Nombre:.....
Titulación:.....
Grupo en el que estás matriculado:.....

-No está permitido utilizar (ni acceder a) ningún tipo de documentación
-Los datos están en la web de la asignatura. Descarga los datos en el Escritorio del ordenador

1. El fichero **FrenosITV.sf3** contiene datos de una muestra aleatoria de vehículos tomados en una estación de ITV. la variable **KM** es la lectura del cuentakilómetros de los automóviles. Obtén la siguiente información.

Media= 125131
Coef. de Asimetría= 1.09
Q1=78879.5
Q2=116089
Q3=164500
Percentil 80=176852
Rango Intercuartílico=85620.5

2. En ese mismo fichero, la variable **Peso** tiene el peso de cada vehículo, y la variable **Potencia** su potencia. ¿Qué peso esperamos observar en un coche de 100 unidades de potencia? Justifica adecuadamente tu respuesta

El gráfico XY muestra una relación lineal entre ambas variables, y su coeficiente de correlación es alto. Por lo tanto, una regresión lineal puede contestar a esa pregunta. Se puede predecir el peso a partir de la potencia con bastante precisión.

*La regresión lineal es: $\text{Peso}=632.17+6.119 \times \text{Potencia}$
para $\text{Potencia}=100$, el peso previsto por la regresión es: 1244.1*

3. Se tiene un lote de 200 artículos procedentes de un sistema productivo que produce, por término medio, un 1% de artículos defectuosos. Los artículos defectuosos se producen de forma independiente, y la proporción de artículos defectuosos que produce el sistema se considera estable. Nos interesa la variable aleatoria X =número de artículos defectuosos en el lote. ¿Cuál es la probabilidad de producir 2 ó menos artículos defectuosos? Justifica brevemente la respuesta.

$$X \text{ es } B(200,0.01). P(X < 3) = P(X \leq 2) = 0.6767$$

4. Los usuarios que acceden a un servidor web de una agencia de viajes lo hacen de forma independiente y con un ritmo medio de 5 accesos por minuto. ¿Cuál es la probabilidad de que accedan más de 10 clientes en un mismo minuto?

$X = \text{accesos/minuto}$ es Poisson(5). $P(X > 10) = 0.0137$

5. Continuamos con el fichero **FrenosITV.sf3**. Ajusta una normal a la variable **KM** utilizando, si es preciso, alguna transformación. A partir de ese modelo calcula la probabilidad de que a una ITV llegue un vehículo que haya recorrido más de 150.000 kilómetros.

Km no se ajusta a una normal. Su histograma muestra una clara asimetría, y el p-valor del test de la chi-cuadrado es muy pequeño (p-valor=0.002).

Podemos corregir la asimetría con alguna transformación del tipo x^c , con $c < 1$. Con la raíz cuadrada ($x^{0.5}$) el histograma es ya muy simétrico y el p-valor del test de la chi-cuadrado (p-valor=0.386) sí nos permite asumir normalidad en los datos transformados. Por tanto asumiremos que $x^{(0.5)}$ es $N(340.5; 95.98)$

*La probabilidad que se pide es, utilizando la normal de los datos transformados:
 $P(X > 150000) = P(X^{0.5} > 387.2983) = 0.3131$*

6. Consideremos la variable **Potencia** del fichero **FrenosITV.sf3**. Escribe, en los casos que sea posible, un intervalo de confianza para media poblacional y la desviación típica poblacional de esta variable. En cada caso, razona brevemente por qué se puede/no se puede proporcionar el intervalo de confianza.

Media:

Como se tienen 200 observaciones ($n > 30$), podemos hacer el intervalo para la media que proporciona el statgraphics sin importarnos qué distribución siguen los datos. Este intervalo es 95,0% confidence interval for mean: 65,1755 +/- 2,79719 [62,3783;67,9727]

Desviación típica.

La variable Potencia no se ajusta a una normal. El histograma es asimétrico y el p-valor del test de la chi-cuadrado es $< 0,05$. Por tanto, la inferencia para la desviación típica que realiza el statgraphics no es válida, pues se basa en la normalidad de la variable.

7. Continuamos con el fichero **FrenosITV.sf3**. La variable DCHA_DEL tiene la fuerza de frenado de la rueda delantera derecha, y la variable DCHA_TRA la de la rueda trasera. ¿Tienen la misma fuerza de frenado? Justifica adecuadamente tu respuesta así como los pasos que sigues.

Vamos a comparar la fuerza media de frenado de ambas ruedas. Tenemos una muestra de 200 datos para cada rueda, por lo que el tamaño muestral es suficiente para hacer inferencia sobre la diferencia de medias sin importarnos la distribución de los datos.

La hipótesis nula es que ambas ruedas tienen la misma fuerza media de frenado, y la alternativa es que ambas medias son diferentes.

Las cuasidesviaciones típicas son muy similares por lo que podemos hacer el contraste asumiendo varianzas iguales. En cualquier caso, como los tamaños muestrales son grandes, podríamos hacer el test de igualdad de medias usando estimaciones de la varianza diferentes en cada grupo.

Asumiendo varianzas iguales, el p-valor del contraste es casi nulo. Hay mucha evidencia de que ambas ruedas no frenan igual por término medio.