26}\left(B - \sum_{t=1}^{11} C_\ell\right)}{\partial C_\ell} = \frac{\partial}{\partial C\ell}\left(\frac{B - (C_1 + C_2 + \dots + C_{11})}{(1+i)^2} + \right.$$

$$\left. + \frac{B - (C_1 + C_2 + \dots + C_{11})}{(1+i)^3} + \dots + \frac{B - (C_1 + C_2 + \dots + C_{11})}{(1+i)^{26}}\right)$$

y esto es una anualidad $= -\left(a_{\overline{26|}} - a_{\overline{1|}}\right)$

para $\ell = 13$ a $\ell = 15$

$$\frac{\partial VPN}{\partial C_{13}} = -\frac{1}{2}\left(1 + \frac{1}{(1+i)}\right) - \left(\frac{1}{(1+i)^6} + \frac{1}{(1+i)^{11}} + \frac{1}{(1+i)^{16}} + \frac{1}{(1+i)^{21}}\right)$$

$$\frac{\partial VPN}{\partial C_{14}} = -\frac{1}{2}\left(1 + \frac{1}{(1+i)}\right) - \left(\frac{1}{(1+i)^{11}} + \frac{1}{(1+i)^{21}}\right)$$

$$\frac{\partial VPN}{\partial C_{15}} = -\frac{1}{2}\left(1 + \frac{1}{(1+i)}\right) - \left(\frac{1}{(1+i)^{16}}\right)$$

y con ello podrá calucularse la tabla de valores totales.

	$E[\cdot]$	$\Gamma[\cdot]$	$\sigma[\cdot]$	$\partial VPN/\partial c[\cdot]$	$\sigma^2_{[\cdot]}$	$\left(\frac{\partial VPN}{\partial[\cdot]}\right)^2$
c_0	.3	.10	.03	−.954	.0009	.9101
c_1	.3	.10	.03	−8.25	.0009	68.06
c_2	.28	.10	.028	−8.25	.000784	68.06
c_3	.072	.05	.0036	−8.25	.00001296	68.06
c_4	.168	.05	.0084	−8.25	.00007056	68.06
c_5	.384	.10	.0384	−8.25	.00147456	68.06
c_6	.240	.01	.0024	−8.25	.00000576	68.06
c_7	.160	.05	.0080	−8.25	.000064	68.06
c_8	.120	.05	.0060	−8.25	.000036	68.06
c_9	.150	.05	.0075	−8.25	.00005625	68.06
c_{10}	.280	.01	.0028	−8.25	.00000784	68.06
c_{11}	.800	.20	.160	−8.25	.0256	68.06
c_{12}	.150	.10	.0150	−.954	.000225	.9101
c_{13}	.5	.10	.05	−1.87	.0025	3.50
c_{14}	2.00	.20	.40	−1.44	.16	2.07
c_{15}	1.300	.30	.39	−1.17	.1521	1.37

$$\sum_{t=0}^{15} \sigma^2_{[\cdot]} \left(\frac{\partial VPN}{\partial c[\cdot]}\right)^2 = 2.52$$

Como

$$\sigma^2(VPN) = \sum_{t=1}^{15}\left(\frac{\partial VPN}{\partial C_t}\bigg|_{E[\cdot]}\right)^2 \sigma^2_{c_t}$$

$$+ 2\sum_{t=1}^{14}\sum_{j=t+1}^{15}\left(\frac{\partial VPN}{\partial C_t}\bigg|_{E[\cdot]}\right)\left(\frac{\partial VPN}{\partial C_j}\bigg|_{E[\cdot]}\right) + \rho_{c_t c_j}\cdot\sigma_{c_j}\cdot\sigma_{c_t}$$

$$= 2.52 + 2\,[(-1.87)(-1.44)(.05)(.4)$$

$$+ (-1.87)(-1.17)(0.05)(.39)$$

$$+ (-1.44)(-1.17)(.4)(.39)] = 3.23\ \overline{M}$$

y, desde el principio supusimos que $VPN \sim N\left(E(VPN), \sigma^2(VPN)\right)$
De acuerdo con el teorema del límite central:

