Minitab®

INFORME TÉCNICO SOBRE EL ASISTENTE DE MINITAB

Este documento forma parte de un conjunto de informes técnicos que explican la investigación llevada a cabo por los especialistas en estadística de Minitab para desarrollar los métodos y las verificaciones de datos que se utilizan en el Asistente de Minitab Statistical Software.

Prueba t de 1 muestra

Introducción

La prueba t de 1 muestra se utiliza para estimar la media de procesos y compararla con un valor objetivo. Esta prueba se considera un procedimiento robusto debido a que es extremadamente sensible al supuesto de normalidad cuando la muestra es moderadamente grande. De acuerdo a la mayoría de los libros de texto de estadística, la prueba t de 1 muestra y el intervalo de confianza t para la media son apropiados para cualquier muestra con un tamaño de 30 o más.

En este artículo, describimos las simulaciones que realizamos para evaluar esta regla general de un mínimo de 30 unidades de muestra. Nuestras simulaciones se enfocaron en el impacto de la no normalidad en la prueba t de 1 muestra. Queríamos evaluar el impacto que tienen los datos poco comunes en los resultados de la prueba.

Con base en nuestra investigación, el Asistente realiza automáticamente las siguientes verificaciones de sus datos y muestra los resultados en la Tarjeta de informe:

- Datos poco comunes
- Normalidad (¿Es la muestra suficientemente grande para que la normalidad no represente un problema?)
- Tamaño de la muestra

Para obtener información general sobre la metodología correspondiente a la prueba t de 1 muestra, consulte Arnold (1990), Casella y Berger (1990), Moore y McCabe (1993) y Srivastava (1958).

Nota Los resultados de este documento también se aplican a la prueba t pareada del Asistente, debido a que la prueba t pareada aplica el método de la prueba t de 1 muestra a una muestra de diferencias pareadas.

Verificaciones de datos

Datos poco comunes

Datos poco comunes son valores de datos extremadamente grandes o pequeños, conocidos también como valores atípicos. Los datos poco comunes pueden tener una fuerte influencia en los resultados del análisis. Cuando la muestra es pequeña, estos pueden afectar las probabilidades de hallar resultados estadísticamente significativos. Los datos poco comunes pueden indicar problemas con la recolección de los datos o pudieran deberse a un comportamiento inusual del proceso bajo estudio. Con frecuencia, resulta útil investigar estos puntos de los datos, los cuales se deberían corregir cuando sea posible.

Objetivo

Queríamos desarrollar un método para verificar los valores de los datos que son muy grandes o muy pequeños en relación con la muestra general, lo cual pudiera afectar los resultados del análisis.

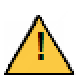
Método

Desarrollamos un método para determinar la presencia de datos poco comunes con base en el método descrito por Hoaglin, Iglewicz y Tukey (1986) para identificar valores atípicos en las gráficas de caja.

Resultados

El Asistente identifica un punto de los datos como poco común si es más de 1.5 veces el rango intercuartil después del cuartil inferior o superior de la distribución. Los cuartiles inferior y superior son los percentiles 25 y 75 de los datos. El rango intercuartil es la diferencia entre los dos cuartiles. Este método funciona adecuadamente cuando existen múltiples valores atípicos debido a que es posible detectar cada valor atípico específico.

Cuando se buscan datos poco comunes, la Tarjeta de informe del Asistente correspondiente a la prueba t de 1 muestra exhibe los siguientes indicadores de estado:

Estado	Condición
	No hay puntos de datos poco comunes.
	Por lo menos un punto de los datos es poco común y pudiera afectar los resultados de la prueba.

Normalidad

La prueba t de 1 muestra se deriva bajo el supuesto de que la población está normalmente distribuida. Afortunadamente, incluso si los datos no están normalmente distribuidos, este método funciona adecuadamente cuando el tamaño de la muestra es suficientemente grande.

Objetivo

Queríamos determinar el efecto de la no normalidad en el error Tipo I de la prueba con el fin de proporcionar directrices sobre el tamaño de la muestra y la normalidad.

Método

Realizamos simulaciones para determinar el tamaño de la muestra para el cual pudiera ignorarse el supuesto de normalidad cuando se realiza una prueba t de 1 muestra o se calcula un intervalo de confianza t para la media de una población.

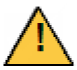
Diseñamos el primer estudio para evaluar el efecto de la no normalidad en la tasa de error Tipo I de la prueba. Específicamente, queríamos deducir el tamaño de la muestra mínimo necesario para que la prueba fuera sensible a la distribución de la población. Realizamos la prueba t de 1 muestra en muestras pequeñas, moderadas y grandes generadas a partir de poblaciones normales y no normales. Las poblaciones no normales incluían poblaciones de ligeramente a severamente asimétricas, poblaciones simétricas con colas ligeras y pesadas, así como poblaciones normales contaminadas. La población normal se utilizó como control para efectos de comparación. Para cada caso, calculamos y comparamos los niveles de significancia simulados con el nivel de significancia objetivo, o nominal, de 0.05. Si la prueba funciona adecuadamente, los niveles de significancia simulados deberían aproximarse a 0.05. Examinamos los niveles de significancia simulados de todas las diferentes condiciones para evaluar el tamaño de la muestra mínimo para el cual estos permanecen cerca del nivel objetivo, independientemente de la distribución. Véase el Apéndice A para obtener más detalles.

