

Conceptos sobre una prueba de hipótesis

- **Hipótesis:** método científico basado en una muestra de datos para evaluar conjeturas sobre la población.
- **Hipótesis nula:** los valores que se conjetura puede tomar un parámetro, que incluye (pero no se limita a eso) la igualdad entre el valor conjeturado y el parámetro probado. Usualmente se denota como

$$H_0 : \text{Valor hipotético del parámetro} (\leq = \geq)$$

- **Hipótesis Alternativa:** el valor conjeturado para el parámetro, que es mutuamente excluyente y colectivamente exhaustivo con respecto a la **Hipótesis Nula** y por lo tanto incluye la relación ($\geq \neq \leq$) entre el valor conjeturado y el parámetro probado. Usualmente se denota como

$$H_1 : \text{Valor hipotético del parámetro} (\geq \neq \leq)$$

Inferencia sobre el vector de medias

Obsérvese que nuestra conclusión se realiza con respecto a la **hipótesis nula**, es decir que puede ser rechazada o no rechazada - es decir, nunca aceptamos la hipótesis nula.

- **Región Crítica:** Área que contiene todos los posibles valores estimados del parámetro para el cual se rechaza la hipótesis nula.
- **Valor (es) crítico (s):** Los valores que separan la región (es) crítica (s) de la región de no rechazo.
- **Estadístico de prueba:** Valor basado en la muestra que se comparará con la región o regiones críticas para decidir si se rechazará o no la hipótesis nula. La forma genérica del estadístico de prueba es

$$\text{Estadístico de prueba} = \frac{\text{Valor observado del estadístico de prueba} - \text{Valor hipotético del estadístico}}{\text{Error estándar del estadístico de prueba}}$$

Inferencia sobre el vector de medias

- **Regla de decisión:** declaración que especifica los valores del estadístico de prueba que dará como resultado el rechazo o no rechazo de la hipótesis nula
- **Prueba de Hipótesis de dos Colas:** Evaluación de una conjetura para la cual los resultados de la muestra son **suficientemente menores o mas grande que** el valor conjeturado del parámetro, lo cual implicará el rechazo de la hipótesis nula. Es decir, para una hipótesis nula que sólo incluye la igualdad entre el valor conjeturado y el parámetro probado.
- **Prueba de hipótesis de una cola:** Evaluación de una conjetura para la cual los resultados de la muestra son sólo suficientemente menores o mayores que el valor conjeturado del parámetro, que resultará en el rechazo de la hipótesis nula. Es decir, para una hipótesis nula que incluye una desigualdad entre el valor conjeturado y el parámetro probado.

- **Prueba de cola superior:** Prueba de hipótesis para la cual los resultados de la muestra que sólo son suficientemente mayores que el valor conjeturado del parámetro resultarán en el rechazo de la hipótesis nula
- **Prueba de cola inferior:** Prueba de hipótesis para la cual los resultados de la muestra que sólo son suficientemente menores que el valor conjeturado del parámetro resultarán en el rechazo de la hipótesis nula

Inferencia sobre el vector de medias

- **Error Tipo I:** Rechazo de la hipótesis nula cuando ésta es verdadera. La probabilidad de esta ocurrencia (dado que la hipótesis nula es verdadera) se denomina como α .
- **Error Tipo II:** No rechazo de la hipótesis nula cuando esta es falsa. La probabilidad de esta ocurrencia (dado que la hipótesis nula es falsa) se denomina como β

		Situación real	
		H_0 es verdad	H_0 es Falsa
C o n c l u s i o n	No se Rechaza H_0	Decisión correcta	Probabilidad del error Tipo II = β
	Reject H_0	Probabilidad del error Tipo I = α	Decisión Correcta

Pasos para probar una Hipótesis

- 1 Definir las hipótesis nula y alternativa
- 2 Seleccionar el estadístico de prueba adecuado
- 3 Definir el nivel de significancia deseado, α , encontrar los valores críticos y dar la regla de decisión
- 4 Calcular el estadístico de prueba
- 5 Utilizar la regla de decisión para evaluar el estadística de prueba y decidir si rechazar o no la hipótesis nula. Interpretar los resultados

Prueba de hipótesis sobre la media en la teoría univariada

Al probar una hipótesis sobre una sola media, el estadístico de prueba apropiado (cuando la población de origen es normal o la muestra es suficientemente grande) es

$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

donde t sigue una distribución t *student* con $n - 1$ grados de libertad y donde

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j, \quad s^2 = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2$$

Ejemplo: Prueba de hipótesis para una media

Suponga que tenemos las siguientes 15 observaciones muestrales de una variable aleatoria x_1

x_{j1}
5.76
6.68
6.79
7.88
2.46
2.48
2.97
4.47
1.62
1.43
7.46
8.92
6.61
4.03
9.42

En un nivel de significancia de $\alpha = 0.10$, ¿estos datos apoyan la afirmación de que fueron extraídos de una población con una media de 4.0?

