



Módulo 1. Asociación y Correlación entre Variables

☰ 1. Asociación entre variables categóricas

☰ 2. Correlación en variables cuantitativas

☰ Referencias

1. Asociación entre variables categóricas

Introducción

En el análisis profesional de datos, una de las preguntas más frecuentes consiste en determinar si dos características se relacionan entre sí o si su comportamiento responde a patrones independientes. En ámbitos como la salud, la gestión empresarial, la investigación social o el análisis de mercados, las variables categóricas aparecen de manera constante: diagnóstico clínico, nivel de satisfacción, tipo de contrato, categoría de producto, zona geográfica, entre muchas otras. Frente a este escenario, surge un interrogante central: ¿cómo evaluar de manera rigurosa si existe asociación entre dos variables cualitativas?

Cuando se cruzan dos variables categóricas en una población, el profesional requiere herramientas que permitan identificar patrones de distribución conjunta. ¿La frecuencia de un determinado resultado es similar en todos los grupos? ¿Se concentran ciertos valores en combinaciones específicas? ¿La

distribución observada responde al azar o revela una estructura subyacente? Estas preguntas orientan la necesidad de construir tablas de contingencia y aplicar coeficientes de asociación que cuantifiquen la intensidad del vínculo.

El análisis de asociación no se limita a verificar la existencia de relación; también permite estimar su magnitud y comprender su sentido práctico. En contextos organizacionales, por ejemplo, identificar asociación entre tipo de capacitación y desempeño laboral aporta información para la toma de decisiones estratégicas. En estudios epidemiológicos, evaluar la relación entre exposición y condición de salud permite describir patrones poblacionales.

En esta unidad abordaremos la construcción e interpretación de tablas de contingencia, los principales coeficientes de asociación para variables cualitativas y los alcances del análisis cuando se trabaja con datos categóricos. El objetivo consiste en desarrollar criterios técnicos que permitan interpretar adecuadamente los resultados y fundamentar decisiones basadas en evidencia cuantitativa.

Tablas de contingencia: construcción e interpretación



El análisis de asociación entre variables categóricas se estructura, en primer término, a partir de la organización sistemática de los datos en una tabla de contingencia. Esta herramienta permite representar la distribución conjunta de dos variables cualitativas mediante el cruce de sus categorías, facilitando la observación de patrones de concentración, dispersión o equilibrio en las frecuencias. Desde el punto de vista profesional, la tabla de contingencia constituye el punto de partida para evaluar si las diferencias observadas entre grupos responden a una estructura sistemática o a variaciones atribuibles al azar.

Cuando se dispone de dos variables categóricas, cada individuo de la muestra pertenece simultáneamente a una categoría de cada variable. La tabla de contingencia organiza esta información en filas y columnas, donde cada celda contiene la frecuencia absoluta correspondiente a la combinación específica de categorías. Tal como expone Moral Peláez (s.f.), las medidas de asociación requieren previamente una estructura de datos que permita evaluar la relación entre variables, lo cual se logra mediante tablas cruzadas que sintetizan la información de forma ordenada y comparable.

Desde una perspectiva operativa, la construcción de una tabla de contingencia implica tres decisiones técnicas: definir correctamente las categorías, asegurar la exhaustividad y exclusividad de las clasificaciones y verificar la coherencia de las frecuencias totales. Este procedimiento resulta especialmente relevante en ámbitos como la gestión hospitalaria, donde se analiza la asociación entre tipo de tratamiento y evolución clínica, o en estudios organizacionales que examinan la relación entre modalidad de capacitación y nivel de desempeño laboral.

Tabla 1. Estructura general de una tabla de contingencia 2×2

	Categoría B1	Categoría B2	Total fila
Categoría A1	a	b	a + b
Categoría A2	c	d	c + d
Total columna	a + c	b + d	n

Fuente: elaboración propia

Esta estructura básica permite identificar la frecuencia conjunta y las frecuencias marginales. El total general (n) representa el

tamaño muestral. A partir de esta organización, el análisis se orienta a examinar si la distribución de las frecuencias revela concentraciones diferenciales entre categorías.