En el segundo estudio, examinamos el efecto de la no normalidad en la tasa de error Tipo II de la prueba. El diseño de la simulación es idéntico al primer estudio. Sin embargo, comparamos los niveles de potencia simulada bajo diferentes condiciones con niveles de potencia objetivo utilizando la función de potencia teórica de la prueba t de 1 muestra. Véase el Apéndice B para obtener más detalles.

Resultados

El efecto de la no normalidad en las tasas de error Tipo I y Tipo II de la prueba es mínimo para tamaños de muestra de por lo menos 20. Sin embargo, cuando la población original de la muestra es severamente asimétrica, se requieren muestras más grandes. Recomendamos un tamaño de la muestra de aproximadamente 40 para esos casos. Véase el Apéndice A y el Apéndice B para obtener más detalles.

Debido a que la prueba funciona adecuadamente con muestras relativamente pequeñas, el Asistente no prueba la normalidad de los datos. En su lugar, verifica el tamaño de la muestra y exhibe los siguientes indicadores de estado en la Tarjeta de informe:

Estado	Condición
	El tamaño de la muestra es por lo menos 20; la normalidad no representa problema alguno.
	El tamaño de la muestra es menor que 20; la normalidad pudiera representar un problema.

Tamaño de la muestra

Por lo general, se realiza una prueba de hipótesis para recolectar evidencia con el fin de rechazar la hipótesis nula de que "no existe diferencia". Si las muestras son demasiado pequeñas, la potencia de la prueba pudiera no ser adecuada para detectar una diferencia entre las medias que realmente existe, lo cual genera un error Tipo II. Es por tanto crucial asegurar que el tamaño de la muestra sea suficientemente grande para detectar diferencias prácticamente importantes con alta probabilidad.

Objetivo

Si los datos no proporcionan evidencia suficiente para rechazar la hipótesis nula, queremos determinar si el tamaño de la muestra era suficientemente grande para que la prueba detectara diferencias prácticas de interés con alta probabilidad. Si bien el objetivo de planificar los tamaños de las muestras es asegurar que sean suficientemente grandes para detectar diferencias importantes con alta probabilidad, tampoco deberían ser excesivamente grandes, debido a que las diferencias insignificantes se volverían estadísticamente significativas con alta probabilidad.

Método

El análisis de potencia y tamaño de la muestra se basa en la función de potencia teórica de la prueba específica utilizada para realizar el análisis estadístico. Tal como se discutió anteriormente, la función de potencia de la prueba t de 1 muestra es sensible al supuesto de normalidad cuando el tamaño de la muestra es por lo menos 20. La función de potencia depende del tamaño de la muestra, la diferencia entre la media objetivo y la media de la población, así como de la varianza de la población. Véase el Apéndice B para obtener más detalles.

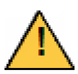
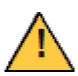
Resultados

Cuando los datos no proporcionan evidencia suficiente para rechazar la hipótesis nula, el Asistente calcula las diferencias prácticas que se pueden detectar con una probabilidad de 80%

y 90% para los tamaños de muestra especificados. Además, si el usuario proporciona una diferencia de interés práctica particular, el Asistente calcula el tamaño de la muestra que ofrezca una probabilidad de 80% y 90% de detectar la diferencia.

No existe un resultado general que informar debido a que los resultados dependen de la muestra específica del usuario. Sin embargo, puede consultar el Apéndice B para obtener más información sobre la potencia de la prueba t de 1 muestra.

Cuando se realizan verificaciones para determinar la potencia y el tamaño de la muestra, la Tarjeta de informe del Asistente correspondiente a la prueba t de 1 muestra exhibe los siguientes indicadores de estado:

Estado	**Condición**
	La prueba halla una diferencia entre la media y el objetivo, de modo que la potencia no representa problema alguno. O La potencia es suficiente. La prueba no halló una diferencia entre la media y el objetivo, pero la muestra es suficientemente grande para proporcionar una probabilidad de por lo menos 90% de detectar la diferencia especificada.
	La potencia pudiera ser suficiente. La prueba no halló una diferencia entre la media y el objetivo, pero la muestra es suficientemente grande para proporcionar una probabilidad de 80% y 90% de detectar la diferencia especificada. Se informa el tamaño de la muestra requerido para alcanzar una potencia de 90%.
	La potencia pudiera no ser suficiente. La prueba no halló una diferencia entre la media y el objetivo, y la muestra es suficientemente grande para proporcionar una probabilidad de 60% a 80% de detectar la diferencia especificada. Se informan los tamaños de las muestras requeridos para alcanzar una potencia de 80% y 90%.
	La potencia no es suficiente. La prueba no halló una diferencia entre la media y el objetivo, y la muestra no es suficientemente grande para proporcionar una probabilidad de por lo menos 60% de detectar la diferencia especificada. Se informan los tamaños de las muestras requeridos para alcanzar una potencia de 80% y 90%.
	La prueba no halló una diferencia entre la media y el objetivo. Usted no especificó una diferencia práctica entre la media y el objetivo que detectar; por lo tanto, el informe indica las diferencias que podría detectar con una probabilidad de 80% y 90%, con base en sus tamaños de muestra, desviaciones estándar y alfa.

