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

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

- **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

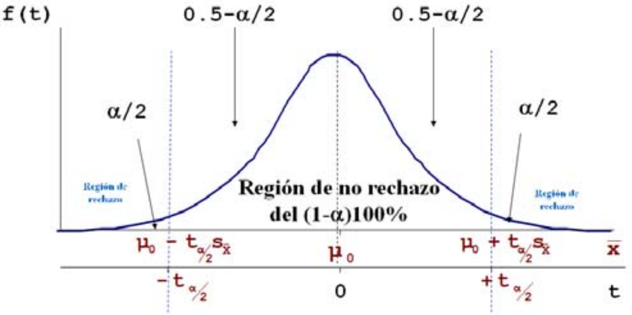
Inferencia sobre el vector de medias

- **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
Conclusion	No se Rechaza H_0	Decisión correcta	Probabilidad del error Tipo II $= \beta$
	Reject H_0	Probabilidad del error Tipo I $= \alpha$	Decisión Correcta

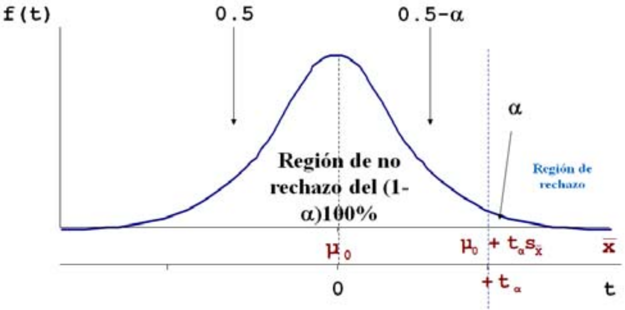
Inferencia sobre el vector de medias

- **Nivel de significancia**: La probabilidad de rechazar la Hipotesis nula cuando es verdadera. Para una prueba de dos colas :



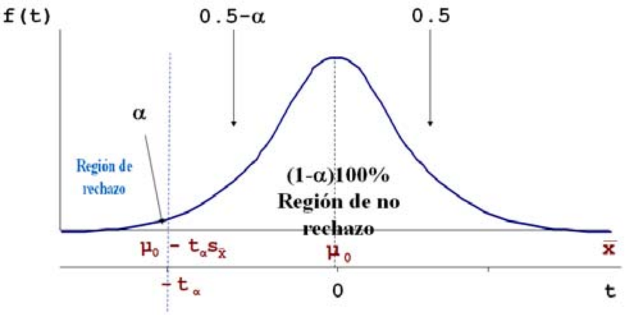
- **Regla de decisión**: no rechazar H_0 si $-t_{\alpha/2} \leq t \leq t_{\alpha/2}$, de otra forma se rechaza H_0

- Para una prueba de cola superior:



- **Regla de decisión**: no rechazar H_0 si $t \leq t_\alpha$, de otra forma se rechaza H_0

- Para una prueba de cola inferior:



- **Regla de decisión**: no rechazar H_0 si $-t_\alpha \leq t$, de otra forma se rechaza H_0

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 hipotesis 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

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

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

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

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:

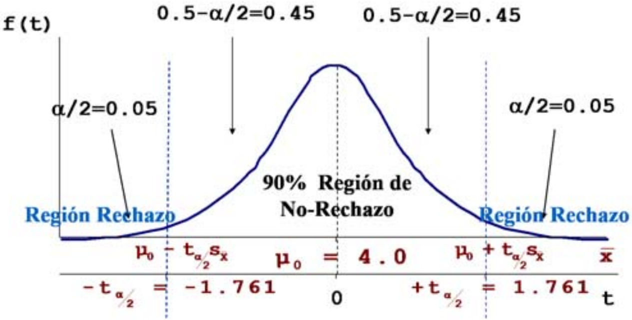
$$H_0 : \mu_1 = 4.0$$
$$H_1 : \mu_1 \neq 4.0$$

