Pruebas de Hipótesis (Dos Muestras)

Este procedimiento prueba hipótesis acerca de cualquiera de los siguientes parámetros:

1. la diferencia entre las **medias** μ_1 y μ_2 de dos distribuciones normales.
2. el radio de la **desviación estándar** σ_1 y σ_2 de dos distribuciones normales.
3. la diferencia entre las **proporciones** θ_1 y θ_2 de dos distribuciones binomiales.
4. la diferencia entre las **tasas** λ_1 y λ_2 de dos distribuciones Poisson.

El desplegado contiene:

1. el estadístico de prueba calculado junto a un P-Valor.
2. un *intervalo de confianza* para la diferencia o radio.
3. una gráfica de la potencia de la prueba de hipótesis.

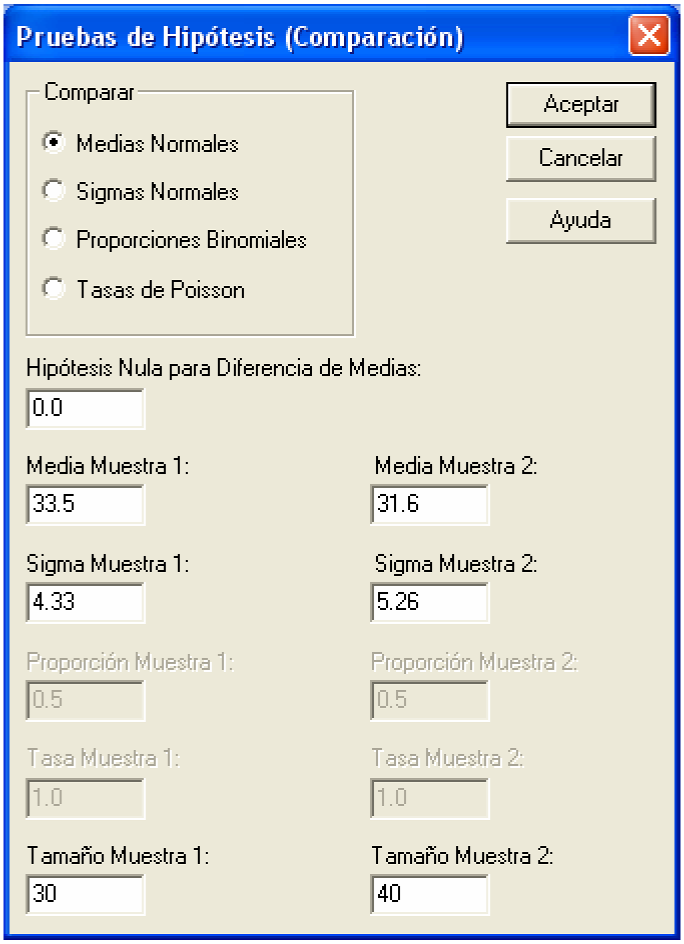
StatFolio Muestra: *hyptest2.sgp*

Datos de Muestra:

Ninguno.

Entrada de Datos

El cuadro de diálogo de entrada de datos requiere información sobre la prueba a realizarse.



- **Comparar:** la comparación a hacerse. Se asume que se han colectado dos muestras de la distribución indicada.
- **Hipótesis Nula**: el valor de la diferencia o radio Δ_0 especificada por la hipótesis nula.
- **Media Muestra 1:** la media de la primera muestra $\bar{x}_1$.
- **Media Muestra 2:** la media de la segunda muestra $\bar{x}_2$.
- **Sigma Muestra 1**: la desviación estándar de la primera muestra s_1.
- **Sigma Muestra 2**: la desviación estándar de la segunda muestra s_2.
- **Proporción Muestra 1**: la proporción calculada a partir de la primera muestra $\hat{\theta}_1$.
- **Proporción Muestra 2**: la proporción calculada a partir de la segunda muestra $\hat{\theta}_2$.
- **Tasa Muestra 1**: la tasa en la primera muestra $\hat{\lambda}_1$.
- **Tasa Muestra 2**: la tasa en la segunda muestra $\hat{\lambda}_2$.

- **Tamaño Muestra 1**: el tamaño de la primera muestra n_1.

- **Tamaño Muestra 2**: el tamaño de la segunda muestra n_2.

Por ejemplo, el cuadro de diálogo de abajo indica que se tomaron dos muestras de distribuciones normales, con los siguientes resultados:

	Muestra 1	*Muestra 2*
Tamaño	30	40
Media	33.5	31.6
Desviación estándar	4.33	5.26

Se desea determinar si las medias de las poblaciones de donde se tomaron los datos difieren significativamente.

Resumen del Análisis

El *Resumen del Análisis* despliega un intervalo de confianza para la diferencia o radio de interés, además de los resultados de la prueba de hipótesis.

