**METODOLOGÍA PARA EL ANÁLISIS DE ENCUESTA DE CAPACIDADES DE INNOVACIÓN EMPRESARIALES**

1. **Introducción**

Al aplicar herramientas estadísticas, es fundamental establecer la naturaleza de las variables involucradas. Esta distinción resulta crucial porque no todas las variables pueden analizarse mediante los mismos métodos estadísticos. En particular, se debe determinar si una variable es cuantitativa o cualitativa, dado que esta clasificación condiciona las técnicas apropiadas para su análisis.

Una variable cualitativa representa una característica o atributo que clasifica a cada unidad de análisis en una categoría. En el contexto de este estudio, todas las variables recolectadas mediante el instrumento de medición corresponden a respuestas en escalas tipo Likert, con opciones como “Nunca”, “Raramente”, “A veces”, “Muy a menudo” y “Siempre”.

Las variables cualitativas se dividen habitualmente en nominales y ordinales. Las nominales no tienen un orden inherente entre sus categorías (por ejemplo, el color de ojos: azul, marrón, verde). En cambio, las ordinales sí presentan una jerarquía clara entre categorías, como ocurre con escalas de satisfacción o desempeño (por ejemplo: Malo, Regular, Bueno, Muy Bueno, Excelente). Aunque estas escalas asignan valores numéricos a las categorías, los números son meramente indicativos del orden y no implican intervalos iguales entre respuestas.

A pesar de que en la práctica es común codificar las respuestas ordinales con valores numéricos consecutivos (por ejemplo, 1 a 5), esta codificación no convierte a la variable en cuantitativa. Así, calcular promedios o aplicar métodos estadísticos convencionales como regresión lineal o análisis de correlación puede conducir a interpretaciones erróneas, ya que estos métodos asumen intervalos constantes y una escala continua.

En este contexto, resulta necesario utilizar modelos diseñados explícitamente para respuestas categóricas ordenadas, como los modelos de la Teoría de Respuesta al Ítem (IRT). Esta familia de modelos permite estimar una variable latente no observable —como la propensión de una empresa a adoptar prácticas asociadas al sensing, seizing o transformation— a partir de patrones de respuesta a ítems observables con escalas ordinales.

En particular, el Modelo de Respuesta Graduada (GRM), propuesto por Samejima (1969), es uno de los más adecuados para este tipo de datos. El GRM permite modelar la probabilidad de que una entidad seleccione una determinada categoría (o una superior) en función de su nivel en la variable latente y de las características del ítem (como su dificultad y capacidad de discriminación).

Este modelo ha sido ampliamente adoptado no solo en psicometría, sino también en áreas como educación, salud pública, análisis organizacional y marketing (Embretson & Reise, 2000; De Ayala, 2009). En nuestro caso, se utiliza el GRM para estimar una variable continua latente que resume la propensión de cada empresa a desarrollar prácticas específicas en una dimensión dada. Esta transformación permite aplicar posteriormente métodos estadísticos convencionales, como pruebas estadísticas de comparación de medias y algoritmos supervisados o no supervisados.

En resumen, el uso del GRM constituye una alternativa teórica y empíricamente sólida frente a la práctica incorrecta de promediar ítems ordinales o aplicar directamente técnicas multivariadas a variables categóricas. Al reconocer la naturaleza ordinal de los datos y modelar explícitamente la estructura latente subyacente, el GRM produce mediciones más precisas, comparables y psicométricamente válidas.

1. **Metodología**

El GRM asume que cada ítem tiene un conjunto de categorías de respuesta ordenadas y que la probabilidad de responder en una categoría determinada o superior depende de un rasgo latente θ. Para un ítem con categorías ordenadas, el modelo especifica umbrales y un parámetro de discriminación .

**Formulación Matemática:**

Para el ítem , la probabilidad de que una empresa con rasgo latente θ responda en la categoría o superior se expresa como:

La probabilidad de responder exactamente en la categoría es:

**Componentes del Modelo:**

* θ: variable latente — representa el nivel subyacente del rasgo que se desea medir (por ejemplo, la capacidad de *sensing* de una empresa).
* : parámetro de discriminación del ítem — indica qué tan bien el ítem distingue entre entidades con diferentes niveles de θ.
* : umbral de dificultad para el ítem y la categoría — representa el nivel de θ requerido para que una entidad tenga al menos una probabilidad del 50 % de seleccionar la categoría o una superior.

Este enfoque permite usar ítems con diferentes niveles de discriminación y dificultad, generando una medición más flexible e informativa.

Con el fin de facilitar la comprensión de los parámetros de discriminación y dificultad, a continuación, se plantean dos ejemplos representativos:

Considérese una encuesta aplicada a empresas que contiene las siguientes preguntas:

* **Ítem A**: "¿Con qué frecuencia su empresa revisa las tendencias del mercado?"
* **Ítem B**: "¿Con qué frecuencia su empresa aplica modelos predictivos avanzados de inteligencia artificial para anticiparse a la demanda?"

Ambos ítems utilizan una escala ordinal de cinco puntos, donde 1 representa "Nunca" y 5 representa "Siempre".

Desde el punto de vista del parámetro de dificultad:

* El Ítem A tiende a recibir respuestas afirmativas (puntuaciones altas) por parte de la mayoría de las empresas, incluso aquellas con un nivel incipiente en capacidades de sensing. Esto sugiere que los umbrales asociados a este ítem se ubican en valores bajos del continuo , por lo que se clasifica como un ítem de baja dificultad.
* El Ítem B, en contraste, exige un mayor nivel de desarrollo organizacional para ser respondido afirmativamente. Es esperable que solo las empresas más avanzadas en prácticas analíticas puedan seleccionar categorías superiores en este ítem, lo que implica umbrales ubicados en niveles altos de . Por tanto, se clasifica como un ítem de alta dificultad.

Respecto al parámetro de discriminación:

* Si el Ítem B permite diferenciar de manera precisa a las empresas con altos niveles de aquellas con bajos niveles —es decir, si pequeñas variaciones en implican cambios significativos en la probabilidad de seleccionar categorías altas—, entonces se le atribuye una alta discriminación.
* Por el contrario, si dicha probabilidad cambia de forma paulatina y empresas con distintos niveles de tienden a responder de manera similar, el ítem presenta una baja capacidad de discriminación.

Este ejemplo ilustra cómo el GRM permite capturar no solo la "exigencia" asociada a cada ítem (su dificultad), sino también su capacidad para aportar información útil en la medición de la variable latente (su discriminación).

1. **Nota metodológica**

Una de las principales limitaciones metodológicas del presente estudio radica en el tamaño muestral empleado para la estimación del modelo de Respuesta Graduada (GRM, por sus siglas en inglés). Con una muestra de 56 participantes, se reconoce que las estimaciones obtenidas pueden estar sujetas a restricciones inherentes al ajuste de modelos IRT (Item Response Theory), en particular cuando el número de observaciones es limitado. La literatura especializada ha documentado que tamaños muestrales pequeños pueden comprometer la estabilidad y precisión de los parámetros estimados en modelos IRT, incrementando la variabilidad de los errores estándar, afectando la convergencia del algoritmo de estimación y dificultando la generalización de los resultados (de Ayala, 2009; Baker & Kim, 2017).