Referencias

Arnold, S.F. (1990). *Mathematical statistics*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Casella, G., & Berger, R. L. (1990). *Statistical inference.* Pacific Grove, CA: Wadsworth, Inc.

Hoaglin, D. C., Iglewicz, B., and Tukey, J. W. (1986). Performance of some resistant rules for outlier labeling. *Journal of the American Statistical Association*, *81*, *991-999*.

Moore, D.S. & McCabe, G.P. (1993). *Introduction to the practice of statistics*, 2nd ed. New York, NY: W. H. Freeman and Company.

Neyman, J., Iwaszkiewicz, K. & Kolodziejczyk, S. (1935). Statistical problems in agricultural experimentation, *Journal of the Royal Statistical Society, Series B, 2, 107-180*.

Pearson, E.S., & Hartley, H.O. (Eds.). (1954). *Biometrika tables for statisticians*, Vol. I. London: Cambridge University Press.

Srivastava, A. B. L. (1958). Effect of non-normality on the power function of t-test, *Biometrika*, *45, 421-429*.

Apéndice A: Impacto de la no normalidad en el nivel de significancia (validez de la prueba)

Bajo el supuesto de normalidad, la prueba t de 1 muestra es una prueba *sin sesgo uniformemente más potente* (UMP) con un tamaño α. Es decir, la prueba es tan potente como, o más que, cualquier otra prueba sin sesgo con tañamo α de la media. Sin embargo, cuando la población original de la muestra no está normalmente distribuida, entonces las anteriores propiedades óptimas se cumplen si el tamaño de la muestra es suficientemente grande. En otras palabras, para muestras suficientemente grandes, el nivel de significancia real de la prueba t de 1 muestra es aproximadamente igual al nivel objetivo para datos tanto normales como no normales, además, la función de potencia de la prueba no es sensible al supuesto de normalidad (Srivastava, 1958).

Queríamos determinar qué tan grande debe ser una muestra para considerarse suficientemente grande para que la prueba t no sea sensible al supuesto de normalidad. Numerosos libros de texto recomiendan que si el tamaño de la muestra $n \geq 30$, entonces se puede ignorar el supuesto de normalidad para la mayoría de los propósitos prácticos (Arnold, 1990; Casella y Berger, 1990; y Moore y McCabe, 1993). El propósito de la investigación descrita en estos apéndices es realizar estudios de simulación para evaluar esta regla general al examinar el impacto de diferentes distribuciones no normales en la prueba t de 1 muestra.

Estudio de simulación A

Queríamos examinar el impacto de la no normalidad en la tasa de error Tipo I de la prueba para evaluar un tamaño de muestra mínimo que sea estable y que se mantenga cerca de la tasa de error objetivo, independientemente de la distribución.

Para ello, se realizaron dos pruebas t bilaterales utilizando $\alpha = 0.05$ con muestras aleatorias de diferentes tamaños ($n = 10, 15, 20,\ 25,\ 30, 35, 40, 50, 60, 80, 100$) provenientes de diferentes distribuciones con distintas propiedades. Estas distribuciones incluían:

- La distribución normal estándar (N(0,1))
- Distribuciones simétricas y con cola pesada, como la distribución t con 5 y 10 grados de libertad (t(5),t(10))
- La distribución de Laplace con ubicación 0 y escala 1 (Lpl))
- Distribuciones asimétricas y con cola pesada representadas por la distribución exponencial con escala 1 (Exp), las distribuciones de Chi-cuadrado con 3, 5 y 10 grados de libertad (Chi(3), Chi(5), Chi(10))

- Distribuciones simétricas y con cola ligera, como la distribución uniforme (U(0,1)), la distribución Beta con ambos parámetros establecidos en 3 (B(3,3))
- Una distribución asimétrica hacia la izquierda y con cola pesada (B (8,1))

Además, para evaluar el efecto directo de los valores atípicos, generamos muestras de distribuciones normales contaminadas, definidas como:

$$\mathrm{CN}(\mathrm{p}, \sigma) = \mathrm{pN}(0{,}1) + (1 - \mathrm{p})\mathrm{N}(0, \sigma)$$

donde p se define como el parámetro de combinación y $1 - \mathrm{p}$ es la proporción de contaminación o proporción de valores atípicos. Seleccionamos dos poblaciones normales contaminadas para el estudio: $\mathrm{CN}(0.9{,}3)$ (10% de los integrantes de la población son valores atípicos) y $\mathrm{CN}(.8{,}3)$ (20% de los integrantes de la población son valores atípicos). Estas dos distribuciones son simétricas y tienen colas largas debido a los valores atípicos.

Para cada tamaño de muestra, se extrajeron 10,000 réplicas de muestra de cada población y se realizó una prueba t de 1 muestra con hipótesis nula $\mu = \mu_o$ e hipótesis alternativa $\mu \neq \mu_o$ para cada una de las 10,000 muestras. Para cada prueba, establecimos la media hipotética μ_o como la media real de la población original de la muestra. Como resultado, para un tamaño de muestra específico, la fracción de las 10,000 réplicas de muestra que produjeron un rechazo de la hipótesis nula representa la tasa de error Tipo I simulada o el nivel de significancia de la prueba. Debido a que el nivel de significancia objetivo es 5%, el error de simulación es aproximadamente 0.2%.

Los resultados de la simulación se muestran en las Tablas 1 y 2, y se representan gráficamente en las Figuras 1 y 2.

