

TÉCNICAS ANÁLISIS ESTADÍSTICO

Cuando hablamos de técnicas de análisis estadístico, nos referimos a diferentes herramientas que se utilizan para observar datos y comprender como son, y en los cuales podemos comparar y saber si existe relación entre variables y hacer predicciones.

Existen diferentes grupos de técnicas:

Grupo 1: para describir los datos

Aquí se incluyen las técnicas que nos ayudan a observar los datos y comprenderlos.

1. **Medidas básicas.** Sirven para conocer los datos que hemos recabado.
 - **Media:** el “promedio” de todos los valores.
 - **Mediana:** el valor que queda justo en medio cuando se ordenan los datos.
 - **Moda:** el valor que más se repite.
 - **Desviación estándar:** nos dice si los datos están muy dispersos o parecidos entre sí.
 - **Tablas de frecuencia:** cuántas personas respondieron cada opción.
2. **Gráficas.** Son dibujos para ver los datos.
 - Gráfica de barras: para respuestas tipo “A, B, C...”.
 - Histograma: para datos numéricos (edades, ingresos).
 - Gráfico de dispersión: puntos que muestran cómo se relacionan dos cosas (por ejemplo: edad e ingresos).

Grupo dos: para comparar grupos

Nos sirven para comprender si los datos son diferentes o al azar.

1. **Prueba t.** Se usa cuando se quiere comparar el promedio de dos grupos, por ejemplo: ¿el promedio de satisfacción es diferente entre hombres y mujeres?
2. **ANOVA.** Es como la prueba t, pero para tres o más grupos. Por ejemplo: ¿el promedio de satisfacción cambia entre sucursal A, B y C?

3. **Chi-cuadrado.** Se usa cuando las variables son categorías (no números). Por ejemplo: ¿el tipo de servicio (online/presencial) está relacionado con “satisfecho / no satisfecho”?

Estas pruebas dan un número llamado p-valor. Si p es pequeño (normalmente menor a 0.05), se dice: “La diferencia probablemente no es por azar”.

Grupo tres: para ver relaciones

Sirven para responder: “¿Cuándo cambia A, también cambia B?” Es decir, podemos determinar si existe relación en cambios que hemos detectado.

1. Correlación

Mide qué tan relacionadas están dos variables numéricas.

Da un número r entre -1 y +1:

- r cercano a +1: cuando A sube, B también sube.
- r cercano a -1: cuando A sube, B baja.
- r cercano a 0: no hay relación clara.

Por ejemplo:

- Años de estudio - salario.
- Tiempo de espera - satisfacción (posiblemente relación negativa).

Es importante identificar que cuando se dice que existe relación, no significa que sea la causa de... es decir, que dos cosas se muevan juntas no quiere decir que una cause la otra.

Grupo cuatro: para predecir

Este tipo de técnicas nos ayudan a determinar si al contar con una cantidad de datos podemos estimar o predecir eventos futuros. “Si conozco ciertos datos, ¿puedo estimar otro?”

1. **Regresión lineal.** Permite predecir un número a partir de otro u otros. Ejemplo: predecir la **satisfacción** usando:
- tiempo de espera
 - trato del personal
 - limpieza

El modelo devuelve:

- Una fórmula que dice cuánto pesa cada factor.
- Un número R^2 que indica qué tanto explican esos factores (0% a 100%).

2. **Regresión logística.** Se usa cuando lo que queremos predecir es sí/no. Por ejemplo: ¿el cliente se dará de baja? (1=Sí, 0=No).

Referencia:

Volk-Jesussek, Hannah. (2024). *Estadística descriptiva y estadística inferencial*. Numiqo. Recuperado de:

<https://numiqo.com/tutorial/descriptive-inferential-statistics>

Redacción de Geeksforgeeks. (2025). *¿Qué es el análisis estadístico en la ciencia de datos?*

Geeksforgeeks. Recuperado de: [https://www.geeksforgeeks.org/data-science/what-is-statistical-analysis-in-](https://www.geeksforgeeks.org/data-science/what-is-statistical-analysis-in-data-science/)

[data-science/](https://www.geeksforgeeks.org/data-science/what-is-statistical-analysis-in-data-science/)