
La Relevancia de los Supuestos Estadísticos en la Investigación Cuantitativa: Guía para la Validez Metodológica

Este documento explora la naturaleza y la importancia de los supuestos estadísticos en la investigación cuantitativa, destacando su rol fundamental en la validez y fiabilidad de los resultados. Se detallan los supuestos más comunes, como normalidad, homogeneidad de varianzas, independencia, linealidad y ausencia de valores extremos, y se discute cómo su cumplimiento o incumplimiento afecta la elección y la interpretación de las pruebas estadísticas. Además, se propone una estrategia de evaluación multinivel y se ofrecen soluciones prácticas para abordar las violaciones de supuestos, enfatizando la necesidad de un enfoque crítico y reflexivo en el análisis de datos.

Alexis E. Rebolledo Carreño

Metodólogo

ORCID: [0000-0003-2728-5723](https://orcid.org/0000-0003-2728-5723)

19 de mayo de 2026

Introducción: La Fundamentación de la Validez en la Investigación Cuantitativa

La investigación cuantitativa, en su búsqueda de objetividad y generalización, se apoya en el rigor metodológico y estadístico. Un aspecto crítico, a menudo subestimado o malinterpretado, es el papel de los supuestos estadísticos. Estos supuestos no son meros formalismos burocráticos, sino condiciones inherentes a las pruebas y modelos estadísticos que garantizan la validez y la interpretabilidad de sus resultados (Field, 2018). Su correcta comprensión y evaluación son pilares esenciales para la construcción de conocimiento científico robusto y para evitar inferencias erróneas que puedan comprometer la credibilidad de un estudio.

La adherencia a estos principios estadísticos es tan crucial como la formulación de una pregunta de investigación clara o la selección adecuada de una muestra. Como señalan Hernández-Sampieri, Fernández-Collado y Baptista-Lucio (2014), la elección de una técnica de análisis de datos debe estar intrínsecamente ligada a las características de los datos y a los requisitos subyacentes del método seleccionado. Ignorar estos supuestos puede llevar a la aplicación de pruebas inadecuadas, lo que a su vez genera conclusiones inválidas o, en el mejor de los casos, menos fiables. Este documento tiene como objetivo proporcionar una guía concisa pero completa sobre la naturaleza, la evaluación y el manejo de los supuestos estadísticos en el contexto de tesis e investigaciones cuantitativas.

¿Qué son los Supuestos Estadísticos y por qué son Cruciales?

Los supuestos estadísticos pueden definirse como las condiciones o propiedades que se espera que cumplan los datos de una muestra para que una determinada prueba estadística o modelo inferencial sea válido y sus resultados puedan ser interpretados correctamente. Estas condiciones no son arbitrarias, sino que se derivan de la teoría matemática que sustenta cada procedimiento estadístico (Kerlinger & Lee, 2002). Por ejemplo, muchas pruebas paramétricas

se basan en la teoría de la probabilidad y requieren que los datos sigan ciertas distribuciones o que las varianzas entre grupos sean similares para que los cálculos de probabilidad asociados a los valores p sean precisos.

La importancia de estos supuestos radica en que constituyen la base metodológica sobre la cual se asienta la validez de las conclusiones obtenidas. Si un supuesto fundamental no se cumple, la prueba estadística empleada puede perder su capacidad para generar inferencias precisas y generalizables a la población de interés. Sin embargo, es fundamental comprender que los supuestos rara vez se cumplen de manera 'perfecta' en datos reales. La cuestión central no es si los datos son inmaculados, sino si las desviaciones de estos supuestos son lo suficientemente significativas como para afectar sustancialmente la robustez de los resultados y la validez de las conclusiones (Field, 2018). La evaluación de los supuestos es, por tanto, un acto de juicio crítico y no una aplicación mecánica de reglas.

