UNIVERSIDAD DE PANAMA
VICERRECTORIA DE INVESTIGACION Y POSTGRADO
PROGRAMA DE MAESTRIA EN MATEMATICA

"UNA MEJOR APROXIMACION AL COEFICIENTE DE CORRELACION DE PEARSON"

Por:

JOSE DEL C. OCHOA B.

Tesis presentada como uno de los requisitos para optar por el grado de Maestro en Ciencias con Especialización en Estadística Matemática.

Panamá, República de Panamá

1 9 8 6

UNIVERSIDAD DE PANAMA

SIBIUP
Biblioteca Int Simon Bolivar
00253490

VICERRECTORIA DE INVESTIGACION Y POSTGRADO

Aprobado por:

Director de Tesis ______________________
M.Sc. RICARDO PARKER DUPUY

Miembro del Jurado ______________________
M.Sc. DIMAS QUIEL

Miembro del Jurado ______________________
M.Sc. PLUTARCO RAMOS

DEDICATORIA

Quiero dedicar este trabajo
con todo cariño a quienes
tuvieron la suficiente
paciencia para conmigo y que
en todo momento se constitu-
yeron en apoyo y estímulo pa-
ra que este trabajo fuese
una realidad

A mis padres Isidora y José
A mis amigos, Julio, Aurora,
Egberto y Jorge Poltronieri,
A la Virgen del Carmen y a
Dios.

José

AGRADECIMIENTO

Quiero agradecer de manera muy especial al profesor RICARDO A. PARKER D., asesor de mi tesis de graduación como también a todas las personas que desinteresadamente me animaron y me dieron confianza para concluir este trabajo.

Muchas Gracias,

INDICE GENERAL

INDICE GENERAL

CAPITULO III

INTERVALO DE CONFIANZA

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

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

Uno de los temas que mayor interés ofrece dentro de los fundamentos representados por variables de carácter bidimensional o multidimensional es aquel que nos da a conocer la forma y el grado de relación que existe entre dichas variables.

Para estudiar este tema se plantean dos técnicas estrechamente relacionadas entre sí que son: la regresión y la correlación.

El análisis de correlación produce un número que resume el grado de relación o asociación entre dos o más variables; y el análisis de regresión da lugar a una ecuación matemática que describe dicha relación.

Considérense, preguntas como las siguientes:

1.- ¿Están relacionadas la edad y la resistencia física?

2.- ¿Tienden a tener mayor nivel de escolaridad las personas con altos ingresos, en comparación con las de bajos ingresos?

3.- ¿Puede el éxito en el trabajo predecirse a partir de calificaciones obtenidas en las pruebas de selección?

4.- ¿Parece influir la temperatura en el índice de criminalidad?

5.- ¿Se desenvolverán mejor en un curso de matemática, los alumnos con mayor habilidad para la lectura que los que muestran menos habilidad?

6.- ¿Está relacionado el aprovechamiento escolar de nivel universitario con el aprovechamiento respectivo a nivel de bachillerato y la calificación del examen de admisión?

7.- ¿Qué relación existe entre elconsumo de carne, el precio de la carne y el precio de la carne de cerdo?

8.- ¿El número de personas protegidas por la C.S.S., guarda relación con el número de asegurados cotizantes y el número de beneficiarios?

Estos problemas y problemas semejantes se presentan en un análisis de correlación. El resultado de un análisis de esta naturaleza es el llamdo coeficiente de correlación, cuyo valor cuantifica el grado de asociación, entre las variables involucradas en dicho análisis.

En este trabajo estudiaremos métodos para medir el grado de dependencia entre variabales aleatorias. Primeramente, estudiaremos el coeficiente de correlación exactamente como medida de dependencia entre dos variables aleatorias.

Esta noción es luego extendida a una medida de dependencia entre una variable y un conjunto de variables por medio del coeficiente de correlación múltiple. Por otra parte, para estudiar la dependencia entre dos variables de un grupo cuando los efectos de las otras variables correlacionadas han sido removidas, se introduce la noción de correlación parcial.

Se estudiará además la estimación y prueba de hipótesis concerniente a los distintos coeficientes de correlación.

CAPITULO I

CORRELACION SIMPLE O TOTAL

Cuando tratamos de establecer el grado de asociación entre dos variables estamos refiriéndonos a su correlación, que podemos entender como el conjunto de causas comunes a ambos. Así, si una variable (por ejemplo Y) se supone "motivada" o "explicada" por otra (por ejemplo X), la correlación entre ambas nos hablará del grado de asociación en que el fenómeno representado por ambas poseen causas comunes.

Estas variables representan un conjunto de puntos, de acuerdo con el tipo de relación. Si el espacio es bidimensional, los puntos son $(x_1, y_1), \ldots, (x_n, y_n)$, que determinan el diagrama de dispersión, sirve como referencia para determinar nuestra curva de ajuste, que puede ser:

1.- $Y = a_o + a_1 X$ Línea Recta

2.- $Y = a_o + a_1 X + a_2 X^2$ Parábola

3.- $Y = a_o + a_1 X + a_2 X^2 + a_3 X^3$ Curva Cubíca

4.- $Y = a_o + a_1 X + a_2 X^2 + a_3 X^3 + a_4 X^4$ Curva Cuártica

5.- $Y = a_o + a_1 X + a_2 X^2 + \ldots + a_n X^n$ Curva de grado n.

Estas ecuaciones se llaman polinómicas de primero, segundo, tercero, cuarto y n grados respectivamente.

Otras posibles ecuaciones que en la práctica aparecen son las siguientes:

1.- $Y = \dfrac{1}{a_o + a_1 X}$ ó $\dfrac{1}{Y} = a_o \neq a_1 X$ Hipérbola

2.- $Y = ab^X$ ó $\log Y = \log a \neq X \log b$ Curva exponencial

3.- $Y = aX^b$ ó $\log Y = \log a \neq b \log X$ Curva geométrica

Cuando se trata de dos variables solamente se habla de correlación simple, ésta puede ser Lineal o no-Lineal.

Para llegar a una posible definición y evitar el juicio individual en la construcción de recta, parábolas u otras curvas de aproximación, usaremos el Método de Mínimos - Cuadrados.

Dado los siguientes puntos (x_1, y_1), $(x_2, y_2), \ldots (x_n, y_n)$, es decir, para cada valor de X, obtendremos el correspondiente valor de Y en la curva, consecuentemente las desviaciones, $D_1, D_2 \ldots D_n$.

Una medida de "bondad del ajuste" de la curva a los datos viene suministrada por la cantidad $D_1^2 + D_2^2 + \ldots + D_n^2$. Si esta es pequeña, el ajuste es bueno, si es grande el ajuste es malo.

Haremos la siguiente definición formal.

<u>Definición</u>:

De todas las curvas de aproximación a una serie de datos puntuales la curva que tiene la propiedad de que

$D_1^2 \neq D_2^2 \neq \ldots D_n^2$ es mínima (1)

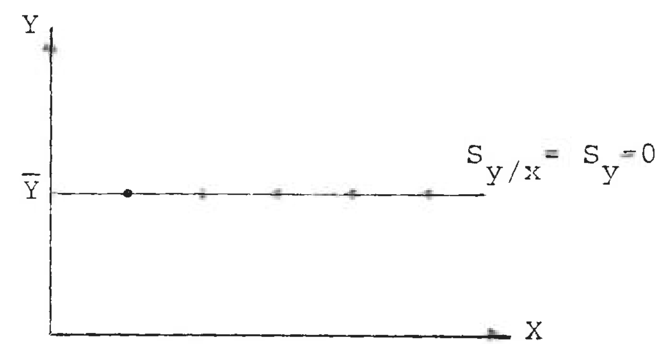
Se conoce como la mejor curva de ajuste.

Una curva que presenta esta propiedad se dice que se ajusta a los datos por mínimos cuadrados y se llama curva de mínimos cuadrados. Así, una recta con esta propiedad se llama recta de Mínimos cuadrados, una parábola con esta propiedad se llama parábola de mínimos cuadrados.

Nuestro objetivo es medir el grado de asociación entre Y y X, antes de estudiar la Recta de Mínimos-Cuadrados, estudiaremos tres casos extremos.

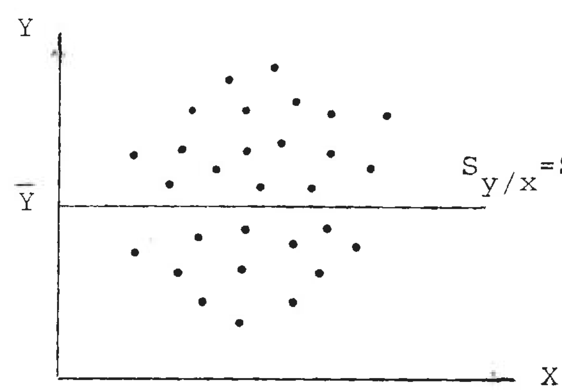
Caso 1. Correlación cero.

Y no esta asociado con X; Y no esta afectada por nada. No hay variación, este caso no es interesante porque no hay variación en Y.



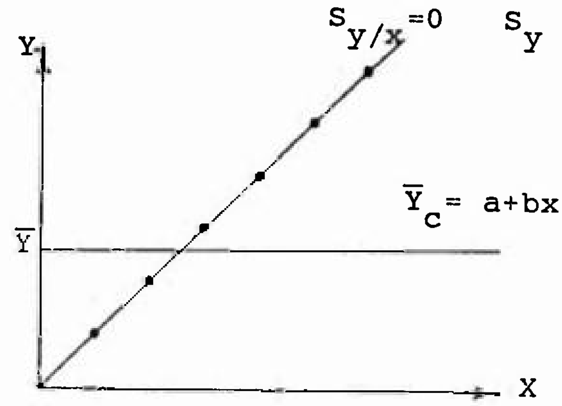
Caso 2. Correlación cero

Y no esta asociada con X, no hay ninguna asociación o la que llamaremos correlación; predecir Y no es trivial pero conocer X no reduce la variación de Y.



Caso 3. Correlación perfecta

La asociación entre Y y X aparece gráficamente como una línea recta. Todos los puntos caen en una línea. Hay variación en Y, pero el conocer X sabemos exactamente lo que es Y.



El caso 1 tiene la característica perturbadora de que una línea recta (con b = 0) describe perfectamente los punto pero no hay "ninguna asociación" porque Y es independiente de X. El caso 1 "no es interesante" desde el punto de vista de esta discusión porque no hay variación en Y.

Recta de Mínimos Cuadrados

La recta de aproximación por mínimos-cuadrados del conjunto de puntos (x_1,y_1), (x_2,y_2),...(x_n,y_n) tiene la ecuación

$$Y = a + bX \qquad (2)$$

Nuestro problema es determinar los parámetros a y b, utilizando el principio de mínimos-cuadrados Sea $d_i = Y_i - Y_c$ la i-ésima desviación, error o residuo entre Y_i(valor observado) y Y_c (valor calculado).

Entonces $\sum_{i=1}^{N} d_i = 0$ y $\sum_{i=1}^{N} d_i^2 \neq 0$

Luego

$$\sum_{i=1}^{N} (y_i - y_c)^2 = \sum_{i=1}^{N} d_i^{\,2}$$

Tenemos

$$Q = \sum_{i=1}^{N} d_i^{\,2} = \sum_{i=1}^{N} (y_i - a - bx_i)^2 \qquad (3)$$

que ha de ser mínima respecto de a y b para la cual es preciso que sus respectivas derivadas parciales con relación a dicho parámetro se anulen

$$\frac{\partial Q}{\partial a} = \sum_{i=1}^{N} 2(y_i - a - bx_i)(-1) = 0$$

$$\frac{\partial Q}{\partial b} = \sum_{i=1}^{N} 2(y_i - a - bx_i)(-x_i) = 0$$

Así las ecuaciones normales quedarán en la forma

$$\left.\begin{aligned} \sum_{i=1}^{N} Y_i &= Na + b \sum_{i=1}^{N} X_i && (4) \\ \sum_{i=1}^{N} X_i Y_i &= a \sum_{i=1}^{N} X_i + b \sum_{i=1}^{N} X_i^{\,2} && (5) \end{aligned}\right\}$$

Dividiendo en ambas ecuaciones por n, y representando por $\overline{X}$ a la media de X y $\overline{Y}$ a la media de Y, nos quedará

$$\left.\begin{aligned} \overline{Y} &= a + b\,\overline{x} && (6) \\ \frac{\sum X_i Y_i}{N} &= a\overline{x} + \frac{b \sum_{i=1}^{N} X_i^{2}}{N} && (7) \end{aligned}\right\}$$

de donde, podemos deducir los valores de a y b multiplicando la primera ecuación por $-\overline{x}$ y sumándola con la segunda tendremos

$$\frac{\sum_{i=1}^{N} X_i Y_i}{N} - \overline{X}\,\overline{Y} = b\left(\frac{\sum x_i^{2}}{N} - \overline{x}^{2}\right) \qquad (8)$$

$$b = \frac{\dfrac{\sum_{i=1}^{N} x_i y_i}{N} - \overline{X}\,\overline{Y}}{\dfrac{\sum_{i=1}^{N} X_i^{2}}{N} - \overline{X}^{2}}$$

En el denominador de (8) tenemos la notación que representa a la Variación de X o sea

$$S^2_X = \frac{\sum_{i=1}^{N} X_i^{\,2}}{N} - \overline{X}^2 = \frac{\sum_{i=1}^{N} X_i^{\,2}}{N} - 2\overline{X}^2 + \overline{X}^2$$

$$= \frac{\sum_{i=1}^{N} X_i^{\,2}}{N} - 2\overline{X}\,\frac{\sum_{i=1}^{N} X_i}{N} + \overline{X}^2$$

$$= \frac{\sum_{i=1}^{N} X_1^{\,2} - 2\overline{X}\sum_{i=1}^{N} X_i + N\,\overline{X}^2}{N}$$

$$= \frac{\sum_{i=1}^{N} (X_i^{\,2} - 2\overline{X}\,X_i + \overline{X}^2}{N}$$

$$S_X^{\,2} = \frac{\sum_{i=1}^{N} (x_i - \overline{x})^2}{N}$$

En el numerador de (8) tenemos la notación que representa a la covarianza de XeY, o sea

$$S^2_{XY} = \frac{\sum X_i Y_i}{N} - \overline{X}\,\overline{Y}$$

$$= \frac{\sum X_i Y_i - N \overline{X}\ \overline{Y}}{N}$$

$$= \frac{\sum X_i Y_i - N \overline{Y}\ \overline{X} - N \overline{X}\ \overline{Y} + N \overline{X}\ \overline{Y}}{N}$$

$$= \frac{\sum (X_i Y_i - \overline{Y} \sum X_i - \overline{X}\ Y_i + \overline{X}\ \overline{Y})}{N}$$

$$= \frac{\sum [(Y_i - \overline{Y}) X_i - (Y_i - \overline{Y}) \overline{X}]}{N} = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{N}$$

entonces

$$b = \frac{S^2_{XY}}{S^2_{\ X}} \qquad (9)$$

de la ecuación (6) podemos obtener que

$$a = \overline{Y} - b\overline{X}$$

remplazando en (2) obtenemos

$$Y = \overline{Y} - b\overline{X} + bX$$

entonces

$$Y - \overline{Y} = b(X - \overline{X})$$

Finalmente, la expresión de la regresión lineal de Y sobre X, se puede escribir así

$$Y - \overline{Y} = \frac{S^2XY}{S^2X}(X-\overline{X}) \qquad (10)$$

De la misma forma, podríamos determinar la regresión lineal mínimo-cuadrática de x sobre y, donde b' sería el coeficiente de regresión de X sobre Y, estableciendo que

$$b' = \frac{S^2X}{S^2YX} \qquad (11)$$

La Correlación Lineal.

Después de obtener la recta de regresión entre las dos variables buscaremos el grado de asociación entre las variables o sea la correlación.

