



ESCUELA SUPERIOR DE INGENIERÍA Y TECNOLOGÍA

**Laboratorio: Estadística Inferencial: del análisis muestral
a la predicción poblacional**

Realizado por:
Nicolás Felipe Trujillo Montero

Asignatura:
Análisis e Interpretación de Datos

Universidad:
Universidad Internacional de La Rioja, 26006, Logroño, España

UNIR, FEBRERO DE 2023

Fecha: 17 de febrero de 2023

Índice

1. Resumen	3
2. Introducción	3
3. Desarrollo	5
4. Conclusión	7

1

Resumen

Se trata de una actividad la cual se enfoca en aprender y poner en práctica los principales aspectos con respecto al tema de la estadística inferencial y los contrastes de hipótesis.

Para ello, se va a concebir dos partes de la actividad, una primera parte donde se va a cuestionar los aspectos más teóricos sobre esta teoría, y una segunda parte donde se va a poner en práctica.

2

Introducción

Siguiendo con lo planteado en el resumen, introduciremos una primera aproximación al empleo del contraste de hipótesis resolviendo un par de preguntas:

- ¿Qué es un contraste de hipótesis? (Minitab, [2023a](#))

Definimos el concepto como un criterio por el cual podemos aceptar o denegar una hipótesis planteada sobre una población en base al estudio de la hipótesis en la muestra general. Para llegar al objetivo de verificar si se cumple o no la hipótesis, se estudiará dos casos: la aceptación de la hipótesis nula (caso en el cual podremos afirmar que la hipótesis planteada no es correcta) y la denegación de ésta (caso en el cual podremos afirmar que la hipótesis planteada es correcta).

Este estudio se realiza estudiando el p-value, tal que:

- $PV < \alpha$: rechazamos hipótesis nula.
- $PV \geq \alpha$: aceptamos hipótesis nula.

- ¿Para qué se utilizan? (Minitab, [2023b](#))

Una de las cosas que suele hacer el ser humano es utilizar el sentido común, es decir, realizar deducciones (o en este caso inferencias) acerca de lo que nos rodea en base a lo que

vemos.

Poniendo un ejemplo: si una persona X desde su nacimiento no tuviera contacto ninguna vez con coches y le enseñáramos un coche Y por primera vez, dicha persona tendería a pensar que en el supuesto caso de que viera más coches serían iguales a ese (Realiza una deducción en base a lo que ha visto: muestra pequeña sobre una muestra grande que serían todos los coches del mundo).

Estas pruebas se encargan de aceptar o denegar una hipótesis verificando la consistencia que presentan con respecto a una muestra de datos. En el caso de la persona X podría plantear la hipótesis de que todos los coches son iguales que el coche Y, planteando el estudio de la inconsistencia de la hipótesis.

- Planteamiento de un algoritmo de Prueba de Hipótesis para una o dos población(es) (Duca, 2023; Barai, 2023; Exam, 2023) En esta sección, lo mejor que se puede plantear es hacer una mezcla de todo lo aprendido en las referencias para plantear problemas.

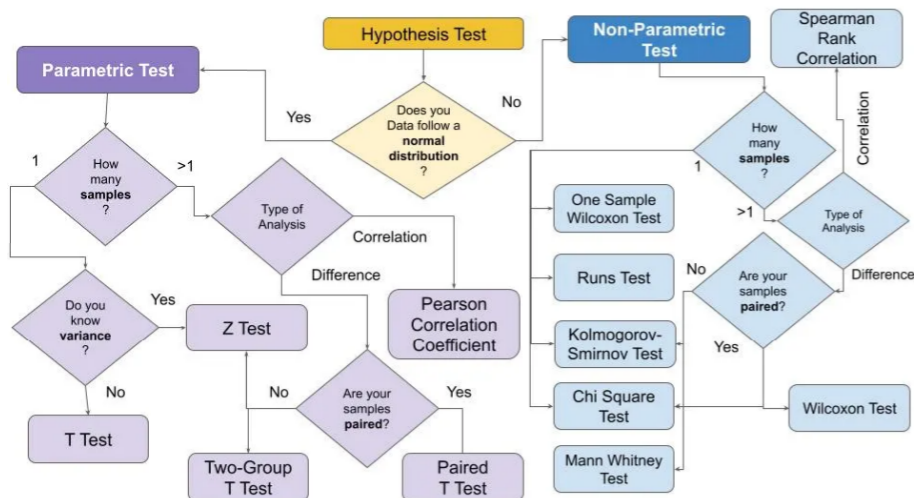


Figura 1: Diagrama de Contraste propuesto en (Duca, 2023)

Para comenzar, primero tenemos que presentar una idea de qué hipótesis plantear (1) analizando de forma descriptiva los datos, una vez realizado observamos si los datos planteados siguen una distribución normal $Data N(0, 1)$ (2) y si partimos de una sola muestra de la población o más (3).

Una vez que ya nos hayamos encauzado en un flujo u otro del diagrama, faltaría plantear

el tipo de análisis que queremos hacer en función de nuestra hipótesis (4) Seguido de esto, declarar el umbral correspondiente para aceptar o rechazar la hipótesis α calculando el estadístico (P-Value) (5) y llegar a una conclusión (6). Es importante en última instancia, comprobar si no nos encontramos en uno de los dos errores (Falso Positivo "Tipo 1.º Falso Negativo "Tipo 2") (7).

3

Desarrollo

Una vez contemplados los conceptos más técnicos de la estadística inferencial, plantearemos el problema.

