****

**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**

**FACULTAD DE CIENCIAS SOCIALES Y HUMANÍSTICAS**

**EXAMEN DE MEJORAMIENTO DE ECONOMÍA ESTADÍSTICA COMPUTARIZADA**

APELLIDOS: NOMBRES:

MATRICULA: PARALELO:

**COMPROMISO DE HONOR**

Yo, \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, al firmar este compromiso, reconozco que el presente examen está diseñado para ser resuelto de manera individual, que puedo usar una calculadora ordinaria para cálculos aritméticos, un lápiz o esferográfico; que solo puedo comunicarme con la persona responsable de la recepción del examen; y, cualquier instrumento de comunicación que hubiere traído, debo apagarlo y depositarlo en la parte anterior del aula, junto con algún otro material que se encuentre acompañándolo. No debo además, consultar libros, notas, ni apuntes adicionales a las que se entreguen en esta evaluación. Los temas debo desarrollarlos de manera ordenada. Firmo al pie del presente compromiso, como constancia de haber leído y aceptar la declaración anterior.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
*Firma de compromiso del estudiante*

**Instrucción general**

* Todos los enunciados presentan una sola respuesta correcta
* Las respuestas serán revisadas en la tabla final de respuestas

**ESTADÍSTICA DESCRIPTIVA, DATOS Y MUESTREO (14 puntos)***Resultados de aprendizaje e, i*

1. ¿Cuál de las siguientes medidas corresponde a un momento estándar?

a) Error estándar b) Desviación estándar c) Coeficiente de correlación d) Curtosis

1. La regla de \_\_\_\_\_\_\_\_\_\_\_\_\_\_ se aplica para calcular el tiempo en que se duplicará un valor con una tasa de crecimiento fluctuante

a) Logaritmo b) 70 c) L’Hopital d) No existe esta regla

1. Una serie de 5 datos (a, b, 2a+2b, b, a) presentará un sesgo \_\_\_\_\_\_\_\_ y una curtosis \_\_\_\_\_\_\_\_

 a) 0, positiva b) positivo, positiva c) 0, negativa d) Depende del los valores de a y b

1. Ante una serie de datos de facturación numerados, el mejor tipo de muestreo sería

a) Aleatorio simple b) Sistemático c) Estratificado d) Serial

1. Para realizar un muestreo en zonas geográficas delimitadas, el mejor tipo de muestreo sería

a) Aleatorio simple b) Demográfico c) Conglomerado d) Censo

1. Un coeficiente de correlación de cero implica con certeza

a) Relación no lineal b) Covarianza de cero c) Independencia lineal d) Ninguna de las anteriores

1. ¿Qué características tiene la matriz de correlación arrojada como resultado en Excel?

a) Triangular inferior b) Diagonal principal presenta unos c) Traza igual al número N de variables d) Todas las anteriores

**REGRESIÓN LINEAL Y ANOVA (28 puntos)***Resultados de aprendizaje e, i*

1. ¿Cuál de los siguientes supuestos no es crucial si vas a estimar una regresión lineal para una muestra de 100 observaciones?

a) Homocedasticidad b) No multicolinealidad c) X no es aleatoria d) Residuos normales

1. Un modelo de regresión lineal es estimado para una muestra de 5 datos. Si 4 de estos están en la línea de regresión, ¿dónde se encuentra el quinto?

a) Por arriba de la línea b) Por debajo de la línea c) En la línea d) Falta información

1. Para un modelo de regresión lineal simple con 100 observaciones, la sumatoria cuadrática de los residuos RSS es 200. ¿Cuánto es la varianza del residuo?

a) 2.00 b) 2.02 c) 1.42 d) 1.72

1. Para un modelo de regresión lineal estimado con 100 muestras, ¿cuál es el valor F crítico para contrastar la ANOVA asociada a esta regresión si α=0.05?

a) 2.4664 b) 2.6993 c) 3.0892 d) 3.4801

1. Para un modelo de regresión lineal con dos variables independientes y una muestra de 100, la sumatoria cuadrática del modelo ESS es 11 y la sumatoria cuadrática de los residuos RSS es 196. ¿Cuánto es el valor P del F calculado de esta regresión?

a) 0.3513 b) 0.3039 c) 0.1495 d) 0.0707

1. Para un modelo de regresión con 3 variables independientes y una muestra de 30 observaciones. Se obtiene que el coeficiente de determinación es 0.80. Determine el F calculado de la ANOVA.

