

# Examen Final 2024 PAOI

⚠ Esta es una vista previa de la versión publicada del examen

Comenzado: 4 de sep en 21:32

## Instrucciones del examen

El siguiente examen consta de tres temas.

TEMA 1. DESARROLLO (PUEDE SER EN EL ESPACIO EN BLANCO PROPORCIONADO EN EL EXAMEN EN LINEA O EN EL PAPEL PROPORCIONADO)

TEMA 2. DESARROLLO EN R (SUBIR ARCHIVO .R o .Rmd)

TEMA 3. DESARROLLO (EN PAPEL PROPORCIONADO Y DEBE RESPONDER LA CONCLUSION EN EL ESPACIO PROPORCIONADO EN EL EXAMEN EN LINEA)

Cuentan con 120 minutos. Distribuya bien su tiempo.

Para este examen pueden utilizar el material de clase (diapositivas, talleres). No se permite uso de internet fuera del material subido en Aula Virtual.

Es un examen individual, está prohibida la comunicación con sus compañeros o cualquier tercera persona durante el examen.

Cualquier acto de deshonestidad será informado a las autoridades competentes.

Al comenzar el examen, el estudiante acepta el siguiente compromiso:

### COMPROMISO DE HONOR

Reconozco que el presente examen está diseñado para ser resuelto de manera individual, y no se permite la ayuda de fuentes no autorizadas ni copiar.

#### Pregunta 1

30 pts

Considere los datos del Estudio de Panel de Dinámica de Ingresos para 1982. Las variables registradas en este estudio se enumeran a continuación:

EXP Años de experiencia laboral a tiempo completo

WKS Semanas trabajadas (en el último año)

OCC Ocupación de cuello azul (sí = 1, no = 0 (cuello blanco))

IND Trabaja en una industria manufacturera (sí = 1, no = 0)

SUR Reside en el Sur (sí = 1, no = 0)

SMSA Reside en un área estadística metropolitana estándar (sí = 1, no = 0)

MS Casado (sí = 1, no = 0)

FEM Mujer (sí = 1, no = 0)

UNIÓN El salario se establece mediante un contrato sindical (sí = 1, no = 0)

ED Años de educación

BLK Negro (sí = 1, no = 0)

SALARIO Salario (en cientos de dólares)

Luego de una selección de variables, el resumen de un modelo de regresión lineal estimado que estudia el efecto de los factores clave en el salario ganado se presenta a continuación.

```
##
## Call:
## lm(formula = WAGE ~ EXP + WKS + OCC + SOUTH + FEM + ED + BLK +
##     EXP:WKS + EXP:OCC + WKS:ED + OCC:ED + OCC:BLK, data = earn)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1186.0  -267.8   -26.3   196.6  3377.0
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  788.7873   956.2641   0.825  0.40979
## EXP           34.7672    17.0997   2.033  0.04249 *
## WKS          -18.8882    20.2352  -0.933  0.35098
## OCC           568.0019   253.0211   2.245  0.02515 *
## SOUTH       -110.1530    39.8239  -2.766  0.00586 **
## FEM         -420.8215    59.0629  -7.125 3.09e-12 ***
## ED           -39.0292    58.8374  -0.663  0.50738
## BLK         -328.0309   116.0881  -2.826  0.00488 **
## EXP:WKS      -0.5155     0.3623  -1.423  0.15537
## EXP:OCC      -6.0547     3.4037  -1.779  0.07578 .
## WKS:ED        2.7580     1.2561   2.196  0.02851 *
## OCC:ED       -50.2433    17.2158  -2.918  0.00365 **
## OCC:BLK      300.1448   145.9993   2.056  0.04025 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 431.5 on 582 degrees of freedom
## Multiple R-squared:  0.3532, Adjusted R-squared:  0.3399
## F-statistic: 26.49 on 12 and 582 DF,  p-value: < 2.2e-16
```

Utilice un nivel de significancia del 5%.

