Pruebas de Hipótesis

Juan Sosa, PhD

Terminología

Hipótesis estadística

Aseveración (afirmación) acerca de uno o más **parámetros** poblacionales.

- Hipótesis **simple**: $H : \mu = \mu_0$.
- Hipótesis **compuesta**: $H : \mu \geq \mu_0$, $H : \mu \leq \mu_0$, $H : \mu > \mu_0$, $H : \mu < \mu_0$, $H : \mu \neq \mu_0$.
- μ_0 se denomina **valor hipotético**.

Sistema de hipótesis

Arreglo de dos hipótesis denominas **hipótesis nula** e **hipótesis alternativa**.

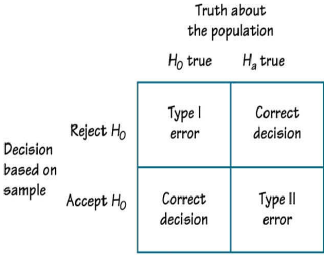
- $H_0 : \mu \leq \mu_0$ frente a $H_1 : \mu > \mu_0$.
- $H_0 : \mu \geq \mu_0$ frente a $H_1 : \mu < \mu_0$.
- $H_0 : \mu = \mu_0$ frente a $H_1 : \mu \neq \mu_0$.

Objetivo

Establecer si hay **suficiente evidencia** en la realización de una muestra aleatoria para **rechazar** o **no rechazar** la **hipótesis nula**. ¿Cuál hipótesis explica mejor los datos?

Tipos de error

- Error tipo 1: **Rechazar** la hipótesis nula cuando ésta es cierta.
- Error tipo 2: **No rechazar** la hipótesis nula cuando ésta es falsa.



Nivel de significativa

- **Probabilidad** de cometer el **error tipo 1**:

$$\alpha = \mathbb{P}\text{r}\left[\text{Error tipo 1}\right] = \mathbb{P}\text{r}\left[\text{Rechazar } H_0 \mid H_0\right]$$

- El nivel de significativa usualmente **se fija de antemano** a 0.05 o 0.01.

Analogía

$$H_0 : \text{El acusado es } \textbf{inocente} \quad \text{frente a} \quad H_1 : \text{El acusado es } \textbf{culpable}$$

El acusado se considera inocente hasta que se demuestre lo contrario más allá de cualquier duda razonable.

Rechazar H_0

Se ha encontrado algo en la muestra tan improbable que ocurra si H_0 es cierta, que obliga al investigador a favorecer H_1 (rechazar H_0).

La ciencia es conservadora!

- Se considera mejor, equivocadamente no rechazar, que equivocadamente rechazar el estado actual del conocimiento (H_0).
- H_0 se mantiene a menos de que haya suficiente evidencia para revocarla.

Procedimiento de prueba (test)

¿Cómo decidir si se debe rechazar o no la hipótesis nula?

1. Establecer el **sistema de hipótesis** de interés y el **nivel de significancia**.
2. Calcular el **estadístico de prueba**.
3. Establecer la **región crítica** o el **valor p**.
4. Tomar la **decisión**.
5. **Interpretar los resultados!**

¿Cómo tomar la decisión?

Si el estadístico de prueba pertenece a la región crítica, entonces se rechaza H_0.
Si el valor p es menor que α, entonces se rechaza H_0.

Modelo

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \mathsf{N}(\mu, \sigma^2)$$

Sistema de hipótesis

$$\begin{array}{lcl} H_0 : \mu \leq \mu_0 & \text{frente a} & H_1 : \mu > \mu_0 \\ H_0 : \mu \geq \mu_0 & \text{frente a} & H_1 : \mu < \mu_0 \\ H_0 : \mu = \mu_0 & \text{frente a} & H_1 : \mu \neq \mu_0 \end{array}$$

Estadístico de prueba y distribución probabilística de referencia

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim \mathsf{t}_{n-1}$$

Nota: Para $n \geq 30$ se tiene que $\mathsf{t}_{n-1} \approx \mathsf{N}(0,1)$ y la población no tiene que ser Normal.