En estudios aplicados, la interpretación comienza con el análisis descriptivo de proporciones. Por ejemplo, si se evalúa la asociación entre tipo de contrato (temporal/permanente) y acceso a capacitación (sí/no), la comparación de proporciones por fila o por columna permite detectar patrones preliminares. La observación de diferencias sistemáticas constituye el primer indicio de posible asociación.

Sin embargo, la sola inspección visual puede inducir interpretaciones imprecisas. Ortiz Pinilla y Ortiz Rico (2021) señalan que el análisis de asociación requiere claridad conceptual respecto del tipo de relación que se pretende describir, evitando asumir equivalencias metodológicas que no corresponden. En el caso de variables categóricas, la interpretación se centra en la estructura de frecuencias y en la magnitud relativa de las diferencias.

Tabla 2. Ejemplo aplicado: Tipo de tratamiento y recuperación clínica

Tipo de tratamiento	Recuperación favorable	Recuperación desfavorable	Total
Tratamiento A	45	15	60
Tratamiento B	30	30	60
Total	75	45	120

Fuente: elaboración propia

En este ejemplo hipotético del ámbito sanitario, se observa que el Tratamiento A presenta una mayor proporción de recuperaciones favorables respecto del Tratamiento B. La tabla permite identificar que el 75% de los pacientes del Tratamiento A presentan evolución favorable, mientras que en el Tratamiento B la proporción alcanza el 50%. Este patrón sugiere una posible asociación entre tratamiento y resultado clínico.

El análisis de tablas de contingencia se vincula históricamente con el desarrollo de métodos de evaluación de asociación en estadística aplicada. La formalización de medidas para variables categóricas permitió avanzar desde la descripción simple hacia la

cuantificación de la intensidad de la relación. La siguiente línea de tiempo sintetiza esta evolución conceptual.

Evolución del análisis de asociación en variables categóricas

- **Finales del siglo XIX:** Desarrollo inicial del análisis de frecuencias cruzadas en estudios demográficos.
- **1900-1930:** Formalización de métodos estadísticos para evaluar independencia entre variables categóricas.
- **Décadas posteriores:** Introducción de coeficientes de asociación como Cramér y medidas para variables ordinales.
- **Actualidad:** Integración del análisis de contingencia en modelos predictivos y sistemas de apoyo a decisiones.

La sistematización del análisis de asociación amplió la capacidad interpretativa de los estudios observacionales. Martínez Ortega et al. (2009) destacan que el análisis de relaciones entre variables requiere comprender la naturaleza de los datos y el tipo de correspondencia que se evalúa. En el caso de variables categóricas, la relación se expresa en términos de coincidencias y distribución relativa de frecuencias.

Tabla 3. Comparación entre frecuencias absolutas y proporciones relativas

Modalidad de capacitación	Desempeño alto (f)	Desempeño alto (%)	Total
Presencial	32	64%	50
Virtual	18	36%	50
Total	50	—	100

Fuente: elaboración propia

Esta tabla ilustra la importancia de complementar la lectura de frecuencias absolutas con proporciones relativas. La interpretación profesional se apoya en porcentajes que permiten comparar grupos con mayor precisión, especialmente cuando los tamaños muestrales difieren.

En términos conceptuales, la tabla de contingencia representa un instrumento descriptivo que organiza la información y prepara el terreno para la estimación de coeficientes de asociación. Camacho-Sandoval (2008) señala que la elección del método de análisis depende de la naturaleza de las variables y de las

condiciones de aplicación. Aunque su trabajo se centra en correlación no paramétrica, el principio de adecuación metodológica resulta igualmente aplicable al análisis de variables cualitativas.

En síntesis, la tabla de contingencia constituye la base estructural del análisis de asociación entre variables categóricas. Permite visualizar patrones, comparar proporciones y fundamentar la aplicación posterior de medidas cuantitativas que estimen la intensidad del vínculo. Su correcta construcción y lectura garantizan que el análisis avance desde la descripción organizada hacia la interpretación rigurosa de la relación entre variables.

Coeficientes de asociación para variables cualitativas

Una vez organizada la información mediante tablas de contingencia, el análisis profesional requiere avanzar hacia la cuantificación de la intensidad de la relación observada. La inspección visual de frecuencias permite identificar patrones preliminares; sin embargo, la comparación sistemática entre

estudios, poblaciones o períodos demanda indicadores numéricos que sintetizen el grado de asociación. En este contexto surgen los coeficientes de asociación para variables cualitativas.