No obstante, con el fin de mitigar y evaluar este posible efecto, se realizó un análisis complementario de los errores estándar asociados a las puntuaciones latentes obtenidas en cada una de las dimensiones estimadas por el GRM (Sensing, Seizing y Transformation). Los resultados mostraron que los errores estándar promedio se mantuvieron en niveles bajos (entre 0.120 y 0.132), con intervalos de confianza del 95% relativamente estrechos. En particular, la dimensión *Transformation* presentó el menor coeficiente de variación del error estándar (8.2%), lo que sugiere una alta estabilidad en la estimación de los puntajes latentes. Aunque en la dimensión *Seizing* se identificó una mayor dispersión relativa (coeficiente de variación del 31.1%), los niveles absolutos de error estándar no resultaron preocupantes en la mayoría de los casos.

En conjunto, si bien el tamaño muestral constituye una limitación relevante, los indicadores analizados respaldan una aceptable precisión de las estimaciones obtenidas a través del modelo GRM. Se recomienda, no obstante, replicar este análisis con muestras más amplias que permitan validar la robustez de los hallazgos y reducir potenciales sesgos derivados de la limitada cantidad de observaciones.

1. **Resultados**

En esta sección se presentarán distintos análisis bivariados entre las variables latentes calculadas para cada una de las dimensiones —sensing, seizing y transformation— y las variables de caracterización de las empresas (tamaño de la empresa, ventas del último año, industria a la que pertenece la compañía, etc.), lo anterior, con el fin de obtener algunas inferencias.

* 1. **Análisis Univariados**

**Tamaño de la compañía**

Se presenta a continuación la distribución porcentual de los tamaños de empresa en la muestra analizada.

Gráfico, Gráfico circular

El contenido generado por IA puede ser incorrecto.

En la gráfica anterior se observa que, dentro de la muestra recolectada, la mayoría de las empresas corresponde al grupo de tamaño mediano, representando aproximadamente el 54 % (n = 15). Por su parte, el 18 % (n = 5) corresponde a empresas pequeñas, mientras que tanto las empresas grandes como las microempresas representan cada una el 14 % (n = 4).

* 1. **Análisis bivariados para sensing**

**4.2.1 Tamaño de la compañía**

Para llevar a cabo este análisis, se plantea la siguiente pregunta central, que orienta el desarrollo del estudio: ¿existen diferencias estadísticamente significativas en las variables latentes de *sensing* entre empresas grandes, medianas, pequeñas y microempresas? En otros términos, **¿la propensión al sensing podría variar en función del tamaño de la empresa?**

Esta pregunta puede abordarse mediante un ejercicio estadístico de comparación de medias de las variables latentes de *sensing* entre los distintos grupos empresariales definidos según su tamaño.

Antes de responder a la pregunta central formulada previamente, es necesario presentar los resultados de una serie de pruebas estadísticas preliminares. En particular, se evalúan dos supuestos fundamentales para la comparación de medias: normalidad (es decir, que las variables latentes de *sensing* sigan una distribución normal dentro de cada grupo) y homocedasticidad (igualdad de varianzas entre los grupos).

Para la verificación del supuesto de normalidad se emplea la prueba de Shapiro-Wilk. A continuación, se reportan los resultados de la prueba de Shapiro-Wilk (SW) aplicada a cada grupo de empresas, con el objetivo de evaluar el cumplimiento del supuesto de normalidad en las variables latentes de *sensing*:

**Tabla 1. Resultados de prueba de normalidad para análisis bivariado entre variable latente *sensing* y tamaño de la compañía**

| **GRUPO** | **ESTADÍSTICO** | **P-VALOR** | **DECISIÓN** |
| --- | --- | --- | --- |
| **Grande** | 0,9166 | 0,5178 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |
| **Mediana** | 0,9561 | 0,6249 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |
| **Pequeña** | 0,9295 | 0,5929 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |
| **Micro** | 0,9628 | 0,7965 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |

**Fuente:** Elaboración propia.

A partir de los resultados reportados en la tabla anterior, es posible observar que, en cada uno de los grupos, el valor p asociado a la prueba de Shapiro-Wilk supera el umbral de significancia convencional (p > 0.05), por lo que no se rechaza la hipótesis nula. Esto indica que los datos de las variables latentes de *sensing* siguen una distribución normal en todos los grupos analizados.

Posteriormente, para evaluar el supuesto de homocedasticidad —es decir, la igualdad de varianzas entre los grupos definidos por tamaño de empresa— se emplea la prueba de Levene (en adelante, LV).

A continuación, se reportan los resultados de la prueba de Levene (LV) aplicada a los grupos definidos según el tamaño de empresa, con el propósito de evaluar la igualdad de varianzas en las variables latentes de *sensing*:

**Tabla 2. Resultados de prueba de homocedasticidad para análisis bivariado entre variable latente *sensing* y tamaño de la compañía**

|  |  |  |
| --- | --- | --- |
| **ESTADÍSTICO LV** | **P-VALOR** | **DECISIÓN** |
| 0,9367 | 0,4383 | No se rechaza la hipótesis nula. Las varianzas de los grupos poblaciones son iguales. |

**Fuente:** Elaboración propia.

A partir de los resultados reportados en la tabla anterior, es posible observar que, conforme al valor p obtenido, no se rechaza la hipótesis nula de homocedasticidad. Esto permite concluir que las varianzas entre los grupos de empresas, definidos según su tamaño, pueden considerarse iguales, cumpliéndose así el supuesto de igualdad de varianzas.

Una vez verificados los supuestos de normalidad y homocedasticidad, se procede con la aplicación del análisis de varianza (ANOVA) para determinar si existen diferencias significativas entre las medias de los grupos. A continuación, se reportan los resultados de la prueba de ANOVA aplicada a las variables latentes de *sensing*, con el fin de evaluar si existen diferencias estadísticamente significativas entre los grupos de empresas según su tamaño:

**Tabla 3. Resultados de prueba ANOVA para análisis bivariado entre variable latente sensing y tamaño de la compañía**

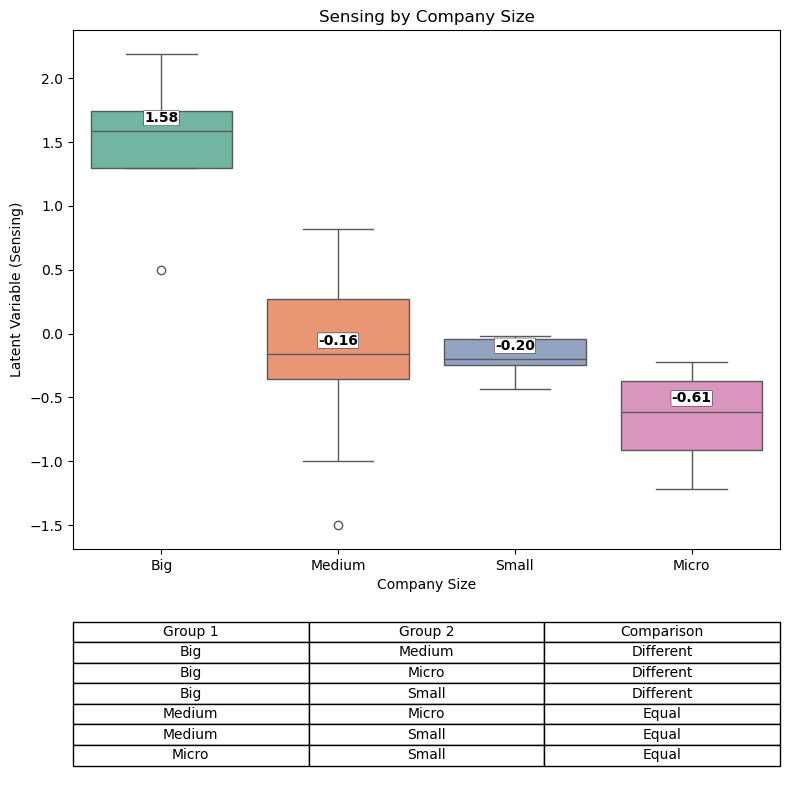
|  |  |  |
| --- | --- | --- |
| **ESTADÍSTICO F** | **P-VALOR** | **DECISIÓN** |
| 12,5298 | 0 | Se rechaza la hipótesis nula. Al menos una de las medias poblaciones difiere |

