

DEFINICIONES

Análisis univariante

La estadística descriptiva univariante se centra en el análisis de una única característica o cualidad del individuo.

Consiste en el análisis de cada una de las variables estudiadas por separado, es decir, el análisis está basado en una sola variable

Análisis multivariante

El análisis multivariante es un método estadístico utilizado para determinar la contribución de varios factores en un simple evento o resultado.

- Los factores de estudio son los llamados *factores de riesgo* (bioestadística), *variables independientes* o *variables explicativas*.
- El resultado estudiado es el *evento*, la *variable dependiente* o la *variable respuesta*.

El análisis multivariante mediante técnicas de proyección sobre variables latentes tiene muchas ventajas sobre los métodos de regresión tradicionales:

- se puede utilizar la información de múltiples variables de entrada, aunque éstas no sean linealmente independientes
- puede trabajar con matrices que contengan más variables que observaciones
- puede trabajar con matrices incompletas, siempre que los valores faltantes estén aleatoriamente distribuidos y no superen un 10%
- puesto que se basan en la extracción secuencial de los factores, que extraen la mayor variabilidad posible de la matriz de las X (variables explicativas, tienen que ser dependientes) pueden separar la información del ruido. Se asume que las X se miden con ruido.

Análisis inferencial

La estadística inferencial es una parte de la estadística que comprende los métodos y procedimientos que por medio de la inducción determina propiedades de una población estadística, a partir de una parte de esta. ... Se usa para modelar patrones en los datos y extraer inferencias acerca de la población bajo estudio.

Correlación

En probabilidad y estadística, la correlación indica la fuerza y la dirección de una relación lineal y proporcionalidad entre dos variables estadísticas.

Datos cualitativos

Estos datos son los relativos a las "cualidades", este tipo de información relacionada con los adjetivos. Por ejemplo, si te pidieran que describieras tu más reciente experiencia de compra en un importante punto de venta, es posible utilizar los términos rápido,

práctico, agradable, caro o inútil. Todos estos son ejemplos de la información cualitativa.

Datos cuantitativos

Aquí todo se trata de números. Imagínate a la información cuantitativa como referente de una determinada "cantidad". Por ejemplo, tu altura, cuánto pesas, o el tiempo que te lleva ir de un lugar a otro.

DATOS CUALITATIVOS	DATOS CUANTITATIVOS
Se utiliza sobre todo para obtener una comprensión de las opiniones y motivaciones de los consumidores.	Datos que se pueden transformar en estadísticas utilizables.
Los métodos para recopilar este tipo de datos son principalmente grupos de enfoque y preguntas de texto abiertas.	Los métodos cuantitativos de recopilación de datos están mucho más estructurados.
Los métodos de recopilación de datos cualitativos varían utilizando técnicas como grupos focales, entrevistas u observaciones.	Los métodos cuantitativos de recopilación de datos incluyen encuestas de papel o encuestas online, encuestas móviles, entrevistas cara a cara o telefónicas, etc.
El volumen de datos obtenidos puede hacer que el análisis y la interpretación consuman mucho tiempo.	Las preguntas para obtener este tipo de datos tienden a ser más bien cerradas, y el tamaño de la muestra más grande.
La presencia del investigador durante la recopilación de datos, puede afectar las respuestas de los sujetos de estudio.	Se puede reunir información de un número relativamente grande de participantes.

Grado de acuerdo o índice de concordancia

En cualquier estudio de investigación una cuestión clave es la fiabilidad de los procedimientos de medida empleados.

Tradicionalmente se ha reconocido una fuente importante de error de medida en la variabilidad entre observadores. Consecuentemente, un objetivo de los estudios de fiabilidad debe consistir en estimar el grado de dicha variabilidad.

En este sentido, dos aspectos distintos entran a formar parte típicamente del estudio de fiabilidad: de una parte, el sesgo entre observadores –dicho con menos rigor, la tendencia de un observador a dar consistentemente valores mayores que otro– y de otra, la concordancia entre observadores –es decir, hasta qué punto los observadores coinciden en su medición–.

Muestras apareadas

Tenemos muestras apareadas o correlacionadas cuando sabemos de antemano que una observación está relacionada con la otra. Pueden ser observaciones tomadas al mismo tiempo, diseño pareado 1, o medidas tomadas en un mismo sujeto o unidad en dos oportunidades o tiempo distintos (diseño pareado 2).

Muestras dependientes

La variable dependiente es aquella cuyo valor depende del valor numérico que adopta la variable independiente en la función.

