**http://www.uam.es/personal_pdi/stmaria/jmurillo/javier_murillo_archives/logo.gifhttp://www.uam.es/personal_pdi/stmaria/jmurillo/javier_murillo_archives/uamlar.gif**

**FACULTAD DE FORMACIÓN DE PROFESORADO Y EDUCACIÓN**

**ESTADÍSTICA APLICADA A LA INVESTIGACIÓN EDUCATIVA**

**Guía de Estudio**

**1. Conceptos básicos de Estadística Descriptiva**

* + **Estadística descriptiva *versus* inferencial**
  + **Medidas de tendencia central: Media, Moda y Mediana**
  + **Medidas de Variabilidad: Rango, Varianza y Desviación típica.**
  + **Asimetría y Curtosis**
  + **Cuartiles y Percentiles**
  + **Relación: Correlación y covarianza**
  + **Predicción: La recta de Regresión: elementos, intercepto y pendiente**

**Actividad: Localizar la definición de cada uno de los conceptos anteriores y comprender su significado, así como entender las implicaciones que tienen los diferentes valores que toman los índices con los que se miden. Cualquier libro de estadística básica o una búsqueda en Internet pueden servir para ello.**

**Por ejemplo: Los apuntes de** [**Estadística Descriptiva**](http://www.unavarra.es/estadistica/I.T.T.Imagen/descriptiva.pdf) **de Ignacio Cascos (Universidad Pública de Navarra)**

**Es necesario tener grabado en el disquete de trabajo las definiciones. La evaluación final dependerá de los contenido en el disquete.**

**2. Utilización del Excel para el análisis estadístico**

* + **Manejo básico del Excel: las funciones**
  + **Funciones de Excel para estadística descriptiva (**[**material en Excel**](http://www.uam.es/personal_pdi/stmaria/jmurillo/Estadistica/Funciones%20en%20Excel.xls)**)**
  + **Las tablas dinámicas**
  + **Representación y estimación de las ecuaciones de regresión**

**Actividades:**

**1. Explorar el Excel para descubrir su funcionamiento básico. Por ejemplo: pegado especial, funciones, ordenación de datos....**

**2. Calcular, sobre la base de** [**datos 1**](http://www.uam.es/personal_pdi/stmaria/jmurillo/Estadistica/Datos_1.xls) **de Excel, las diferentes medidas de tendencia central y variabilidad de las diferentes variables.**

**3. Elaborar la tabla de frecuencias absolutas y relativas de la variable "Rendimiento en Matemáticas"**

**4. Estimar la asimetría y curtosis de la variable "Rendimiento en Matemáticas" e interpretar sus resultados.**

**5. Calcular el cuartil 1, 2 y 3 y los percentiles 10, 25 y 35 sobre la variable "Rendimiento en Matemáticas"**

**6. Estimar el índice de correlación entre "Rendimiento en Matemáticas" y diferentes variables e interpretar los resultados. ¿Cuál es la variable que predice mejor el Rendimiento en Matemáticas?**

**7. Calcular el intercepto y la pendiente de la ecuación de regresión lineal que predice el rendimiento en Matemáticas a partir de la variable "Situación socio-económica".**

**8. Explorar la utilización de la "Tabla dinámica" para realizar tablas de contingencia.**

**9. Posibilidades gráficas de Excel y Word. Presentar los datos en a) Gráfico de barras, b) grafico circular, c) gráfico de líneas, y d) gráfico de dispersión... y señalar las ventajas y limitaciones de cada tipo**

**10. A partir de un gráfico de dispersión, dibujar y estimar la ecuación de regresión que prediga el rendimiento en matemáticas a partir de la situación sociocultural.**

**3. Introducción al SPSS**

**Nota: los datos para trabajar este programa los podéis encontrar en:** [**Datos**](http://www.uam.es/personal_pdi/stmaria/jmurillo/Estadistica/Datos%20curso.sav)

* + **Manejo básico del SPSS.**
  + **Introducción de datos y definición de las propiedades de las variables.**
  + **Calculo, recodificación y categorización**
  + **Frecuencias**
  + **Descriptivos: Media, moda, mediana, asimetría, curtosis, rango, percentiles, varianza, desviación típica...**
  + **Correlación y regresión lineal simple**
  + **T de Student**
  + **Análisis de Varianza**
  + **Chi cuadrado**
  + **Análisis de regresión múltiple**

**Actividades:**

**1. Explorar el SPSS para descubrir sus funcionamiento básico.**

**2. Definir una base de datos (propiedades de las variables) e introducir datos.**

**3. Sobre la base de datos de trabajo, cambiar las etiquetas, la definición de los valores, los valores perdidos y el tipo de medida de varias variables.**

