

ANÁLISIS ESTADÍSTICO DE UN ARTÍCULO CIENTÍFICO

Jesús María Mora Mur, Noa Fernández Garrido

06/05/2023

Afiliación: UNIR.

Resumen

Se trabaja en el presente artículo el problema de la fuerza de una red de polímeros dependiendo del número de enlaces de dicha red. Partiendo de los estudios de Tao et al. (2023), se realiza un análisis estadístico para comprobar su corrección y se utilizan procedimientos de estadística inferencial para dilucidar un modelo matemático que permita, con relativa precisión, conocer los límites de la red de polímeros en lo que a las características de los enlaces respecta. Se detallan, asimismo los procedimientos seguidos y la obtención y el tratamiento de los datos experimentales.

Contenido

Introducción y estado del tema.	3
Enunciado del problema.	3
Descripción del modelo.	3
Algoritmos para el cálculo.	3
Metodología y resultados obtenidos.	5
Base de datos.	5
Cálculos estadísticos.	6
Representación y ajuste de los datos obtenidos.	7
Conclusiones y líneas futuras.	9
Referencias.....	9

Introducción y estado del tema.

Enunciado del problema.

El problema que nos ocupa aborda el tema de la fuerza del enlace covalente de una red de polímeros. Teóricamente, los enlaces de la red deberían ser mucho más fuertes de lo que se observa tras el análisis experimental. El modelo muestra que la fuerza de la red disminuye significativamente incluso con una pequeña variación en el número de enlaces por cadena, y la fuerza de la red se relaciona con la variación del número de enlaces según una ley de potencia. (Tao et al., 2023)

Descripción del modelo.

Para realizar el análisis estadístico de los datos recolectados, se emplearán los siguientes métodos de cálculo estadístico:

- **Medidas de tendencia central**: media aritmética (para resumir el conjunto de mediciones de cada variable) y mediana (para hallar el valor central del conjunto de datos y eliminar posibles influencias de valores atípicos).
- **Medidas de dispersión**: desviación media, varianza (para comprobar cuánto se alejan los valores de la media aritmética del conjunto de datos al que pertenecen), desviación estándar y cuartiles (para comprender la distribución de los datos e identificar valores atípicos).
- **Gráficas**: representación visual de los datos obtenidos (para visibilizar posibles correlaciones entre variables y estudiar su linealidad).

Algoritmos para el cálculo.

Para realizar el análisis, se ha procedido tras la pauta de los siguientes algoritmos:

1. ALGORITMO PARA LA TOMA, MEDICIÓN Y ANÁLISIS DE DATOS.

Para comenzar, es necesario tomar una muestra debidamente aislada de la red de polímeros con la que se va a trabajar y sobre la que se realizarán las mediciones. Una vez elegida la muestra, será transportada a un laboratorio competente que contenga todos los dispositivos que nos permitan observar las distintas variables que nos interesan, para así discutir cuáles nos son útiles a la hora de realizar el análisis estadístico y obtener resultados reveladores acerca del objeto de estudio. En este caso, hemos elegido las siguientes variables: fuerza del enlace, fuerza de rotura del enlace, distancia o longitud del enlace, energía de enlace y número o cantidad de enlaces por medición.

Tras configurar y calibrar el equipo de medición adecuado a cada variable, se dispone la muestra de polímeros y se ajustan las condiciones necesarias para llevar a cabo el proceso de toma de datos. En función de la variable elegida y del número de enlaces seleccionados, se opta por un determinado nivel de presión y temperatura, así como otras condiciones que puedan resultar concluyentes. A continuación, para cada variable se realizan 20 mediciones diferentes, estas en función de la cantidad de enlaces con las que estemos trabajando en cada medición. Para cada cantidad fija de enlaces se evalúa cada variable, y así observar posibles relaciones entre parámetros. Cuando se llevan a cabo todas las mediciones acordadas, procedemos al análisis estadístico de los datos mediante el empleo de herramientas de cálculo; en este caso, medidas de distribución de tendencia central, de dispersión y de correlación. Para finalizar, se interpretan los resultados y discutimos qué conclusiones ha traído la información obtenida sobre las variables elegidas.