$$P(VPN > 0) = 1 - P(VPN < 0)$$

$$= 1 - P\left(\frac{VPN - 3.99}{1.79} \leq \frac{0 - 3.99}{1.79}\right) = 1 - P(Z \leq -2.22)$$

$$P(VPN > 0) = .986$$

De acuerdo con Chebyshev:

$$\beta = \frac{E(VPN)}{\sigma(VPN)} = \frac{3.99}{1.79} \Rightarrow \beta = 2.22$$

$$P[VPN < 0] \leq \frac{1}{\beta^2} = .11$$

$$\Rightarrow P[VPN \geq 0] \geq 1 - \frac{1}{\beta^2} = .89$$

Con las dos pruebas, el proyecto es rentable.

Todas las diferentes alternativas deberán ser tratadas con el mismo método para que haya una buena base de comparación entre ellas.

El criterio se basa en funciones de producción econo-métricas y es muy aceptado en el mercado. En la forma en que fue usado para el ejemplo podemos tener una idea precisa de cómo usar la regresión lineal múltiple para planeación, y de lo estrechamente ligada a conceptos financieros para poder tener un punto de vista más completo con respecto al mercado que nos espera.

APENDICE

INVERSO DE MATRICES.

Consideremos el siguiente problema. Si nos dan una transformación A y un vector y tal que $A x = y$

¿ Podemos encontrar x ? Si asumimos que $Ax = y$ tiene una y sólo una solución, es natural, desde el punto de vista algebráico designar la solución:

$$x = \frac{1}{A} y = A^{-1} y$$

El símbolo A^{-1} representa el inverso de A con respecto de la multiplicación. Definimos la inversa, entonces, como la transformación que hace que y sea x .

Para asegurar la existencia de A^{-1} , requerimos una y solo una solución al problema de encontrar x tal que $Ax = y$ donde A e y están dadas. Las condiciones que establecemos son:

1.- $Ax = 0$ tiene una y solo una solución.

2.- $Ax = y \neq 0$ tiene solución (i. e. al menos una).

Satisfaciendo estas dos condiciones, $Ax = y$ tiene solución única. Si $Ax = 0$ tiene sólo una solución, sea $x = 0$. Entonces supongamos que $Ax = y$ tiene más de una solución, x_1 y x_2. Así que:

$$Ax_1 = y$$
$$Ax_2 = y$$
$$A(x_1 - x_2) = 0$$
$$\therefore x_1 = x_2$$

Decimos que las condiciones (1) y (2) definen una inversa; i.e., si se satisfacen, A^{-1} existe. Si A^{-1} existe se dice que A es no singular. Observemos que la existencia del inverso significa que hay una correspondencia uno a uno entre S y S_1 (los espacios donde están A y A^{-1}).

Algunas Propiedades Importantes.

Si tomamos $Ax = y$, y, asumiendo la existencia de A^{-1}, lo aplicamos a ambos lados, tenemos

$$A^{-1}Ax = A^{-1}y = x \quad . \; . \; . \; . \; . \; .$$

o

$$A^{-1}A = I \qquad . \qquad (a)$$

Si empezamos con $x = A^{-1}y$

$$A A^{-1} y = Ax = y$$

o

$$A A^{-1} = I \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; (b)$$

El saber que se cumple (a) ó (b) no demuestra la existencia del inverso, pero el hecho de que ambas se cumplan <u>sí</u> establece su existencia.

Demostración: Sean A , B y C operadores tales que

$$AB = CA = I$$

entonces, si A^{-1} existe y $A^{-1} = B = C$, para demostrar la existencia de A^{-1}, debemos mostrar que se satisfacen las condiciones (1) y (2).

$$Ax=0$$

con C operando en ambos lados

$$CAx=C(0)=0$$

Como, por hipótesis $CA=I$, $Ix=0$

de ahí que la X que satisface $Ax=0$ es $X=0$, por lo que se cumple la condición (1).

