



TRABAJO PRACTICO INTEGRADOR

PROBABILIDAD Y ESTADISTICA

ALUMNO: FABIAN IGNACIO CARDOZO

TRABAJO PRÁCTICO INTEGRADOR

Consignas a resolver primera entrega

1. Haciendo uso del vocabulario técnico, identificar con relación al caso propuesto.
 - a. Población y Muestra
 - b. Unidad Elemental
 - c. Variables en estudio, detallando su tipo.

a. Población y muestra:

En el contexto del caso planteado, se entiende por **Población** al conjunto total de estudiantes que cursan el segundo año de la Tecnicatura Universitaria en Programación en la Universidad INNOVA XXII; de dicho grupo se desea obtener conclusiones generales.

Por otro lado, la **Muestra** corresponde al subconjunto de esa población, ósea a los alumnos que participaron en la encuesta realizada por la consultora STUDIO X. Este subconjunto fue seleccionado para relevar información específica y, mediante técnicas estadísticas, permitirá inferir resultados válidos para la población en su totalidad.

b. Unidad elemental:

La unidad elemental es el objeto individual de estudio a partir del cual se recolectan los datos. En este caso, cada estudiante encuestado del segundo año de la carrera constituye una unidad elemental, ya que sobre él se registran las características personales, académicas y socioeconómicas incluidas en el cuestionario.

c. Variables en estudio y su tipología:

El cuestionario aplicado permite identificar diversas variables estadísticas, cada una con naturaleza y escala de medición específicas. A continuación, se detallan junto con su clasificación técnica:

- Edad (en años cumplidos): variable cuantitativa discreta (numérica entera).
- Género (Femenino, Masculino, Otro): variable cualitativa nominal (categórica sin orden).
- Peso (kg): variable cuantitativa continua (magnitud con posibles valores en un rango real).
- Estatura (cm): variable cuantitativa continua.
- Número de hermanos: variable cuantitativa discreta.

- Condición de fumador (Sí/No): variable cualitativa nominal dicotómica.
- Horas de estudio semanal en el hogar: variable cuantitativa continua.
- Condición laboral (Trabaja: Sí/No): variable cualitativa nominal dicotómica.
- Gastos semanales en alimentación (pesos): variable cuantitativa continua.
- Cantidad de materias aprobadas: variable cuantitativa discreta.
- Nivel de satisfacción con la carrera (Muy satisfecho, Satisfecho, Insatisfecho, Muy insatisfecho): variable cualitativa ordinal (posee categorías con un orden implícito).

2. Construir la/s Tabla/s de Frecuencias y calcular todas las frecuencias de las siguientes variables:

- a. Tiempo en horas semanales dedicadas al estudio. (Determinar la cantidad optima de intervalos a utilizar)

```
CARDOZO FABIAN IGNACIO - P1y2.R* x
library(readxl)
# a). Vamos a determinar la cantidad optima de intervalos a utilizar
# Empezamos guardando archivo Excel en una Variable y Leyendola
Archivo <- file.choose()
datos <- read_excel(Archivo)
# En una variable guardamos los datos extraídos del vector "TIEMPO SEMANAL en HS. DEDIC. EST"
tiempo <- datos$`TIEMPO SEMANAL en HS. DEDIC. EST.`
# Usamos la fórmula de Sturges
n <- length(tiempo)
k <- ceiling(1 + 3.322*log10(n))
k
```

