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 del % de defectuosos para 1 muestra

Revisión general

Una prueba de 1 proporción se utiliza para determinar si una proporción difiere de un valor objetivo. En el análisis de calidad, la prueba se utiliza frecuentemente cuando un producto o servicio se caracteriza como defectuoso o no defectuoso con el fin de determinar si el porcentaje de artículos defectuosos difiere significativamente de un % de defectuosos objetivo.

El Asistente de Minitab incluye una Prueba del % de defectuosos para 1 muestra. Los datos recopilados para la prueba son el número de artículos defectuosos en una muestra, que se presume que es el valor observado de una variable aleatoria binomial. El Asistente utiliza métodos exactos para calcular los resultados de las pruebas de hipótesis y los intervalos de confianza; por lo tanto, la tasa de error Tipo I real debería aproximarse al nivel de significancia (alfa) especificado para la prueba y no se requieren más investigaciones. Sin embargo, el análisis de potencia y tamaño de la muestra para la prueba del % de defectuosos para 1 muestra se basa en una aproximación y necesitamos evaluar su precisión.

En este documento, investigamos la metodología utilizada para evaluar la potencia y el tamaño de la muestra de la prueba del % de defectuosos para 1 muestra, comparando la potencia teórica del método de aproximación con la potencia real de la prueba exacta.

También describimos cómo establecimos una directriz para ayudar al usuario a evaluar si su muestra es suficientemente grande para detectar si el porcentaje de artículos defectuosos difiere de un % de defectuosos objetivo. El Asistente realiza automáticamente una verificación del tamaño de la muestra y notifica los resultados en la Tarjeta de informe.

La Prueba del % de defectuosos para 1 muestra también depende de otros supuestos. Para mayor información, véase el Apéndice A.

Método del % de defectuosos para 1 muestra

Rendimiento de la función de potencia teórica

El Asistente realiza la prueba de hipótesis para una sola proporción (% de defectuosos) de una población de Bernoulli utilizando métodos exactos (tasa de verosimilitud). Sin embargo, debido a que la función de potencia de la prueba exacta no se deriva con facilidad, se obtiene una aproximación de la función de potencia utilizando la función de potencia teórica de la prueba de aproximación a la normal correspondiente.

Objetivo

Queríamos determinar si podíamos utilizar la función de potencia teórica partiendo de la aproximación a la normal para evaluar los requisitos de potencia y tamaño de la muestra para la prueba del % de defectuosos para 1 muestra en el Asistente. Para ello, necesitábamos evaluar si esta función de potencia teórica refleja de manera correcta la potencia real de la prueba exacta (tasa de verosimilitud).

Método

El estadístico de la prueba, el valor p, y el intervalo de confianza de la prueba exacta (tasa de verosimilitud) se definen en el Apéndice B. La función de potencia teórica basada en la aproximación a la normal se define en el Apéndice C. Con base en estas definiciones, realizamos simulaciones para estimar los niveles de potencia real (a los que nos referimos como niveles de potencia simulada) utilizando la prueba exacta.

Para realizar simulaciones, generamos muestras aleatorias de diferentes tamaños provenientes de distintas poblaciones de Bernoulli. Para cada población de Bernoulli, realizamos la prueba exacta a cada una de las 10,000 réplicas de las muestras. Para cada tamaño de muestra, calculamos la potencia simulada de la prueba para detectar una diferencia específica como la fracción de los 10,000 pares de muestras para los que la prueba es significativa. Para las comparaciones, también calculamos la potencia teórica correspondiente utilizando la prueba de aproximación a la normal. Si la aproximación funciona adecuadamente, los niveles de potencia teórica y simulada deberían ser similares. Para mayor información, véase el Apéndice D.

Resultados

Nuestras simulaciones demostraron que, en general, la función de potencia teórica de la prueba de aproximación a la normal y la función de potencia simulada de la prueba exacta (tasa de verosimilitud) son casi iguales. Por lo tanto, el Asistente utiliza la función de potencia teórica de la prueba de aproximación a la normal para calcular los tamaños de las muestras necesarios para asegurar que la prueba exacta tenga suficiente potencia para detectar diferencias prácticamente importantes en el porcentaje de defectuosos.