Es claro que

$$\sum_{i=1}^{N}(y_i-\overline{y})^2 = \sum_{i=1}^{N}\left[(y_i-y_o)+(y_o-\overline{y})\right]^2 \qquad (12)$$

Desmostraremos que (12) es equivalente a la

$$\sum_{i=1}^{N}(y_i-\overline{y})^2 = \sum_{i=1}^{N}(y_i-y_o)^2 + \sum_{i=1}^{N}(y_o-\overline{y})^2$$

Demostración:

Desarrollando el producto notable, obtendremos

$$\sum_{i=1}^{N}(y_i-\bar{y})^2 = \sum_{i=1}^{N}(y_i-y_o)^2 + 2\sum_{i=1}^{N}(y_i-y_o)(y_o-\bar{y}) + \sum_{i=1}^{N}(y_o-\bar{y})^2$$

La demostración queda hecha, si se consigue demostrar que el segundo sumando es cero, y esto es evidente ya que

$$\sum_{i=1}^{N}(y_i-y_o)(y_o-\bar{y}) = \sum_{i=1}^{N}(y_i-a-bx)(a+bx-\bar{y})$$

$$= a\sum(y_i-a-bx) + b\sum X(y-a-bx)$$

$$- \bar{y}\sum(y-a-bx)$$

las ecuaciones normales

$$\sum(y_i-a-bx) = 0$$

y

$$\sum X(y-a-bx) = 0$$

por lo que

$$\sum_{i=1}^{N}(y_i-y_o)(y_o-\bar{y}) = 0$$

De alli

$$\sum_{i=1}^{N} (y_i-\bar{y})^2 = \sum_{i=1}^{N} (y_i-y_o)^2 + \sum_{i=1}^{N} (y_o-y)^2 \qquad (13)$$

de la ecuación (13) podemos determinar que y_o es el valor de Y según la recta de regresión de Y sobre X, entonces podemos deducir la ecuación de regresión

$$Y_o - \bar{Y} = \frac{S^2{}_{xy}}{S_X^2} (X-\bar{X}) \qquad (14)$$

Elevándolos al cuadrado obtenemos

$$(y_o-\bar{y})^2 = \frac{(S^2{}_{xy})^2}{(s_x^2)^2} (x-\bar{x})^2 \qquad (15)$$

Dividiendo ambos miembros entre $(y_i-\bar{y})^2$ obtenemos

$$\frac{(y_o-\bar{y})^2}{(y_i-\bar{y})^2} = \frac{(x-\bar{x})^2}{(y_i-\bar{y})^2} \frac{(S^2_{xy})^2}{(s_x^2)^2} \qquad (16)$$

Como este es para el i-ésimo elemento entonces

$$\frac{\sum_{i=1}^{N} (y_o-\bar{y})^2}{\sum_{i=1}^{N} (y_i-\bar{y})^2} = \frac{(S^2{}_{xy})^2}{(s^2{}_x)^2} \frac{\sum_{i=1}^{N} (x_i-\bar{x})^2}{\sum_{i=1}^{N} (y_i-\bar{y})^2} \qquad (17)$$

representando $\dfrac{\sum_{i=1}^{N}(y_o-\bar{y})^2}{\sum_{i=1}^{N}(y_i-\bar{y})^2}$ por ρ^2xy, quedaría (17) transformado en

$$\rho^2_{xy} = \frac{\dfrac{(S^2_{xy})^2}{S_x^2} \sum_{i=1}^{N}(x_i-\bar{x})^2}{\sum_{i=1}^{N}(y_i-\bar{y})^2}$$

$$= \frac{\dfrac{(S^2_{xy})^2}{(s^2x)^2} \quad \dfrac{1}{N} \sum_{i=1}^{N} (x_i-\bar{x})^2}{\dfrac{1}{N} \sum_{i=1}^{N} (y_i-\bar{y})^2}$$

$$= \frac{\dfrac{(S^2xy)^2}{(s^2x)^2} \quad S^2X}{S^2Y}$$

$$= \frac{(S^2_{xy})^2}{(s_x^2)(s^2y)} \qquad (18)$$

Así pués tendremos el coeficiente de correlación lineal xy o sea

$$\rho_{xy} = \frac{s^2{}_{xy}}{\sqrt{s^2{}_x}\ \sqrt{s^2{}_y}} \qquad (19)$$

$$\rho_{xy} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2/N}\ \sqrt{\sum_{i=1}^{N}(y_i-\bar{y})2/N}} = \frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i-\bar{y})^2}}$$

Correlación Simple (No Lineal)

Las definiciones de coeficiente de correlación son completamente generales y pueden utilizarse, tanto para regresiones No-Lineales, como para lineales; la única diferencia es que Y_o se calcula a partir de una ecuación de regresión no-lineal en lugar de una ecuación de regresión lineal.

Entre las curvas de regresión no-Lineal que estudiaremos, se encuentran la parábola.

Parábola de Mínimos Cuadrados

La parábola de aproximación de mínimos cuadrados a la serie de puntos (x_1,y_1), (x_2,y_2),...,(x_n,y_n) tiene la ecuación

$$Y_o = a + bX + cX^2 \qquad (20)$$

sea

$$d_i = Y_i - Y_o$$

el residuo o error entre Y_i (valor observado) y Y_o (valor calculado)

Entonces

$$\sum_{i=1}^{N} d_i = 0 \quad y \quad \sum_{i=1}^{N} d_i^2 \neq 0 \tag{21}$$

Luego

$$\sum_{i=1}^{N} (y_i - y_o)^2 = \sum_{i=1}^{N} d_i^2$$

$$Q = \sum_{i=1}^{N} d_i^2 = \sum_{i=1}^{N} (y_i - a - bx_i - cx_i^2)^2$$

Derivando con respecto a los parámetros a,b,c, tenemos que:

$$\frac{\partial Q}{\partial a} = \sum_{i=1}^{N} 2(y_1 - a - bx_i - cx_i^2)(-1) = 0$$

$$\frac{\partial Q}{\partial b} = \sum_{i=1}^{N} 2(y_i - a - bx_i - cx_i^2)(-x_i) = 0$$

$$\frac{\partial Q}{\partial c} = \sum_{i=1}^{N} 2(y_i - a - bx_i - cx_i^2)(-x_i^2) = 0$$

Así las ecuaciones normales quedarán en la forma

$$\left.\begin{aligned} \sum_{i=1}^{N} y_i &= Na + b \sum_{i=1}^{N} x_i + c \sum_{i=1}^{N} X_i^2 \\ \sum_{i=1}^{N} X_i Y_i &= a \sum_{i=1}^{N} X_i + b \sum_{i=1}^{N} X_i^2 + c \sum_{i=1}^{N} X_i^3 \\ \sum_{i=1}^{N} X_i^2 Y_2 &= a \sum_{i=1}^{N} X_i^2 + b \sum_{i=1}^{N} X^3 + c \sum_{i=1}^{N} X^4 \end{aligned}\right\} \quad (22)$$

Las constantes a,b,c se determinan resolviendo el sistema de ecuación (22). Una vez obtenido los valores a,b,c especificamos la curva de regresión (20). Utilizando los mismos principios que en (12) y (13), obtendremos que

$$\sum_{i=1}^{N} (y_i - \bar{y})^2 = \sum_{i=1}^{N} (y_i - y_o)^2 + \sum_{i=1}^{N} (y_o - \bar{y})^2 \qquad (23)$$

realizando el mismo trabajo que en (14), (15), (19) y (17), obtendremos que:

$$r^2 = \frac{\sum_{i=1}^{N}(y_i-y_o)^2}{\sum_{i=1}^{N}(y_i-\bar{y})^2}$$

$$r = \sqrt{\frac{\sum_{i=1}^{N}(y_i-y_o)^2}{\sum_{i=1}^{N}(y_i-\bar{y})^2}} = \sqrt{\frac{\text{Variación Explicada}}{\text{Variación Total}}} \qquad (24)$$

Regresión Lineal en Notación Matricial

Introducción:

Entre los instrumentos matemáticos que se suelen emplear las matrices ocupan un lugar importante, como ayuda, aunque los problemas sean complejos, su presentación en esta forma es concisa. Sucede lo mismo con las soluciones.

La ecuación general para el modelo de regresión lineal.

$Y_i = a+bx_i + ei, \; i = 1,...,n$

Entonces podemos escribir el conjunto de observaciones como:

$$
\begin{aligned}
Y_1 &= a + bx_1 + e_1 \\
Y_2 &= a + bx_2 + e_2 \\
- &\;\; - \; - \; - \; - \;\; - \; - \\
- &\;\; - \; - \; - \; - \;\; - \; - \\
Y_n &= a + bx_n + e_n
\end{aligned}
$$

Estas n ecuaciones tienen elementos que caen naturalmente dentro de categorías. A la izquierda de la igualdad están las observaciones propiamente dicha. A la derecha los coeficientes de los parámetros y las componentes de error (e_i). Podemos escribir en notación Vectorial y Matricial respectivamente, como

$$
Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_n \end{bmatrix} \qquad X = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & X_n \end{bmatrix} \qquad B = \begin{bmatrix} a \\ b \end{bmatrix} \qquad e = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_n \end{bmatrix}
$$

Donde Y es el vector aleatorio de observación. La matriz X, llamada también matriz de diseño, es una matriz de parámetros observable. Aquí la columna de los 1 puede considerarse valores de una variable ficticia $X_{oi}=1, i=1,\ldots,n$.

El vector B es un vector de parámetros desconocidos y e_i es un vector de componentes aleatorias desconocidas.

Ahora las ecuaciones pueden escribirse como

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \qquad (25)$$

0 en forma general $Y = XB \neq e_i$

<u>La Matriz de Covarianza</u>

Consideremos la distribución normal bivariante donde el vector media es

$$E \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \qquad (26)$$

En el caso bivariante la matriz de covarianza puede escribirse

$$\Sigma_1 = E \begin{bmatrix} (x_1-\mu_1)^2 & (x_1-\mu_1)(x_2-\mu_2) \\ (x_2-\mu_2)(x_1-\mu_1) & (x_2-\mu_2)^2 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

donde σ_{11} y σ_{22} son las varianzas y σ_{12} y σ_{21} son las covarianzas.

En el caso multivariante, la matriz de varianza-covarianza se puede denotar

$$\Sigma = (\sigma_{ij})i,j=1,\ldots,n = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & & \ddots & \vdots \\ \sigma_{n_1} & & & \sigma_{nn} \end{bmatrix} \qquad (28)$$

Como Σ es definida semi-positiva, entonces el determinante del i-j ésimo elemento de la matriz es mayor que cero entonces

$$\begin{vmatrix} \sigma_{ii}^2 & \sigma_i & \sigma_j & \sigma_{ij} \\ \sigma_{jj} & \sigma_{ii} & \sigma_{ij} & \sigma_{jj}^2 \end{vmatrix} = \sigma_i \; \sigma_j (1-\rho_{ij})^2 \qquad (29)$$

es positivo

Luego $1-\rho_{ij}^2 > 0 \Longrightarrow -\rho_{ij}^2 > -1$

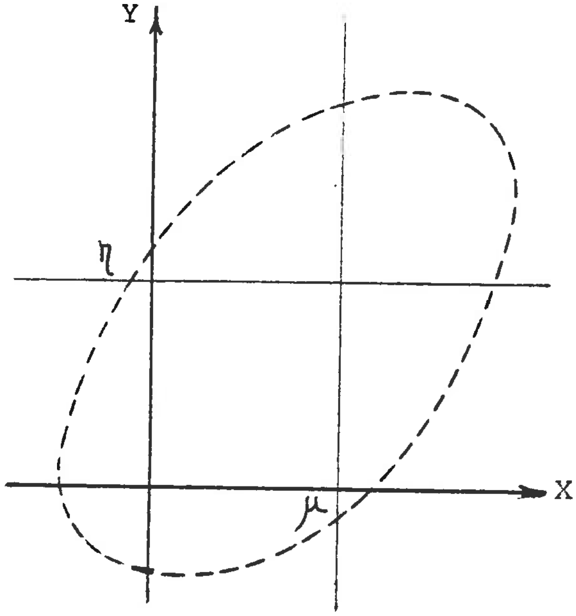
$$\Longrightarrow \rho_{ij}^2 < 1$$

$$\Longrightarrow |\rho_{ij}| < 1 \qquad (30)$$

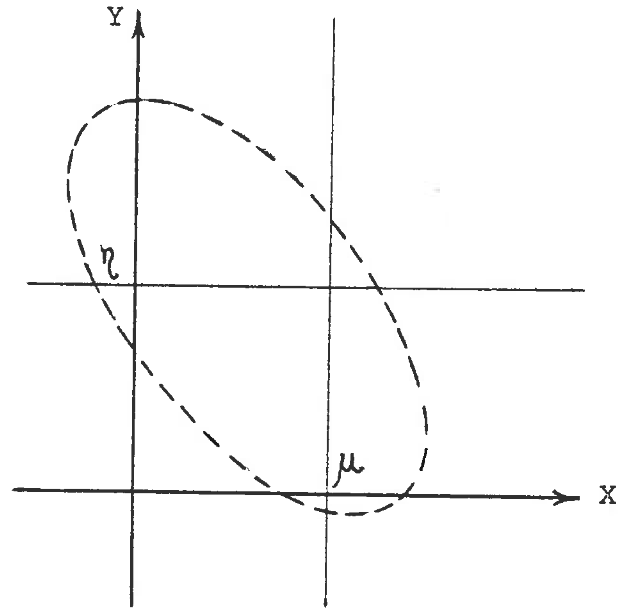
El coeficiente de correlación simple o total lo escribimos

$$\rho_{ij} = \frac{E(x_i-\mu)(y_j-\eta)}{\sqrt{E(x_i-\mu)^2 E(y-\eta)^2}} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} \qquad (31)$$

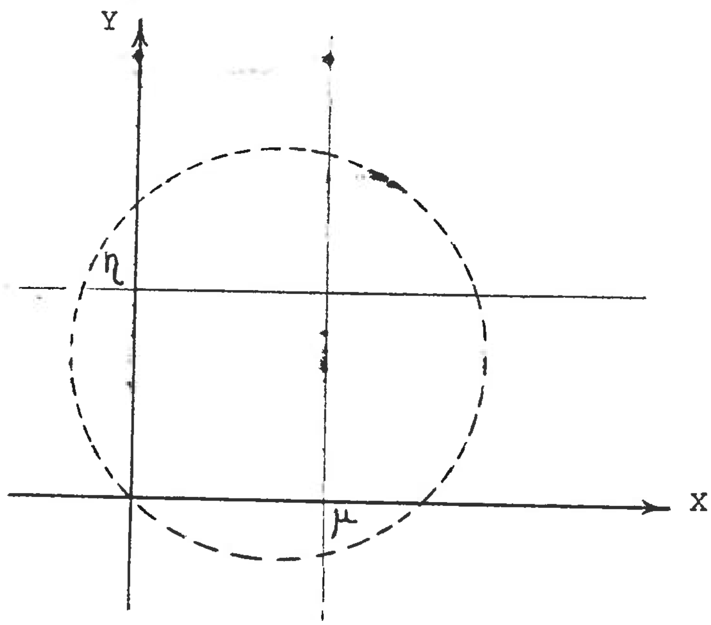
como sabemos, ésta es una medida de asociación entre dos variables aleatorias X_i e Y; La covarianza σ_{ij} en el numerador es la esperanza del producto de las desviaciones de las variables respecto de sus medias $(x_i-\mu)(y_j-\eta)$, por lo tanto el hecho de que una covarianza sea grande y positiva lo entendemos entonces como aquel caso en que ρ_{ij} está cerca de 1. Las siguientes figuras ejemplifican como se distribuye la gran mayoría de la masa de probabilidad correspondiente a la variable (x,y) de acuerdo al valor de la correlación



ρ_{ij} cercano a 1



ρ_{ij} cercano a -1



ρ_{ij} cercano a 0

Es importante notar que ρ_{ij} alcanza los valores extremos en (30), cuando $\rho_{ij} = 1$ o -1 las variables X y Y están sobre una recta con probabilidad 1. Es fácil verificar que si $\rho_{ij} = -1$ ésta es negativa. Teniendo en cuenta esta discusión, podemos afirmar que la correlación es una medida del grado de linealidad de la distribución (x,y).

Proposición

El coeficiente de Correlación no depende de las unidades que se hallan adoptado para medir la X o Y.

Demostración

Verificaremos que ρ_{xy} es invariante bajo cambios de escala o sea que la correlación de $T = ax \neq b$ y $P = cy \neq d$ es igual a la de X e Y

Dado

$$\rho_{xy} = \frac{E(x-\mu)(y-\eta)}{\sqrt{E(x-\mu)^2 \; E(y-\eta)^2}}$$

Sea

$$T = ax + b \quad y \quad P = cy + d$$

$$E[T] = aE[x] + b \quad y \quad E[P] = CE[y] + d$$

$$\text{como } E[x] = \mu \quad y \quad E[y] = \eta$$

Entonces

$$E\left[T\right] = a\mu + b \quad y \quad E\left[P\right] = c\eta + d$$

Luego

$$\rho_{TP} = \frac{E(T-E(T))E(P-E(P))}{\sqrt{E(T-E(T))^2 E(P-E(p))^2}}$$

$$= \frac{E(ax+b-a\mu-b)(cy+d-c\eta-d)}{\sqrt{E(ax+b-a\mu-b)^2 E(cy+d-c\eta-d)^2}}$$

$$= \frac{acE(x-\mu)(y-\eta)}{ac\sqrt{E(x-\mu)^2 E(y-\eta)^2}}$$

$$\rho_{TP} = \frac{E(x-\mu)(y-\eta)}{\sqrt{E(x-\mu)^2 \; E(y-\eta)^2}} = \rho_{xy} \quad \text{L.Q.Q.D}$$

<u>Propiedad</u>

Si X e Y son independientes entonces $\rho_{ij} = 0$.