- 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 distribucion t *student*

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

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

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

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

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

$$-1.761 \leq t \nleq 1.761,\ i.e., -1.761 \leq 1.84 \nleq 1.761$$

- 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

$$H_0 : \mu_1 = \mu_0 \; vs \; H_1 : \mu_1 \neq \mu_0$$

$$H_0 : \mu_1 = -1.5 \; vs \; H_1 : \mu_1 \neq -1.5$$

Ejemplo 2: Prueba de hipótesis para una media

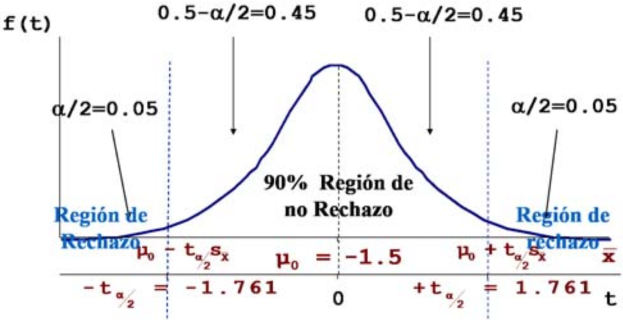
- Las Hipótesis Nula y Alternativa son:

$$H_0 : \mu_1 = -1.5$$
$$H_1 : \mu_1 \neq -1.5$$

- 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*

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

- 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

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

  Así

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

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

$$-1.761 \leq t \leq 1.761,\ i.e. -1.761 \leq -1.748 \leq 1.761$$

- 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.

Visualización del estadístico t

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

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

  ¡es grande!
- Obsérvese también que t^2 representa la distancia desde la media de la muestra $\overline{x}$ a la media hipotética de la población μ_0, expresada en términos de $s_{\overline{x}}$ (lo cual es equivalente a la distancia cuadrada generalizada de $\overline{x}$ a μ en el caso univariado)
- Si la hipótesis nula es verdadera, $t^2 \sim F_{1,n-1}$.

Una generalización natural de la distancia cuadrada univarida t es su análogo multivariado T^2 de Hotelling

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

donde

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{j=1}^{n} \mathbf{x}_j = \frac{1}{n} \mathbf{X}' \mathbf{1} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix}, \ \boldsymbol{\mu}_\mathbf{0} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$$

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

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

- 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:

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

Inferencia sobre un vector de medias μ

- 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 $W_{p,n-1}(\Sigma))$ - esto es una generalización completa del caso univariante, donde podríamos escribir el estadístico de prueba cuadrada t^2 como

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

- Podriamos 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

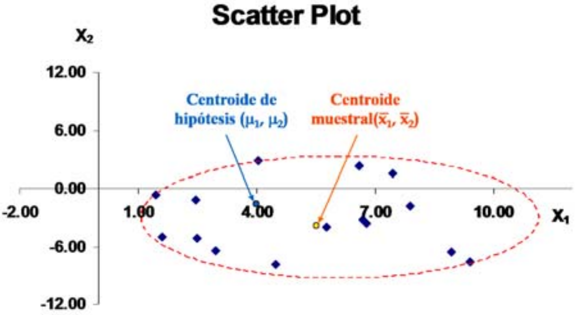
En un nivel de significancia de $\alpha = 0.10$, ¿estos datos apoyan la afirmación de que fueron extraídos de una población con un centroide (4.0,-1.5)?
Entonces, queremos probar la hipótesis nula

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

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

Ejemplo: Inferencia sobre un vector de medias μ

- El diagrama de dispersión de los pares (x_1, x_2), el centroide muestral $(\overline{x}_1, \overline{x}_2)$ 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

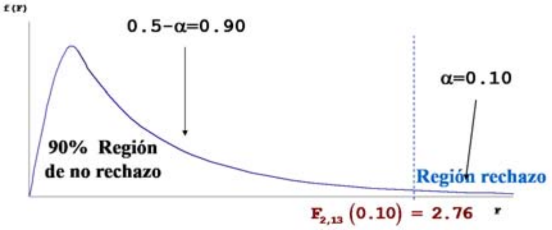
$$H_0 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix} \; vs \; H_1 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

- 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

$$T^2 = n(\overline{\mathbf{x}} - \mu_{\mathbf{0}})' \mathbf{S}^{-\mathbf{1}}(\overline{\mathbf{x}} - \mu_{\mathbf{0}})$$