Ejemplo: Prueba de hipótesis para una media

- Las Hipótesis Nula y Alternativa son:

$$H_0 : \mu_1 = 4.0$$

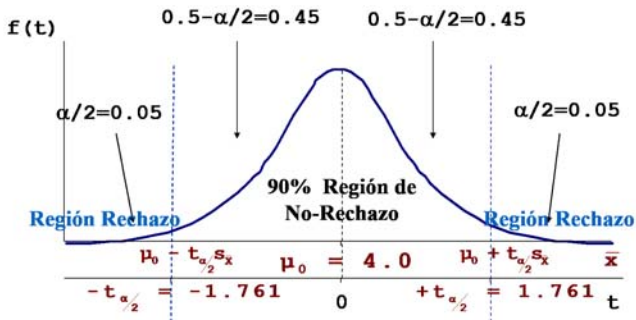
$$H_1 : \mu_1 \neq 4.0$$

- Se selecciona el estadístico de prueba adecuado, aunque $n = 15 < 30$, los datos parecen ser normales, por tanto usamos el estadístico de prueba t que sigue una distribución *t student*

$$t = \frac{\bar{x}_1 - \mu_0}{s_{\bar{x}_1}} = \frac{\bar{x}_1 - \mu_0}{s/\sqrt{n}}$$

Ejemplo: Prueba de hipótesis para una media

- El nivel de significancia es $\alpha = 0.10$, tenemos una prueba de dos colas, así que $t_{\alpha/2} = \pm 1.761$.



Regla de decisión: no rechazar H_0 si $-1.761 \leq t \leq 1.761$,
de lo contrario rechazar H_0

Ejemplo: Prueba de hipótesis para una media

- Calculamos el estadístico de prueba t
Tenemos que

$$\bar{x}_1 = 5.26, \mu_0 = 4.0, s_1^2 = s_{11} = 7.12 \rightarrow s_1 = 2.669$$

$$t = \frac{\bar{x}_1 - \mu_0}{s_{\bar{x}_1}} = \frac{5.26 - 4.0}{2.669/\sqrt{15}} = 1.84$$

- Utilizamos la regla de decisión para evaluar el estadístico de prueba y decidir si rechazamos o no la hipótesis nula

$$-1.761 \leq t \not\leq 1.761, \text{ i.e., } -1.761 \leq 1.84 \not\leq 1.761$$

- Por lo tanto se rechaza H_0 con $\alpha = 0.10$. La muestra de datos no nos da una evidencia para apoyar la afirmación de que la media de x_1 es 4.0.

Ejemplo 2: Prueba de hipótesis para una media

Supongamos que tenemos una muestra de quince observaciones tomadas de alguna variable aleatoria x_2

x_{j2}
-3.97
-3.24
-3.56
-1.87
-1.13
-5.20
-6.39
-7.88
-5.00
-0.69
1.61
-6.60
2.32
2.87
-7.64

En un nivel de significancia de $\alpha = 0.10$, ¿estos datos apoyan la afirmación de que fueron extraídos de una población con una media de -1.5?

En otras palabras, queremos probar la hipótesis nula

$$H_0 : \mu_1 = \mu_0 \text{ vs } H_1 : \mu_1 \neq \mu_0$$

$$H_0 : \mu_1 = -1.5 \text{ vs } H_1 : \mu_1 \neq -1.5$$

Ejemplo 2: Prueba de hipótesis para una media

- Las Hipótesis Nula y Alternativa son:

$$H_0 : \mu_1 = -1.5$$

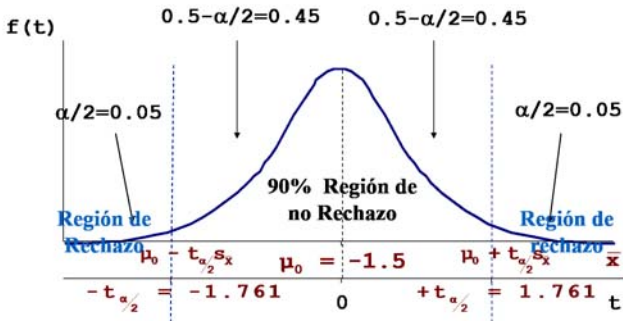
$$H_1 : \mu_1 \neq -1.5$$

- Seleccionamos el estadístico de prueba adecuado, $n = 15 < 30$, pero los datos parecen seguir una distribución normal, entonces utilizamos el estadístico de prueba t que sigue una distribución t de *student*

$$t = \frac{\bar{x}_2 - \mu_0}{s_{\bar{x}_2}} = \frac{\bar{x}_2 - \mu_0}{s/\sqrt{n}}$$

Ejemplo 2: Prueba de hipótesis para una media

- El nivel de significancia es $\alpha = 0.10$, tenemos una prueba de dos colas, así que $t_{\alpha/2} = \pm 1.761$



Regla de decisión: no rechazar H_0 si $-1.761 \leq t \leq 1.761$,
de lo contrario rechazar H_0

Ejemplo 2: Prueba de hipótesis para una media

- Calculamos el estadístico de prueba
Tenemos que

$$\bar{x}_2 = -3.09, \mu_0 = -1.5, s_2^2 = s_{22} = 12.43 \rightarrow s_2 = 3.526$$

Así

$$t = \frac{\bar{x}_2 - \mu_0}{s_{\bar{x}_2}} = \frac{-3.09 - (-1.5)}{3.526/\sqrt{15}} = -1.748$$

- Utilizamos la regla de decisión para evaluar el estadístico de prueba y decidir si rechazamos o no la hipótesis nula,

$$-1.761 \leq t \leq 1.761, \text{ i.e. } -1.761 \leq -1.748 \leq 1.761$$

- Por tanto no rechazamos H_0 con $\alpha = 0.10$. La muestra tiene evidencia para no rechazar la afirmación de que la media de x_2 es -1.5.

