

Práctica 1: Introducción

Problema 1 *En el archivo C96_altura_M.dat aparecen las alturas en cm de 55 mujeres. Realiza las siguientes operaciones escribiendo todos los comandos que utilices en un script llamado practica1.R .*

- i) Introduce los datos del archivo en el programa R utilizando el comando scan. Comprueba que los datos son los mismos que aparecen en el archivo abriéndolo con el bloc de notas.*
- ii) Guarda los datos del archivo en un vector dentro del programa R.*
- iii) Ordena los datos de menor a mayor y guarda los datos ordenados en un vector.*
- iv) Guarda el vector de los datos ordenados en un archivo llamado datos_ordenados.dat .*
- v) Calcula la suma total, la media, la varianza, la desviación típica, la mediana y los cuartiles de los datos de la muestra.*
- vi) Resta al vector de datos su media y divídelo por la desviación típica.*
- vii) Guarda el vector que acabas de obtener en un archivo llamado datos_normalizados.dat .*
- viii) Calcula qué porcentaje tiene una altura superior a 175cm. Utiliza el comando which y la ayuda si es necesario.*
- ix) Calcula el máximo y el mínimo de los datos.*
- x) Genera un vector en forma de secuencia desde 0 hasta 10 con paso 0.5 y otro con paso 1.*

Práctica 2: Gráficas

Problema 2 Realiza los gráficos adecuados que se piden a continuación y guarda las figuras en archivos jpeg. Comenta en cada caso las conclusiones que sacas acerca de los datos al ver la gráfica.

- i) Estudia gráficamente la distribución de las alturas de las 55 mujeres del archivo *C96_altura_M.dat*
- ii) Compara gráficamente las alturas de 55 mujeres y 38 hombres de los archivos *C96_altura_M.dat* y *C96_altura_V.dat* respectivamente.
- iii) Estudia la relación entre la altura y el peso de las 55 mujeres cuyos datos figuran en los archivos *C96_altura_M.dat* y *C96_peso_M.dat*
- iv) Compara gráficamente los tiempos de reacción medidos para un grupo en el que los participantes ingieren alcohol hasta alcanzar un nivel de alcoholemia de 0.6g/l y otro con alcoholemia 0.0g/l. Los datos figuran en los archivos *tiempo_A.dat* y *tiempo_B.dat* respectivamente.
- v) Compara gráficamente los siguientes datos de nacimientos de un hospital. De un total de 500 nacimientos 398 fueron normales, 54 necesitaron forceps, y el resto cesárea.

Práctica 3: Gráficas Avanzadas

Problema 3 *Realiza los gráficos adecuados que se piden a continuación y guarda las figuras en archivos jpeg.*

- i) Añade títulos y colores a todas las gráficas de la práctica 2.*
- ii) Ajusta los tamaños y ejes de manera apropiada en todas las gráficas de la práctica 2.*
- iii) Junta las 3 primeras gráficas de la práctica 2 en una sola gráfica con 3 paneles.*
- iv) El porcentaje de bebedores de riesgo en España en los años 1987, 1993, 1997, 2001 y 2007 figura en los archivos bebedores_M.dat para hombres y bebedores_W.dat para mujeres. Representa estos datos en una sola gráfica y utiliza una leyenda.*

Práctica 4: Distribuciones

Problema 4 De una determinada enfermedad se espera que se produzcan 20 casos en Cantabria este año. ¿Cuál es la probabilidad de que se produzcan 10 casos o menos?

Problema 5 Se dispone de una muestra de 8 sujetos que proviene de una población con una prevalencia de tabaquismo del 40%.

- i) ¿Qué probabilidad hay de que haya algún fumador?
- ii) ¿Qué probabilidad hay de que haya más de dos fumadores?
- iii) ¿Qué probabilidad hay de que haya menos de 5 fumadores?

Problema 6 La eliminación urinaria total por día de α -cetoesteroides medida en adultos jóvenes tiene una media de 12 UI y una desviación típica de 1.9 UI.

- i) Calcula el porcentaje de población que tiene una eliminación superior a 9.8 UI.
- ii) Calcula el intervalo de eliminación dentro del cuál se sitúa el 90% de la población.

Problema 7 Un medicamento se expide en tabletas de 100 unidades de peso cada una, que contienen una sustancia S en un promedio de 22 unidades con una desviación estándar de 4 (se supone una distribución Normal). Sabiendo que una tableta es efectiva solo si contiene entre 18 y 28 unidades de la sustancia S ¿Qué porcentaje de tabletas serán efectivas?

Problema 8 La probabilidad de que un recién nacido sea varón es 0.5

- i) Calcula las probabilidades de que de 10 recién nacidos haya 0, 1, ..., 10 varones.
- ii) Calcula las probabilidades acumuladas correspondientes.
- iii) Representa todas las probabilidades calculadas en una gráfica con dos paneles.

Problema 9 En Cantabria se producen en promedio 4 casos al año de infección neumocócica invasiva en niños.

