TEMA 3

REGRESIÓN Y CORRELACIÓN

Regresión mínimo-cuadrática bidimensional

Planteamiento del problema

Dadas dos variables aleatorias X e Y definidas sobre un mismo espacio de probabilidad (asociadas a un mismo experimento aleatorio), el problema de la regresión consiste en determinar relaciones entre X e Y que permitan predecir o aproximar el valor de una de ellas a partir del valor tomado por la otra.
Concretamente, si se quiere predecir los valores de Y a partir de los de X, se considera una función $d(X)$ de forma que cuando X tome el valor x, la predicción para el valor de Y será $y = d(x)$.
El problema consiste entonces en elegir la función d, que se denomina función de regresión, de forma que las predicciones sean óptimas en algún sentido.
El criterio de optimalidad usado más frecuentemente está basado en el PRINCIPIO DE MÍNIMOS CUADRADOS, que consiste en elegir d de forma que la media de las desviaciones cuadráticas entre los valores reales y los aproximados sea mínima.
Se trata, por tanto, de encontrar la función $d(X)$ que minimice

$$\mathrm{E}[Y - d(X)]^2 \equiv \text{ ERROR CUADRÁTICO MEDIO asociado a la función de regresión } d$$

El problema de regresión abordado según este criterio se denomina REGRESIÓN MÍNIMO CUADRÁTICA y, evidentemente, requiere que $\exists \mathrm{E}Y^2$ y $\exists \mathrm{E}(d(X))^2$.

Solución del problema

Vamos a obtener tal función $d(X)$: Por las propiedades de la esperanza condicionada se tiene

$$\mathrm{E}[Y - d(X)]^2 = \mathrm{E}\left[\mathrm{E}\left[(Y - d(X))^2/X\right]\right]$$

Por tanto, la función $d(X)$ que minimiza el primer miembro es la que minimiza

$$\mathrm{E}\left[(Y - d(X))^2/X\right]$$

y, recordando una propiedad vista para variables aleatorias ($\mathrm{E}[X-c]^2$ se hace mínima para $c = \mathrm{E}X$), claramente en este caso, $d(X) = \mathrm{E}[Y/X]$.
El error cuadrático medio de asociado a ella es

$$\mathrm{ECM} = \mathrm{E}\left[Y - \mathrm{E}[Y/X]\right]^2 = \mathrm{E}\left[\mathrm{E}\left[(Y - \mathrm{E}[Y/X])^2/X\right]\right] = \mathrm{E}[\mathrm{Var}[Y/X]]$$

Nota

Observemos que si queremos predecir el valor de Y sin ninguna información sobre el valor de X (sin observar X), según el Principio de Mínimos Cuadrados será el valor c que haga mínimo $\mathrm{E}[Y-c]^2$ que, evidentemente, es $c = \mathrm{E}Y$ y el ECM cometido en la predicción es $\mathrm{E}[Y - \mathrm{E}Y]^2 = \mathrm{Var}Y$.

Teniendo en cuenta la descomposición de la varianza

$$\text{Var}Y = \text{E}[\text{Var}[Y/X]] + \text{Var}[\text{E}[Y/X]]$$

resulta que $\text{Var}Y \geq \text{E}[\text{Var}[Y/X]]$ y, además,

$$\text{Var}Y = \text{E}[\text{Var}[Y/X]] \Leftrightarrow \text{Var}[\text{E}[Y/X]] = 0 \Leftrightarrow \text{E}[Y/X] = c$$