- Considerando la prueba de dos colas, nótese que rechazar H_0 cuando $|t|$ es grande, equivale a rechazar H_0 si su cuadrado

$$t^2 = \frac{(\bar{x} - \mu_0)^2}{s^2/n} = n(\bar{x} - \mu_0)(s^2)^{-1}(\bar{x} - \mu_0)$$

es grande

- Obsérvese que t^2 representa la distancia desde la media de la muestra, \bar{x} , a la media hipotética de la población, μ_0 , expresada en términos de $s_{\bar{x}}$ (lo cual es equivalente a la distancia cuadrada generalizada de \bar{x} a μ para el caso multivariado)
- Cuando el \bar{x} y s^2 son observados, entonces se rechaza H_0 en favor de H_1 , en un nivel de significancia α , si

$$n(\bar{x} - \mu_0)(s^2)^{-1}(\bar{x} - \mu_0) > t_{n-1}^2(\alpha/2)$$

- Si la hipótesis nula es verdadera, $t^2 \sim F_{1,n-1}$.

Inferencia sobre un vector de medias μ

Una generalización natural de la distancia cuadrada univariada t^2 al caso multivariado está dada por T^2 de Hotelling

$$T^2 = (\bar{\mathbf{x}} - \mu_0)' \left(\frac{1}{n} \mathbf{S} \right)^{-1} (\bar{\mathbf{x}} - \mu_0) = n(\bar{\mathbf{x}} - \mu_0)' \mathbf{S}^{-1} (\bar{\mathbf{x}} - \mu_0)$$

donde

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{j=1}^n \mathbf{x}_j = \frac{1}{n} \mathbf{X}' \mathbf{1} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix}, \mu_0 = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$$

$$\mathbf{S} = \frac{1}{n-1} \sum_{j=1}^n (\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})' = \frac{1}{n-1} \mathbf{X}' \left(\mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}' \right) \mathbf{X} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{12} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{1p} & s_{2p} & \cdots & s_{pp} \end{bmatrix}$$

note que $n^{-1} \mathbf{S}$ es la matriz de covarianza estimada de $\bar{\mathbf{x}}$.

Inferencia sobre un vector de medias μ

- Esto nos da un marco de referencia para probar hipótesis sobre el vector de medias, donde las hipótesis nula y alternativa son,

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

- El estadístico T^2 se puede reescribir como:

$$T^2 = \underbrace{\sqrt{n}(\bar{x} - \mu_0)'}_{\text{Variable aleatoria Normal multivariada } N(0, \Sigma)} \underbrace{\left(\frac{\sum_{j=1}^n (x_j - \bar{x})(x_j - \bar{x})'}{n-1} \right)^{-1}}_{\text{Matriz aleatoria Wishart } W_{p, n-1}(\Sigma)/g.l.} \underbrace{\sqrt{n}(\bar{x} - \mu_0)}_{\text{Variable aleatoria Normal m}}$$

- Por lo tanto, cuando la hipótesis nula es verdadera, el estadístico T^2 puede escribirse como el producto de dos distribuciones normales multivariadas $N_p(\mu, \Sigma)$ y una distribución de Wishart $W_{p, n-1}(\Sigma)$,

Inferencia sobre un vector de medias μ

- Lo anterior es una generalización completa del caso univariante, donde podríamos escribir el estadístico de prueba t^2 como

$$T^2 = \underbrace{\sqrt{n}(\bar{\mathbf{X}} - \mu_0)'}_{\text{Variable aleatoria Normal univariada } N(0, s^2)} \underbrace{(s)^{-1}}_{\text{Variable aleatoria chi cuadrada/g.l.}} \underbrace{\sqrt{n}(\bar{\mathbf{x}} - \mu_0)}_{\text{Variable aleatoria Normal univariada}}$$

- Entonces para probar la hipótesis $H_0 : \mu = \mu_0 \text{ Vs } H_1 : \mu \neq \mu_0$, en un nivel de significancia α , rechazamos H_0 en favor de H_1 si el valor observado

$$T^2 = n(\bar{\mathbf{x}} - \mu_0)' \mathbf{S}^{-1}(\bar{\mathbf{x}} - \mu_0) > \text{percentil de una } T^2$$

- Podríamos usar tablas especiales para encontrar los valores críticos de T^2 para diferentes combinaciones de α y grados de libertad, pero esto no es necesario porque

$$T^2 \sim \frac{(n-1)p}{n-p} F_{p, n-p}$$

Ejemplo: Inferencia sobre un vector de medias μ

Suponga que tenemos las siguientes 15 observaciones muestrales obtenidas de variables aleatorias bivariadas (x_1, x_2)

x_{j1}	x_{j2}
1.43	-0.69
1.62	-5.00
2.46	-1.13
2.48	-5.20
2.97	-6.39
4.03	2.87
4.47	-7.88
5.76	-3.97
6.61	2.32
6.68	-3.24
6.79	-3.56
7.46	1.61
7.88	-1.87
8.92	-6.60
9.42	-7.64

En un nivel de significancia de $\alpha = 0.10$, ¿estos datos apoyan la afirmación de que fueron extraídos de una población con un centroide $(4.0, -1.5)$?

