

Módulo: ING1909 - Análisis exploratorio de datos - (A52)

Actividad: Reto de aprendizaje 5

Nombre: Roberto Mora Balderas

Asesor: Rafael Vazquez Perez

Fecha: 11 de septiembre de 2023



Objetivo de la actividad

• Reforzar el conocimiento adquirido del material de estudio correspondiente a poder identificar los 7 abusos más comunes en el uso de la regresión lineal.

Instrucciones

 A partir de lo visto en el Tema 2 crear un mapa conceptual sobre los siete abusos en el uso de la regresión lineal. Finalizar la actividad con una conclusión personal sobre el posible daño que uno de los mencionados puntos que puede llegar a causar en la realización de algún experimento social.

Desarrollo:

1 ESPECIFICACIÓN INCORRECTA DEL MODELO

- Definición: Utilización de un modelo inapropiado para los datos, ya sea por incluir variables irrelevantes, omitir variables cruciales o elegir una forma funcional incorrecta.
- Consecuencia: Predicciones incorrectas y conclusiones erróneas sobre las relaciones entre variables.

2 OBSERVACIONES INUSUALES O ATÍPICAS

- Definición: Observaciones atípicas o puntos influyentes que pueden afectar de manera desproporcionada la línea de regresión.
- Consecuencia: Puede llevar a una interpretación errónea de la relación general si no se identifican y tratan adecuadamente.

3 SIGNO INCORRECTO DEL COEFICIENTE DE REGRESIÓN

- Definición: Cuando el signo del coeficiente de regresión no coincide con lo que se espera teóricamente.
- Consecuencia: Puede indicar variables omitidas, multicolinealidad o problemas en la especificación del modelo.

4 MODELOS SIMPLIFICADO

- Definición: Presentar modelos simplificados que no capturan la complejidad real de los datos o las relaciones.
- Consecuencia: Puede llevar a decisiones basadas en una comprensión superficial o incorrecta de los datos

5 INTERPRETACIÓN DEL COEFICIENTE

- Definición: Confundir el coeficiente de determinación con una medida de "bondad de ajuste" o la calidad general del modelo.
- Consecuencia: Confiar demasiado en un modelo que puede no ser prácticamente significativo o útil.

6 ESTADÍSTICAS SIGNIFICATIVAS

- Definición: Enfatizar en estadísticas avanzadas o complicadas en lugar de aquellas que tienen un significado práctico claro.
- Consecuencia: Puede llevar a decisiones erróneas o confusiones al presentar los resultados a personas que no son expertas en estadísticas.

7 MODELOS ESTADÍSTICOS

- Definición: La recomendación de utilizar modelos estadísticos más establecidos y conocidos.
- Consecuencia: Al no seguir esta recomendación, se corre el riesgo de utilizar modelos que no han sido probados o que podrían no ser apropiados para el tipo de datos en cuestión.



Conclusión:

El abuso de "Model Misspecification" me parece particularmente preocupante, especialmente en experimentos sociales. Si un investigador especifica erróneamente un modelo, las conclusiones derivadas pueden ser totalmente erróneas. Esto no solo puede llevar a la difusión de información incorrecta, sino que en el ámbito social, donde las políticas y las intervenciones se basan a menudo en resultados de investigación, puede tener repercusiones reales y dañinas en la vida de las personas. Por ejemplo, si se omite una variable clave en un estudio sobre los determinantes de la educación, podríamos malinterpretar la importancia de otros factores y, como resultado, implementar políticas educativas ineficaces o incluso perjudiciales. Es crucial que los investigadores sean meticulosos y reflexivos al especificar sus modelos para garantizar la validez y relevancia de sus conclusiones.

El análisis de regresión es una herramienta poderosa en investigación, especialmente en las ciencias sociales. Sin embargo, su potencial solo puede ser aprovechado completamente cuando se utiliza con precisión y responsabilidad. Los siete abusos mencionados reflejan las trampas comunes en las que los investigadores pueden caer, llevando a interpretaciones erróneas o a decisiones mal informadas.

Es fundamental que, al emplear técnicas estadísticas, los investigadores se acerquen a los datos con un sentido crítico, siempre cuestionando la validez de sus modelos y siendo conscientes de las limitaciones de sus análisis. En última instancia, los resultados de estos estudios tienen el poder de influir en políticas, opiniones y decisiones de gran alcance. Por lo tanto, es imperativo que se manejen con el debido cuidado y rigor, garantizando que la información que se difunde sea tanto precisa como relevante.

Referencias:

- Anáhuac Online. (2019). Regresión lineal en la práctica. [Contenido creado para Anáhuac Online]. México: Anáhuac Online
- Tarafder, S. (2014). How to Avoid Common Mistakes in Linear Regression. https://acortar.link/1SOcGN
- Anáhuac Online. (2019). Interpretando la regresión lineal. [Contenido creado para Anáhuac Online]. México: Anáhuac Online



- Duflo, E. (2010, 3 mayo). Esther Duflo: Experimentos sociales para luchar contra la pobreza [Vídeo]. TED Talks. https://acortar.link/4WODuM
- Anáhuac Online. (2019). Entendiendo el análisis de regresión. [Contenido creado para Anáhuac Online]. México: Anáhuac Online
- Dave Your Tutor. (2013, 12 abril). The Easiest Introduction to Regression Analysis! Statistics Help [Vídeo]. YouTube. https://www.youtube.com/watch?v=k_OB1tWX9PM