

Reporte proyecto parcial 2

Equipo 5

29/10/2020

0.1 Índice

Índice 1

Introducción 1

Descripción del problema 1

Resultados 1

Conclusión 3

0.2 Introducción

A lo largo del semestre, hemos estudiado diversa variedad de métodos numéricos con los cuales se pueden resolver distintas problemáticas de la vida laboral y cotidiana, y hoy toca el turno de utilizar un método denominado 'Regresión lineal'.

El análisis por medio de la regresión lineal tiene como objetivo el modelar de forma matemática el comportamiento de variables de respuesta en función de una o más variables independientes.

Un ejemplo de esto sería, suponiendo que el rendimiento de un proceso químico está relacionado con la temperatura de operación. Generando un modelo matemático es posible describir la relación, por lo que el modelo sería bien utilizado en el proceso de predicción y optimización, los cuales son los principales usos y aplicaciones de la regresión lineal.

0.3 Descripción del problema

En una empresa cosmética, se analiza el estado de las sustancias para detectar impurezas en su estado y estudiar qué tan factible es el emplear estos compuestos. Se sospecha que en estos procesos químicos hay una relación lineal entre temperatura y el porcentaje de impureza.

Nuestras variables a utilizar son temperatura en grados Celsius, que identificaremos con la letra 'x' y el porcentaje de impureza que se representará con la letra 'y'.

Se utilizará software como MatLab para la comprobación de nuestros procedimientos y Excel para hacer los cálculos pertinentes y la graficación de nuestro método para una mejor visualización.

0.4 Resultados

A continuación presentamos la metodología utilizada para obtener los resultados que serán destacados en esta sección.

En primera instancia, se realizó una tabla de valores en Excel, donde se dividieron los datos de las distintas pruebas, en el eje X contamos con la temperatura y en el eje Y contamos con el porcentaje de impureza:

PASO 6 Obtener IC para B0 y B1	
X=temperatura	Y=% de impureza
100	88.5
110	72.2
110	72.9
125	59.9
125	59.1
130	58.7
130	47.4
140	35.6
140	35
150	30.4

Seguido de este orden, presentamos las distintas iteraciones para conseguir las estimaciones de error y de nuestra Y_i , ed igual forma adjuntamos como se va a estimar la varianza:

Xi	Yi	Xi^2	Yi^2	XiYi	Yi Estimada	Error (Yi-YEstimada)
100	88.5	10000	7832.25	8850	86.8	1.7
110	72.2	12100	5212.84	7942	74.93	-2.73
110	72.9	12100	5314.41	8019	74.93	-2.03
125	59.9	15625	3588.01	7487.5	57.125	2.775
125	59.1	15625	3492.81	7387.5	57.125	1.975
130	58.7	16900	3445.69	7631	51.19	7.51
130	47.4	16900	2246.76	6162	51.19	-3.79
140	35.6	19600	1267.36	4984	39.32	-3.72
140	35	19600	1225	4900	39.32	-4.32
150	30.4	22500	924.16	4560	27.45	2.95
1260	559.7	160950	34549.29	67923		
126	55.97	16095				

— Estimación de la Varianza de ε (σ^2) —

$$\hat{\sigma}^2 = \frac{S_{YY} - \hat{\beta}_1 S_{XY}}{n - 2}$$

$$\hat{\sigma} = \sqrt{CME}$$

Var CME= 17.203825
CME, DE= 4.14774939

A continuacion, se hizo la evaluacion para determinar la B0 y B1 de nuestra problematica, ademas presentaremos apoyo visual y un breve resumen de como se realiza esta evaluacion:

REGRESIÓN LINEAL SIMPLE

Estimación de β_0 y β_1

Suponer una muestra de tamaño n .
 $(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)$. Sea :