El análisis de asociación en variables categóricas se orienta a responder una pregunta central: ¿en qué medida la distribución conjunta de dos variables se aparta de lo que cabría esperar si fueran independientes? Moral Peláez (s.f.) expone que, cuando se desea evaluar la relación entre variables cualitativas nominales, el coeficiente de Cramér constituye una herramienta adecuada para estimar la magnitud del vínculo. Este coeficiente permite expresar la intensidad de la asociación en una escala comprendida entre 0 y 1.

Desde una perspectiva técnica, los coeficientes de asociación cumplen tres funciones principales:

- Sintetizar la información contenida en la tabla de contingencia.
- Permitir la comparación entre estudios con diferentes tamaños muestrales.
- Proporcionar un indicador estandarizado de intensidad de relación.

En el caso de variables nominales, el coeficiente V de Cramér se calcula a partir del estadístico chi-cuadrado y del tamaño de la muestra. Su interpretación se basa en la magnitud del valor obtenido: valores próximos a 0 indican asociación débil, mientras que valores cercanos a 1 reflejan asociación fuerte.

Tabla 4. Coeficientes de asociación según tipo de variables

Tipo de variables	Medida de asociación sugerida	Rango de valores	Interpretación general
Nominal × Nominal	V de Cramér	0 a 1	Intensidad sin dirección
Ordinal × Ordinal	Gamma de Goodman-Kruskal	-1 a 1	Intensidad y dirección
Ordinal × Ordinal / Intervalo	Tau de Kendall	-1 a 1	Concordancia ordinal

Fuente: elaboración propia

La elección del coeficiente depende de la naturaleza de las variables. Martínez Ortega et al. (2009) enfatizan que el análisis de asociación requiere considerar el tipo de relación que se evalúa y la estructura de los datos. Cuando las variables poseen orden inherente, resulta pertinente emplear medidas que incorporen dirección y sentido del vínculo.

El coeficiente gamma y el coeficiente tau de Kendall introducen el concepto de concordancia y discordancia entre pares de observaciones. En términos operativos, se examina si los ordenamientos relativos de las variables coinciden o divergen. Esta lógica permite capturar asociaciones monotónicas entre categorías ordenadas.

Ortiz Pinilla y Ortiz Rico (2021) subrayan que la interpretación de un coeficiente debe estar alineada con el tipo de relación conceptual que se pretende describir. En variables ordinales, el interés se centra en identificar tendencias crecientes o decrecientes entre categorías, lo cual exige medidas que incorporen dirección.

Para ilustrar su aplicación, consideremos un estudio organizacional en el que se analiza la relación entre nivel de experiencia laboral (bajo, medio, alto) y evaluación de

desempeño (insuficiente, adecuado, sobresaliente). En este caso, ambas variables presentan orden jerárquico.

Tabla 5. Ejemplo aplicado: Experiencia laboral y desempeño

Nivel de experiencia	Insuficiente	Adecuado	Sobresaliente	Total
Bajo	20	10	5	35
Medio	8	22	10	40
Alto	2	15	28	45
Total	30	47	43	120

Fuente: elaboración propia

En esta distribución se observa una tendencia creciente entre experiencia y desempeño. Las categorías superiores de experiencia concentran mayores frecuencias en niveles altos de evaluación. Un coeficiente gamma positivo elevado reflejaría esta concordancia ordinal.

En términos analíticos, el procedimiento general para seleccionar y aplicar un coeficiente de asociación puede organizarse en los siguientes pasos:

- Identificar el nivel de medición de cada variable.
- Construir la tabla de contingencia correspondiente.
- Determinar si las categorías poseen orden inherente.
- Seleccionar el coeficiente adecuado según la estructura de los datos.
- Interpretar la magnitud del valor obtenido en coherencia con el contexto profesional.

Camacho-Sandoval (2008) señala que la adecuación metodológica resulta determinante para obtener interpretaciones consistentes. Aunque su trabajo se centra en correlación no paramétrica, el principio de correspondencia entre método y naturaleza de los datos se aplica igualmente a variables cualitativas.