a) 25.00 b) 34.67 c) 36.00 d) 38.45

1. Para un modelo de regresión lineal simple y una muestra de 30 observaciones, ¿cuál es el valor t crítico para contrastar la significancia individual de los parámetros?

a) 2.0484 b) 2.3595 c) 2.3638 d) 2.3684

1. Para un modelo de regresión lineal simple con 30 observaciones, si se conoce que el F calculado = 120. ¿Cuánto es el t calculado asociado al beta estimado de la regresión?

a) 9.95 b) 10.95 c) 11.95 d) 12.95

1. Para un modelo de regresión lineal siendo la variable dicotómica 1 para mujer y 0 para hombre. Si la muestra es de 30 observaciones, y el error estándar de ese coeficiente es 2, ¿existe un punto de corte diferencial significativo en el modelo por el atributo de ser mujer?

a) Sí b) No c) Falta información

1. Para un modelo de regresión lineal siendo la variable dicotómica 1 para mujer y 0 para hombre. Si la muestra es de 30 observaciones y el error estándar de ese coeficiente es 1.5, ¿existe un punto de corte significativo en el modelo?

a) Sí b) No c) Falta información

1. Para un modelo de regresión lineal siendo la variable dicotómica 1 para mujer y 0 para hombre. Si la muestra es de 30 observaciones y el error estándar de ese coeficiente es 2.2, ¿existe una pendiente diferencial significativa en el modelo?

a) Sí b) No c) Falta información

1. Para una regresión lineal simple con nivel de significancia al 5%, el estadístico Chi cuadrado de Breusch-Pagan presenta un valor P asociado de 0.07563. Se concluye que el modelo presenta:

a) Normalidad residual b) Falta de variables c) Homocedasticidad d) Ninguna de las anteriores

1. ¿Cuál de las siguientes es una prueba de normalidad?

a) Cook-Weisberg b) Normalex c) Jarque-Bera d) Durbin-Watson

1. En un modelo de regresión lineal simple, si la inflación de la varianza VIF es mayor a 10. Se concluye la presencia de:

a) Multicolinealidad b) Autocorrelación c) Normalidad residual d) Ninguna de las anteriores

**REGRESIÓN LOGÍSTICA (28 puntos)***Resultados de aprendizaje e, i*

1. ¿Cuál de los siguientes no es un problema del modelo lineal de probabilidad?

a) Probabilidad estimada negativa b) Probabilidad estimada mayor que 1 c) Acotada por un solo lado d) Probabilidad marginal constante

1. Si los coeficientes del componente lineal Z de la regresión logit y las variables independientes son todos positivos. La probabilidad es

a) Mayor que 0.5 b) Menor que 0.5 c) 0.5 d) Indeterminado

1. Los coeficientes en el modelo logit al ser estimados por máxima verosimilitud aseguran cumplir con las características MELI, mejor estimador lineal insesgado.

a) Verdadero b) Falso c) Falta información

1. La regresión logística está estimada como , donde Y denota la probabilidad de tener una casa (Y=1) o no tenerla (Y=0), X denota el ingreso familiar en miles de dólares, and L es la función de la distribución logística acumulada. Busque la probabilidad de tener una casa para las familias con ingreso de $7000 (X=7).

a) 0.562 b) 0.572 c) 0.582 d) 0.592

1. La regresión probit está estimada como , donde Y denota la probabilidad de tener una casa (Y=1) o no tenerla (Y=0), X denota el ingreso familiar en miles de dólares, and L es la función de la distribución normal estándar acumulada. Busque la probabilidad de tener una casa para las familias con ingreso de $7000 (X=7).

a) 0.329 b) 0.429 c) 0.529 d) 0.629

1. Aplicando una regresión logística, si el logaritmo de verosimilitud en la primera iteración es -100 y la iteración final es -75. ¿Cuánto es el ratio de verosimilitud Chi-cuadrado de la regresión?

a) 0.75 b) 5 c) 25 d) 50

1. ¿Cuál de los siguientes Pseudo-R² utilizaría para comparar modelos logit?

a) Cox-Snell b) Cuenta c) McFadden d) Nagelkerke

1. Agregar *,nolog* al comando *logit* en STATA oculta

a) Iteraciones b) Gráfica c) Intervalos de confianza d) Pseudo- R² excepto el de McFadden

1. La curva ROC compara la sensibilidad del modelo frente a

a) Valor predictivo positivo b) Valor predictivo negativo c) Tasa de falsos positivos d) Especificidad