c. A partir de la tabla obtenida en el punto **a.** realizar la interpretación de todas las frecuencias correspondientes al cuarto intervalo en el contexto del caso planteado.

```
8 # c).Determinaremos el rango de valores de la variable "Tiempo de estudio semanal"
9 rango <- range(tiempo)
10
11 # Calculamos la amplitud (ancho de clase) utilizando la fórmula:
12 # (máximo - mínimo) / k , donde k es el número de intervalos según Sturges
13 # Usamos ceiling() para redondear hacia arriba y asegurar que todos los datos queden cubiertos
14 amplitud <- ceiling((rango[2] - rango[1]) / k)
15
16 # Visualizamos el rango y la amplitud obtenida
17 rango
18 amplitud
19
20 # Construimos los puntos de corte (breaks) para los intervalos de clase
21 # seq() genera una secuencia desde el valor mínimo redondeado hacia abajo (floor)
22 # hasta el máximo redondeado hacia arriba (ceiling) + amplitud,
23 # con un paso igual a la amplitud calculada
24 breaks <- seq(floor(rango[1]), ceiling(rango[2]) + amplitud, by= amplitud)
25 # Clasificamos los datos en intervalos utilizando cut()
26 # right = FALSE indica que los intervalos son de la forma [a, b)
27 # es decir, incluyen el límite inferior y excluyen el superior (excepto el último)
28 clases <- cut(tiempo, breaks = breaks, right = FALSE)
29 # Visualizamos las primeras clasificaciones de intervalos para verificar el resultado
30 head(clases)
31
32 # Calculamos la tabla de frecuencias absolutas (fa) para cada intervalo
33 Tabla_Tiempo <- table(clases)
34 # Calculamos la frecuencia absoluta acumulada (faa)
35 f_acum <- cumsum(Tabla_Tiempo)
36 # Calculamos la frecuencia relativa (fr) como proporción sobre el total
37 f_rel <- prop.table(Tabla_Tiempo)
38 # Calculamos la frecuencia relativa acumulada (fra)
39 f_rel_acum <- cumsum(f_rel)
40
41 # Construimos un Data Frame que consolida toda la información en una tabla de frecuencias
42 Tiempo_Semanal <- data.frame(
43   Tiempo_Estad = levels(clases),           # Intervalos de clase
44   Frecuencia = as.vector(Tabla_Tiempo),    # Frecuencia absoluta
45   Frec_Acum = as.vector(f_acum),           # Frecuencia absoluta acumulada
46   Frec_Relativa = round(as.vector(f_rel), 4), # Frecuencia relativa
47   Frec_Rel_Acum = round(as.vector(f_rel_acum), 4) # Frecuencia relativa acumulada
48 )
49 Tiempo_Semanal[4, ]
```

b. Nivel de satisfacción con la Carrera.

d. A partir de la tabla obtenida en el punto **b.** realizar la interpretación de todas las frecuencias correspondientes a la categoría “Satisfecho”

```
62 # d). Transformamos la variable "SATISFACCIÓN CON LA CARRERA" en un factor con etiquetas descriptivas
63 # levels = define los códigos originales (1, 2, 3, 4)
64 # labels = asigna el nombre real de cada categoría (Muy satisfecho, Satisfecho, etc.)
65 satisfaccion <- factor(datos$`SATISFACCIÓN CON LA CARRERA`,
66                       levels = c(1, 2, 3, 4),
67                       labels = c("Muy satisfecho", "Satisfecho", "Insatisfecho", "Muy insatisfecho"))
68
69
70 # b). Calculamos la tabla de frecuencias absolutas (fa) para cada categoría de satisfacción
71 Tabla_satisfaccion <- table(satisfaccion)
72 # Calculamos la frecuencia absoluta acumulada (faa) sumando las frecuencias categoría por categoría
73 frec_acum <- cumsum(Tabla_satisfaccion)
74 frec_rel <- prop.table(Tabla_satisfaccion)
75 frec_rel_acum <- cumsum(frec_rel)
76
77 # Construimos un Data Frame que consolida toda la información en una tabla de frecuencias
78 Tabla_final <- data.frame(
79   Satisfaccion= names(Tabla_satisfaccion),
80   Frecue = as.vector(Tabla_satisfaccion),
81   Frecu_Acum = as.vector(frec_acum),
82   Frecu_Rela = round(as.vector(frec_rel), 4),
83   Frecu_Rela_Acum = round(as.vector(frec_rel_acum), 4)
84 )
85
86
87 # Mostramos la tabla de frecuencias final
88 Tabla_final
89
```

Link Repositorio de GitHub: <https://github.com/fabian24cf/Repositorio-Estad-stica.git>