Prueba de Hipótesis (Dos Muestras)
Medias muestrales = 33.5 y 31.6
Desviaciones estándar muestrales = 4.33 y 5.26
Tamaños de muestra = 30 y 40

Intervalos de confianza del 95.0% para la diferencia entre medias: 1.9 +/- 2.35438 [-0.454378,4.25438]

Hipótesis Nula: diferencia entre medias = 0.0
Alternativa: no igual
Estadístico t calculado = 1.61036
Valor-P = 0.111952
No rechazar la hipótesis nula para alfa = 0.05.

(Asumiendo varianzas iguales).

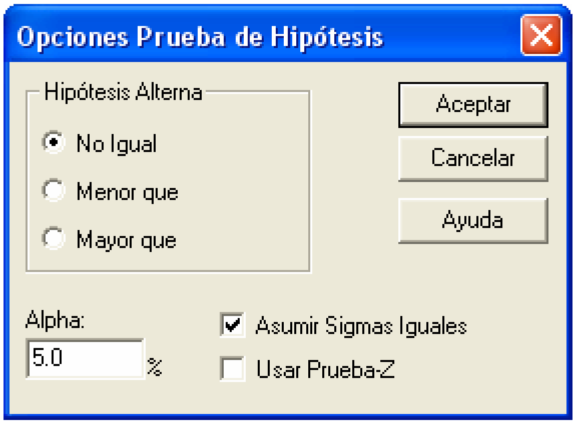
En este ejemplo, el intervalo de confianza del 95% establece que la diferencia entre las medias poblacionales $\mu_1 - \mu_2$ puede estar entre -0.454 y 4.254. También despliega una prueba formal de la hipótesis:

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_A: \mu_1 - \mu_2 \neq 0$$

El *P-Valor* mide la verosimilitud de obtener una diferencia igual a la observada si la hipótesis nula fuese cierta. A un nivel de significancia del 5% rechazaríamos la hipótesis nula sólo si $P < 0.05$. En este caso, el P-Valor está muy por encima de 0.05, así que la hipótesis nula no se rechaza, indicando que *no* hay diferencia significativa entre las dos medias.

En el caso de comparar dos medias normales, usted tiene la opción de asumir que las dos muestras vinieron de poblaciones con varianzas iguales, en cuyo caso se realiza una t-prueba exacta, o no hacer esa suposición y correr una prueba aproximada. Tal vez usted quiera correr una F-prueba para comparar las dos desviaciones estándar antes de tomar esa decisión.

Opciones de Análisis



- **Hipótesis Alterna:** selecciona la hipótesis alterna de interés.

No Igual	H_A: $\mu_1 - \mu_2 \neq 0.5$
Menor que	H_A: $\mu_1 - \mu_2 < 0.5$
Mayor que	H_A: $\mu_1 - \mu_2 > 0.5$

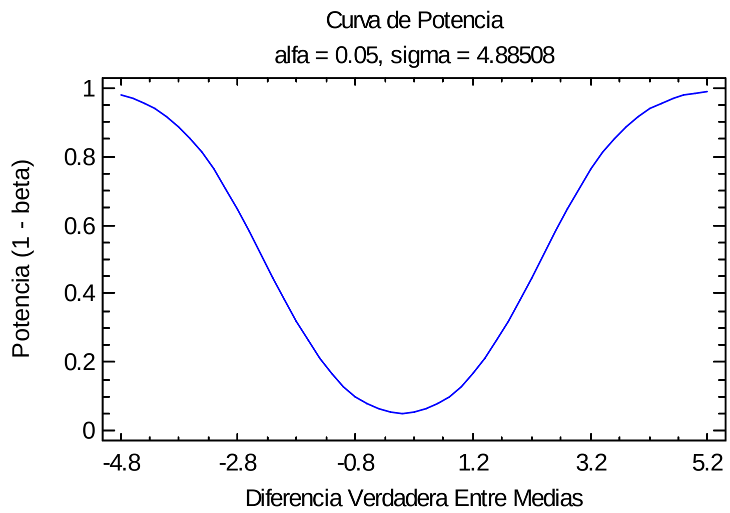
- **Alfa:** la probabilidad de un error tipo I (rechazar una hipótesis nula cierta). Esto afecta la conclusión sentada en el resultado, no el estadístico de prueba ni el P-Valor.

- **Asumir Sigmas Iguales:** al comparar dos medias, asume que las muestras vienen de poblaciones con la misma desviación estándar.

- **Usar Prueba–Z**: si está seleccionada, corre una prueba Z al probar una media normal. De otro manera se realiza una prueba t.

