Pruebas estadísticas

Fabiola Vázquez

12 de octubre de 2020

---

1. Introducción

El objetivo de este estudio es mostrar la realización de distintas pruebas estadísticas [6] con el software R versión 4.0.2 [7]. Estas pruebas se realizaron a datos extraídos de la página oficial del INEGI [2] en un cuaderno de Jupyter [5]. Algunas preguntas respecto a la selección y aplicación de pruebas estadísticas se contestan en el Apéndice A.

2. Análisis

Se trabaja con dos conjuntos de datos, un conjunto consiste de la cantidad de mujeres por cada cien hombres, dividido por año y entidad federativa. El cuadro 1 muestra un fragmento de los datos con los que se trabaja. El otro conjunto considera los nacimientos totales por entidad federativa en los años de 2017 a 2019, un fragmento de los datos se muestra en el cuadro 2.

2.1. Prueba de Shapiro

Para el uso de algunas pruebas estadísticas, es necesario saber si las muestras siguen o no una distribución normal. Para ello se utiliza la prueba de Shapiro, donde las hipótesis son las siguientes:

$$H_0 : \text{La muestra tiene una distribución normal.}$$

$$H_1 : \text{La muestra no tiene una distribución normal.}$$

Por ejemplo, aplicando esta prueba a la cantidad de mujeres por cada cien hombres en 1990, se obtiene un valor p de 0.5985. Como dicho valor es mayor que 0.05, no se rechaza la hipótesis H_0. Esta prueba se realiza para todos los años involucrados en los datos (ver cuadro 1). El cuadro 3a muestra los valores p obtenidos, que en todos los casos son mayores a 0.05, por lo que en ninguno se rechaza la hipótesis nula.

Adicionalmente, se realiza una prueba de Shapiro a el total de nacimientos en el año 2017 por entidad federativa, obteniendo un valor p de $8{,}086 \times 10^{-5}$, por lo que se rechaza la hipótesis nula y se concluye que la muestra no tiene una distribución normal.

Cuadro 1: Cantidad de mujeres por cada cien hombres.

	Estado	1990	1995	2000	2005	2010	2015
1	Total	96.50	97.10	95.40	94.80	95.40	94.40
2	Aguascalientes	94.80	95.90	93.60	93.70	94.80	95.20
3	Baja California	100.40	101.60	101.40	101.40	101.80	99.10
4	Baja California Sur	103.80	103.50	104.10	104.20	104.40	101.80
5	Campeche	100.90	101.10	99.40	98.00	98.30	96.20
6	Coahuila	98.60	98.80	98.50	98.30	98.60	98.00

Cuadro 2: Nacimientos por entidad federativa.

	Estado	Total 2017	Hombres 2017	Mujeres 2017	Total 2018	Hombres 2018	Mujeres 2018
1	Total	2234039	1134349	1099674	2162535	1098674	1063826
2	AGS	26955	13730	13225	25938	13074	12864
3	BC	61840	31263	30576	60174	30692	29479
4	BCS	12573	6353	6220	11917	6076	5841
5	Campeche	17034	8678	8356	16247	8266	7980
6	Coahuila	58393	29618	28775	56718	29018	27700

Cuadro 3: Resultado de algunas pruebas estadísticas.

(a) Pruebas de normalidad de Shapiro-Wilk.

Datos	Valor p
1990	0.3155
1995	0.9237
2000	0.1590
2005	0.1904
2010	0.2152
2015	0.9516

(b) Pruebas de t de Student.

Datos	μ	Valor p
1990	97.5	0.5985
1995	97.5	0.1429
2000	97.5	0.1546
2005	95.5	0.5686
2010	96.5	0.9596
2015	95.5	0.8607

(c) Pruebas de Wilcoxon.

Datos	μ	Valor p
2017	62500	0.9923
2018	62500	0.8848
2019	62500	0.7352

2.2. Prueba t de Student

Es una prueba paramétrica usada para comprobar si es razonable que la media de una muestra que sigue una distribución normal es un valor μ. Las hipótesis son las siguientes:

$$H_0 : \text{La media real es igual a } \mu.$$

$$H_1 : \text{La media real no es igual a } \mu.$$

Por ejemplo, se considera nuevamente el año de 1990 del cuadro 1 y se quiere comprobar si la media μ es igual a 97.5. Para esto se realiza la prueba `t.test` con dicho valor μ, y se obtiene un valor p igual a 0.5985 por lo que se acepta la hipótesis H_0. El cuadro 3b muestra el valor μ y los valores p que se obtienen al aplicar la prueba a los demás años. Como los valores p son siempre mayores a 0.05, se acepta la hipótesis H_0.