1. En un modelo de regresión logística de pacientes contagiados del Chikungunya, si la tasa de falsos positivos es 16%, y el total de no contagiados de la muestra es 200 de 1000. ¿Cuánto es el error tipo I del experimento?

a) 0,02 b) 0,032 c) 0,04 d) 0,05

1. En los criterios de información bayesianos y de Akaike se busca \_\_\_\_\_\_\_\_ los valores

a) Maximizar b) Minimizar c) Igual a cero d) Igual a 0,05

1. ¿Qué significa el acrónimo ROC en la curva ROC en referencia a los modelos logit?

a) *Responding Over Credibility* b) *Receiver Operating Characteristic* c) *Relationship On Criteria* d) *Reliability On Credibility*

1. En un modelo de regresión logística, con dos variables independientes tal que una es continua y la segunda es categórica con tres categorías. El ratio de verosimilitud es 7. ¿Cuál es su conclusión con respecto a los betas estimados del modelo?

a) Existe al menos uno significativo b) No significativos c) No hay suficiente información

1. En un modelo de regresión logística, con una sola variable independiente. El coeficiente beta estimado presenta un intervalo de confianza (0.75, 3.4). ¿Cuál es su conclusión con respecto al beta estimado del modelo?

a) Significativo b) No significativo c) No hay suficiente información

**ANÁLISIS FACTORIAL Y DE COMPONENTE PRINCIPAL (10 puntos)***Resultados de aprendizaje e, i*

1. En un análisis factorial con 2 factores comunes retenidos y 4 ítems, los factor loadings son (0.36, 0.7), (0.9, 0.58), (0.30, 0.65), (0.25, 0.55) respectivamente. El investigador debería

a) Eliminar ítem 4 b) Aplicar KMO c) Rotar ortogonalmente d) Comparar comunalidades

1. En un análisis factorial Varimax con 2 factores comunes retenidos y 4 ítems, los factor loadings son (0.38, 0.65), (0.9, 0.4), (0.1, 0.65), (0.25, 0.55) respectivamente. El investigador debería

a) Eliminar ítem 1 b) Eliminar ítem 3 c) Eliminar ítem 4 d) Comparar comunalidades

1. En un análisis factorial Varimax con 2 factores comunes retenidos y 4 ítems, los factor loadings son (0.38, 0.65), (0.9, 0.4), (0.1, 0.65), (0.25, 0.75) respectivamente. El investigador debería

a) Eliminar ítem 1 b) Eliminar ítem 3 c) Eliminar ítem 4 d) Comparar comunalidades

1. En un análisis factorial Varimax con 2 factores comunes retenidos y 4 ítems, los factor loadings son (0.38, 0.65), (0.9, 0.4), (0.42, 0.65), (0.25, 0.75) respectivamente. El investigador debería

a) Eliminar ítem 1 b) Eliminar ítem 3 c) Eliminar ítem 4 d) Comparar comunalidades

1. En un análisis factorial Varimax con 2 factores comunes retenidos y 4 ítems, los factor loadings son (0.95, 0.1), (0.85, 0.15), (0.75, 0.05), (0.65, 0.20) respectivamente. El investigador debería

a) Aplicar Promax b) Revisar el Cronbach alfa c) Revisar la varianza explicada d) Obtener Factor Score para cada factor

**SIMULACIONES (20 puntos)***Resultados de aprendizaje e, i*

*Al acabar el examen, el profesor pasará revisando este literal. Espere pacientemente a que pase por su equipo. Utilizando el método de Montecarlo para 1000 números aleatorios, demuestre que la integral de una función cúbica entre 0 y 1 es 0.25,*

*Tabla de respuestas*

*Utilice mayúsculas para escribir sus respuestas: ABCD*

|  |  |  |  |
| --- | --- | --- | --- |
| 1 |  | 24 |  |
| 2 |  | 25 |  |
| 3 |  | 26 |  |
| 4 |  | 27 |  |
| 5 |  | 28 |  |
| 6 |  | 29 |  |
| 7 |  | 30 |  |
| 8 |  | 31 |  |
| 9 |  | 32 |  |
| 10 |  | 33 |  |
| 11 |  | 34 |  |
| 12 |  | 35 |  |
| 13 |  | 36 |  |
| 14 |  | 37 |  |
| 15 |  | 38 |  |
| 16 |  | 39 |  |
| 17 |  | 40 |  |
| 18 |  | **Nota Simulación** |
| 19 |  |  |
| 20 |  |
| 21 |  |
| 22 |  |
| 23 |  |