

Práctica Calificada 1

Profesores del curso

05/09/2023

Indicaciones Generales:

- Esta práctica debe ser desarrollada por un grupo de máximo 3 alumnos del mismo horario.
- Justifique claramente sus soluciones, presente los códigos y resultados obtenidos en R, interprete estos resultados y escriba respuestas completas teniendo en cuenta el contexto de los ejercicios propuestos; todo esto influirá en la calificación.
- La presentación, la ortografía y la gramática de los trabajos influirán en la calificación.
- La solución debe ser subida en único archivo PDF en la tarea PC1 de su correspondiente horario del curso. La primera hoja debe presentar el código PUCP, nombre completo y correo electrónico de todos los integrantes del grupo.

Cuestionario

La base de datos `Censo_Universitario.Rdata` contiene las siguientes variables medidas en 227540 alumnos universitarios

Variable	Descripción
universidad	Universidad
facultad	Facultad
especialidad	Especialidad
sexo	Sexo
	Hombre
	Mujer
pre	Tipo de preparación para la universidad
	Por su cuenta
	Profesor particular
	Academia
	Centro Pre-Universitario
modalidad	Modalidad de ingreso a la universidad
	Examen de admisión
	Primeros puestos
	Convenio universitario
	Centro Pre-universitario
	Traslado externo
	Graduado o titulado
	Disposiciones especiales
	Otra
trabajo	Trabaja actualmente?
	Si
	No
tipo	Tipo de universidad
	Pública
	Privada

Lea este conjunto de datos mediante la función `load()`,

```
load("Censo_Universitario.Rdata")
```

esto generará un objeto `d` que contendrá los datos.

Pregunta 1 (3.0 puntos)

Se desea comparar la distribución de frecuencias relativas del tipo de preparación para la universidad (`pre`) entre alumnos de universidades públicas (`tipo=="Pública"`) y privadas (`tipo=="Privada"`). Con este fin construya la tabla de distribución de frecuencias para cada grupo y represéntela gráficamente. Presente un análisis comparativo de sus resultados.

Pregunta 2 (3.0 puntos)

Elabore un diagrama de cajas (boxplot) para comparar el tiempo de desplazamiento de casa a la universidad, entre los alumnos de universidades públicas y privadas. A partir de sus resultados, realice una comparación en términos de tendencia central, dispersión y asimetría.

Pregunta 3 (3.0 puntos)

Para un nuevo estudio se van a considerar solamente la información de mujeres universitarias cuyas horas de trabajo por semana hayan estado dentro del 25 % más altas de su grupo, y de hombres universitarios cuyas horas de trabajo por semana hayan estado dentro del 20 % más altas de su grupo.

¿Qué número de horas de trabajo por semana debe superar como mínimo una mujer para que sea incluida en el estudio? ¿Qué número de horas de trabajo por semana debe superar como mínimo un hombre para que sea incluida en el estudio? Justifique sus respuestas.

Pregunta 4 (6.0 puntos)

Para la variable ingreso monetario mensual del hogar considere que

- **Hasta** 500 corresponde a un ingreso en el intervalo $[0, 500]$.
- **Más de** 20000 corresponde a un ingreso en el intervalo $(20000, 25000]$.
- En los otros casos, asuma que es un intervalo abierto a la izquierda y cerrado a la derecha, por ejemplo, para **De** 501 a 1000 corresponde a un ingreso en el intervalo $(500, 1000]$.

- a) (2.0 puntos) Construya un polígono de frecuencias para la distribución de los ingresos mensuales de los hogares de alumnos de universidades públicas y otro sobre el mismo eje x para los alumnos de universidades privadas. Interprete sus resultados. **Sugerencia:** en R para que un segundo gráfico se muestre sobre un primer gráfico elaborado previamente: utilice la función `plot` para crear el primer gráfico y luego use la función `lines` para crear el segundo.
- b) (2.0 puntos) Calcule las estadísticas necesarias y realice una comparación en términos de tendencia central, dispersión y asimetría del ingreso mensual de los hogares de alumnos por el tipo de universidad a la que asisten.
- c) (2.0 puntos) Algunos autores han propuesto la siguiente medida de asimetría robusta ante valores atípicos para una distribución

$$A_0 = \frac{q_{0,875} - 2q_{0,5} + q_{0,125}}{q_{0,875} - q_{0,125}},$$

donde q_p denota el cuantíl $0 < p < 1$ de la distribución. Calcule e interprete esta medida para la variable ingreso monetario mensual del hogar para los alumnos de universidades públicas y privadas.

Pregunta 5 (5.0 puntos)

Evalúe la veracidad o falsedad de cada una de las siguientes afirmaciones. Justifique su respuesta (presente sus códigos y resultados en R como parte de su justificación)

- a) (1.0 punto) La variable tiempo en minutos de desplazamiento de casa a la universidad presenta una menor dispersión que el gasto anual en matrícula.
- b) (1.0 punto) La mediana del tipo de preparación para la universidad es **Academia**.
- c) (1.0 punto) El tercer cuartil del número de horas de trabajo a la semana es la misma para hombres y mujeres.
- d) (1.0 punto) Se aplicó la función **summary** a la variable número de horas de trabajo a la semana y a partir de los resultados obtenidos, se llegó a la conclusión que aproximadamente el 50 % de todos los alumnos trabaja más de 36 horas.
- e) (1.0 punto) Para representar la tendencia central de la variable gasto anual en matrícula es adecuado usar la mediana.