Región de rechazo

$$\text{Rechazar } H_0 \text{ si } \begin{cases} t > \mathsf{t}_{1-\alpha,n-1}, & \text{si } H_1 : \mu > \mu_0; \\ t < \mathsf{t}_{\alpha,n-1}, & \text{si } H_1 : \mu < \mu_0; \\ |t| > \mathsf{t}_{1-\alpha/2,n-1}, & \text{si } H_1 : \mu \neq \mu_0. \end{cases}$$

Definición

$$p = \mathbb{P}\text{r}\left[\text{Observar datos } \textbf{tan o más extremos} \text{ en dirección de } H_1 \mid H_0\right]$$

Probabilidad de que el azar produzca una **coincidencia al menos tan extraordinaria** como el fenómeno observado, bajo la hipótesis nula.

Observación

Rechazar H_0 siempre que $p < \alpha$.

Se ha observado algo que es **tan poco probable que ocurra** si la hipótesis nula es cierta que hace dudar seriamente que realmente sea cierta.

Calculo del valor p para la media poblacional

$$p = \begin{cases} \mathbb{P}\text{r}\left[\text{t}_{n-1} > t\right], & \text{si } H_1 : \mu > \mu_0; \\ \mathbb{P}\text{r}\left[\text{t}_{n-1} < t\right], & \text{si } H_1 : \mu < \mu_0; \\ 2\,\mathbb{P}\text{r}\left[\text{t}_{n-1} > |t|\right], & \text{si } H_1 : \mu \neq \mu_0. \end{cases}$$

Nota: Para $n \geq 30$ se tiene que $\text{t}_{n-1} \approx \text{N}(0,1)$.

Ejemplo

Pruebas de hipótesis para la media poblacional

Con propósitos de acreditación institucional, se quiere comprobar si la media de las calificaciones promedio de todos los estudiantes de una Facultad difiere de 4.0.

Para tal fin, se toma una muestra aleatoria de estudiantes de la Facultad y se registra la calificación promedio del semestre inmediatamente anterior de cada estudiante (ver los datos correspondientes en `PROM.txt`).

Usando un nivel de significancia de 5%, probar el sistema de hipótesis correspondiente.

Modelo

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \text{Ber}(\pi)$$

Sistema de hipótesis

$$\begin{array}{lcl} H_0 : \pi \leq \pi_0 & \text{frente a} & H_1 : \pi > \pi_0 \\ H_0 : \pi \geq \pi_0 & \text{frente a} & H_1 : \pi < \pi_0 \\ H_0 : \pi = \pi_0 & \text{frente a} & H_1 : \pi \neq \pi_0 \end{array}$$

Estadístico de prueba y distribución probabilística de referencia

$$z = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} \approx \text{N}(0, 1)$$

Nota: Esta aproximación es apropiada para $n \geq 30$, $n\hat{\pi} \geq 5$ y $n(1 - \hat{\pi}) \geq 5$.

Región de rechazo

$$\text{Rechazar } H_0 \text{ si } \begin{cases} z > z_{1-\alpha}, & \text{si } H_1 : \pi > \pi_0; \\ z < z_{\alpha}, & \text{si } H_1 : \pi < \pi_0; \\ |z| > z_{1-\alpha/2}, & \text{si } H_1 : \pi \neq \pi_0. \end{cases}$$

Ejemplo

Pruebas de hipótesis para la proporción poblacional

Hacer trampa ha sido una preocupación del decano de la Facultad de Negocios de Bayview University por muchos años. Algunos miembros de la Facultad de negocios en la Universidad creen que hacer trampa es más extendido en Bayview que en otras universidades, mientras que otros miembros de la Facultad piensan que hacer trampa no es un problema importante en la universidad. Para resolver algunos de estos problemas, el decano encargó un estudio para evaluar el comportamiento ético actual de los estudiantes de negocios en Bayview. Como parte de este estudio, se administró una encuesta de salida anónima a una muestra de 90 estudiantes de la clase de graduación de 2009. Las respuestas a las preguntas están dadas en `BAYVIEW.txt` en relación a tres tipos de trampas.

