

coeficiente de asociación – coefficient of association

Authored by
memjavad

November 17, 2025

RECOMMENDED CITATION

memjavad (2025). *coeficiente de asociación – coefficient of association*. Spanish Psychological Databases. Retrieved from <https://spanish.arabpsychology.com/?p=4932>

Coeficiente de Asociación

Primary Disciplinary Field(s): Estadística, Sociología, Psicometría

1. Definición Central

El coeficiente de asociación es una medida estadística diseñada para cuantificar la fuerza y, en algunos casos, la dirección de la relación entre dos o más variables, particularmente cuando estas variables son de naturaleza cualitativa o categórica. A diferencia de los coeficientes de correlación tradicionales, como el coeficiente de correlación de Pearson, que miden la relación lineal entre variables continuas, los coeficientes de asociación están específicamente adaptados para el análisis de [tablas de contingencia](#). Estas tablas son estructuras fundamentales que resumen la distribución conjunta de las frecuencias de las categorías de las variables analizadas. La esencia del coeficiente de asociación radica en determinar si la distribución observada de las frecuencias difiere significativamente de la distribución que se esperaría si las variables fueran completamente independientes. Un valor alto del coeficiente indica que la pertenencia a una categoría de una variable está fuertemente ligada a la pertenencia a una categoría específica de la otra variable.

La necesidad de estos coeficientes surge del hecho de que muchas investigaciones en ciencias sociales, biológicas y de mercado manejan datos que no cumplen con los supuestos de normalidad o linealidad requeridos por métodos paramétricos. Por ejemplo, al estudiar la relación entre el sexo (masculino/femenino) y la preferencia política (partido A/partido B), las variables son nominales o dicotómicas, haciendo inapropiada la correlación de Pearson. En este contexto, el coeficiente de asociación proporciona una métrica estandarizada que permite comparar la intensidad de la relación entre diferentes conjuntos de datos o estudios. Aunque el término es amplio, a menudo se asocia con medidas derivadas del estadístico de prueba [Chi-cuadrado \(\$\chi^2\$ \)](#), ya que este último evalúa la existencia de una relación, mientras que el coeficiente de asociación mide la magnitud de dicha relación.

Es crucial diferenciar la asociación de la causalidad. Un coeficiente de asociación elevado indica una fuerte relación estadística entre las variables, pero no implica necesariamente que una variable cause la otra. La interpretación debe limitarse a la co-ocurrencia sistemática de las categorías. Además, la elección del coeficiente de asociación apropiado depende críticamente del nivel de medición de las variables involucradas (nominal, ordinal o dicotómica), así como del tamaño y la estructura de la tabla de contingencia (por ejemplo, una tabla 2x2 frente a una tabla RxC). La estandarización de estos coeficientes generalmente asegura que sus valores se encuentren dentro de un rango definido (típicamente entre -1 y +1, o 0 y +1), facilitando su interpretación universal.

2. Contexto Histórico y Desarrollo

El desarrollo formal de los coeficientes de asociación está intrínsecamente ligado al surgimiento de la estadística moderna a finales del siglo XIX y principios del XX, particularmente con la necesidad de analizar datos censales y biológicos categorizados. Antes de esta época, la medición de la relación se centraba predominantemente en variables continuas. La formalización de la estadística de [Karl Pearson](#), especialmente su trabajo en el coeficiente de contingencia y la prueba Chi-cuadrado (1900), sentó las bases para evaluar la independencia en tablas de frecuencias. Sin embargo, el estadístico χ^2 en sí mismo es sensible al tamaño de la muestra y al número de categorías, lo que limitaba su utilidad como medida de fuerza de la asociación que pudiera ser comparable entre estudios.