En síntesis, los coeficientes de asociación para variables cualitativas permiten cuantificar la intensidad del vínculo identificado en las tablas de contingencia. La elección correcta de

la medida depende del tipo de variables y del carácter ordinal o nominal de las categorías. Su aplicación sistemática fortalece el análisis comparativo y proporciona una base cuantitativa sólida para la toma de decisiones en contextos profesionales.

Alcances y límites del análisis de asociación en variables categóricas

El análisis de asociación entre variables categóricas proporciona herramientas robustas para describir la estructura conjunta de datos cualitativos. Sin embargo, su aplicación profesional exige reconocer tanto sus posibilidades interpretativas como sus restricciones metodológicas. Comprender estos alcances y límites permite evitar conclusiones que excedan la información efectivamente contenida en las tablas de contingencia y en los coeficientes de asociación.

En primer término, el análisis de asociación permite identificar patrones de dependencia estadística entre variables. Cuando las distribuciones conjuntas difieren sistemáticamente de lo esperado bajo independencia, se obtiene evidencia cuantitativa de relación. Esta evidencia resulta particularmente valiosa en estudios observacionales, donde se describen comportamientos poblacionales sin intervención experimental.

Martínez Ortega et al. (2009) señalan que la interpretación de los coeficientes debe realizarse en coherencia con el fenómeno estudiado, evitando atribuir significados que no se derivan directamente del análisis estadístico. En el caso de variables categóricas, la asociación describe estructura de distribución, no causalidad.

Alcances del análisis de asociación

Desde una perspectiva aplicada, el análisis de tablas de contingencia y coeficientes de asociación permite:

- Detectar concentraciones diferenciales entre categorías.
- Comparar patrones de distribución entre grupos.
- Estimar intensidad del vínculo mediante indicadores estandarizados.
- Fundamentar decisiones basadas en evidencia descriptiva.

En ámbitos organizacionales, por ejemplo, la identificación de asociación entre modalidad de contratación y permanencia laboral permite describir patrones de estabilidad. En estudios epidemiológicos, la asociación entre exposición y condición clínica orienta la formulación de hipótesis posteriores.

Sin embargo, estos resultados deben interpretarse dentro de su marco conceptual. Ortiz Pinilla y Ortiz Rico (2021) advierten que la utilización de herramientas estadísticas requiere claridad sobre el tipo de relación que se analiza y el alcance interpretativo del indicador empleado. En variables categóricas nominales, la asociación cuantifica intensidad, pero no establece dirección causal ni magnitud de efecto en términos funcionales.

Tabla 6. Alcances y límites del análisis de asociación en variables categóricas

Dimensión evaluada	Alcance del análisis	Límite metodológico
Dependencia estadística	Detecta distribución conjunta no aleatoria	No demuestra causalidad
Intensidad del vínculo	Cuantifica magnitud relativa de asociación	No indica impacto práctico directo
Comparación entre grupos	Permite análisis comparativo entre	Sensible al tamaño muestral

	categorias	
Variables ordinales	Puede incorporar dirección de tendencia	No identifica forma funcional de la relación

Fuente: elaboración propia

Uno de los límites centrales consiste en la imposibilidad de inferir causalidad a partir de asociación. La coexistencia sistemática de categorías no implica que una variable determine a la otra. Moral Peláez (s.f.) recuerda que las medidas de asociación describen relación estadística, lo cual exige complementar el análisis con fundamentos teóricos y contextuales.

Asimismo, el tamaño muestral influye en la estabilidad de los coeficientes. En muestras reducidas, pequeñas variaciones en frecuencias pueden generar cambios relevantes en los valores obtenidos. Este aspecto requiere considerar la consistencia de los resultados y su replicabilidad.

Desde el punto de vista técnico, el análisis de asociación puede representarse mediante el siguiente esquema conceptual:

Figura 1. Proceso analítico en asociación de variables categóricas



Fuente: elaboración propia

Este esquema evidencia que el coeficiente constituye una etapa dentro de un proceso más amplio. Camacho-Sandoval (2008) destaca que la adecuación metodológica depende de la naturaleza de las variables y del objetivo del análisis. La coherencia entre diseño, tipo de datos y herramienta estadística garantiza interpretaciones sólidas.