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

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

- 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

- Calculamos el estadístico de prueba T^2

  Tenemos que

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

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

Entonces

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

- 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

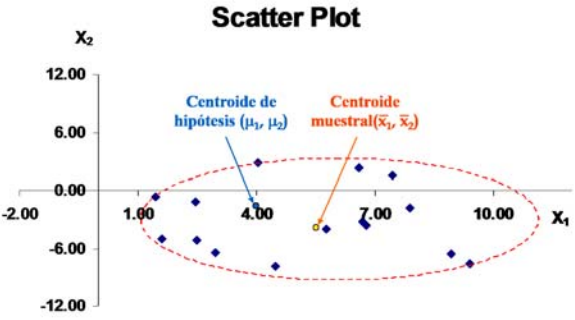
$$T^2 = 5.970 \nleq 5.95$$

- 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 los test univariados originales y el test bivariado son tan consistentes? ¡Revisa los datos !



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

- ¿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

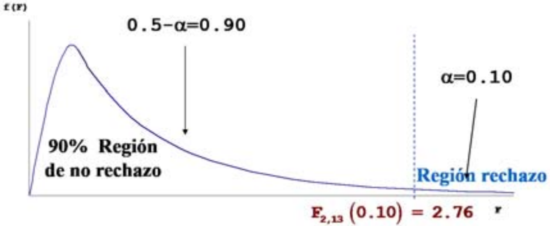
Por supuesto, los resultados de las pruebas univariadas no cambiarían, pero a un nivel de significancia de $\alpha = 0.10$, ¿estos datos apoyan ahora la afirmación de que fueron extraídos de una población con un centroide (4,0, -1,5)?
Es decir, probamos de nuevo la hipótesis nula

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

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

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 critico 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

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

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

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

Entonces

$$\begin{aligned} T^2 &= n(\bar{\boldsymbol{x}} - \mu_{\mathbf{0}})' \mathbf{S}^{-\mathbf{1}} (\bar{\boldsymbol{x}} - \mu_{\mathbf{0}}) \\ &= 15 \begin{bmatrix} 5.26 - 4.0 \\ -3.09 + 1.5 \end{bmatrix}' \begin{bmatrix} 7.12 & -9.19 \\ -9.19 & 12.43 \end{bmatrix}^{-1} \begin{bmatrix} 5.26 - 4.0 \\ -3.09 + 1.5 \end{bmatrix} \\ &= 15 \begin{bmatrix} 1.26 & -1.59 \end{bmatrix} \begin{bmatrix} 3.0374 & 2.2452 \\ 2.2452 & 1.74 \end{bmatrix} \begin{bmatrix} 1.26 \\ -1.59 \end{bmatrix} = 3.411 \end{aligned}$$

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

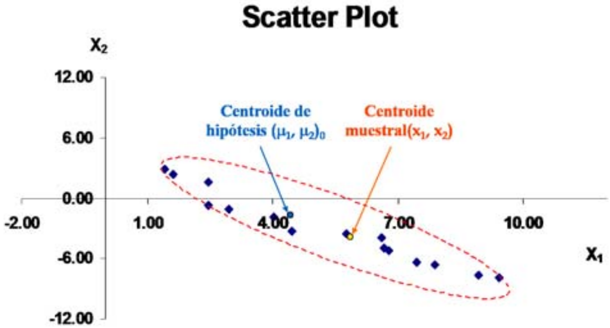
$$T^2 = 3.411 \leq 5.95$$

- 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?

- ¿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	-3.56
1.62	-7.64
4.03	-5.00
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	-5.20
2.48	-6.39

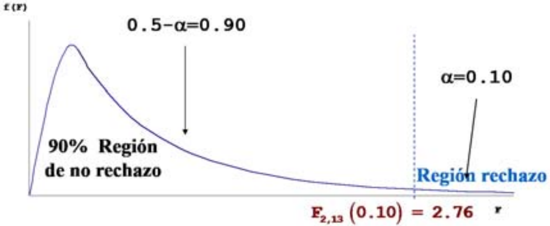
De nuevo, los resultados de las pruebas univariadas no cambiarían, pero a un nivel de significancia de $\alpha = 0.10$, ¿estos datos apoyan ahora la afirmación de que fueron extraídos de una población con un centroide (4,0, -1,5)?
En otras palabras, probamos de nuevo la hipótesis nula

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

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

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 critico 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