La limitación del χ^2 motivó a otros estadísticos a buscar medidas estandarizadas. El estadístico británico [George Udny Yule](#) (1871-1951) jugó un papel fundamental en este desarrollo. Yule se centró en el análisis de variables dicotómicas (tablas 2x2), introduciendo en 1900 el famoso [coeficiente Q de Yule](#) (también conocido como coeficiente de asociación) y el coeficiente de coligación (Yule's Y). Estos coeficientes se basaron en las frecuencias de las casillas de la tabla y fueron diseñados específicamente para alcanzar los límites de +1 y -1 en caso de asociación perfecta positiva o negativa, respectivamente, independientemente del tamaño de la muestra. El trabajo de Yule fue decisivo para establecer la distinción entre la prueba de independencia (χ^2) y la medida de la intensidad de la asociación (el coeficiente).

Posteriormente, la necesidad de un coeficiente aplicable a tablas R x C (con múltiples filas y columnas) llevó al desarrollo de medidas como el [Coeficiente V de Cramér](#), introducido por el estadístico sueco Harald Cramér en 1946. Este coeficiente es una adaptación del estadístico χ^2 que lo normaliza para que su valor máximo sea 1, permitiendo la comparación de la fuerza de asociación entre tablas de diferentes dimensiones. La evolución histórica muestra una progresión desde la simple prueba de independencia hasta la creación de métricas robustas y estandarizadas que son esenciales para la investigación empírica en la actualidad.

3. Tipos Clave de Coeficientes de Asociación

Existe una variedad de coeficientes de asociación, cada uno optimizado para un tipo específico de variables y estructura de tabla. La selección adecuada del coeficiente es vital para obtener una interpretación válida de los datos. Entre los más utilizados se encuentran el Coeficiente Phi, el Coeficiente V de Cramér y los coeficientes específicos para variables ordinales.

El **Coeficiente Phi (ϕ)** es la medida de asociación más utilizada para tablas de contingencia 2x2, es decir, cuando ambas variables son dicotómicas. Se calcula tomando la raíz cuadrada del estadístico χ^2 dividido por el tamaño de la muestra (N). Su rango de valores es de 0 a 1, si las variables son puramente dicotómicas y no hay una dirección inherente (como en el caso de

dos variables nominales). Sin embargo, si las variables son binarias con un orden implícito, puede oscilar entre -1 y +1. Phi es conceptualmente similar al coeficiente de correlación de Pearson aplicado a datos binarios. Su principal limitación es que solo es interpretable y comparable en el contexto de tablas 2x2.

Para tablas de contingencia más grandes ($R \times C$, donde R o C son mayores que 2), el **Coeficiente V de Cramér** se convierte en la herramienta preferida. Este coeficiente es una medida de asociación de post-hoc que corrige el estadístico Chi-cuadrado dividiéndolo por su valor máximo posible y ajustándolo por el número mínimo de filas o columnas ($\min(R-1, C-1)$). Esto asegura que V de Cramér siempre varíe entre 0 (ausencia total de asociación) y 1 (asociación perfecta). V de Cramér es robusto y permite comparar la fuerza de la asociación entre tablas de diferentes tamaños, lo cual es una ventaja significativa sobre otros coeficientes como el Coeficiente de Contingencia (C), que raramente alcanza el valor de 1.

Cuando las variables son de naturaleza ordinal (es decir, categóricas con un orden lógico, como niveles de acuerdo: bajo, medio, alto), se utilizan coeficientes que tienen en cuenta este orden para medir la concordancia o discordancia. Ejemplos prominentes incluyen el **Tau de Kendall (\$tau\$)** y el **Gamma de Goodman y Kruskal (\$gamma\$)**. Estos coeficientes miden la probabilidad de que dos pares de observaciones estén en el mismo orden (pares concordantes) versus órdenes inversos (pares discordantes). Gamma, en particular, es popular porque ignora los empates y puede ser interpretado como la diferencia entre la probabilidad de concordancia y discordancia, dado que no hay empates.

4. Interpretación y Rango de Valores

La interpretación de los coeficientes de asociación depende fundamentalmente de su rango y de si miden únicamente la fuerza (magnitud) o también la dirección (sentido). La mayoría de los coeficientes están estandarizados para operar dentro de límites claros, facilitando la comprensión de la intensidad de la relación observada.