Se deduce fácilmente de (31) en caso de X e Y independiente

$$\begin{aligned} E(x-\mu)(y-\eta) &= E(x-\mu)\; E\,(y-\eta) \\ &= (E(x)-\mu)(E(Y)-\eta) \\ &= (\mu-\mu)(\eta-\eta) \end{aligned}$$

$$= 0 \cdot 0$$

$$= 0$$

entonces $\rho_{ij} = 0$

Observación

1.- El recíproco de esta propiedad no es en general cierta o sea que existen variables dependientes con $\rho_{ij}=0$

2.- Si las variables son normales implica la independencia cuando $\rho_{ij} = 0$.

Correlación Parcial

Consideremos un vector aleatorio $X = (x_1, \ldots, x_p)$ con vector de media μ y matriz de covarianza Σ.

Consideremos una partición de dimensión q y (p - q), respectivamente, $X = (x^{(1)}, x^{(2)})$.

La función densidad condicional de $X^{(1)}$ dado $X^{(2)}$ es

$$f(x^{(1)}/x^{(2)}) = \frac{f(x^{(1)}, x^{(2)})}{g(x^{(2)})} \qquad (32)$$

en donde g representa la función de densidad marginal de $X^{(2)}$.

La distribución condicional tiene medias, varianzas y convarianzas que son funciones de $X^{(2)}$.

A las medias de esa distribución condicional se les llama función de regresión:
denotémosles por

$$E(x_i/x_{q+1},\ldots,x_p) = \mu_i\ (x_{q+1},\ldots,x_p) \tag{33}$$

$$i=1,2,\ldots,q$$

Esta función de regresión tienen la propiedad de ser aquellas que mejor estiman a X_i en el sentido de que minimiza

$$E\left[(x_i-g(x_{q+1},\ldots,x_p))^2\right] \tag{34}$$

Esto se demuestra tomando esperanzas condicionales dado $x^{(2)} = (x_{q+1},\ldots,x_p)$, lo cual permite dar la siguiente expresión equivalente a (34)

$$E\left\{E\left[(x_i-g(x_{q+1},\ldots,x_p))^2/x^{(2)}\right]\right\} \tag{35}$$

Minimizando la esperanza de adentro se logra el mínimo de (34); el cual por lo tanto se alcanza cuando

$$g(x_{q+1},\ldots,x_p) = \mu_i(x_{q+1},\ldots,x_p) \tag{36}$$

Consideremos ahora el problema de minimizar una expresión similar a (34) cuando la función g es un hiperplano. Supongamos de ahora en adelante para facilitar la escritura que $E(x) = \mu = 0$.

Las funciones g que consideramos son del tipo

$$g = \beta_{q+1}X_{q+1} + \beta_{q+2}X_{q+2} + \cdots + \beta_p X_p = \beta' X^{(2)} \qquad (37)$$

Definición:

Plano de regresión media cuadrática.

El plano de regresión media cuadrática para X_i con respecto a $X^{(2)} = (x_{q+1}, \ldots, x_p)$ es el hiperplano que minimiza

$$E\left[(x_i - \beta' x^{(2)})^2 \right] \qquad (38)$$

Diremos que $\beta' X^{(2)}$ es el mejor estimador lineal de X_i en término de $X^{(2)}$ en el sentido de minimizar (38); es una representación lineal óptima de X_i en términos de las componentes de $X^{(2)}$

Consideremos ahora la variable aleatoria que se obtiene como diferencia entre la variable X_i y el plano de regresión media cuadrática.

$$\eta_i,\ q \neq 1, \ldots, p = X_i - \beta' X^{(2)} \qquad (39)$$

que denominaremos residuo de X_i con respecto a los componentes de $X^{(2)}$, donde β' está dado por las N-1 ecuaciones

$$C_{q+1,q+1}\ \beta_{q+1} + C_{q+1,q+2}\ \beta_{q+2}+\ldots+C_{q+1,p}\ \beta_p = C_{i,q+1} \quad (x)$$

$$C_{q+2,q+1}\ \beta_{q+1}+C_{q+2,q+2}\ \beta_{q+2}+\ldots+C_{q+2,p}\ \beta_p = C_{i,q+2}$$

$$C_{p,q+1}\ \beta_{q+1}+C_{p,q+2}\ \beta_{q+2}+\ldots+C_{p,p}\ \beta_p = C_{ip}$$

que se obtienen derivando (38) con respecto a los p,q coeficientes β_{ij}, y donde $C_{h,k}$ denota la covarianza de X_h con X_k; $C_{hk} = E(x_h, x_k)$

La idea es que el residuo representa aquella parte de X_i que no puede ser explicada a través de $X^{(2)}$, o bien dicho de otro modo, aquella parte de X_i que queda luego de remover la influencia de $X^{(2)}$ sobre X_i.

Consideremos

$$E\ [(x_i - \beta' x^{(2)})\ x^{(2)})]$$

$$= E\ [X_i X^{(2)}) - \beta'\ X^{(2)}\ X^{(2)})]$$

$$= E(x_i x^{(2)'}) - E(\beta' x^{(2)} x^{(2)'})$$

$$= E(x_i x_{q+1}) - E(\beta' x^{(2)} x_{q+1})$$

$$= C_{i,q+1} - E(\beta_{q+1,}X_{q+1,}X_{q+1} + \dots + \beta_p \; x_p x_{q+1})$$

$$= C_{i,q+1} - \beta_{q+1} C_{q+1,q+1} - \dots - \beta_p \; C_{p,q+1}$$

$$= C_{i,q+1} - (\beta_{q+1} \; C_{q+1,q+1} + \dots + \beta_p C_{p,q+1})$$

$$= C_{i,q+1} - C_{i,q+1} = 0 \quad \text{de acuerdo con la ecuación } (*)$$

Esto demuestra que los residuos de los componentes de $X^{(1)}$ respecto a $X^{(2)}$ no están correlacionados con $X^{(2)}$

Consideremos dos variables X_i y X_j $i, j < q$ conjuntamente con el vector $X^{(2)} = (x_{q+1}, \dots, x_p)$. La covarianza de X_i y X_j, pueden estar influenciadas por la variación de $X^{(2)}$ e interesa en muchos casos a estudiar dicha dependencia separando el efecto que sobre ella efectúa $X^{(2)}$.

Ahora bien los residuos $\eta_{i,\, q \neq 1, \dots, p}$ y $\eta_{j,q \neq 1, \dots, p}$ representan aquella parte de X_i y X_j que queda luego de restar el mejor estimador lineal en términos de $X^{(2)}$. Entonces el coeficiente será llamado coeficiente de Correlación Parcial de X_i y X_j con respecto a $X^{(2)}$ y se denotará

mediante $$\rho_{ij,q+1,\dots,p} = \frac{E\,[(\eta_{i,q+1,\dots,p})(\eta_{j,q+1,\dots,p})]}{\sqrt{E(\eta_i^2{}_{,q+1,\dots,p})E(\eta_j^2{}_{q+1,\dots,p})}} \tag{40}$$

Siendo esta expresión un coeficiente de correlación simple entre dos variables, tendrá todas las propiedades enunciadas en particular.

$$-1 \leqslant \rho_{ij,q+1,\ldots,p} \leqslant 1$$

Examinaremos ahora el caso particular en que X tiene distribución normal multivariante.

Teorema (Distribución Condicional)

Sea $X \sim N(\mu\ ,\ \Sigma)$ donde $X=(X^{(1)},X^{(2)})$ es un vector distribuido con media

$$= (\mu^{(1)}, \mu^{(2)})$$

y matriz de covarianzas

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \qquad (41)$$

Entonces la distribución de $X^{(1)}$ dado $X^{(2)} = x^{(2)}$ es normal con media $\mu^{(1)} + \Sigma_{12}\ \Sigma_{22}^{-1}(x^{(2)} - \mu^{(2)}$ y matriz de convarianzas.

$$\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\ \Sigma_{22}^{-1}\ \Sigma_{21} \qquad (42)$$

Demostración

Sea la función de densidad condicional (o de cuantía) de $X^{(1)}$ dado $X^{(2)}$ es

$$g(x^{(1)}/x^{(2)} = x^{(2)}) = \frac{f(x^{(1)}\ x^{(2)})}{f(x^{(2)})}$$

Entonces

$$h(x_2) = \frac{1}{(2\pi)^{q/2}\ |\Sigma_{22}|^{1/2}} \exp\left[-\tfrac{1}{2}(x_2-\mu_2)'\ \Sigma_{22}^{-1}(x_2-\mu_2)\right]$$

Luego como

$$\Sigma^{-1} = \begin{bmatrix} (\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')^{-1} & -(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\ -\Sigma_{22}^{-1}\Sigma_{12}^{-1}(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')^{-1} & \Sigma_{22}^{-1}+\Sigma_{22}^{-1}\Sigma_{12}'(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')\Sigma_{12}\Sigma_{22}^{-1} \end{bmatrix}$$

La determinante es

$$|\Sigma| = |\Sigma_{22}|\ \ |\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}'|$$

$$f(x_1,x_2) = \frac{1}{(2\pi)^{1/2(p+q)} \left|\Sigma_{22}\right|^{1/2} \left|\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}'\right|^{1/2}}\,.$$

$$. \operatorname{Exp}\left\{-\frac{1}{2}\left[(x_1-\mu_1)'(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')^{-1}(x_1-\mu_1)-\right.\right.$$

$$(x_2-\mu_2)'\Sigma_{22}^{-1}\Sigma_{12}'(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')^{-1}(x_1-\mu_1) -$$

$$(x_1-\mu_1)'(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')^{-1}\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)$$

$$+ (x_2-\mu_2)'\Sigma_{22}^{-1}\Sigma_{12}'(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')^{-1}\Sigma_{12}\Sigma_{22}^{-1}$$

$$\left.\left.(x_2-\mu_2) + (x_2-\mu_2))\Sigma_{22}^{-1}(x_2-\mu_2)\right]\right\}$$

Luego

$$g(x_1/x_2) = \frac{1}{(2\pi)^{p/2}\left|\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}'\right|^{1/2}}\,.$$

$$. \operatorname{Exp}\left\{-\frac{1}{2}\left[x_1-\mu_1-\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)\right]'(\Sigma_{11}-\Sigma_{12}\right.$$

$$\Sigma_{22}^{-1}\Sigma_{12}')^{-1}.$$

$$\left.\left[x_1-\mu_1-\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)\right]\right\}$$

h representa la función de densidad (o cuantia marginal de $X^{(2)}$).

La distribución condicional

$$X^{(1)}/X^{(2)} = x^{(2)} \sim N\left[\mu_1 + \Sigma_{12}\,\Sigma_{22}^{-1}(x_2 - \mu_2),\right.$$

$$\left.\Sigma_{11} - \Sigma_{12}\,\Sigma_{22}^{-1}\,\Sigma_{12}'\right] \qquad (43)$$

Este teorema nos dice que en el caso normal, las funciones de regresión que habiamos llamado $\mu_i(x_{q+1},\dots,x_p)$ son los componentes del vector $\mu^{(1)} + \Sigma_{12}\,\Sigma_{22}^{-1}\cdot$ $\cdot(x^{(2)} - \mu^{(2)})$.

Estas funciones son lineales en $X^{(2)}$ y coinciden entonces con los planos de regresión media cuadrática; o sea que la solución del problema de minimizar (34) coincide con la de minimizar (38). Por consiguiente, los residuos $\eta_{i,q+1,\dots,p}$ son los componentes del vector

$$X^{(1)} - \mu^{(1)} - \Sigma_{12}\,\Sigma_{22}^{-1}(X^{(2)} - \mu^{(2)}).$$

A partir de esta expresión de los residuos, puede verse fácilmente que las correlaciones parciales pueden construirse a partir de la matriz $\Sigma_{11.2}$ por el método que se construye una correlación simple o total a partir de una matriz de covarianza entonces

$$\rho_{ij,p+1,\dots,p+q} = \frac{\sigma_{ij,p+1,\dots,p+q}}{\sqrt{\sigma_{ii,p+1,\dots,p+q}\ \sigma_{jj,p+1,\dots,p+q}}} \qquad (44)$$

El Coeficiente de Correlación Múltiple.

Definición:

El coeficiente de correlación múltiple es una medida de dependencia existente una variable X_1 y el conjunto $X_2,\dots,X_p = X^{(2)}$

Sea $X = (x_1,\dots,x_p)$ un vector aleatorio y una matriz de covarianza $\Sigma > 0$. Particionando, X y Σ respectivamente así:

$$X = \begin{bmatrix} x_1 \\ \\ x^{(2)} \end{bmatrix} \quad y \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12}' \\ \\ \sigma_{21} & \Sigma_{22} \end{bmatrix} \qquad (45)$$

Como $X^{(2)} = (x_2,\dots,x_p)'$ y Σ_{22} es $(p-1)x(p-1)$ entonces $Var(x_1) = \sigma_{11}$ y $Cov(x_2) = \Sigma_{22}$ y σ_{12} es el vector de covarianza $(p-1)x1$ comprendido entre X_1 y $X^{(2)'}$, llamaremos coeficiente de correlación múltiple entre X_1 y $X^{(2)}$ a la máxima correlación entre X_1 y una combinación lineal de

$\alpha' x^{(2)}$, que se denota por $\bar{R}\ 1,2,\ldots,p$.

Usando la definición se tiene

$$\bar{R}\ 1,2,\ldots,p = \max_{\alpha} \frac{\operatorname{Cov}(x_1,\ \alpha' x^{(2)})}{\sqrt{\operatorname{Var}(x_1)\operatorname{Var}(\alpha' x^{(2)})}} \tag{46}$$

$$= \max_{\alpha} \frac{\alpha' \sigma_{12}}{(\sigma_{11}\ \alpha' \Sigma_{22}\ \alpha)^{1/2}}$$

Luego

$$\frac{\alpha' \sigma_{12}}{(\sigma_{11}\ \alpha' \Sigma_{22}\ \alpha)^{1/2}} = \frac{\alpha' \Sigma_{22}^{-1/2} \Sigma_{22}^{-1/2}\ \sigma_{12}}{(\sigma_{11}\ \alpha' \Sigma_{22}\ \alpha)^{1/2}}$$

$$= \frac{u'v}{(\sigma_{11}\ \alpha' \Sigma_{22}\ \alpha)^{1/2}} \quad \text{cuando} \quad u = \Sigma_{22}^{1/2}\alpha$$

$$v = \Sigma_{22}^{-1/2}\ \sigma_{12}$$

Entonces

$$\frac{u'v}{(\sigma_{11}\ \alpha' \Sigma_{22}\ \alpha)^{1/2}} \leqslant \frac{(u'u)^{1/2}(v'v)^{1/2}}{(\sigma_{11}\ \alpha' \Sigma_{22}\alpha)^{1/2}}$$

por la desigualdad de Cauchy-Schwars.

$$= \frac{(\alpha' \Sigma_{22} \alpha)^{1/2} (\sigma_{12}' \ \Sigma_{22}^{-1} \ \sigma_{12})^{1/2}}{(\sigma_{11} \alpha' \Sigma_{22} \alpha)^{1/2}} = \left[\frac{\sigma_{12}' \ \Sigma_{22}^{-1} \ \sigma_{12}}{\sigma_{11}} \right]^{1/2}$$

Entonces

$$\bar{R}i,q+1,\ldots,p = \max_{\alpha} \frac{\alpha' \ \sigma_{12}}{(\sigma_{11} \alpha' \Sigma_{22} \alpha)^{1/2}} \leqslant \left[\frac{\sigma_{12}' \ \Sigma_{22}^{-1} \ \sigma_{12}}{\sigma_{11}} \right]^{1/2}$$