Otro límite relevante se vincula con la naturaleza categórica de la información. Las categorías nominales carecen de orden intrínseco; por lo tanto, los coeficientes aplicables describen intensidad sin dirección. En cambio, cuando las variables son ordinales, la asociación puede interpretarse en términos de tendencia creciente o decreciente, aunque sin especificar forma funcional.

En síntesis, el análisis de asociación en variables categóricas constituye una herramienta descriptiva que permite identificar y cuantificar patrones de dependencia estadística. Sus resultados adquieren valor cuando se integran con fundamentos teóricos y análisis contextual. Reconocer sus límites metodológicos fortalece la calidad interpretativa y evita extrapolaciones indebidas.

CONTINUAR

2. Correlación en variables cuantitativas

Introducción

En múltiples contextos profesionales surge una pregunta recurrente: ¿qué tan estrechamente se relacionan dos variables numéricas? Cuando se analiza la relación entre edad y presión arterial, inversión en capacitación y productividad, horas de estudio y rendimiento académico, o consumo energético y temperatura ambiental, el interés se centra en identificar patrones de comportamiento conjunto. ¿A medida que una variable aumenta, la otra también lo hace? ¿La relación mantiene una proporción constante? ¿La intensidad del vínculo permite realizar estimaciones confiables?

En el ámbito científico, los coeficientes de correlación se han convertido en herramientas ampliamente difundidas. Resulta llamativo que el coeficiente de correlación de Pearson haya sido propuesto a fines del siglo XIX y continúe siendo uno de los indicadores más utilizados en investigación empírica. De manera similar, el coeficiente de Spearman, desarrollado a comienzos del

siglo XX, introdujo una alternativa basada en rangos que amplió las posibilidades de análisis cuando las condiciones de linealidad estricta no se verifican.

Un dato curioso radica en que ambos coeficientes pueden alcanzar valores idénticos en situaciones de relación lineal perfecta, pero divergen significativamente cuando la relación es monótona sin ser lineal. ¿Qué ocurre cuando los datos siguen una curva creciente pero no se ajustan a una recta? ¿Puede un valor alto de correlación garantizar proporcionalidad en los cambios? ¿Qué implica que un coeficiente se aproxime a 1 o a -1?

La correlación constituye un punto de partida para comprender la estructura de relación entre variables cuantitativas. Permite evaluar dirección, intensidad y consistencia del vínculo observado. Sin embargo, su interpretación requiere distinguir entre asociación monótona y relación lineal, así como entre descripción estadística y explicación causal.

En esta unidad abordaremos los fundamentos del coeficiente de correlación de Pearson, el coeficiente de Spearman y los criterios técnicos que orientan su interpretación. El objetivo consiste en desarrollar una lectura rigurosa de estos indicadores, comprendiendo tanto su potencial descriptivo como las condiciones que delimitan su uso profesional.

Correlación de Pearson: fundamentos y aplicación

El coeficiente de correlación de Pearson constituye una medida estadística diseñada para evaluar la intensidad y dirección de la relación lineal entre dos variables cuantitativas. Su desarrollo formal se vincula con los estudios sobre regresión y herencia realizados a fines del siglo XIX. Desde entonces, se consolidó como uno de los indicadores más utilizados en análisis de datos numéricos.

El interés central de esta medida consiste en responder una pregunta específica: ¿qué tan ajustable es una recta al conjunto de puntos que representan las parejas de valores observados? Ortiz Pinilla y Ortiz Rico (2021) señalan que el coeficiente de Pearson se encuentra íntimamente ligado al concepto de linealidad, dado que su valor máximo absoluto se alcanza únicamente cuando todos los puntos se ubican exactamente

sobre una recta. Esta característica define su alcance interpretativo.

Desde el punto de vista matemático, el coeficiente se construye a partir de la covarianza entre variables estandarizadas. Esto implica que el indicador:

- Considera la variabilidad conjunta de ambas variables.
- Ajusta las diferencias respecto de sus medias.
- Elimina la influencia de las unidades de medida.

El resultado se expresa en un intervalo comprendido entre -1 y 1 . Los valores positivos indican relación directa; los negativos, relación inversa; y los valores cercanos a 0 reflejan ausencia de relación lineal.

