

Nombre alumno: Paula Inés Pelayo Morales

Correo alumno: paula.pelayo@iteso.mx

Instituto tecnológico de Estudios Superiores de Occidente

Repaso para primer examen de la clase Series de Tiempo

Docente: Daniel Nuño, danielnuno@iteso.mx

Otoño 2024

Fecha de entrega: miércoles 18 de septiembre, 2024

El repaso está diseñado en base a los conocimientos teóricos adquiridos en la clase y tomando como referencia el libro Forecasting: Principles and Practice (3rd ed), capítulos 1 a 7.

La entrega es en digital en un archivo de Word. Las preguntas y respuestas deben diferenciarse claramente.

En tus palabras, o subrayando si se presentan opciones, responde las siguientes preguntas:

1. Componentes de una serie de tiempo descompuesta:

- **Tendencia-ciclo:** Es la **dirección** general **que sigue** la serie en el tiempo, ya sea de **incremento o decrecimiento**, con fluctuaciones a largo plazo.
- **Estacionalidad:** Es un **patrón que se repite en intervalos regulares** (periodos) a lo largo del tiempo, como cambios estacionales en ventas o producción.
- **Residuos:** Son las **variaciones o errores que NO pueden explicarse** por la tendencia o estacionalidad. Se consideran "ruido" o variabilidad aleatoria.
- **Media:** El valor **promedio** de la serie.
- **Varianza cero:** Significa que **no hay variabilidad en los datos**; todos los valores son iguales.
- **Autocorrelación:** Es la **correlación de una serie consigo misma** en distintos retrasos temporales (lags).
- **Normalidad:** Implica que los **residuos** de una serie de tiempo están **distribuidos normalmente** (curva de Gauss).

2. Ventajas de usar la descomposición STL (Seasonal and Trend decomposition using Loess):

- STL puede **manejar cualquier tipo de estacionalidad**, más allá de datos mensuales o trimestrales.
- El **componente estacional puede cambiar con el tiempo**.
- Se puede **ajustar la suavidad de la tendencia**.
- Es **robusta ante valores atípicos**.
- Puede **manejar fechas faltantes**.
- Todas las anteriores.

3. ¿Qué mide la autocorrelación?

Mide la relación lineal entre valores rezagados de una serie de tiempo.

4. ¿Para qué se utiliza la función de autocorrelación?

Se utiliza para identificar la dependencia o relación entre los valores de la serie de tiempo en diferentes momentos (lags). Es útil para detectar patrones como estacionalidad o determinar si existe correlación significativa entre los retrasos temporales de la serie.

5. En las pruebas de hipótesis, ¿qué es el valor-p?

Indica la probabilidad de obtener un resultado al menos tan extremo como el observado, bajo la suposición de que la hipótesis nula es verdadera. Se utiliza para decidir si rechazar o no la hipótesis nula. Un valor-p bajo generalmente sugiere que es poco probable que los datos se hayan generado bajo la hipótesis nula, lo que lleva a rechazarla.

6. ¿Para qué sirve la prueba Ljung-Box? ¿Qué significa aceptar la hipótesis nula y cuándo la aceptas?

La prueba Ljung-Box se utiliza para verificar si los residuos de un modelo de serie de tiempo están correlacionados en diferentes rezagos (es decir, para comprobar la independencia de los residuos).

- **Hipótesis nula:** Los residuos no están autocorrelacionados.

- Aceptamos la hipótesis nula cuando no hay evidencia de autocorrelación en los residuos, lo que significa que el modelo ajusta bien la serie.

7. ¿Para qué sirve la prueba Shapiro-Wilk? ¿Qué significa aceptar la hipótesis nula y cuándo la aceptas?

La prueba Shapiro-Wilk evalúa si una muestra de datos sigue una distribución normal.

- **Hipótesis nula:** Los datos siguen una distribución normal.

- Aceptamos la hipótesis nula cuando no hay evidencia suficiente para rechazar que los datos son normales, lo que indica que los residuos o la variable evaluada podrían seguir una distribución normal.

8. ¿Para qué sirve la prueba Breusch-Pagan o White? ¿Qué significa aceptar la hipótesis nula y cuándo la aceptas?