- i) Calcula las probabilidades de que se produzcan 0, 1, ..., 10 casos el año que viene.
- ii) Calcula las probabilidades acumuladas correspondientes.
- iii) Representa todas las probabilidades calculadas en una gráfica con dos paneles.

Problema 10 El colesterol se mide en una población con media 196 mg/dL y desviación estándar 46 mg/dL.

- i) ¿Cuál es la probabilidad de que un individuo tenga menos de 166 mg/dL?
- ii) ¿Cuál es la probabilidad de que tenga más de 166 mg/dL?
- iii) Dibuja en una gráfica con dos paneles, la curva de Gauss correspondiente entre 0 y 400 mg/dL y las probabilidades acumuladas en el mismo intervalo.

Práctica 5: Tests de Hipótesis e Intervalos de Confianza

Problema 11 *En una muestra de 584 personas, se encuentra que 60 tienen algún tipo de arritmia. Calcula la proporción de pacientes con arritmia, con confianzas del 90, 95 y 99%. Somete a prueba la hipótesis nula que afirma que la proporción de pacientes con arritmia es del 13%.*

Problema 12 *En el archivo peso_nacimiento.dat se recogen los pesos al nacer de 820 recién nacidos. Calcula los intervalos de confianza de la media al 80, 90 y 95%. Somete a prueba la hipótesis nula que afirma que el peso medio al nacer es de 3220 gramos.*

Problema 13 *En un estudio con 20 pacientes, se mide la diuresis en 24 horas que figura en el archivo diuresis.dat siendo las unidades ml. Calcula los intervalos de confianza de la media al 90% y 95%.*

Problema 14 *Mediante el tratamiento habitual, una enfermedad se cura en el plazo de una semana en un 60% de los pacientes. Aplicamos un tratamiento adicional a 250 pacientes de los cuales se curan 170. ¿Es efectivo el tratamiento adicional?*

Problema 15 *En un grupo de 120 pacientes con enfermedad pulmonar obstructiva crónica, el volumen espiratorio forzado en un segundo tenía una media de 2.5 litros antes de recibir tratamiento. Tras el tratamiento con salbutamol, medimos el volumen espiratorio que figura en el archivo volumen.dat. ¿Es efectivo el tratamiento?*

Práctica 6: Comparación de proporciones

Problema 16 *Se desea comprobar si un nuevo tratamiento “A” para una enfermedad grave mejora los resultados de la terapia aplicada hasta el momento “B”. Para ello se realizó un estudio en el que se administra A a diez pacientes y B a otros diez, registrándose en cada caso si la respuesta ha sido curación (+) o fracaso terapéutico (-). Con el tratamiento A la curación se da en 8 de los 10 casos, mientras que con el tratamiento B la curación se da en 3 de los 10 casos.*

Representa gráficamente los datos. ¿Crees que hay diferencias significativas entre los dos tratamientos? Realiza los cálculos necesarios para responder a la pregunta.

Problema 17 *Para mejorar la asistencia a un programa de deshabituación alcohólica se propone incorporar una llamada el día previo a la cita como recordatorio. Antes de poner en marcha esta medida se quiere valorar si realmente es efectiva. Para ello se realiza un estudio con 45 pacientes de los que 25 reciben la llamada para recordar la cita y otros 20 actúan como controles. Al final del estudio 20 de los 25 pacientes con recordatorio acudieron a la cita frente a solo 8 en el grupo control.*

Representa gráficamente los datos. ¿Debe incorporarse el recordatorio telefónico? Realiza los cálculos necesarios para responder a la pregunta.

Problema 18 *Para valorar la eficacia de una vacuna frente a una determinada enfermedad “X” se realizó un ensayo clínico en el que se seleccionó a 350 personas de los que a 175 se les asignó la vacuna, actuando los otros 175 como grupo control. Tras un periodo de seguimiento de 6 meses el 15% del grupo vacunado padeció la enfermedad frente al 25% del grupo control.*

i) *¿Es eficaz la vacuna?*

ii) *Calcula el intervalo de confianza de la diferencia de proporciones de enfermos entre los dos tratamientos.*

Problema 19 *Se desea comprobar si determinada sustancia produce reacción alérgica en el ser humano. Para ello se reclutan 45 voluntarios y se les administra dicha sustancia por vía intramuscular en cada uno de sus brazos. La reacción resulta positiva en el brazo derecho de 15 sujetos, al mismo tiempo que es negativa en el brazo izquierdo. A su vez da positiva en el brazo izquierdo de 30 sujetos al tiempo que da negativa en el brazo derecho.*

¿Podemos afirmar que existen diferencias estadísticamente significativas en el porcentaje de reacciones alérgicas por el hecho de administrar la sustancia en un brazo o el otro?

Práctica 7: Comparación de medias

Problema 20 Para valorar el incremento del tiempo de reacción producido por una alcoholemia de 0.6g/l se realiza un ensayo clínico con una muestra de 12 sujetos que se divide al azar en dos grupos: el grupo A en el que los participantes ingieren alcohol hasta alcanzar un nivel de alcoholemia de 0.6g/l y el grupo B con alcoholemia 0.0g/l. Los tiempos de reacción medidos para ambos grupos que figuran en los archivos tiempo_A.dat y tiempo_B.dat .

