Modelamiento estadístico

Trabajo módulo 1

DANIEL EDGARDO GARCIA BONILLA

2022

**Cálculo de costos cubiertos por seguro médico**

**Introducción**

Los seguros son una herramienta de gran valor que permiten cubrir a las personas o bienes materiales acerca de la ocurrencia de incidentes que perjudiquen el bienestar de el objeto o persona asegurado.

Para realizar el cálculo de la prima de los seguros médicos se toman en cuenta diversos factores que inciden en la probabilidad de ocurrencia de estos siniestros.

**Problemática**

Un término importante para cada seguro es la prima, la cual se define como el costo que el asegurado deberá pagar al adquirir una póliza de seguro sin importar de qué tipo es. Antes de poder estar asegurado y recibir las coberturas, el titular debe pagar esta cantidad en su totalidad.

Este pago se cubre a partir de los acuerdos que se establezcan previamente en el contrato que se firma antes de adquirir el producto, en este caso, el seguro. Puede hacerse de forma anual, semestral, trimestral, mensual o en una sola exhibición (pago único). La decisión es del asegurado.

**Objeto de estudio**

Poder identificar cuando una persona será un buen cliente para hacer uso de un seguro médico es una tarea que puede llevarse a cabo a través de realizar un estudio detallado de las variables que pueden incidir en el nivel de salud de las personas aseguradas. De esta manera incluso, se puede tomar una pauta para calcular la prima del seguro a pagar.

**Metodología o Estrategia Empírica**

**Descripción de los datos y variables.**

La data utilizada para este proyecto, si bien no especifica el tipo de seguro en base al cual se tomaron los datos, se detallan 7 variables incluyendo el cargo calculado como prima de seguro para 1064 personas.

Las columnas o variables son las siguientes:

|  |  |
| --- | --- |
| **Columna** | **Descripción** |
| Age | Edad del principal beneficiario |
| Sex | Genero del asegurado: female, male |
| Bmi | Indice de masa muscular. El peso puede ser relativamente alto o bajo dependiendo de la estatura. En este conjunto de datos se estima la relación entre estatura y peso (bmi) la cual es idealmente se encuentra entre los rangos de 18.5 y 24.9 kg/m^2 |
| Children | Numero de hijos cubiertos por el seguro médico / Numero de dependientes |
| Smoker | Fumador o no fumador |
| Región | La residencia de los beneficiarios de los datos pertenece a diferentes áreas de los estados unidos (northeast, southeast, southwest, northwest.) |
| Charges | Costos médicos individuales facturados por el seguro médico |

**Elección del modelo estadístico**

Ya que se busca la relación entre la variable de costos médicos y las demás con las que cuenta el dataset, se decidió hacer uso de un modelo de regresión lineal.

**Discusión de los resultados**

**Análisis gráfico**

EL siguiente gráfico muestra la relación existente entre los costos médicos individuales facturados por el seguro médico y el índice de masa muscular de las personas. Cabe recalcar que el índice de masa muscular para una persona común oscila entre 18.5 y 24.9 kg/m^2. Por lo que valores arriba o debajo pueden representar o personas con deficiencias de salud o niños.

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

El bmi parece no afectar en la cantidad de gastos médicos facturados por individuo, a menos que la persona sea un fumador activo. Esto debido a que el fumar daña casi cada órgano del cuerpo y sistema de órganos del cuerpo y disminuye la salud general de la persona.

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

A medida aumenta la edad, hay un aumento en los gastos médicos en los que incurre el asegurado. Esto es debido a el deterioro natural del cuerpo humano y la necesidad de atención médica que incrementa con el pasar de los años. Sin embargo, existe una diferencia entre en los gastos médicos de las personas que fuman y las que no fuman sin importar las edades.

Por simple intuición podríamos deducir que la media de los gastos de los fumadores y los no fumadores es diferente. Para comprobar este postulado observamos en el siguiente gráfico de cajas que en efecto sí existe una diferencia significativa entre las medias de gastos médicos de las personas que fuman y las que no fuman.

Gráfico, Gráfico de cajas y bigotes

Descripción generada automáticamente

Además de esto se realizó el análisis gráfico de los gastos médicos por sexo, de lo cual se obtuvo el siguiente resultado:

Gráfico, Gráfico de cajas y bigotes

Descripción generada automáticamente

Con esto concluimos que la media de los gastos médicos es similar para ambos sexos.

**Estimación del modelo**

De acuerdo con todo lo planteado se procedió a realizar la elaboración del modelo de regresión lineal, del cual se obtuvieron los siguientes resultados:

Texto

Descripción generada automáticamente con confianza media

Del modelo anterior obtenemos que tanto el índice de masa muscular, como si la persona es un fumador activo, la edad y si tiene hijos o no son variables determinantes para la cantidad de gastos médicos de la muestra.

Ya que el hecho de que las personas fumen o no, es uno de los factores que hace una mayor diferencia en los gastos médicos por persona, se decidió subdividir el dataset en los fumadores activos y otro para las personas que no fuman.

Para los fumadores activos se encontró el siguiente modelo:

Texto

Descripción generada automáticamente

Si la persona fuma, los costos médicos aumentan a medida que el índice de masa muscular aumenta y también la edad.

**Conclusiones**

* Los gastos médicos aumentan considerablemente a si las personas fuman. El daño que este hábito puede causar a la salud de las personas se ve representado en los datos graficados en este estudio.
* La media de los gastos médicos no varía de manera significativa entre el sexo masculino y femenino.
* Una persona que sobrepasa el límite del índice de masa muscular saludable (24.9 kg/m^2), y además fuma, es propensa a problemas de salud, y por lo tanto aumentar sus gastos médicos.