Estas pruebas se utilizan para detectar heterocedasticidad en un modelo de regresión (cuando la varianza de los errores no es constante a lo largo del tiempo o los niveles de predicción).

- **Hipótesis nula:** Los residuos son homocedásticos (tienen varianza constante).

- Aceptamos la hipótesis nula si no hay evidencia de heterocedasticidad, lo que significa que el modelo es adecuado en cuanto a la distribución homogénea de los errores.

9. ¿Qué es heterocedasticidad y homocedasticidad?

Heterocedasticidad: Es cuando los **residuos** de un modelo de regresión tienen **varianza no constante**, lo que puede causar problemas en la estimación de parámetros.

Homocedasticidad: Es cuando los **residuos tienen varianza constante** a lo largo de las observaciones.

10. ¿En qué casos cree usted que sea útil realizar un pronóstico a partir de una descomposición?

Cuando los datos muestran **patrones claros de estacionalidad y tendencia**, la descomposición permite modelar y ajustar esos componentes de forma separada, lo que puede mejorar la precisión del pronóstico.

11. ¿Cuándo trabajar sobre datos desestacionalizados y cuándo con la serie original?

Desestacionalizados: Cuando **se desea** eliminar el efecto estacional para **entender mejor la tendencia subyacente** o cuando la estacionalidad interfiere con el análisis. En este caso, se obtiene una imagen más clara de la tendencia o de los efectos a largo plazo.

Serie original: Cuando la **estacionalidad es importante para el análisis** o el pronóstico, por ejemplo, para datos de **ventas mensuales** donde los patrones estacionales son clave.

12. ¿Cuáles son otras formas de modificar tus datos para mejorar tu ajuste y pronóstico?

- **Transformación matemática:** Por ejemplo, una transformación logarítmica para **estabilizar la varianza**.

- **Trabajar sobre las diferencias y retornos:** Esto **elimina** la tendencia para modelar solo los **cambios** entre periodos.

- **Modificar observaciones atípicas:** **Ajustar o eliminar valores atípicos** que distorsionan el modelo.

- **Ajustes de incremento poblacional o inflación:** Si los datos reflejan crecimiento poblacional o económico.

- **Recortar tus datos:** En ocasiones, **datos antiguos** no son útiles para modelar el presente o futuro.

13. ¿Por qué harías una transformación matemática?

Para **estabilizar la varianza**, hacer que los datos sigan una **distribución más normal** o mejorar la **linealidad entre variables** en un modelo de regresión.

14. ¿Cómo identificar los datos atípicos?

Mediante gráficos de dispersión, boxplots, o usando métodos estadísticos como puntajes z (z-scores), que miden la desviación de un valor respecto a la media.

15. Describe intuitivamente los modelos base y cuándo los usarías:

- **Media:** Promedio de los valores pasados. Útil cuando no hay tendencia o estacionalidad clara.
- **Ingenuo:** Usa el último valor observado como pronóstico. Se usa para series altamente volátiles o sin tendencia.
- **Ingenuo estacional:** Usa el valor del mismo periodo anterior para predecir el futuro. Se usa cuando hay una estacionalidad clara.
- **Deriva:** Usa la tendencia de los datos pasados para proyectar al futuro. Se usa cuando hay una tendencia clara.

16. Ventajas/desventajas de las métricas de error:

- **RMSE:** Penaliza más los errores grandes, es fácil de interpretar pero sensible a valores atípicos.
- **MAE:** No penaliza errores grandes tanto como RMSE, es robusto ante outliers pero menos interpretativo que RMSE.
- **MAPE:** Expresa el error en términos porcentuales, fácil de entender pero problemático cuando hay valores cercanos a cero.
- **MASE:** Normaliza el error en función de un modelo ingenuo, lo que permite comparar entre series de diferentes escalas.

17. Un buen método de pronóstico producirá errores con las siguientes propiedades: No están correlacionados, media es cero, varianza constante a través del tiempo.

18. ¿Cuál es la afectación en un modelo o sus pronósticos cuando los errores no se distribuyen de manera normal?

Si los errores no son normales, puede afectar la estimación de los intervalos de confianza y la validez de ciertas pruebas de hipótesis, lo que lleva a resultados inexactos.

19. ¿Cuál es la afectación en un modelo o sus pronósticos cuando los errores son heterocedásticos?