Tabla 8. Interpretación general del coeficiente de correlación de Pearson

Valor de r	Interpretación lineal
$r = 1$	Relación lineal creciente perfecta

$0,8 \leq r < 1$	Relación lineal alta
$0,4 \leq r < 0,8$	Relación lineal moderada
$0 < r < 0,4$	Relación lineal baja
$r = 0$	Ausencia de relación lineal
$-1 \leq r < 0$	Relación lineal inversa (según magnitud)

Fuente: elaboración propia

La interpretación de estos valores debe contextualizarse. Martínez Ortega et al. (2009) explican que el coeficiente mide intensidad de relación matemática, sin implicar causalidad. La presencia de un valor elevado indica proximidad a un modelo lineal, pero no identifica el mecanismo subyacente que produce la asociación.

En el ejercicio profesional, la correlación de Pearson se aplica en situaciones donde ambas variables son cuantitativas continuas y existe fundamento conceptual para suponer proporcionalidad en los cambios. Por ejemplo, en un estudio industrial se podría analizar la relación entre

horas de funcionamiento de una máquina y nivel de producción diaria.

Si la relación mantiene un patrón proporcional, el coeficiente de Pearson ofrece un indicador consistente.

Sin embargo, el análisis exige verificar ciertas condiciones técnicas:

- Existencia de relación aproximadamente lineal.
- Ausencia de valores extremos que distorsionen la tendencia.
- Distribución conjunta compatible con los supuestos del modelo.

Ortiz Pinilla y Ortiz Rico (2021) advierten que un valor alto de correlación no garantiza necesariamente linealidad estricta si la estructura de los datos presenta curvaturas o patrones complejos. Esta observación resulta relevante cuando se interpretan resultados sin apoyo gráfico.

Para ilustrar su aplicación, consideremos un ejemplo en el ámbito educativo donde se analiza la relación entre horas de estudio semanal y calificación obtenida en un examen.

Tabla 9. Ejemplo aplicado: Horas de estudio y calificación

Estudiante	Horas de estudio	Calificación
1	5	60
2	8	72
3	10	78
4	12	85
5	15	92

Fuente: elaboración propia

En este conjunto hipotético, se observa una tendencia creciente proporcional. El cálculo del coeficiente de Pearson produciría un valor cercano a 1, indicando fuerte relación lineal positiva. Esta magnitud respalda la posibilidad de modelar la relación mediante una recta.

Camacho-Sandoval (2008) destaca que la selección del método de correlación depende de la naturaleza de los datos y de la

estructura de la relación. Cuando la proporcionalidad resulta consistente, la correlación de Pearson proporciona una estimación adecuada de la intensidad lineal.

En síntesis, el coeficiente de correlación de Pearson permite cuantificar la intensidad de una relación lineal entre variables cuantitativas. Su interpretación exige verificar coherencia entre modelo y datos observados, así como reconocer que la asociación matemática describe comportamiento conjunto sin establecer causalidad.

Correlación de Spearman: análisis no paramétrico

El coeficiente de correlación de Spearman surge como una medida destinada a evaluar la asociación monotónica entre dos variables cuantitativas u ordinales. A diferencia del coeficiente de Pearson, su cálculo se basa en los rangos de los datos y no en sus valores originales. Esta característica modifica sustancialmente el tipo de relación que describe. Mientras Pearson cuantifica linealidad, Spearman evalúa concordancia en el ordenamiento de las observaciones.

Martínez Ortega et al. (2009) explican que el coeficiente de Spearman, también denominado rho de Spearman, se construye

a partir de la diferencia entre los rangos asignados a cada observación en ambas variables. El procedimiento consiste en ordenar cada variable de menor a mayor, asignar posiciones numéricas y luego analizar la coherencia entre dichos ordenamientos. Cuando los valores más altos de una variable coinciden sistemáticamente con los valores más altos de la otra, la asociación es directa; cuando coinciden con los más bajos, la asociación es inversa.