Resultados y resumen

Los resultados (véanse la Tabla 1 y la Figura 1) muestran que cuando las muestras se generan de poblaciones simétricas, los niveles de significancia simulados de la prueba se aproximan al nivel de significancia objetivo cuando las muestras tienen tamaños pequeños. Sin embargo, los resultados son ligeramente conservadores para las distribuciones simétricas con colas pesadas cuando las muestras son pequeñas, incluidas las muestras pequeñas generadas de distribuciones normales contaminadas. Aparentemente, los valores atípicos también reducen el nivel de significancia de la prueba cuando las muestras son pequeñas. Sin embargo, este efecto se revierte cuando las muestras pequeñas se generan de poblaciones originales simétricas con colas ligeras (las distribuciones Beta (3,3) y uniforme). Los niveles de significancia simulados son ligeramente mayores en estos casos.

Tabla 1 Niveles de significancia simulados de la prueba t de 1 muestra bilateral para muestras generadas de poblaciones simétricas. El nivel de significancia objetivo es $\alpha = 0.05$.

Dist.	N(0,1)	t(5)	t(10)	Lpl	CN(.9,3)	CN(.8,3)	B(3,3)	U(0,1)
N	Simétrica y colas pesadas						Simétrica y colas ligeras	
10	0.050	0.046	0.048	0.044	0.043	0.039	0.057	0.057
15	0.051	0.050	0.049	0.049	0.043	0.043	0.053	0.054
20	0.047	0.051	0.051	0.047	0.043	0.044	0.051	0.052
25	0.050	0.047	0.050	0.046	0.046	0.046	0.048	0.050
30	0.053	0.050	0.048	0.043	0.049	0.046	0.050	0.048
35	0.052	0.047	0.049	0.050	0.047	0.045	0.051	0.054
40	0.046	0.052	0.054	0.048	0.046	0.049	0.044	0.050
50	0.050	0.049	0.051	0.048	0.047	0.051	0.053	0.050
60	0.052	0.049	0.053	0.050	0.051	0.056	0.054	0.052
80	0.049	0.050	0.051	0.047	0.047	0.052	0.049	0.049
100	0.050	0.052	0.049	0.051	0.052	0.054	0.051	0.054

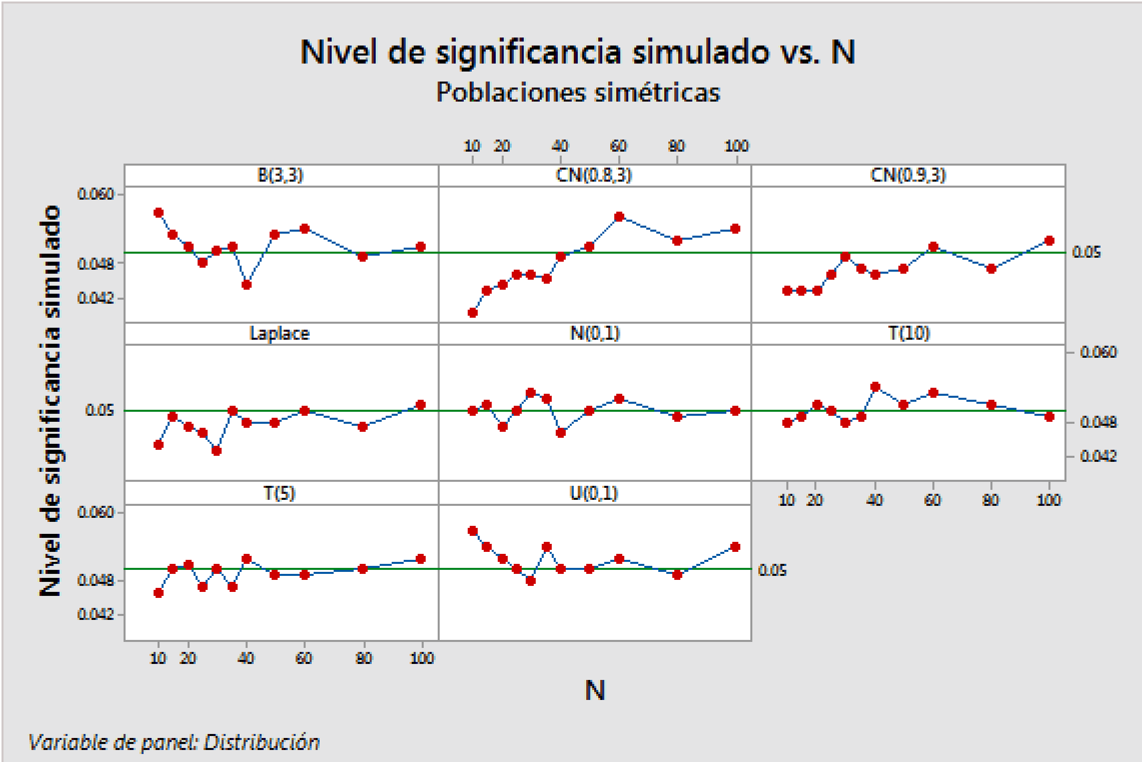


Figura 1 Gráfica de niveles de significancia simulados de las pruebas t de 1 muestra bilaterales en función del tamaño de las muestras generadas de poblaciones simétricas. El nivel de significancia objetivo es $\alpha = 0.05$.