- i) ¿Podemos afirmar que la alcoholemia de 0.6 g/l incrementa el tiempo de reacción?
- ii) Calcula el intervalo de confianza para la diferencia de medias al nivel de confianza del 95%.

Problema 21 Imagina ahora que en los datos del estudio anterior no se cumpliera el supuesto de homocedasticidad de varianzas y repite el ejercicio.

Problema 22 Se ha realizado un estudio sobre tabaco y carboxihemoglobina (CH) en sangre con objeto de verificar la hipótesis de que los fumadores tienen valores superiores de CH que los no fumadores. Para ello se ha seleccionado una muestra de 28 voluntarios a los que se ha clasificado en 2 grupos: fumador (16 personas) y no fumador (12 personas), midiéndose el nivel de CH. Los resultados aparecen en los archivos CH_Fum.dat y CH_NoFum.dat .

- i) ¿Podemos afirmar que en los fumadores el nivel de CH es mayor que en los no fumadores?
- ii) Calcula el intervalo de confianza para la diferencia de medias al nivel de confianza del 95%.

Problema 23 Para comprobar si después de respirar una atmósfera cargada de humo de tabaco aumenta la CH en sangre, a los 28 sujetos de la muestra del problema anterior se les encierra en una sala sin ventilación durante 5 horas permitiéndoles fumar para lograr un ambiente cargado de humo. Al finalizar el experimento se analiza el nivel de CH y se obtienen los resultados que aparecen en los archivos CH_Fum2.dat y CH_NoFum2.dat . Los niveles iniciales, antes del experimento son los dados en los archivos CH_Fum.dat y CH_NoFum.dat .

- i) ¿Podemos afirmar que después de respirar durante 5 horas en una atmósfera cargada de humo de tabaco aumenta la CH de los sujetos de la población?
- ii) Calcula el intervalo de confianza para la diferencia de medias al nivel de confianza del 95%.

Practica 8: ANOVA

Problema 24 *Para conocer cuál es el mejor tratamiento de la anemia ferropénica, se comparan tres preparados comerciales. Los resultados obtenidos son los que figuran en el archivo anemia.dat . ¿Hay diferencias entre los tratamientos?*

Problema 25 *Se desea conocer si el consumo de anfetaminas o ginseng aumenta el rendimiento académico. Para ello, se seleccionan 15 alumnos de bioestadística; durante la semana anterior al examen 5 toman anfetamina, 5 toman ginseng y 5 no toman nada. Los resultados obtenidos figuran en el archivo drogas.dat . Con estos datos, ¿puede demostrarse que alguno de los grupos tuvo mejor rendimiento?*

Problema 26 *Para conocer el efecto de tres antiarrítmicos en la frecuencia cardíaca, se procede de la siguiente forma: (1) se seleccionan 9 voluntarios; (2) se les administra el antiarrítmico A y 30 minutos después se mide la frecuencia cardíaca; (3) una semana más tarde se procede igual con el antiarrítmico B; (4) otra semana después se procede igual con el antiarrítmico C. ¿Puede decirse que los tres antiarrítmicos tienen el mismo efecto?*

Práctica 9: Regresión y Correlación

Problema 27 *El número de trabéculas en hueso trabecular es un parámetro para cuantificar el grado de osteoporosis. Dicho parámetro se estimó en necropsias de cresta ilíaca en un grupo de mujeres y en otro de hombres. Las edades de las 16 mujeres figuran en el archivo Edad.dat y el correspondiente número de trabéculas por milímetro cúbico en el archivo Conectividad.dat .*

- i) Representa los datos en un gráfico adecuado.*
- ii) Ajusta una recta de regresión lineal a los datos, y representala. Escribe la ecuación de la recta.*
- iii) Calcula el coeficiente de correlación lineal entre edad y conectividad para la muestra de mujeres estudiada.*
- iv) Dar una medida de hasta qué punto la conectividad viene explicada por la edad.*
- v) ¿Podemos afirmar que el coeficiente de correlación lineal es significativamente distinto de cero?*

Problema 28 *En un estudio epidemiológico se reclutan 10 pacientes con insuficiencia renal crónica, a los que se miden los valores de hemoglobina (g/dl), creatinina (mg/dl) que figuran en los archivos Hemoglobina.dat y Creatinina.dat respectivamente.*

- i) Representa los datos en un gráfico adecuado.*
- ii) Ajusta una recta de regresión lineal a los datos, y representala. Escribe la ecuación de la recta.*
- iii) Calcula el coeficiente de correlación lineal entre hemoglobina y creatinina.*
- iv) Dar una medida de hasta qué punto el descenso de hemoglobina viene explicado por el aumento de creatinina.*
- v) ¿Podemos afirmar que el coeficiente de correlación lineal es significativamente distinto de cero?*