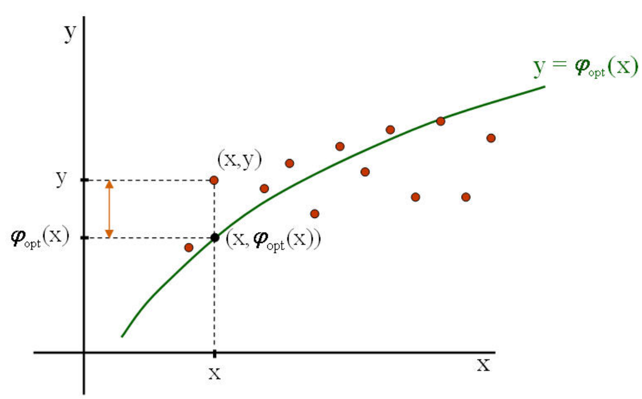
y, evidentemente, $c = \text{E}Y$ $(c = \text{E}[\text{E}[Y/X]]$.
Por tanto, salvo que $\text{E}[Y/X]$ sea constante (lo que ocurre, por ejemplo, si X e Y son independientes) el ECM de la predicción al observar X es menor que el ECM de la predicción de Y sin observar X.
Esto indica que siempre que se tenga la posibilidad de observar una variable aleatoria relacionada con Y antes de dar una predicción, debe hacerse. Sin embargo, en la práctica, las observaciones previas pueden conllevar algún coste (económico, temporal, etc.) que habrá que sopesar con la ganancia en la predicción.

Curvas de regresión mínimo-cuadráticas
Notemos que al considerar una función de regresión d para aproximar los valores de Y a partir de los de X, estamos considerando una curva en el plano

$$y = d(x)$$

de forma que cada punto de la distribución (x, y) se transforma en $(x, d(x))$. Esta curva se ajustará más o menos bien a la distribución según que las diferencias $y - d(x)$ sean menos o más grandes. Una medida global de la magnitud de estas diferencias es la media de los cuadrados.
Así, tomando $d(X) = \text{E}[Y/X]$, la curva correspondiente, $y = \text{E}[Y/X = x]$ es la que mejor se ajusta a la distribución en el sentido de que hace mínima la media de las desviaciones verticales al cuadrado. Se denomina **curva de regresión mínimo cuadrática de Y sobre X**.

Curva de regresión mínimo cuadrática de Y sobre X: $y = \text{E}[Y/X = x]$



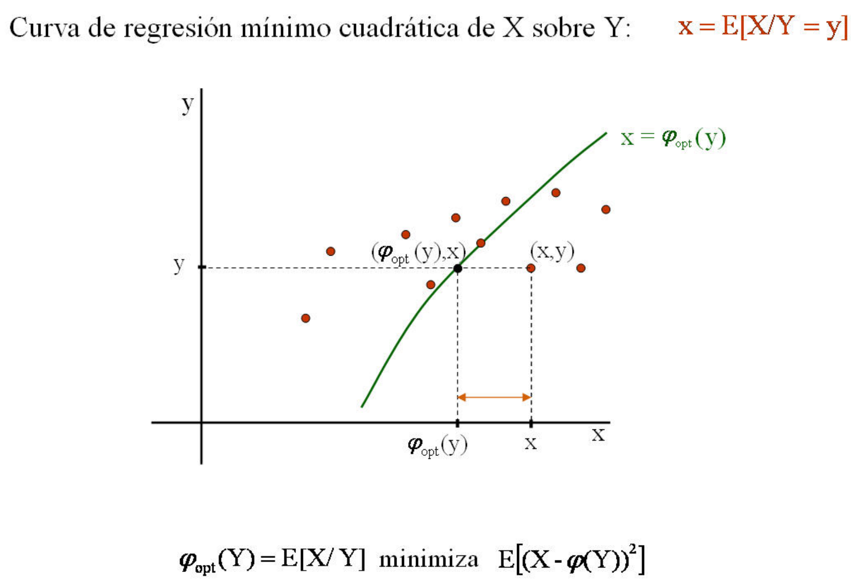
$\varphi_{opt}(X) = \text{E}[Y/X]$ minimiza $\text{E}\left[(Y - \varphi(X))^2\right]$

Si se conoce la curva de regresión, dado cualquier valor concreto de X, $X = x$, el valor aproximado para Y será $\mathrm{E}[Y/X = x]$ y este valor coincide con la predicción mínimo cuadrática de Y dado $X = x$, esto es

$$\mathrm{E}\left[(Y - d(x))^2/X = x\right]$$

es mínimo para $d(x) = \mathrm{E}[Y/X = x]$. Además el ECM de esta predicción concreta es $\mathrm{Var}[Y/X = x]$

Un desarrollo paralelo puede hacerse para aproximar la variable X por Y y, en este caso, ya que lo que se intenta minimizar es $\mathrm{E}[(X - d(Y))^2]$, la curva de regresión de X sobre Y es la que mejor se ajusta a los puntos de la distribución en el sentido de que hace mínima la media de las desviaciones al cuadrado cuando éstas se toman paralelas al eje de abcisas.