Si hay heterocedasticidad, las predicciones pueden ser ineficientes y los errores estándar de los coeficientes pueden estar mal calculados, lo que afecta la interpretación del modelo.

20. Si los residuos parecen estar correlacionados, ¿qué alternativas tiene para intentar mejorar el modelo?

- Utilizar un modelo ARIMA u otro modelo que tenga en cuenta la autocorrelación.
- Incluir variables rezagadas o realizar una transformación sobre los datos para eliminar la autocorrelación.

21. ¿Cuál es el proceso iterativo de un flujo de pronóstico?

Generalmente implica los siguientes pasos: recolectar datos, descomponer o transformar los datos, ajustar modelos, validar los resultados, pronosticar y ajustar según el feedback.

22. ¿Cuál es el propósito de separar tus datos entre entrenamiento y prueba?

Evaluar la capacidad de generalización de un modelo, evitando que se ajuste demasiado a los datos históricos (overfitting).

23. ¿Cuál es la proporción recomendada para dividir los datos entre entrenamiento y prueba?

Generalmente, entre 70% y 80% para entrenamiento y el resto para prueba (20%-30%).

24. ¿Para qué sirve el proceso de cross-validation (o back-testing)?

Sirve para evaluar la estabilidad y capacidad predictiva de un modelo mediante la división repetida de los datos en subconjuntos de entrenamiento y prueba.

25. Si las métricas de error del proceso de cross-validation apuntan a distintos modelos, ¿qué harías para decidir cuál modelo utilizar?

Revisar el contexto y los requerimientos del pronóstico (por ejemplo, precisión frente a simplicidad). Podrías optar por el modelo con menor error absoluto o relativo, o por el modelo que mejor se ajuste a las necesidades del negocio o fenómeno a predecir.

26. ¿Qué es una regresión lineal?

Es un método para modelar la relación entre una variable dependiente y una o más variables independientes.

27. ¿Qué es el método de mínimos cuadrados?

Es un método que minimiza la suma de los errores cuadráticos (diferencias al cuadrado entre los valores predichos y los observados) para encontrar la mejor línea de ajuste en un modelo de regresión.

28. ¿Cuál es la función objetivo y pasos para obtener los parámetros β de una regresión lineal simple?

La función objetivo es minimizar la suma de los errores cuadráticos:

$$\sum (y = \beta_0 + \beta_1 x)$$

Los parámetros se obtienen resolviendo las ecuaciones normales para minimizar los errores.

29. ¿Cuáles son los supuestos de la regresión lineal simple?

Relación lineal entre las variables, independencia de los errores, homocedasticidad de los errores, y normalidad de los errores.

30. ¿Qué es R^2 ?

Es el coeficiente de determinación que mide el porcentaje de la variabilidad de la variable dependiente que se puede explicar con la variable independiente.

31. Diferencia entre regresión lineal simple y múltiple:

La simple tiene una variable independiente, mientras que la múltiple tiene varias.

32. Multicolinealidad y por qué es mala para la regresión lineal múltiple:

Es cuando dos o más variables independientes están altamente correlacionadas, lo que dificulta la estimación precisa de los coeficientes.

33. Métricas AIC, BIC y R^2 ajustada:

AIC: Evalúa el ajuste del modelo penalizando la complejidad. Un AIC más bajo indica un mejor equilibrio entre simplicidad y ajuste.

BIC: Similar al AIC pero con una penalización más fuerte para modelos más complejos. Un BIC más bajo indica un modelo mejor ajustado y más simple.

R^2 ajustada: Penaliza la inclusión de predictores innecesarios. Es útil para comparar modelos con diferentes números de predictores.

34. Técnicas de selección de predictores:

- **Stepwise:** Método automático que añade o elimina predictores uno por uno, evaluando su contribución al modelo.

- **Lasso:** Regresión regularizada que selecciona automáticamente predictores importantes al forzar coeficientes irrelevantes a cero.

- **Ridge:** Similar a Lasso pero no selecciona predictores, sino que reduce la varianza de los coeficientes para manejar multicolinealidad.

- **Elastic Net:** Combina Lasso y Ridge, ideal cuando hay predictores altamente correlacionados, seleccionando variables pero manteniendo algunas que Lasso eliminaría.