2.3. Prueba de los rangos con signo de Wilcoxon

La diferencia entre esta prueba y la prueba t de Student reside en el hecho que en la última la muestra a la que se le aplica la prueba tiene que seguir una distribución normal, y en la prueba de Wilcoxon no necesariamente. Sin embargo, ambas sirven para verificar si la media de una muestra puede ser un valor específico μ. Las hipótesis son iguales a la prueba t.

Por ejemplo, considerando los datos del cuadro 2, se aplica la prueba a los nacimientos totales del año 2017 para verificar si tiene una media de μ igual a 62,500. En esta prueba se obtiene un valor p de 0.3846, el cual es mayor que 0.05, por lo que se acepta la hipótesis H_0. Esta prueba se realiza a los otros dos años (2018 y 2019), los valores p se muestran en el cuadro 3c.

2.4. Para dos muestras

Las dos anteriores pruebas, se pueden aplicar a dos muestras diferentes. Las diferencias siguen siendo las mismas: la prueba t necesita que las muestras sigan una distribución normal y la prueba de Wilcoxon no necesariamente. Las hipótesis son,

$$H_0 : \text{La diferencia entre las medias es igual a cero.}$$

$$H_1 : \text{La diferencia entre las medias no es igual a cero.}$$

Por ejemplo, considerando los datos del cuadro 1, se quiere comparar las medias de los años 1990 y 1995, se realiza una prueba t de Student para dos muestras y se obtiene un valor p de 0.5385 por lo cual, se concluye que la diferencia entre las medias es igual a cero.

En otro ejemplo, considerando ahora, los datos del cuadro 2, se quiere comparar las medias de los nacimientos del año 2017 y 2018, realizando una prueba de Wilcoxon para dos muestras. Se obtiene un valor p de 0.3846, el cual es mayor que 0.05, por lo que se concluye que la diferencia entre las medias de dichos años es igual a cero.

2.5. Prueba de Kolmogorov-Smirnov

Esta prueba es para verificar si dos muestras siguen la misma distribución, las hipótesis son,

$$H_0 : X \text{ y } Y \text{ siguen la misma distribución.}$$

$$H_1 : X \text{ y } Y \text{ no siguen la misma distribución.}$$

Como ejemplo, se considera los nacimientos por año y se quiere verificar si los datos del año 2017 y los del 2018 siguen la misma distribución, usando `ks.test` a dichos datos, se obtiene un valor p igual a 0.9991 el cuál es mayor que 0.05, por lo que se acepta la hipótesis H_0.

2.6. Prueba F de Fisher

Esta prueba sirve para verificar si dos muestras tienen la misma varianza. Las hipótesis son,

$$H_0 : X \text{ y } Y \text{ tienen la misma varianza.}$$

$$H_1 : X \text{ y } Y \text{ no tienen la misma varianza.}$$

Como ejemplo, se consideran los nacimientos de hombres y mujeres en el año 2017 por entidad federativa y se verifica si tienen la misma varianza. Se utiliza la función `var.test` y se obtiene un valor p de 0.8592, por lo cual se concluye que tienen la misma varianza.

2.7. Prueba χ^2

El objetivo de esta prueba es verificar si dos variables categóricas son dependientes. Por ejemplo, consideramos la cantidad total de nacimientos por entidad federativa en el año de 2017, y el total de nacimientos de mujeres por entidad federativa en el mismo año. Al aplicar una prueba χ^2 se obtiene un valor p igual a 0.2373, el cual es mayor que 0.05, por lo que se concluye que las variables son dependientes.

2.8. Correlación

Sirve para probar la relación lineal entre dos variables continuas. Como ejemplo, se considera los nacimientos de hombres y mujeres en el año 2018 por entidad federativa. Al aplicar la función `cor.test` se obtiene un valor p igual a $2{,}2 \times 10^{-16}$, el cual es menor que 0.05, por lo que se concluye que no hay correlación entre las variables.

A. Preguntas

Las siguientes son algunas preguntas frecuentes que surgen a la hora de escoger una prueba estadística, o interpretar el resultado de la misma. A su vez, se muestra en el cuadro 4 una guía sobre cómo escoger la prueba estadística adecuada a la ocasión.