**Fuente:** Elaboración propia.

A partir de los resultados reportados en la tabla anterior, es posible observar que, conforme al valor p obtenido, se rechaza la hipótesis nula. Esto permite concluir que al menos una de las medias poblacionales de la variable latente de *sensing* difiere significativamente entre los grupos de empresas definidos según su tamaño.

Como siguiente paso, se aplica la prueba post-hoc de Tukey HSD, la cual permite identificar específicamente entre qué pares de grupos se presentan diferencias estadísticamente significativas en las medias de *sensing*.

**Gráfico 1. Resultados prueba post-hoc de Tukey HSD y visualización de distribución de índices por tamaño de compañía**



**Fuente:** Elaboración propia.

En la gráfica anterior se observa que las empresas grandes de la muestra recolectada presentan los valores más altos en la variable latente de *sensing*, lo que sugiere una mayor propensión al desarrollo de capacidades asociadas al *sensing* en comparación con los demás grupos: medianas, pequeñas y microempresas. No obstante, el objetivo central del análisis no es únicamente identificar si esta diferencia está presente en la muestra, sino determinar si dicha diferencia puede generalizarse a la población. Es decir, interesa establecer si, a nivel poblacional, las empresas grandes presentan una mayor propensión al *sensing* en comparación con los otros grupos definidos por tamaño.

La prueba post-hoc de Tukey permite responder a esta cuestión, ya que evalúa las diferencias entre las medias poblacionales de cada par de grupos, controlando el error tipo I asociado a las comparaciones múltiples. Según los resultados mostrados en la tabla anexa al gráfico, se evidencia que la población de empresas grandes presenta valores significativamente más altos en la variable latente de *sensing* en comparación con los grupos de empresas medianas, pequeñas y micro. En contraste, no se encontraron diferencias estadísticamente significativas entre las medias de las variables latentes de *sensing* de las empresas medianas, pequeñas y micro, lo que sugiere un comportamiento similar entre estos tres grupos en términos de propensión al *sensing*.

**4.2.2 Ventas del último año**

Para llevar a cabo este análisis, se plantea la siguiente pregunta central, que orienta el desarrollo del estudio: ¿existe una relación lineal entre la propensión al *sensing* y las ventas del último año? En otros términos, **¿la propensión al sensing podría afectar las ventas de la empresa en el último año?**

Inicialmente podemos revisar la correlación existente entre las variables latentes calculadas para *sensing* y las ventas de la empresa en el último año. Al calcular la correlación siguiendo el método de Pearson encontramos lo siguiente:

**Gráfico 2: Correlación entre variable latente *sensing* y ventas del último año**

Gráfico, Gráfico de dispersión

El contenido generado por IA puede ser incorrecto.

**Fuente:** Elaboración propia.

Como primera medida, encontramos que el coeficiente de correlación de Pearson es igual a 0,58 (r=0,58), lo que indica una relación positiva entre las ventas del último año de las empresas incluidas en la muestra y la variable *sensing*. Adicionalmente, el valor p asociado a esta prueba de Pearson confirma que esta correlación es significativa (generalizable a la población).

Para complementar estos hallazgos y obtener resultados más robustos, se estima posteriormente un modelo de regresión lineal simple, donde la variable dependiente corresponde a las ventas del último año y la variable independiente al índice de *sensing*. Con este modelo se busca inferir dos aspectos principales: (1) determinar si existe una relación estadísticamente significativa entre ambas variables, y (2) estimar el efecto del índice de *sensing* sobre las ventas del último año, es decir, identificar cómo varían las ventas en función de los cambios en dicho índice. A continuación, se presenta la tabla con los resultados del modelo:

**Tabla 4. Resultados de modelo de regresión lineal simple con variable dependiente de ventas en el último año y variable independiente de variable latente de *sensing***

Interfaz de usuario gráfica

El contenido generado por IA puede ser incorrecto.

**Fuente:** Elaboración propia.

Antes de presentar los resultados del modelo, es pertinente señalar que la estimación se realizó utilizando errores estándar robustos, debido a la presencia de heterocedasticidad en los residuales. Asimismo, se aplicó una transformación logarítmica a la variable dependiente (ventas del último año), con el fin de mitigar los desvíos respecto al supuesto de normalidad de los errores.

El uso de errores estándar robustos permitió obtener estimaciones consistentes y válidas para la inferencia estadística, a pesar de la violación del supuesto de homocedasticidad. Por otro lado, la transformación logarítmica de la variable dependiente contribuyó a corregir la no normalidad en la distribución de los residuales, fortaleciendo así la solidez del modelo estimado. Por último, se comprobó el cumplimiento de la no autocorrelación de los residuales.

Al analizar los resultados del modelo de regresión, se pueden destacar los siguientes aspectos:

* El valor-p asociado a la variable latente de *sensing* (Indice\_Sensing) es 0.013, lo cual es menor al umbral convencional de significancia del 5%. Esto indica que el coeficiente es estadísticamente significativo al 95% de confianza. En otras palabras, existe evidencia suficiente para afirmar que las actividades de innovación empresarial relacionadas con la dimensión *sensing* tienen un impacto significativo sobre las ventas del último año.
* Una vez establecida la significancia estadística de la variable latente de *sensing*, es relevante examinar la naturaleza y la magnitud de dicha relación. El coeficiente estimado para esta variable es de 1.0938, lo que permite extraer dos conclusiones claves:
  + La relación entre la variable latente de *sensing* y las ventas es positiva, es decir, un mayor valor en la propensión a realizar actividades de innovación en la dimensión *sensing* se asocia con mayores niveles de ventas en el último año. Esto sugiere que las empresas con mayor énfasis en detectar oportunidades, cambios en el entorno y necesidades del mercado tienden a obtener mejores resultados comerciales.
  + Dado que la variable dependiente se encuentra en escala logarítmica, el coeficiente puede interpretarse como un cambio porcentual aproximado. En este caso, un incremento de una unidad en la variable latente de sensing se asocia, en promedio, con un incremento del 109.4% en las ventas del último año.
* El valor del R-cuadrado ajustado es de 0.211, lo que indica que aproximadamente el 21.1% de la variabilidad en las ventas del último año puede ser explicada por el modelo a partir de la propensión de las empresas a desarrollar actividades de innovación en la dimensión *sensing*. Si bien este porcentaje es moderado, resulta razonable considerando que el modelo incluye una sola variable explicativa. En todo caso, es claro que otras variables pudieron haber influenciado las ventas en el último año y un modelo que las contenga podría tener mayor capacidad explicativa (mayor R cuadrado).