Son iguales cuando $\alpha = \Sigma_{22}^{-1} \ \sigma_{12}$, reemplazandolo en (46) obtenemos

$$\bar{R}i,q+1,\ldots,p = \frac{\sigma_{12}' \ \Sigma_{22}^{-1} \ \sigma_{12}}{[\sigma_{11} \ \sigma_{12}' \ \Sigma_{22}^{-1} \ \Sigma_{22}^{-1} \Sigma_{22}^{-1} \ \sigma_{12}]^{1/2}}$$

$$= \frac{\sigma_{12}' \ \Sigma_{22}^{-1} \ \sigma_{12}}{(\sigma_{11})^{1/2} [\sigma_{12}' \ \Sigma_{22}^{-1} \sigma_{12}]^{1/2}} ,$$

$$= \left[\frac{\sigma_{12}' \ \Sigma_{22}^{-1} \ \sigma_{12}}{\sigma_{11}} \right]^{1/2}$$

BIBLIOTECA
UNIVERSIDAD DE PANAMA

Si X es $N_p(\mu, \Sigma)$ y X es particionada en X_1 y X_2 la distribución condicional de X_1 dado X_2 tiene como media

$$E(x_1/x_2) = \mu_1 + \sigma_{12}' \Sigma_{22}^{-1}(x_2 - \mu_2)$$

y la varianza

$$Var(x_1/x_2) = \sigma_{11,2,\ldots,p} = \sigma_{11} - \sigma_{12}' \Sigma_{22}^{-1} \sigma_{12}$$

El coeficiente de Correlación $\bar{R}^2 1,2,\ldots,p$ es

$$\bar{R}^2 1,2,\ldots,p = \frac{\sigma_{11} - \sigma_{11,2\ldots p}}{\sigma_{11}}$$

Demostración

Sea $$\bar{R}^2 1,2,\ldots,p = \frac{\sigma_{12}' \Sigma_{22}^{-1} \sigma_{12}}{\sigma_{11}}$$

Entonces

$$1 - \bar{R}^2 1,2,\ldots,p = 1 - \frac{\sigma_{12}' \Sigma_{22}^{-1} \sigma_{12}}{\sigma_{11}}$$

$$= \frac{\sigma_{11} - \sigma_{12}' \Sigma_{22}^{-1} \sigma_{12}}{\sigma_{11}}$$

$$1 - \bar{R}^2\ 1,2,\ldots,p = \frac{\sigma_{11,2,\ldots,p}}{\sigma_{11}}$$

$$\bar{R}^2\ 1,2,\ldots,p = 1 - \frac{\sigma_{11,2,\ldots,p}}{\sigma_{11}}$$

$$= \frac{\sigma_{11} - \sigma_{11,2\ldots,p}}{\sigma_{11}} \qquad (47)$$

En el caso bivariante la matriz de covarianza se reduce

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\ \sigma_2 \\ \rho\sigma_1\ \sigma_2 & \sigma_2^2 \end{bmatrix}$$

Entonces $Var(x_1/x_2) = \sigma_{11.2} = \sigma_1^2(1-\rho^2)$

Luego $\bar{R}^2\ 1,2 = \dfrac{\sigma_1^2 - \sigma_1^2(1-\rho^2)}{\sigma_1^2} = \rho^2$ (48)

Entonces

$\bar{R}\ 1,2, = |\rho|$, que podemos observar que es el valor absoluto del coeficiente de correlación ordinaria.

Si particionamos a X y Σ en

$$X = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

Cuando $X^{(1)}$ es el vector kxl, $X^{(2)}$ es el vector (p-k)x1 Σ_{11} es k x k y Σ_{22} es (p-k)x(p-k).

Una forma de medir la dependencia es considerar correlaciones entre X_i y combinaciones lineales de las variables en $X^{(2)}$ o sea $\alpha' X^{(2)}$, luego como sabemos por demostración anterior el máximo valor de α es $\alpha = \Sigma_{22}^{-1} \sigma_i$ entonces el coeficiente de correlación múltiple es

$$\bar{R}_{i,k+1,\ldots,p} = \left[\frac{\sigma_i' \Sigma_{22}^{-1} \sigma_i}{\sigma_{ii}} \right]^{1/2}$$

Es equivalente

$$\bar{R}^2{}_{i,k+1,\ldots,p} = \frac{\sigma_{ii} - \sigma_{ij,k+1,\ldots p}}{\sigma_{ii}} \qquad (49)$$

cuando $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} = (\sigma_{ij,k+1,\ldots,m})$

en el caso cuando X es normal $\Sigma_{11.2}$ es la matriz de covarianza en la distribución condicional de $X^{(1)}$ dado $X^{(2)}$.

Es importante destacar que de acuerdo con la definición dada el coeficiente de correlación múltiple mide el grado de dependencia lineal existente entre X_i y los componentes de $X^{(2)}$.

El hecho de que $\overline{R}$ sea pequeña significa que las aproximaciones lineales del tipo $\alpha' X^{(2)}$ no son apropiadas para X_i.

CAPITULO II

ESTIMACION DEL VECTOR DE MEDIA Y MATRIZ DE COVARIANZA DE UNA DISTRIBUCION NORMAL.

La distribución normal multivariante está completamente especificada por su vector de media μ y su matriz de covarianzas Σ.

Veremos a continuación que el estimador de maxima verosimilitud de μ es la media muestral ($\bar{x}$) y el de Σ ' es proporcional a la matriz de covarianzas muestral. Una varianza muestral es la suma de cuadrados de las desviaciones de las observaciones respecto de la media muestral dividido por el tamaño muestral menos uno; una covarianza muestral es similarmente definida en término de productos cruzados. Demostraremos que la matriz de covarianzas muestrales es un estimador insesgado de Σ.

Consideremos una muestra de N observaciones de X distribuidas como $N_p(\mu, \Sigma)$, es $X_1,\ldots,X_n$ donde $N > p$ (el número de observaciones debe ser mayor que el número de caracteres).

La función de maxima verosimilitud es

$$\mathcal{L}(x_1,\ldots,x_N, \mu, \Sigma) = \prod_{\alpha=1}^{N} \left[\frac{1}{(2\pi)^{p/2}} \frac{1}{|\Sigma|^{1/2}} \cdot \right.$$

$$\left. \cdot \exp\left\{ -\frac{1}{2}(x_\alpha - \mu)'\Sigma^{-1}(x_\alpha - \mu)\right\} \right] =$$

$$= \frac{1}{(2\pi)^{1/2PN}} \; \frac{1}{\left|\Sigma\right|^{1/2N}} \exp\left\{-\frac{1}{2} \sum_{\alpha=1} (x_\alpha - \mu)\Sigma \; (x_\alpha - \mu)\right\}$$

Para encontrar el Estimador de Maxima verosimilitud de necesitamos del siguiente Lema.

Lema 1:

Sea $X_1, \ldots, X_N$, N vectores (de p componentes) y sea $\overline{X}$ el vector de media. Entonces para cualquier vector b

$$\sum_{\alpha=1}^{N}(x_\alpha - b)(x_\alpha - b)' = \sum_{\alpha=1}^{N}(x_\alpha - \overline{x})(x_\alpha - \overline{x})' + N(\overline{x}-b)(\overline{x}-b)'$$

Demostración

$$\sum_{\alpha=1}^{N}(x_\alpha - b)(x_\alpha - b)' = \sum_{\alpha=1}^{N}\left[(x_\alpha - \overline{x}) + (\overline{x}-b)\right]\left[(x_\alpha - \overline{x}) + (x_\alpha - b)\right]'$$

$$= \sum_{\alpha=1}^{N}(x_\alpha - \overline{x})(x_\alpha - \overline{x})' + \sum_{\alpha=1}^{N}(x_\alpha - \overline{x})(x_\alpha - b)'$$

$$+ \sum_{\alpha=1}^{N}(\overline{x}-b)(x_\alpha - \overline{x})' + \sum_{\alpha=1}^{N}(\overline{x}-b)(x_\alpha - b)'$$

Donde $\sum_{\alpha=1}^{N}(x_\alpha - \bar{x}) = \sum_{\alpha=1}^{N} x_\alpha - \sum_{\alpha=1}^{N}\bar{x} = N\bar{X} - N\bar{X} = 0$

Entonces

$$\sum_{\alpha=1}^{N}(x_\alpha - b)(x_\alpha - b)' = \sum_{\alpha=1}^{N}(x_\alpha - \bar{x})(x_\alpha - \bar{x})' + N(\bar{x}-b)(\bar{x}-b)' \text{L.Q.Q.D}$$

denotamos $\sum_{\alpha=1}^{N}(x_\alpha - \bar{x})(x_\alpha - \bar{x})' = A$ entonces

$$\sum_{\alpha=1}^{N}(x_\alpha - b)(x_\alpha - b)' = A + N(\bar{x}-b)(\bar{x}-b)'$$

Usando este resultado y la propiedad de la traza de una matriz tenemos

$$\sum(x_\alpha - \mu)' \Sigma^{-1}(x_\alpha - \mu) = \mathrm{Tr}(\Sigma^{-1}A) + N(\bar{x}-\mu)' \Sigma^{-1}(\bar{x}-\mu)$$

Demostración

Sea $\psi^* = \Sigma^{-1}$ entonces

$$\sum_{\alpha=1}^{N}(x_\alpha - \mu^*)' \psi^*(x_\alpha - \mu^*) = \mathrm{tr} \sum_{\alpha=1}^{N}(x_\alpha - \mu^*) \psi^*(x_\alpha - \mu^*)'$$

$$= \mathrm{tr} \sum_{\alpha=1}^{N} \psi^*(x_\alpha - \mu^*)(x_\alpha - \mu^*)'$$

$$= tr\ \psi^* \left[A+N(\bar{x}-\mu^*)(\bar{x}-\mu^*)' \right]$$

$$= tr\ \psi^* A + tr\ \psi^* N(\bar{x}-\mu^*)(\bar{x}-\mu^*)'$$

$$= tr\ \psi^* A + tr\ N(\bar{x}-\mu^*)\ \psi^*(\bar{x}-\mu^*)$$

como sabemos que $\psi^* = \Sigma^{-1}$, entonces la densidad de $X_1,\ldots,X_p$ puede ser escrita

$$K \exp\left\{ -\frac{1}{2}\left[N(\bar{x}-\mu)'\,\Sigma^{-1}(\bar{x}-\mu) + tr(\Sigma^{-1}A) \right] \right\}$$

$$= K_1 \exp\left\{ -\frac{1}{2} N(\bar{x}-\mu)'\,\Sigma^{-1}(\bar{x}-\mu) \right\} + K_2 \exp\left[-\frac{1}{2} tr(\Sigma^{-1}A) \right]$$

Así $\bar{x}$ y $(\frac{1}{N})A$ forma un conjunto de estadística suficiente para μ y Σ

Entonces

$$\log \mathcal{L} = \frac{N_p}{2} \log 2\pi + \frac{N}{2} \log |\psi^*| - \frac{1}{2} tr\ \psi^* A - \frac{1}{2} N(\bar{x}-\mu^*)'\psi^*(\bar{x}-\mu)$$

Derivando primero con respecto a μ, tenemos

$$\frac{\partial \log[\mathcal{L}(\mu , \Sigma)]}{\partial\mu} = N\, 2\left[\tfrac{1}{2} \quad \psi^{*}(\bar{x}- \mu)\right]$$ y ψ^{*} es

definida positiva.

Igualando a cero tenemos

$$\hat{\mu} = \bar{X}$$

derivando con respecto a ψ^{*}

$$\frac{\partial \operatorname{Log} \mathcal{L}(\bar{x}, \ \psi^{*})}{\partial\psi^{*}} = - \frac{1}{2} \left[N\, \psi^{-1} \quad \psi^{*-1} A \quad \psi^{*-1} \right]$$

igualando a cero tenemos

$$N \quad \hat{\psi}^{-1} - \hat{\psi}^{-1} A \ \hat{\psi}^{-1} = 0$$

$$\hat{\psi}^{-1} \left[\quad N\hat{\psi} \quad - A\right] \hat{\psi}^{-1} \quad = 0$$

$$\hat{\psi}\ \hat{\psi}^{-1} \left[N\hat{\psi} \quad - A\right] \ \hat{\psi}^{-1}\hat{\psi} \quad = 0$$ multiplicando por $\hat{\psi}$

entonces

$$\hat{\psi} = \frac{1}{N} \ A$$

es ekquivalente $\hat{\Sigma} = \frac{1}{N} \sum_{\alpha=1}^{N} (x_{\alpha} - \bar{x})(x_{\alpha} - \bar{x})'$

Luego el elemento σ_{ij} de la matriz Σ esta determinado por

$$\hat{\sigma}_{ij} = \frac{1}{N} \sum_{\alpha=1}^{N} (x_{\alpha i} - \bar{x}_i)(x_{\alpha j} - \bar{x}_j)$$

Se puede demostrar que $E(\hat{\sigma}_{ij}) = \frac{N-1}{N} \sigma_{ij}$

<u>Demostración</u>:

Se $A = \sum_{\alpha=1}^{N} X_\alpha X'_\alpha - N\bar{X}\bar{X}'$

Utilizando el lema que dice que si $C = (C_{\alpha\beta})$ es ortogonal entonces

$$\sum_{\alpha=1}^{N} X_\alpha X'_\alpha = \sum_{\alpha=1}^{N} Y_\alpha Y'_\alpha$$

cuando $Y_\alpha = \sum_{\beta=1}^{N} c_{\alpha\beta} \; x_\beta$

Prueba

$$\sum_{\alpha} Y_\alpha Y'_\alpha = \sum_{\alpha} \left[\sum_{\beta} c_{\alpha\beta} \; x_\beta \sum_{y} c_{\alpha y} \; X'_y \right]$$

$$= \sum_{\beta y} \left(\sum_{\alpha} c_{\alpha\beta} \; c_{\alpha y} \right) x_\beta \; x_y'$$

$$= \sum_{\beta y} \delta_{\beta y} \, X_\beta \, X'_y$$

cuando $\beta = y$ entonces $\delta_{\beta\beta} = 1$, luego,

$$= \sum_{\beta} X_\beta \, X'_\beta$$

Utilizando este principio tenemos que

$$A = \sum_{\alpha=1}^{N} X_\alpha \, X'_\alpha - N \, \overline{X} \, \overline{X}' = \sum_{\alpha=1}^{N} Z_\alpha \, Z'_\alpha - Z_N Z_N'$$

Suponiendo

$$Z_N = \sqrt{N} \, \overline{X}$$

Entonces

$$N \hat{\Sigma} = \sum_{\alpha=1}^{N-1} Z_\alpha \, Z'_\alpha$$

$$\hat{\Sigma} = \frac{1}{N} \sum_{\alpha=1}^{N-1} Z_\alpha \, Z'_\alpha \quad , \quad \text{con} \quad Z_\alpha \sim N(0, \Sigma)$$

Luego

$$E[\hat{\Sigma}] = \frac{1}{N} \, E\left[\sum_{\alpha=1}^{N-1} Z_\alpha \, Z'_\alpha \right]$$

$$E[\hat{\Sigma}] = \frac{N-1}{N} \Sigma$$

Luego $\hat{\Sigma}$ es un estimador sesgado de Σ

Definimos ahora

$$S = \frac{N}{N-1}\hat{\Sigma} = \frac{1}{N-1} A = \frac{1}{N-1} \sum_{\alpha=1}^{N} (x_\alpha - \bar{x})(x_\alpha - \bar{x})'$$

como una matriz de covarianza muestral.

Ella es un estimador insesgado de Σ y los elementos diagonales son usualmente, Las varianzas muestrales de las componentes de X(insesgada).