Se tiene, por tanto

	SIN OBSERVAR X	A PARTIR DE X	A PARTIR DE UN VALOR CONCRETO $X = x$
PREDICIÓN MÍNIMO CUADRÁTICA DE Y	$\mathrm{E}Y$	$\mathrm{E}[Y/X]$	$\mathrm{E}[Y/X = x]$
ERROR CUADRÁTICO MEDIO DE LA PREDICCIÓN Y	$\mathrm{Var}Y$	$\mathrm{E}[\mathrm{Var}[Y/X]]$	$\mathrm{Var}[Y/X = x]$

Casos particulares.

1. Si las variables aleatorias X **e** Y **son independientes**, las distribuciones condicionadas son idénticas y coinciden con la distribución marginal correspondiente, por lo que las medias condicionadas son iguales a la media marginal, esto es

$$\mathrm{E}[X/Y] = \mathrm{E}X, \quad \mathrm{E}[Y/X] = \mathrm{E}Y$$

por lo tanto las curvas de regresión son rectas paralelas a los ejes (ninguna variable proporciona información sobre la otra y la predicción es constante)

$$y = \mathrm{E}Y \qquad x = \mathrm{E}X$$

que, evidentemente, se cortan en el punto $(\mathrm{E}X, \mathrm{E}Y)$. Además los ECM asociados son $\mathrm{Var}Y$ (para el caso Y/X) y $\mathrm{Var}X$ (para el caso X/Y)

2. Si Y depende funcionalmente de X pero X no depende funcionalmente de Y, la curva de regresión de Y sobre X es $y = f(x)$ (coincide con la curva de dependencia). Evidentemente, el ECM es cero.

3. Si X e Y tienen **dependencia funcional recíproca** $Y = f(X),\ X = f^{-1}(Y)$, la curva de regresión de Y sobre X es

$$y = \mathrm{E}[Y|X = x] = \mathrm{E}[f(X)|X = x] = f(x)$$

y la curva de regresión de X sobre Y es

$$x = \mathrm{E}[X|Y = y] = \mathrm{E}[f^{-1}(Y)|Y = y] = f^{-1}(y)$$

que coinciden y coinciden con la curva de dependencia. Los ECM son evidentemente nulos.

Regresión mínimo-cuadrática lineal

Planteamiento y solución del problema

En la práctica es a veces útil restringirse a un tipo determinado de curvas para hacer la regresión. Esto suele hacerse si se tiene idea de que los puntos de la distribución se ajustan a una determinada forma funcional: funciones lineal, exponenciales, logarítmicas, etc. En tales casos, el problema de regresión puede simplificarse restringiéndose a funciones del tipo deseado. Un caso particularmente importante es la REGRESIÓN LINEAL que es al que dedicamos este apartado.
Intentamos ahora resolver el problema de predecir, por mínimos cuadrados, los valores de una variable (Y) a partir de una función lineal de otra (X). O sea, encontrar una recta $y = ax + b$ en torno a la cual se ajusten lo mejor posible los puntos de la distribución.
Según el Principio de Mínimos Cuadrados, la mejor aproximación de Y mediante una función lineal de X, viene dada por la función $aX + b$ que minimice

$$\mathrm{E}[Y - (aX + b)]^2$$

(para ello hace falta suponer la existencia de $\mathrm{E}X^2$ y $\mathrm{E}Y^2$).