**4.2.3 Industria de la empresa**

Para llevar a cabo este análisis, se plantea la siguiente pregunta central, que orienta el desarrollo del estudio: ¿existen diferencias estadísticamente significativas en las variables latentes de *sensing* entre empresas de manufactura, comercio y servicios? En otros términos, **¿la propensión al *sensing* podría variar en función del sector de la industria en el que está la empresa?**

Esta pregunta puede abordarse mediante un ejercicio estadístico de comparación de medias de las variables latentes de *sensing* entre los distintos grupos empresariales definidos según la industria a la que pertenece.

Antes de responder a la pregunta central formulada previamente, es necesario presentar los resultados de una serie de pruebas estadísticas preliminares. En particular, se evalúan dos supuestos fundamentales para la comparación de medias: normalidad (es decir, que las variables latentes de *sensing* sigan una distribución normal dentro de cada grupo) y homocedasticidad (igualdad de varianzas entre los grupos).

Para la verificación del supuesto de normalidad se emplea la prueba de Shapiro-Wilk. A continuación, se reportan los resultados de la prueba de Shapiro-Wilk (SW) aplicada a cada grupo de empresas, con el objetivo de evaluar el cumplimiento del supuesto de normalidad en las variables latentes de *sensing*:

**Tabla 5. Resultados de prueba de normalidad para análisis bivariado entre variable latente *sensing* y sector de la industria**

|  |  |  |  |
| --- | --- | --- | --- |
| **GRUPO** | **ESTADÍSTICO** | **P-VALOR** | **DECISIÓN** |
| **Manufactura** | 0,9226 | 0,2399 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |
| **Servicios** | 0,9524 | 0,7355 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |
| **Comercio** | 0,6173 | 0,0007 | Se rechaza la hipótesis nula. Los datos no siguen una distribución normal. |

**Fuente:** Elaboración propia.

Se observa que los datos presentan una distribución aproximadamente normal en los sectores de manufactura y servicios; sin embargo, este supuesto no se cumple en el sector comercio. La violación del supuesto de normalidad en uno de los grupos imposibilita el uso de pruebas paramétricas como ANOVA para evaluar diferencias en las medias de las variables latentes asociadas al constructo de *sensing* entre los distintos sectores industriales. En consecuencia, se recurre a una prueba no paramétrica, específicamente la prueba de Kruskal-Wallis[[1]](#footnote-1), la cual constituye una alternativa robusta al ANOVA cuando no se cumple la normalidad (o homogeneidad de varianzas). Esta prueba permite determinar si existen diferencias significativas en las medianas (o distribuciones de rangos) entre los grupos. A continuación, se presentan los resultados obtenidos tras su aplicación:

**Tabla 6. Resultados de prueba Kruskal-Wallis para análisis bivariado entre variable latente *sensing* y sector de la industria**

|  |  |  |
| --- | --- | --- |
| **ESTADÍSTICO H** | **P-VALOR** | **DECISIÓN** |
| 0.5757 | 0.7499 | No se rechaza la hipótesis nula. Las distribuciones de los grupos son iguales. |

**Fuente:** Elaboración propia.

A partir de los resultados presentados en la tabla anterior, se observa que, dado el valor p obtenido, no se rechaza la hipótesis nula de la prueba de Kruskal-Wallis. Esto indica que no existen diferencias estadísticamente significativas en las distribuciones de la variable latente asociada al *sensing* entre los tres sectores de la industria. En consecuencia, no se justifica la aplicación de pruebas post-hoc, como la prueba de Dunn, ya que estas solo se implementan cuando se detectan diferencias globales significativas que requieran identificar los pares de grupos responsables de dichas diferencias.

En este caso, con base en la muestra analizada, no es posible afirmar la existencia de diferencias estadísticamente significativas —y por tanto generalizables a la población— en las medias de la variable latente asociada al *sensing* entre los tres sectores de la industria evaluados.

**4.2.4 Ubicación geográfica de la empresa**

Para llevar a cabo este análisis, se plantea la siguiente pregunta central, que orienta el desarrollo del estudio: ¿existen diferencias estadísticamente significativas en las variables latentes de *sensing* entre empresas internacionales, nacionales, regionales y locales? En otros términos, **¿la propensión al *sensing* podría variar en función de la ubicación geográfica de la empresa?**

Esta pregunta puede abordarse mediante un ejercicio estadístico de comparación de medias de las variables latentes de *sensing* entre los distintos grupos de empresas según su ubicación geográfica.

Antes de responder a la pregunta central formulada previamente, es necesario presentar los resultados de una serie de pruebas estadísticas preliminares. En particular, se evalúan dos supuestos fundamentales para la comparación de medias: normalidad (es decir, que las variables latentes de *sensing* sigan una distribución normal dentro de cada grupo) y homocedasticidad (igualdad de varianzas entre los grupos).

Para la verificación del supuesto de normalidad se emplea la prueba de Shapiro-Wilk. A continuación, se reportan los resultados de la prueba de Shapiro-Wilk (SW) aplicada a cada grupo de empresas, con el objetivo de evaluar el cumplimiento del supuesto de normalidad en las variables latentes de *sensing*:

**Tabla 7. Resultados de prueba de normalidad para análisis bivariado entre variable latente *sensing* y ubicación geográfica de la empresa[[2]](#footnote-2)**

|  |  |  |  |
| --- | --- | --- | --- |
| **GRUPO** | **ESTADÍSTICO** | **P-VALOR** | **DECISIÓN** |
| **Internacional** | 0,834 | 0,19 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |
| **Nacional** | 0,846 | 0,06 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |
| **Local** | 0,917 | 0,20 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |

**Fuente:** Elaboración propia.

A partir de los resultados reportados en la tabla anterior, es posible observar que, en cada uno de los grupos, el valor p asociado a la prueba de Shapiro-Wilk supera el umbral de significancia convencional (p > 0.05), por lo que no se rechaza la hipótesis nula. Esto indica que los datos de las variables latentes de *sensing* siguen una distribución normal en todos los grupos analizados.

Posteriormente, para evaluar el supuesto de homocedasticidad —es decir, la igualdad de varianzas entre los grupos definidos por ubicación geográfica de la empresa— se emplea la prueba de Levene (en adelante, LV).

A continuación, se reportan los resultados de la prueba de Levene (LV) aplicada a los grupos definidos, con el propósito de evaluar la igualdad de varianzas en las variables latentes de *sensing*:

**Tabla 8. Resultados de prueba de homocedasticidad para análisis bivariado entre variable latente *sensing* y ubicación geográfica de la empresa**

|  |  |  |
| --- | --- | --- |
| **ESTADÍSTICO LV** | **P-VALOR** | **DECISIÓN** |
| 1,0330 | 0,3719 | No se rechaza la hipótesis nula. Las varianzas de los grupos poblaciones son iguales. |

**Fuente:** Elaboración propia.

A partir de los resultados reportados en la tabla anterior, es posible observar que, conforme al valor p obtenido, no se rechaza la hipótesis nula de homocedasticidad. Esto permite concluir que las varianzas entre los grupos de empresas, definidos según su tamaño, pueden considerarse iguales, cumpliéndose así el supuesto de igualdad de varianzas.