Como corolario de la estimación de M.V. de μ y Σ y utilizando la propiedad de invarianza se deduce que el estimador M.V. del coeficiente de correlación simple o total es

$$\hat{\rho}_{ij} = \frac{\sum_{\alpha} (x_{\alpha i} - \bar{x}_i (x_{\alpha j} - \bar{x}_j))}{\sqrt{\sum_{\alpha} (x_{\alpha i} - \bar{x}_i)^2 \sum (x_{\alpha j} - \bar{x}_j)^2}}$$

$$= \frac{\sum_{\alpha} x_{\alpha i} x_{\alpha j} - N \bar{x}_i \bar{x}_j}{\sqrt{\sum_{\alpha} X^2_{\alpha i} - N \bar{X}^2_i} \sqrt{\sum_{\alpha} X^2_{\alpha j} - N\bar{X}^2_j}}$$

Demostración de la propiedad de invarianza $\hat{\rho}_{ij}$ es invariante con respecto a la transformación escalar es decir $\hat{\rho}_{ij} = \hat{\rho}_{ij}^*$ caundo $\hat{\rho}^*_{ij}$ se computa en forma básica, cuando $X^*_{i\alpha} = C_i X_{i\alpha} + d_i$ cuando $c_i > 0$

Solución:

Sea $X^*_{i\alpha}$, la i-ésima componente de X_α y sea

$\bar{X}^*_i$, la i-ésima componente de $\bar{X}$

Dado

$$\hat{\rho}^*_{ij} = \frac{\sum_\alpha (x^*_{i\alpha} - \bar{x}^*_i)(x^*_{i\alpha} - \bar{x}^*_j)'}{\sqrt{\sum_\alpha (x^*_{i\alpha} - \bar{x}^*)^2}\sqrt{\sum_\alpha (x^*_{j\alpha} - \bar{x}_j)^2}}$$

Sea

$$X^*_{i\alpha} = C_i X_{i\alpha} + d_i$$

y $$\bar{X}^*_i = C_i \bar{X}_i + d_i$$

Entonces

$$\hat{\rho}^*_{ij} = \frac{\sum_\alpha (c_i X_{i\alpha} + d_i - C_i \bar{x}_i - d_i)(c_j x_{j\alpha} + d_j - c_j \bar{x}_j - d_j)}{\sqrt{\sum_\alpha (c_i x_{i\alpha} + d_i - c_i \bar{x}_i - d_i)^2}\sqrt{\sum_\alpha (c_j x_{j\alpha} - c_j \bar{x}_j)^2}}$$

$$\hat{\rho}^{*}_{ij} = \frac{\sum_{\alpha} c_i (x_{i\alpha} - \bar{x}_i)\; c_j (x_j - \bar{x}_j)}{\sqrt{\sum_{\alpha} c_i^2 (x_{i\alpha} - \bar{x}_i)^2}\;\sqrt{\sum_{\alpha} c_j^2 (x_j - \bar{x}_j)^2}}$$

$$= \frac{\cancel{c_i c_j} \sum_{\alpha} (x_{i\alpha} - \bar{x}_i)(x_{j\alpha} - \bar{x}_j)}{\cancel{c_i c_j} \sqrt{\sum_{\alpha} (x_{i\alpha} - \bar{x}_i)^2}\;\sqrt{\sum_{\alpha} (x_{j\alpha} - \bar{x}_j)^2}}$$

$$\hat{\rho}^{*}_{ij} = \frac{\sum_{\alpha} (x_{\alpha} - \bar{x}_i)(x_{j\alpha} - \bar{x}_j)}{\sqrt{\sum_{\alpha} (x_{i\alpha} - \bar{x}_i)^2}\;\sqrt{\sum_{\alpha} (x_{j\alpha} - \bar{x}_j)^2}} = \hat{\rho}_{ij} \quad \text{L.Q.Q.D}$$

Coeficiente de Correlación Parcial.

El coeficiente de correlación parcial es un coeficiente de correlación con distribuciones condicionales normales.

Sea $g(x_1/x_2) = \dfrac{f(x_1, x_2)}{h(x_2)}$, la función de (1) densidad condicional de $X^{(1)}$ (con q componentes) conociendo $X^{(2)} = x^{(2)}$.

Entonces para calcular el vector de media y la matriz de covarianza, calcularemos (1).

Luego

$$h(x_2) = \frac{1}{(2\pi)^{q/2}|\Sigma_{22}|^{1/2}} \exp\left[-\frac{1}{2}(x_2-\mu_2)'\,\Sigma_{22}^{-1}(x_2-\mu_2)\right]$$

y por teoría de matrices

$$\Sigma^{-1} = \begin{bmatrix} (\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')^{-1} & -(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\ -\Sigma_{22}^{-1}\Sigma_{12}'(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')^{-1} & \Sigma_{22}^{-1}+\Sigma_{22}^{-1}\Sigma_{12}'(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')^{-1}\cdot \\ & \cdot\,\Sigma_{12}\Sigma_{22}^{-1} \end{bmatrix}$$

donde el determinante es

$$|\Sigma| = |\Sigma_{22}| \cdot |\Sigma_{11}-\Sigma_{12}\;\Sigma_{22}^{-1}\;\Sigma_{12}'|$$

Entonces

$$f(x_1,x_2) = \frac{1}{(2\pi)^{1/2(p+q)}|\Sigma_{22}|^{1/2}|\Sigma_{11}-\Sigma_{12}\;\Sigma_{22}^{-1}\;\Sigma_{12}'|}\cdot$$

$$. \exp \left\{ - \frac{1}{2} \left[(x_i - \mu_i)' (\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}')(x_1 - \mu_1) \right.\right.$$

$$- (x_2 - \mu_2)' \Sigma_{22}^{-1} \Sigma_{12}' (\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}')^{-1} (x_1 - \mu_1)$$

$$- (x_1 - \mu_1)' (\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}')^{-1} \Sigma_{12} \Sigma_{22}^{-1} (x_2 - \mu_2)$$

$$+ (x_2 - \mu_2)' \Sigma_{22}^{-1} \Sigma_{12}' (\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}')^{-1} \Sigma_{12} \Sigma_{22}^{-1} (x_2 - \mu_2)$$

$$\left. + (x_2 - \mu_2)' \Sigma_{22}^{-1} (x_2 - \mu_2) \right\}$$

Luego

$$g(x_1/x_2) = \frac{1}{(2\pi)^{p/2} | \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}' |^{1/2}} \cdot$$

$$. \exp \left\{ - 1/2 \left[x_1 - \mu_1 - \Sigma_{12} \Sigma_{22}^{-1} (x_2 - \mu_2) \right] (\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}')^{-1} \right.$$

$$\left. \left[x_1 - \mu_1 - \Sigma_{12} \Sigma_{22}^{-1} (x_2 - \mu_2) \right] \right\}$$

El vector de media es $\mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (x_2 - \mu_2)$ y la matriz de covarianza $\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}'$.

Entonces si $X \sim N(\mu\ , \ \Sigma\ )$; $f(x^{(1)}/x^{(2)})$ es

$$x^{(2)}/x^{(2)} \sim N\left[(\mu\ ' + \beta\ (x^{(2)} - \mu^{(2)}),\ \Sigma_{11.2} \right]$$

donde $\beta = \Sigma_{12}\ \Sigma_{22}^{-1}$ (1)

$$\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\ \Sigma_{22}^{-1}\ \Sigma_{21} \qquad (2)$$

El estimador de maxima verosimilitud del coeficiente de correlación parcial de $X^{(1)}$ (con q componentes) es $\hat{\rho}_{ij}.q+1,\ldots,p$, por teoría sabemos que el estimador de máxima verosimilitud de Σ es $\hat{\Sigma} = \frac{1}{N} \sum_{\alpha=1}^{N} (x_\alpha - \bar{x})(x_\alpha - \bar{x})'$

cuando $\bar{x} = \frac{1}{N} \sum X_\alpha$. En virtud de (1) y (2) podemos deducir que

$$\Sigma_{12} = \beta\ \Sigma_{22}$$

$$\Sigma_{11} = \Sigma_{11.2} + \beta\ \Sigma_{22}\ \beta'$$

Los estimadores de $\Sigma_{11.2}$, β y Σ_{22} son

$$\hat{\Sigma}_{11.2} = \hat{\Sigma}_{11} - \hat{\Sigma}_{12}\ \hat{\Sigma}_{22}^{-1}\ \hat{\Sigma}_{21},\quad \hat{\beta} = \hat{\Sigma}_{12}\ \hat{\Sigma}_{22}^{-1} \text{ y } \hat{\Sigma}_{22}$$

El estimador de máxima verosimilitud de el coeficiente de correlación paracial es

$$\hat{\rho}_{ij.q+1,\dots,p} = \frac{\hat{\sigma}_{ij,q+1,\dots,p}}{\sqrt{\hat{\sigma}_{ii,q+1,\dots,p}}\sqrt{\hat{\sigma}_{jj,q+1,\dots,p}}}$$

$$i,j = 1,\dots,q$$

Luego el elemento ij-ésimo ($i \neq j$) de $\Sigma_{11.2}$ de representa por $\hat{\sigma}_{ij,q+1,\dots,p}$; donde $\hat{\Sigma}_{11.2} = \hat{\Sigma}_{11} - \hat{\Sigma}_{12}\hat{\Sigma}_{22}^{-1}\hat{\Sigma}_{21}$

y el elemento ii-ésimo de $\Sigma_{11.2}$ se representa por $\sigma_{ii(k+1),\dots,p}$.

Analogamente conocido el coeficiente de correlación múltiple,

$$\bar{R} = \frac{\beta\Sigma_{22}\beta'}{\sqrt{\sigma_{11}\beta\Sigma_{22}\beta'}} = \sqrt{\frac{\beta\Sigma_{22}\beta'}{\sigma_{11}}}$$

$$= \sqrt{\frac{\sigma_{(1)}\Sigma_{22}^{-1}\sigma_{(1)}'}{\sigma_{11}}}$$

donde β , $\sigma_{(1)}$ y Σ_{22} son definidas por

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{(1)} \\ \sigma_{(1)} & \Sigma_{22} \end{bmatrix} \quad y \quad \beta = \sigma_{(1)}\Sigma_{22}^{-1}$$

por teoría el estimador de maxima verosimilitud de Σ_{22}^{-1} es

$$\hat{\Sigma} = \frac{1}{N} A = \frac{1}{N} \left[\sum_{\alpha=1} (x_\alpha - \bar{x})(x_\alpha - \bar{x})' \right]$$

$$= \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{(1)} \\ \hat{\sigma}'_{(1)} & \hat{\Sigma}_{22} \end{bmatrix}$$

y el estimador de β es

$$\hat{\beta} = \hat{\sigma}_{(1)} \quad \hat{\Sigma} A_{22}^{-1} = a_{(1)} \quad {}^{-1}_{22}$$

el estimador del coeficiente de correlación múltiple es

$$R = \sqrt{\frac{\hat{\beta}\ \hat{\Sigma}_{22}^{-1}\ \hat{\beta}'}{\hat{\sigma}_{(1)}}} = \sqrt{\frac{\hat{\sigma}_{(1)}\ \hat{\Sigma}_{22}^{-1}\ \hat{\sigma}'_{(1)}}{\hat{\sigma}_{(1)}}}$$

$$= \sqrt{\frac{a_{(1)}\ A_{22}^{-1}\ a'_{(1)}}{a_{11}}}$$

CAPITULO III

DOCIMAS E INTERVALOS DE CONFIANZA PARA LOS COEFICIENTES DE CORRELACION.

Hasta ahora sólo nos hemos interesado por la obtención de un estimador puntual para un parámetro desconocido. Existe otro enfoque que conduce a menudo a resultados muy significativos, para proceder a la estimación. Puede seguirse el criterio de definir un intervalo al que presumiblemente, pertenezca el parámetro a estimar; vamos a ocuparnos, por ello, de obtener estimación de este carácter, a cuyos efectos definiremos como intervalo de confianza a aquel intervalo del espacio paramétrico al que, con una cierta probabilidad, pertenece el parámetro.

Para determinar la función de r utilizaremos el siguiente lema.

Lema

Si $(z_{1\alpha}, z_{2\alpha})$, $\alpha = 1, \ldots, n$ son independientes cada par con distribución

$$N\left[\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_1^2 & \sigma_1\sigma_2\rho\\ \sigma_2\sigma_1\rho & \sigma_2^2\end{bmatrix}\right]$$

entonces la distribución condicional de

$$b = \sum_{\alpha} z_{2\alpha} \quad z_{1\alpha} \Big/ \sum_{\alpha} z_{1\alpha}^{2} \quad y$$

$$\frac{V}{\sigma^2} = \sum_{\alpha} (z_{2\alpha} - bz_{1\alpha})^2 / \sigma$$

dado $z_{1\alpha} = z_{1\alpha}$ (α =1,...,n) es $N(\beta\, ,\ \sigma^2/c^2)$ donde

$(c^2 = \sum_{\alpha} z_{1\alpha}^{2})$ y de χ^2 con n-1 grados de libertad respectivamente y b y V son independiente

Si ρ =0 entonces β = 0 y b tiene distribución condicional de acuerdo a $N(0,\ \sigma^2/c^2)$.

Luego $\frac{cb}{\sigma} \sim N(0,1)$ y $\frac{V/\sigma^2}{n-1} \sim \chi^2$ con n-1 grado de libertad.

Luego por teoría conocida

$$\frac{cb/\sigma}{\sqrt{\frac{V/\sigma^2}{n-1}}} = \frac{cb}{\sqrt{\frac{V}{n-1}}} \sim t_{n-1} \text{ grado de libertad}$$

como $b = \frac{a_{12}}{a_{11}}$ y $V = a_{22} - a_{12}^2/a_{11}$ entonces

$$\sqrt{n-1} \quad \frac{cb}{\sqrt{V}} = \sqrt{n\ 1} \quad \frac{a_{11}\frac{a_{22}}{a_{11}}}{\sqrt{a_{22}-a_{12}^2/a_{11}}}$$

$$= \sqrt{n-1} \qquad \frac{a_{12}/\sqrt{a_{11}a_{22}}}{\sqrt{1 - \left[a_{12}^2/a_{11}a_{22}\right]}}$$

$$= \sqrt{n-1} \quad \frac{r}{\sqrt{1-r^2}} \sim t_{n-1}$$

Así $\sqrt{n-1} \quad \dfrac{r}{\sqrt{1-r^2}}$ tiene una t-distribución condicional con n-1 grados, la densidad de t es

$$\frac{\Gamma(\frac{1}{2}\, n)}{\sqrt{n-1}\; \Gamma\left[\frac{1}{2}(n-1)\right]\sqrt{\pi}} \left[1 + \frac{t^2}{n-1}\right]^{-1/2\, n}$$

La función de densidad de $w = r/\sqrt{1-r^2}$

$$f(w) = \frac{\Gamma(1/2\ n)}{\Gamma\left[\frac{1}{2}(n-1)\right]\sqrt{\pi}} \quad (1 + w^2)^{-1/2\ n}$$

$w = r(1-r^2)^{-1/2}$ y $dw/dr = (1-r^2)^{-3/2}$

Entonces

$$f(r) = \frac{\Gamma\left[\frac{1}{2}(n-1)\right]}{\Gamma\left[\frac{1}{2}(n-1-1)\right]\sqrt{\pi}} \left[1 + \frac{r^2}{1-r^2}\right]^{-1/2n} (1-r^2)^{-3/2}$$

por ser

$$w = r\,(1-r^2)^{-1/2} \quad \text{entonces} \quad w^2 = \frac{r^2}{1-r^2}$$

Luego

$$f(r) = \frac{\Gamma\left[\frac{1}{2}(N-1)\right]}{\Gamma\left[\frac{1}{2}(N-2)\right]\sqrt{\pi}} \left[\frac{1-r^2+\ r^2}{1-r^2}\right]^{-\frac{1}{2}n} (1-r^2)^{-3/2}$$

$$f(r) = \frac{\Gamma\left[\frac{1}{2}(N-1)\right]}{\Gamma\left[\frac{1}{2}(N-2)\right]\sqrt{\pi}} \left[\frac{1}{(1-r^2)}\right]^{-\frac{1}{2}n} (1-r^2)^{-3/2}$$

$$= \frac{\Gamma\left[\frac{1}{2}(N-1)\right]}{\Gamma\left[\frac{1}{2}(N-2)\right]\sqrt{\pi}} (1-r^2)^{\frac{1}{2}n} \quad (1-r^2)^{-3/2}$$

$$= \frac{\Gamma\left[\frac{1}{2}(N-1)\right]}{\Gamma\left[\frac{1}{2}(N-2)\right]\sqrt{\pi}} (1-r^2)^{\frac{1}{2}(n-3)}$$

$$= \frac{\Gamma\left[\frac{1}{2}(N-1)\right]}{\Gamma\left[\frac{1}{2}(N-2)\right]\sqrt{\pi}} (1-r^2)^{\frac{1}{2}(N-1-3)}$$

$$f(r) = \frac{\Gamma\left[\frac{1}{2}(N-1)\right]}{\Gamma\left[\frac{1}{2}(N-2)\right]\sqrt{\pi}} (1-r^2)^{\frac{1}{2}(N-4)}$$

Utilizando la demostración anterior podemos determinar que si $X_1,\ldots,X_n$ es una muestra aleatoria de una distribución multivariante $N(\mu, \Sigma)$. Llamemos r_{ij} al coeficiente de correlación muestral entre las componentes i-esima y j-ésima del vector observado. Analogamente llamemos ρ_{ij} al valor poblacional de la correlación.

En estas condiciones es válido el siguiente Teorema

Si $\rho_{ij} = 0$, entonces la densidad de r_{ij} es

$$f_N(r) = \frac{\Gamma\left[\frac{1}{2}(N-1)\right]}{\Gamma\left[\frac{1}{2}(N-2)\right]\sqrt{\pi}} (1-r^2)^{\frac{1}{2}(N-4)} \qquad (50)$$

Además en ese caso, la variable $w = \sqrt{N-2}\,r / \sqrt{1-r}$ tiene una distribución t con N-2 grados de libertad.