Sin embargo, cuando las muestras se generan de distribuciones asimétricas, el rendimiento de la prueba depende de la severidad de la asimetría. Los resultados de la Tabla 2 y la Figura 2 muestran que la prueba t de 1 muestra es sensible a la asimetría en muestras pequeñas. En el caso de poblaciones severamente asimétricas (exponencial, Chi(3) y Beta(8,1)), se requieren muestras más grandes para que los niveles de significancia simulados se aproximen al nivel de significancia objetivo. Sin embargo, en el caso de poblaciones moderadamente asimétricas (Chi(5) y Chi(10)), un tamaño de muestra mínimo de 20 es suficiente para que los niveles de significancia simulados se aproximen al nivel objetivo. Con un tamaño de la muestra de 20, el nivel de significancia simulado es aproximadamente 0.063 para la distribución de Chi-cuadrado con 5 grados de libertad y es aproximadamente 0.056 para la distribución de Chi-cuadrado con 10 grados de libertad.

Tabla 2 Niveles de significancia simulados de la prueba t de 1 muestra bilateral para muestras generadas de poblaciones asimétricas. El nivel de significancia objetivo es $\alpha = 0.05$.

N	**Exp**	**Chi(3)**	**B(8,1)**	**Chi(5)**	**Chi(10)**
	Asimetría de la población				
	2.0	**1.633**	**-1.423**	**1.265**	**0.894**
	Niveles de significancia simulados				
10	0.101	0.089	0.087	0.069	0.060
15	0.088	0.076	0.072	0.068	0.057
20	0.083	0.073	0.069	0.063	0.056
25	0.075	0.068	0.067	0.067	0.056
30	0.069	0.067	0.066	0.058	0.054
35	0.075	0.067	0.062	0.062	0.056
40	0.067	0.067	0.061	0.059	0.056
50	0.064	0.057	0.062	0.057	0.054
60	0.063	0.056	0.061	0.054	0.055
80	0.059	0.058	0.053	0.052	0.052
100	0.060	0.055	0.055	0.047	0.053

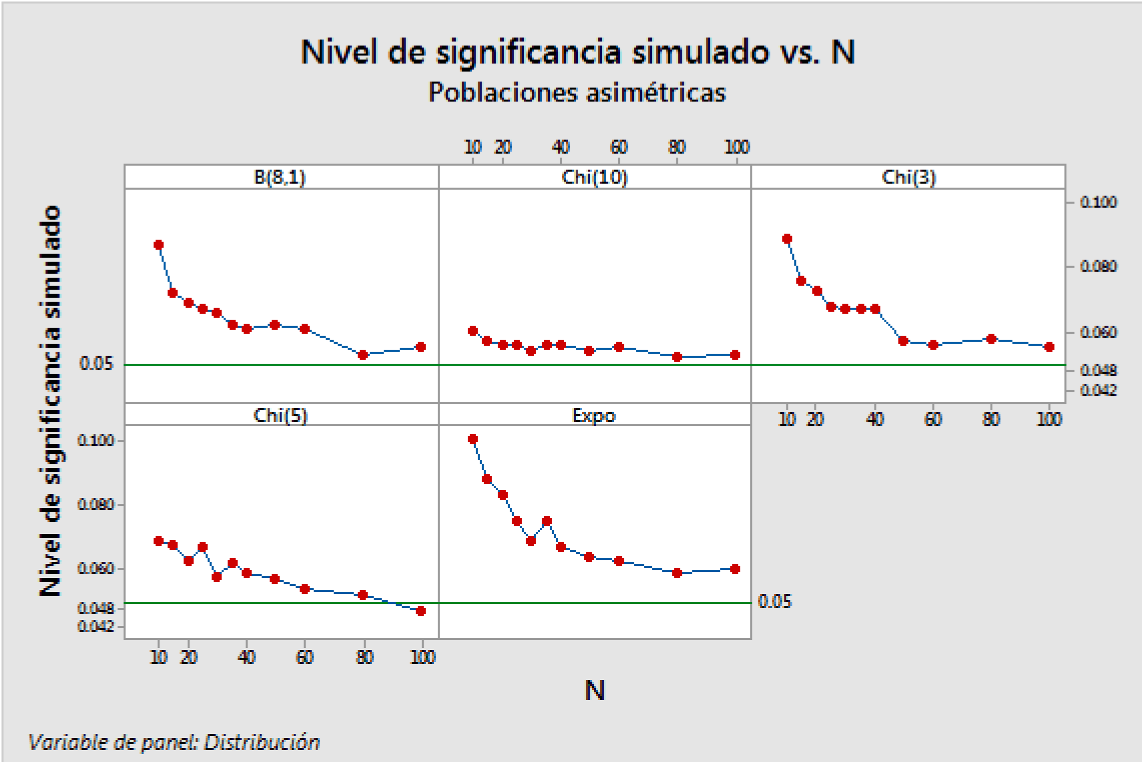


Figura 2 Gráfica de niveles de significancia simulados de las pruebas t de 1 muestra bilaterales en función del tamaño de las muestras generadas de poblaciones asimétricas. El nivel de significancia objetivo es $\alpha = 0.05$.

