

**UNIVERSIDAD TECNOLÓGICA DE**

**SAN LUIS RIO COLORADO**

**Investigación, parcial 3**

**MTRA. YOHANI PAOLA VALDEZ AYON**

**ALUMNO: VICTOR MANUEL GALVAN COVARRUBIAS**

San Luis Rio Colorado, Sonora Diciembre, 2020

En inferencia estadística se llama estimación al conjunto de técnicas que permiten dar un valor aproximado de un parámetro de una población a partir de los datos proporcionados por una muestra.

Estimación puntual

Una estimación es puntual cuando se usa un solo valor extraído de la muestra para estimar el parámetro desconocido de la población. Al valor usado se le llama estimador.

La media de la población se puede estimar puntualmente mediante la media de la muestra:

La proporción de la población se puede estimar puntualmente mediante la proporción de la muestra:

La desviación típica de la población se puede estimar puntualmente mediante la desviación típica de la muestra, aunque hay mejores estimadores:

Estimación por intervalos

A veces es conveniente obtener unos límites entre los cuales se encuentre el parámetro con un cierto nivel de confianza, en este caso hablamos de estimación por intervalos.

Nivel de confianza

El nivel de confianza, C, indica, en porcentaje, con qué proporción el intervalo de confianza contiene el parámetro estimado. El coeficiente de confianza, c, es la misma proporción en tanto por uno, c = C/100. En otras palabras, c es la probabilidad de que el intervalo de confianza contenga el parámetro estimado.

Una prueba de hipótesis es una regla que especifica si se puede aceptar o rechazar una afirmación acerca de una población dependiendo de la evidencia proporcionada por una muestra de datos.

Una prueba de hipótesis examina dos hipótesis opuestas sobre una población: la hipótesis nula y la hipótesis alternativa. La hipótesis nula es el enunciado que se probará. Por lo general, la hipótesis nula es un enunciado de que "no hay efecto" o "no hay diferencia". La hipótesis alternativa es el enunciado que se desea poder concluir que es verdadero de acuerdo con la evidencia proporcionada por los datos de la muestra.

Con base en los datos de muestra, la prueba determina si se puede rechazar la hipótesis nula. Usted utiliza el valor p para tomar esa decisión. Si el valor p es menor que el nivel de significancia (denotado como α o alfa), entonces puede rechazar la hipótesis nula.

Usted puede seguir seis pasos básicos para configurar y realizar correctamente una prueba de hipótesis. Por ejemplo, el gerente de una fábrica de tuberías desea determinar si el diámetro promedio de los tubos es diferente de 5 cm. El gerente sigue los pasos básicos para realizar una prueba de hipótesis.

**NOTA**

Debe determinar los criterios para la prueba y el tamaño de muestra necesario antes de recolectar los datos.

1. Especificar las hipótesis.

En primer lugar, el gerente formula las hipótesis. La hipótesis nula es: la media de la población de todos los tubos es igual a 5 cm. Formalmente, esto se escribe como: H0: μ = 5

Luego, el gerente elige entre las siguientes hipótesis alternativas:

| **Condición que se probará** | **Hipótesis alternativa** |
| --- | --- |
| La media de la población es menor que el objetivo. | unilateral: μ < 5 |
| La media de la población es mayor que el objetivo. | unilateral: μ > 5 |
| La media de la población es diferente del objetivo. | bilateral: μ ≠ 5 |

Como tiene que asegurarse de que los tubos no sean más grandes ni más pequeños de 5 cm, el gerente elige la hipótesis alternativa bilateral, que indica que la media de la población de todos los tubos no es igual a 5 cm. Formalmente, esto se escribe como H1: μ ≠ 5

1. Elegir un nivel de significancia (también denominado alfa o α).

El gerente selecciona un nivel de significancia de 0.05, que es el nivel de significancia más utilizado.

