

Ejercicios EYP1113L Control 1

Primera parte

04-04-2024

Introducción

Este documento contiene ejercicios para probar algunas características del cálculo numérico en R. En particular se encuentra el manejo de datos y cálculo de sumatorias y probabilidades a partir de modelos probabilísticos conocidos.

Sumatorias

Calcule las siguientes sumatorias:

1. $\sum_{k=1}^{10} k$
2. $\sum_{k=10}^{73} k^4$
3. $\sum_{k=3}^7 6\sqrt{3k}$
4. $\log 0.1 + \log 0.2 + \log 0.3 + \dots + \log 2.1$
5. $1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{100}$
6. $\sin^2(\pi/5) + \sin^2(2\pi/5) + \sin^2(3\pi/5) + \dots + \sin^2(4\pi)$ (usar variable *pi* definida en *R*)
7. $-\frac{0.6^3}{3} + \frac{0.6^4}{4} - \dots + \frac{0.6^{20}}{20} = \sum_{k=3}^{20} \frac{(-0.6)^k}{k}$
8. $\sum_{k=11}^{33} \frac{1}{\sin^2\left(1 + \frac{\pi k}{34}\right)}$
9. $\frac{\ln 2}{1.5} + \frac{\ln 3}{3} + \frac{\ln 4}{4.5} + \dots + \frac{\ln 23}{33}$
10. $\frac{1}{19} \sum_{j=1}^{20} \left(\sin^3(2j) - \frac{1}{20} \sum_{k=1}^{20} \sin^3(2k) \right)^2$

Respuestas:

(1) 55, (2) 4.2892778×10^8 , (3) 114.9738081, (4) -2.9741481, (5) 5.1873775, (6) 10, (7) -0.050003, (8) 4117.3927316, (9) 3.9920938, (10) 0.3298818

Bases de datos

Abalon.xlsx

Base de datos consiste en 400 locos (Abalón Rojo) observados en 4 centros de estudio de Chile. Se registran varias características físicas

- **largo**: El largo del loco (en cm)
- **diametro**: El diámetro del loco (en cm)
- **alto**: El alto del loco (en cm)
- **pesot**: Peso total, cuerpo más concha (en g)
- **pesocu**: Peso cuerpo (en g)
- **pesoco**: Peso concha (en g)
- **anillos**: Número de anillos en la concha
- **centro**: Nombre del centro de estudio

Responda las siguientes preguntas

1. ¿Cuántos locos se registraron del **centro** Caldera?
2. ¿Cuál es el **largo** promedio de todos los locos registrados?
3. ¿Cuál es la desviación estándar del peso total (**pesot**) de los locos registrados?
4. ¿Cuál es el peso total (**pesot**) promedio de los locos registrados en el centro Coquimbo?
5. Considera sólo los locos de más de 10 cm de **largo**, ¿cuánto es su **largo** promedio?
6. Considera los locos registrados en Puerto Montt y Chiloé, ¿qué **alto** promedio tienen los locos registrados en estos dos centros en conjunto?
7. ¿Cuántos locos se registraron con no menos de 12 **anillos**?
8. ¿Cuántos locos se registraron con al menos 12 **anillos** en el **centro** Caldera?
9. ¿Qué porcentaje de locos pesaron (**pesot**) entre 200 y 600 gramos (inclusive)?
10. ¿Qué porcentaje de locos fueron registrados en el **centro** Chiloé?

Respuestas:

(1) 100, (2) 13.23975, (3) 220.1695729, (4) 385.845, (5) 14.1606509, (6) 3.4565, (7) 86, (8) 22, (9) 0.5875 o 58.75, (10) 0.25 o 25.

BandaAncha.xlsx

Esta base de datos consiste la observación de velocidades promedio de transmisión de Internet Banda Ancha en hogares de varias ciudades de Chile, provista por cuatro compañías telefónicas: Claro, Entel, Movistar y WOM. Nombres de variables se explican por sí solas. **Velocidad** en MB/s.

Responda las siguientes preguntas:

1. ¿Qué velocidad promedio se registró para todo Chile y todas las compañías en conjunto?
2. ¿Cuántas ciudades fueron observadas en el estudio?
3. ¿Cuál es la velocidad de internet provista por Claro (C) observada en Iquique?
4. ¿Cuál es la velocidad promedio de internet en Illapel entre todas las compañías?
5. ¿Cuál es la desviación estándar de velocidad de internet en Osorno entre todas las compañías?