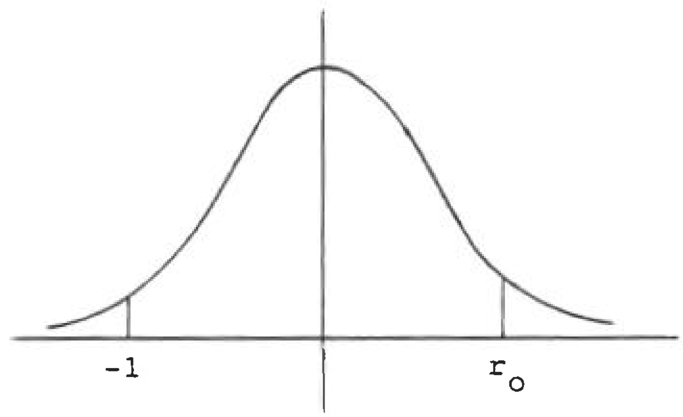
El uso más importante del teorema es en la docimasia de la hipótesis de que dos variables no están correlacionadas

$$H_o : \rho_{ij} = 0$$

Hay tres casos a considerar

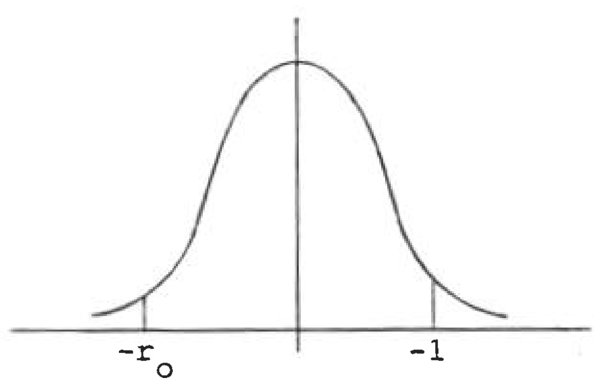
i) la hipótesis alternativa es H_1: $\rho_{ij} > 0$. En este caso rechazamos H_o si el coeficiente de correlación muestral es mayor que cierto número r_o. El nivel de significación viene dado en este caso por

$$\alpha = \int_{r_o}^{1} f_N(r)dr \qquad (51)$$



donde $f_N(r)$ viene dada por la fórmula (50)

ii) La hipótesis alternativa es H_1: $\rho_{ij} < 0$. Este problema es análogo al anterior, dada la simetría respecto al origen de $f_N(r)$, una región critica con nivel α es $\{ r_{ij} < -r_o \}$ donde r_o fue calculado en (51)

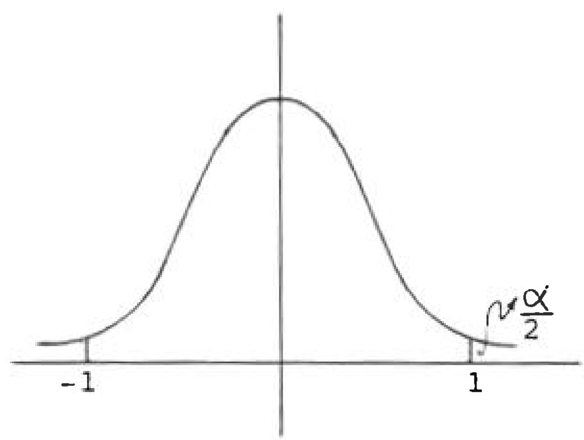


iii) la hipótesis alternativa es H_1: $\rho_{ij} \neq 0$. En este caso rechazamos H_o si $r_{ij} > r_o$ ó $r_{ij} < -r_o$ ó sea $|r_{ij}| > r_o$, donde

$$\int_{r_o}^{1} f(r)dr = \frac{\alpha}{2}$$

Luego el nivel de significación es

$$\alpha = \int_{-1}^{-r_o} f_N(r)dr + \int_{r_o}^{1} f_N(r)dr \qquad (52\ )$$



El número r_o debera calcularse de modo que cumpla con (52).

Los puntos de significación r_o están dados en muchos libros; en particular en la tabla VI de Fisher y Yates, "Statics Tables for Biological, Agricultural and Medical Research" (2da edición 1942, Oliver and Boyd); el indice n en la tabla V es igual a nuestro N-2. Dado que w= $\sqrt{N-2}\, r/\sqrt{1-r}$ tiene distribución t con N-2 g.L, también puede usarse las tablas t, contra alternativas $\rho_{ij} \neq 0$ se rechaza H_o si

$$\sqrt{N-2} \quad \frac{|r_{ij}|}{\sqrt{1-r_{ij}^2}} > t_{N-2}(\alpha/2)$$

donte $t_{N-2}(\alpha/2)$ es el punto de significación para una docima t a dos ramas con nivel $\alpha/2$

Si la hipótesis alternativa es $\rho_{ij} > 0$, rechazaremos H_o si

$$\sqrt{N-2} \quad \frac{r_{ij}}{\sqrt{1-r_{ij}^2}} > t_{N-2}(\alpha)$$

donde $t_{N-2}(\alpha)$ es el punto de significación dócimas a una sola rama.

Si la hipótesis alternativa es $\rho_{ij} < 0$, rechazaremos H_o si

$$\sqrt{N-2} \; \frac{r_{ij}}{\sqrt{1-r_{ij}^2}} > -t_{N-2}(\alpha)$$

donde $-t_{N-2}(\alpha)$ también es el punto de significación para dócimas a una sola rama.

Distribución del Coeficiente de Correlación muestral cuando $\rho \neq 0$; Pruebas de Hipótesis.

Teorema:

Consideremos una muestra aleatoria tamaño N de una distribución normal bivariante con correlación ρ . Entonces el coeficiente de correlación muestral calculado a partir de dichas observaciones tiene una función de densidad

$$f_{\rho,N}(r) = \frac{2^{n-2}(1-\rho^2)^{n/2}(1-r^2)^{\frac{1}{2}(n-3)}}{(n-2)!\,\pi} .$$

$$\cdot \sum_{\alpha=0}^{\infty} \frac{(2 \rho r)^{\alpha}}{\alpha !} \Gamma^{2}\left[\frac{1}{2}(n+\alpha)\right]$$

donde $n = N-1$ (53)

Demostración:

Sea la función de densidad de $a_{11} \sim g_n(a_{11}/ \ _1^{\ 2})$ distribuida χ_n^2, la de b distribuida $N(b/\beta \ , \sigma^2/c^2)$ y la de V distribuida $g_{N-1}(v/\sigma^2)/\sigma^2$ como χ^2_{n-1}, entonces

$$f(b,V,a_{11}) = g_n\left[a_{11}/\sigma^2\right] \quad N(b/\beta \ , \sigma^2/c^2) g_{N-1}(v/\sigma^2)/\sigma$$

Luego

$$f(b,V,a_{11}) = \frac{(a_{11})^{\frac{1}{2}n-1}}{(2\sigma_1^{\ 2})^{\frac{1}{2}}\Gamma(\frac{1}{2}n)} \exp(-\frac{1}{2\sigma_2} a_{11}) \frac{\sqrt{a_{11}}}{\sqrt{2\pi\sigma^2}} \exp\left[\right.$$

$$\left. - \frac{a_{11}}{2\sigma^2}(b-\beta)^2\right] \cdot \frac{1 \qquad\qquad v^{\frac{1}{2}(n-2)}}{(2\sigma^2)^{\frac{1}{2}(n-1)}\Gamma\left[\frac{1}{2}(n-1)\right]}$$

$$\exp(-\frac{1}{2\sigma^2} V)$$

con

$$b = \frac{a_{12}}{a_{11}}, \quad V = a_{22} - \frac{a^2_{12}}{a_{11}} \quad y \quad \beta = \rho \frac{\sigma_2}{\sigma_1}$$

donde

$$\frac{\partial b}{\partial a_{12}} = \frac{1}{a_{11}}, \quad \frac{\partial b}{\partial a_{22}} = 0 \; ; \quad \frac{\partial V}{\partial a_{12}} = - 2\frac{a_{12}}{a_{11}} \quad y \quad \frac{\partial V}{\partial a_{22}} = 1$$

El Jacobiano

$$\frac{\partial (b,V)}{\partial (a_{12},a_{22})} = \begin{vmatrix} \frac{1}{a_{11}} & 0 \\ - \frac{2a_{12}}{a_{11}} & 1 \end{vmatrix} = \frac{1}{a_{11}}$$

Entonces

$$f(a_{11},a_{12},a_{22}) = f\,(b,V,a_{11}) \cdot \frac{1}{a_{11}}$$

$$f(a_{11},a_{12},a_{22}) = \frac{a_{11}^{\frac{1}{2}n-1} e^{-\frac{1}{2\sigma_1^2} a_{11}} \sqrt{a_{11}}}{(2\sigma_1^2)^{\frac{1}{2}n}\, \Gamma(\frac{1}{2}n) \sqrt{2\pi\sigma^2}}\, e^{-\frac{a_{11}}{2\sigma^2}(b-\beta)^2} .$$

$$\cdot \frac{\left[\dfrac{a_{11}a_{22}-a_{12}^2}{a_{11}}\right]^{\frac{1}{2}(n-3)}}{(2\,\sigma^2)^{\frac{1}{2}(n-1)}\Gamma\left[\frac{1}{2}(n-1)\right]}\, e^{-\frac{1}{2\sigma^2}\left[\frac{(a_{11}a_{22}-a_{12}}{a_{11}}\right]} \cdot \frac{1}{a_{11}}$$

donde

$$a_{11}^{\frac{1}{2}n-1} \cdot a_{11}^{1/2} \cdot a_{11}^{-1} = a_{11}^{\frac{1}{2}n-1-1+1/2} = a_{11}^{\frac{1}{2}n-3/2} = a_{11}^{\frac{1}{2}(n-3)}$$

$$e^{-\frac{a_{11}}{2\,\sigma^2}} \cdot e^{-\frac{a_{11}}{2\sigma^2}(b-\beta)^2} \cdot e^{-\frac{1}{2\,\sigma^2}(a_{22}-\frac{a_{12}^2}{a_{11}})}$$

Luego

$$\frac{a_{11}}{\sigma^2} + \frac{a_{11}}{\sigma^2}(b-\beta)^2 + \frac{1}{\sigma^2}(a_{22} - \frac{a_{12}^2}{a_{11}})$$

$$\frac{a_{11}}{\sigma^2} + \frac{a_{11}}{\sigma^2}(b^2-2b\beta + \beta^2) + \frac{1}{\sigma^2}(a_{22}-\frac{a_{12}^2}{a_{11}})$$

Entonces

$$Q = \frac{a_{11}}{\sigma^2} + \frac{a_{11}}{\sigma^2}\left[\frac{a_{12}^2}{a_{11}^2} - 2\,\frac{a_{11}}{a_{11}}\rho\,\frac{\sigma_1\;\sigma_2}{\sigma_1^2} + \right.$$

$$+ \quad \rho^2 \frac{\sigma_1^2 \; \sigma_2^2}{\sigma_1^2} \Bigg] + \frac{1}{\sigma^2}\left(a_{21} - \frac{a_{22}^2}{a_{11}}\right)$$

$f(a_{11}, a_{12}, a_{22}) =$

$$= \frac{a_{11}^{\frac{1}{2}(n-3)} \left[\dfrac{a_{11}a_{22}-a_{12}^2}{a_{11}} \right]^{\frac{1}{2}(n-3)} e^{-\frac{1}{2}Q}}{2^n \sigma_1^n \left[\dfrac{\sigma_1^2 \sigma_2^2 - \rho^2 \sigma_1^2 \sigma_2^2}{\sigma_1^2} \right]^{1/2n} \sqrt{\pi}\, \Gamma\left[\frac{1}{2}n\right] \Gamma\left[\frac{1}{2}(n-1)\right]}$$

conocida f (a_{11}, a_{12}, a_{22}), buscaremos $f(a_{11}, a_{22}, r)$ cuando

$$r = a_{12} \Big/ \sqrt{a_{11}a_{22}} \quad y \quad \frac{d\, a_{12}}{dr} = \sqrt{a_{11}a_{22}}$$

Entonces

$f(a_{11}, a_{22}, r) =$

$$= \frac{a_{11}^{\frac{1}{2}(n-3)} \, a_{11}^{-\frac{1}{2}(n-3)} (a_{11}a_{22} - a_{12}^2)^{\frac{1}{2}(n-3)} e^{-\frac{1}{2}Q} \, a_{11}^{\frac{1}{2}} a_{22}^{1/2}}{2^n \sigma_1^n \, \sigma_1^{-n} (\sigma_1^2 \sigma_2^2 - \rho^2 \sigma_1^2 \sigma_2^2)^{\frac{1}{2}n} \sqrt{\pi}\, \Gamma\left[\frac{1}{2}n\right] \Gamma\left[\frac{1}{2}(n-1)\right]}$$

$$= \frac{a_{11}^{\frac{1}{2}(n-2)} \; a_{22}^{\frac{1}{2}(n-2)} \; (1-r^2)^{\frac{1}{2}(n-3)} \; e^{-\frac{1}{2}\left[\frac{1}{(1-\rho^2)}\left(\frac{a_{11}}{\sigma_1^2} - \frac{2\rho r\sqrt{a_{11}}\sqrt{a_{22}}}{\sigma_1\sigma_2} + \frac{a_{22}}{\sigma_2^2}\right.\right.}}{2^n \left[\sigma_1^2 \; \sigma_2^2(1-\rho^2)\right]^{\frac{1}{2}n} \sqrt{\pi} \; \Gamma\left[\frac{1}{2}n\right]\Gamma\left[\frac{1}{2}(n-1)\right]}$$

por ser una distribución bivariante

$$f(r) = \int_0^\infty \int_0^\infty \frac{a_{11}^{\frac{1}{2}n-1} a_{22}^{\frac{1}{2}\,n-1} (1-r^2)^{\frac{1}{2}(n-3)} \quad \frac{1}{2}\left[\frac{1}{(1-\rho^2)} \cdot Q^*\right]}{2^n \left[\sigma_1^2 \; \sigma_2^2(1-\rho^2)\right]^{\frac{1}{2}n} \sqrt{\pi} \; \Gamma\left[\frac{1}{2}n\right]\Gamma\left[\frac{1}{2}(n-1)\right]}$$

Donde

$$Q^* = \frac{a_{11}}{\sigma_1^2} - 2\rho r \frac{\sqrt{a_{11}}\sqrt{a_{22}}}{\sigma_1 \; \sigma_2} + \frac{a_{22}}{\sigma_2^2}$$

Como

$$\exp\left[\frac{\rho r \sqrt{a_{11}}\sqrt{a_{22}}}{(1-\rho^2)\,\sigma_1\,\sigma_2}\right] = \sum_{\alpha=0}^{\infty} \frac{(\rho r \sqrt{a_{11}}\sqrt{a_{22}})}{\alpha!\left[\sigma_2 \; \sigma_1(1-\rho^2)\right]^{\alpha}}$$

