

Tutorial de Estadística ^[1]

[Explorable.com](#) ^[2]43.8K visitas

Este tutorial de estadística es una guía para ayudarle a entender los conceptos clave de la estadística y cómo dichos conceptos se relacionan con el método científico y la investigación.

Los científicos suelen utilizar la estadística para analizar ^[3] sus resultados. ¿Por qué los investigadores utilizan la estadística? ^[4] Porque gracias a esta pueden comprender un fenómeno al confirmar o rechazar una hipótesis. Es de vital importancia para la mayoría de las teorías científicas y para la forma en que adquirimos el conocimiento

Sin embargo, usted no necesita ser un científico; cualquiera que desee aprender acerca de cómo la estadística ayuda a los investigadores, querrá leer este tutorial de estadística para el método científico.

¿Qué es estadística? ^[5]

Investigación de Datos

Esta sección del tutorial de estadística consiste en comprender cómo los datos son adquiridos y usados.

A menudo, los resultados de una investigación científica contienen muchos más datos o información de la que los investigadores necesitan. Estos datos o información, son llamados datos en bruto.

Para poder analizar los datos de manera eficiente, los datos en bruto son procesados ^[6] en "datos de salida ^[7]". Hay muchos métodos para procesarlos pero, básicamente, el científico organiza y resume los datos en bruto en un grupo más eficiente. Cualquier tipo de información organizada puede ser llamada un "conjunto de datos ^[8]".

Posteriormente, los investigadores pueden aplicar diferentes métodos estadísticos para analizar y comprender mejor los datos (y de forma más exacta). Dependiendo de la investigación, el científico también puede utilizar la estadísticas descriptiva ^[9] o para la investigación exploratoria.

Lo mejor de los datos en bruto es que se puede retroceder y comprobar las cosas si tal vez sospecha que algo que pensaba en un principio ha cambiado. Esto sucede después de haber analizado el significado de los resultados.

Los datos en bruto pueden dar ideas para nuevas hipótesis dado que se obtiene una mejor perspectiva de lo que está pasando. Además, también puede controlar las variables que podrían influir en la conclusión [10] (por ejemplo terceras variables [11]). En estadística, un parámetro [12] es cualquier cantidad numérica que caracteriza a una población determinada o algún aspecto de ella.

Tendencia Central y Distribución Normal

Esta parte del tutorial de estadística le ayudará a entender la distribución, la tendencia central y su relación con los conjuntos de datos [8].

Muchos datos del mundo real se distribuyen normalmente [13], es decir, una curva de frecuencia, o una distribución de frecuencia [14], que tiene el número más frecuente cerca del medio. Muchos experimentos se basan en asunciones de una distribución normal [15] y esta es una de las razones por las cuales, muy a menudo, los investigadores hacen mediciones de tendencia central [16] en las investigaciones estadísticas, tales como la media [17] (media aritmética [18] o media geométrica [19]), mediana [20] o la moda [21].

La tendencia central puede dar una idea bastante buena acerca de la naturaleza de los datos (la media, mediana y moda muestran el "valor medio"), especialmente cuando se combina con mediciones acerca de cómo los datos son distribuidos. Los científicos normalmente calculan la desviación estándar [22] para medir dicha distribución.

Sin embargo, existen varios métodos para esta medición de datos: la varianza [23], desviación estándar [24], error estándar de la media [25], error estándar de estimación o el "rango [26]" (el cual establece los extremos en los datos).

Para crear el gráfico de la distribución normal, usualmente tendrá que usar la media aritmética [18] de una "muestra lo suficientemente grande [27]" y tendrá que calcular la desviación estándar.

Sin embargo, la distribución de la muestra [28] no se distribuirá normalmente si la distribución está sesgada (claro esta) o si tiene valores atípicos [29] (frecuentes resultados atípicos o errores de medición) que dañen los datos. Un ejemplo de una distribución que no se distribuye normalmente es la distribución F [30], que está sesgada a la derecha.

Es por este motivo que generalmente los investigadores tienden a comprobar que sus resultados se distribuyan normalmente usando el rango, mediana y moda. Si la distribución no tiene una distribución normal, esto influirá en la prueba / método a elegir para el análisis.

Otras herramientas

- Cuartil [31]
- Trimean [32]

Comprobación de hipótesis - Tutorial de Estadística

¿Cómo podemos saber si una hipótesis es correcta o no?