2. ALGORITMO PARA EL ANÁLISIS ESTADÍSTICO.

Para comenzar, llevamos a cabo la recolecta de datos según lo explicado en el algoritmo previo, para así obtener los valores de las distintas variables en función de la cantidad de enlaces seleccionada. A continuación, ejecutamos la consola de cálculo del software de cálculo estadístico “R” y generamos un vector (por variable) que recoja los 20 valores que le pertenecen. Primero se realiza la media aritmética para cada variable, esto mediante el uso del comando “*mean()*”. Continuamos con el cálculo de la mediana, para ello empleamos la función “*median()*”. A la hora de hallar la desviación media nos ayudamos de la función “*mean(abs(x - mean(x)))*”. Proseguimos con la varianza, por lo que utilizaremos el comando “*var()*”. Con la desviación estándar empleamos la función “*sd()*” y, por último, para obtener los cuartiles, hacemos uso del comando “*quantile()*”.

Para finalizar el análisis, se representarán gráficamente las distintas variables con sus respectivas mediciones frente al número de enlaces que le corresponde a cada valor. Esto permite enfocar el estudio desde una perspectiva de correlación entre variables, por lo que, a la hora de discutir los resultados obtenidos, contaremos con más información y veracidad. Tras dicha discusión y puesta de conclusiones, daremos por finalizado el análisis estadístico del asunto que nos concierne.

Metodología y resultados obtenidos.

Base de datos.

Fuerza del enlace (nN)	Fuerza de rotura (nN)	Distancia de enlace (pm)	Energía de enlace (eV)	Número de enlaces
66,00	75	60	1,07	2
66,75	75	60	1,08	2
88,11	99	64	1,42	5
124,08	141	73	2,07	8
128,52	153	76	2,09	9
131,04	156	77	2,11	10
139,04	176	82	2,16	12
232,29	261	89	2,23	17
244,00	305	91	2,52	19
256,80	321	93	2,64	21
269,10	345	95	2,69	24
327,99	377	98	3,01	27
330,00	440	107	3,08	32
460,08	568	117	3,44	36
525,99	591	123	3,75	39
501,48	597	132	3,63	43
585,00	750	163	3,95	48
689,62	841	179	4,36	54
781,44	888	191	4,56	56
769,44	916	200	4,99	61

Cálculos estadísticos.

- MEDIA ARITMÉTICA:**

Fuerza de enlace (nN)	Fuerza de rotura (nN)	Distancia de enlace (pm)	Energía de enlace (eV)	Número de enlaces
335.54	403.75	108.50	2.84	26.25

- MEDIANA:**

Fuerza de enlace (nN)	Fuerza de rotura (nN)	Distancia de enlace (pm)	Energía de enlace (eV)	Número de enlaces
262.95	333.00	94.00	2.67	22.50

- DESVIACIÓN MEDIA:**

Fuerza de enlace (nN)	Fuerza de rotura (nN)	Distancia de enlace (pm)	Energía de enlace (eV)	Número de enlaces
196.43	315.79	32.62	1.02	20.76

- VARIANZA:**

Fuerza de enlace (nN)	Fuerza de rotura (nN)	Distancia de enlace (pm)	Energía de enlace (eV)	Número de enlaces
55850.30	79158.30	1887.95	1.26	358.09

- DESVIACIÓN ESTÁNDAR:**

Fuerza de enlace (nN)	Fuerza de rotura (nN)	Distancia de enlace (pm)	Energía de enlace (eV)	Número de enlaces
236.33	281.35	43.45	1.13	18.92