Coeficientes como el V de Cramér o el Coeficiente de Contingencia (C) tienen un rango de 0 a +1. Un valor de 0 indica que las variables son estadísticamente independientes, lo que significa que la distribución de una variable es la misma en todas las categorías de la otra variable. Un valor de +1 indica una asociación perfecta, donde el conocimiento de la categoría de una variable permite predecir con certeza la categoría de la otra. Sin embargo, en la práctica, los valores de asociación perfecta son raros en datos reales, y la mayoría de los hallazgos se sitúan en valores intermedios.

Otros coeficientes, como el Coeficiente Q de Yule, el Phi, el Tau de Kendall y el Gamma, tienen un rango de -1 a +1. En este caso, el signo es vital. Un valor cercano a +1 indica una fuerte asociación positiva (las categorías altas de una variable tienden a ocurrir con las categorías altas de la otra, o la presencia de A implica la presencia de B). Un valor cercano a -1 indica una fuerte

asociación negativa o inversa (la presencia de A implica la ausencia de B, o las categorías altas de una variable se asocian con las categorías bajas de la otra). Un valor cercano a 0 indica una asociación débil o nula. La interpretación de la magnitud (fuerza) se basa en el valor absoluto del coeficiente.

Aunque no existen reglas universales estrictas, las guías comunes para interpretar la fuerza de la asociación (basadas en el valor absoluto) a menudo sugieren que valores entre 0.00 y 0.10 representan una asociación insignificante o muy débil; 0.10 a 0.30, una asociación débil; 0.30 a 0.50, una asociación moderada; y valores superiores a 0.50, una asociación fuerte o muy fuerte. Es fundamental que esta interpretación se realice siempre en el contexto disciplinario específico. Por ejemplo, en física, un coeficiente de 0.5 puede ser considerado bajo, mientras que en sociología o psicología, puede representar un hallazgo sustancial debido a la alta variabilidad de los fenómenos humanos.

5. Aplicaciones en Diversas Disciplinas

Los coeficientes de asociación son herramientas metodológicas esenciales con amplias aplicaciones en prácticamente todas las áreas de la investigación empírica que maneja datos categóricos. Su versatilidad permite transformar la información cualitativa en métricas cuantitativas interpretables, lo que es crucial para la toma de decisiones basada en evidencia.

En las **ciencias sociales**, especialmente en sociología y ciencia política, estos coeficientes se utilizan para analizar la relación entre variables demográficas y actitudes o comportamientos. Por ejemplo, se puede emplear el V de Cramér para determinar la fuerza de la asociación entre el nivel educativo (bajo, medio, alto) y la elección de un candidato político. El Coeficiente Q de Yule es frecuentemente utilizado en estudios de opinión para correlacionar respuestas binarias, como la aprobación o desaprobación de dos políticas distintas.

En la **epidemiología y la salud pública**, los coeficientes de asociación son vitales para evaluar la relación entre factores de riesgo categóricos (ej. tabaquismo: sí/no) y la presencia de enfermedades (ej. cáncer de pulmón: sí/no). Aunque en este campo se utilizan a menudo medidas específicas como la razón de probabilidades (Odds Ratio), los coeficientes de asociación proporcionan una medida de la magnitud de la dependencia estadística en tablas 2x2.

En el **análisis de datos de mercado y la minería de datos**, los coeficientes de asociación, a menudo bajo el nombre de "reglas de asociación", se utilizan para descubrir patrones de co-ocurrencia en grandes bases de datos. Por ejemplo, pueden identificar qué productos tienden a ser comprados juntos (análisis de la cesta de mercado), lo que tiene implicaciones directas para la estrategia de colocación de productos y las campañas de marketing. La capacidad de estos coeficientes para manejar datos categóricos los hace indispensables para el análisis exploratorio inicial en campos de datos masivos.

6. Limitaciones y Advertencias Metodológicas

A pesar de su utilidad, los coeficientes de asociación presentan varias limitaciones metodológicas que deben ser consideradas cuidadosamente por el investigador para evitar interpretaciones erróneas.