Ahora, para demostrar que $Ax=y$ tiene una solución, escribimos

$$Iy=y$$

en la forma

$$ABy=y$$

que es cierto por la hipótesis de que $AB=I$. Pero entonces $X=By$ es justamente la X que satisface $Ax=y$; lo que significa que se alcanzó la condición (2).

Como se satisfacen (1) y (2), A^{-1} existe.

Para demostrar la otra parte del resultado, que $A^{-1}=B=C$, usamos el hecho de que A^{-1} existe. A la identidad

$$A^{-1}=A^{-1}$$

aplicamos CA

$$CAA^{-1}=IA^{-1}=A^{-1}$$

da ahí que

$$CI=C=A^{-1}$$

queda ahora por demostrar que $A^{-1}=B$, empezamos con $AB=I$

aplicando A^{-1}

$$A^{-1}AB=A^{-1}$$

ó

$$IB=B=A^{-1}$$

VECTORES CARACTERISTICOS Y RAICES.

El problema del valor característico se define como encontrar los valores de un escalar λ y un vector asociado $x \neq 0$, que satisfacen:

$$Ax = \lambda x$$

donde A es una matríz de $n \times n$, λ es la raíz característica de A y x el vector característico.

Si tomamos el caso presentado en el capítulo 1, en el apartado de " Regresión Ortogonal", un caso 2×2 :

$$(a_{11}-\lambda) + a_{12}x_2 = 0$$

$$a_{21}x_1 + (a_{22}-\lambda)x_2 = 0$$

que podemos regresar a la forma matricial:

$$(A - \lambda I)x = 0$$

y la ecuación solamente tiene una solución no trivial, $x \neq 0$,si $(A - \lambda I)$ es singular, esto es, si

$$|A - \lambda I| = 0$$

Esto nos conduce a un polinomio con λ como desconocida, que puede ser resuelto y con ello, obtenerse los valores de el vector característico:

$$(a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21} = 0$$

esto es:

$$\lambda^2 - (a_{11}+a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0$$

con raíces:

$$\lambda_1 = \frac{1}{2}\left[(a_{11}+a_{22}) + \sqrt{(a_{11}+a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}\right.$$

$$\lambda_2 = \frac{1}{2}\left[(a_{11}+a_{22}) - \sqrt{(a_{11}+a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}\right.$$

En el caso especial de matríz simétrica, $a_{12} = a_{21}$, las raíces son:

$$\lambda = \frac{1}{2}(a_{11}+a_{22}) \pm \sqrt{(a_{11}-a_{22})^2 + 4a_{12}^2}$$

en donde λ_1 y λ_2 serán necesariamente reales.

INTERPRETACION GEOMETRICA DEL METODO DE AJUSTE POR MINIMOS CUADRADOS.

En la estadística clásica, usamos el criterio de los mínimos cuadrados para definir el mejor ajuste, en esta parte veremos cómo se relaciona este ajuste con ideas geométricas. Una vez que sea clara esa relación, resultará obvio ver que el ajuste por mínimos cuadrados ofrece un ajuste conveniente y unificado a varias situaciones que son manejadas frecuentemente por fórmulas individuales. Los principios de inferencia justifican el ajuste por mínimos cuadrados, pero para muchos propósitos pedagógicos, su simplicidad es más que suficiente para que se piense en ellos. Será conveniente ver los datos en forma de una matríz con n filas (una por cada observación o unidad experimental) y una columna por cada característica o variable de la manera:

$$\begin{array}{ccccc} 1 & z_1 & x_1 & y_1 & \cdots \\ 1 & z_2 & x_2 & y_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \\ 1 & z_n & x_n & y_n & \cdots \end{array}$$

e interpretar las n-adas que forman las columnas de la matríz, como vectores. El vector constante J =(1, 1,..., 1) lleva una parte muy importante en lo que sigue. Interpretaremos las n-tuplas geométricamente.