¿Por qué utilizar la estadística para determinar esto? [4]

Emplear la estadística en la investigación implica mucho más que utilizar fórmulas estadísticas o conocer este tipo de softwares.

Hacer uso de la estadística en la investigación consiste básicamente en

1. Aprender estadística básica [33]
2. Comprender la relación entre probabilidad y estadística [34]
3. Comprensión de las dos principales ramas de la estadística [35] : la estadística descriptiva [9] y estadística inferencial [36] .
4. El conocimiento de cómo la estadística se relaciona con el método científico [4] .

La estadística en la investigación no se trata sólo de fórmulas y cálculos. (Muchas conclusiones erróneas se han llevado a cabo por no entender los conceptos básicos de estadística)

La Estadística Inferencial nos ayuda a sacar conclusiones [10] a partir de muestras de una población.

Durante la realización de experimentos [37] , una de las partes más críticas es comprobar hipótesis [38] entre ellas. Por lo tanto, es una parte importante del tutorial de estadística para el método científico.

La comprobación de hipótesis se lleva a cabo mediante la formulación de una hipótesis alternativa [39] la cual se prueba contra la hipótesis nula [40] , el punto de vista común. Las hipótesis se comprueban estadísticamente [41] entre ellas.

El investigador puede obtener un intervalo de confianza [42] , el cual define los límites para considerar un resultado como de apoyo a la hipótesis nula o a la hipótesis alternativa de investigación [39].

Esto significa que no todas las diferencias entre el grupo experimental y el grupo control [43] pueden ser aceptadas como apoyo a la hipótesis alternativa - el resultado debe diferir de forma estadísticamente significativa [44] para que el investigador pueda aceptar la hipótesis alternativa. Esto se hace usando una prueba de significación [45] (otro artículo [46]).

Pero cuidado, la tergiversación de datos [47] , la manipulación o la búsqueda de datos sin comprobar posteriormente su hipótesis en un experimento controlado, puede conducir a una conclusión de causa y efecto [48] a pesar de que no exista una relación con la verdad [49] .

Dependiendo de la hipótesis, tendrá que elegir entre pruebas de una o dos colas.

A veces, el grupo de control se sustituye con probabilidad experimental [50] - frecuentemente si la investigación trata un fenómeno que es éticamente problemático [51] , económicamente muy costoso o que consume demasiado tiempo, entonces el verdadero diseño experimental [52] es sustituido por un enfoque cuasi-experimental [53] .

A menudo hay un sesgo de publicación cuando el investigador encuentra que la hipótesis

alternativa es correcta, en lugar de tener un "resultado nulo", concluyendo así que la hipótesis nula es la mejor explicación.

Si se aplica correctamente, la estadística se puede utilizar para entender la causa y efecto entre las variables de investigación [54] .

También puede ayudar a identificar las terceras variables, aunque, por otro lado, la estadística también puede ser usada para manipular y encubrir terceras variables [11] si la persona que presenta los números no tiene intenciones honestas (o el conocimiento suficiente) con sus resultados.

El mal uso de la estadística es un fenómeno común, y esto probablemente continuará mientras la gente tenga la intención de influenciar a los demás. Sin embargo, un adecuado tratamiento estadístico [55] de los datos experimentales puede ayudar a evitar el uso no ético de la estadística. La Filosofía de la estadística [56] consiste en justificar su uso apropiado, asegurando la validez estadística [57] y estableciendo en ella la ética [58] .

Aquí está otro gran tutorial de estadística que integra tanto la estadística como el método científico.

Fiabilidad y Error Experimental

Las pruebas estadísticas hacen uso de los datos de las muestras. Estos resultados son generalizados [59] a la población. ¿Cómo podemos saber que esto refleja la conclusión [10] correcta?

Contrariamente a lo que algunos podrían creer, los errores en la investigación [60] son una parte esencial de las pruebas de significación [46] . Irónicamente, la posibilidad de un error de investigación es lo que hace que la investigación sea científica. Si una hipótesis no puede ser falsificada [61] (por ejemplo, si la hipótesis tiene una lógica circular), entonces no es comprobable [62] y, por definición, no es científico.