En esta investigación, nos enfocamos en las pruebas de hipótesis y no en los intervalos de confianza. Sin embargo, los resultados se extienden naturalmente hacia los intervalos de confianza debido a que tanto las pruebas de hipótesis como los intervalos de confianza se pueden utilizar para determinar significancia estadística.

Apéndice B: Tamaño de la muestra y potencia de la prueba

Queríamos examinar la sensibilidad de la función de potencia al supuesto de normalidad bajo el cual se deriva. Tenga en cuenta que si β es el error Tipo II de una prueba, entonces $1 - \beta$ es la potencia de la prueba. Como resultado, se determina el tamaño de la muestra planificado, de modo que la tasa de error Tipo II sea pequeña o, equivalentemente, el nivel de potencia sea razonablemente alto.

Las funciones de potencia de las pruebas t son ampliamente conocidas y están bien documentadas. Pearson y Hartley (1954) y Neyman, Iwaszkiewicz, y Kolodziejczyk (1935) proporcionan gráficas y tablas de las funciones de potencia.

Para una prueba t de 1 muestra bilateral con tamaño α, se puede utilizar una expresión matemática de esta función en términos del tamaño de la muestra y la diferencia δ entre la media real μ y la media hipotética μ_o tal que

$$\pi(n, \delta) = 1 - F_{n-1,\lambda}\left(t_{n-1}^{\alpha/2}\right) + F_{n-1,\lambda}\left(-t_{n-1}^{\alpha/2}\right)$$

donde $\mathrm{F}_{\mathrm{d},\lambda}(.)$ es la C.D.F de la distribución t no central con $\mathrm{d} = \mathrm{n} - 1$ grados de libertad y parámetro de no centralidad

$$\lambda = \frac{\delta\sqrt{n}}{\sigma}$$

y donde $\mathrm{t}_{\mathrm{d}}^{\alpha}$ denota el punto percentil 100α superior de la distribución t con d grados de libertad.

En el caso de las alternativas unilaterales, la potencia se expresa como

$$\pi(n, \delta) = 1 - F_{n-1,\lambda}(t_{n-1}^{\alpha})$$

para probar la hipótesis nula frente $\mu > \mu_o$ y se expresa como

$$\pi(n, \delta) = F_{n-1,\lambda}(-t_{n-1}^{\alpha})$$

cuando se prueba la hipótesis nula frente $\mu < \mu_o$.

Estas funciones se derivan bajo el supuesto de que los datos están normalmente distribuidos y de que el nivel de significancia de la prueba se fijó en un valor α.

Estudio de simulación B

Diseñamos esta simulación para evaluar el efecto de la no normalidad en la función de potencia teórica de la prueba t de 1 muestra. Para evaluar el efecto de la no normalidad, comparamos los niveles de potencia simulada con los niveles de potencia objetivo, que se calcularon con la función de potencia teórica de la prueba.

Realizamos pruebas t bilaterales con $\alpha = 0.05$ en muestras aleatorias de diferentes tamaños ($n = 10, 15, \ 20, 25, 30, 35, 40, 50, 60, 80, 100$) generadas de las mismas poblaciones descritas en el primer estudio de simulación (véase el Apéndice A).

Para cada una de las poblaciones, la hipótesis nula de la prueba es $\mu = \mu_o - \delta$ y la hipótesis alternativa es $\neq \mu_o - \delta$, donde μ_o se establece en la media real de la población y $\delta = \sigma/2$ (σ es la desviación estándar de la población original). Por lo tanto, la diferencia entre la media real y la media hipotética es 0, de modo que la decisión correcta es rechazar la hipótesis nula.

Para cada tamaño de la muestra especificado, se extrajeron 10,000 réplicas de muestra de cada una de las distribuciones. Para cada tamaño de la muestra especificado, la fracción de las 10,000 réplicas por la cual se rechaza la hipótesis nula representa el nivel de potencia simulada de la prueba con el tamaño de la muestra y la diferencia δ que se especificaron. Tenga en cuenta que elegimos este particular valor para la diferencia debido a que produce valores de potencia que son relativamente pequeños cuando los tamaños de las muestras son pequeños.

Además, los valores de potencia teórica correspondientes, conocidos como valores de potencia objetivo, se calcularon con la diferencia δ y los distintos tamaños de muestra para establecer una comparación con los niveles de potencia simulada.

Los resultados de la simulación se muestran en la Tabla 3 y la Tabla 4, y se representan gráficamente en la Figura 3 y la Figura 4.

Resultados y resumen

Los resultados confirman que la potencia de la prueba t de 1 muestra no es, generalmente, sensible al supuesto de normalidad cuando el tamaño de la muestra es suficientemente grande. Para las muestras generadas de poblaciones simétricas, los resultados de la Tabla 3 muestran que los niveles de potencia objetivo y de potencia simulada se aproximan, incluso con muestras pequeñas. Las curvas de potencia correspondientes que se muestran en la Figura 3 son prácticamente indistinguibles. Para muestras generadas de distribuciones normales contaminadas, los valores de potencia son un tanto conservadores para tamaños de muestra de pequeño a moderado. Esto pudiera ser debido a que el nivel de significancia real de la prueba para dichas poblaciones es ligeramente mayor que el nivel de significancia α establecido.