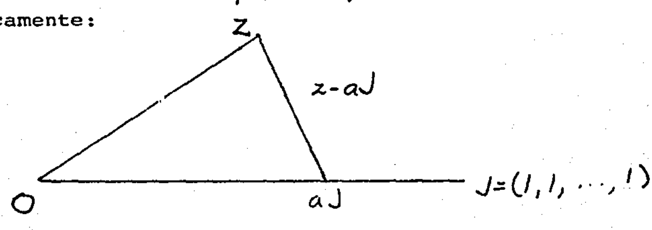
De acuerdo con el principio de los mínimos cuadrados, si queremos resumir los datos $(z_1, z_2, \ldots, z_n)$ en un sólo número a, el mejor a es aquel que minimiza

$$\sum_{i=1}^{n} (z_i - a)^2$$

y usando producto interior esto puede escribirse:

$$|z - aJ|^2$$

Geométricamente:



podemos considerar como posibles modelos a todos los múltiplos aJ de J y elegir el más cercano a z. "Mínimos Cuadrados" significa "distancia más corta" usando producto interior en el Espacio Euclidiano.

El punto más cercano es la proyección $P_J(z)$ de z sobre J. Aplicando esto para obtener el valor de a :

$$P_J(z) = aJ \quad \text{para alguna } a$$

$$\langle z - P_J(z), J \rangle = 0$$

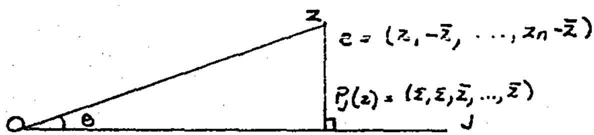
de ahí que

$$\langle z, J \rangle = a \langle J, J \rangle$$

y esto es

$$\sum_{i=1}^{n} z_i = an \qquad \text{o} \qquad a = \bar{z} = \frac{1}{n} \sum_{i=1}^{n} z_i$$

por eso el mejor número que resume los datos es la media $\bar{z}$. La Geometría Euclidiana nos sugiere la solución y la Geometría Analítica nos permite calcularla. La solución que se muestra:



indica cómo hemos descompuesto z en sus dos componentes ortogonales: $\bar{z}J$, el componente sobre J, y $e = z - \bar{z}J$, el residuo o error. Si todos los z's fueran idénticos, z caería exactamente sobre J y tendríamos $|e|^2 = 0$. De ahí que podamos pensar en $\bar{z}J$ como la parte de z que puede ser explicada satisfactoriamente por una constante. Para $|e|^2 > 0$ esta explicación no es buena porque los datos varían. Interpretamos e como las variaciones en z. Entonces z = componente constante más variaciones $= \bar{z}J + (z - \bar{z}J)$.

De acuerdo con Pitágoras:

$$|z|^2 = |\bar{z}J|^2 + |z - \bar{z}J|^2$$

que se reduce en este caso a:

$$\sum_{i=1}^{n} z_i^2 = n\bar{z}^2 + \sum_{i=1}^{n} (z_i - \bar{z})^2$$

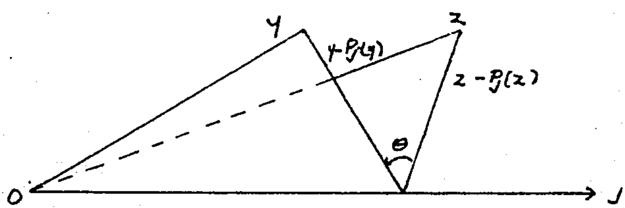
la descomposición usual de la suma de cuadrados.