Una vez verificados los supuestos de normalidad y homocedasticidad, se procede con la aplicación del análisis de varianza (ANOVA) para determinar si existen diferencias significativas entre las medias de los grupos. A continuación, se reportan los resultados de la prueba de ANOVA aplicada a las variables latentes de *sensing*, con el fin de evaluar si existen diferencias estadísticamente significativas entre los grupos de empresas según su ubicación geográfica:

**Tabla 9. Resultados de prueba ANOVA para análisis bivariado entre variable latente *sensing* y ubicación geográfica de la empresa**

|  |  |  |
| --- | --- | --- |
| **ESTADÍSTICO F** | **P-VALOR** | **DECISIÓN** |
| 0,6970 | 0,5630 | No se rechaza la hipótesis nula. Las medias poblacionales de todos los grupos son iguales. |

**Fuente:** Elaboración propia.

A partir de los resultados presentados en la tabla anterior, se observa que, dado el valor p obtenido, no se rechaza la hipótesis nula de la prueba de ANOVA. Esto indica que no existen diferencias estadísticamente significativas en las medias de la variable latente asociada al *sensing* entre los tres grupos de empresas según su ubicación geográfica. En consecuencia, no se justifica la aplicación de pruebas post-hoc, como Tukey-HSD, ya que estas solo se implementan cuando se detectan diferencias globales significativas que requieran identificar los pares de grupos responsables de dichas diferencias.

En este caso, con base en la muestra analizada, no es posible afirmar la existencia de diferencias estadísticamente significativas —y por tanto generalizables a la población— en las medias de la variable latente asociada al *sensing* entre los tres grupos de empresas según su ubicación geográfica.

**4.2.5 Edad de la empresa**

Para llevar a cabo este análisis, se plantea la siguiente pregunta central, que orienta el desarrollo del estudio: ¿existe una relación lineal entre la propensión al *sensing* y la edad de la empresa? En otros términos, **¿empresas con más o menos años en la industria tienen mayor o menor propensión al sensing?**

Inicialmente podemos revisar la correlación existente entre las variables latentes calculadas para *sensing* y la edad de la empresa. Al calcular la correlación siguiendo el método de Pearson encontramos lo siguiente:

**Gráfico 3: Correlación entre variable latente *sensing* y edad la empresa**

Gráfico, Gráfico de dispersión

El contenido generado por IA puede ser incorrecto.

**Fuente:** Elaboración propia.

Como primer resultado, se obtuvo un coeficiente de correlación de Pearson de 0,04 (r = 0,04), lo cual indica una relación positiva extremadamente débil (prácticamente nula) entre la edad de las empresas incluidas en la muestra y la variable latente de *sensing*. Adicionalmente, el valor p asociado a esta prueba (0.8216) indica que dicha correlación no es estadísticamente significativa, por lo que no puede generalizarse a la población. Es importante señalar que la ausencia de una correlación lineal significativa no implica que no pueda existir algún otro tipo de relación, por ejemplo, no lineal. Esta posibilidad puede explorarse empíricamente mediante la inspección gráfica de la nube de puntos, en la cual se observa una dispersión sin patrón lineal claro, lo que refuerza la evidencia de una débil o nula relación lineal entre las variables analizadas.

Como complemento al análisis de correlación de Pearson, se calcularon también las correlaciones no paramétricas de Spearman y Kendall, obteniéndose coeficientes de 0.2078 y 0.1678, respectivamente. No obstante, al examinar los valores p asociados a ambas pruebas (0.2888 para Spearman y 0.2123 para Kendall), se concluye que no existe una asociación estadísticamente significativa entre las variables analizadas, ya que en ambos casos los valores p superan el umbral convencional de significancia (α = 0,05).

Dado que no se evidencia una relación lineal estadísticamente significativa entre las variables edad de la empresa y la variable latente *sensing*, no se considera necesario avanzar hacia un modelo de regresión lineal simple. Esta decisión se fundamenta en que la regresión lineal busca precisamente modelar la relación lineal entre dos variables, bajo el supuesto de que la variable independiente explica, al menos en parte, la variabilidad de la dependiente. Sin embargo, los resultados obtenidos a través de la prueba de correlación de Pearson indican que no existe evidencia suficiente para afirmar la existencia de dicha relación lineal, lo cual hace inapropiado continuar con un análisis basado en ese supuesto.

**4.2.6 Gasto de Investigación y Desarrollo**

Para llevar a cabo este análisis, se plantea la siguiente pregunta central, que orienta el desarrollo del estudio: ¿existe una relación lineal entre la propensión al *sensing* y gasto en I+D? En otros términos, **¿empresas con mayor gasto en I+D tienen mayor propensión al *sensing*?**

Inicialmente podemos revisar la correlación existente entre las variables latentes calculadas para *sensing* y el gasto en I+D. Al calcular la correlación siguiendo el método de Pearson encontramos lo siguiente:

**Gráfico 4: Correlación entre variable latente *sensing* y gasto en I+D**

Gráfico

El contenido generado por IA puede ser incorrecto.

**Fuente:** Elaboración propia.

Como primer resultado, se obtuvo un coeficiente de correlación de Pearson de -0.21 (r = -0.21), lo cual indica una relación negativa débil entre el gasto en I+D de las empresas incluidas en la muestra y la variable latente de *sensing*. Adicionalmente, el valor p asociado a esta prueba (0.2918) indica que dicha correlación no es estadísticamente significativa, por lo que no puede generalizarse a la población. Es importante señalar que la ausencia de una correlación lineal significativa no implica que no pueda existir algún otro tipo de relación, por ejemplo, no lineal. Esta posibilidad puede explorarse empíricamente mediante la inspección gráfica de la nube de puntos, en la cual se observa una dispersión sin patrón lineal claro, lo que refuerza la evidencia de una débil o nula relación lineal entre las variables analizadas.

Por otro lado, es importante señalar que, según el coeficiente de correlación de Pearson, la relación entre las variables presenta una naturaleza negativa; es decir, a mayor propensión al *sensing*, se observa una menor inversión en I+D. Este resultado es, desde una perspectiva intuitiva, contradictorio, ya que cabría esperar lo contrario: que una mayor orientación al *sensing* esté asociada con un mayor gasto en innovación.

Sin embargo, es fundamental recordar que, como se evidenció anteriormente, el coeficiente de Pearson no fue estadísticamente significativo, por lo que no se puede afirmar la existencia de una relación lineal entre ambas variables.

En este contexto, es posible que exista una relación no lineal, la cual podría estar distorsionando la pendiente negativa observada en la línea de regresión (línea roja). Por tanto, no es válido concluir que existe una relación lineal negativa real entre las dos variables.

Como complemento al análisis de correlación de Pearson, se calcularon también las correlaciones no paramétricas de Spearman y Kendall, obteniéndose coeficientes de 0.1454 y 0.1294, respectivamente. No obstante, al examinar los valores p asociados a ambas pruebas (0.4604 para Spearman y 0.3468 para Kendall), se concluye que no existe una asociación estadísticamente significativa entre las variables analizadas, ya que en ambos casos los valores p superan el umbral convencional de significancia (α = 0,05). Nótese que las correlaciones calculadas para Spearman y Kendall son positivas, y no negativas como la de Pearson. Esto podría estar reflejando justamente que la relación entre ambas no es lineal.