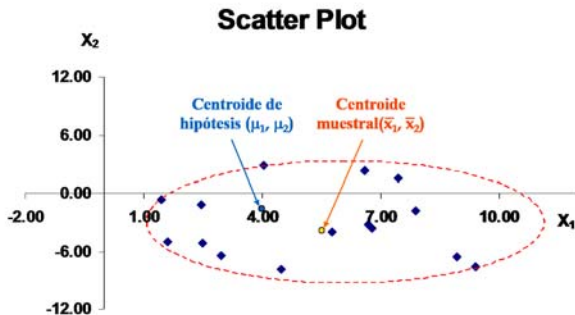
Entonces, queremos probar la hipótesis nula

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu \neq \mu_0$$

$$H_0 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix} \text{ vs } H_1 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

Ejemplo: Inferencia sobre un vector de medias μ

- El diagrama de dispersión de los pares (x_1, x_2) , el centroide muestral (\bar{x}_1, \bar{x}_2) y centroide de la hipótesis $\mu_0 = (\mu_1, \mu_2)_0$ se muestran a continuación:



- Estos datos parecen apoyar nuestra hipótesis nula?

Ejemplo: Inferencia sobre un vector de medias μ

- Las Hipótesis Nula y Alternativa son

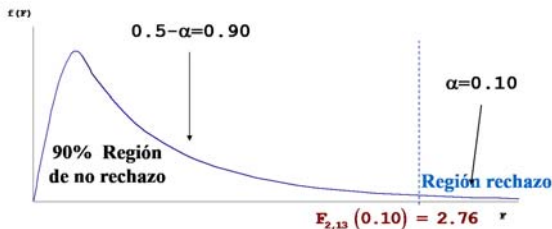
$$H_0 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix} \text{ vs } H_1 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

- Seleccionamos el estadístico de prueba adecuado,
 $n - p = 15 - 2 = 13$ lo cual no es muy grande, pero los datos parecen aproximarse una normal bivariada, así que utilizamos

$$T^2 = n(\bar{\mathbf{x}} - \mu_0)' \mathbf{S}^{-1}(\bar{\mathbf{x}} - \mu_0)$$

Ejemplo: Inferencia sobre un vector de medias μ

- El nivel de significancia dado es $\alpha = 0.10$, $p = 2$, $n - p = 13$ son los grados de libertad de F , entonces $F_{2,13}(0.10) = 2.76$



- Pero todavía no tenemos una regla de decisión pues

$$T^2 \sim \frac{(n-1)p}{n-p} F_{p, n-p}$$

Ejemplo: Inferencia sobre un vector de medias μ

- Así, nuestro valor crítico es

$$\frac{(n-1)p}{n-p} F_{p, n-p}(\alpha) = \frac{(15-1)2}{15-2} F_{2, 15-2}(0.10) = 5.95$$

- Entonces Regla de decisión es: no rechazar H_0 si $T^2 \leq 5.95$, de otra manera se rechaza H_0

Ejemplo: Inferencia sobre un vector de medias μ

- Calculamos el estadístico de prueba T^2
Tenemos que

$$\bar{\mathbf{x}} = \begin{bmatrix} 5.26 \\ -3.09 \end{bmatrix}, \mu_0 = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} 7.12 & -0.72 \\ -0.72 & 12.43 \end{bmatrix} \longrightarrow \mathbf{S}^{-1} = \begin{bmatrix} 0.1412 & 0.0082 \\ 0.0082 & 0.0809 \end{bmatrix}$$

Entonces

$$\begin{aligned} T^2 &= n(\bar{\mathbf{x}} - \mu_0)' \mathbf{S}^{-1} (\bar{\mathbf{x}} - \mu_0) \\ &= 15 \begin{bmatrix} 5.26 - 4.0 \\ -3.09 + 1.5 \end{bmatrix}' \begin{bmatrix} 7.12 & -0.72 \\ -0.72 & 12.43 \end{bmatrix}^{-1} \begin{bmatrix} 5.26 - 4.0 \\ -3.09 + 1.5 \end{bmatrix} \\ &= 15 \begin{bmatrix} 1.26 & -1.59 \end{bmatrix} \begin{bmatrix} 0.1412 & 0.0082 \\ 0.0082 & 0.0809 \end{bmatrix} \begin{bmatrix} 1.26 \\ -1.59 \end{bmatrix} = 5.970 \end{aligned}$$

Ejemplo: Inferencia sobre un vector de medias μ

- Utilizamos la regla de decisión para evaluar el estadístico de prueba y decidimos si se rechaza o no se rechaza la hipótesis nula. En este caso

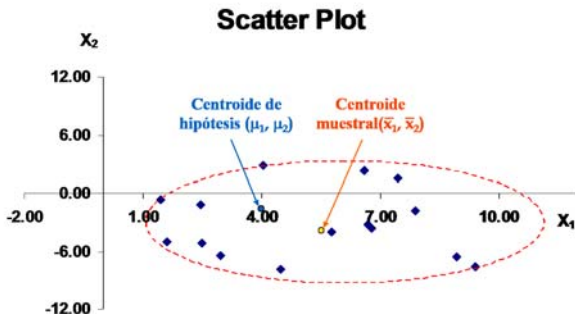
$$T^2 = 5.970 \not\leq 5.95$$

- Por lo tanto rechazamos H_0 con $\alpha = 0.10$. La muestra tiene evidencia para apoyar la afirmación de que el vector de medias μ difiere de

$$\mu_0 = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

Ejemplo: Inferencia sobre un vector de medias μ

- ¿Por qué los resultados de las pruebas univariadas originales y la prueba bivariada son tan consistentes? ¡Revisa los datos!