Entonces

$$f(r) = \frac{(1-r^2)^{\frac{1}{2}(n-3)}}{\sigma_1^n \sigma_2^n (1-\rho^2)^{1/2n} \, 2^n \sqrt{\pi}\, \Gamma(\frac{1}{2}n)\, \Gamma[\frac{1}{2}(n-1)]} \cdot$$

$$\cdot \sum_{\alpha=0}^{\infty} \frac{(\rho r \sqrt{a_{11}}\sqrt{a_{22}})}{\alpha! [\sigma_2 \sigma_1 (1-\rho^2)]^{\alpha}} \cdot \int_0^{\infty} \int_0^{\infty} a_{11}^{\frac{1}{2}n-1} a_{22}^{\frac{1}{2}n-1} \cdot$$

$$\cdot e^{-\frac{a_{11}}{2(1-\rho^2)\sigma_2^2}} e^{-\frac{a_{22}}{2(1-\rho^2)\sigma_2^2}} da_{11} da_{22}$$

$$f(r) = \frac{(1-r^2)^{\frac{1}{2}(n-3)}}{\sigma_1^n \sigma_2^n (1-\rho^2)^{\frac{1}{2}n} \, 2^n \sqrt{\pi}\, \Gamma(\frac{1}{2}n)\, \Gamma[\frac{1}{2}(n-1)]} \cdot$$

$$\sum_{\alpha=0}^{\infty} \frac{(\rho r \sqrt{a_{11}}\sqrt{a_{22}})^{\alpha}}{\alpha! [\sigma_2 \sigma_1 (1-\rho^2)]^{\alpha}} \int_0^{\infty} a_{11}^{\frac{1}{2}n-1} \exp\left[-\frac{a_{11}}{2(a-\rho^2)\sigma_1^2}\right] da_{11} \cdot$$

$$\cdot \int_0^{\infty} a_{22}^{\frac{1}{2}n-1} \exp\left[-\frac{a_{22}}{2(1-\rho^2)\sigma_2^2}\right] da_{22}$$

$$f(r) = \frac{(1-r^2)^{\frac{1}{2}(n-3)}}{\sigma_1^{\,n}\ \sigma_2^{\,n}(1-\rho^2)^{\frac{1}{2}n}\ 2^n\sqrt{\pi}\,\Gamma(\tfrac{1}{2}n)\,\Gamma\left[\tfrac{1}{2}(n-1)\right]}\ \cdot$$

$$\sum_{\alpha=0}^{\infty} \frac{(\rho\, r)^{\alpha}}{\alpha!\left[\sigma_2\ \sigma_1(1-\rho^2)\right]^{\alpha}} \cdot \int_0^{\infty} a_{11}^{\frac{1}{2}(n+\alpha)-1}\ \cdot$$

$$\cdot \exp\left[-\frac{a_{11}}{2(1-\rho^2)\ \sigma_1^{\,2}}\right] da_{11} \cdot \int_0^{\infty} a_{22}^{\frac{1}{2}(n+\alpha)-1}\ \cdot$$

$$\cdot \exp\left[-\frac{a_{22}}{2(1-\rho^2)\ \sigma_2^{\,2}}\right] da_{22}$$

Como

$$\int_0^{\infty} a_{11}^{\frac{1}{2}(n+\alpha)-1} \exp\left[-\frac{a_{11}}{2(1-\rho^2)\ \sigma_1^{\,2}}\right] da_{11} =$$

$$= \Gamma\left[\tfrac{1}{2}(n+\alpha)\right]\left[2\sigma_1^{2}(1-\rho^{2})\right]^{\frac{1}{2}(n+\alpha)}$$

$$y \int_0^{\infty} a_{22}^{\frac{1}{2}(n+\alpha)-1} \exp\left[-\frac{a_{22}}{2(1-\rho^2)\sigma_2^{2}}\right] da_{22} =$$

$$= \left[\tfrac{1}{2}(n+\alpha)\right]\left[2\sigma_2^{2}(1-\rho^{2})\right]^{\frac{1}{2}(n+\alpha)}$$

Entonces

$$f(r) = \frac{(1-r^2)^{\frac{1}{2}(n-3)}}{\sigma_1^{n}\,\sigma_2^{n}(1-\rho^2)^{\frac{1}{2}n}\,2^{n}\sqrt{\pi}\;\Gamma(\tfrac{1}{2}n)\,\Gamma\left[\tfrac{1}{2}(n-1)\right]} \cdot$$

$$\sum_{\alpha=0}^{\infty} \frac{(\rho r)^{\alpha}}{\alpha!(1-\rho^2)^{\alpha}\,\sigma_1^{\alpha}\sigma_2^{\alpha}} \cdot \Gamma^{2}\left[\tfrac{1}{2}(n+\alpha)\right] \cdot$$

$$\cdot\; 2^{n+\alpha}\,\sigma_1^{n+\alpha}\sigma_2^{n+\alpha}\,(1-\rho^2)^{n+\alpha}$$

$$f(r) = \frac{(1-\rho^2)^{\frac{1}{n}n}(1-r^2)^{\frac{1}{2}(n-3)}}{\sqrt{\pi}\,\Gamma(\frac{1}{2}n)\,\Gamma\left[\frac{1}{2}(n-1)\right]} \sum_{\alpha=0}^{\infty} \frac{(2\rho r)}{\alpha!}\,\Gamma^2\left[\frac{1}{2}(n+\alpha)\right]$$

Utilizando el principio de que

$$\Gamma(\tfrac{1}{2}z)\,\Gamma(z+1/2) = \sqrt{\pi}\,\Gamma(2z)/2^{2z-1}$$

lo aplicamos en

$$\Gamma(\tfrac{1}{2}n)\,\Gamma\left[\tfrac{1}{2}(n-1)\right] = \Gamma\left[\tfrac{1}{2}(n-1)+\tfrac{1}{2}\right]\Gamma\left[\tfrac{1}{2}(n-1)\right]$$

$$= \sqrt{\pi}\,\Gamma\left[2\cdot\tfrac{1}{2}(n-1)\right]\Big/2^{2\cdot\frac{1}{2}(n-1)-1}$$

$$= \sqrt{\pi}\,\Gamma\left[(n-1)\right]\Big/2^{n-2} = \frac{\sqrt{\pi}(n-2)!}{2^{n-2}}$$

$$f(r) = \frac{(1-\rho^2)^{\frac{1}{2}n}(1-\sigma^2)^{\frac{1}{2}(n-3)}}{\sqrt{\pi}\,\frac{\sqrt{\pi}(n-2)!}{2^{n-2}}} \sum_{\alpha=0}^{\infty}\frac{(2\rho r)^2}{\alpha!}\cdot$$

$$\cdot\,\Gamma^2\left[\tfrac{1}{2}(n+\alpha)\right]$$

$$f(r) = \frac{2^{n-2}(1-\rho^2)^{\frac{1}{2}n}(1-r^2)^{\frac{1}{2}(n-3)}}{\pi(n-2)!} .$$

$$\cdot \sum_{\alpha=0}^{\infty} \frac{(2\rho r)^2}{\alpha!} \quad \Gamma^2\left[\tfrac{1}{2}(n+\alpha)\right] \quad \text{L.Q.Q.D}$$

Si r es una variable aleatoria, entonces, para cualquier valor de r^* existe la probabilidad $Pr(r \leqslant r^*)$.

Claro que $Pr(r \leqslant r^*)$ depende de la elección de r^*, esta es una función de r^* que se llama función de distribución acumulativa de r y se representa por $F(r^*/N, \rho)$. Así que

$$F(r^*/N, \rho) = Pr(r \leqslant r^*)$$

es muy importante observar que la densidad (3-4) toma el mismo valor en r, ρ que en $-r, -\rho$

Consideremos ahora el problema de probar la hipótesis $Ho: \rho = \rho o$ a partir de la información proporcionada por una muestra tamaño N de una distribución normal bivariante.

Si la hipótesis alternativas es $H_1 : \rho > \rho_o$, rechazaremos H_o caundo $r > r_o$ donde r_o es elegido de modo que

$$1 - F(r_o/N, \rho_o) = \alpha$$

CONCLUSION

donde α es el nivel de significación requerido.

Si la alternativa es H_1 ; $\rho < \rho_o$ rechazaremos H_o cuando $r < r_o'$ donde $F(r_o'/N , \rho_o) =$

Finalmente si la alternativa es $H_1 : \rho \neq \rho_o$, la región crítica, será del tipo $\{ r > r_o, r < r_o' \}$ donde r_o y r_o' deberan elegirse de modo que

$$\{ 1-F(r_o/N, \rho_o) \} + F(r_o'/N, \rho_o) = \alpha$$

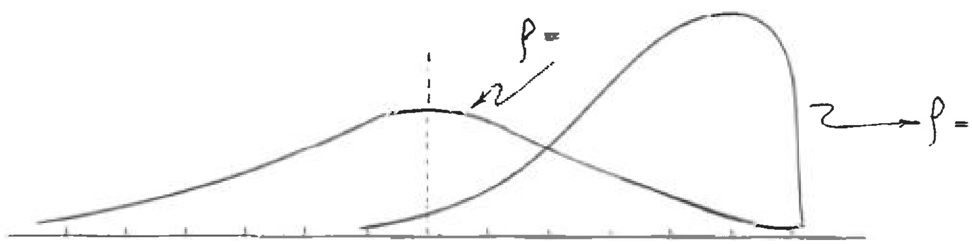
Distribución Asintótica del Coeficiente de Correlación.

Los N pares de valores (x,y) de dos variables pueden ser concebidos como una muestra de una población de todos los posibles pares.

Se puede pensar en un coeficiente de correlación poblacional teórico denotado por ρ , que se estima por el coeficiente de correlación r de la muestra. Los ensayos de significación o hipótesis concernientes a distintos valores de ρ requiere el conocimiento de la distribución muestral de r. Para ρ =0 esta distribución es simétrica y puede utilizarse un estadistico con una distribución de Student.

Para $\rho \neq 0$ la distribución es sesgada. En tal caso una transformación debido a Fisher origina un estadistico que se distribuye aproximadamente normal.

La siguiente gráfica demuestra la distribución de coeficiente de correlación muestrales en muestra de 8 pares sacados de dos poblaciones bivariantes normalmente distribuidas y con los valores indicados para ρ .



Se ha comprobado que la distribución del coeficiente de correlación muestral r, tiende a la normal cuando el tamaño de la muestra crece. Estudiaremos algunos teoremas que me permiten tomar esta decisión.

Teorema 1:

Dados los vectores $y_1, y_2, \ldots, y_m$ de m-componentes independientes e identicamente distribuidos con media $E(y_\alpha) = \nu$

y matriz de covarianza $E(y_\alpha - \nu)(y_\alpha - \nu)' = \Sigma$. Entonces el límite de la distribución

$$\left[\frac{1}{\sqrt{N}}\right] \sum_{\alpha=1}^{N} (y_\alpha - \nu) \text{ cuando } n \longrightarrow \infty \text{ es } N(0, \Sigma)$$

Donde $\frac{1}{\sqrt{n}} \sum_{\alpha=1}^{N} (y_\alpha - \nu) \xrightarrow{\text{Ley}} N_m(0,\Sigma)$

Teorema2:

Dado $A(n) = \sum_{\alpha=1}^{N} (x_\alpha - \bar{x}_N)(x_\alpha - \bar{x}_N)'$ donde

$x_1, x_2, \ldots,$ son distribuidas independientemente de acuerdo a $N(\mu, \Sigma)$ y $n = N-1$. Entonces la distribución asintótica de $B_{(N)} = \left[\frac{1}{\sqrt{n}}\right] \left[A_{(n)} - n\Sigma \right]$ es normal de media $\bar{0}$ y la matriz de Varianza-Covarianza $E\left[b_{ij}(n) b_{kl}(n) \right] =$

$$= n\sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk}$$

Demostración:

Como mostraremos primero A(n) es distribuida como

$$A_{(n)} = \sum_{\alpha=1}^{n} z_\alpha \quad z'_\alpha$$ donde $z_1, z_2, \ldots,$ son distribuidos

independientemente de acuerdo a $N(0, \Sigma)$

Los elementos de $z_\alpha \; z'_\alpha$ se pueden arreglar en el vector

$$Y_\alpha = \begin{bmatrix} z_{1\alpha}^2 \\ z_{1\alpha} \; z_{2\alpha} \\ \cdot \\ \cdot \\ \cdot \\ z_{i\alpha}^2 \\ \cdot \\ \cdot \\ \cdot \\ z_{p\alpha}^2 \end{bmatrix} \qquad (54)$$

Los momentos de Y_α pueden ser deducidos de los momentos de z_α tenemos

$$E(z_{i\alpha} \; z_{j\alpha}) = \sigma_{ij}$$

$$E(z_{i\alpha} \; z_{j\alpha} \; z_{k\alpha} \; z_{l\alpha}) = \sigma_{ij} \; \sigma_{kl} + \sigma_{ik} \; \sigma_{jl} + \\ + \sigma_{il} \; \sigma_{jk}$$

Así los vectores Y_α definidos por (3-5) satisfacen las condiciones del teorema 4 con los elementos de γ , donde los elementos de Σ son arreglos en un vector de forma similar a (54). Si los elementos de $A_{(n)}$ son arreglos de un vector n de forma similar a (54) o sea el vector W(n) enton-

$$W(n) - n\gamma = \sum_{\alpha=1}^{n} (y_{\alpha} - \gamma) \text{ por teorema (1)}$$

$$\left[\frac{1}{\sqrt{n}}\right]\left[W_{(n)} - n\gamma\right] \text{ es asintoticamente distribuida como}$$

una normal con media 0 y matriz de covarianza de Y_{α}(L.Q.Q.D).

Como el coeficiente de correlación

$$r_{(n)} = \frac{A_{ij}(n)}{\sqrt{A_{ii}(n)A_{jj}(n)}}$$

para algún i y j ($i \neq j$). ESto también puede escribirse.

$$r_{(n)} = \frac{c_{ij}(n)}{\sqrt{c_{ii}(n)c_{jj}(n)}} \text{ y sabemos } z^*_{i\alpha} = \frac{z_{i\alpha}}{\sqrt{\sigma_{ii}}}$$

Entonces el conjunto $c_{ii}(n)$, $c_{jj}(n)$ y $c_{ij}(n)$ es distribuido semejante a

$$\sum_{\alpha=1}^{n} \begin{bmatrix} z_{i\alpha}{}^* \\ z_{j\alpha}{}^* \end{bmatrix} (z^*_{i\alpha} \; z^*_{j\alpha}) =$$

$$= \sum_{\alpha=1}^{N} \begin{bmatrix} z_{i\alpha} / \sqrt{\sigma_{ii}} \\ z_{j\alpha} / \sqrt{\sigma_{jj}} \end{bmatrix} \begin{bmatrix} \frac{z_{i\alpha}}{\sqrt{\sigma_{ii}}} , & \frac{z_{j\alpha}}{\sqrt{\sigma_{jj}}} \end{bmatrix}$$

donde $(z^*_{i\alpha}, z^*_{j\alpha})$ son independientes cada cual con distribución N $\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right]$ y si $\sigma_{ij} = \rho \sqrt{\sigma_{ii}\ \sigma_{jj}}$

entonces $\rho = \dfrac{\sigma_{ij}}{\sqrt{\sigma_{ii}\ \sigma_{jj}}}$

Dado

$$U_{(n)} = \frac{1}{n} \begin{bmatrix} c_{ii}(n) \\ c_{jj}(n) \\ c_{ij}(n) \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ 1 \\ \rho \end{bmatrix}$$

$$E\left[U_{(n)}\right] = \frac{1}{n} \begin{bmatrix} E\left[c_{ii}(n)\right] \\ E\left[c_{jj}(n)\right] \\ E\left[c_{ij}(n)\right] \end{bmatrix}$$

Como $E(c_{ij}) = \dfrac{1}{\sqrt{\sigma_i\ \sigma_j}} E(A_{ij}) = \dfrac{(N-1)\ \rho \sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}{\sqrt{\sigma_{ii}\ \sigma_{jj}}} = n\rho$

$$E(c_{ii}) = \frac{1}{\sigma_{ii}} E(A_{ii}) = (N-1)\frac{\sigma_{ii}}{\sigma_{ii}} = n$$

Entonces $E\,(U_{(n)}) = \frac{1}{n} \begin{bmatrix} n \\ n \\ n\rho \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \rho \end{bmatrix} = b$

Entonces

$$\sqrt{n}(\;U_{(n)} - b) \xrightarrow{\text{Ley}} N(0,\; \phi)$$

Donde

ϕ es la matriz de covarianza

$$\phi = \begin{bmatrix} 2 & 2\rho^2 & 2\rho \\ 2\rho^2 & 2 & 2\rho \\ 2\rho & 2\rho & 1+\rho^2 \end{bmatrix}$$

<u>Teorema 3</u>:

Sea $U(n)$ un vector aleatorio m-dimensional y sea b vector fijo, $\sqrt{n}(\;U_{(n)}-b) \xrightarrow{\text{ley}} N(0,\;\Sigma)$

Sea f(u), función con 1º y 2º derivada en μ=b.

