
GUIÓN del TRABAJO de Estadística Descriptiva y Análisis de Datos

El objetivo de este trabajo es recoger una serie de datos para realizar un estudio estadístico descriptivo completo. Los datos originales se pueden elegir libremente, con un mínimo de 100 observaciones y 6 variables (siendo 3 cuantitativas y 3 categóricas), o bien simular. En este último caso, se proporcionarán a l@s alumn@s dichos datos simulados, grabados en un disquete.

El tema sobre el que traten los datos es de libre elección. Si los datos han sido recogidos de manera original por l@s alumn@s (no han sido simulados), se tiene que consultar primero para ver si los datos se adaptan a las especificaciones generales del trabajo.

Los resultados del informe se tienen que entregar en papel, con las conclusiones obtenidas, y **no** se pueden entregar en un disquete ni en un CD. Se tiene que usar un procesador de texto estándar (como *Word*) y no se escribirá a mano.

Ejemplo de un posible estudio (no se puede usar este ejemplo para el trabajo):

Se toman los valores de las alturas de los jugadores de baloncesto de la *NBA* (extraídos de la página oficial de la *NBA*, www.nba.com), elegidos al azar de entre la población que forman todos los jugadores de dicha liga durante una temporada. Se consideran, también, otras variables medidas sobre cada jugador que indiquen, por ejemplo, la zona de la cancha en la que juega habitualmente, el país de procedencia etc.

Tareas a realizar

- Construir las tablas de frecuencias para las diferentes variables recogidas en la muestra.
- Hacer lo mismo que en el punto anterior, pero segmentando según los diferentes grupos formados al considerar las variables categóricas. En los datos de la *NBA* significaría construir tablas de frecuencias por separado, por ejemplo, para el grupo de jugadores que son aleros, o que son bases y comparar las distribuciones obtenidas. Comentar y explicar, en la muestra recogida, si se observan diferencias entre los grupos a nivel de tablas de frecuencias.
- Calcular las medidas de centralización, de dispersión y de forma para cada una de las variables de la muestra.
- Repetir lo mismo, sólo con las variables cuantitativas, para los grupos formados al considerar las variables categóricas. Comentar y explicar si se observan diferencias entre dichos grupos a nivel de las medidas calculadas.
- Dibujar los histogramas, diagramas de barras, gráficas de sectores y diagramas de cajas para las variables que les correspondan.

- Repetir lo mismo, sólo con las variables cuantitativas, para los grupos formados al considerar las variables categóricas. Comentar y explicar si se observan diferencias entre dichos grupos.
- Realizar transformaciones lineales (tipificación) de las variables cuantitativas. Realizar transformaciones no lineales de las variables cuantitativas que no sean simétricas.
- Estudiar los *outliers* y justificar su existencia.
- Configurar una tabla de doble entrada (contingencia) con dos de las variables cuantitativas. Calcular medidas de centralización y de dispersión marginales. Calcular las medidas de centralización y de dispersión condicionadas a, por lo menos, una de las categorías de una de las variables (se puede hacer este apartado entero *a mano*, porque con un programa de ordenador es más complicado).
- Calcular las covarianzas y coeficientes de correlación para, al menos, tres variables. Comentar los resultados obtenidos.
- Calcular los coeficientes de la recta de regresión entre dos de las variables cuantitativas. Dibujar el diagrama de dispersión asociado, y la recta de regresión obtenida.
- Repetir lo mismo para los grupos formados al considerar una de las variables categóricas. Comentar y explicar si se observan diferencias entre dichos grupos.

Exponer las conclusiones definitivas y generales sobre el trabajo realizado.