Principales Supuestos Estadísticos en la Investigación Cuantitativa

Aunque la lista de supuestos puede variar según la complejidad del análisis estadístico, existen cinco supuestos fundamentales que son recurrentes en la mayoría de las investigaciones cuantitativas y que merecen una atención especial. La comprensión de cada uno de ellos es vital para seleccionar y aplicar correctamente las herramientas estadísticas. Estos supuestos son el pilar de gran parte de los análisis paramétricos y su evaluación es un paso ineludible en el proceso de análisis de datos (Creswell & Creswell, 2018).

Normalidad

El supuesto de normalidad postula que los datos de una variable, o los residuos de un modelo, siguen una distribución aproximadamente simétrica, similar a la curva de campana (distribución normal). Este supuesto es particularmente relevante en pruebas paramétricas que comparan medias, como la prueba t de Student o el análisis de varianza (ANOVA), y en modelos de regresión lineal (Field, 2018). Aunque la normalidad estricta es rara en la práctica, muchas pruebas paramétricas son robustas a pequeñas desviaciones, especialmente con

tamaños de muestra grandes debido al Teorema del Límite Central.

Homogeneidad de Varianzas (Homocedasticidad)

La homogeneidad de varianzas, o homocedasticidad, implica que la dispersión de los datos es similar entre los diferentes grupos o niveles de una variable independiente. Este supuesto es crucial en pruebas que comparan las medias de dos o más grupos, como el ANOVA, donde se asume que la variabilidad dentro de cada grupo es aproximadamente igual (Hernández-Sampieri et al., 2014). La violación de este supuesto, conocida como heterocedasticidad, puede llevar a estimaciones de error estándar sesgadas y, consecuentemente, a valores p incorrectos, afectando la validez de las inferencias estadísticas.

Independencia de las Observaciones

El supuesto de independencia es quizás uno de los más fundamentales y abarcadores, aplicable a prácticamente todos los análisis estadísticos inferenciales. Establece que cada observación o dato en la muestra es independiente de las demás, es decir, que la medición de un individuo no influye ni está influenciada por la medición de otro (Babbie, 2021). La violación de este supuesto, como ocurre en diseños con medidas repetidas o datos anidados, requiere el uso de modelos estadísticos más complejos, como los modelos lineales jerárquicos, para evitar errores en la estimación de los parámetros y sus errores estándar.

Linealidad

El supuesto de linealidad es esencial en análisis de correlación y, sobre todo, en modelos de regresión lineal. Implica que la relación entre la variable dependiente y las variables independientes puede ser adecuadamente aproximada por una línea recta (Field, 2018). Si la relación subyacente es curvilínea y se aplica un modelo lineal, el modelo no capturará adecuadamente la verdadera relación entre las variables, lo que resultará en un ajuste deficiente y predicciones sesgadas. La evaluación visual a través de diagramas de dispersión es una herramienta clave para verificar este supuesto.

Ausencia de Valores Extremos (Outliers)

La ausencia de valores extremos, o outliers, es un supuesto importante, especialmente en modelos de regresión y otros análisis que son sensibles a observaciones inusuales. Los valores extremos son datos que se desvían significativamente del patrón general de otras observaciones y pueden tener una influencia desproporcionada en los resultados estadísticos, alterando las medias, las desviaciones estándar y los coeficientes de regresión (Patton, 2015). Su identificación y manejo adecuado son cruciales para evitar distorsiones en los hallazgos y para asegurar que el modelo refleje la tendencia general de los datos.

La Selectividad de los Supuestos: No Todos Aplican a Todos los Estudios

Es crucial entender que no todos los supuestos estadísticos son aplicables a todas las investigaciones o a todos los análisis. La relevancia de un supuesto particular está intrínsecamente ligada a la naturaleza específica del estudio, el tipo de análisis estadístico elegido, el diseño de investigación implementado y los objetivos del estudio (Creswell & Plano Clark, 2018). Una evaluación reflexiva de estos factores es indispensable antes de proceder con el análisis de datos, evitando así la aplicación indiscriminada de pruebas de supuestos que podrían ser irrelevantes o engañosas.