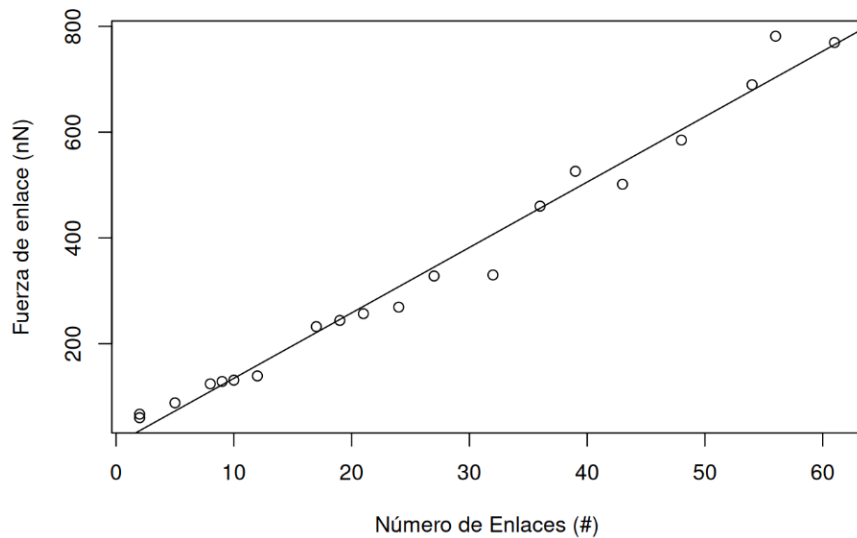
- CUARTILES:**

Fuerza de enlace (nN)	Fuerza de rotura (nN)	Distancia de enlace (pm)	Energía de enlace (eV)	Número de enlaces	<u>CUARTILES</u>
60.00	75.00	60.00	1.07	2.00	0%
130.41	155.25	76.75	2.11	9.75	25%
262.95	333.00	94.00	2.67	22.50	50%
507.61	592.50	125.25	3.66	40.00	75%
781.44	916.00	200.00	4.99	61.00	100%

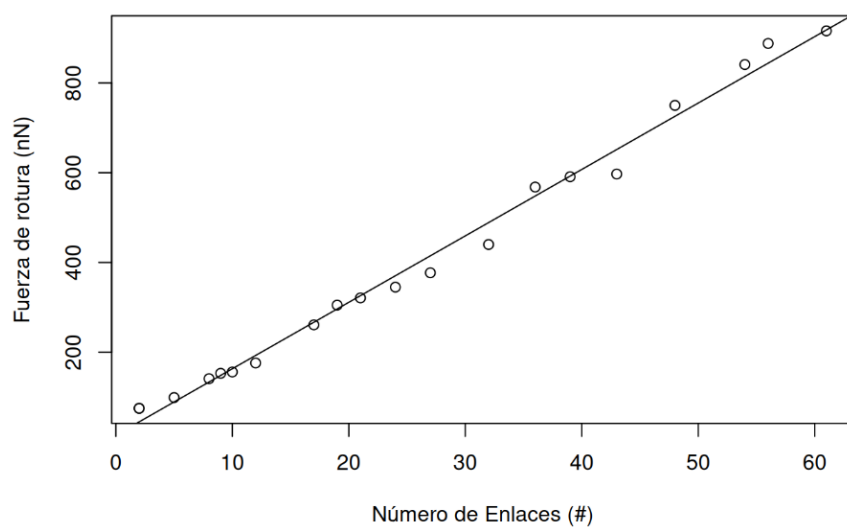
Representación y ajuste de los datos obtenidos.

Al representar las distintas variables analizadas frente al número de enlaces se observa cómo los datos que obtenemos siguen un modelo prácticamente lineal. Así, ajustamos gráficamente entorno a una recta y se extraen los siguientes resultados:

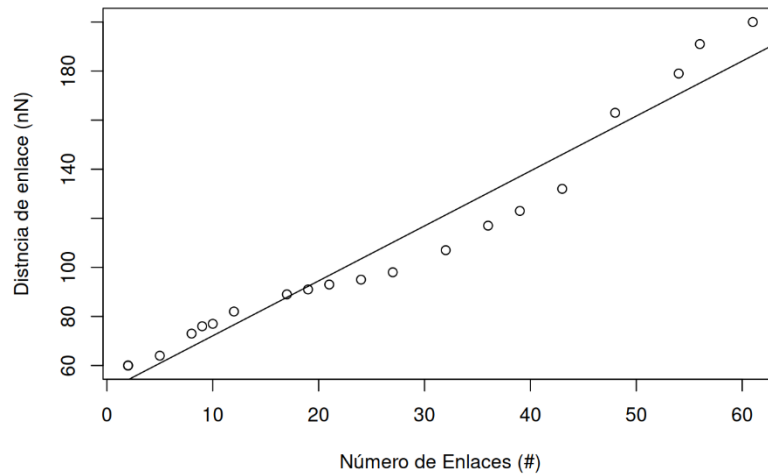
- **RELACIÓN ENTRE FUERZA DE ENLACE Y EL NÚMERO DE ENLACES:**



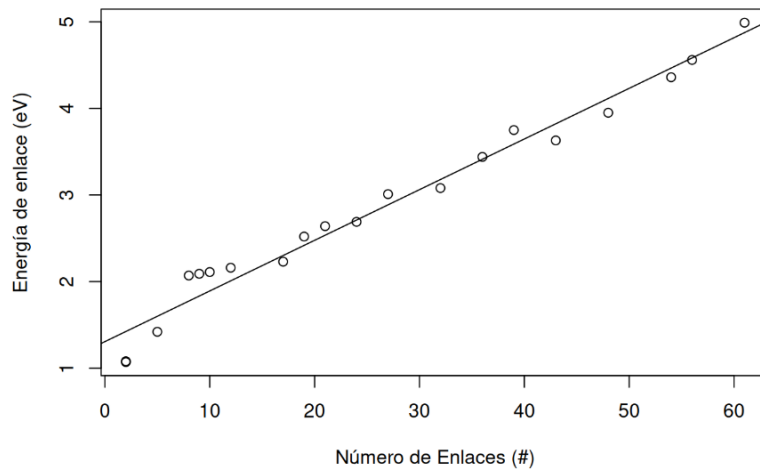
- **RELACIÓN ENTRE FUERZA DE ROTURA Y EL NÚMERO DE ENLACES:**



- **RELACIÓN ENTRE DISTANCIA DE ENLACE Y EL NÚMERO DE ENLACES:**



- **RELACIÓN ENTRE FUERZA DE ENLACE Y EL NÚMERO DE ENLACES:**



Como se observa a la vista de los gráficos anteriores, la recta se ajusta al conjunto de datos experimentales. Puede comprobarse este hecho mediante un parámetro que indica cuán ajustada está la recta con respecto al conjunto experimental. Observamos, en todos los casos, unas correlaciones altas, según lo obtenido al analizar los datos con el lenguaje R:

R^2			
Fuerza de enlace	Fuerza de rotura	Distancia de enlace	Energía de Enlace
0.98	0.99	0.95	0.97

Conclusiones y líneas futuras.

Discusión de resultados.

A la vista del análisis realizado y especificado en los párrafos anteriores vemos como todas las magnitudes se relacionan con el número de enlaces en regresión lineal con una correlación alta. Este hecho implica que podemos ajustar los datos a una recta cuya ecuación podemos deducir. Nos permitiría esto inferir datos no conocidos a partir de una predicción, obteniendo en último término, un modelo de trabajo para el caso que nos ocupa.

Obteniendo estos datos mediante *software*, tenemos las siguientes relaciones:

$$F_{enlace} \approx 12.37 \cdot N + 10.79$$

$$F_{rotura} \approx 14.80 \cdot N + 15.23$$

$$D_{enlace} \approx 2.24 \cdot N + 49.72$$

$$E_{enlace} \approx 0.06 \cdot N + 1.31$$

Siendo N el número de enlaces y las variables dependientes la fuerza de enlace, fuerza de rotura, distancia de enlace y energía de enlace con las unidades explicitadas en la presentación de los datos.

Por otra parte, los parámetros estadísticos descriptivos obtenidos dan cuenta de la distribución de los datos obtenidos. Al tener un número de datos alto y con valores crecientes, no supondría un problema tener desviaciones altas, al no tener los datos distribuidos en torno a una media.

Referencias.

Tao, M., Lavoie, S., Suo, Z., & Cameron, M. K. (2023). *The effect of scatter of polymer chain length on strength*. arXivID: 2304.12815