- **¿Cuál es la relación entre contraste de hipótesis y pruebas estadísticas?** Ambas son un procedimiento para evaluar la evidencia que los datos proporcionan para probar o rechazar una hipótesis [1].
- **¿Qué indicaría rechazar la hipótesis nula?** Los datos proporcionan suficiente evidencia contra la hipótesis nula H_0 y se considera la hipótesis alternativa H_1 [1].
- **¿Cómo se interpreta la salida de una prueba estadística?** Primeramente, al diseñar un estudio, se específica un nivel de significación α, que está entre 0 y 1, por encima del cual H_0 no debería rechazarse. La prueba estadística aplicada, produce un valor, denominado valor p entre 0 y 1, si el valor $p < \alpha$, se rechaza la hipótesis nula y se acepta la hipótesis alternativa con un riesgo de ser errónea. En caso contrario, si el valor $p > \alpha$ no se rechaza H_0, pero no significa que debamos aceptarla [1].
- **¿Cómo seleccionar el alfa?** La elección del alfa depende de cuán peligroso sea rechazar H_0 en el caso de que sea verdadera [1].
- **¿Cuáles son los errores frecuentes de interpretación del valor p?**
- **¿Qué es la potencia estadística y para qué sirve?** Es la capacidad de un experimento o una prueba para conducir al rechazo de la hipótesis nula [1].
- **Ejemplos de pruebas estadísticas paramétricas y no paramétricas.** Las prueba t de Student, el coeficiente de correlación de Pearson, regresión lineal, ANOVA, son ejemplos de pruebas paramétricas. Pruebas como la de χ^2, coeficientes de correlación e independencia para tabulaciones cruzadas, coeficientes de correlación por rangos ordenados Spearman y Kendall [3].
- **Resume LA GUÍA para encontrar la prueba estadística que buscas.**
  - Definir de forma clara el objetivo del análisis.
  - Identificar el tipo de variables.
  - Identificar si las muestras son independientes o no.
  - Analizar los supuestos para verificar si se puede emplear técnicas parámetricas.
  - Seleccionar una prueba adecuada según el cuadro 4 [4].
- **¿Cuáles son los supuestos para aplicar técnicas paramétricas?** Las observaciones deben ser independientes entre sí, las poblaciones deben hacerse en poblaciones distribuidas normalmente y deben tener la misma varianza, las variables deben haberse medido por lo menos en una escala de intervalo de manera que sea posible utilizar las operaciones aritméticas [3].

Cuadro 4: Resumen de las funciones a utilizar en R para cada tipo de prueba [4].

Objetivo	Gaussiana	No gaussiana	Numéricos	Nominal binaria
Comparar 2 grupos independientes	`t.test(y~g)`	`wilcox.test( y~g)`	`yuen(y~g)`	`fisher.test(M)` `chisq.test(M)`
Comparar 2 grupos relacionados	`t.test(y~g, paired=T)`	`wilcox.test( y~g, paired=T)`	`yuen(y1~y2)`	`mcnemar.test(M)`
Comparar 3 o más grupos independientes	`aov(y~g)` `pairwise.t.test( y,g)`	`kruskal.test(y~g)` `kruskalmc(y ~g)`	`t1way(y~g)` `lincon(y~g)`	`chisq.test(M)` `fisher.multcomp(M)`
Comparar 3 o más grupos relacionados	`ezANOVA(dv, wid, within)` `pairwise.t. test(y,x)`	`friedman.test (y~g|id)` `pairwise.wilcoxon. test(y,g)`	`rmanova(y,g, block)` `rmmcp(y, g, block)`	`mantelhaen.test(M)`
Asociar 2 variables	`cor.test(x,y)`	`cor.test(x,y, metod="spearman")`	`pbcor(x,y)`	`assocstats(M)`

Referencias

[1] Addinsoft. ¿Qué es una prueba estadística?. https://help.xlstat.com/s/article/que-es-una-prueba-estadistica?language=es#:~:text=Una%20prueba%20estad%C3%ADstica%20es%20una,nula%2C%20y%20suele%20denominarse%20H0.&text=H0%20normalmente%20se%20opone%20a,alternativa%2C%20denominada%20H1%20o%20Ha.

[2] Instituto Nacional de Estadística y Geografía. Relación hombres-mujeres por entidad federativa, 1990 a 2015. https://www.inegi.org.mx/app/tabulados/interactivos/?px=Poblacion_02&bd=Poblacion.

[3] EcuRed. Pruebas estadísticas. https://www.ecured.cu/Pruebas_estad%C3%ADsticas.

[4] Rosana Ferrero. Guía definitiva para encontrar las pruebas estadísticas que buscas. https://www.maximaformacion.es/blog-dat/guia-para-encontrar-tu-prueba-estadistica/.

[5] Thomas Kluyver, Benjamin Ragan-Kelley, Pérez, et al. Jupyter notebooks—a publishing format for reproducible computational workflows. In *Positioning and Power in Academic Publishing: Players, Agents and Agendas: Proceedings of the 20th International Conference on Electronic Publishing*, page 87. IOS Press, 2016.

[6] Selva Prabhakaran. Statistical tests. http://r-statistics.co/Statistical-Tests-in-R.html.

[7] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020.