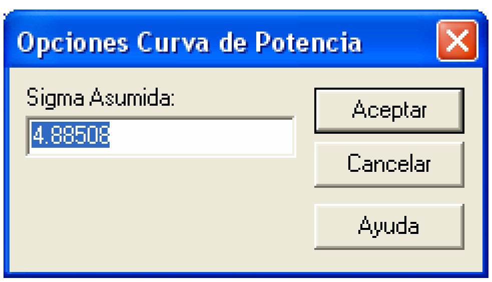
Curva de Potencia

La *Curva de Potencia* muestra la potencia de la prueba de hipótesis específica.



Potencia se define como la probabilidad de rechazar correctamente la hipótesis nula cuando esta es falsa. Esta prueba actual tiene una probabilidad de aproximadamente 70% de rechazar correctamente la hipótesis nula si la verdadera diferencia absoluta entre las poblaciones es menor a 3.0.

Panel de Opciones



- **Sigma Asumida:** si se prueba una hipótesis concerniente a una media normal, el valor asumido de la desviación estándar de la población.

Cálculos

- **Medias Normales – Intervalo de Confianza**

Si se asume que σ_1 y σ_2 son conocidas:

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \tag{1}$$

Si σ_1 y σ_2 se estiman a partir de los datos y se asume que son iguales:

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,\nu}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \tag{2}$$

donde

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \tag{3}$$

y

$$\nu = n_1 + n_2 - 2 \tag{4}$$

Si σ_1 y σ_2 se estiman a partir de los datos y no se asume que sean iguales:

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,m}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \tag{5}$$

donde

$$\frac{1}{m} = \frac{c^2}{n_1 - 1} + \frac{(1-c)^2}{n_2 - 1} \tag{6}$$

y

$$c = \frac{s_1^2 / n_1}{s_1^2 / n_1 + s_2^2 / n_2} \tag{7}$$

- **Medias Normales – Prueba de Hipótesis**

Si se asume que σ_1 y σ_2 son conocidas:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad (8)$$

Si σ_1 y σ_2 se estiman a partir de los datos y se asume que son iguales:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_v \qquad (9)$$

Si σ_1 y σ_2 se estiman a partir de los datos y no se asume que sean iguales:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_m \qquad (10)$$

- **Sigmas Normales – Intervalo de Confianza**

$$\left[\frac{s_1^2}{s_2^2}\frac{1}{F_{\alpha/2,n_1-1,n_2-1}}, \frac{s_1^2}{s_2^2}F_{\alpha/2,n_2-1,n_1-1}\right] \qquad (11)$$

- **Sigmas Normales – Prueba de Hipótesis**

$$F = \frac{s_2^1 / s_2^2}{\Delta_0} \sim F_{n_1-1,n_2-1} \qquad (12)$$

- **Proporciones Binomiales – Intervalo de Confianza**

$$\hat{\Delta} \pm z_{\alpha/2}\sqrt{\hat{\theta}_1(1-\hat{\theta}_1)/n_1 + \hat{\theta}_2(1-\hat{\theta}_2)/n_2} \qquad (13)$$

- **Proporciones Binomiales – Prueba de Hipótesis**

Si $\Delta_0 = 0$:

$$z = \hat{\Delta} / \sqrt{\hat{\theta}(1-\hat{\theta})/n_1 + \hat{\theta}\ (1-\hat{\theta}\ )/n_2} \qquad (14)$$

donde

$$\hat{\theta} = \left(n_1\hat{\theta}_1 + n_2\hat{\theta}_2\right)/\left(n_1 + n_2\right) \qquad (15)$$

If $\Delta_0 \neq 0$:

$$z = \left(\hat{\Delta} - \Delta_0\right) / \sqrt{\hat{\theta}_1 (1-\hat{\theta}_1)/n_1 + \hat{\theta}_2 (1-\hat{\theta}_2)/n_2} \tag{16}$$

- **Tasas de Poisson – Intervalo de Confianza**

$$\hat{\Delta} \pm z_{\alpha/2} \sqrt{\hat{\lambda}_1 / T_1 + \hat{\lambda}_2 / T_2} \tag{17}$$

donde el periodo de muestra $T_1 = n_1$ y $T_2 = n_2$.

- **Tasas de Poisson – Prueba de Hipótesis**

Si $\Delta_0 = 0$:

$$z = \hat{\Delta} / \sqrt{\hat{\lambda} / T_1 + \hat{\lambda} / T_2} \tag{18}$$

donde

$$\hat{\lambda} = \left(T_1 \hat{\lambda}_1 + T_2 \hat{\lambda}_2\right) / \left(T_1 + T_2\right) \tag{19}$$

Si $\Delta_0 \neq 0$:

$$z = \left(\hat{\Delta} - \Delta_0\right) / \sqrt{\hat{\lambda}_1 / T_1 + \hat{\lambda}_2 / T_2} \tag{20}$$