- ¿Qué sugieren los datos sobre la relación entre x_1 y x_2 ?

Ejemplo: Inferencia sobre un vector de medias μ

- ¿Qué pasaría si tomamos los mismos valores de x_1 y x_2 pero cambiamos los pares de valores?

x_{j1}	x_{j2}
4.03	-1.87
6.61	-3.97
7.46	-6.39
9.42	-7.88
6.79	-5.20
1.43	2.87
8.92	-7.64
1.62	2.32
2.48	-0.69
5.76	-3.56
7.88	-6.60
4.47	-3.24
2.46	1.61
2.97	-1.13
6.68	-5.00

Por supuesto, los resultados de las pruebas univariadas no cambiarían, pero a un nivel de significancia de $\alpha = 0.10$, ¿estos datos apoyan ahora la afirmación de que fueron extraídos de una población con un centroide $(4, 0, -1, 5)$?

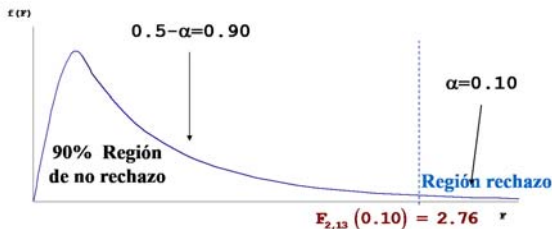
Es decir, probamos de nuevo la hipótesis nula

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu \neq \mu_0$$

$$H_0 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix} \text{ vs } H_1 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

Ejemplo: Inferencia sobre un vector de medias μ

- Ya que estamos considerando el mismo nivel de significancia ($\alpha = .10$), y los mismos grados de libertad, el valor crítico y regla de decisión no cambian



- Entonces nuestro valor crítico es

$$\frac{(n-1)p}{n-p} F_{p,n-p}(\alpha) = \frac{(15-1)2}{15-2} F_{2,15-2}(0.10) = 5.95$$

- Entonces la regla de decisión: no rechazar H_0 si $T^2 \leq 5.95$, de otra manera se rechaza H_0

Ejemplo: Inferencia sobre un vector de medias μ

- Sin embargo, tendremos que recalcular el estadístico de prueba (¿por qué?)

Tenemos que

$$\bar{\mathbf{x}} = \begin{bmatrix} 5.26 \\ -3.09 \end{bmatrix}, \mu_0 = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} 7.12 & -9.19 \\ -9.19 & 12.43 \end{bmatrix} \longrightarrow \mathbf{S}^{-1} = \begin{bmatrix} 3.0374 & 2.2452 \\ 2.2452 & 1.74 \end{bmatrix}$$

Entonces

$$\begin{aligned} T^2 &= n(\bar{\mathbf{x}} - \mu_0)' \mathbf{S}^{-1} (\bar{\mathbf{x}} - \mu_0) \\ &= 15 \begin{bmatrix} 5.26 - 4.0 \\ -3.09 + 1.5 \end{bmatrix}' \begin{bmatrix} 7.12 & -9.19 \\ -9.19 & 12.43 \end{bmatrix}^{-1} \begin{bmatrix} 5.26 - 4.0 \\ -3.09 + 1.5 \end{bmatrix} \\ &= 15 \begin{bmatrix} 1.26 & -1.59 \end{bmatrix} \begin{bmatrix} 3.0374 & 2.2452 \\ 2.2452 & 1.74 \end{bmatrix} \begin{bmatrix} 1.26 \\ -1.59 \end{bmatrix} = 3.411 \end{aligned}$$

Ejemplo2_Inferencia sobre un vector de medias μ

- Utilizando la regla de decisión evaluamos el estadístico de prueba, para decidir si rechazamos o no la hipótesis nula. En este caso

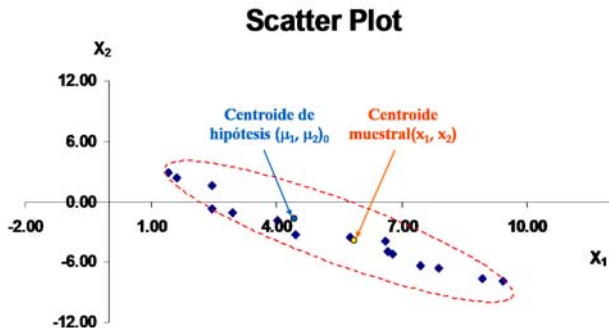
$$T^2 = 3.411 \leq 5.95$$

- Por lo que se no rechaza H_0 con $\alpha = 0.10$. La muestra no apoya la afirmación de que el vector de medias difiere de

$$\mu_0 = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

Ejemplo: Inferencia sobre un vector de medias μ

- ¿Por qué ahora no rechazamos en el caso bivariado? ¡De nuevo, Revisa los datos !



- ¿Qué sugieren los datos sobre la relación entre x_1 y x_2 ?