Verificaciones de datos

Tamaño de la muestra

Generalmente, se realiza una prueba de hipótesis para recolectar evidencia para rechazar la hipótesis nula de que “no existe diferencia”. Si la muestra es muy pequeña, la potencia de la prueba pudiera no ser adecuada para detectar una diferencia que realmente existe, lo cual produce un error Tipo II. Es por lo tanto crucial asegurarse de que los tamaños de las muestras sean lo suficientemente grandes para detectar diferencias parcialmente importantes con una alta probabilidad.

Objetivo

Si los datos no proporcionaban evidencia suficiente para rechazar la hipótesis nula, queríamos determinar si los tamaños de las muestras son suficientemente grandes para que la prueba detecte diferencias prácticas de interés con alta probabilidad. Si bien el objetivo de planificar los tamaños de las muestras es asegurar que estos sean suficientemente grandes para detectar diferencias importantes con alta probabilidad, el tamaño tampoco debería ser excesivo, ya que ello provocaría que las diferencias despreciables se vuelvan estadísticamente significativas.

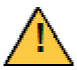
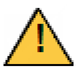
Método

El análisis de potencia y tamaño de la muestra para la prueba del % de defectuosos para 1 muestra se basa en la función de potencia teórica utilizando la aproximación a la normal, que proporciona una excelente estimación de la potencia real de la prueba exacta (véase la sección Método del % de defectuosos para 1 muestra arriba). Cuando se especifica el % de defectuosos, la función de potencia teórica depende del tamaño de la muestra y la diferencia que se desee detectar.

Resultados

Cuando los datos no proporcionan evidencia suficiente contra la hipótesis nula, el Asistente calcula diferencias prácticas que se pueden detectar con una probabilidad del 80% y 90% para el tamaño de la muestra dado. Además, si el usuario proporciona una diferencia práctica particular de interés, el Asistente calcula los tamaños de muestra que ofrezcan una probabilidad del 80% y 90% de detectar la diferencia.

Para ayudar a interpretar los resultados, la tarjeta de informe del Asistente correspondiente a la prueba del % de defectuosos para 1 muestra exhibe los siguientes indicadores de estado cuando se verifican la potencia y el tamaño de la muestra:

Estado	Condición
	La prueba halla una diferencia entre el % de defectuosos y el valor objetivo, de modo que la potencia no representa problema alguno. O La potencia es suficiente. La prueba no halló una diferencia con respecto al valor objetivo, pero la muestra es suficientemente grande para proporcionar una probabilidad de por lo menos 90% para detectar la diferencia especificada (potencia ≥ 90).
	La potencia pudiera ser suficiente. La prueba no halló una diferencia con respecto al valor objetivo, pero la muestra es suficientemente grande para proporcionar una probabilidad de 80% a 90% para detectar la diferencia especificada (80 ≤ potencia ≥ 90). Se informa el tamaño de la muestra que se requiere para alcanzar una potencia del 90%.
	La potencia pudiera no ser suficiente. La prueba no halló una diferencia con respecto al valor objetivo y la muestra es suficientemente grande para proporcionar una probabilidad de 60% a 80% para detectar la diferencia especificada (60 ≤ potencia ≥ 80). Se informan los tamaños de las muestras que se requieren para alcanzar una potencia del 80% y una potencia del 90%.
	La potencia no es suficiente. La prueba no halló una diferencia con respecto al valor objetivo y la muestra no es suficientemente grande para proporcionar una probabilidad de por lo menos 60% para detectar la diferencia especificada (potencia ≥ 60). Se informan los tamaños de las muestras que se requieren para alcanzar una potencia del 80% y una potencia del 90%.
	La prueba no halló una diferencia con respecto al valor objetivo. No se especificó la detección de una diferencia práctica; por lo tanto, el informe indica las diferencias que se pudieran detectar con una potencia del 80% y 90%, con base en el tamaño de la muestra 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.