En primer lugar, el tipo de análisis estadístico es el principal determinante de los supuestos exigibles. Las pruebas paramétricas, por ejemplo, suelen requerir supuestos más estrictos (como normalidad y homocedasticidad) que sus contrapartes no paramétricas. En segundo lugar, el diseño de investigación influye directamente en supuestos como la independencia de las observaciones; un diseño de medidas repetidas, por ejemplo, viola inherentemente la independencia y requiere modelos estadísticos específicos. Finalmente, el objetivo del estudio (describir, relacionar, predecir) guía la rigurosidad necesaria en la evaluación de los supuestos, ya que un estudio exploratorio podría ser más indulgente que uno predictivo o causal (Maxwell, 2013).

Estrategias para Evaluar los Supuestos Estadísticos

La evaluación de los supuestos estadísticos debe ser un proceso reflexivo y no una mera aplicación automática de pruebas. Se recomienda un enfoque multinivel que combine la revisión conceptual, el análisis descriptivo visual y, solo entonces, el uso de pruebas formales (Miles, Huberman & Saldaña, 2014). Este enfoque jerárquico permite una comprensión más profunda de los datos y evita la obsesión por el cumplimiento perfecto de los supuestos, que a menudo es inalcanzable en la práctica.

Nivel 1: Revisión Conceptual y Lógica

Antes de cualquier análisis gráfico o formal, es fundamental realizar una revisión conceptual de los datos y el diseño del estudio. Esto implica formular preguntas básicas como: ¿Qué tipo de variables se están analizando? ¿Cómo se recolectaron los datos? ¿El diseño de la investigación garantiza la independencia de las observaciones o, por el contrario, introduce dependencia inherente? (Creswell, 2014). Esta etapa inicial permite identificar posibles violaciones de supuestos desde la concepción del estudio y orientar la elección de las pruebas estadísticas más adecuadas. Por ejemplo, si se sabe que los datos provienen de una escala ordinal, aplicar una prueba paramétrica que asume datos de intervalo/razón requerirá una justificación muy sólida o la elección de una alternativa no paramétrica.

Nivel 2: Revisión Descriptiva y Visual

La exploración visual de los datos es una herramienta poderosa y a menudo más informativa que las pruebas formales. Histogramas, diagramas de caja (box plots), diagramas de dispersión y gráficos Q-Q pueden revelar patrones, asimetrías, valores extremos y relaciones no lineales de manera intuitiva (Field, 2018). 'Leer los datos' a través de estas representaciones gráficas permite al investigador desarrollar una comprensión más profunda de la estructura de sus datos y detectar desviaciones de los supuestos que podrían ser significativas o, por el contrario, triviales. Esta etapa es crucial para formar un juicio informado antes de recurrir a pruebas estadísticas más complejas.

Nivel 3: Pruebas Formales de Supuestos

Las pruebas formales de supuestos, como la prueba de Shapiro-Wilk o Kolmogorov-Smirnov para normalidad, o la prueba de Levene para homogeneidad de varianzas, deben utilizarse como apoyo al juicio del investigador, no como su reemplazo. Es importante recordar que en muestras grandes, incluso desviaciones mínimas e irrelevantes desde un punto de vista práctico pueden resultar estadísticamente significativas (Field, 2018). Por lo tanto, un valor p significativo en una prueba de supuestos no siempre indica una violación grave que invalide el análisis. La interpretación de estas pruebas debe hacerse en conjunto con la revisión conceptual y descriptiva, considerando el tamaño de la muestra y la robustez de la prueba principal a las violaciones de supuestos.