Para medir qué tan buen resultado es se usan varias formas. $|e|^2$ es frecuentemente llamada la suma de los errores al cuadrado (SSE). Una variación de $|e|^2$ se llama media cuadrática del error (MSE). Esperamos que $|e|^2$ dependa de n, mientras más observaciones, más variación, cuando menos en sentido absoluto. Para tomarlo en cuenta, podemos devidir SSE por la dimensión del sub-espacio relevante obteniendo un promedio de la suma de errores al cuadrado "por dimensión". Como e está restringida a caer en el sub-espacio de dimensión (n-1) ortogonal a J, dividimos por (n-1) obteniendo:

$$MSE = |e|^2/(n-1) = (n-1)^{-1} \sum_{i=1}^{n} (z_i - \bar{z})^2$$

que es la varianza usual. Ya sea SSE o MSE es una medida, entonces, de la extensión del mejor número que resume los datos, $\bar{z}$. Incidentalmente, la dimensión del subespacio dentro del cual un vector está restringido o pertenece se llama " grados de libertad" asociados con ese vector.

En el caso en que se tengan $(y_1, y_2, \ldots, y_n)$ y $(z_1, z_2, \ldots, z_n)$ y se quiera saber qué tan fuerte es la relación entre ellas, usando la misma interpretación geométrica de y y z como antes, esto significa preguntarnos qué tan cerca están y y z el uno del otro.



Por lo visto anteriormente, podríamos medir esto con $|y-z|^2$ o con alguna función del ángulo entre y y z. En la prác tica, se interpreta como ¿ Qué tan fuertes son las variaciones en y relacionadas con las variaciones en z ? De la misma manera que en el caso anterior, tomamos y y z como sus componentes ortogonales a $\mathcal{J}$; $y-\bar{y}\mathcal{J} = y - P_{\mathcal{J}}(y)$ y $z-\bar{z}\mathcal{J} = z - P_{\mathcal{J}}(z)$. La medida usual de la relación de las variaciones en y y z es cos (θ), regularmente llamado " correlación" de y y z . Cuando $\cos^2(\theta) = 1$ (θ=0° ó 180°) las variaciones en y son exactamente proporcionales a las variaciones en z, $y-\bar{y}\mathcal{J}$ y $z-\bar{z}\mathcal{J}$ son colineales. Cuando cos $(\theta) = 0$ (θ=90°) las variaciones en y no están relacionadas con las variaciones en z , $y-\bar{y}\mathcal{J}$ y $z-\bar{z}\mathcal{J}$ son ortogonales.

El producto interior de $y-\bar{y}\mathcal{J}$ y $z-\bar{z}\mathcal{J}$, cuando se divide por los grados de libertad apropiados (n-1), es la covarianza de y y z . De una manera más general, si primero proyectamos y y z sobre el plano L determinado por $\mathcal{J}$, w, x, ..., entonces el coseno del ángulo entre $y - P_L(y)$ y $z - P_L(z)$ se llama " correlación parcial" de y y de z , quitando los efectos de w, x, ... La mayor parte de los análisis de correlación tienen interpretación geométrica natural, la atención está restringida regularmente al subespacio ortogonal a $\mathcal{J}$, el subespacio de variación en las variables.

CONCLUSIONES.

Los modelos de regresión en general, y los utilizados en ---- esta tesis en particular, son una herramienta valiosa en la estadística, ya que permiten describir la asociación entre las variables dependientes e independientes al sumarizar la tendencia de los datos y encontrar la forma de asociación entre las variables. No obstante esto, no se pretende el establecimiento de relaciones causales, en el sentido de que los valores de las variables independientes produzcan cambios en los valores de las variables dependientes, sino que indican únicamente la forma de asociación entre las variables.

Por otro lado, puede suceder que existan variables intervinientes pero no controlables que modifiquen los valores de las variables, y basados en otros conocimientos ajenos a la estadística se establezca una relación causa-efecto. Es aquí, donde los modelos de regresión se convierten en poderosos auxiliares que nos permitirán simplificar y/o estudiar esta relación.