Apéndice A: Supuestos adicionales para % de defectuosos de 1 muestra

La Prueba del % de defectuosos para 1 muestra se basa en los siguientes supuestos:

- Los datos constan de n diferentes artículos, de los cuales cada uno se clasificó como defectuoso o no defectuoso.
- La probabilidad de que un artículo sea defectuoso es la misma para cada artíuclo dentro de una muestra.
- La probabilidad de que un artículo sea defectuoso no ve afectada por si otro artículo es o no defectuoso.

Estos supuestos no se pueden comprobar en las verificaciones de datos de la Tarjeta de informe debido a que, para esta prueba, se ingresaron datos de resumen en lugar de datos sin procesar.

Apéndice B: Prueba exacta (tasa de verosimilitud)

Supongamos que observamos una muestra aleatoria $X_1, \dots, X_n$ proveniente de una distribución de Bernoulli con una probabilidad de éxito $p = \Pr(X_i = 1) = 1 - \Pr(X_i = 0)$.

Los métodos exactos para obtener una conclusión sobre p se describen abajo.

Fórmula B1: Prueba exacta y valor p

Considere una prueba de la hipótesis nula $H_0: p = p_0$ en función de cualquiera de las siguientes hipótesis alternativas: $H_A: p > p_0$, $H_A: p < p_0$ o $H_A: p \neq p_0$.

Supongamos que $X = \sum_{i=1}^{n} X_i$

Entonces, X es una variable aleatoria binomial con una cantidad de ensayos n y una probabilidad de éxito p.

Una muestra unilateral basada en X es UMP (uniformemente más potente) y una prueba de la tasa de verosimilitud. Para pruebas bilaterales, la prueba de la tasa de verosimilitud también se basa en X y el estadístico de la prueba es

$$\Lambda(X) = \left(\frac{\hat{p}}{p_0}\right)^X \left(\frac{1-\hat{p}}{1-p_0}\right)^{n-X}$$

(véase Arnold, 1990).

Los valores P para las pruebas unilaterales se pueden obtener directamente con base en la distribución exacta de X. Para las pruebas bilaterales, los valores p se calculan como la probabilidad, bajo la hipótesis nula, de observar una tasa de verosimilitud (o tasa de verosimilitud logarítmica) por lo menos tan grande como la realmente observada. Para calcular esta probabilidad, se genera un algoritmo numérico de búsqueda de raíces.

Fórmula B2: intervalo de confianza exacto

Un intervalo de confianza bilateral exacto de $100(1-\alpha)\%$ para p es

$$\frac{1}{1 + \frac{n-x+1}{x} F_{2(n-x+1),2x,\alpha/2}} \leq p \leq \frac{\frac{x+1}{n-x} F_{2(x+1),2(n-x),\alpha/2}}{1 + \frac{x+1}{n-x} F_{2(x+1),2(n-x),\alpha/2}}$$

donde x es el número observado de éxitos y $F_{\nu_1,\nu_2,\alpha}$ es el punto percentil α superior de la distribución F con ν_1 y ν_2 grados de libertad (véase Casella y Berger, 1990). Adoptamos la convención de que el límite inferior es 0 si $x = 0$ y el límite superior es 1 si $x = n$.

Apéndice C: Función de potencia teórica

Es muy complejo derivar una función de potencia teórica de la prueba exacta. Por lo tanto, aproximamos la función de potencia de la prueba utilizando la función de potencia teórica de la prueba con base en la aproximación a la normal. Esta prueba de aproximación se basa en el hecho de que la variable aleatoria

$$Z = \frac{n^{1/2}(\hat{p} - p)}{(p(1-p)^{1/2}}$$

está asintóticamente distribuida como la distribución normal estándar. La función de potencia teórica de esta prueba es bien conocida y está documentada. Para la hipótesis alternativa bilateral, la función de potencia se expresa como:

$$\pi(n, \delta) = 1 - \Phi\left(\frac{-\delta + z_{\alpha/2}\sqrt{p_o(1-p_o)/n}}{\sqrt{p(1-p)/n}}\right) + \Phi\left(\frac{-\delta - z_{\alpha/2}\sqrt{p_o(1-p_o)/n}}{\sqrt{p(1-p)/n}}\right)$$

donde $p = \delta + p_o$, $\Phi(.)$ es la distribución acumulativa de probabilidad de la distribución normal estándar y z_α es el percentil superior de la distribución normal estándar.