Si una hipótesis es comprobable, está abierta a la posibilidad de salir mal. Estadísticamente, esto también abre la posibilidad de obtener errores experimentales [63] en los resultados debido a errores aleatorios u otros problemas en la investigación. Los errores experimentales también puede ser divididos en dos error de tipo I y error tipo II [60] . Las curvas ROC [64] se utilizan para calcular la sensibilidad entre los verdaderos positivos y falsos positivos.

Un análisis de potencia [65] de una prueba estadística puede determinar el número de muestras que una prueba necesitará para tener un valor de p [66] aceptable, y así rechazar una falsa hipótesis nula [40] .

El margen de error [67] está relacionado con el intervalo de confianza [42] y la relación entre la significación estadística, tamaño de la muestra y los resultados esperados [68] . El tamaño del efecto [69] estima la fuerza de la relación entre dos variables de una población. Puede que determinar el tamaño de la muestra [27] necesaria ayude a generalizar [59] los resultados a toda la población.

Replicar [70] la investigación de otros también es esencial para entender si los resultados de la investigación fueron un resultado que puede ser generalizado o sólo se debe a un

"experimento atípico" al azar. La replicación puede ayudar a identificar los errores aleatorios [71] y errores sistemáticos [72] (validez de la prueba [73]).

El Alfa de Cronbach [74] se utiliza para medir la consistencia interna o fiabilidad [75] de la calificación de una prueba.

Replicar el experimento / investigación asegura la fiabilidad estadística de los resultados [76] .

Cuando ves frecuentemente si los resultados tienen valores atípicos, es una regresión a la media [77] , lo que a su vez hace que el resultado no sea estadísticamente diferente entre el grupo experimental y el de control.

Pruebas Estadísticas

Aquí vamos a introducir algunas pruebas estadísticas / métodos frecuentemente usados ??por los investigadores.

Relación entre las variables

La relación entre las variables [78] es muy importante para los científicos. Esto les ayudará a comprender la naturaleza de lo que están estudiando. Una relación lineal [79] es cuando dos variables varían proporcionalmente, es decir, si una variable aumenta, la otra variable también lo hará. Una relación no lineal [80] es cuando las variables no varían proporcionalmente. La Correlación [81] es una manera de expresar la relación entre dos conjuntos de datos o entre dos variables.

Las Escalas de medición [82] se utilizan para clasificar, categorizar y (si es aplicable) cuantificar variables.

El Coefficiente de correlación de Pearson [83] (o correlación producto-momento de Pearson) sólo expresará la relación lineal entre dos variables. El rho de Spearman [84] es mayormente utilizado para relaciones lineales cuando se trata con variables ordinales. El coeficiente Tau de Kendall (?) puede ser empleado para medir las relaciones no lineales.

La correlación parcial [85] (y la correlación múltiple) se pueden usar cuando se controla una tercera variable [11] .

Predicciones

El objetivo de las predicciones es entender las causas, teniendo en cuenta que Correlación no necesariamente significa causa [86] . Con la regresión lineal, regularmente se mide una variable manipulada [87] .

¿Cuál es la diferencia entre correlación y regresión lineal [88] ? Básicamente, un estudio de correlación [89] observa la fuerza entre las variables , [90] mientras que la regresión lineal consiste en el mejor ajuste de línea en un gráfico.

El análisis de regresión y otras herramientas de modelado

- Regresión Lineal [91]
- Regresión Múltiple [92]

- Un análisis de trayectoria es una extensión del modelo de regresión
- Un análisis factorial [93] intenta descubrir los factores subyacentes de algo.
- El meta-análisis [94] con frecuencia hacen uso del tamaño del efecto [69]

La probabilidad bayesiana [95] es una manera de predecir la probabilidad de eventos futuros en una forma interactiva, en lugar de empezar a medir y luego obtener resultados / predicciones.

Probar Hipótesis Estadísticamente

La prueba de t de Student [96] es una prueba que puede indicar si la hipótesis nula [40] es correcta o no. En las investigaciones es comúnmente utilizada para probar las diferencias entre dos grupos (por ejemplo, entre un grupo control [43] y un grupo experimental).

La prueba t supone que los datos están más o menos normalmente distribuidos y que la varianza es igual (esto puede ser comprobado por la prueba F [97]).

La prueba t de Student [98] :

- Prueba T para una muestra única [99]
- Prueba T para dos muestras independientes [100]
- Prueba T dependiente para muestras apareadas [101]

La prueba de Wilcoxon de rango signado [102] se puede utilizar para datos no paramétricos.