Ejemplo: Inferencia sobre un vector de medias μ

- ¿Qué pasaría si tomamos los mismos valores de x_1 y x_2 pero cambiamos de nuevo los pares de valores para que su correlación sea positiva

x_{j1}	x_{j2}
2.46	-6.60
5.76	-3.66
1.62	-7.64
4.03	-6.00
7.46	-0.69
9.42	2.87
8.92	2.32
6.68	-1.87
7.88	1.61
6.61	-3.24
1.43	-7.88
4.47	-3.97
6.79	-1.13
2.97	-6.20
2.48	-6.39

De nuevo, los resultados de las pruebas univariadas no cambiarían, pero a un nivel de significancia de $\alpha = 0.10$, ¿estos datos apoyan ahora la afirmación de que fueron extraídos de una población con un centroide $(4,0, -1,5)$?

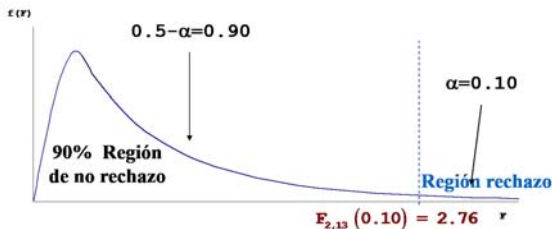
En otras palabras, probamos de nuevo la hipótesis nula

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu \neq \mu_0$$

$$H_0 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix} \text{ vs } H_1 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

Ejemplo: Inferencia sobre un vector de medias μ

- Ya que estamos considerando el mismo nivel de significancia ($\alpha = .10$), y los mismos grados de libertad, el valor crítico y regla de decisión no cambian



- Entonces nuestro valor crítico es

$$\frac{(n-1)p}{n-p} F_{p,n-p}(\alpha) = \frac{(15-1)2}{15-2} F_{2,15-2}(0.10) = 5.95$$

- Regla de decisión: no rechazar H_0 si $T^2 \leq 5.95$, de otra manera se rechaza H_0

Ejemplo3_Inferencia sobre un vector de medias μ

- Sin embargo, tenemos que volver a calcular la estadística de prueba

Tenemos que

$$\bar{\mathbf{x}} = \begin{bmatrix} 5.26 \\ -3.09 \end{bmatrix}, \mu_0 = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} 7.12 & 9.16 \\ 9.16 & 12.43 \end{bmatrix} \longrightarrow \mathbf{S}^{-1} = \begin{bmatrix} 2.7407 & -2.0206 \\ -2.0206 & 1.5701 \end{bmatrix}$$

Entonces

$$\begin{aligned} T^2 &= n(\bar{\mathbf{x}} - \mu_0)' \mathbf{S}^{-1} (\bar{\mathbf{x}} - \mu_0) \\ &= 15 \begin{bmatrix} 5.26 - 4.0 \\ -3.09 + 1.5 \end{bmatrix}' \begin{bmatrix} 7.12 & 9.16 \\ 9.16 & 12.43 \end{bmatrix}^{-1} \begin{bmatrix} 5.26 - 4.0 \\ -3.09 + 1.5 \end{bmatrix} \\ &= 15 \begin{bmatrix} 1.26 & -1.59 \end{bmatrix} \begin{bmatrix} 2.7407 & -2.0206 \\ -2.0206 & 1.5701 \end{bmatrix} \begin{bmatrix} 1.26 \\ -1.59 \end{bmatrix} = 247.4 \end{aligned}$$

Ejemplo: Inferencia sobre un vector de medias μ

- Utilizando la regla de decisión evaluamos el estadístico de prueba, para decidir si rechazamos o no la hipótesis nula, en este caso

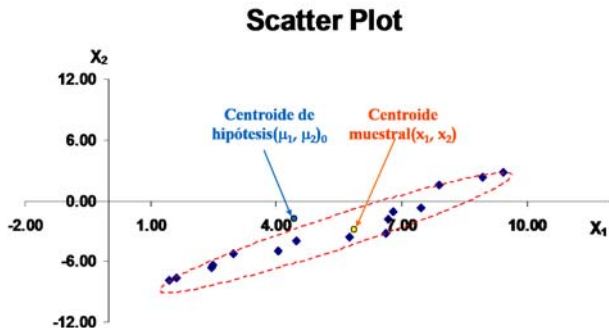
$$T^2 = 247.470 \not\leq 5.95$$

- Por lo que se rechaza H_0 con $\alpha = 0.10$. La muestra apoya la afirmación de que el vector de medias difiere de

$$\mu_0 = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

Ejemplo: Inferencia sobre un vector de medias μ

- ¿Por qué ahora la decisión es completamente diferente? ¡De nuevo, Revisa los datos!



- ¿Qué sugieren los datos sobre la relación entre x_1 y x_2 ?

Método de la razón de verosimilitud: principio general para la construcción de procedimientos para probar hipótesis:

- tienen varias propiedades óptimas para muestras grandes
- son particularmente convenientes para probar hipótesis formuladas en términos de parámetros de normales multivariadas.

Recordemos que el máximo de la función de verosimilitud normal multivariada (sobre todos valores posibles de μ y Σ) está dado por

$$\max_{\mu, \Sigma} L(\mu, \Sigma) = \frac{1}{(2\pi)^{np/2} |\hat{\Sigma}|^{n/2}} e^{-np/2}$$

T² de Hotelling y pruebas de razón de verosimilitud

Donde

$$\hat{\mu} = \bar{\mathbf{X}} = \frac{1}{n} \sum_{j=1}^n \mathbf{x}_j$$

$$\hat{\Sigma} = \frac{1}{n} \sum_{j=1}^n (\mathbf{x}_j - \bar{\mathbf{X}})(\mathbf{x}_j - \bar{\mathbf{X}})' = \frac{n-1}{n} \mathbf{S}$$

- Bajo $H_0 : \mu = \mu_0$, la verosimilitud normal se especifica como

$$L(\mu_0, \Sigma) = \frac{1}{(2\pi)^{np/2} |\Sigma|^{n/2}} e^{-\frac{1}{2} \sum_{j=1}^n (\mathbf{x}_j - \mu_0) \Sigma^{-1} (\mathbf{x}_j - \mu_0)'}$$

- Es decir, la media μ se fija a μ_0 , pero Σ puede variar hasta encontrar el valor que nos proporcione la probabilidad mas alta de obtener la muestra de datos.

