

Departamento de Estadística y Matemáticas
Facultad de Ciencias Económicas
Estadística II
Parcial III

Nombre: _____ Cédula: _____

Antes de iniciar con el desarrollo del trabajo **LEA ATENTAMENTE:**

- La calificación de cada punto desarrollado estará conformado por tres partes:
 1. El planteamiento del ejercicio, que abarca la complejidad en la creación del enunciado, el correcto planteamiento del juego de hipótesis, y/o intervalo de confianza, la correcta evaluación de las condiciones que conducen al empleo de una u otra ecuación.
 2. El desarrollo del ejercicio planteado.
 3. Las interpretaciones resultantes en el contexto de los datos y del enunciado planteado.
- Si se presenta una situación en donde al momento de la calificación, surge alguna duda o inconsistencia con los resultados debido a similitud con algún compañero, el docente solicitará sustentación por parte de los estudiantes que desarrollaron el punto, y **hasta que no ocurra la sustentación tendrán un 0 en ese punto.** >:c

1. **(5 puntos)** El Ministerio de Salud de Colombia ha decidido realizar un estudio epidemiológico a nivel nacional para evaluar la prevalencia de enfermedades cardiovasculares en la población, y por ello, ha decidido recopilar la información de una muestra representativa personas de diferentes regiones, edades, género y condiciones socioeconómicas.

La base de datos contiene los siguientes campos:

- **Region:** Región del paciente, donde 'Andina', 'Caribe', 'Pacífica', 'Orinoquía' y 'Amazonía', representan la región de Colombia donde vive el paciente.
- **Edad:** Edad del paciente en años.
- **Altura:** Altura del paciente en centímetros.
- **Peso:** Peso del paciente en kilogramos.
- **Genero:** Género del paciente, donde 'Hombre' representa un paciente de género masculino y 'Mujer' representa un paciente de género femenino.
- **Niv_Soc:** Nivel socioeconómico del paciente, donde 'Bajo', 'Medio' y 'Alto' son los niveles bajo los cuales se cataloga la capacidad adquisitiva y estado de la vivienda donde vive el paciente.
- **Pre_S:** Presión arterial sistólica del paciente en mmHg.
- **Pre_D:** Presión arterial diastólica del paciente en mmHg.
- **Col:** Nivel de colesterol del paciente, donde 'normal' representa un nivel normal de colesterol, 'por encima de lo normal' representa un nivel de colesterol por encima de lo normal y 'muy por encima de lo normal' representa un nivel de colesterol muy por encima de lo normal.
- **Col_sangre:** Nivel de colesterol del paciente en la sangre en mg/dL.
- **Glu:** Nivel de glucosa del paciente, donde 'normal' representa un nivel normal de glucosa, 'por encima de lo normal' representa un nivel de glucosa por encima de lo normal y 'muy por encima de lo normal' representa un nivel de glucosa muy por encima de lo normal.
- **Az_sangre:** Nivel de glucosa del paciente en forma numérica en mg/dL.
- **Fumador:** Si el paciente es fumador o no, donde 'No' representa que el paciente no es fumador y 'Sí' representa que el paciente es fumador.
- **Con_Alc:** Si el paciente consume alcohol o no, donde 'No' representa que el paciente no consume alcohol y 'Sí' representa que el paciente consume alcohol.

- **Act_Fis:** Si el paciente realiza actividad física o no, donde 'No' representa que el paciente no realiza actividad física y 'Sí' representa que el paciente realiza actividad física.
- **Enf_Car:** Si el paciente tiene enfermedad cardiovascular o no, donde 'No' representa que el paciente no tiene enfermedad cardiovascular y 'Sí' representa que el paciente tiene enfermedad cardiovascular.
- **IMC:** Índice de masa corporal del paciente, calculado como el peso en kilogramos dividido por el cuadrado de la altura en metros.
- **Niv_Estres:** Nivel de estrés del paciente, donde 'Bajo' representa un nivel bajo de estrés, 'Moderado' representa un nivel moderado de estrés y 'Alto' representa un nivel alto de estrés.
- **Con_Tabaco:** Frecuencia de consumo de tabaco del paciente, donde 'Nunca' representa que el paciente nunca consume tabaco, 'Ocasional' representa que el paciente consume tabaco ocasionalmente y 'Diario' representa que el paciente consume tabaco a diario.
- **Dieta:** Tipo de dieta del paciente, donde 'Saludable' representa una dieta saludable y 'No saludable' representa una dieta no saludable.
- **Act_Social:** Nivel de actividad social del paciente, donde 'Baja' representa un nivel bajo de actividad social, 'Moderada' representa un nivel moderado de actividad social y 'Alta' representa un nivel alto de actividad social.
- **Niv_Act:** Nivel de actividad física del paciente, donde 'Bajo' representa un nivel bajo de actividad física, 'Moderado' representa un nivel moderado de actividad física y 'Alto' representa un nivel alto de actividad física.
- **Con_Grasas:** Nivel de consumo de grasas del paciente, donde 'Bajo' representa un nivel bajo de consumo de grasas, 'Moderado' representa un nivel moderado de consumo de grasas y 'Alto' representa un nivel alto de consumo de grasas.
- **Con_Sal:** Nivel de consumo de sal del paciente, donde 'Bajo' representa un nivel bajo de consumo de sal, 'Moderado' representa un nivel moderado de consumo de sal y 'Alto' representa un nivel alto de consumo de sal.
- **Con_Azucar:** Nivel de consumo de azúcar del paciente, donde 'Bajo' representa un nivel bajo de consumo de azúcar, 'Moderado' representa un nivel moderado de consumo de azúcar y 'Alto' representa un nivel alto de consumo de azúcar.
- **Frec_Card:** Frecuencia cardíaca del paciente en latidos por minuto.
- **Con_FV:** Consumo de frutas y verduras del paciente, donde 5 representa un consumo alto, 3 representa un consumo moderado y 1 representa un consumo bajo.
- **Horas_Sueño:** Horas de sueño del paciente al día.
- **Con_Agua:** Consumo de agua del paciente en litros.

El objetivo del estudio es analizar los factores de riesgo asociados con las enfermedades cardiovasculares, como la edad, el género, el estilo de vida (fumador, obesidad), los niveles de presión arterial, colesterol y diabetes. Además, se busca evaluar la prevalencia de estas enfermedades en diferentes regiones y grupos socioeconómicos.

Basado en la información contenida en la base de datos suministrada,

NOTA: DEBE APLICAR AL MENOS UN FILTRO EN CADA PUNTO >:c

- a) **(1 punto)** Plantee un intervalo de confianza del 99 % para la diferencia de dos proporciones.
- b) **(1 punto)** Plantee una prueba de hipótesis para diferencia de medias empleando un nivel de significancia del 4 %.
- c) **(0.5 puntos)** Plantee una prueba de hipótesis para una media empleando un nivel de significancia del 1 % .
- d) **(0.5 puntos)** Basado en el inciso c), plantee una hipótesis alternativa específica para calcular el Error Tipo II.
- e) **(1 punto)** Plantee una prueba de independencia empleando un nivel de significancia del 5 %.

- f)* **(1 punto)** Plantee un intervalo de confianza del 95 % para razón de varianzas, en donde el objetivo es saber si hay o no diferencias significativas entre las varianzas.