El problema se reduce a obtener "a" y "b" que minimicen $L = \mathrm{E}[Y - (aX + b)]^2$

$$L = \mathrm{E}[Y - (aX + b)]^2 = \mathrm{E}[Y^2 + (aX + b)^2 - 2Y(aX + b)] =$$

$$= \mathrm{E}Y^2 + a^2\mathrm{E}X^2 + b^2 + 2ab\mathrm{E}X - 2a\mathrm{E}XY - 2b\mathrm{E}Y$$

Derivando con respecto a "a" y "b" e igualando a cero, se obtiene

$$\left.\begin{array}{r} a\mathrm{E}X^2 + b\mathrm{E}X = \mathrm{E}XY \\ b + a\mathrm{E}X = \mathrm{E}Y \end{array}\right\} \text{ ECUACIONES NORMALES}$$

cuya solución es

$$a = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X}, \qquad b = \mathrm{E}Y - \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X}\mathrm{E}X$$

NOTA.- Se puede comprobar que la solución obtenida es un mínimo

NOTA.- Observemos que X debe ser tal que $\mathrm{Var}X \neq 0$. Si $\mathrm{Var}X = 0$, X es constante y el problema es predecir Y por una constante, cuya solución es $y = \mathrm{E}Y$

Rectas de regresión mínimo-cuadráticas

A la recta de ecuación

$$y = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X}x + \mathrm{E}Y - \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X}\mathrm{E}X$$

o bien,

$$y - \mathrm{E}Y = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X}(x - \mathrm{E}X)$$

se le llama **recta de regresión de Y sobre X**.

De igual forma, se procede para la predicción de X mediante una función lineal de Y, $cY + d$ y y **la recta de regresión de X sobre Y** es

$$x - \mathrm{E}X = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}Y}(y - \mathrm{E}Y).$$

Suponiendo la existencia de $\mathrm{E}X^2$ y $\mathrm{E}Y^2$ y que las variables son no degeneradas, se definen los **coeficientes de regresión** como

COEFICIENTE DE REGRESIÓN DE Y SOBRE X: $\gamma_{Y/X} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X}$

COEFICIENTE DE REGRESIÓN DE X SOBRE Y: $\gamma_{X/Y} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}Y}$

Notas.-

1.- El Principio de Mínimos Cuadrados se puede utilizar de igual forma para obtener predicciones mediante cualquier otro tipo de funciones (polinomios de cierto grado, funciones exponenciales, etc).

2.- Evidentemente si una curva de regresión es una recta coincide con la recta de regresión correspondiente.

3.- Las rectas de regresión se cortan en el punto $(\mathrm{E}X, \mathrm{E}Y)$.

4.- $\gamma_{Y/X}$ y $\gamma_{X/Y}$ tienen el mismo signo (el de la $\mathrm{Cov}(X,Y)$).

5.- La recta de regresión de Y sobre X tiene menor pendiente, en valor absoluto, que la de X sobre Y. En efecto, la pendiente, en valor absoluto, de la recta de regresión de Y sobre X es

$$\frac{|\mathrm{Cov}(X,Y)|}{\mathrm{Var}X}$$

y la de X sobre Y

$$\frac{\mathrm{Var}Y}{|\mathrm{Cov}(X,Y)|}$$

y, por la desigualdad de Cauchy-Schwarz

$$\frac{|\mathrm{Cov}(X,Y)|}{\mathrm{Var}X} \leq \frac{\mathrm{Var}Y}{|\mathrm{Cov}(X,Y)|}$$

Algunos casos particulares.

1. Si las variables aleatorias X **e** Y **son independientes**, $\mathrm{Cov}(X,Y) = 0$ y, por tanto, $\gamma_{Y/X} = \gamma_{X/Y} = 0$. En este caso las rectas de regresión son rectas paralelas a los ejes

$$y = \mathrm{E}Y \qquad x = \mathrm{E}X$$

que, evidentemente, se cortan en el punto $(\mathrm{E}X, \mathrm{E}Y)$.

2. Y (X) depende linealmente de X (Y), entonces la dependencia lineal es recíproca ($Y = aX + b \;\Rightarrow\; X = -\frac{b}{a} + \frac{Y}{a}$). En este caso, la recta de regresión de Y sobre X es

$$y = ax + b$$

y la de X sobre Y

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

Las rectas de regresión coinciden y coinciden con la recta de dependencia.