Manejo de las Violaciones de Supuestos: Decisiones Informadas

Cuando un supuesto estadístico no se cumple, no significa automáticamente que el estudio sea inválido o que deba abandonarse el análisis. La clave reside en tomar decisiones informadas y metodológicamente justificadas. El principio fundamental es que la estrategia analítica debe dialogar con los datos reales, en lugar de forzar los datos para que se ajusten a una prueba predeterminada (Tashakkori & Teddlie, 2010). Existen diversas estrategias para abordar las violaciones de supuestos, cada una con sus propias implicaciones.

- ****Justificar el uso del análisis si la desviación es leve:**** Muchas pruebas paramétricas son robustas a violaciones moderadas de sus supuestos, especialmente con muestras grandes. Si la desviación es menor y no parece afectar sustancialmente la distribución o la relación, el investigador puede justificar el uso de la prueba original, haciendo explícitas las limitaciones en la interpretación de los resultados (Field, 2018).
- ****Utilizar pruebas o modelos alternativos más robustos:**** Algunos procedimientos estadísticos están diseñados para ser menos sensibles a las violaciones de ciertos supuestos. Por ejemplo, en lugar de un ANOVA estándar, se podría usar un ANOVA de Welch si se viola la homocedasticidad, o un bootstrap para estimar errores estándar más robustos.
- ****Transformar variables cuando tenga sentido:**** Las transformaciones matemáticas

(logarítmica, raíz cuadrada, inversa, etc.) pueden ayudar a normalizar distribuciones asimétricas, estabilizar varianzas o linealizar relaciones. Sin embargo, las transformaciones deben aplicarse con precaución, ya que pueden alterar la interpretabilidad de los resultados y deben justificarse teóricamente (Hair et al., 2014).

- ****Optar por análisis no paramétricos:**** Cuando las violaciones de supuestos son severas y no pueden ser mitigadas por otras estrategias, las pruebas no paramétricas (como la U de Mann-Whitney en lugar de la t de Student, o la H de Kruskal-Wallis en lugar del ANOVA) son una alternativa viable. Estas pruebas no requieren supuestos sobre la distribución de los datos, aunque pueden tener menor potencia estadística (Siegel & Castellan, 1988).
- ****Ajustar la interpretación de los resultados:**** Independientemente de la estrategia elegida, es fundamental ajustar la interpretación de los resultados y reconocer las limitaciones impuestas por las violaciones de los supuestos. Esto implica ser transparente en la sección de metodología y discusión, explicando cómo se abordaron las violaciones y cómo podrían afectar la generalizabilidad o la precisión de las conclusiones.

Errores Frecuentes en Tesis Cuantitativas y su Impacto

La calidad de una tesis cuantitativa a menudo se ve más mejorada por la evitación de errores metodológicos básicos que por la aplicación de técnicas estadísticas excesivamente sofisticadas. Una comprensión sólida de los supuestos y un enfoque crítico en el análisis de datos son más valiosos que la mera habilidad para ejecutar comandos de software (Bryman, 2016). Los errores comunes en el manejo de supuestos pueden minar la validez interna y externa de un estudio, llevando a conclusiones erróneas o no justificadas.

- ****Elegir la prueba estadística antes de explorar los datos:**** Este es un error fundamental. La exploración de datos y la evaluación de supuestos deben preceder a la selección de la prueba, ya que las características de los datos dictan la técnica apropiada.

- ****Confiar exclusivamente en pruebas de supuestos sin análisis descriptivo:**** Como se mencionó, las pruebas formales pueden ser engañosas, especialmente en muestras grandes. El análisis visual y el juicio del investigador son insustituibles.
- ****Ignorar la dependencia entre observaciones:**** Este error es crítico, ya que viola el supuesto de independencia y puede llevar a errores estándar subestimados y, por ende, a conclusiones falsamente significativas. Es común en diseños longitudinales o con datos anidados.
- ****Tratar escalas ordinales como si fueran variables continuas sin justificación:**** Aunque a veces se hace por conveniencia, aplicar pruebas paramétricas a datos ordinales sin una justificación sólida (e.g., escala Likert con muchos puntos y distribución cercana a la normal) puede ser inapropiado y llevar a interpretaciones incorrectas (Vasilachis de Gialdino, 2006).
- ****No explicitar las decisiones metodológicas en el apartado correspondiente:**** Toda decisión sobre el manejo de supuestos, incluyendo las violaciones y las estrategias adoptadas, debe ser documentada de manera transparente en la sección de metodología de la tesis. Esto demuestra rigor y permite la replicabilidad y evaluación crítica del estudio.

