

Test (o Prueba) de Hipótesis

Ejemplo: Una muestra de 36 datos tiene una media igual a 4.64 ¿ Qué puede deducirse acerca de la población de donde fue tomada ?

Se necesita contestar a dos tipos de preguntas :

1. Tomando como base la muestra, ¿ que estimación puede hacerse acerca del valor de la media poblacional ?
2. ¿ El valor de la media muestral es significativamente distinto del valor hipotético 4.5 del media poblacional ?

La 1º pregunta implica *una estimación*, mientras que la 2º corresponde a *una decisión*. En este capítulo se verá la manera como se utiliza una prueba de hipótesis para tomar una decisión estadística.

La manera formal de hacer inferencia estadística a partir de una muestra, es a través del test o prueba de hipótesis (otra forma es a través de los I.C.). Esta principalmente la podemos emplear para comprobar una afirmación (o hipótesis) acerca de un valor del parámetro, y no es nuestro interés , la ubicación ni la longitud del intervalo en sí. Por ejemplo, si se garantiza que un componente electrónico tiene una vida útil promedio de por lo menos 200 horas. entonces a un investigador tan sólo puede interesarle comprobar la hipótesis de que el promedio es en realidad de 200 horas. o más, contra ala alternativa de que el promedio sea menor que 200 horas.

La prueba de hipótesis surge como una consecuencia natural del método científico. El científico observa la naturaleza, formula una teoría y luego contrasta la teoría y la observación. En el contexto de nuestros problemas, el experimentador formula una teoría acerca de que un parámetro de la población toma determinado valor o conjunto de valores. A continuación selecciona una muestra de la población en cuestión y se compara la observación con la teoría. Si las observaciones se apartan mucho de la teoría, entonces quien experimenta puede rechazar la teoría (hipótesis). Si las observaciones son compatibles con la teoría , entonces el experimentador puede proseguir como si la hipótesis fuera cierta .

Nótese que la prueba de hipótesis necesita una decisión al comparar la muestra observada con la teoría .

Las pruebas de hipótesis estadísticas es un procedimiento de 4 pasos bien organizados

Paso 1: Formular las hipótesis nula y alternativa

Paso 2: Determinar el criterio de prueba o contraste

Paso 3: Obtener de los datos muestrales (evidencia) y calcúlese el valor de la estadística de prueba

Paso 4: Tomar una decisión e interpretarla

Paso 1: Formular las hipótesis nula y alternativa

- **Hipótesis :** Aseveración acerca de la verdad de algo. Se consideran dos hipótesis.
- **Hipótesis Nula, H_0 :** Hipótesis por contrastar. Generalmente es una aseveración en el sentido de que un parámetro poblacional tiene un valor específico. Esta hipótesis nula recibe tal nombre debido a que es el “punto de partida “ de la investigación. A menudo se utiliza en su interpretación las frases :
“no (sí) existe diferencia significativa para ... “
- **Hipótesis Alternativa, H_1 :** Esta hipótesis, sobre la cual se enfoca la atención, es una aseveración sobre el mismo parámetro poblacional que se utiliza en la hipótesis nula. Generalmente se especifica que el parámetro poblacional tiene un valor diferente, de alguna manera, al establecido en la hipótesis nula (es lo que sugieren los datos)

Nota : - El rechazo de la hipótesis nula implicará la aceptación de la hipótesis alternativa

La hipótesis nula y la alternativa se formulan examinando el problema o aseveración buscada planteándose primeramente dos afirmaciones opuestas.

Ilustración:

1. Un presidente de una empresa afirma que el nº de llamadas solicitando servicio no es más de 15 por semana, en promedio. Para comprobar su afirmación , se revisaron los registros de servicio para $n = 36$ semanas seleccionadas al azar, el resultado fue que $\bar{X} = 17$ y $S^2 = 9$ para los datos de la muestra . ¿ Contradice la evidencia de la muestra la afirmación del presidente ?

Solución

Como el presidente afirma que según sus especificaciones, el promedio debe ser 15 o menos. Entonces se está comparando :

$$H_0 : \mu \leq 15 \quad \text{v/s} \quad H_1 : \mu > 15$$

Etc...

Decisiones y errores de prueba

Al concluir el contraste de la hipótesis se tomará una de dos decisiones posibles. Se estará de acuerdo con la hipótesis nula y se dirá que “ no se puede rechazar H_0 “. O bien se decidirá en contra de la hipótesis nula y se dirá que “ se rechaza H_0 “. sin embargo esto esta relacionado con la veracidad de la hipótesis nula y la correcto de la decisión que se tome, es decir se puede llegar a que la hipótesis nula sea verdadera o falsa, y que la decisión sea “ no rechazar “ o bien “ rechazar “ por lo que se tiene cuatro posibilidades.

