1. Presentación
2. Título: análisis exploratorio de datos climatológicos
3. Partes
4. Información general:

Lugar: Como se mencionó la presentación anterior, los datos son de una estación en Chachapoyas, la cual nos da información cada 15 minutos, a partir del procesamiento de estos datos obtuvimos datos mensuales tanto de precipitación como de temperatura.

Periodo de estudio: 2022-2023

Ciclo hidrológico:

Periodo de avenida (lluvia): enero-abril

Periodo de estiaje (no llueve): mayo-septiembre

Periodo de recarga (donde las lluvias empiezan gradualmente): octubre-diciembre

Teniendo en cuenta las mencionadas consideraciones, se seleccionaron en conjunto los 12 meses del año 2022 para compararlos con los 12 meses del año 2023 utilizando la prueba t de Student. Este análisis se llevó a cabo tanto para la temperatura como para la precipitación.

1. Antes de empezar a mostrar los resultados aclararemos unos conceptos

**T Student:**

La distribución t student es una distribución de probabilidad que estima el valor de la media de una muestra pequeña extraída de una población que sigue una distribución normal en la cual su desviación típica es desconocida.

El gráfico que vemos en la diapositiva es de la **función de densidad de probabilidad** de la distribución **t-Student** para diferentes grados de libertad (df = 1, 5, 10, 30). Cada curva representa cómo cambia la distribución t a medida que aumentan los grados de libertad. A medida que los grados de libertad aumentan, la distribución t se aproxima a una distribución normal estándar.

**Serie estacionaria**

Una serie estacionaria es una serie temporal cuyas propiedades estadísticas, como la media, varianza y autocorrelación, permanecen constantes a lo largo del tiempo. En otras palabras, el comportamiento de la serie no cambia dependiendo del momento en el que se analice, lo que la hace predecible en ciertos aspectos.

Para representar una serie estacionaria de manera gráfica, lo mejor es generar una serie de tiempo que tenga una media constante y una varianza constante a lo largo del tiempo, ya que esas son características clave de una serie estacionaria y es lo que vemos en el gráfico de la diapositiva, pues tiene una media constante en torno a 0 y una varianza constante a lo largo del tiempo.

También este gráfico no presenta correlación, es decir no hay una relación diferenciadora entre los valores de la serie a lo largo del tiempo, o sea, no se puede predecir un valor futuro a partir de los valores anteriores. Este gráfico podría ser ejemplo de fluctuaciones de precios en un mercado sin tendencia, o errores de predicción.

**Serie no estacionaria**

Una serie no estacionaria es una serie temporal cuyas propiedades estadísticas, como la media, varianza y autocorrelación, cambian a lo largo del tiempo. Esto significa que los datos no siguen un comportamiento constante ni predecible y pueden mostrar tendencias, estacionalidad o cambios en la variabilidad.

Las características clave de una serie no estacionaria son

Tendencia: La serie no tiene una media constante, ya que esta varía con el tiempo (en este caso del gráfico, la serie tiene una tendencia ascendente).

Varianza no constante: La dispersión alrededor de la tendencia puede cambiar con el tiempo.

Autocorrelación: Los valores de la serie están relacionados con el tiempo; los valores futuros dependen de los valores anteriores debido a la tendencia.

El gráfico que vemos en la diapositiva representa una serie no estacionaria con tendencia. A diferencia de una serie estacionaria, una serie no estacionaria muestra una tendencia o un patrón a lo largo del tiempo. En este caso, la serie tiene una tendencia lineal ascendente, y puede representar el crecimiento o la caída de valores financieros a lo largo del tiempo o un crecimiento poblacional

**Diferencia estadística significativa**

Una diferencia estadística significativa implica que la diferencia observada entre dos conjuntos de datos no se debe al azar, sino a una causa específica, cuando se dice que una diferencia es significativa, se infiere que es muy probable que esa diferencia sea real y no solo el resultado de fluctuaciones aleatorias en los datos. El concepto “significación estadística” se relaciona con la necesidad de “probar hipótesis” comparando la hipótesis nula en la que niega que hay relación de variables, con la hipótesis alternativa que afirma que existe asociación entre variables

En el gráfico se representan dos grupos de datos aleatorios, representando un grupo de control y un grupo experimental, y luego ilustraré la diferencia entre sus medias.

El grupo de control representado con barra azul tiene una media alrededor de 50 y el grupo experimental representado con barra naranja tiene una media alrededor de 55.

El valor numérico sobre las barras de error indica el tamaño del error estándar en cada caso. Si las barras de error no se superponen, como en este caso, eso sugiere una posible diferencia significativa entre los grupos.

1. José
2. Conclusiones

No se hallaron pruebas estadísticas suficientes para afirmar que hubo una diferencia en la precipitación mensual promedio entre 2022 y 2023. En general, parece que la cantidad de lluvia fue similar en ambos años. El análisis con la prueba t respalda la idea de que no hubo cambios importantes en las lluvias mensuales durante ese tiempo.

Con un valor p de 0.09971, no hay suficiente evidencia para decir que la diferencia en las temperaturas promedio de 2022 y 2023 sea significativa al nivel del 5%. Es posible que la diferencia observada se deba al azar, por lo que no se puede concluir que hubo un cambio importante en las temperaturas entre esos dos años con los datos disponibles.

La prueba t es útil cuando quieres ver si las diferencias entre los grupos son reales o simplemente producto del azar. Pero, hay que tener en cuenta el tamaño de la muestra y si los datos siguen ciertas reglas, para que los resultados sean confiables.

Dado que no se encontraron diferencias significativas en la precipitación ni en la temperatura, se sugiere seguir monitoreando estas variables climáticas en los próximos años, ya que el clima es un sistema complejo y cambiante, por lo que realizar un análisis a más largo plazo, cubriendo varios años, ayudaría a detectar posibles tendencias. Además, sería útil complementar estos estudios con otros factores climáticos o usar pruebas más precisas para captar cambios más sutiles.