Dado que no se evidencia una relación lineal estadísticamente significativa entre la variable de gasto en I+D y la variable latente *sensing*, no se considera necesario avanzar hacia un modelo de regresión lineal simple. Esta decisión se fundamenta en que la regresión lineal busca precisamente modelar la relación lineal entre dos variables, bajo el supuesto de que la variable independiente explica, al menos en parte, la variabilidad de la dependiente. Sin embargo, los resultados obtenidos a través de la prueba de correlación de Pearson indican que no existe evidencia suficiente para afirmar la existencia de dicha relación lineal, lo cual hace inapropiado continuar con un análisis basado en ese supuesto.

* 1. **Análisis bivariados para Seizing**

En esta sección se presenta el análisis de los resultados asociados a la capacidad que tienen las organizaciones para movilizar recursos, tomar decisiones estratégicas y ejecutar acciones orientadas a aprovechar oportunidades en su entorno. En el marco del modelo conceptual, *Seizing* representa una fase indispensable para las empresas para transformar la intención estratégica en intervenciones concretas y, por tanto, se relaciona con la estructura de las operaciones y la asignación efectiva de capacidades internas. Como en la sección anterior, se aplicó un modelo de GRM para la estimación del índice latente () y se llevaron a cabo diversos análisis comparativos entre las empresas siguiendo los mismos criterios de validación estadística utilizados en la dimensión de *Sensing*.

Ahora bien, el propósito de este análisis está enfocado en identificar si existen diferencias estadísticamente significativas en los niveles de *Seizing* según características estructurales de las organizaciones, y de esta manera poder aportar evidencia acerca de los posibles factores que facilitan o limitan la capacidad de ejecución estratégica en el contexto empresarial evaluado.

**Tamaño de la compañía**

Continuando con el enfoque analítico aplicado en la dimensión anterior, se llevó a cabo una evaluación de si existen diferencias significativas en los niveles latentes de *Seizing* según el tamaño empresarial, esto es, la verificación del cumplimiento de los supuestos que permiten realizar comparaciones válidas entre grupos haciendo uso de ANOVA.

Dicho eso se evidenció que los datos asociados a la dimensión analizada presentaron una distribución consistente con la normalidad según los resultados arrojados por la prueba de Shapiro-Wilk. Por su parte, a través del test de Levene se encontró que no hay suficiente evidencia para rechazar la hipótesis de igualdad de vairanzas (p=0.0901 frente a un estadístico de 2.4280), lo que respalda la aplicación del análisis de varianza.

**Tabla 4. Prueba de Normalidad y Levene para Seizing**

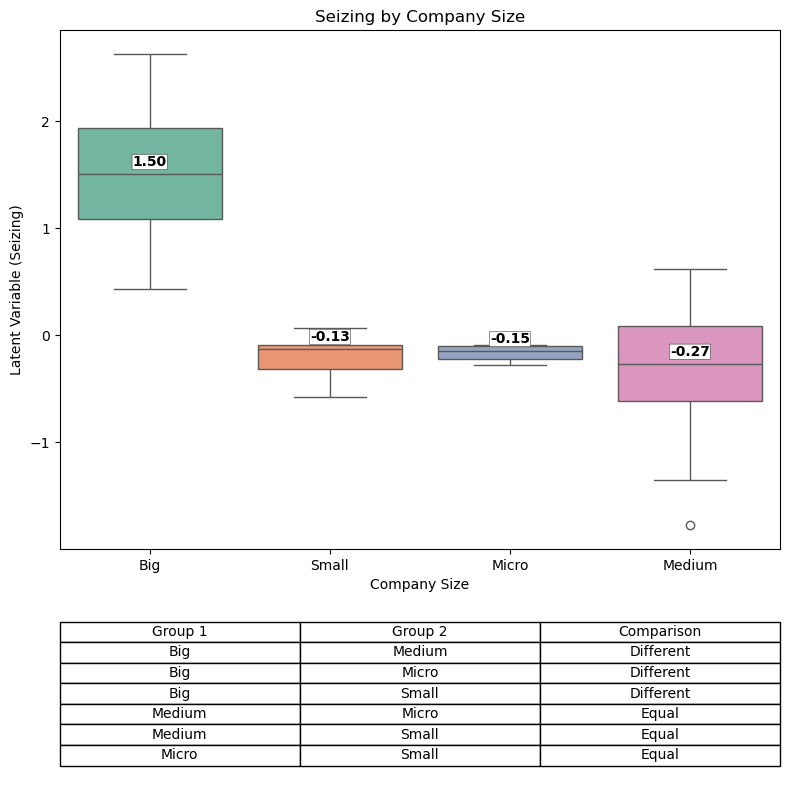
| **GRUPO** | **ESTADÍSTICO SW** | **P-VALOR** | **DECISIÓN** |
| --- | --- | --- | --- |
| **Grande** | 0,9963 | 0,9872 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |
| **Mediana** | 0,9396 | 0,3776 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |
| **Pequeña** | 0,9577 | 0,7922 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |
| **Micro** | 0,8819 | 0,3469 | No se rechaza la hipótesis nula. Los datos siguen una distribución normal. |

**Fuente: Elaboración propia**

Una vez verificados los supuestos estadísticos necesarios, se llevó a cabo la aplicación de análisis de varianza (ANOVA) para evaluar si existen diferencias significativas en los niveles latentes de *Seizing* entre empresas clasificadas por tamaño. El resultado obtenido (F = 9.6422; p = 0.0002) muestra una diferencia estadísticamente significativa, lo que indica que al menos uno de los grupos empresariales presenta un nivel de *Seizing* distinto a los demás.

Este resultado refuerza la idea de que el tamaño organizacional puede estar vinculado a la forma en que las empresas movilizan recursos, toman decisiones y ejecutan acciones estratégicas. Sin embargo, no especifica qué grupos difieren entre sí, por tanto, se recurre a una prueba posterior (Tukey) para el análisis por pares, lo cual se evidenciará con mayor claridad en el gráfico número 4, donde se ilustran visualmente las comparaciones de medias entre los diferentes tamaños de empresa.

**Gráfico 4. Índice Seizing según tamaño de la compañía**



**Fuente: Elaboración propia**

Ahora bien, enfocando el análisis en el gráfico número 4, es posible observar diferencias claras en los niveles de *Seizing* según el tamaño empresarial. En particular, las empresas grandes destacan con un valor medio de 1.50 en la variable latente, muy por encima del resto de grupos. Este resultado sugiere que, en términos de capacidad de ejecución estratégica, las grandes organizaciones de la muestra cuentan con condiciones estructurales o procesos internos que favorecen una mayor respuesta frente a oportunidades identificadas explicadas por la primera dimensión.

De otro lado, los otros grupos (micro, pequeñas y medianas) evidencian valores considerablemente más bajos y relativamente similares entre sí. Esta proximidad en el índice, a pesar de sus estructuras organizacionales en términos del tamaño, sugieren que los tres grupos enfrentan barreras comunes en términos de movilización de recursos, coordinación operativa y/o poca capacidad de implementación de estrategias.