Error cuadrático medio asociado a la recta de regresión

El **error cuadrático medio** asociado a la predicción (centrándonos en la recta de regresión de Y sobre X) es

$$\mathrm{Var}Y - \frac{(\mathrm{Cov}(X,Y))^2}{\mathrm{Var}X}$$

Dem

$$\mathrm{ECM} = \mathrm{E}\left[Y - \left(\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X}X + \mathrm{E}Y - \mathrm{E}X\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X}\right)\right]^2 =$$

$$= \mathrm{E}\left[(Y - \mathrm{E}Y) - \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X}(X - \mathrm{E}X)\right]^2 =$$

$$\mathrm{E}[Y-\mathrm{E}Y]^2+\frac{(\mathrm{Cov}(X,Y))^2}{(\mathrm{Var}X)^2}\mathrm{E}(X-\mathrm{E}X)^2-2\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X}\mathrm{E}[(X-\mathrm{E}X)(Y-\mathrm{E}Y)]=$$

$$\mathrm{Var}Y+\frac{(\mathrm{Cov}(X,Y))^2}{\mathrm{Var}X}-2\frac{(\mathrm{Cov}(X,Y))^2}{\mathrm{Var}X}=\mathrm{Var}Y-\frac{(\mathrm{Cov}(X,Y))^2}{\mathrm{Var}X}$$

Correlación

La correlación estudia la bondad del ajuste de la función de regresión encontrada mediante el método de mínimos cuadrados, es decir, evalúa en qué medida la función de regresión explica a una variable a partir de la otra.

Razón de correlación

Centrándonos en la curva de regresión de Y sobre X ($y=\mathrm{E}[Y|X=x]$) se trata de definir un coeficiente para medir el grado en que la variable Y depende de X (puede determinarse a partir de X). Dicha dependencia se medirá en función del grado de concentración de la distribución en torno a dicha curva.
De hecho, si Y depende funcionalmente de X (Y está totalmente determinada por X y el grado de dependencia es máximo), $Y=\mathrm{E}[Y/X]$ y la curva de regresión proporciona exactamente los puntos de la distribución. En general

$$Y=Y-\mathrm{E}[Y/X]+\mathrm{E}[Y/X]=\text{RESIDUO }+\mathrm{E}[Y/X]$$

y la variable Y depende más de X cuanto menor sea el residuo. Ya que $Y-\mathrm{E}[Y/X]$ es una variable aleatoria centrada en cero, para obtener una medida global de sus valores puede usarse su varianza

$$\mathrm{Var}[Y-\mathrm{E}[Y/X]]=\mathrm{E}[(Y-\mathrm{E}[Y/X])^2]=\text{ECM de la curva de regresión}=\mathrm{E}[\mathrm{Var}[Y/X]]$$

que se denomina VARIANZA RESIDUAL.

Así, la variable Y depende más de X cuanto menor sea el ECM asociado a la curva de regresión o, equivalentemente, teniendo en cuenta la descomposición de la varianza

$$\mathrm{Var}Y=\mathrm{Var}[\mathrm{E}[Y/X]]+\mathrm{E}[\mathrm{Var}[Y/X]]$$

cuanto mayor sea la varianza de la función de regresión, $\mathrm{Var}[\mathrm{E}[Y/X]]$.
Esta ecuación suele especificarse diciendo que la varianza de Y queda explicada por la varianza de la función de regresión y la varianza residual, y la variable Y queda mejor explicada por X cuanto mayor sea la varianza de la función de regresión.

Sin embargo, tomar $\mathrm{Var}[\mathrm{E}[Y/X]]$ como medida del grado de dependencia tiene algunos inconvenientes:

- No es una medida adimensional, por lo que no es válida para efectuar comparaciones (la unidad de medida es la de Y al cuadrado).
- No es invariante frente a cambios de escala en la unidad de medida por lo que puede conducir a conclusiones engañosas. Por ejemplo, si la variable Y mide tiempo en horas y la variable Y^* tiempo en minutos, $Y^* = 60Y$ y dado que Y e Y^* son la misma variable (medidas en distintas unidades) cualquier otra variable X explica igual de bien a Y que a Y^*. Sin embargo,

$$\mathrm{Var}[\mathrm{E}[Y^*/X]] = (60)^2 \mathrm{Var}[\mathrm{E}[Y/X]]$$

y se deduciría que Y^* depende de X mucho más que Y.

