



Pasar de la estadística descriptiva a la inferencial. Empezar a entender cuales son las herramientas estadísticas de trabajo de un analista de datos.

## Estadística inferencial para novatos

Los desarrollos de IA tienen un denominador común: necesitan tener datos para "aprender". PERO... sobre todo necesitan de "analistas de datos" capaces de entender e interpretar los datos con los que se pretende trabajar para que posteriormente se puedan modelar buenos algoritmos para que las computadoras "aprendan" correctamente.

Para explotar correctamente los datos es necesario sumergirse en la estadística descriptiva (tarea #46), "disciplina" que se no se ocupa de sacar conclusiones, hacer interpretaciones ni proponer hipótesis. Abarca los métodos idóneos para la organización de la información que contiene el conjunto representativo de datos y procura poner en evidencia las características principales.

A diferencia de la estadística descriptiva, la **estadística inferencial** se encarga de hacer deducciones, inferir propiedades, llegar a conclusiones y describir tendencias a partir de una muestra representativa. Trata de interpretar los datos y hacer proyecciones. Es muy útil en el análisis de poblaciones y tendencias, para hacerse una idea posible de las acciones y reacciones de la misma de cara a condiciones específicas. Esto no significa que se puedan predecir fielmente, ni que estemos en presencia de una ciencia exacta, pero sí de una aproximación posible al resultado final.

Esta rama de la estadística emplea habitualmente diferentes técnicas para poder hacer deducciones, como por ejemplo:

- 1.- Pruebas de estimación puntual (o de intervalos de confianza):** El objetivo es aproximar el valor del parámetro desconocido (tiempo medio de ejecución de un algoritmo, altura media de una población, diferencia del resultado entre dos tratamientos médicos, ?). Una estimación es puntual cuando se usa un solo valor extraído de la muestra para estimar el parámetro desconocido de la población. A veces es conveniente obtener unos límites entre los cuales se encuentre el parámetro con un cierto nivel de confianza, en este caso es una estimación por intervalos.
- 2.- Pruebas paramétricas y no paramétricas:** Las paramétricas asumen distribuciones estadísticas subyacentes a los datos (como de media, de diferencia de medias, proporciones, etc.). Las no paramétricas no deben ajustarse a ninguna distribución (como la prueba del chi-cuadrado, etc.).
- 3.- Otros:** Análisis de correlación y de regresión, las series cronológicas, el análisis de varianza, ...

En la estadística inferencial es vital entender el concepto de muestra y población y qué son las distribuciones de variables, los intervalos de confianza, la distribución de densidad de probabilidad, la distribución normal, la esencia de los modelos predictivos, ...

Ser metódico en el estudio de los datos nos ayudará a conseguir avances significativos en relativamente poco tiempo. La metodología de trabajo común en este tipo de estudios estadísticos suele ser el siguiente (Fuente: Wikipedia):

**1.- Planteamiento del problema:** un problema de inferencia estadística suele iniciarse con una fijación de objetivos o algunas preguntas del tipo:

¿Cuál será la media de esta población respecto a tal característica?

¿Se parecen estas dos poblaciones?

¿Hay alguna relación entre...?

En el planteamiento se definen con precisión la población, la característica a estudiar, las variables, etc.

**2.- Elaboración de un modelo:** en caso de establecer un modelo teórico, se replantea el procedimiento y se llega a una conclusión lógica. Los posibles modelos son distribuciones de probabilidad.

**3.- Extracción de la muestra:** se usa alguna técnica de muestreo o un diseño experimental para obtener información de una pequeña parte de la población.

**4.- Tratamiento de los datos:** en esta fase se eliminan posibles errores, se depura la muestra, se tabulan los datos y se calculan los valores que serán necesarios en pasos posteriores, como la media muestral, la varianza muestral.

Los métodos de esta etapa están definidos por la estadística descriptiva.

**5.- Estimación de los parámetros:** con determinadas técnicas se realiza una predicción sobre cuáles podrían ser los parámetros de la población.

**6.- Contraste de hipótesis:** los contrastes de hipótesis son técnicas que permiten simplificar el modelo matemático bajo análisis. Frecuentemente el contraste de hipótesis recurre al uso de estadísticos muestrales.

**7.- Conclusiones:** se critica el modelo y se hace un balance. Las conclusiones obtenidas en este punto pueden servir para tomar decisiones o hacer predicciones.

El estudio puede comenzar de nuevo a partir de este momento, en un proceso cíclico que permite conocer cada vez mejor la población y características de estudio.

Se recomienda leer estos links e incluso hacer algunos de los ejercicios propuestos:

1.- [https://calculo.cc/temas/temas\\_estadistica/estadistica/teoria/estadistica.html](https://calculo.cc/temas/temas_estadistica/estadistica/teoria/estadistica.html)

2.- [http://formacion.intef.es/pluginfile.php/246706/mod\\_resource/content/1/index.html](http://formacion.intef.es/pluginfile.php/246706/mod_resource/content/1/index.html)

3.- [https://www.superprof.es/apuntes/escolar/matematicas/estadistica/inferencia/problemas-de-inferencia-estadistica.html#tema\\_muestreo](https://www.superprof.es/apuntes/escolar/matematicas/estadistica/inferencia/problemas-de-inferencia-estadistica.html#tema_muestreo)

4.- [https://calculo.cc/temas/temas\\_estadistica/muestreo/problemas/prob\\_tama.html](https://calculo.cc/temas/temas_estadistica/muestreo/problemas/prob_tama.html)

5.- <https://www.superprof.es/diccionario/matematicas/estadistica/estadistica-inferencial.html>

No hace falta subir ninguna documentación al repositorio.

**#HASHTAGS** *(etiquetas de ayuda para búsqueda de información relevante)*

#estadística-inferencial #machine-learning #predicción #muestra-población #modelos-predictivos

## LINKS DE INTERÉS

<https://www.youtube.com/watch?v=nbJU4iS-LEg>

[https://www.youtube.com/watch?v=N\\_Bnk9Wq7E4](https://www.youtube.com/watch?v=N_Bnk9Wq7E4)

[https://es.wikipedia.org/wiki/Estad%C3%ADstica\\_inferencial](https://es.wikipedia.org/wiki/Estad%C3%ADstica_inferencial)

<https://www.superprof.es/diccionario/matematicas/estadistica/estadistica-inferencial.html>

## DICCIONARIO

estadística-inferencial | variable | muestra | predicción | estimación | hipótesis | varianza | correlación | p-valor | distribución-normal

## PUNTUACIÓN

Programación: 2

Redes: 1

Seguridad: 1