Para la $H_A: p > p_0$ alternativa bilateral, la función de potencia se puede expresar como

$$\pi(n, \delta) = 1 - \Phi\left(\frac{-\delta + z_{\alpha}\sqrt{p_o(1-p_o)/n}}{\sqrt{p(1-p)/n}}\right)$$

Cuando se realizan pruebas para rechazar la $H_A: p < p_0$ alternativa bilateral, la función de potencia también se puede expresar como

$$\pi(n, \delta) = \Phi\left(\frac{-\delta - z_{\alpha}\sqrt{p_o(1-p_o)/n}}{\sqrt{p(1-p)/n}}\right)$$

Apéndice D: Comparación de potencia real con potencia teórica

Simulación D1: estimación de potencia real utilizando la prueba de exacta

Diseñamos una simulación para comparar los niveles de potencia real estimados (a los que se hace referencia como niveles de potencia simulada) con los niveles de potencia teórica con base en la función de potencia de la prueba de aproximación a la normal (a los que se hace referencia como niveles de potencia aproximada). En cada experimento, generamos 10,000 muestras, cada una con un tamaño n, provenientes de una población de Bernoulli con una probabilidad de éxito p dada. Consideramos dos casos para la probabilidad de éxito: (1) una probabilidad de éxito moderada, con un valor p cercano a 0.5 (específicamente, $p = 0.45$) y una (2) a probabilidad de éxito pequeña o grande, con un valor p cercano a 0 o 1(específicamente, $p = 0.85$). Consideramos estos dos casos debido a que se sabe que la aproximación de la normal a la distribución binomial de DeMoivre-Laplace, de la cual se deriva la prueba de aproximación a la normal, es exacta cuando el tamaño de la muestra de Bernoulli es mayor que 10 y la probabilidad de éxito se aproxima a 0.5. Sin embargo, para las probabilidades de éxito más grande o más pequeña, se requieren muestras de Bernoulli más grandes para que la aproximación sea exacta.

En cada experimento, fijamos el tamaño de la muestra en un único valor de n, donde $n = 10, 15, 20, 30, \ldots, 100$. En todos los experimentos, fijamos la diferencia que se detectará $\delta = p - p_0$ en en 0.2 para asegurar que los valores de potencia resultantes no fueran demasiado pequeños o grandes a medida que aumentara el tamaño de la muestra hasta llegar a 100. Para calcular la potencia real de la prueba con base en los resultados de cada simulación, calculamos la fracción de las 10,000 réplicas de muestra para la cual la prueba exacta era significativa al nivel de significancia de $\alpha = 0.05$, utilizando tanto la prueba unilateral y bilateral. Finalmente, para efectos de comparación, calculamos los niveles de la potencia teórica correspondientes con base en la prueba de aproximación a la normal. Los resultados se muestran en la Tabla 1, a continuación.

Tabla 1 Niveles de potencia simulada y aproximada (ap.) de las pruebas exactas bilateral y unilateral. El nivel de significancia objetivo es $\alpha = 0.05$.

n	Prueba bilateral				Prueba unilateral			
	$p = 0.45$		$p = 0.85$		$p = 0.45$		$p = 0.85$	
	Potencia simulada	Potencia ap.	Potencia simulada	Potencia ap.	Potencia simulada	Potencia ap.	Potencia simulada	Potencia ap.
10	0.101	0.333	0.20	0.199	0.257	0.436	0.20	0.335
15	0.339	0.441	0.322	0.327	0.339	0.55	0.322	0.489
20	0.406	0.537	0.409	0.455	0.59	0.643	0.648	0.621
30	0.632	0.69	0.708	0.674	0.632	0.779	0.708	0.808
40	0.781	0.799	0.863	0.822	0.781	0.867	0.863	0.911
50	0.877	0.872	0.874	0.91	0.877	0.921	0.933	0.961
60	0.878	0.92	0.942	0.957	0.922	0.954	0.969	0.984
70	0.925	0.951	0.972	0.981	0.953	0.973	0.987	0.994
80	0.954	0.971	0.986	0.992	0.986	0.985	0.993	0.998
90	0.971	0.982	0.993	0.996	0.991	0.991	0.996	0.999
100	0.989	0.99	0.998	0.999	0.994	0.995	0.999	1.00