Conclusión: La Estadística como Cimiento de la Validez

La estadística en la investigación cuantitativa no es un proceso que comienza y termina en el software. Es una disciplina que exige un pensamiento crítico, una comprensión profunda de los datos y un compromiso con la validez metodológica en cada etapa. La revisión y el manejo de los supuestos estadísticos son componentes integrales de este proceso, no meros trámites técnicos (Guba & Lincoln, 1994). Al cuidar la validez de los supuestos, los investigadores no solo aseguran la precisión de sus análisis, sino que también fortalecen la credibilidad y el impacto de sus afirmaciones científicas.

Una metodología sólida, cimentada en una evaluación rigurosa de los supuestos estadísticos y en una toma de decisiones informada sobre el análisis de datos, es el pilar de toda

investigación cuantitativa de calidad. La estadística, en su esencia, comienza con la lectura crítica de los datos y con la reflexión sobre cómo estos datos pueden ser analizados de la manera más honesta y precisa posible, garantizando así que las conclusiones obtenidas sean un reflejo fiel de la realidad estudiada y contribuyan de manera significativa al cuerpo de conocimiento existente.

Referencias bibliográficas

- Babbie, E. R. (2021). **The practice of social research** (15th ed.). Cengage Learning.
- Bryman, A. (2016). **Social research methods** (5th ed.). Oxford University Press.
- Creswell, J. W. (2014). **Research design: Qualitative, quantitative, and mixed methods approaches** (4th ed.). SAGE Publications.
- Creswell, J. W., & Creswell, J. D. (2018). **Research design: Qualitative, quantitative, and mixed methods approaches** (5th ed.). SAGE Publications.
- Creswell, J. W., & Plano Clark, V. L. (2018). **Designing and conducting mixed methods research** (3rd ed.). SAGE Publications.
- Field, A. (2018). **Discovering statistics using IBM SPSS Statistics** (5th ed.). SAGE Publications.
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. En N. K. Denzin & Y. S. Lincoln (Eds.), **Handbook of qualitative research** (pp. 105-117). SAGE Publications.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). **Multivariate data analysis** (7th ed.). Pearson.
- Hernández-Sampieri, R., Fernández-Collado, C., & Baptista-Lucio, P. (2014). **Metodología de la investigación** (6ta ed.). McGraw-Hill Education.
- Kerlinger, F. N., & Lee, H. B. (2002). **Foundations of behavioral research** (4th ed.). Wadsworth Thomson Learning.
- Maxwell, J. A. (2013). **Qualitative research design: An interactive approach** (3rd ed.). SAGE Publications.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). **Qualitative data analysis: A methods sourcebook** (3rd ed.). SAGE Publications.

Patton, M. Q. (2015). **Qualitative research & evaluation methods: Integrating theory and practice** (4th ed.). SAGE Publications.

Siegel, S., & Castellan, N. J. (1988). **Nonparametric statistics for the behavioral sciences** (2nd ed.). McGraw-Hill.

Tashakkori, A., & Teddlie, C. (2010). **Handbook of mixed methods in social and behavioral research** (2nd ed.). SAGE Publications.

Vasilachis de Gialdino, I. (2006). **Estrategias de investigación cualitativa**. Gedisa.

Cómo citar este documento

Alexis E. Rebolledo Carreño (2026). La Relevancia de los Supuestos Estadísticos en la Investigación Cuantitativa: Guía para la Validez Metodológica. Documento de trabajo metodológico. ORCID: <https://orcid.org/0000-0003-2728-5723>