

PRÁCTICA 4: AJUSTES DE DISTRIBUCIONES

Atanasov Angelov, Daniel – daniel.atanasov24@estudiantes.uva.es

Sanz Tomé, Raúl – raul.sanz24@estudiantes.uva.es

Resumen:

En esta práctica de estadística con Statgraphics, hemos ajustado una serie de datos a un modelo de probabilidad, comprobado que el modelo es válido y transformado los datos para hacer que se ajusten al modelo, en caso de que fuera necesario. Por último, hemos empleado los modelos para distintos cálculos.

1. Introducción

En esta práctica hemos aprendido a ajustar un conjunto de datos estadísticos a los modelos de probabilidad para, a continuación, comprobar que el modelo es válido para nuestros datos y, en este caso, utilizarlo para el cálculo de probabilidades. Además, en el caso de que fuera necesario, hemos realizado las correspondientes transformaciones de los datos para que éstos se ajustaran al modelo de probabilidad.

2. Metodología

Hemos seguido los siguientes procesos: tras generar un conjunto de números aleatorios de una distribución binomial con $n = 100$ y $p = 0,2$, hemos usado estos datos para ajustarlos al modelo de la normal. Una vez hecho esto, hemos comprobado por los métodos de la chi-cuadrado y de kolmogorov-smirnov que estos se ajustaban al modelo correctamente.

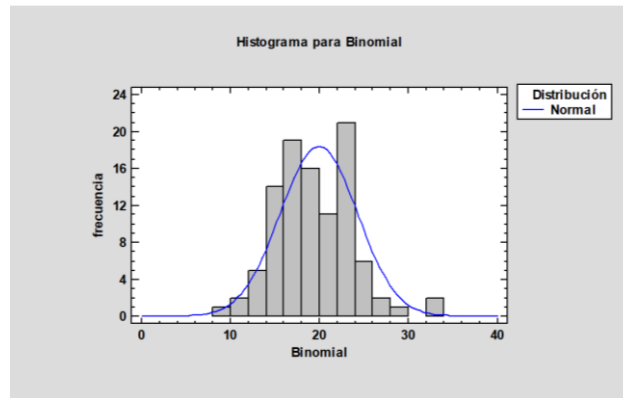
A continuación, hemos trabajado con los datos del fichero “AlumnosIndustriales.sf3”, con los que hemos realizado un análisis descriptivo y también los hemos ajustado al modelo de la normal. Hecho esto, con los resultados obtenidos del modelo, hemos calculado una probabilidad.

Más tarde, empleando los 100 datos que hemos recopilado, hemos realizado un nuevo ajuste para comprobar que se ajustaban, por la prueba de la chi-cuadrado, al modelo.

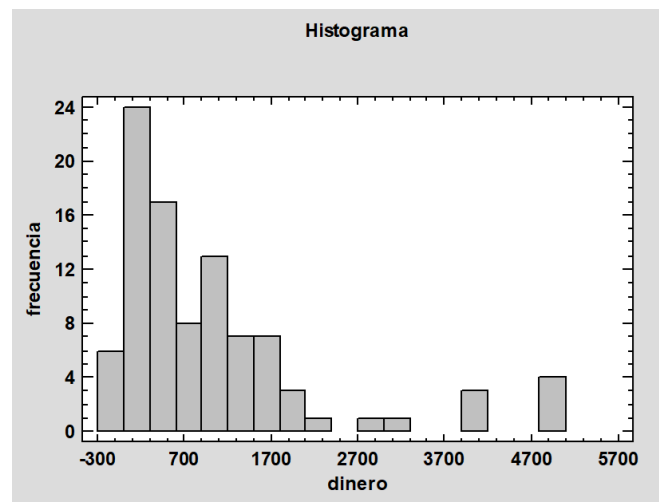
Para finalizar, hemos realizado el informe StatReporter con los análisis realizados en cada uno de los pasos anteriormente descritos. Este informe lo hemos exportado en formato .rtf y está adjunto en la entrega de la práctica.

3. Resultados

Para comenzar, en cuanto a los valores de la prueba de chi-cuadrado para los números de la distribución binomial, hemos obtenido un p de valor 0.09, el cual demuestra que el modelo escogido es válido para estos números. Además, la prueba de Kolmogorov-Smirnov también nos muestra un resultado válido.