Desde el punto de vista conceptual, el coeficiente de Spearman mide monotonía. Ortiz Pinilla y Ortiz Rico (2021) desarrollan con precisión esta distinción al señalar que una relación puede ser estrictamente creciente sin ser lineal. En estos casos, la correlación de Pearson puede producir valores intermedios que sugieren una relación débil o moderada, mientras que Spearman detecta coherencia ordinal perfecta. La diferencia radica en que la proporcionalidad de los cambios no constituye requisito para Spearman; basta con que el orden relativo de las observaciones se conserve.

Esta propiedad adquiere relevancia en situaciones donde los datos presentan curvaturas, distribuciones no normales o valores extremos que alteran la covarianza. Camacho-Sandoval (2008) destaca que la prueba de Spearman resulta apropiada cuando las variables no cumplen el supuesto de normalidad bivariada o cuando se trabaja con escalas ordinales. En estos contextos, el

uso de rangos atenúa la influencia de valores atípicos y proporciona un indicador más estable de tendencia.

La interpretación del coeficiente de Spearman conserva el intervalo entre -1 y 1 . Valores próximos a 1 indican asociación monótona creciente; valores cercanos a -1 , asociación monótona decreciente; valores alrededor de 0 reflejan ausencia de tendencia ordinal dominante. Sin embargo, su significado difiere del coeficiente de Pearson. Un valor elevado de Spearman garantiza coherencia en el ordenamiento, aunque no asegura que una recta describa adecuadamente la relación.

En el ejercicio profesional, el coeficiente de Spearman se aplica en diversas situaciones. En investigación social, cuando se analizan escalas tipo Likert, la naturaleza ordinal de los datos hace pertinente su utilización. En estudios clínicos, cuando las mediciones presentan asimetrías o distribuciones alejadas de la normalidad, la correlación por rangos permite evaluar asociación sin imponer estructura paramétrica estricta.

Ortiz Pinilla y Ortiz Rico (2021) advierten que considerar a Spearman como una simple "alternativa" de Pearson puede conducir a interpretaciones imprecisas. Ambos coeficientes responden a preguntas distintas: uno examina proporcionalidad lineal; el otro, consistencia monótona. La elección debe

sustentarse en la naturaleza del fenómeno y en el tipo de relación que se pretende describir.

En síntesis, el coeficiente de correlación de Spearman constituye una herramienta adecuada para evaluar asociación monotónica cuando la linealidad no se verifica o cuando los datos poseen carácter ordinal. Su fortaleza reside en la estabilidad frente a valores extremos y en la flexibilidad interpretativa respecto de la forma funcional de la relación. Comprender esta diferencia conceptual permite aplicar la medida de manera coherente con el diseño del análisis y con los objetivos profesionales del estudio.

Interpretación, supuestos y limitaciones de las medidas de correlación

El análisis de correlación constituye una herramienta potente para describir la relación entre variables cuantitativas. Sin embargo, su interpretación exige una comprensión precisa de los supuestos que sustentan cada coeficiente y de los límites metodológicos que condicionan su alcance. Tanto el coeficiente

de Pearson como el de Spearman proporcionan información sobre la estructura de asociación, pero responden a preguntas diferentes y presentan condiciones específicas de aplicación.

En el caso del coeficiente de Pearson, el supuesto central consiste en la existencia de una relación lineal entre las variables. Ortiz Pinilla y Ortiz Rico (2021) señalan que «la linealidad es un requisito conceptual para garantizar que el coeficiente de Pearson sea plenamente válido e interpretable». Esta afirmación delimita claramente su campo de aplicación: cuando la proporcionalidad de los cambios no se verifica, la magnitud del coeficiente puede inducir interpretaciones parciales.

Además de la linealidad, el análisis paramétrico tradicional asocia la inferencia del coeficiente de Pearson con el supuesto de normalidad bivariada. Martínez Ortega et al. (2009) describen el procedimiento de contraste estadístico bajo dicha condición. Si bien en la práctica profesional se utilizan aproximaciones robustas, la coherencia entre modelo y datos fortalece la interpretación.

Por su parte, el coeficiente de Spearman exige como condición conceptual la existencia de una relación monotónica. Camacho-Sandoval (2008) destaca que este coeficiente se basa en rangos y resulta apropiado cuando no se cumple el supuesto de normalidad. En este caso, la estabilidad frente a valores extremos

constituye una ventaja operativa, aunque la medida no describe proporcionalidad de cambios.