DECISION	HIPÓTESIS NULA	
	VERDADERA	FALSA
NO SE RECHAZA H_0	Decisión correcta Tipo A	Error Tipo II
SE RECHAZA H_0	Error Tipo I	Decisión correcta Tipo B

Ocurre:

- Una decisión correcta tipo A, cuando la hipótesis nula es verdadera y se decide a favor de ella
- Una decisión correcta tipo B, cuando la hipótesis nula es falsa y la decisión tomada es contraria a esta hipótesis
- Error Tipo I, cuando es cierta la hipótesis nula pero se decide en su contra
- Error Tipo II, cuando se decide a favor de una hipótesis nula que en realidad es falsa.

La probabilidad asignada al error tipo I se conoce como *nivel de significación* y se denota por α , y la probabilidad de cometer el error tipo II se denota por la letra β

Paso 2: Determinar el criterio de prueba o contraste

- **Criterio de contrastes** : consiste en la especificación de : (1) un nivel de significación α , (2) una estadística de prueba o contraste, (3) la(s) región(es) crítica(s), y (4) el (los) valor(es) crítico(s).
- **Nivel de significación** : es la probabilidad de cometer el error tipo I (α)

Usualmente se utiliza $\alpha = 0.05$ (5%).

Típicamente quien dirige una acción determina la cantidad (probabilidad de) riesgo que parece razonable en vista de la gravedad de cometer el error tipo I.

Nota : Existe una relación entre las probabilidades de los errores tipo I y II, y el tamaño de la muestra. En esta etapa de la discusión , en el problema se especificarán α la probabilidad de cometer el error tipo I, así como el tamaño muestral n .

- **Estadística de la Prueba o contraste** : Es una variable aleatoria para tomar la decisión “ no se rechaza H_0 ” o bien “ se rechaza H_0 “
- **Región Crítica** : Conjunto de valores de la estadística de contraste que causan el rechazo de la hipótesis nula

Nota : Recordar que se está trabajando con el supuesto de que la hipótesis nula es verdadera

- **Valor Crítico (Valor – p)** : “ Primer “ valor (o valor “ frontera “) en la región crítica, esto es el valor en el cual no se conoce la decisión a tomar (aceptar o rechazar H_0) o el punto donde se cambia de decisión.

Paso 3: Obtener de los datos muestrales (evidencia) y calcúlese el valor de la estadística de prueba

En este momento se reúnen los datos muestrales y se calcula el valor de la estadística muestral que corresponde al parámetro referido en la hipótesis nula. A partir de tal estadística se determina el valor de la estadística de prueba

Paso 4: Tomar una decisión e interpretarla

Adóptese una decisión comparando el valor calculado de la estadística de prueba encontrado en el paso 3, con el valor crítico de dicha estadística obtenido en el paso 2.

- **Regla de decisión :** Si el valor calculado de la estadística de prueba queda localizado dentro de la región crítica, se rechazará H_0 . En caso contrario no se podrá rechazar H_0 .

Nota : El conjunto de valores que no están en la región crítica se llama *región de aceptación*.

- **Regla para la conclusión :** Si la decisión es “ rechazar H_0 “, entonces la conclusión debe se redactada aproximadamente como “ Existe evidencia suficiente al nivel de significación α para indicar que ...(el significado de la hipótesis alternativa) “. Si la decisión es “ no se puede rechazar H_0 “, entonces la conclusión debe ser más o menos como “ No existe suficiente evidencia al nivel de significación α que indique que...(el significado de la hipótesis alternativa “

Nota : Algunas personas prefieren utilizar la frase “ se acepta H_0 “ en vez de “ no se puede rechazar H_0 “. Sin embargo, esta última forma describe la decisión más correctamente.

Ejemplo: Un presidente de una empresa afirma que el n° de llamadas solicitando servicio no es más de 15 por semana, en promedio. Para comprobar su afirmación , se revisaron los registros de servicio para $n = 36$ semanas seleccionadas al azar, el resultado fue que $\bar{X} = 17$ y $S^2 = 9$ para los datos de la muestra . ¿ Contradice la evidencia de la muestra la afirmación del presidente al nivel de significación del 5% ?

Solución

Paso 1 : Plantear las hipótesis

$$H_0 : \mu \leq 15 \quad \text{v/s} \quad H_1 : \mu > 15 \quad (\text{Aquí } (\mu_0 = 15)).$$

Paso 2 : como la varianza poblacional es desconocida (σ^2) hay que usar la tabla N°2
(ver fotocopias, parte Prueba de Hipótesis)

Paso 3 : $T = \sqrt{n} (\bar{X} - \mu_0) / S = 4$ y $T^* = t (n-1, 1-\alpha)$ y como $n > 30$,
entonces $T^* = Z_{1-\alpha} = Z_{0.95} = 1.645$

Paso 4 : Como $T > T^*$,estoy en la región de rechazo, por lo que hay evidencias
suficientes para rechazar la hipótesis nula

Conclusión : El número promedio de llamadas solicitando servicio es mayor que 15.