CORRELACIÓN

• Usar datos pareados para encontrar el valor del coeficiente de correlación lineal r.
• Determinar si hay evidencia suficiente para respaldar la conclusión de que existe una correlación lineal entre dos variables.

Definición: *Existe una **correlación** entre dos variables cuando los valores de una variable están de alguna manera asociados con los valores de la otra variable. Existe una **correlación lineal** entre dos variables cuando existe una correlación y los puntos graficados de los datos pareados dan como resultado un patrón que se puede aproximar mediante una línea recta.*

Medi. de fuerza la corre. lineal de r

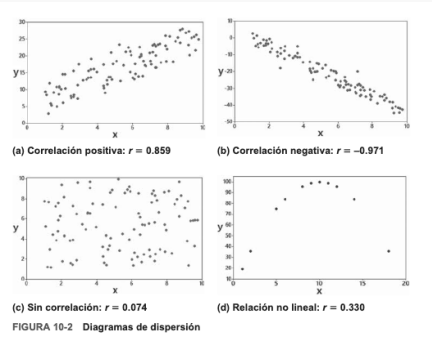


FIGURA 10-2 Diagramas de dispersión

Debido a que las conclusiones basadas en exámenes visuales de diagramas de dispersión son subjetivas en gran medida, necesitamos mediciones más objetivas. Usaremos el coeficiente de correlación lineal r, que es un número que mide la fuerza de la asociación lineal entre las dos variables.

REGRESIÓN

Definición: *Dada una colección de datos muestrales pareados, la línea de regresión (o línea de mejor ajuste, o línea de mínimos cuadrados) es la línea recta que "mejor" se ajusta al diagrama de dispersión de los datos. (El criterio específico para la línea recta de "mejor ajuste" es la propiedad de los "mínimos cuadrados" que se describirá posteriormente). La ecuación de regresión: y = b0 + b1x . Describe algebraicamente la línea de regresión. La ecuación de regresión expresa una relación entre x (llamada variable explicativa, variable predictora, o variable independiente) y (llamada variable de respuesta o variable dependiente).*

INTER. DE LA ECUA. DE REGRESI.

Valores Atípicos y Puntos Influyentes: *En un diagrama de dispersión, un valor atípico es un punto que cae lejos de los demás puntos de datos. Los datos muestrales pareados pueden incluir uno o más puntos influyentes, que son puntos que afectan fuertemente la gráfica de la línea de regresión.*

Cambio Marginal: *Al trabajar con dos variables relacionadas por una ecuación de regresión, el cambio marginal de una variable es la cantidad que cambia cuando la otra se modifica en exactamente una unidad. La pendiente b1 en la ecuación de regresión representa el cambio marginal en y que ocurre cuando x cambia en una unidad.*

INTERV. DE PREDICCIÓN Y VARIACIÓN

• Usar datos muestrales pareados para determinar el valor del coeficiente de determinación r2 e interpretar ese valor. • Utilizar un valor dado de una variable para encontrar un intervalo de predicción para la otra variable.

Definición: *Un intervalo de predicción es un rango de valores utilizados para estimar una variable (como un valor predicho de y en una ecuación de regresión). Un intervalo de confianza es un rango de valores utilizados para estimar un parámetro poblacional (como r, m o s).*

Variación Explicable e Inexplicable

- Hay suficiente evidencia para respaldar la afirmación de una correlación lineal entre x y y.
- La ecuación de la recta de regresión es $\hat{y} = 3 + 2x$.
- La media de los valores de y está dada por $\bar{y} = 9$.
- Uno de los pares de datos muestrales es $x = 5$ y $y = 19$.
- El punto (5, 13) es uno de los puntos en la línea de regresión, porque al sustituir $x = 5$ en la ecuación de regresión $\hat{y} = 3 + 2x$ se obtiene $\hat{y} = 13$.

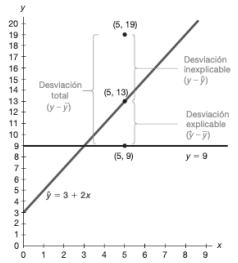


FIGURA 10-7 Desviaciones total, explicable e inexplicable

Coeficiente de Determinación

El valor de r^2 es la proporción de la variación en y que se explica por la relación lineal entre x y y.

Esta declaración sobre la variación explicable se formaliza con la siguiente definición.

DEFINICIÓN

El **coeficiente de determinación** es la proporción de la variación en y que se explica por la línea de regresión. Se calcula como

$$r^2 = \frac{\text{variación explicable}}{\text{variación total}}$$

Podemos calcular r^2 usando la definición que acabamos de dar con la fórmula 10-7, o simplemente podemos elevar al cuadrado el coeficiente de correlación lineal r. Se recomienda usar el cuadrado de r.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com

REGRESIÓN MÚLTIPLE

Definición: *Una ecuación de regresión múltiple expresa una relación lineal entre una variable de respuesta y y dos o más variables predictoras (x1, x2, ..., xk). La forma general de una ecuación de regresión múltiple obtenida a partir de datos muestrales es: y = b0+ b1x1 + b2x2 + ••• + bkxk*

1º. Interpretar los resultados de la tecnología para determinar si una ecuación de regresión múltiple es adecuada para hacer predicciones. 2º. Comparar los resultados de diferentes combinaciones de variables predictoras e identificar la combin ación que resulta en la mejor ecuación de regresión múltiple.

Variables Ficticias y Regresión Logística

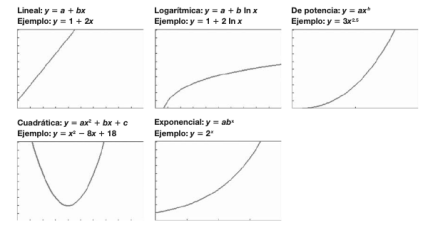
Definición: *Una variable ficticia es una variable que tiene sólo los valores de 0 y 1, utilizados para representar las dos categorías diferentes de una variable cualit ativa.*

REGRESIÓN NO LINEAL

• Usar datos pareados para identificar los modelos lineales, cuadráticos, logarítmicos, exponenciales y de potencia. • Determinar qué modelo se ajusta mejor a los datos pareados.

Definición: *Las secciones anteriores de este capítulo tratan sólo con relaciones lineales, pero no todas en el mundo son lineales. Esta sección es una breve introd ucción a los métodos para encontrar algunas funciones no lineales que se ajustan a los datos muestrales. Nos enfocamos en el uso de la tecnología porque los cálculos requeridos son bastante complejos.*

5 Modelos Genéricos



Cada uno de los modelos se proporciona con una fórmula genérica junto con un ejemplo de una función específica y su gráfica.

Sponsored by **ApolloPad.com**
Everyone has a novel in them. Finish Yours!
https://apollopad.com