Tabla 3 Niveles de potencia simulada con una diferencia $\delta = \sigma/2$ para una prueba t de 1 muestra bilateral con tamaño $\alpha = 0.05$ cuando las muestras se generan de poblaciones simétricas. Los niveles de potencia simulada se comparan con los niveles de potencia teórica objetivo derivados bajo el supuesto de normalidad.

n	Potencia objetivo	N(0,1)	t(5)	t(10)	Lpl	CN(.9,3)	CN(.8,3)	B(3,3)	U(0,1)
		Nivel de potencia simulada con $\delta = \sigma/2$ (poblaciones simétricas)							
10	0.293	0.299	0.334	0.311	0.357	0.361	0.385	0.28	0.269
15	0.438	0.438	0.48	0.45	0.491	0.512	0.511	0.423	0.421
20	0.565	0.57	0.603	0.578	0.60	0.629	0.623	0.557	0.548
25	0.67	0.674	0.695	0.68	0.691	0.712	0.70	0.665	0.67
30	0.754	0.756	0.77	0.756	0.767	0.768	0.765	0.754	0.75
35	0.82	0.819	0.827	0.815	0.82	0.819	0.812	0.822	0.818
40	0.869	0.87	0.871	0.868	0.862	0.869	0.868	0.875	0.867
50	0.934	0.933	0.929	0.93	0.929	0.923	0.925	0.932	0.94
60	0.968	0.967	0.963	0.965	0.964	0.955	0.955	0.968	0.971
80	0.993	0.993	0.989	0.992	0.991	0.988	0.989	0.994	0.994
100	0.999	0.998	0.996	0.998	0.999	0.998	0.996	0.999	0.999

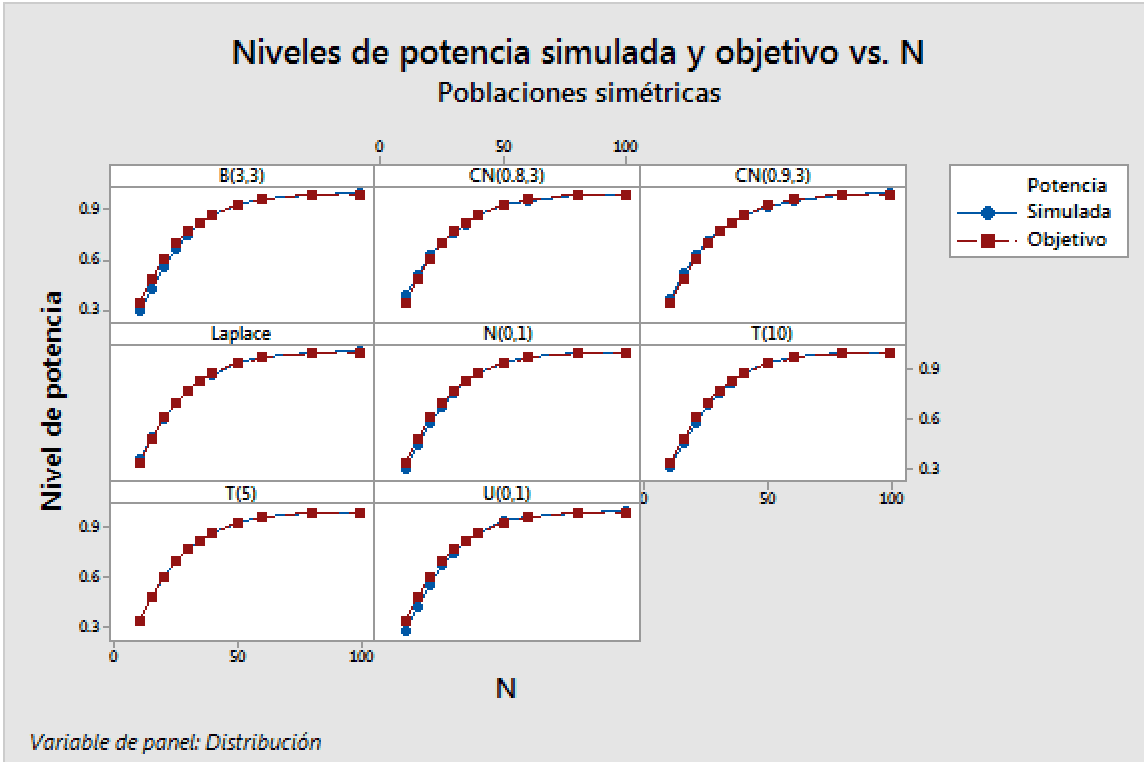


Figura 3 Curvas de potencia simulada comparadas con las curvas de potencia teórica objetivo con $\alpha = 0.05$, prueba t de 1 muestra bilateral cuando las muestras se generan de poblaciones simétricas. Los valores de potencia se evalúan con una diferencia de $\delta = \sigma/2$.

Sin embargo, cuando las muestras provienen de poblaciones asimétricas, los valores de potencia simulada se ubican fuera del objetivo para muestras pequeñas, tal como se muestra en la Tabla 4 y la Figura 4. Para poblaciones moderadamente asimétricas, tal como la distribución de Chi-cuadrado con 5 grados de libertad y la distribución de Chi-cuadrado con 10 grados de libertad, cuando el tamaño de la muestra es por lo menos 20, se aproximan los niveles de potencia objetivo y simulada. Por ejemplo, para $n = 20$, el nivel de potencia objetivo es 0.565 cuando los niveles de potencia simulada son 0.576 y 0.577 para las distribuciones de Chi-cuadrado 5 y Chi-cuadrado 10, respectivamente. En el caso de distribuciones extremadamente asimétricas, se requieren muestras más grandes para que los niveles de potencia simulada se aproximen al nivel de significancia objetivo. Esto pudiera ser porque la prueba t de 1 muestra no controla el error Tipo I de manera apropiada cuando los tamaños de las muestras son pequeños y las poblaciones originales son extremadamente asimétricas.