El patrón encontrado en esta comparación refuerza la idea de que el tamaño empresarial no solo marca una diferencia en los recursos, sino también en la capacidad de traducir decisiones estratégicas en acciones efectivas.

**Ventas Último año**

Dado que la dimensión *Seizing* se relaciona con la capacidad de una organización para movilizar recursos, tomar decisiones y ejecutar acciones estratégicas, resulta pertinente explorar si esta capacidad guarda alguna relación con los resultados del volumen de ventas de las compañías registradas en el último año. Para ello, como en la dimensión anterior, se calculó el coeficiente de correlación de Pearson entre la variable latente de *Seizing* y las ventas reportadas, obteniendo un valor de r = 0.54.

Este resultado indica una correlación positiva moderada, lo que sugiere que las empresas con mayores niveles de *Seizing* tienden a registrar mayores ventas. Aunque la relación no es perfectamente lineal, el patrón observado en el gráfico respalda esta asociación: a medida que aumentan los valores de *Seizing*, se incrementa también el rango de ventas, especialmente en los niveles más altos.

**Gráfico 5. Relación entre la variable latente (Seizing) y ventas del último año**

Gráfico, Gráfico de dispersión

El contenido generado por IA puede ser incorrecto.

**Fuente: Elaboración propia**

* 1. **Análisis bivariados para Transformation**

Este índice sintetiza la propensión de una empresa a implementar prácticas organizacionales asociadas con la dimensión de *Transformation*, a partir de un conjunto de ítems ordinales con diferente nivel de dificultad y capacidad de discriminación. Gracias a esta estimación, se dispone de una métrica psicométricamente válida y comparable entre empresas, que supera las limitaciones de un simple promedio de ítems.

Con el objetivo de comprender de manera más integral los niveles de transformación organizacional y su impacto en el desempeño empresarial, se realizaron las siguientes validaciones:

**Tamaño de Empresa**

Se comprobó que el Índice de Transformación presenta distribución normal dentro de cada grupo empresarial (Shapiro-Wilk) y varianzas homogéneas (Levene), lo que habilita el uso de ANOVA. Los resultados mostraron diferencias estadísticamente significativas entre grupos (p < 0.001), siendo el grupo de **empresas grandes** el que reporta valores más altos en transformación, con una mediana de 1.73, frente a valores negativos en micro, pequeñas y medianas empresas.

**Gráfico 6. Relación entre la variable latente (Transformation) y ventas del último año**

A chart with different colored squares

AI-generated content may be incorrect.

**Fuente: Elaboración propia**

Las pruebas post-hoc (Tukey HSD) confirmaron que las empresas grandes difieren significativamente de los otros grupos, mientras que no se observaron diferencias relevantes entre micro, pequeñas y medianas. Esto respalda la capacidad del índice para **distinguir niveles de madurez transformacional en función del tamaño organizacional**, coherente con lo reportado en la literatura sobre capacidades tecnológicas y recursos disponibles.

Este hallazgo se refuerza con el coeficiente de correlación obtenido (r = 0.78), el cual evidencia una fuerte relación positiva, indicando que las empresas de mayor tamaño tienden a presentar niveles más altos de transformación.

**Ventas Último año**

El análisis de regresión lineal simple entre el índice de transformación y las **ventas del último año** arrojó un modelo significativo (p < 0.01), con un coeficiente de determinación R² = 0.255. Esto implica que cerca del **25.5% de la variabilidad en las ventas** puede ser explicada por el nivel de transformación de la empresa.

Además, el coeficiente estimado sugiere que, **por cada punto adicional en el índice**, las ventas anuales aumentan, en promedio, en aproximadamente **20 millones de unidades monetarias**, lo cual representa un impacto económico tangible.

**Gráfico 7. Relación entre la variable latente (Transformation) y ventas del último año**

A graph with green dots and red line

AI-generated content may be incorrect.

**Fuente: Elaboración propia**

Aunque hay dispersión, el patrón es claro: a mayor transformación, mayores ventas. Es un predictor **significativo** de las ventas del último año, aunque no es el único factor.

La correlación moderada positiva (r = 0.51) entre transformación y ventas refuerza esta asociación, confirmando que **la transformación organizacional no es solo un constructo deseable, sino un factor estratégicamente relevante en el desempeño comercial**.

A close-up of a color chart

AI-generated content may be incorrect.

**Implicaciones Estratégicas**

* Este índice ofrece una forma **válida, objetiva y continua** de medir la transformación en organizaciones, especialmente útil para **segmentar empresas**, **identificar brechas**, y **priorizar acciones de política o inversión**.
* Al estar alineado con variables de desempeño reales como las ventas, **consolida su utilidad práctica como indicador de madurez organizacional**.

**Consideraciones**

Aunque el tamaño muestral (n = 56) representa una limitación para la generalización, los errores estándar de estimación obtenidos fueron bajos y estables, particularmente en la dimensión Transformation. Esto respalda la **precisión aceptable del modelo**, pero se recomienda replicar el análisis con muestras mayores para reforzar la robustez de los hallazgos.

1. **Anexos (pruebas estadísticas)**

**Prueba de correlación de Pearson**

La prueba de correlación de Pearson es una técnica estadística paramétrica utilizada para evaluar la fuerza y dirección de la relación lineal entre dos variables cuantitativas continuas. Fue desarrollada por Karl Pearson y es una de las herramientas más ampliamente utilizadas en estadística inferencial para analizar asociaciones entre variables.

Esta prueba tiene las siguientes hipótesis:

Para decidir si existe una correlación lineal (y distinta de cero) entre las dos variables, se aplica la siguiente regla de decisión. Se selecciona un nivel de significancia α (0,05). Luego, se compara el valor-p obtenido en la prueba con el valor α:

* Si p-valor > α, no se rechaza , por lo que no hay evidencia suficiente para afirmar la existencia de una correlación lineal significativa entre las variables.
* Si p-valor < α, se rechaza , por lo que existe una correlación lineal estadísticamente significativa entre las variables.

**Prueba de correlación de Spearman**

La prueba de correlación de Spearman es una técnica estadística no paramétrica utilizada para evaluar la asociación monótona entre dos variables. A diferencia de la correlación de Pearson, Spearman no requiere que las variables sigan una distribución normal ni que la relación sea lineal. En su lugar, esta prueba se basa en los rangos de los datos y es útil cuando se trabaja con variables ordinales o cuantitativas no normales.

Esta prueba tiene las siguientes hipótesis:

Para decidir si existe una correlación lineal (y distinta de cero) entre las dos variables, se aplica la siguiente regla de decisión. Se selecciona un nivel de significancia α (0,05). Luego, se compara el valor-p obtenido en la prueba con el valor α:

* Si p-valor > α, no se rechaza , lo que indica que no se dispone de evidencia suficiente para afirmar la existencia de una asociación significativa.
* Si p-valor < α, se rechaza , lo que indica que existe evidencia estadística de una relación significativa entre las variables.

**Prueba de correlación de Kendall**

La prueba de correlación de Kendall, también conocida como tau de Kendall, es otra técnica no paramétrica que mide la asociación ordinal entre dos variables. Esta prueba se basa en el número de pares concordantes y discordantes entre las observaciones y es especialmente útil en muestras pequeñas o en presencia de muchos empates. Al igual que Spearman, Kendall no asume normalidad ni linealidad.