- Este valor para Σ se obtiene maximizando $L(\mu_0, \Sigma)$ con respecto a Σ
- Por las propiedades de la traza, podemos reescribir el exponente en $L(\mu_0, \Sigma)$ de tal forma que

$$\begin{aligned} L(\mu_0, \Sigma) &= \frac{1}{(2\pi)^{np/2} |\Sigma|^{n/2}} e^{-\frac{1}{2} \sum_{j=1}^n \text{tr}[\Sigma^{-1}(\mathbf{x}_j - \mu_0)(\mathbf{x}_j - \mu_0)']} \\ &= \frac{1}{(2\pi)^{np/2} |\Sigma|^{n/2}} e^{-\frac{1}{2} \text{tr}[\Sigma^{-1} \sum_{j=1}^n (\mathbf{x}_j - \mu_0)(\mathbf{x}_j - \mu_0)']} \end{aligned}$$

- Recordemos un resultado anterior: Para una matriz simétrica $p \times p$ positiva definida \mathbf{B} y escalar $b > 0$, se tiene que

$$\frac{1}{|\mathbf{\Sigma}|^b} e^{-\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{B})/2} \leq \frac{1}{|\mathbf{B}|^b} (2b)^{pb} e^{-pb}$$

para toda $\mathbf{\Sigma}$, positiva definida, de dimensión $p \times p$, y donde la igualdad se cumple únicamente para

$$\mathbf{\Sigma} = \frac{1}{2b} \mathbf{B}$$

Pruebas de razón de verosimilitud

- Aplicando este resultado con

$$\mathbf{B} = \sum_{j=1}^n (\mathbf{x}_j - \mu_0)(\mathbf{x}_j - \mu_0)' \quad \text{y} \quad b = \frac{n}{2}$$

tenemos

$$\max_{\Sigma} L(\mu_0, \Sigma) = \frac{1}{(2\pi)^{np/2} |\hat{\Sigma}_0|^{n/2}} e^{-np/2}$$

donde

$$\hat{\Sigma}_0 = \frac{1}{n} \sum_{j=1}^n (\mathbf{x}_j - \mu_0)(\mathbf{x}_j - \mu_0)'$$

Pruebas de razón de verosimilitud

- Para determinar si μ_0 es un valor posible de μ , compararemos el máximo de $L(\mu_0, \Sigma)$ con el máximo de $L(\mu, \Sigma)$ cuando no se imponen restricciones sobre μ y Σ .
- La relación resultante se denomina Estadístico de la Razón de Verosimilitud:

$$\text{Razón de verosimilitud} = \Lambda = \frac{\max L(\mu_0, \Sigma)}{\max L(\mu, \Sigma)} = \frac{\frac{1}{(2\pi)^{np/2} |\hat{\Sigma}_0|^{n/2}} e^{-np/2}}{\frac{1}{(2\pi)^{np/2} |\hat{\Sigma}|^{n/2}} e^{-np/2}} = \left(\frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} \right)^{n/2}$$

O de manera equivalente

$$\Lambda_n^2 = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} = \frac{\left| \frac{1}{n} \sum_{j=1}^n (\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})' \right|}{\left| \frac{1}{n} \sum_{j=1}^n (\mathbf{x}_j - \mu_0)(\mathbf{x}_j - \mu_0)' \right|} = \frac{\left| \sum_{j=1}^n (\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})' \right|}{\left| \sum_{j=1}^n (\mathbf{x}_j - \mu_0)(\mathbf{x}_j - \mu_0)' \right|}$$

Λ_n^2 que se le denomina Lambda de Wilks

Relación de T^2 de Hotelling y la prueba de razón de verosimilitud

- Obsérvese que un valor pequeño de la lambda de Wilks sugiere que la hipótesis nula $H_0 : \mu = \mu_0$ no es probable que sea cierta (y conduce a su rechazo), es decir, rechazamos H_0 si

$$\Lambda^{\frac{2}{n}} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} = \frac{|\sum_{j=1}^n (\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})'|}{|\sum_{j=1}^n (\mathbf{x}_j - \mu_0)(\mathbf{x}_j - \mu_0)'|} < c_\alpha$$

donde c_α es el percentil inferior $(100\alpha)\%$ de la distribución Λ .

- En teoría debemos calcular los percentiles de la distribución Λ para obtener las regiones de rechazo y aceptación
- Sin embargo existe una relación entre T^2 y Λ que puede ayudarnos a calcularlos.