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

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

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

Entonces

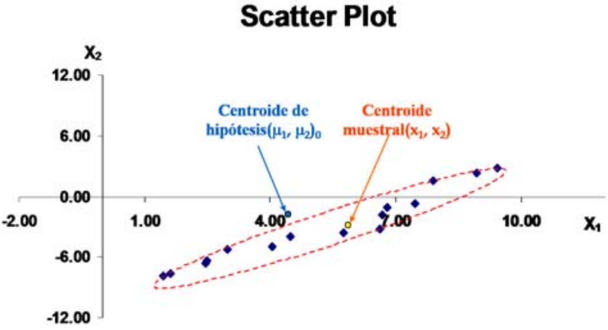
$$T^2 = n(\bar{\boldsymbol{x}} - \mu_{\mathbf{0}})'\mathbf{S}^{-\mathbf{1}}(\bar{\boldsymbol{x}} - \mu_{\mathbf{0}})$$
$$= 15 \begin{bmatrix} 5.26 - 4.0 \\ -3.09 + 1.5 \end{bmatrix}' \begin{bmatrix} 7.12 & 9.16 \\ 9.16 & 12.43 \end{bmatrix}^{-1} \begin{bmatrix} 5.26 - 4.0 \\ -3.09 + 1.5 \end{bmatrix}$$
$$= 15 \begin{bmatrix} 1.26 & -1.59 \end{bmatrix} \begin{bmatrix} 2.7407 & -2.0206 \\ -2.0206 & 1.5701 \end{bmatrix} \begin{bmatrix} 1.26 \\ -1.59 \end{bmatrix} = 247.4$$

- 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 $\mathbf{\Sigma}$) está dado por

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

Donde

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

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

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

$$L(\mu_{\mathbf{0}}, \mathbf{\Sigma}) = \frac{1}{(2\pi)^{np/2} \left|\hat{\mathbf{\Sigma}}\right|^{n/2}} e^{-\frac{1}{2}\sum_{j=1}^{n}(\mathbf{x}_j - \mu_{\mathbf{0}})\mathbf{\Sigma}^{-1}(\mathbf{x}_j - \mu_{\mathbf{0}})'}$$

- Es decir, la media μ se fija a $\mu_{\mathbf{0}}$, pero $\mathbf{\Sigma}$ puede variar hasta encontrar el valor que nos de la probabilidad mas alta de obtener la muestra de datos

- Este valor se obtiene maximizando $L(\mu_{\mathbf{0}}, \boldsymbol{\Sigma})$ con respecto a $\boldsymbol{\Sigma}$
- Por un resultado anterior podemos reescribir el exponente en $L(\mu_{\mathbf{0}}, \boldsymbol{\Sigma})$ de tal forma que

$$
\begin{aligned}
L(\mu_{\mathbf{0}}, \boldsymbol{\Sigma}) &= \frac{1}{(2\pi)^{np/2}\,|\boldsymbol{\Sigma}|^{n/2}} e^{-\frac{1}{2}\sum_{j=1}^{n} tr\left[\boldsymbol{\Sigma}^{-1}(\mathbf{x}_j-\mu_{\mathbf{0}})(\mathbf{x}_j-\mu_{\mathbf{0}})'\right]} \\
&= \frac{1}{(2\pi)^{np/2}\,|\boldsymbol{\Sigma}|^{n/2}} e^{-\frac{1}{2} tr\left[\boldsymbol{\Sigma}^{-1}\sum_{j=1}^{n}(\mathbf{x}_j-\mu_{\mathbf{0}})(\mathbf{x}_j-\mu_{\mathbf{0}})'\right]}
\end{aligned}
$$

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

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

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

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

- Aplicando este resultado con

$$\boldsymbol{B} = \sum_{j=1}^{n} (\mathbf{x}_j - \boldsymbol{\mu}_0)(\mathbf{x}_j - \boldsymbol{\mu}_0)' \quad \text{y} \quad b = \frac{n}{2}$$

tenemos

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

donde

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

- Para determinar si $\boldsymbol{\mu_0}$ es un plausible valor de $\boldsymbol{\mu}$, compararamos el máximo de $L(\boldsymbol{\mu_0}, \boldsymbol{\Sigma})$ con el máximo de $L(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ cuando no se imponen restricciones sobre $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$.
- La relación resultante se denomina Estadístico de la Razón de Verosimilitud:

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

O de manera equivalente

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

$\Lambda^{\frac{2}{n}}$ que se le denomina Lambda de Wilk

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

- Obsérvese que un valor pequeño de la lambda de Wilk 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

$$\Lambda^{\frac{2}{n}} = \frac{\left|\hat{\boldsymbol{\Sigma}}\right|}{\left|\hat{\boldsymbol{\Sigma}}_{\mathbf{0}}\right|} = \frac{\left|\sum_{j=1}^{n}(\mathbf{x}_j - \overline{\boldsymbol{x}})(\mathbf{x}_j - \overline{\boldsymbol{x}})'\right|}{\left|\sum_{j=1}^{n}(\mathbf{x}_j - \boldsymbol{\mu}_{\mathbf{0}})(\mathbf{x}_j - \boldsymbol{\mu}_{\mathbf{0}})'\right|} < c_\alpha$$

  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 prueba de razón de verosimilitud

Resultado: Suponga $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ es una muestra aleatoria de una población $N_p(\mu, \boldsymbol{\Sigma})$. 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

$$\Lambda^{\frac{2}{n}} = \frac{\left|\hat{\boldsymbol{\Sigma}}\right|}{\left|\hat{\boldsymbol{\Sigma}}_{\mathbf{0}}\right|} = \left(1 + \frac{T^2}{n-1}\right)^{-1}$$

- 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 demostrar que T^2 se puede calcular como el cociente de dos determinantes (evitando así la necesidad de calcular S^{-1}). Resolviendo para T^2 se obtiene

$$T^2 = \frac{(n-1)\left|\hat{\boldsymbol{\Sigma}}_{\mathbf{0}}\right|}{\left|\hat{\boldsymbol{\Sigma}}\right|} - (n-1)$$

$$= \frac{(n-1)\left|\sum_{j=1}^{n}(\mathbf{x}_j - \boldsymbol{\mu}_{\mathbf{0}})(\mathbf{x}_j - \boldsymbol{\mu}_{\mathbf{0}})'\right|}{\left|\sum_{j=1}^{n}(\mathbf{x}_j - \overline{\boldsymbol{x}})(\mathbf{x}_j - \overline{\boldsymbol{x}})'\right|} - (n-1)$$

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 α = 0.10) de 5.95.

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

- Definimos la hipótesis nula y alternativa

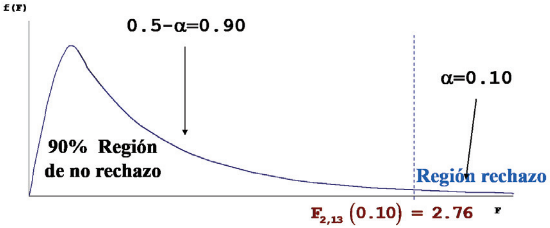
$$H_0 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix} \; vs \; H_1 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} 4.0 \\ -1.5 \end{bmatrix}$$

- 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

$$\Lambda^{\frac{2}{n}} = \frac{\left|\hat{\boldsymbol{\Sigma}}\right|}{\left|\hat{\boldsymbol{\Sigma}}_{\mathbf{0}}\right|} = \left(1 + \frac{T^2}{n-1}\right)^{-1}$$

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 a un valor crítico de la razón de verosimilitud de

$$\Lambda^{\frac{2}{n}} = \frac{\left|\hat{\boldsymbol{\Sigma}}\right|}{\left|\hat{\boldsymbol{\Sigma}}_{\mathbf{0}}\right|} = \left(1 + \frac{T^2}{n-1}\right)^{-1} = \left(1 + \frac{5.95}{15-1}\right)^{-1} = (1.412)^{-1} = 0.7018$$

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

$$\Lambda^{\frac{2}{n}} = \frac{\left|\hat{\boldsymbol{\Sigma}}\right|}{\left|\hat{\boldsymbol{\Sigma}}_{\mathbf{0}}\right|} = \left(1 + \frac{T^2}{n-1}\right)^{-1} = \left(1 + \frac{5.970}{15-1}\right)^{-1} = (1.426)^{-1} = 0.7011$$