Realizar una prueba de hipótesis para determinar si la proporción de estudiantes de negocios en La Universidad Bayview que estuvo involucrada en algún tipo de trampa es menor que la de estudiantes de negocios en otras instituciones según lo informado por *Chronicle of Higher Eudcation* (el artículo asegura que 56% de los estudiantes de negocios hacen trampa en algún momento de su carrera académica).

Pruebas de hipótesis para la varianza poblacional

Modelo

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \mathsf{N}(\mu, \sigma^2)$$

Sistema de hipótesis

$$\begin{array}{lcl} H_0 : \sigma^2 \leq \sigma_0^2 & \text{frente a} & H_1 : \sigma^2 > \sigma_0^2 \\ H_0 : \sigma^2 \geq \sigma_0^2 & \text{frente a} & H_1 : \sigma^2 < \sigma_0^2 \\ H_0 : \sigma^2 = \sigma_0^2 & \text{frente a} & H_1 : \sigma^2 \neq \sigma_0^2 \end{array}$$

Estadístico de prueba y distribución probabilística de referencia

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$$

Región de rechazo

$$\text{Rechazar } H_0 \text{ si } \begin{cases} \chi^2 > \chi^2_{1-\alpha,n-1}, & \text{si } H_1 : \sigma^2 > \sigma_0^2; \\ \chi^2 < \chi^2_{\alpha,n-1}, & \text{si } H_1 : \sigma^2 < \sigma_0^2; \\ \chi^2 < \chi^2_{\alpha/2,n-1} \text{ o } \chi^2 > \chi^2_{1-\alpha/2,n-1}, & \text{si } H_1 : \sigma^2 \neq \sigma_0^2. \end{cases}$$

Ejemplo

Pruebas de hipótesis para la varianza poblacional

A fines de 2008, la variación en los rendimientos semestrales de bonos gubernamentales en el extranjero era $\sigma^2 = 0.70$. Un grupo de inversores se reunió en ese momento para analizar las tendencias futuras de rendimientos en el exterior. Algunos esperaban que la variabilidad en los rendimientos de los bonos en el extranjero aumentara y otros tomaron la posición opuesta. Los datos dados en `YIELDS.txt` corresponden a los rendimientos semestrales de 12 países desde el 6 de marzo de 2009 (*Barron's*, 9 de marzo de 2009).

Desarrollar y probar la hipótesis para comprobar si los datos indican que la variabilidad en los los rendimientos han cambiado desde finales de 2008. Usar un nivel de significancia del 5%.

Modelo

Poblaciones independientes:

$$X_1, X_2, \ldots, X_{n_1} \overset{\text{iid}}{\sim} \text{Ber}(\pi_1) \qquad Y_1, Y_2, \ldots, Y_{n_2} \overset{\text{iid}}{\sim} \text{Ber}(\pi_2)$$

Sistema de hipótesis

$$\begin{array}{lcl} H_0 : \pi_1 - \pi_2 \leq \delta_0 & \text{frente a} & H_1 : \pi_1 - \pi_2 > \delta_0 \\ H_0 : \pi_1 - \pi_2 \geq \delta_0 & \text{frente a} & H_1 : \pi_1 - \pi_2 < \delta_0 \\ H_0 : \pi_1 - \pi_2 = \delta_0 & \text{frente a} & H_1 : \pi_1 - \pi_2 \neq \delta_0 \end{array}$$

Estadístico de prueba y distribución probabilística de referencia

$$z = \frac{\hat{\pi}_1 - \hat{\pi}_2 - \delta_0}{\sqrt{\hat{\pi}(1-\hat{\pi})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \approx \text{N}(0,1) \qquad \hat{\pi} = \frac{x_1 + x_2}{n_1 + n_2}$$

Nota: Esta aproximación es apropiada para

$$n_1, n_2 \geq 30 \quad n_1\hat{\pi}_1, n_1\hat{\pi}_2 \geq 5 \quad n_1(1-\hat{\pi}_1), n_2(1-\hat{\pi}_2) \geq 5$$

Pruebas de hipótesis para la diferencia de proporciones poblacional

Para comparar la aceptación de una medida económica en dos universidades se toma una muestra aleatoria en cada una obteniéndose los dados en `ESTEC.txt`.

Probar si hay diferencias significativas entre la proporción de estudiantes que aprueban la medida. Usar un nivel de significancia de 5%.