Para salvar estos inconvenientes, se define el siguiente coeficiente normalizado
Definición.- Se define la **razón de correlación** de Y sobre X y se notará $\eta^2_{Y/X}$ como

$$\eta^2_{Y/X} = \frac{\mathrm{Var}[\mathrm{E}[Y/X]]}{\mathrm{Var}Y} = 1 - \frac{\mathrm{E}[\mathrm{Var}[Y/X]]}{\mathrm{Var}Y}$$

- Proporción de varianza de la variable ($\mathrm{Var}Y$) que corresponde a la función de regresión ($\mathrm{E}[Y/X]$).
- Proporción de la varianza de la variable que queda explicada por la función de regresión

Es una medida adimensional que sirve para efectuar comparaciones y es invariante frente a cambios de escala $\eta^2_{aY/X} = \eta^2_{Y/X}$.

De forma análoga se define la **razón de correlación** de X sobre Y y se notará $\eta^2_{X/Y}$ como

$$\eta^2_{X/Y} = \frac{\mathrm{Var}[\mathrm{E}[X/Y]]}{\mathrm{Var}X} = 1 - \frac{\mathrm{E}[\mathrm{Var}[X/Y]]}{\mathrm{Var}X}$$

En resumen, $\eta^2_{Y/X}$ ($\eta^2_{X/Y}$) proporciona una medida del grado de dependencia de Y respecto de X (X respecto de Y) o, equivalentemente, del grado de concentración de la distribución en torno a la curva de regresión, esto es, una medida de la bondad del ajuste de la distribución a la curva de regresión correspondiente.

Las propiedades que vemos a continuación reflejan que esta es una medida para ello.

Propiedades

1) $0 \leq \eta^2_{Y/X} \leq 1$

2) $\eta^2_{Y/X} = 0 \Longleftrightarrow$ La curva de regresión de Y sobre X coincide con la recta de regresión de Y sobre X y es $y = \mathrm{E}Y$.

3) $\eta^2_{Y/X} = 1 \Longleftrightarrow$ Y depende funcionalmente de X. (Dicha dependencia no tiene por qué ser lineal, ni X tiene que depender funcionalmente de Y)

4) $\eta^2_{Y/X} = \eta^2_{X/Y} = 1 \iff$ Hay dependencia funcional recíproca.

5) ECM $= \mathrm{Var}Y[1 - \eta^2_{Y/X}]$ (para la curva de regresión de Y sobre X).

Coeficiente de determinación lineal

De la expresión del ECM asociado con la recta de regresión (caso de Y/X)

$$\text{ECM (Recta de regresión)} = \mathrm{Var}Y - \frac{(\mathrm{Cov}(X,Y))^2}{\mathrm{Var}X}$$

se tiene la siguiente descomposición de la varianza de Y

$$\mathrm{Var}Y = \frac{(\mathrm{Cov}(X,Y))^2}{\mathrm{Var}X} + \text{ECM (Recta de regresión)}$$

Notemos que el primer sumando es la varianza de la función de regresión (lineal) que estamos considerando

$$\mathrm{Var}[aX+b] = a^2\mathrm{Var}X = \frac{(\mathrm{Cov}(X,Y))^2}{(\mathrm{Var}X)^2}\mathrm{Var}X$$

Por lo tanto, como en el caso de las curvas de regresión, tenemos descompuesta la varianza de Y en la varianza debida a la función de regresión y el ECM.
Ya que, de nuevo, la recta proporciona un mejor ajuste a los puntos de la distribución cuanto menor sea el ECM y, por tanto, cuanto mayor sea el otro sumando, y buscando de nuevo que la medida sea adimensional, se mide ahora el grado de concentración por

$$\frac{(\mathrm{Cov}(X,Y))^2}{\mathrm{Var}X\mathrm{Var}Y}$$

que, observemos es el mismo para la recta de regresión de X sobre Y.

Se define el **Coeficiente de determinación lineal de** X **e** Y como

$$\rho^2_{X,Y} = \frac{(\mathrm{Cov}(X,Y))^2}{\mathrm{Var}X\mathrm{Var}Y}$$

que verifica

$$\rho^2_{X,Y} = \gamma_{Y/X}\gamma_{X/Y}$$

Es, de forma análoga a $\eta^2_{Y/X}$ en la curva, la proporción de la Var$Y$ que corresponde a la función de regresión considerada.