1. Determinar la potencia y el tamaño de la muestra para la prueba.

El gerente utiliza un cálculo de potencia y tamaño de la muestra para determinar cuántos tubos tiene que medir para tener una buena probabilidad de detectar una diferencia de 0.1 cm o más con respecto al diámetro objetivo.

1. Recolectar los datos.

Recoge una muestra de tubos y mide los diámetros.

1. Comparar el valor p de la prueba con el nivel de significancia.

Después de realizar la prueba de hipótesis, el gerente obtiene un valor p de 0.004. El valor p es menor que el nivel de significancia de 0.05.

1. Decidir si rechazar o no rechazar la hipótesis nula.

El gerente rechaza la hipótesis nula y concluye que el diámetro medio de todos los tubos no es igual a 5 cm.

**Hipótesis nula (H0)**

La hipótesis nula indica que un parámetro de población (tal como la media, la desviación estándar, etc.) es igual a un valor hipotético. La hipótesis nula suele ser una afirmación inicial que se basa en análisis previos o en conocimiento especializado.

**Hipótesis alternativa (H1)**

La hipótesis alternativa indica que un parámetro de población es más pequeño, más grande o diferente del valor hipotético de la hipótesis nula. La hipótesis alternativa es lo que usted podría pensar que es cierto o espera probar que es cierto.

Ninguna prueba de hipótesis es 100% cierta. Puesto que la prueba se basa en probabilidades, siempre existe la posibilidad de llegar a una conclusión incorrecta. Cuando usted realiza una prueba de hipótesis, puede cometer dos tipos de error: tipo I y tipo II. Los riesgos de estos dos errores están inversamente relacionados y se determinan según el nivel de significancia y la potencia de la prueba. Por lo tanto, usted debe determinar qué error tiene consecuencias más graves para su situación antes de definir los riesgos.

**Error de tipo I**

Si usted rechaza la hipótesis nula cuando es verdadera, comete un error de tipo I. La probabilidad de cometer un error de tipo I es α, que es el nivel de significancia que usted establece para su prueba de hipótesis. Un α de 0.05 indica que usted está dispuesto a aceptar una probabilidad de 5% de estar equivocado al rechazar la hipótesis nula. Para reducir este riesgo, debe utilizar un valor menor para α. Sin embargo, usar un valor menor para alfa significa que usted tendrá menos probabilidad de detectar una diferencia si esta realmente existe.

**Error de tipo II**

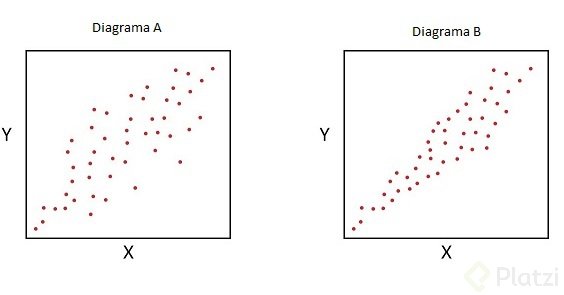
Cuando la hipótesis nula es falsa y usted no la rechaza, comete un error de tipo II. La probabilidad de cometer un error de tipo II es β, que depende de la potencia de la prueba. Puede reducir el riesgo de cometer un error de tipo II al asegurarse de que la prueba tenga suficiente potencia. Para ello, asegúrese de que el tamaño de la muestra sea lo suficientemente grande como para detectar una diferencia práctica cuando esta realmente exista.