Pruebas de hipótesis para el cociente de varianzas poblacional

Modelo

Poblaciones independientes:

$$X_1, X_2, \ldots, X_{n_1} \overset{\text{iid}}{\sim} \mathsf{N}(\mu_1, \sigma_1^2) \qquad Y_1, Y_2, \ldots, Y_{n_2} \overset{\text{iid}}{\sim} \mathsf{N}(\mu_2, \sigma_2^2)$$

Sistema de hipótesis

$$\begin{array}{lcl} H_0 : \sigma_1^2/\sigma_2^2 \leq 1 & \text{frente a} & H_1 : \sigma_1^2/\sigma_2^2 > 1 \\ H_0 : \sigma_1^2/\sigma_2^2 \geq 1 & \text{frente a} & H_1 : \sigma_1^2/\sigma_2^2 < 1 \\ H_0 : \sigma_1^2/\sigma_2^2 = 1 & \text{frente a} & H_1 : \sigma_1^2/\sigma_2^2 \neq 1 \end{array}$$

Estadístico de prueba y distribución probabilística de referencia

$$f = \frac{s_1^2}{s_2^2} \sim F_{n_1-1, n_2-1}$$

Nota: Esta prueba es indispensable antes de realizar una prueba de hipótesis sobre la diferencia de medias.

Modelo

Poblaciones independientes:

$$X_1, X_2, \ldots, X_{n_1} \overset{\text{iid}}{\sim} \mathsf{N}(\mu_1, \sigma_1^2) \qquad Y_1, Y_2, \ldots, Y_{n_2} \overset{\text{iid}}{\sim} \mathsf{N}(\mu_2, \sigma_2^2)$$

Sistema de hipótesis

$$\begin{array}{lcl} H_0 : \mu_1 - \mu_2 \leq \delta_0 & \text{frente a} & H_1 : \mu_1 - \mu_2 > \delta_0 \\ H_0 : \mu_1 - \mu_2 \geq \delta_0 & \text{frente a} & H_1 : \mu_1 - \mu_2 < \delta_0 \\ H_0 : \mu_1 - \mu_2 = \delta_0 & \text{frente a} & H_1 : \mu_1 - \mu_2 \neq \delta_0 \end{array}$$

Estadístico de prueba y distribución probabilística de referencia

$$t = \frac{\bar{x_1} - \bar{x}_2 - \delta_0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

Nota: Este estadístico de prueba está construido bajo el supuesto de homogeneidad $(\sigma_1^2 = \sigma_2^2)$

Modelo

Poblaciones independientes:

$$X_1, X_2, \ldots, X_{n_1} \overset{\text{iid}}{\sim} \mathsf{N}(\mu_1, \sigma_1^2) \qquad Y_1, Y_2, \ldots, Y_{n_2} \overset{\text{iid}}{\sim} \mathsf{N}(\mu_2, \sigma_2^2)$$

Sistema de hipótesis

$$\begin{array}{lcl} H_0 : \mu_1 - \mu_2 \leq \delta_0 & \text{frente a} & H_1 : \mu_1 - \mu_2 > \delta_0 \\ H_0 : \mu_1 - \mu_2 \geq \delta_0 & \text{frente a} & H_1 : \mu_1 - \mu_2 < \delta_0 \\ H_0 : \mu_1 - \mu_2 = \delta_0 & \text{frente a} & H_1 : \mu_1 - \mu_2 \neq \delta_0 \end{array}$$

Estadístico de prueba y distribución probabilística de referencia

$$t = \frac{\bar{x_1} - \bar{x}_2 - \delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_\nu \qquad \nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$

Nota: Este estadístico de prueba está construido bajo el supuesto de heterogeneidad ($\sigma_1^2 \neq \sigma_2^2$)

Pruebas de hipótesis para la diferencia de medias poblacional

Millones de inversores compran fondos de inversión, eligiendo entre miles de posibilidades. Algunos fondos se pueden comprar directamente desde bancos u otras instituciones financieras, mientras que otros deben comprarse a través de intermediarios, que cobran una tarifa por este servicio. Esto plantea la pregunta: ¿pueden los inversores mejorar comprando fondos directamente que mediante la compra de fondos a través de corredores?