Sea $\left.\frac{\partial f(\;u\;)}{\partial u_1},\ldots,\frac{\partial f(\;u\;)}{\partial_{u_m}}\right|_{a=b} = \phi_b$, a si $\left.\frac{\partial f(\;u\;)}{\partial\;u_i}\right|_{a=b}$

es la i-ésima componente de ϕ_b, Entonces la distribución límite de

$$\sqrt{n}\left[f(u_{(n)})-f(b)\right] \quad \text{es } N(0,\ \phi_b' \Sigma \phi_b)$$

por teoría

$$\sqrt{n}(f(u_{(n)})-f(b)) \xrightarrow{\text{Ley}} N(0,\ \phi_b' \Sigma \phi_b)$$

Como

$$U = (u_1,u_2,u_3) = f(u) = \frac{u_3}{\sqrt{u_1 u_2}} = u_3 u_1^{-1/2} u_2^{-1/2} = r$$

Los elementos de ϕ_b son

$$\left.\frac{\partial r}{\partial u_1}\right|_{u=b} = -\frac{1}{2} u_3 u_1^{-3/2} u_2^{-1/2}\Big|_{u=b} = -\frac{1}{2}\rho$$

$$\left.\frac{\partial r}{\partial u_2}\right|_{u=b} = -\frac{1}{2}\rho \quad , \quad \left.\frac{\partial r}{\partial u_3}\right|_{u=b} = 1$$

y $f(b) = \rho$, Luego la varianza asintótica de $\sqrt{n}(r(n-\rho)$ es

$$\left(-\frac{1}{2}\rho\ ,\ -\frac{1}{2}\rho\ ,\ 1\right)\begin{bmatrix} 2 & 2\rho^2 & 2\rho \\ 2\rho^2 & 2 & 2\rho \\ 2\rho & 2\rho & 1+\rho^2 \end{bmatrix}\begin{bmatrix} -\frac{1}{2}\rho \\ -\frac{1}{2}\rho \\ 1 \end{bmatrix}$$

$$= (\rho-\rho^3, \rho-\rho^3, 1-\rho^2)\begin{bmatrix} -\frac{1}{2}\rho \\ -\frac{1}{2}\rho \\ 1 \end{bmatrix}$$

$$= 1 - 2\rho^2 + \rho^4$$

$$= (1-\rho^2)^2$$

Luego

$$\sqrt{n}(r(n)-\rho) \xrightarrow{\text{Ley}} N(0,(1-\rho^2)^2)$$

<u>Teorema 4</u>:

Si $r(n)$ es el coeficiente de correlación muestral de una muestra de $N(=n+1)$ de una distribución normal con correlación ρ entonces

$\sqrt{n}(r(n) - \rho)/(1-\rho^2) \xrightarrow{Ley} N(0,1)$, entonces se demuestra que tiene distribución asintotica N(0,1)

Del Teorema (3), se puede deducir que si f(x) es una función con primera y segunda derivada en $x=x_o$, entonces

$$\sqrt{n}\left[f(r) - f(\rho)\right]$$

es asintoticamente distribuida como normal con media cero y varianza

$$\left[\frac{df}{dx}\Big|_{x=\rho}\right]^2 (1-\rho^2)^2$$

Si f es una función real con derivada 1º y 2º en $x=\rho$ tal que $\sqrt{n}(f(r) - f(\rho)) \longrightarrow N(0,f'(u)^2(1- \ ^2)^2)$.

. Si esta en $\mathbb{R}^n$ la derivada es una gradiente y si esta en $\mathbb{R}$, la derivada $\frac{\partial r}{\partial u} = \frac{\partial r}{\partial u_1}\Big|_{u=b}$

Luego $f'(\rho) = \frac{1}{1-\rho^2} = \frac{1}{2}\left[\frac{1}{1+\rho} + \frac{1}{1-\rho}\right]$

$$f(u) = \frac{1}{2}\log\frac{1-u}{1+u}$$

entonces

$$\sqrt{n}\left(\frac{1}{2}\log\frac{1+r}{1-r} - \frac{1}{2}\log\frac{1+\rho}{1-\rho}\right) \xrightarrow{Ley} N(0,1)$$

Luego

(55) $z = \frac{1}{2} \log \frac{1+r}{1-r}$, determina como la z de Fisher

Así

$$\xi = \frac{1}{2} \log \frac{1+\rho}{1-\rho} \quad \text{entonces} \quad \sqrt{N-1}(z-\xi) \xrightarrow{\text{Ley}} N(0,1)$$

<u>Teorema 4</u>:

Dado z definido por $z = \frac{1}{2} \log \frac{1+r}{1-r}$, donde r es el coeficiente de correlación de una muestra de N(= n+1) de una distribución normal bivariante con correlación ρ , sea ξ definida por $\xi = \frac{1}{2} \log \frac{1+\rho}{1-\rho}$. Entonces $\sqrt{n}(z-\xi)$ es distribuido asintoticamente como normal con media 0 y varianza 1.

Es importante notar que la estadística z tiene la propiedad de converger más rápidamente a la normal que r. Además interesa saber que una mejor aproximación fue sugerida por Fisher, donde, la media es

$$\mu_1 = E(z) = \frac{1}{2} \ n \left[\frac{1+\rho}{1-\rho}\right] + \frac{\rho}{2(n-1)} \left\{ 1 + \frac{5+\rho^2}{4(n-1)} + \dots \right.$$

y la varianza

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

$$\sigma_1^2 = E(z-\xi)^2 = \frac{1}{n-1}\left\{1 + \frac{4-\rho^2}{2(n-1)} + \frac{22-6\rho^2-3\rho^4}{6(n-1)^2} + \dots\right\}$$

$$\cong \frac{1}{n-1}\left[1 + \frac{2}{n-1} + \frac{4}{(n-1)^2} + \frac{8}{(n-1)^3} + \dots\right] \cong \frac{1}{n-3}$$

cuando $n > 3$

Veremos a continuación algunas aplicaciones de los teoremas enunciados

a) Supongamos que deseamos probar la hipótesis $\rho = \rho_o$ contra la alternativa $\rho \neq \rho_o$. Para ello debemos primero computar r y luego z según (3-6) Entonces si

$$\xi_o = \frac{1}{2} \log \frac{1+\rho_o}{1-\rho_o}$$

tenemos que una región crítica de nivel 5% es

$$\sqrt{N-3}\,|z-\xi_o| > 1.96$$

Una región mejor es

$$\sqrt{N-3}\,| z - \xi_o - \tfrac{1}{2}\,\rho_o / n \,| > 1.96$$

b) Supongamos que tenemos una muestra de tamaño N_1 de una población y otra tamaño N_2 de otra. El problema que nos interesa es cómo probar la hipótesis $\rho_1 = \rho_2$ de que los coeficientes de correlación de ambas poblaciones coinciden.

Para ello debemos computar r_1 y r_2, y luego a partir de (55) computar z_1 y z_2. Entonces, como a partir del último teorema se deduce que la distribución asintóticaa de $z_1 - z_2$ si $\rho_1 - \rho_2$ es normal con media cero y varianza $\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}$ tenemos que una región crítica de nivel 5% es

$$\frac{|z_1 - z_2|}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}} > 1.96$$

c) Bajo las condiciones de (b) supongamos que $\rho_1 = \rho_2 = \rho$.

Cómo usar ambas muestras para dar un estimador conjunto de ρ. Dado que Z_1 y Z_2 tienen varianzas respectivas $\frac{1}{N_1-3}$ y $\frac{1}{N_2-3}$, puedo estimar Insesgadamente a ξ con mínima varianza por

$$\frac{(N_1-3)Z_1 + (N_2-3)Z_2}{N_1 + N_2 - 6}$$

y luego convertir esto en un estimador de ρ por la fórmula inversa de la (55).

d) Sea r la correlación muestral de las observaciones. Como obtengo un intervalo de confianza para ρ. Sabemos que aproximadamente

$$\Pr\left\{-1.96 < \sqrt{N-3}(Z-\xi) < 1.96\right\} = 0.95.$$

De aquí se deduce el intervalo

$$\left[-1.96/\sqrt{N-3} + Z,\ 1.96/\sqrt{N-3} + Z\right]$$

para el parámetro ξ, Entonces, usando el hecho de que

$$\rho = \tan h = \frac{e^{\xi} - e^{-\xi}}{e^{\xi} + e^{-\xi}}$$

es una transformación monotona, se deduce el siguiente intervalo para ρ.

$$\tan h(z-1.96/\sqrt{N-3}) < \rho < \tan h(z+1.96/\sqrt{N-3})$$

Docimas e Intervalos para Coeficientes de Correlación Parcial.

Enunciamos a continuación un teorema que afirma que la distribución de un coeficiente de correlación parcial muestral $r_{ij,a+1,\ldots,p}$, basada en una muestra tamaño N de una distribución con correlación poblacional.

$\rho_{ij,q+1,\ldots,p}$ igual a cierto valor ρ , es la misma que la distribucióon de una correlación simple o total, r, basada en una muestra tamaño N-(p-q) de una población con correlación ρ .

Por lo tanto, todos los métodos de inferencia desarrollados para el coeficiente de correlación poblacional pueden ser usados para correlación parciales cambiando N por N-(p-q)

Teorema

Si la función de distribución de r_{ij}, basado en una muestra tamaño N de una distribución normal con correlación ρ_{ij} en $F(r/N_1 \rho_{ij})$, entonces la función de distribución de $r_{ij,q+1,\ldots,p}$ basado en una muestra tamaño N, es

$F(r/N-(p-q), \quad \rho_{ij,a+1,\ldots,p})$.

Veamos un ejemplo de aplicación

Supongamos que en base a una muestra tamaño N, deseamos usar la estadistica Z de Fisher para probar la hipótesis $\rho_{ij,q+1,\ldots,p} = \rho_o$ contra la alternativa a dos ramas o colas. Para ello calculamos

$$Z = \frac{1}{2} \log \frac{1+r_{ij,a+1,\ldots,p}}{1-r_{ij,a+1,\ldots,p}}$$

y

$$\xi_o = \frac{1}{2} \log \frac{1+\rho_o}{1-\rho_o}$$

Entonces $\sqrt{N-(p-q)-3}\,(z-\xi_o)$ debería ser comparado con los puntos de significación de la distribución normal estándard.

Distribución del Coeficiente de Correlación Múltiple Muestral cuando el Poblacional es cero.

Habiamos demostrado que

$$R = \sqrt{\frac{\hat{\sigma}_{(1)}\, \hat{\Sigma}_{22}^{-1}\, \hat{\sigma}'_{(1)}}{\hat{\sigma}_{11}}}$$

De aquí surge que

$$\frac{R^2}{1-R^2} = \frac{\hat{\sigma}_{(1)} \hat{\Sigma}_{22}^{-1} \hat{\sigma}_{(1)}}{\hat{\sigma}_{11} - \hat{\sigma}_{(1)} \hat{\Sigma}_{22}^{-1} \hat{\sigma}_{(1)}} = \frac{\hat{\sigma}_{(1)} \hat{\Sigma}_{22}^{-1} \hat{\sigma}_{(1)}'}{\hat{\sigma}_{11.2}}$$

Entonces, si $X_{pxl}^{pxl} = (x_{(p-q)xl}^{(1)}, X_{q+i}^{(2)})$ es el vector variable aleatoria del cual se tomó una muestra tamaño N, en base a la cual se calcularon las estimadores, tenemos que

$$\frac{(N-1)\,\hat{\sigma}_{11.2}}{\sigma_{11.1}} \quad \text{y} \quad \frac{(N-1)\,\hat{\sigma}_{(1)} \hat{\Sigma}_{22}^{-1} \hat{\sigma}_{(1)}}{\sigma_{11.2}}$$

están distribuidas independientemente como χ^2 con, N-q-1 y q grados de libertad respectivamente.

Entonces

$$\frac{R^2}{1-R^2} \cdot \frac{N-q-1}{q} = f_{q,N-q-1}$$

tiene distribución f con q y (N-q-1) grados de libertad.

En particular es válido el siguiente:

Teorema

Si R es el coeficiente de correlación múltiple entre

X_1 y $X^{(2)'} = (x_2,\ldots,x_p)$ basado en una muestra tamaño N de una distribución $N(\mu, \Sigma)$ si $\bar{R} = 0$, (es decir si $(\sigma_{12},\ldots, \sigma_{1p}) = \sigma_{(1)} = 0)$, entonces

$$\frac{R^2}{1-r^2} \cdot \frac{N-p}{p-1} \sim F_{p-1,N-p}$$

Puede demostrarse que esta variable es la que se usa en teoría de regresión para probar la hipótesis de que la regresión de X_1 en $= (X_2,\ldots,X_p)$ es cero, o sea que planteado el modelo

$$X_{1i} = \mu + X_i^{(2)'} \beta + e_i, \quad i = 1,\ldots,N$$

la hipótesis que se prueba es $H_o: \beta = 0$. Conviene, a los efectos de demostrar esta propiedad repametrizar el modelo restándole el vector $X_i^{(2)}$ el vector de promedio $\bar{X}^{(2)}$.

Además, puede demostrarse que la prueba de cociente de verosimitudes para la hipótesis

$$H_o: \bar{R}_{1\ q+1,\ldots,p} = 0$$

contra la alternativa $\bar{R} > 0$, en el caso normal tiene como región crítica

$$\left\{ \frac{R^2}{1-R^2} \cdot \frac{N-p}{p-1} > K \right\}$$

Esta prueba por lo tanto rechazara $H_o: \overline{R} = 0$ en los mismos casos en que se rechaza $H_o: \beta = 0$ en el modelo anteriormente mencionado.

Es importante notar el signigicado de la Hipótesis $H_o : \overline{R}_{1,q+1,\ldots,p} = 0$.

Para ello conviene recordar que el coeficiente de correlación múltiple es la máxima correlación existente entre X_1 y una combinación lineal $\beta' X^{(2)}$, en particular; interesa en este caso notar que dicho máximo coincide con el valor absoluto de dichas correlaciones. Entonces decir que $\overline{R} = 0$, implica que la correlación entre X_1 y cualquier combinación lineal $\beta' X^{(2)}$ es nula y esto en el caso normal implica que X_1 es independiente de $X^{(2)}$.

Las conclusiones de este trabajo la hemos dividido en dos aspectos.

A.- Sobre la utilidad del coeficiente de correlación en la toma de decisiones.

1.- El coeficiente de correlación de Fisher nos permite tomar una decisión más confiable para una muestra más pequeña que utilizando el coeficiente de correlación de Pearson.

2.- A través del coeficiente de correlación de Fisher se logra una mayor confiabilidad en los datos calculados y esto es de gran importancia para el mejor desarrollo nacional toda vez que existe una gran cantidad de investigaciones en donde es aplicable dichos coeficientes de correlación.

3.- El coeficiente de Fisher, nos permite utilizar una prueba más potente en la toma de decisiones.

4.- El coeficiente de correlación de Fisher tiene una gran utilidad en estudios exploratorios, nos permite con una muestra más pequeña tomar decisiones confiables y esto minimizar el costo de la investigación.

B.- Sobre el desarrollo teórico analizado en el Coeficiente de Correlación de Fisher.

1.- Dado dos variables aleatorias una de ellas X y la otra Y que depende X, existe una matriz de covarianza Σ semi-definida positiva que nos permite obtener los coeficientes de Correlación Simple o total, Correlación Parcial y Correlación Múltiple.

2.- El coeficiente de correlación es invariante bajo cambios de escala.

3.- Los estimadores del coeficiente de correlación se pueden obtener a partir del estimador máxima verosimilitud de la matriz de covarianza.

4.- La función de densidad cuando $\rho \neq 0$, es un recurso matemático básico utilizado por Fisher (1915) para obtener la esperanza y la varianza de r.

5.- La distribución asintótica del coeficiente de correlación debido a Fisher origina un estadistico que se distribuye aproximadamente normal cuando $\rho \neq 0$.

1.- Que en las investigaciones estadísticas que tienen un componente sobre correlación, se deberían emplear estadísticos que brinden una información más confiable para la toma de decisiones. En este sentido, el coeficiente de correlación de Fisher brinda una serie de elementos que lo hacen más confiable (disminuye el sesgo, pruebas más potentes etc.)

2.- Que estos trabajos no reposen sólo en las bibliotecas y centro de documentación de la Universidad, sino que se haga un esfuerzo para brindar este tipo de conferencias a los diferentes centros de investigaciones del país, así como a oficinas que manejan y procesan informaciones estadísticas.

1 Anderson T.W., Introduction to Multivariante Statistical Analysis; John Wiley, 1958.

2 Robert G.D. Steel y James H. Torrie. Principles and Procedures of Statistics. Editora McGraw-Hill, 1960.

3 Morrison, D.F. Multivariante Statistical Methodos, McGraw-Hill, 1967.

4 Gheorghe Mihoc y Virgil Crain. Tratat de Statistica Matematica; Selectic si Estimatie. Editura Academiei Republicii Socialiste Romania. Bucarest, 1976.

5 Alexander M.Mood y Franklin A. Garybill; Introducción a la Teoría de la Estadística; Aguilar, 1969.

6 Shayle R. Seaele. Matrix Algebra Useful for Statistics, Editora John Wiley & Sons.

7 Robb J. Muirhead. Aspects of Multivariate Statistical Theory. Editora John Wiley & Son Inc.

8 Samuel S. Wilks. Mathematical Statistics. Editora, John Wiley & Sons, Inc. 1972.

9 Henry Scheffe. The Analysis of Variance. John Wiley & Sons, Inc. 1959.