Una Prueba Z [103] es similar a una prueba t, pero generalmente no se utiliza en muestras de tamaños inferiores a 30.

Un Chi-cuadrado [104] se puede utilizar si los datos son cualitativos en lugar de cuantitativos.

Al Comparar Más de Dos Grupos

Un ANOVA [105], o análisis de la varianza, se utiliza cuando se desea probar si hay diferente variabilidad entre grupos en lugar de medias diferentes. El Análisis de Varianza también se puede aplicar a más de dos grupos. Por otro lado, la distribución F [30] se puede emplear para calcular los valores de p [66] para el ANOVA.

Análisis de varianza

- ANOVA de una vía [106]
- ANOVA de dos vías [107]
- ANOVA factorial [108]
- ANOVA de medidas repetidas [109]

Estadística No Paramétrica

Algunos métodos comunes utilizando estadística no paramétrica [110] :

- Kappa de [111]Cohen [111]
- Prueba U de Mann-Whitney [112]
- Coeficiente de Correlación por Rangos de Spearman [84]

Otros términos importantes en materia de estadística

- Variables Discretas [113]

Fuente URL: <https://staging.explorable.com/es/tutorial-de-estadistica>

Enlaces

- [1] <https://staging.explorable.com/es/tutorial-de-estadistica>
- [2] <https://staging.explorable.com/en>
- [3] <https://staging.explorable.com/statistical-analysis>
- [4] <https://staging.explorable.com/why-statistics-matter>
- [5] <https://staging.explorable.com/what-is-statistics>
- [6] <https://staging.explorable.com/raw-data-processing>
- [7] <https://staging.explorable.com/data-output>
- [8] <https://staging.explorable.com/statistical-data-sets>
- [9] <https://staging.explorable.com/descriptive-statistics>
- [10] <https://staging.explorable.com/es/sacar-conclusiones>
- [11] <https://staging.explorable.com/es/confounding-variables-es>
- [12] <https://staging.explorable.com/parameters-and-statistics>
- [13] <https://staging.explorable.com/normal-probability-distribution>
- [14] <https://staging.explorable.com/frequency-distribution>
- [15] <https://staging.explorable.com/normal-distribution-assumptions>
- [16] <https://staging.explorable.com/measures-of-central-tendency>
- [17] <https://staging.explorable.com/statistical-mean>
- [18] <https://staging.explorable.com/arithmetical-mean>
- [19] <https://staging.explorable.com/geometric-mean>
- [20] <https://staging.explorable.com/calculate-median>
- [21] <https://staging.explorable.com/statistical-mode>
- [22] <https://staging.explorable.com/calculate-standard-deviation>
- [23] <https://staging.explorable.com/statistical-variance>
- [24] <https://staging.explorable.com/measurement-of-uncertainty-standard-deviation>
- [25] <https://staging.explorable.com/standard-error-of-the-mean>
- [26] <https://staging.explorable.com/range-in-statistics>
- [27] <https://staging.explorable.com/es/tamano-de-la-muestra>
- [28] <https://staging.explorable.com/es/distribucion-de-muestreo>
- [29] <https://staging.explorable.com/statistical-outliers>
- [30] <https://staging.explorable.com/f-distribution>
- [31] <https://staging.explorable.com/quartile>
- [32] <https://staging.explorable.com/trimean>
- [33] <https://staging.explorable.com/learn-statistics>
- [34] <https://staging.explorable.com/probability-and-statistics>
- [35] <https://staging.explorable.com/branches-of-statistics>
- [36] <https://staging.explorable.com/inferential-statistics>
- [37] <https://staging.explorable.com/es/la-realizacion-de-un-experimento>
- [38] <https://staging.explorable.com/es/prueba-de-la-hipotesis>
- [39] <https://staging.explorable.com/es/hipotesis-de-investigacion>
- [40] <https://staging.explorable.com/es/hipotesis-nula>
- [41] <https://staging.explorable.com/statistical-hypothesis-testing>
- [42] <https://staging.explorable.com/statistics-confidence-interval>
- [43] <https://staging.explorable.com/es/grupo-de-control-cientifico>