**1. Interprete la estimación del intercepto. (10 pts)**

**2. ¿La raza influye en el salario? ¿Depende de algo más? Indique claramente en cuánto se estima que aumente o disminuya el salario promedio de un trabajador negro con respecto a quien no lo es (tome en cuenta interacciones). (10 pts)**

**3. Proporcione la ecuación estimada para el valor esperado de la diferencia del salario de un trabajador de ocupación de cuello azul vs uno de cuello blanco Sustente. (10 pts)**

0 palabras 

## Pregunta 2

50 pts

Para esta pregunta se utilizará los mismos datos del Estudio de Panel de Dinámica de Ingresos para 1982 de la pregunta anterior.

Suponga que la estimación del modelo de la pregunta anterior se realizó en todo el set de datos de 595 observaciones y ahí mismo se realizó la selección de variables. Como usted sabe que esa práctica puede causar problemas para la inferencia, decide hacer su propio análisis para confirmar.

Los datos se encuentran en el archivo: [earn.csv](#).

- 1. Divida la base de datos en 3 partes para selección, validación y estimación final e inferencias, tomando en cuenta que tenga suficientes observaciones en cada set.. Fije las semillas para la división aleatoria, al menos una semilla debe ser su número de matrícula. (5 pts)**
- 2. En el set de selección, realice una exploración gráfica de la relación entre el salario y las variables explicativas y proponga un modelo lineal aditivo. (10 pts)**
- 3. En el set de selección, seleccione dos modelos por medio de dos búsquedas automáticas stepwise, una con el criterio AIC y la otra con el criterio BIC. Utilice el**

modelo del literal 2 como punto de partida. Su modelo máximo debe incluir todas las interacciones de dos variables. Su modelo mínimo es el modelo con el intercepto solo. (15 pts)

4. En el set de validación, compare los dos modelos con MSEP y escoja el mejor (10 pts)

5. Estime el modelo escogido en el literal 4 en el set de estimación. Asuma que los supuestos del modelo se cumplen. Proporcione un intervalo de predicción del 95% para un trabajador hombre, negro, casado, con 15 años de educación y 20 años de experiencia, de cuello blanco, que trabaja en la industria manufacturera, que trabajó 48 semanas el año pasado, cuyo salario se establece SIN contrato sindical y que reside al norte en un área estadística metropolitana estándar. Interprete el intervalo. (10 pts)

Cargar

Elegir un archivo

### Pregunta 3

20 pts

Usted va a emprender en una cafetería. Como quiere que sus productos sean de alta calidad, usted está buscando mucha información de como preparar el café ideal. Ha encontrado que existen 3 métodos para preparar espresso, los cuales son: Método 1=Bar Machine(BM), Método 2 =Hyper-Espresso Method (HIP), y Método 3= I-Espresso System (IT). Además leyó que existe un índice que mide la razón del volumen de la espuma para el volumen del líquido y que dicho índice debe ser al menos del 10% para que la preparación sea correcta. Para seleccionar el método que va a usar en su cafetería hizo un experimento donde preparó 9 espressos con los métodos 1 y 2 y 8 espressos con el método 3 y calculó el índice (%) en cada uno.

1. El siguiente resumen es de un modelo de regresión lineal donde la codificación de la variable categórica método es dummy.

```
##  
## Call:  
## lm(formula = foamIndx ~ method, data = espresso1)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -14.6200  -5.7312  -0.0306   5.0650  17.5687   
##
```

```
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  32.400      2.809  11.535 4.84e-11 ***
## method2     28.900      3.972   7.275 2.10e-07 ***
## method3      6.221      4.095   1.519  0.142
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.426 on 23 degrees of freedom
## Multiple R-squared:  0.7168, Adjusted R-squared:  0.6922
## F-statistic: 29.11 on 2 and 23 DF,  p-value: 5.002e-07
```

Presente la ecuación estimada de un modelo ANOVA de EFECTOS de un factor para el método 2. (10 pts)

$$\mu_{2j} = \mu + \alpha_2 + \epsilon_{2j}$$

2. Complete la tabla ANOVA y concluya en el contexto de la pregunta. (10 pts)

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## method           4133              5e-07 ***
## Residuals       1633
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

[Editor HTML](#)

**B** *I* U A ▾ A ▾ I<sub>x</sub> ≡ ≡ ≡ ≡ ≡ ×<sup>2</sup> ×<sub>2</sub> ≡ ≡ ≡  
        12pt ▾ Párrafo ▾ 

No guardado

Entregar examen