Los resultados demuestran que los niveles de la potencia simulada y los niveles de la potencia aproximada son generalmente muy consistentes. Esta consistencia se puede observar con mayor claridad cuando los resultados se representan gráficamente como curvas de potencia, tal como se muestra en las Figuras 1 y 2 a continuación.

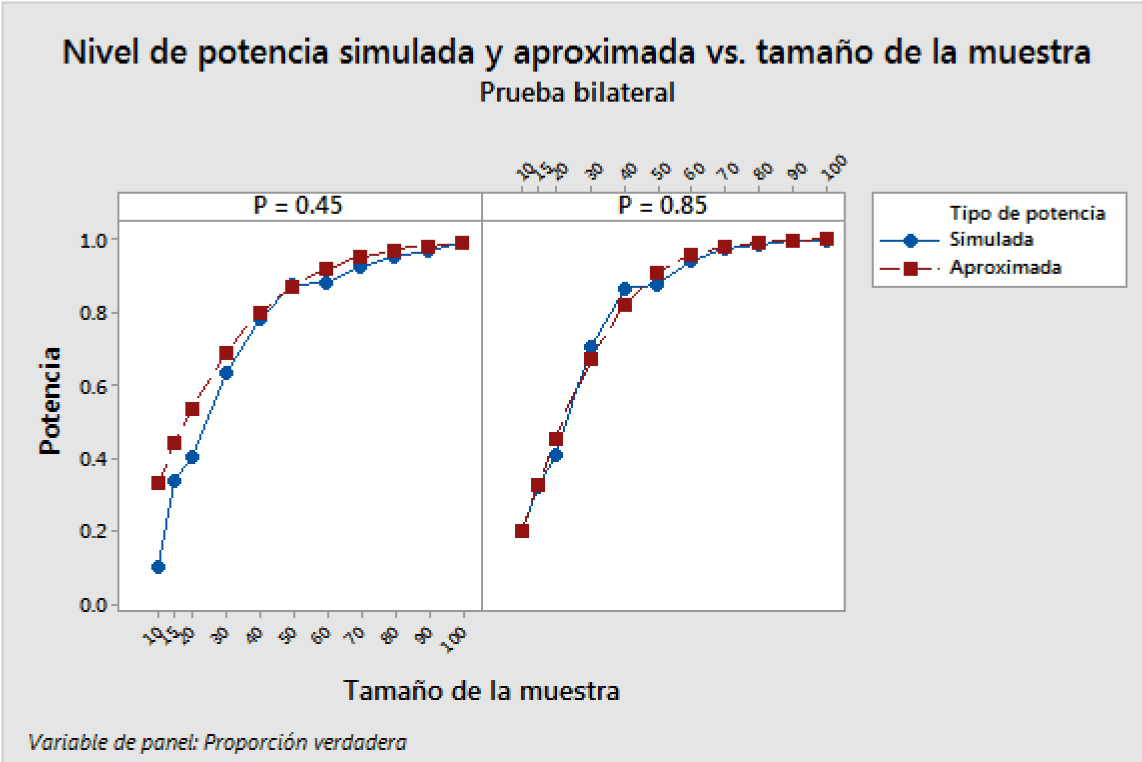


Figura 1 Gráficas de los niveles de potencia simulada y aproximada de la prueba exacta bilateral en función del tamaño de la muestra.