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

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

$$\Lambda^{\frac{2}{n}} = 0.7011 \ngtr 0.7018$$

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

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

Método general de la razón de verosimilitud

- La prueba anterior fue un ejemplo específico de aplicación del Método de Razón de Verosimilitud:
- En general sea θ un vector que consiste en todos los parámetros desconocidos de la población, y $L(\theta)$ la función de verosimilitud obtenida evaluando la densidad conjunta de $\boldsymbol{X_1}, \ldots, \boldsymbol{X_n}$ en sus valores observados $\boldsymbol{x_1}, \ldots \boldsymbol{x_n}$.
- θ toma sus valores en el conjunto de parámetros $\boldsymbol{\Theta}$. Por ejemplo, en el caso normal p-dimensional $\theta = [\mu_1, \ldots, \mu_p; s_{11}, \ldots, s_{1p}; s_{21}, \ldots, s_{2p}; \ldots; s_{p1}, \ldots, s_{pp}]$.
- Entonces $\boldsymbol{\Theta}$ consiste del espacio p-dimensional de las μ_i, $-\infty \leq \mu_i \leq \infty$, $i = 1, \ldots, p$, combinado con el espacio $p(p+1)/2$-dimensional de las varianzas y covarianzas tal que $\boldsymbol{\Sigma}$ es positivo definido. Po lo tanto, $\boldsymbol{\Theta}$ tiene una dimensión $v = p + p(p+1)/2$.

Método general de la razón de verosimilitud

- Bajo la hipótesis nula $H_0 : \theta = \boldsymbol{\theta_0}$, θ está restringido a estar en algun subconjunto $\boldsymbol{\Theta_0} \subseteq \boldsymbol{\Theta}$.
- Por ejemplo, en el caso normal p-dimensional con $\mu = \boldsymbol{\mu_0}$ y $\boldsymbol{\Sigma}$ no especificada, tenemos que

$$\boldsymbol{\Theta_0} = [\mu_1 = \mu_{1_0}, .., \mu_p = \mu_{p_0}; s_{11}, .., s_{1p}; s_{21}, .., s_{2p}; ...; s_{p1}, .., s_{pp}, \text{ con } \boldsymbol{\Sigma} \text{ p}$$

- Así que $\boldsymbol{\Theta_0}$ tiene dimensión $v_0 = 0 + p(p+1)/2 = p(p+1)/2$
- Entonces una prueba de razón de verosimilitud de $H_0 : \theta \in \boldsymbol{\Theta_0}$ rechaza H_0 en favor de $H_1 : \theta \ni \boldsymbol{\theta_0}$ si

$$\Lambda = \frac{\max_{\theta \in \boldsymbol{\Theta_0}} L(\boldsymbol{\Theta})}{\max_{\theta \in \boldsymbol{\Theta}} L(\Theta)} < c$$

- Para una constante c escogida adecuadamente (algún percentil de la distribución de Λ).

- En cada aplicación del método de la razón de verosimilitud, se debe obtener la distribucion muestral del estadístico de la razon de verosimilitud.
- Entonce c puede ser seleccionado para producir una prueba con un específico α
- Sin embargo, cuando n es suficientemente grande, la distribución

$$-2ln(\Lambda) = -2ln\left(\frac{\max_{\theta\in\boldsymbol{\Theta_0}} L(\boldsymbol{\Theta})}{\max_{\theta\in\boldsymbol{\Theta}} L(\Theta)}\right)$$

bajo la hipótesis nula $H_0 : \theta = \boldsymbol{\Theta_0}$ sigue aproximadamente una $\chi^2_{v-v_0}$.
- donde v es el número de parámetros que se estiman sin considerar restricciones y v_0 el número de parámetros que se estiman cuando algunos de los parámetros son fijados

Comentarios sobre método de la razón de verosimilitud

- La aproximación de la razón de verosimilitud a una χ^2 cuando n es grande nos permite obtener de manera directa las regiones de rechazo y aceptación en prueba de hipótesis, de acuerdo a los percentiles de la χ^2
- Evitando así el cálculo de sus valores críticos partir de los valores criticos de una T^2 que a su vez obtiene los valores criticos a partir de una F
- Entre todas las pruebas estadísticas para hacer inferencia sobre los parámetros de la población, las pruebas de la razón de verosimilitud son la que tiene mayor potencia, cuando la muestra es grande.