[44] <https://staging.explorable.com/statistically-significant-results>
[45] <https://staging.explorable.com/significance-test-2>
[46] <https://staging.explorable.com/significance-test>
[47] <https://staging.explorable.com/data-dredging>
[48] <https://staging.explorable.com/es/cause-and-effect-es>
[49] <https://staging.explorable.com/truth-and-theory>
[50] <https://staging.explorable.com/experimental-probability>
[51] <https://staging.explorable.com/es/etica-en-la-investigacion>
[52] <https://staging.explorable.com/es/diseno-experimental-verdadero>
[53] <https://staging.explorable.com/es/diseno-cuasi-experimental>
[54] <https://staging.explorable.com/es/variables-de-investigacion>
[55] <https://staging.explorable.com/statistical-treatment-of-data>
[56] <https://staging.explorable.com/philosophy-of-statistics>
[57] <https://staging.explorable.com/es/validez-estadistica>
[58] <https://staging.explorable.com/ethics-in-statistics>
[59] <https://staging.explorable.com/es/que-es-la-generalizacion>
[60] <https://staging.explorable.com/es/type-1-error-es>
[61] <https://staging.explorable.com/falsifiability>
[62] <https://staging.explorable.com/testability>
[63] <https://staging.explorable.com/experimental-error>
[64] <https://staging.explorable.com/roc-curve-analysis>
[65] <https://staging.explorable.com/statistical-power-analysis>
[66] <https://staging.explorable.com/p-value>
[67] <https://staging.explorable.com/statistics-margin-of-error>
[68] <https://staging.explorable.com/statistical-significance-sample-size>
[69] <https://staging.explorable.com/effect-size>
[70] <https://staging.explorable.com/es/estudio-de-replicacion>
[71] <https://staging.explorable.com/random-error>
[72] <https://staging.explorable.com/systematic-error>
[73] <https://staging.explorable.com/es/validez-de-la-prueba>
[74] <https://staging.explorable.com/cronbachs-alpha>
[75] <https://staging.explorable.com/es/statistics-reliability-es>
[76] <https://staging.explorable.com/es/fiabilidad-estadistica>
[77] <https://staging.explorable.com/regression-to-the-mean>
[78] <https://staging.explorable.com/relationship-between-variables>
[79] <https://staging.explorable.com/linear-relationship>
[80] <https://staging.explorable.com/non-linear-relationship>
[81] <https://staging.explorable.com/es/la-correlacion-estadistica>
[82] <https://staging.explorable.com/measurement-scales>
[83] <https://staging.explorable.com/pearson-product-moment-correlation>
[84] <https://staging.explorable.com/spearman-rank-correlation-coefficient>
[85] <https://staging.explorable.com/partial-correlation-analysis>
[86] <https://staging.explorable.com/es/correlacion-y-causalidad>
[87] <https://staging.explorable.com/es/variable-independiente>
[88] <https://staging.explorable.com/correlation-and-regression>
[89] <https://staging.explorable.com/es/estudio-correlacional>
[90] <https://staging.explorable.com/variables-and-statistics>
[91] <https://staging.explorable.com/linear-regression-analysis>
[92] <https://staging.explorable.com/multiple-regression-analysis>
[93] <https://staging.explorable.com/factor-analysis>
[94] <https://staging.explorable.com/es/metaanalisis>
[95] <https://staging.explorable.com/bayesian-probability>
[96] <https://staging.explorable.com/students-t-test>
[97] <https://staging.explorable.com/f-test>
[98] <https://staging.explorable.com/students-t-test-2>
[99] <https://staging.explorable.com/independent-one-sample-t-test>

- [100] <https://staging.explorable.com/independent-two-sample-t-test>
- [101] <https://staging.explorable.com/dependent-t-test-for-paired-samples>
- [102] <https://staging.explorable.com/wilcoxon-signed-rank-test>
- [103] <https://staging.explorable.com/z-test>
- [104] <https://staging.explorable.com/chi-square-test>
- [105] <https://staging.explorable.com/es/anova-0>
- [106] <https://staging.explorable.com/one-way-anova>
- [107] <https://staging.explorable.com/two-way-anova>
- [108] <https://staging.explorable.com/factorial-anova>
- [109] <https://staging.explorable.com/repeated-measures-anova>
- [110] <https://staging.explorable.com/nonparametric-statistics>
- [111] <https://staging.explorable.com/cohens-kappa>
- [112] <https://staging.explorable.com/mann-whitney-u-test>
- [113] <https://staging.explorable.com/discrete-variables>