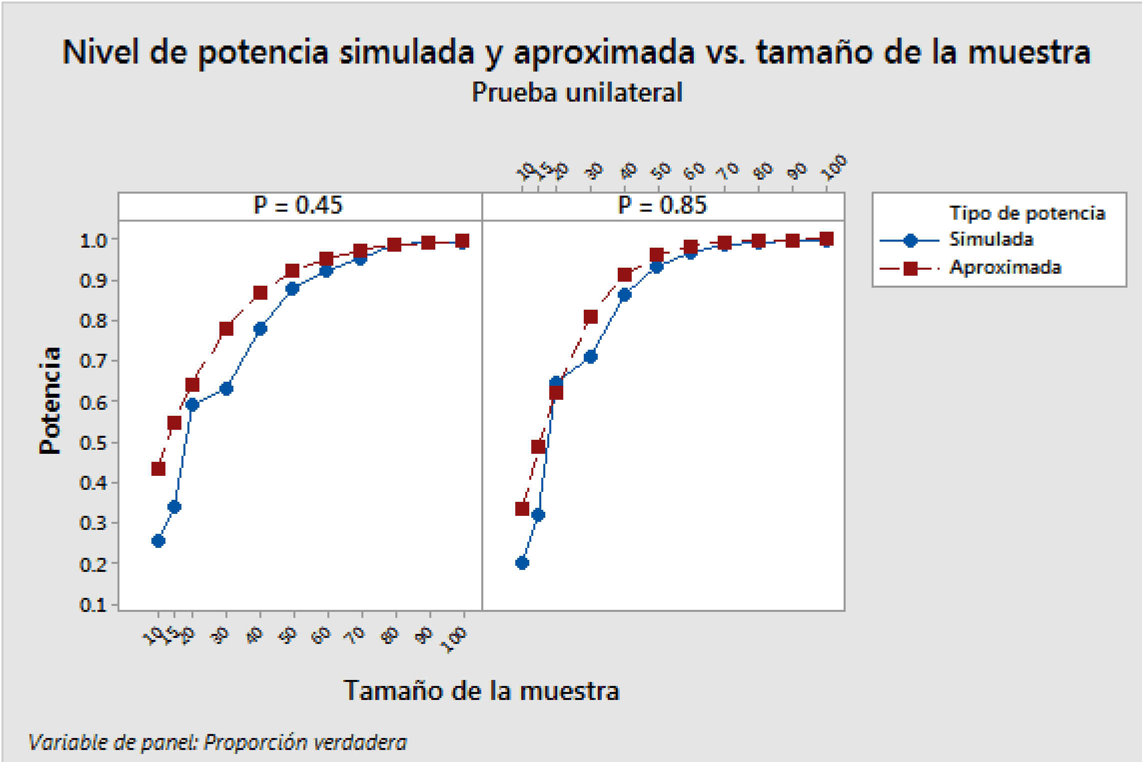


Figura 2 Gráficas de los niveles de potencia simulada y aproximada de la prueba exacta unilateral en función del tamaño de la muestra.

Las dos curvas de pontecia mostradas en cada sección en las Figuras 1 and 2 se aproximan, excepto algunas pocas veces cuando el tamaño de la muestra es pequeño. La cercanía de las curvas indica que la función de potencia aproximada coincide estrechamente con la potencia simulada cuando la prueba exacta se aplica en la práctica. Por lo tanto, es apropiado utilizar la función de potencia aproximada para calcular el tamaño de la muestra.

Las Figuras 1 y 2 también muestran que las curvas de potencia teórica (aproximada) están generalmente más elevadas que las curvas de potencia simulada. Las curvas de potencia aproximada están más elevadas debido a que los niveles de la potencia teórica se calculan asumiendo un valor exacto para el nivel de significancia objetivo (0.05). En comparación, la prueba exacta tiende a ser conservadora, particularmente con muestras pequeñas, y por lo tanto produce niveles de significancia mayores que el nivel objetivo. Como resultado, los niveles de potencia simulada tienden a ser más pequeños cuando los tamaños de las muestras son pequeños.

En conclusión, nuestras simulaciones demuestran que la función de potencia teórica de la prueba de aproximación a la normal se aproxima estrechamente a la potencia de la prueba exacta (tasa de verosimilitud). Como resultado, la función de potencia teórica de la prueba de aproximación a la normal proporciona una base razonable para calcular los tamaños de las muestras necesarios para asegurar que la prueba exacta tenga suficiente potencia para detectar diferencias prácticamente importantes.