Regiones de confianza para el vector de medias μ

- Cuando H_0 no se rechaza, concluimos que μ_0 es un posible valor para la media de la poblacion normal. Sin embargo, existen otros conjuntos de valores de μ que también pueden ser consistentes con los datos.
- Un intervalo de confianza del $100(1-\alpha)\%$ para la media μ tiene la forma

$$P(\overline{x} - t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}} \leq \mu \leq \overline{x} + t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}}) = 1-\alpha$$

- Un intervalo de confianza univariado del $100(1-\alpha)\%$ para un parámetro θ de la población se define de manera general como

$$P(\widehat{\theta} - \varepsilon \leq \theta \leq \widehat{\theta} + \varepsilon) = P(|\widehat{\theta} - \theta| \leq \varepsilon) = 1-\alpha$$

donde $\widehat{\theta}$ es el estimador del parámetro y ε mide la desviación del parámetro respecto a su estimador

Regiones de confianza para el vector de medias μ

- El concepto de intervalo de confianza univariado se puede extender a un espacio multivariado p-dimensional, que son las **regiones de confianza multivariadas**.
- Sea θ un vector de parámetros desconocidos de una poblacion y sea $\mathbf{\Theta}$ el conjunto de todos los posibles valores de θ.
- Una **región de confianza** que llamaremos $R(\boldsymbol{X})$ es una región de probables valores de θ, y se determina a partir de la muestra de datos $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_1]$.
- La región $R(\boldsymbol{X})$ es una región de confianza del $100(1-\alpha)\%$, si antes de que la muestra sea seleccionada

$$P(R(\boldsymbol{X}) \text{ cubra al verdadero } \theta) = 1 - \alpha$$

esta probabilidad es calculada bajo el verdadero pero desconocido valor de θ

Regiones de confianza para el vector de medias μ

- Recordando que el estadístico para probar $H_0 : \mu = \mu_\mathbf{0}$ está dado por

$$T^2 = n(\overline{\mathbf{x}} - \mu_\mathbf{0})' \mathbf{S}^{-\mathbf{1}} (\overline{\mathbf{x}} - \mu_\mathbf{0})$$

- No se rechazaba H_0 si $T^2 \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$, es decir si

$$n(\overline{\mathbf{x}} - \mu_\mathbf{0})' \mathbf{S}^{-\mathbf{1}} (\overline{\mathbf{x}} - \mu_\mathbf{0}) \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$$

- Por tanto la región de confianza para μ de una población normal p-variada está dado por

$$P\left(n(\overline{\mathbf{x}} - \mu)' \mathbf{S}^{-\mathbf{1}} (\overline{\mathbf{x}} - \mu) \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha) \right) = 1 - \alpha$$

- Más formalmente, una región de confianza $R(\boldsymbol{X})$ del $100(1-\alpha)\%$ para el vector de medias μ de una distribución normal p-dimensional es el elipsoide determinado por todos los puntos posibles μ que satisfacen

$$n(\overline{\boldsymbol{x}}-\mu)^{'}\mathbf{S}^{-\mathbf{1}}(\overline{\boldsymbol{x}}-\mu) \leq \frac{(n-1)p}{n-p}F_{p,n-p}(\alpha)$$

donde

$$\overline{\mathrm{x}} = \frac{1}{n}\sum_{j=1}^{n}\mathrm{x}_j \text{ y } \boldsymbol{S} = \frac{1}{n-1}\sum_{j=1}^{n}(\mathrm{x}_j-\overline{\mathrm{x}})(\mathrm{x}_j-\overline{\mathrm{x}})^{'}$$

y $\boldsymbol{x_1},\ldots\boldsymbol{x_n}$ son las muestras de observaciones

Ejemplo: Construcción de una region de confianza del $100(1-\alpha)\,\%$ para el vector de medias

Utilizar la muestra original de datos bivariados para construir una región bidimensional de confianza del 90 % y determine si el punto $(4.0, -1.5)$ se encuentra en esta región

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

Resumen de los estadísticos

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

$$S = \begin{bmatrix} 7.12 & -.72 \\ -.72 & 12.43 \end{bmatrix}, S^{-1} = \begin{bmatrix} .1412 & .0082 \\ .0082 & .0809 \end{bmatrix}$$

y

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