Tabla 4 Valores de potencia simulada con una diferencia $\delta = \sigma/2$ para una prueba t de 1 muestra bilateral con tamaño $\alpha = 0.05$ cuando las muestras se generan de poblaciones asimétricas. Los valores de potencia simulada se comparan con los valores de potencia objetivo derivados bajo el supuesto de normalidad.

N	Potencia objetivo	Exp		Chi(3)	B(8,1)	Chi(5)	Chi(10)
			Asimetría de la población				
		2.0		1.633	-1.423	1.265	0.894
			Niveles de potencia simulada				
10	0.293	0.206		0.212	0.39	0.225	0.238
15	0.438	0.416		0.413	0.484	0.409	0.407
20	0.565	0.604		0.591	0.566	0.576	0.577
25	0.67	0.763		0.734	0.657	0.709	0.695
30	0.754	0.859		0.834	0.729	0.808	0.785
35	0.82	0.917		0.895	0.776	0.874	0.835
40	0.869	0.955		0.935	0.823	0.925	0.905
50	0.934	0.987		0.981	0.90	0.973	0.96
60	0.968	0.997		0.994	0.937	0.991	0.985
80	0.993	1.00		0.999	0.98	0.999	0.997
100	0.999	1.00		1.00	0.994	1.00	1.00

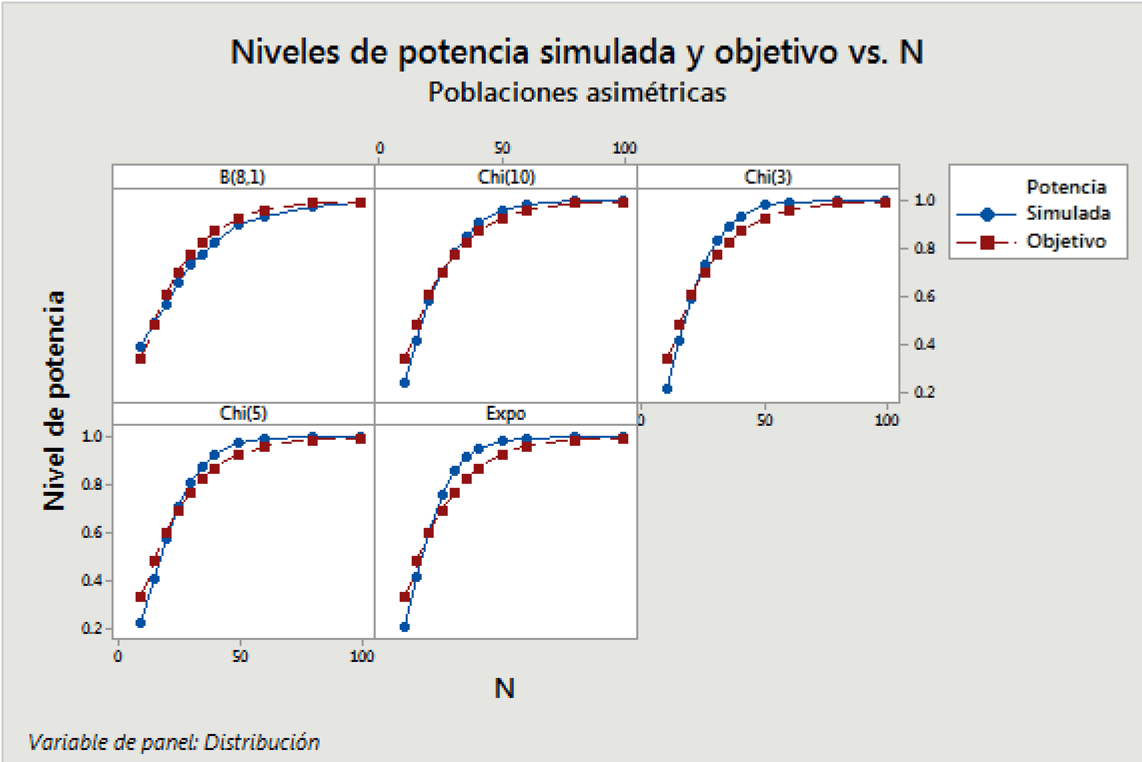


Figura 4 Curvas de potencia simulada comparadas con las curvas de potencia teórica objetivo con $\alpha = 0.05$, prueba t de 1 muestra bilateral cuando las muestras se generan de poblaciones simétricas. Los valores de potencia se evalúan con una diferencia de $\delta = \sigma/2$.

En resumen, para distribuciones moderadamente asimétricas, la función de potencia es fiable si el tamaño de la muestras es por lo menos 20, independientemente de la población original de la que se extrae la muestra. Para poblaciones extremadamente asimétricas, se requiere un tamaño de la muestra mayor (aproximadamente 40) para que la potencia simulada se aproxime a la potencia objetivo.