Propiedades

1) $0 \leq \rho^2_{X,Y} \leq 1$

2) $\rho^2_{X,Y} = 0$ (variables incorreladas (linealmente)) $\iff$ Las rectas de regresión son paralelas a los ejes de coordenadas, esto es $x = \mathrm{E}X$ y $y = \mathrm{E}Y$.

3) Si una de las rectas es paralela al eje correspondiente, entonces $\rho^2_{X,Y} = 0$ y, por tanto, la otra también.

4) $\rho^2_{X,Y} = 1 \iff \exists \alpha \neq 0, \beta \in \mathbb{R}$, tal que$\mathrm{P}\{Y = \alpha X + \beta\} = 1$, es decir, existe una relación lineal entre las variables.

5) $\rho^2_{X,Y} = 1 \iff$ Las dos rectas de regresión coinciden.

6) $0 \leq \rho^2_{X,Y} \leq \eta^2_{Y/X} \leq 1$
$0 \leq \rho^2_{X,Y} \leq \eta^2_{X/Y} \leq 1$

7) $\rho^2_{X,Y} = \eta^2_{Y/X} \Leftrightarrow$ La curva de regresión de Y sobre X coincide con la recta.
$\rho^2_{X,Y} = \eta^2_{X/Y} \Leftrightarrow$ La curva de regresión de X sobre Y coincide con la recta.

Coeficiente de correlación lineal

El coeficiente de determinación lineal entre dos variables, $\rho^2_{X,Y}$, es el análogo, en las rectas de regresión, a las razones de correlación en las curvas, proporcionando la medida del grado de dependencia lineal entre las variables.
Sin embargo, en la regresión lineal, la medida de correlación por excelencia es el denominado coeficiente de correlación lineal que proporciona, no sólo una medida del grado de relación lineal entre las variables (bondad del ajuste por las rectas de regresión), sino el sentido de la relación. Se define como

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}X\mathrm{Var}Y}}$$

para X e Y variables aleatorias con momento de segundo orden y no degeneradas.

Notemos que $\rho_{X,Y}$ tiene el signo de la covarianza y, por tanto, el de los coeficientes de regresión, de manera que $\rho_{X,Y} > 0$ indica que las pendientes de ambas rectas son positivas (cada variable crece con la otra), mientras que $\rho_{X,Y} < 0$ indica pendientes negativas (cada variable decrece con la otra). Además, puesto que $|\rho_{X,Y}| = \sqrt{\rho^2_{X,Y}}$, $|\rho_{X,Y}|$ mide (al igual que $\rho^2_{X,Y}$) la bondad del ajuste lineal.

Propiedades

1) $-1 \leq \rho_{X,Y} \leq 1$

2) Si X e Y son independientes, entonces $\rho_{X,Y} = 0$. El recíproco no es cierto. (El mismo contraejemplo que $\mathrm{Cov}(X,Y) = 0$ no implica X e Y independientes.)
3) $\rho^2_{X,Y} = 0$ (variables incorreladas (linealmente)) $\iff$ Las rectas de regresión son paralelas a los ejes de coordenadas, esto es $x = \mathrm{E}X$ y $y = \mathrm{E}Y$.
4) $|\rho_{X,Y}| = 1 \iff \exists \alpha \neq 0, \beta \in \mathbb{R}$, $\mathrm{P}\{Y = \alpha X + \beta\} = 1$, es decir, existe una relación lineal entre las variables.
5) $|\rho_{X,Y}| = 1 \iff$ Las dos rectas de regresión coinciden y coinciden con la recta de dependencia. Además también coinciden con las curvas de regresión
6) Si $\rho_{X,Y} = 1$, la dependencia funcional lineal exacta es positiva (las rectas de regresión coinciden con la de dependencia que tiene pendiente positiva).
7) Si $\rho_{X,Y} = -1$, la dependencia funcional lineal exacta es negativa (las rectas de regresión coinciden con la de dependencia que tiene pendiente negativa).