Las muestras dependientes son mediciones pareadas de un conjunto de elementos. Las muestras independientes son mediciones realizadas en dos conjuntos de elementos distintos. ... Si los valores de una muestra no revelan información sobre los valores de la otra muestra, entonces las muestras son independientes.

Muestras independientes

Las muestras independientes son mediciones realizadas en dos conjuntos de elementos distintos.

Modelos estadísticos lineales

En estadística, el término modelo lineal es usado en diferentes maneras de acuerdo al contexto. La manera más frecuente es en conexión con modelos de regresión y el término a menudo se toma como un sinónimo del modelo de regresión lineal.

Probabilidad $p \leq 0,05$.

Se rechaza la hipótesis nula si el valor de p asociado al resultado observado es igual o menor que el nivel de significación establecido, convencionalmente 0,05 ó 0,01. Es decir, el valor p nos muestra la probabilidad de haber obtenido el resultado que hemos obtenido si suponemos que la hipótesis nula es cierta. Significa que tenemos un 5% de probabilidades de error en las conclusiones

Prueba para una muestra

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. ... La hipótesis nula es el enunciado que se probará.

Prueba para dos muestras

Un test de hipótesis para dos muestras es similar en muchos aspectos al test para una muestra.

- Se especifica una hipótesis nula, en la mayoría de los casos se propone que las medias de las dos poblaciones son iguales y se establece la hipótesis alternativa (uni o bilateral).
- Se especifica un nivel de significación α .
- Se calcula el p -valor: la probabilidad de obtener datos cuyas medias muestrales difieren tanto o más que la diferencia observada cuando H_0 es verdadera. Si esta probabilidad es pequeña (menor que α) se rechaza H_0 y se concluye que la diferencia observada no es atribuible al azar y las medias de las dos poblaciones son diferentes.

El estadístico del test dependerá de la estructura de los conjuntos de datos. En particular es importante establecer si los datos corresponden a muestras apareadas o independientes.

Pruebas bilaterales o de dos colas

La prueba de dos colas es una prueba estadística utilizada en la inferencia, en la que se da una hipótesis estadística, H_0 (la hipótesis nula), rechazada cuando el valor de la estadística de prueba es suficiente ya sea pequeño o grande lo suficiente. ... Los nombres alternativos son pruebas de una y dos caras.

Una prueba de dos colas se asocia a una hipótesis alternativa para la cual se desconoce el signo de la potencial diferencia. Por ejemplo, supongamos que deseamos comparar las medias de dos muestras A y B. Antes de diseñar el experimento y ejecutar la prueba, esperamos que si se resalta una diferencia entre las dos medias, realmente no sabemos si A debería ser superior a B o a la inversa. Esto nos lleva a elegir una prueba de dos colas, asociada a la siguiente hipótesis alternativa: H_a : media (A) \neq media (B). Las pruebas de dos colas son con diferencia las más utilizadas.

Pruebas unilaterales o de una cola

Una prueba de una cola normalmente está asociada a una hipótesis alternativa para la cual se conoce el signo de la potencial diferencia antes de ejecutar el experimento y la prueba. En el ejemplo descrito más arriba, la hipótesis alternativa referida a una prueba de una cola podría redactarse así: media(A) < media(B) o media(A) > media(B), dependiendo de la dirección esperada de la diferencia.

Regresión lineal

En estadística la regresión lineal o ajuste lineal es un modelo matemático usado para aproximar la relación de dependencia entre una variable dependiente Y , las variables independientes X_i y un término aleatorio ε .

Variable dependiente o de resultado

La variable dependiente es aquella cuyo valor depende del valor numérico que adopta la variable independiente en la función. Una magnitud, de este modo, es función de otra cuando el valor de la primera magnitud depende de forma exclusiva del valor que evidencia la segunda magnitud. La primera magnitud es la variable dependiente; la segunda magnitud, la variable independiente.

Una variable dependiente representa una cantidad cuyo valor depende de cómo se modifica la variable independiente.

A menudo y es la variable que se utiliza para representar la variable dependiente en una ecuación.

¿Cuál es la variable dependiente?

La variable dependiente es la cantidad de dinero que obtienes, pues la cantidad de dinero que ganas depende del número de tareas que hagas.

Variable Independiente o predictora

Una variable independiente es una variable que representa una cantidad que se modifica en un experimento.

A menudo x es la variable que se utiliza para representar la variable independiente en una ecuación.

La variable independiente es la cantidad de tareas que haces, pues esta es la variable sobre la que tienes control.

Variable ordinal y nominal

Una variable cualitativa es ordinal, si sus valores se pueden ordenar

Una variable cualitativa es nominal, si sus valores no se pueden ordenar