Relación de T^2 de Hotelling y la λ de Wilks

Resultado: Suponga $\mathbf{x}_1, \dots, \mathbf{x}_n$ es una muestra aleatoria de una población $N_p(\mu, \Sigma)$. Entonces la prueba basada en T^2 es equivalente a la prueba de la razón de verosimilitud para probar que $H_0 : \mu = \mu_0$ Vs $H_1 : \mu \neq \mu_0$ debido a que

$$\Lambda^{\frac{2}{n}} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} = \left(1 + \frac{T^2}{n-1}\right)^{-1}$$

- Por tanto sustituimos el valor crítico apropiado de T^2 en la ecuación anterior para encontrar el valor crítico de la razón de verosimilitud.
- Se rechaza H_0 para valores pequeños de $\Lambda^{\frac{2}{n}}$ o equivalentemente, para valores grandes de T^2

Relación de T^2 de Hotelling y la prueba de razón de verosimilitud

- La ecuación anterior tiene la ventaja de que T^2 se puede calcular como el cociente de dos determinantes (evitando así calcular \mathbf{S}^{-1}). Resolviendo para T^2 se obtiene

$$T^2 = \frac{(n-1) |\hat{\Sigma}_0|}{|\hat{\Sigma}|} - (n-1)$$
$$= \frac{(n-1) \left| \sum_{j=1}^n (\mathbf{x}_j - \mu_0)(\mathbf{x}_j - \mu_0)' \right|}{\left| \sum_{j=1}^n (\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})' \right|} - (n-1)$$

Ejemplo: Inferencia sobre la media mediante la prueba de razón de verosimilitud

Para nuestros datos bivariados anteriores, realizar la prueba de razón de verosimilitud para la hipótesis de que el centroide es (4,0, -1.5):

x_{j1}	x_{j2}
1.43	-0.69
1.62	-5.00
2.46	-1.13
2.48	-5.20
2.97	-6.39
4.03	2.87
4.47	-7.88
5.76	-3.97
6.61	2.32
6.68	-3.24
6.79	-3.56
7.46	1.61
7.88	-1.87
8.92	-6.60
9.42	-7.64

En otras palabras, probar la hipótesis nula

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu \neq \mu_0$$

$$H_0 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix} \text{ vs } H_1 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

Obtuvimos un valor del estadístico de prueba de T^2 de Hotelling de 5.997 y un valor crítico (en $\alpha = 0.10$) de 5.95.

Ejemplo: Inferencia sobre la media mediante la prueba de razón de verosimilitud

- Definimos la hipótesis nula y alternativa

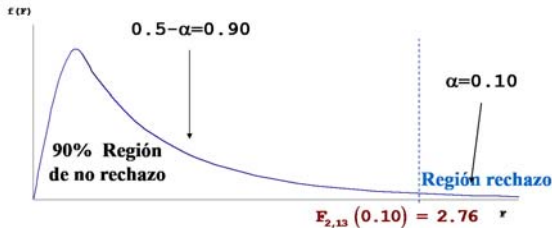
$$H_0 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix} \text{ vs } H_1 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

- Seleccionamos el estadístico de prueba adecuado, $n - p = 15 - 2 = 13$ lo cual no es muy grande, pero los datos parecen ser aproximadamente normales bivariados. Por tanto usamos

$$\Lambda^{\frac{2}{n}} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} = \left(1 + \frac{T^2}{n-1}\right)^{-1}$$

Ejemplo: Inferencia sobre la media mediante la prueba de razón de verosimilitud

- El nivel de significancia es $\alpha = 0.10$ y $p = 2, n - p = 13$ son los grados de libertad. Tenemos que $F_{2,13}(0.10) = 2.76$



- Pero no tenemos una regla de decisión pues

$$T^2 \sim \frac{(n-1)p}{n-p} F_{p,n-p}$$

Ejemplo: Inferencia sobre la media mediante la prueba de razón de verosimilitud

- Así, el valor crítico de T^2 es

$$\frac{(n-1)p}{n-p} F_{p, n-p}(\alpha) = \frac{(15-1)2}{15-2} F_{2, 15-2}(0.10) = 5.95$$

- Lo cual nos conduce a un valor crítico de la razón de verosimilitud de

$$\Lambda_n^{\frac{2}{n}} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} = \left(1 + \frac{T^2}{n-1}\right)^{-1} = \left(1 + \frac{5.95}{15-1}\right)^{-1} = (1.412)^{-1} = 0.7018$$

- Entonces la regla de decisión: no rechazar H_0 si $\Lambda_n^{\frac{2}{n}} \geq 0.7018$
de otra manera rechazar H_0

Ejemplo: Inferencia sobre la media mediante la prueba de razón de verosimilitud

- Calculamos el estadístico de prueba. De nuestros resultados anteriores tenemos $T^2 = 5.970$, por lo que el valor calculado del estadístico de prueba de la razón de verosimilitud para esta muestra es

$$\Lambda_{\frac{2}{n}} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} = \left(1 + \frac{T^2}{n-1}\right)^{-1} = \left(1 + \frac{5.970}{15-1}\right)^{-1} = (1.426)^{-1} = 0.7011$$

Ejemplo: Inferencia sobre la media mediante la prueba de razón de verosimilitud

- Usamos la regla de decisión para evaluar el estadístico de prueba y decidimos si se rechaza o no se rechaza la hipótesis nula

$$\Lambda^{\frac{2}{n}} = 0.7011 \not\geq 0.7018$$

- Por lo que se rechaza H_0 con $\alpha = 0.10$. La muestra apoya la afirmación de que el vector de medias difiere de

$$\mu_0 = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$