Uno de los tópicos más comunes en los últimos tiempos en España es la natalidad de los extranjeros, con respecto a los españoles, por lo que nos vamos a centrar en el estudio de la hipótesis de creer que los extranjeros tienen más hijos que los españoles en los últimos años.

Para dicho estudio, primero partimos de unos datos obtenidos del Instituto Nacional de Estadística INE (INE, [2023](#)). Dichos datos están filtrados de forma que nos muestra el porcentaje, desde el año 2002 hasta el 2021, de nacimientos por parte de españoles y por parte de extranjeros.

Viendo la distribución de los nacimientos tanto de españoles como de los extranjeros, y sus respectivos valores, podemos ver como:

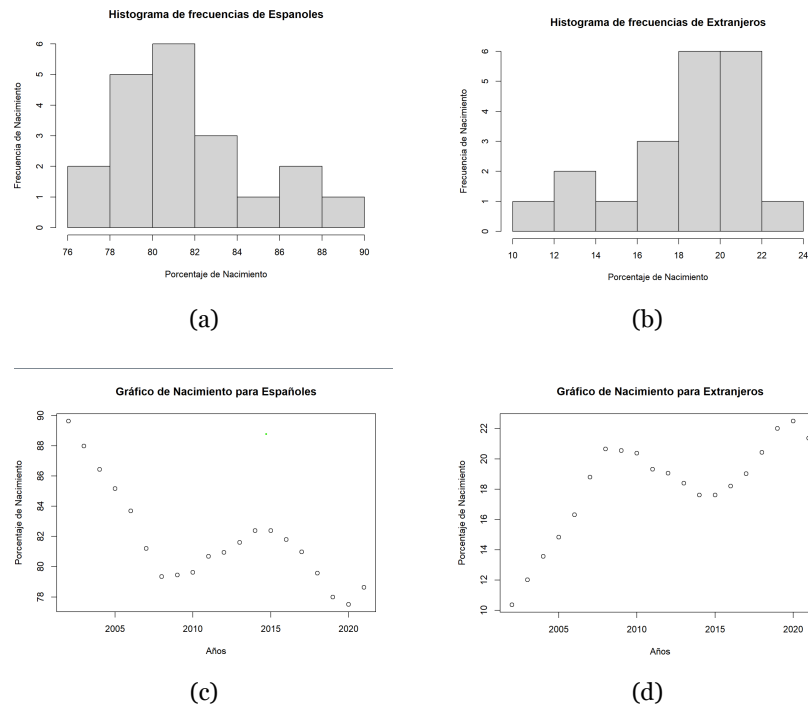


Figura 2: Distribuciones y gráficas

Esto es nuestro análisis principal descriptivo, ya que no hay mucho más que estudiar según el algoritmo propuesto con anterioridad.

Una vez planteado las gráficas, es momento de estudiar la hipótesis planteada con anterioridad. Antes de eso, es importante definir el flujo que se va a llevar respecto al algoritmo. Se ha realizado el paso 1, el paso 2 es que siguen distribuciones no normales, el paso 3 es que partimos de dos muestras y el paso 4 es que si muestras paired Pro, [2023](#) ya que cuando un porcentaje de la población decrece, el otro aumenta, y viceversa. Una vez que sabemos eso, vamos a plantear dos formas de ver la hipótesis ($\alpha = 0,05$):

- **Spearman Rank Correlation**: Aplicando Spearman podemos ver si las variables están correlacionadas cuando no siguen una distribución normal, dándonos un P-Value de $5.976e-06$ se rechaza la hipótesis nula, o lo que es lo mismo se verifica que están correlacionadas.
- **Chi-Square Test**: Aplicando Chi podemos apreciar de otra forma si las variables están relacionadas, dándonos un P-Value de 0.2358 por lo que se acepta la hipótesis nula que en

este caso es que son dependientes entre sí.

4

Conclusión

Una vez hecho todos los cálculos y contrastes de hipótesis pertinentes podemos ver como para ambos test se cumple la hipótesis planteada de que los extranjeros están teniendo más hijos que los españoles conforme va pasando el tiempo.

Por último, el estudio de errores de tipo I y II no se han planteado ya que se verifica correctamente todas las estimaciones oportunas.

Referencias

- Minitab. (2023a). *¿Qué es una prueba de hipótesis?* <https://www.addlink.es/noticias/minitab/2852-que-es-una-prueba-de-hipotesis>
- Minitab. (2023b). *¿Por qué es necesario utilizar pruebas de hipótesis en estadística?* <https://www.addlink.es/noticias/minitab/2870-comprendamos-las-pruebas-de-hipotesis-por-que-es-necesario-utilizar-pruebas-de-hipotesis-en-estadistica>
- Duca, A. L. (2023). *Hypothesis Tests Explained*. <https://towardsdatascience.com/hypothesis-tests-explained-8a070636bd28>
- Barai, A. (2023). *Hypothesis Testing | All a beginner needs to know*. <https://medium.com/analytics-vidhya/hypothesis-testing-all-a-beginner-needs-to-know-dd1bdeebdbf1>
- Exam, P. (2023). *Outline of Hypothesis Testing*. <https://www.processexam.com/six-sigma-tools-hypothesis-testing>
- INE. (2023). *Datos Históricos de Nacionalidad en España*. <https://www.ine.es/jaxiT3/Tabla.htm?t=2777&L=0>
- Pro, Q. (2023). *Paired Samples vs Independent Samples: The Differences*. <https://www.questionpro.com/blog/paired-samples/>