$$S_{XY} = \sum (x_i y_i) - \frac{(\sum x_i)(\sum y_i)}{n}$$

$$S_{XX} = \sum (x_i)^2 - \frac{(\sum x_i)^2}{n}$$

$$\hat{\sigma}^2 = \frac{S_{YY} - \hat{\beta}_1 S_{XY}}{n - 2}$$

$$SCE = S_{YY} - \frac{(S_{xy})^2}{S_{xx}}$$

Se le conoce también como error estándar de estimación

$\hat{\sigma} = \sqrt{CME}$

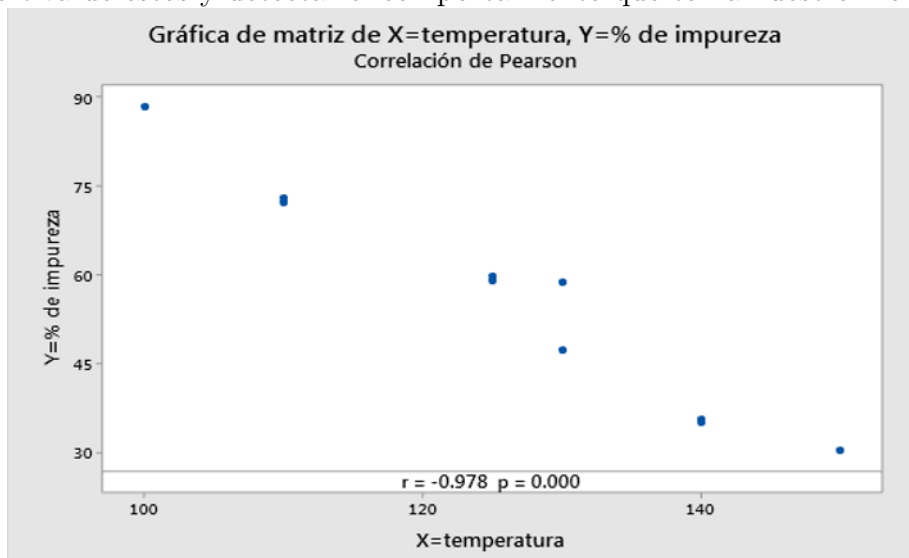
Se le conoce también como suma del cuadrado medio del error

OJO:
Obtenemos S_{yy} para posteriormente saber la SCE y el CME

Recta Estimada Y=	205.5	-1.187	x	SXY=	-2599.2
		B1=	-1.187	SXX	2190
		B0=	205.51	SYY=	3222.881

Continuando con el analisis, realizamos la graficacion de nuestros datos obtenidos para la

visualización acertiva de estos y detectar el comportamiento que toma nuestro modelado y



variable.

Después de un extenso análisis, realizando hasta 10 observaciones o iteraciones, llegamos a predecir el porcentaje de impureza que se va a encontrar en nuestros compuestos químicos, así como también desglosamos la cantidad de residuos y residuos estándares, como se puede observar a continuación:

Observación	Pronóstico Y=% de impureza	Residuos	Residuos estándares
1	86.82808219	1.671917808	0.426934854
2	74.95958904	-2.759589041	-0.704678626
3	74.95958904	-2.059589041	-0.52592917
4	57.15684932	2.743150685	0.700480987
5	57.15684932	1.943150685	0.496195895
6	51.22260274	7.47739726	1.909400982
7	51.22260274	-3.82260274	-0.97612594
8	39.35410959	-3.754109589	-0.958635778
9	39.35410959	-4.354109589	-1.111849596
10	27.48561644	2.914383562	0.744206392

0.5 Conclusion

Como ya o estuvimos mencionando, la regresión lineal simple tiene como objetivo el ajustar una recta a un conjunto de puntos en un plano, los cuales reciben el nombre de 'gráfica de dispersión' que está conformada por los datos de una muestra aleatoria. Respecto al caso, hay ciertas ideas que debemos dejar en claro después de haber estudiado el tema:

A diferencia del análisis de correlación, la RLS tiene como variable independiente X (la controlamos, es decir, no es aleatoria) y la variable Y es dependiente y aleatoria, con distribución normal con media $\mu_{y/x}$ y varianza elevada a la 2da potencia y es por esto que resulta más práctico el emplear esta debido a nuestro caso a estudiar.

1) El coeficiente de determinación indica que la variabilidad de los datos es poca, siguen una tendencia, se ajustan muy bien a la línea recta y esto indica que es una muy buena relación lineal.

2) Con un NS se rechaza la H_0 , es decir la pendiente poblacional de la temperatura respecto al porcentaje de impureza si es diferente a 0, hay relación lineal entre estas dos variables.

3) El coeficiente de determinación indica que la variabilidad de los datos observados es poca, siguen una tendencia, se ajustan muy bien a la línea recta, es una muy buena y esto indica también que es una muy buena relación lineal.

Para concluir, es importante resaltar que este proyecto ha sido de gran utilidad, principalmente para nuestros colegas Bioquímicos y Biotecnólogos ya que pudieron observar la aplicación

en un caso extremadamente posible de un metodo numerico, detectando la importancia y eficacia de contar con estos conocimientos.