**4. Crear una variable "edad" a partir de los datos del año de nacimiento.**

**5. Recodificar la variable "comienzo de la escolaridad" pasando de los valores 1 a 6  a 0 a 5 .**

**6. Categorizar la variable "rendimiento en matemáticas" en 5 grupos.**

**7. Calcular el número de alumnos en cada centro, y el porcentaje de alumnos que tienen diferente número de hermanos, y que tienen diferente número de libros en casa..**

**8. Calcular la media, moda, mediana, varianza, rango, desviación típica, asimetría y curtosis de las variables "rendimiento", "actitud" y "satisfacción".**

**9. ¿Cuál es el percentil 15 de la variable rendimiento?**

**10. Calcular el rendimiento medio para cada grupo de alumnos generado por el número de libros que tienen. ¿Cuál es el centro con mayor rendimiento promedio? ¿Y con los datos más dispersos?**

**11. Estima el intervalo de confianza de la variable rendimiento de los alumnos. ¿Cómo se interpreta?**

**12. Formula las hipótesis estadísticas nula y alternativa de la significación de una correlación.**

**13. ¿Qué dice el teorema central del límite?**

**14. Qué es el error típico de la media? ¿Qué relación tiene con la desviación típica?**

**15. ¿Qué es error de tipo I y error de tipo II? ¿Cuál debemos evitar prioritariamente?**

**16. ¿Para qué sirve la t de Student? Utilízala con los datos de trabajo y explica los resultados.**

**17. ¿Qué diferencias hay entre la t de Student para datos independientes y para datos relacionados? Pon un ejemplo de cada uno con los datos de trabajo, estímalos e interpreta los resultados.**

**18. ¿Para qué sirve el Análisis de Varianza? ¿En qué se parece y en qué se diferencia de la t de Student?**

**19. ¿Qué son y para qué son los contrastes a posteriori del ANOVA?**

**20. Estimar la ecuación de regresión lineal simple con rendimoiento en matemáticas como VD y situación socio-económica como VI.¿Cuánta varianza explica esa relación.**

**21. ¿Qué variable explica mayor porcentaje de varianza de la variable rendimiento en matemàticas?**

**22. Estima la ecuación de regresión múltiple con rendimiento en mates como VD y nº de libros, situación socio-económica y satisfacción del centro ¿Cuánta varianza explica?**

**23. Escoge las variables que mejor expliquen en rendimiento en matemáticas y elabora la ecuación de regresión que más varianza explique.**

Ejercicio 2. *Se tiene anotado los metros de crecimiento de un conjunto de ´arboles de un*

*bosque en cierto periodo de tiempo.*

*N´umero de ´arboles 2 6 10 5 10 3 2 2*

*Crecimiento(m) 1 2 3 4 5 6 7 8*

*Introduce los datos y haz un resumen descriptivo de ellos. Si se quieren talar los que*

*m´as han crecido y exactamente un 15% del total de ´arboles. ¿A partir de qu´e metros de*

*crecimiento deberemos talar?*

*Nota: para crear con rapidez el correspondiente vector de datos en R, introduce los*

*valores haciendo uso de* rep *del siguiente modo c(rep(1,2),rep(2,6),...)*

Ejercicio 3. *Se quiere estudiar el tiempo de arranque de un modelo de ordenador. Para*

*ello se han anotado los tiempos de arranque en segundos de 42 ordenadores de ese modelo.*

*Haz un an´alisis descriptivo los datos. ¿Por debajo de cu´antos segundos arrancan el 70%*

*de los ordenadores?*

*Segundos de arranque 37 37.2 37.5 38 38.1 38.5*

*N´umero de ordenadores 1 5 15 6 10 5*

3

Caracter¶³sticas comparativas:

Media:

*Ventajas*: Usa toda la informaci¶on que existe en los datos

*Inconvenientes*: Cualquier valor extremos puede provocar un desplazamiento

de la media, hasta el punto de que deje de ser una medida representativa de

la mayor¶³a de los datos.

Mediana:

*Ventajas*: No se ve afectada por la aparici¶on de observaciones an¶omalas. Por

ello, en tales casos la podemos considerar como una medida m¶as representativa

de la mayor parte de los datos que la media

*Inconvenientes*: No utiliza toda la informaci¶on de los datos (s¶olo los valores

centrales)

Moda:

*Ventajas*: Se puede calcular para todo tipo de variables, incluidas las cualita-

tivas

*Inconvenientes*: Pueden existir varias modas

En distribuciones unimodales campaniformes, la mediana est¶a con frecuencia

comprendida entre la mediana y la moda.

*Moda '* 3*Med ¡* 2¹*x*