**Cómo hacer un diagrama de dispersión paso a paso**

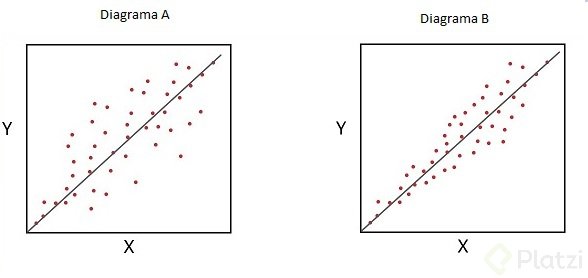
* **Paso 1**: Determina cuál es la situación. Si no entendemos qué es lo que esta ocurriendo, no podremos establecer las variables a estudiar.
* **Paso 2**: Determina las variables a estudiar. Si ya determinaste las variables a estudiar, es porque crees que puede existir una relación entre ellas que te permita caracterizar la situación.
* **Paso 3**: Recolecta los datos de las variables: Si ya los tienes, perfecto. Si no, definimos un período de tiempo para conseguir los datos de las variables antes definidas. Recuerda que los datos de las dos variables deben estar dados en el mismo período de tiempo.
* **Paso 4**: Ubica los valores en el eje respectivo. Por lo general, la variable independiente es aquella que no está influenciada por la otra y se ubica en el *eje x*. La variable dependiente que es la que se ve afectada por la otra variable se ubica en el eje y. Así pues, procedemos a ubicar los valores en el plano cartesiano de acuerdo a su variable (x, y)
* **Paso 5**: Determina el coeficiente de correlación: El coeficiente de correlación debe verse reflejado en la forma que toma el gráfico de dispersión. Es el cociente de la covarianza y la multiplicación de la desviación típica de las dos variables. Con excel logramos calcularlo de manera muy simple.
* **Paso 6**: Analizamos: Con base en el coeficiente y en el gráfico, definimos cuál es la relación de las dos variables y tomamos las decisiones pertinentes.

El Coeficiente de correlación es una medida que permite conocer el grado de asociación lineal entre dos variables cuantitativas (X, Y).

En los siguientes Diagramas de dispersión se puede observar que existe una relación lineal entre la variable X y la variable Y.



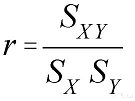
Sin embargo si trazamos una linea recta en los diagramas.



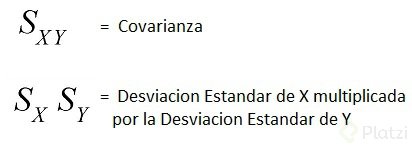
Podemos observar que en un diagrama B los puntos se acercan más a la recta, caso contrario en el diagrama A, los puntos están más alejados. Entonces podemos decir que la relación lineal del diagrama A es más débil con comparación a la relación que existe en el diagrama B.

Un diagrama dispersión no nos da certeza de que tan débil o fuerte es la relación lineal, necesitamos una medida que nos de la fuerza de la asociación y la dirección que toma esta relación.

Para esto sirve el coeficiente de correlación que está dado por la siguiente formula.

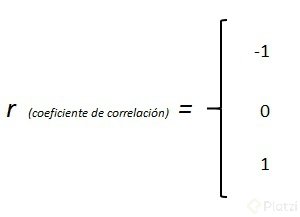


¿De dónde sacamos estos valores?

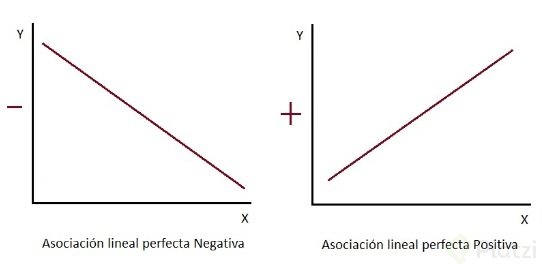


Recordar entonces que el coeficiente de relación lineal, mide la fuerza y el sentido de la relación lineal entre 2 variables cuantitativas.

Luego de haber aplicado la formula, según el resultado se puede clasificar en este rango.



Entre más cercano es a 1 es más fuerte, entre más cercano a 0 es débil hasta llegar hacerse nula, si el valor del coeficiente de relación es -1 es una **Asociación lineal perfecta Negativa**, si es 0 no existe relación y si es 1 es una **Asociación Lineal perfecta Positiva**.



¿Entonces como identificamos cuan dispersos son los puntos de la línea recta?

Con la siguiente tabla podemos clasificar nuestros resultados y responder esta interrogante.