La interpretación de ambos coeficientes puede organizarse comparativamente en la siguiente tabla.

Tabla 10. Comparación interpretativa entre Pearson y Spearman

Dimensión analizada	Pearson (r)	Spearman (ρ)
Tipo de relación evaluada	Lineal	Monotónica
Datos utilizados	Valores originales	Rangos
Sensibilidad a valores extremos	Alta	Moderada
Supuesto conceptual	Proporcionalidad en los cambios	Consistencia en el ordenamiento
Alcance interpretativo	Intensidad y dirección lineal	Intensidad y dirección monotónica

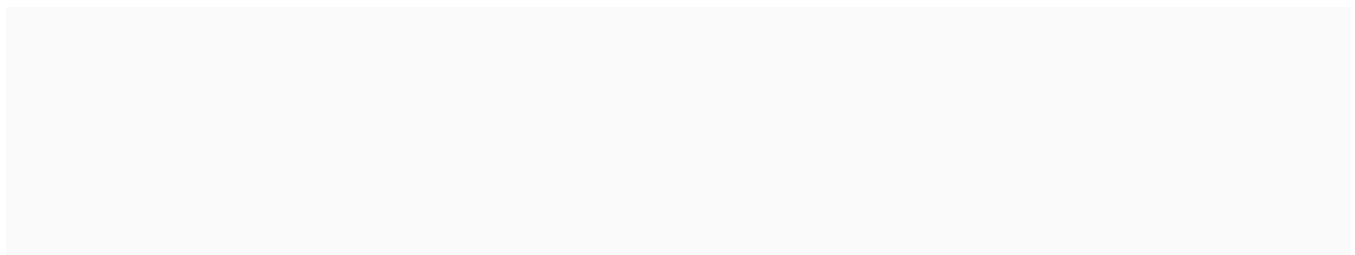
Fuente: elaboración propia

Esta comparación evidencia que los coeficientes no son intercambiables de manera automática. La elección debe responder a la estructura de los datos y al objetivo analítico. Un valor elevado de Pearson indica proximidad a una recta; un valor elevado de Spearman indica coherencia en el orden de las observaciones.

Entre las limitaciones más relevantes se encuentra la imposibilidad de establecer causalidad. La correlación describe comportamiento conjunto, pero no identifica mecanismos explicativos. Asimismo, un coeficiente alto no garantiza adecuación del modelo en todos los intervalos de la variable independiente. Ortiz Pinilla y Ortiz Rico (2021) muestran que pueden existir relaciones no lineales con valores elevados de correlación, lo cual exige análisis gráfico complementario.

Desde una perspectiva metodológica, el proceso interpretativo puede representarse mediante el siguiente esquema conceptual:

Figura 2. Proceso de interpretación de un coeficiente de correlación





Made with [Mapkin](#)

Fuente: elaboración propia

Este esquema destaca que el coeficiente constituye una etapa dentro de un proceso analítico más amplio. La interpretación rigurosa requiere integrar evidencia numérica, análisis gráfico y fundamentación conceptual.

En síntesis, las medidas de correlación permiten cuantificar intensidad y dirección de asociación entre variables cuantitativas. Su correcta aplicación depende de la coherencia entre supuestos, naturaleza de los datos y objetivo del estudio. Reconocer sus límites metodológicos fortalece la calidad interpretativa y

previene conclusiones que excedan el alcance descriptivo del análisis.

CONTINUAR

Referencias

Camacho-Sandoval, J. (2008). Asociación entre variables: correlación no paramétrica. *Acta Médica Costarricense*, 50(3), 144–146.

Martínez Ortega, R. M., Tuya Pendás, L. C., Martínez Ortega, M., Pérez Abreu, A., & Cánovas, A. M. (2009). El coeficiente de correlación de los rangos de Spearman: caracterización. *Revista Habanera de Ciencias Médicas*, 8(2).

Moral Peláez, I. (s.f.). Medidas de asociación. En *Capítulo 13* (pp. 185–194).

Ortiz Pinilla, J., & Ortiz Rico, A. F. (2021). ¿Pearson y Spearman, coeficientes intercambiables? *Comunicaciones en Estadística*, 14(1), 53–63.

CONTINUAR