Para responder esta pregunta, un grupo de investigadores tomaron muestras al azar de las declaraciones anuales de fondos de inversión que se pueden adquirir directamente y fondos de inversión que se compran a través de corredores y registró los rendimientos anuales netos, que corresponden los rendimientos de la inversión después de deducir todas las tarifas relevantes (ver los datos en `PUR.txt`).

¿Es posible concluir con un nivel de significancia del 5% que los fondos comprados directamente superan fondos mutuos comprados a través de corredores?

Prueba chi cuadrado de bondad de ajuste

Experimento multinomial

- Número fijo n de repeticiones.
- El resultado de cada ensayo se puede clasificar en una de las k categorías.
- La probabilidad π_i de que el resultado caerá en la categoría i permanece constante para cada prueba. Además, $\pi_1 + \pi_2 + \ldots + \pi_k = 1$.
- Cada ensayo del experimento es independiente de los otros ensayos.

Sistema de hipótesis

$$H_0 : \pi_1 = \pi_1^0, \ldots, \pi_k = \pi_k^0$$

frente a

$$H_1 : \text{ al menos un } \pi_i \text{ no es igual al valor especificado}$$

Estadístico de prueba y distribución probabilística de referencia

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - e_i)^2}{e_i} \sim \chi^2_{k-1}$$

donde las f_i y las $e_i = n\pi_i^0$ son las frecuencias (absolutas) observadas y esperadas, respectivamente.

Ejemplo

Prueba chi cuadrado de bondad de ajuste

Los gerentes financieros están interesados en el velocidad con la cual los clientes que hacen compras a crédito paguen sus cuentas. Además de calcular el número promedio de días que las cuentas pendientes de pago (llamadas cuentas por cobrar) permanecen pendientes, a menudo preparan un cronograma de seguimiento. Un cronograma de seguimiento clasifica las cuentas por cobrar pendientes de acuerdo al tiempo que ha transcurrido desde la facturación y registra la proporción de cuentas por cobrar perteneciente a cada clasificación.

Una gran empresa determinó su cronograma de seguimiento durante los últimos 5 años:

Número de días	Por cobrar (%)
0-14	.72
15-29	.15
30-59	.10
60 y más	.03

A la compañía le gustaría saber si en el presente año el cronograma de seguimiento a cambiado significativamente. Para tal fin, se tomó muestra aleatoria de 250 cuentas y fueron clasificadas de acuerdo con estas categorías. El conjunto de datos esta disponible en `ACOUNTS.txt`. Usar una confiabilidad del 95%.

Sistema de hipótesis

$$H_0 : \text{Las variables son independientes}$$
frente a
$$H_1 : \text{Las variables son dependientes}$$

Estadístico de prueba y distribución probabilística de referencia

$$\chi^2 = \sum_{i=1}^{f} \sum_{j=1}^{c} \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2_{(f-1)(c-1)}$$

donde las f_{ij} y las $e_{ij} = n_{i.}\, n_{.j}/n$ son las frecuencias (absolutas) observadas y esperadas, respectivamente.

Ejemplo

Prueba chi cuadrado de independencia

Después de un análisis exhaustivo del mercado, un editor ha segmentado el mercado en tres general enfoques para enseñar estadística aplicada. Estos son (1) uso de una computadora y software estadístico con sin cálculos manuales, (2) enseñanza tradicional de conceptos y solución de problemas a mano, y (3) enfoque matemático con énfasis en derivaciones y pruebas. El editor quería saber si este mercado podría ser segmentado con los antecedentes educativos del instructor.

Como resultado, el editor organizó una encuesta que preguntó a 195 profesores de negocios y estadística de economía para informar su enfoque a enseñanza y cuál de las siguientes categorías representa su más alto grado: 1. Negocios (MBA o doctorado en negocios) 2. Economía 3. Matemáticas o ingeniería 4. Otro ¿Existen diferencias significativas en el tipo de grado entre los tres enfoques enseñanzas? Si es así, ¿cómo puede el editor usar esto? El conjunto de datos esta disponible en `INDEP.txt`. Usar una confiabilidad del 95%.