Esta prueba tiene las siguientes hipótesis:

Para decidir si existe una correlación lineal (y distinta de cero) entre las dos variables, se aplica la siguiente regla de decisión. Se selecciona un nivel de significancia α (0,05). Luego, se compara el valor-p obtenido en la prueba con el valor α:

* Si p-valor > α, no se rechaza , lo que indica que no se dispone de evidencia suficiente para afirmar la existencia de una asociación significativa.
* Si p-valor < α, se rechaza , lo que indica que existe evidencia estadística de una relación significativa entre las variables.

**Prueba de normalidad (Shapiro-Wilk)**

La prueba de Shapiro-Wilk es una de las pruebas estadísticas más utilizadas para evaluar si una muestra proviene de una distribución normal. Fue propuesta por Samuel Shapiro y Martin Wilk en 1965 y es especialmente adecuada para tamaños muestrales pequeños o moderados (n < 50), aunque también se puede aplicar a tamaños mayores. Esta prueba tiene las siguientes hipótesis:

Para decidir si los datos siguen una distribución normal, se aplica la siguiente regla de decisión. Se selecciona un nivel de significancia α (0,05). Luego, se compara el valor-p obtenido en la prueba con el valor α:

* Si p-valor > α, no se rechaza , por lo que se asume normalidad (los datos siguen una distribución normal).
* Si p-valor < α, se rechaza , por lo que se rechaza normalidad (los datos no siguen una distribución normal).

**Prueba de Levene**

La prueba de Levene es una prueba estadística utilizada para evaluar la igualdad de varianzas (homocedasticidad) entre dos o más grupos. Es una condición importante para aplicar pruebas paramétricas como ANOVA. A diferencia de otras pruebas como la de Bartlett, la prueba de Levene es más robusta frente a desviaciones de la normalidad, por lo que se recomienda cuando no se puede garantizar que los datos sean normales. Esta prueba tiene las siguientes hipótesis:

Para decidir si existe igualdad de varianzas entre dos o más grupos, se aplica la siguiente regla de decisión. Se selecciona un nivel de significancia α (0,05). Luego, se compara el valor-p obtenido en la prueba con el valor α:

* Si p-valor > α, no se rechaza , por lo que se asume homogeneidad de varianzas (existe igualdad de varianzas entre dos o más grupos).
* Si p-valor < α, se rechaza , por lo que se rechaza homogeneidad de varianzas (no existe igualdad de varianzas entre dos o más grupos).

**Prueba ANOVA**

La prueba ANOVA (*Analysis of Variance*) de un factor es una técnica estadística paramétrica utilizada para comparar las medias de tres o más grupos independientes. Evalúa si las diferencias observadas entre las medias grupales son estadísticamente significativas o si pueden explicarse por la variabilidad interna del error aleatorio. Fue desarrollada por Ronald Fisher y es ampliamente utilizada en contextos experimentales y comparativos. Esta prueba tiene las siguientes hipótesis:

Las hipótesis de esta prueba se plantean de la siguiente manera:

Para decidir si al menos una de las medias poblaciones difiere, se aplica la siguiente regla de decisión. Se selecciona un nivel de significancia α (0,05). Luego, se compara el valor-p obtenido en la prueba con el valor α:

* Si p-valor > α, no se rechaza , por lo que no hay evidencia de diferencias significativas en las medias poblaciones entre grupos.
* Si p-valor < α, se rechaza , por lo que existe evidencia de diferencias significativas en las medias poblacionales entre grupos.

**Prueba Kruskal-Wallis**

La prueba de Kruskal-Wallis es una técnica estadística no paramétrica utilizada para comparar tres o más grupos independientes cuando no se cumplen los supuestos necesarios para aplicar una prueba ANOVA de un factor, como la normalidad de los datos o la homogeneidad de varianzas. Esta prueba fue desarrollada por William Kruskal y W. Allen Wallis y es particularmente útil cuando los datos son ordinales o cuando se sospecha que las distribuciones no son normales.

En lugar de comparar las medias de los grupos, como en el ANOVA, la prueba de Kruskal-Wallis evalúa si las distribuciones de los rangos de los grupos difieren significativamente entre sí.

Las hipótesis de esta prueba se plantean de la siguiente manera:

Para decidir si al menos una de las medias poblaciones difiere, se aplica la siguiente regla de decisión. Se selecciona un nivel de significancia α (0,05). Luego, se compara el valor-p obtenido en la prueba con el valor α:

* Si p-valor > α, no se rechaza , por lo que no hay evidencia suficiente para afirmar diferencias significativas entre las distribuciones de los grupos.
* Si p-valor < α, se rechaza , por lo que existe al menos un grupo cuya distribución difiere significativamente de las demás.

**Prueba Tukey HSD**

La prueba de Tukey HSD (*Honest Significant Difference*) es un procedimiento post hoc utilizado después de obtener un resultado significativo en un ANOVA. Su propósito es identificar cuáles pares de grupos presentan diferencias significativas entre sus medias, controlando el error tipo I asociado a realizar múltiples comparaciones. Fue propuesta por John Tukey y es una de las pruebas más utilizadas cuando se cumple la homogeneidad de varianzas y se tiene un diseño balanceado o moderadamente desbalanceado. Las hipótesis para esta prueba son las siguientes (se testea cada hipótesis en cada par de grupos de ; en nuestro caso, sería equivalente a comparar cada tamaño de empresa entre sí)

Para decidir si la media poblaciones difiere en cada par de grupo se aplica la siguiente regla de decisión. Se selecciona un nivel de significancia α (0,05). Luego, se compara el valor-p obtenido en la prueba con el valor α:

* Si p-valor > α, no se rechaza , por lo que no hay evidencia de diferencias significativas en las medias poblaciones entre el grupo .
* Si p-valor < α, se rechaza , por lo que existe evidencia de diferencias significativas en las medias poblacionales entre el grupo .

**Referencias**

* Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores*. Psychometrika Monograph Supplement.
* Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Lawrence Erlbaum Associates.
* De Ayala, R. J. (2009). *The Theory and Practice of Item Response Theory*. Guilford Press.
* Baker, F. B., & Kim, S.-H. (2017). *The Basics of Item Response Theory Using R* (2nd ed.). Springer. <https://doi.org/10.1007/978-3-319-54205-8>
* Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52(3/4), 591–611.
* Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.), Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling (pp. 278–292). Stanford University Press.
* Fisher, R. A. (1925). Statistical Methods for Research Workers. Oliver and Boyd.
* Tukey, J. W. (1949). Comparing individual means in the analysis of variance. Biometrics, 5(2), 99–114.
* Pearson, K. (1895). Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58, 240–242.
* Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101.
* Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, 30(1/2), 81–93.
* Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583–621.

1. La prueba de Kruskal-Wallis no asume existencia de normalidad de datos y homogeneidad en las varianzas. [↑](#footnote-ref-1)
2. Se omite del análisis el grupo de empresas “Regional” puesto que solamente tiene dos empresas, y la prueba de normalidad exige un mínimo de tres observaciones para ser ejecutada. [↑](#footnote-ref-2)