6. ¿Cuál fue el operador de mayor velocidad observado en Arauco?
7. ¿Cuál fue la ciudad con mayor velocidad de internet provista por WOM (W)?
8. ¿En cuántas ciudades se observó mayor velocidad provista por Claro (C) que por WOM (W)?

Respuestas:

(1) 11.23052, (2) 50, (3) 7.864, (4) 10.674, (5) 4.4443772, (6) M, (7) Chiloe, (8) 27.

ENS.xlsx

La base de datos consiste en un extracto de la Encuesta Nacional de la Salud (ENS) tomada el 2009. Consiste en encuestar y medir ciertos indicadores corporales de cada persona, junto con su región de origen, edad, y otros factores. Datos se encuentra en hoja "ENS2009", y descripción de las variables se encuentra en la segunda hoja "Variables".

Responda las siguientes preguntas:

1. ¿Cuántos pacientes registrados tenían 21 años?
2. ¿Cuál es el nivel de colesterol total (COLES) promedio en todos los registros?
3. ¿Cuál es la desviación estándar del PESO de los pacientes registrados?
4. ¿Cuántos pacientes registraron un nivel de colesterol HDL superior a 40 mg/dL?
5. ¿Cuántos pacientes registraron un nivel de triglicéridos (TGD) entre 120 a 200 mg/dL (inclusive)?
6. ¿Qué porcentaje de pacientes pesa menos de 60 kg?
7. ¿Qué porcentaje de pacientes fuma? (FUMADOR == 1)
8. ¿Cuál es el nivel de azúcar promedio (GLUBASAL), en mg/dL en los hombres de la muestra (SEXO == 1)?
9. ¿Cuál es el PESO promedio de los pacientes registrados en la Región de Valparaíso (REGION == 5)?
10. ¿Qué porcentaje de los pacientes registrados tiene menores niveles de colesterol "bueno" (HDL) que sus respectivos niveles de colesterol "malo" (LDL)?

Respuestas:

(1) 34, (2) 194.7220676, (3) 14.8686919, (4) 1747, (5) 803, (6) 0.2111332 o 21.1133201, (7) 0.3252485 o 32.5248509, (8) 98.2368169, (9) 71.447541, (10) 0.9821074 o 98.2107356.

PCR.txt

Base de datos consiste en 112 pruebas de PCR realizadas el 2020 en Chile durante la pandemia de Covid-19. Cada observación describe 3 variables más un indicador numérico.

- CASO: El número del caso
- EXTRANJERO: Indicador si paciente es extranjero (1) o no (0).
- PCR: Resultado del test: Positivo (1) o negativo (0).
- EDAD: Edad del paciente (en años)

Responda las siguientes preguntas

1. ¿Cuál es el porcentaje de extranjeros testeados en la muestra?
2. ¿Qué edad tiene el paciente más longevo (mayor edad)?
3. ¿Qué porcentaje de los pacientes extranjeros registró un test PCR positivo?
4. ¿Qué porcentaje de pacientes tiene 40 años o menos?

5. ¿Qué edad tuvo la persona más joven en registrar un resultado positivo del test PCR?

6. ¿Cuántas personas de 59 años dio positivo en el test PCR?

Respuestas:

(1) 0.4375, (2) 67, (3) 0.2857143, (4) 0.1785714, (5) 44, (6) 1.

Sismos2021.xlsx

La siguiente base de datos contiene 410 sismos ocurridos en la zona norte y central de Chile. Sus variables son,

- **Fecha:** Día y hora del sismo
- **Tiempo:** Tiempo desde el último sismo (en minutos)
- **Latitud:** Ángulo de latitud del foco del sismo
- **Longitud:** Ángulo de longitud del foco del sismo
- **Profundidad:** Profundidad del sismo (en km)
- **Magnitud:** Magnitud del sismo en grados Richter
- **Referencia:** Texto que indica algún lugar de referencia al foco del sismo.

Responda las siguientes preguntas:

1. ¿Cuántos sismos ocurrieron al menos una hora (60 minutos) después del último sismo?
2. ¿Qué porcentaje de los sismos en la muestra fueron registrados a latitud 30° S o más al sur? (-30 o menor)?
3. ¿Cuántos sismos tienen magnitud de hasta 3 grados y ocurrieron a más de 3 horas después del último sismo?
4. ¿Qué magnitud tiene el sismo registrado a mayor profundidad?
5. ¿Cuál es el percentil 90% del tiempo transcurrido entre sismos?
6. ¿Cuál es el percentil 80% de la profundidad de los sismos con magnitud de al menos 4 grados?

Respuestas:

(1) 170, (2) 0.3365854, (3) 17, (4) 3.6, (5) 157.38, (6) 217.1.