Una limitación fundamental es su sensibilidad a la **distribución marginal** de las variables. El valor de muchos coeficientes, como el Phi o el Coeficiente de Contingencia, puede verse artificialmente reducido si las distribuciones de las variables están muy sesgadas (es decir, cuando la mayoría de las observaciones caen en una sola categoría). Una asociación perfecta teórica puede no alcanzarse en la práctica si las distribuciones marginales son desiguales, lo que dificulta la comparación directa de la fuerza de asociación entre diferentes estudios que manejan distribuciones marginales muy diferentes.

Otra advertencia importante es la **dependencia del tamaño de la tabla**. Aunque V de Cramér fue diseñado para mitigar esto, algunos coeficientes más antiguos, como el Coeficiente de Contingencia (C), tienen un valor máximo que depende del número de filas y columnas. Por ejemplo, en una tabla 2x2, el máximo teórico de C es aproximadamente 0.707, mientras que en una tabla 3x3 es 0.816. Esto significa que un valor de $C = 0.70$ en una tabla 2x2 representa una asociación mucho más fuerte (casi perfecta) que el mismo valor en una tabla 3x3. Por lo tanto, se requiere cautela al comparar coeficientes entre tablas de diferentes dimensiones.

Finalmente, la interpretación de la asociación en variables ordinales mediante coeficientes como Gamma o Tau puede ser engañosa si existe una alta proporción de **empates**. El Coeficiente Gamma, por ejemplo, ignora explícitamente los pares empatados, lo que puede sobreestimar la fuerza de la asociación si el número de empates es grande. En estos casos, es esencial reportar también medidas que ajusten o incorporen los empates para ofrecer una imagen completa de la estructura de la relación. La validez de cualquier coeficiente de asociación está intrínsecamente ligada a la validez de la prueba Chi-cuadrado subyacente, lo que implica que las frecuencias esperadas no deben ser demasiado pequeñas (generalmente, no menos de 5 en la mayoría de las celdas).

7. Comparación con Otras Medidas de Relación

Es fundamental distinguir los coeficientes de asociación de otras medidas estadísticas de relación, principalmente el coeficiente de correlación de Pearson y los coeficientes de regresión. Esta distinción se basa principalmente en el nivel de medición de las variables y el propósito analítico.

El **Coeficiente de Correlación de Pearson (r)** es la medida de relación más común para variables continuas (intervalo o razón). Mide la fuerza y dirección de una relación lineal. Si bien el coeficiente Phi es matemáticamente equivalente a r cuando se aplica a dos variables binarias

codificadas, Pearson requiere que las variables se distribuyan normalmente y que la relación sea lineal. Los coeficientes de asociación, en cambio, no asumen normalidad ni linealidad, siendo adecuados para datos nominales u ordinales donde el concepto de "línea recta" carece de significado.

Las **Medidas de Regresión**, como el coeficiente beta en la regresión lineal o los coeficientes logísticos en la regresión logística, tienen un propósito diferente. Mientras que un coeficiente de asociación solo mide la fuerza de la relación mutua, los coeficientes de regresión miden el cambio esperado en una variable dependiente (Y) resultante de un cambio unitario en una variable independiente (X). Es decir, la regresión introduce una asimetría causal (X predice Y), mientras que la mayoría de los coeficientes de asociación son simétricos, tratando a ambas variables por igual. Solo en el contexto de variables ordinales (como el Tau de Kendall), la dirección puede ser interpretada, pero sin la inferencia causal de la regresión.

Finalmente, la comparación con el propio estadístico χ^2 es clave. El valor del χ^2 indica la probabilidad de que la asociación observada no se deba al azar (significatividad estadística), pero no su magnitud. Un χ^2 grande en una muestra grande puede corresponder a un coeficiente de asociación muy pequeño (relación débil pero estadísticamente significativa). Por lo tanto, los coeficientes de asociación actúan como medidas de efecto (effect size), complementando la prueba de hipótesis del χ^2 al proporcionar una magnitud estandarizada de la fuerza de la relación.

Further Reading

[Coeficiente de Asociación \(Wikipedia\)](#)

[Yule, George Udny. Contribuciones a la estadística de asociación.](#)

[Cramér, Harald. Mathematical Methods of Statistics \(1946\).](#)

[Coeficiente Phi y su uso en tablas 2x2.](#)