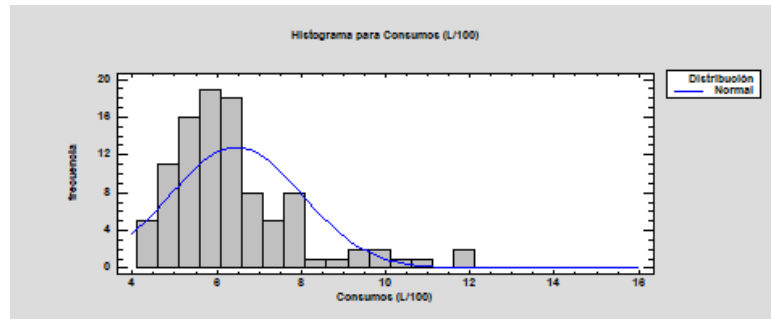


Pasando a los ejercicios con los datos del fichero, en concreto, el dinero que portaban los alumnos en el momento de ser preguntados: con el análisis descriptivo de los datos observamos los valores de la media = 1039.23, desviación estándar = 1200.14, coeficiente de variación = 115.48%, rango = 5000, coeficiente de asimetría = 7,95 y el coeficiente de curtosis = 7,60. También hemos representado los datos en el siguiente histograma:

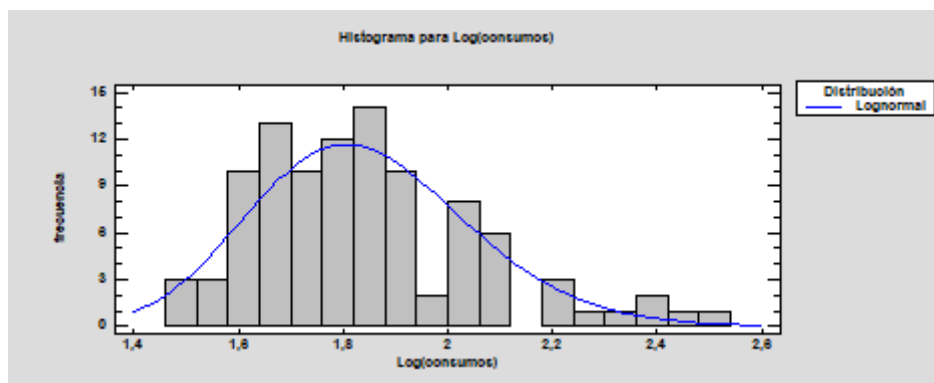


Estos mismos datos los hemos ajustado a la distribución normal después de aplicarles una transformación logarítmica. Hecho esto, hemos obtenido la estimación de los parámetros del modelo, cuyos valores son 6.41473 para la media y 1.22683 para la desviación estándar. Para terminar con estos valores, utilizando estos últimos datos, hemos desarrollado la distribución de probabilidad con la que hemos calculado $P(X > 1500)$, pero como los datos fueron transformados, tendríamos que calcular $P(X > \ln(1500)) = 0.232773 = 23\%$.

Continuamos empleando nuestros 100 datos, los cuales tratan sobre el consumo de 100 coches de gasolina en L/100(km), los cuales hemos tratado de ajustar al modelo de la normal, para lo que les hemos aplicado una transformación logarítmica, aunque aún no se ajustaban correctamente al modelo.



Por lo que hemos tratado de ajustar los valores transformados al modelo lognormal, al cual sí que se ajustaban correctamente como nos indica la prueba de la chi-cuadrado, que nos daba un valor de $p = 0.138$.



Por último, en cuanto al informe, en este hemos añadido todos los ajustes realizados junto con el análisis de los datos del dinero de los alumnos y la distribución de probabilidad correspondiente al modelo al que hemos ajustado esos mismos datos.

4. Conclusiones

Esta práctica nos ha ayudado a comprender lo útiles que son los modelos de probabilidad para analizar datos y que, si los ajustamos correctamente, podemos hacer predicciones bastante cercanas a la realidad. Además, hemos comprendido lo importante que es comprobar si el modelo es válido para evitar errores y nos ha servido para entender el uso de las transformaciones de los valores.