Un ejemplo de la utilización de los modelos de regresión - lo encontramos en las relaciones entre factores que pueden pasar a - ser hipótesis científicas al suponerse la relación causa-efecto, provisionalmente.

Entre los métodos más conocidos para la estimación de parámetros (y con ello obtener las ecuaciones), se encuentra el de mínimos cuadrados, que nos permite determinar los parámetros de tal manera que al hacer el ajuste, el error sea el menor posible (ver interpretación geométrica pág. 117). Además, este método ofrece propiedades estadísticas muy atractivas (los estimadores obtenidos por este método cumplen con las propieda des deseables en ellos: sin tendencia, varianza mínima, error cuadrático medio mínimo y consistencia), razón por la cual se ha convertido en uno de los métodos más eficaces no sólo para el análisis de regresión, sino para otras muchas -- aplicaciones de diferentes disciplinas entre las que podemos mencionar la óptica y la biomedicina.

Asimismo, en el caso de las aplicaciones económicas, casi siempre se tienen tamaños de muestra pequeños, y al calcularse, por ejemplo, por otros métodos, se sesga, cosa que no sucede con el método de mínimos cuadrados.

La manera en la que este trabajo se presenta propone un orden lógico para estudiar métodos estadísticos (muy en especial mínimos cuadrados) uniéndolos en todo lo posible esperando que se vea la relación directa entre ellos. En el capítulo 1, donde se tratan los "modelos no estocásticos", las estimaciones hechas solamente con estas herramientas pueden ser "buenas" en algunos casos, pero; tal vez no lo suficiente en la mayoría de ellos. Es debido a eso que se introduce el capítulo 2, que indica un estudio sobre modelos estocásticos en donde se nos ayuda a obtener mayores bases para calcular mejores estimadores que los anteriores.

Una vez que se han obtenido estos estimadores, lo más natural parece ser preguntarse ¿ Qué sucederá cuando el fenómeno en estudio quiera ser visto en diferentes momentos y no en uno solo? Es por ello que se incluyen los modelos autorregresivos en el capítulo 3 con el propósito de presentar algunas maneras de calcularlos y teniendo como ejemplo de ello el " Ajuste por mínimos cuadrados en muestras pequeñas" (pág 81) ya que después de ver la sección " Método de mínimos cuadrados" para modelos autorregresivos (pág 77), es notoria la dificultad para calcularlos y, como se mencionó en su oportunidad, sólo se ha logrado obtener unos cuantos resultados.

Después de que estos tres capítulos estudian las partes más importantes de algunos modelos de regresión, lo ideal de acuerdo con los propósitos de esta tesis es la manera en que algunos conceptos básicos son interpretados en economía haciendo uso de ellos, dando motivos para esclarecer dudas sobre el porqué de su tratamiento en una forma específica, como se ve en el capítulo 4.

Una de las razones más poderosas para efectuar este tipo de investigaciones es el buen uso que puede hacerse de ellas para llevar a cabo planeaciones en distintos tipos de empresas, que fue lo que motivó el desarrollo del capítulo 5.

Las investigaciones en Economía son las únicas que se tratan en forma diferente de las que corresponden a las otras ciencias sociales, ya que; como pudimos notar, son susceptibles de creación de modelos matemáticos para la conceptualización de diferentes fenómenos, pero no se les puede ver en forma demasiado abstracta porque nos estaríamos alejando de comportamientos del ser humano y esta es una materia humanística.

Existen todavía muchas conductas que son difíciles de identificar como funciones debido a que en teoría pueden seguir un tipo de comportamiento pero las conclusiones a las que uno llega con ello pudieran ser totalmente erróneas si cambiamos un aspecto que no consideramos relevante y que en realidad sí lo es, pero puede no ser susceptible de medición. Sin embargo, se ha avanzado mucho en este campo durante el presente siglo.

El Actuario, cuyo ámbito de acción se basa en la aplicación de teorías matemáticas a diferentes aspectos de la vida cotidiana, debe entonces conocer la forma de utilizar los distintos métodos aquí presentados y que son considerados básicos, en bien de su desarrollo profesional futuro, y que, a fín de cuentas, no son más que una de las muchas aplicaciones de la Estadística.

BIBLIOGRAFIA.

-Allard, R. J., " Introducción a la Econometría". Ed. Limusa, México 1980.

-Campbell, Stephen K., " Equívocos y Falacias en la interpretación de Estadísticas", Ed. Limusa, México, 1981.

-Colin Glass, J., " Métodos Matemáticos para Economistas", McGraw-Hill, Colombia, 1982.

- Cramer, J. S., " Econometría Empírica", Fondo de Cultura Económica, México, 1981.

- Croxton, F. E., y Crowden, D. J., " Estadística General Aplicada", Fondo de Cultura Económica, México, 1957.

-Ferguson y Gould, " Teoría Microeconómica", Fondo de Cultura Económica, México, 1984.

- Gujarati, Damodar; " Econometría Básica", McGraw-Hill, México, 1985.

-Harnecker, Marta; " Los Conceptos fundamentales del Materialismo Histórico". Ed. Siglo XXI, México, 1980.

- Hernández Lerma, Onésimo. " Métodos Estadísticos y Aplicaciones",I. P. N., México, 1981.

- Hogg, Robert V., " Introduction to Mathematical Statistics". Macmillan Publishing Co., EE. UU. 1978.

- Johnston, J., " Econometric Methods". McGraw-Hill, EE. UU., 1960.

- Lange, Oscar; "Introducción a la Econometría", Fondo de Cultura Económica, México, 1978.

- Maddala, G. S. " Econometría", McGraw-Hill, México, 1985.

- Malinvaud, Edmond;"Métodos Estadísticos de la Econometría", Editorial Ariel, España.

- Malinvaud, Edmond; "Cahiers de séminaire d'économétrie", Centre National de la Recherche Scientifique, Francia. 1984.

- Martínez Garza, Angel; " Métodos Econométricos". Colegio de Postgraduados, México, 1982.

- Mills, Richard, L. " Estadística para Economía y Administración", McGraw-Hill, México, 1980.

- Monfort, Alain. " Cours de Statistique Mathematique", Economica, Francia, 1982.

- Núñez del Prado, Arturo. " Estadística Básica para Planificación". Ed. Siglo XXI, México, 1985.
- O. N. U., "Manual de Proyectos de Desarrollo Económico", EE. UU.
- O. N. U., " Manual para la Preparación de Estudios de Viabilidad-- Industrial", EE. UU., 1978.
- Patat, Jean-Pierre; " Monnaie, institutions financieres et politique monétaire", Economica, Francia, 1984.
- Pérez, Luis Antonio. " Estadística Matemática", I. P. N., México, -- 1983.
- Salvatore, D. " Econometría", McGraw-Hill, México, 1983.
- Samuelson, Paul, A. " Curso de Economía Moderna". Ed. Aguilar, España, 1979.
- Sierra Bravo, Restituto. "Análisis Estadístico y Modelos Matemáticos" Ed. Paraninfo, España, 1981.
- Spiegel, M. " Estadística". McGraw-Hill, México, 1976.
- Tinbergen, J. " La Planeación del Desarrollo", Fondo de Cultura Económica. México, 1982.
- Tobin, James. " Reflexions sur la theorie Macroeconomique Contempo-- raine", Economica, Francia, 1983.
- Vuskovic, Pedro. " Los Instrumentos Estadísticos del Análisis Económico", C. I. D. E., México, 1984.
- Wexler, Charles. " Geometría Analítica: Un Enfoque Vectorial". Editorial Montaner y Simon. España, 1977.
- Wonnacott, R. J., y Wonnacott, T. H., " Econometrics", John Wiley -- and Sons, EE. UU. 1979.