

II Jornadas Internacionales de Estadística Aplicada 5 y 6 de diciembre de 2019

Análisis del Rendimiento Académico en relación al IMC (Índice de Masa Corporal) y otras variables, utilizando las técnicas del AED (Análisis Exploratorio de Datos) y RLM (Regresión Lineal Múltiple)

Autores: Octavio Daniel Coro – Adelma Beatriz Grágeda

Institución: Facultad de Ingeniería, Universidad Nacional de Jujuy. San Salvador de Jujuy.

Datos de contacto: Mail: odcoro@fi.unju.edu.ar Celular: +54-388-4081142

RESUMEN.

En este trabajo mediante un modelo de regresión lineal múltiple se realiza el análisis del rendimiento académico de alumnos universitarios, variable dependiente, estudiando el comportamiento de las materias aprobadas del primer año, variables independientes, denominadas variables regresoras: tomando de datos muestrales el índice de masa corporal, las materias regularizadas y las materias rendidas del primer año. Para esto se utilizará la técnica de Análisis Exploratorio de Datos y la Regresión Lineal Múltiple utilizando el paquete de software estadístico y EXCEL.

Palabras Claves: Regresión Lineal Múltiple - Análisis Exploratorio de Datos – Rendimiento Académico – Índice de Masa Corporal.

INTRODUCCIÓN:

En la vida cotidiana y en los medios audiovisuales (internet, televisión, radio, etc) vemos como es motivo de bullying, grooming y de acosos diversos, el tema de la delgadez extrema o de obesidad en nuestra sociedad. Por ese motivo analizamos en este trabajo, tomando como población a los estudiantes de segundo año que cursan la asignatura Probabilidad y Estadística en la Facultad de Ingeniería de Jujuy, la influencia del estado físico, a través de la determinación del índice de masa corporal (IMC) de los estudiantes, y considerando también las variables materias regularizadas y materias rendidas; con el rendimiento académico asociado a la variable materias aprobadas.

Quedando para un posterior trabajo la inclusión de la variable Coeficiente Intelectual que seguramente también afecta en el rendimiento académico. Teníamos pensado incluir esta variable, que básicamente es el cociente entre la Edad Mental y la Edad Cronológica, pero por razones de tiempo no se pudo realizarlos, porque necesitaríamos de un test para obtener la edad mental del alumno con la ayuda de un psicólogo.

En este trabajo se intenta demostrar el nivel de relación entre las variables IMC; materias regularizadas y materias rendidas del primer año con la variable materias aprobadas del primer año.

Para realizar este trabajo utilizamos una base de datos obtenida de una encuesta que se realiza en la cátedra de Probabilidad y Estadística de la Facultad de Ingeniería de la UNJu, en una actividad de Evaluación por Competencias, que se realizó en el mes abril del corriente año. Para eso seleccionamos solamente las variables que necesitamos para este trabajo, que fueron mencionadas en el párrafo anterior.

METODOLOGIA

Este trabajo está organizado de la siguiente manera:

1. Aplicamos el Análisis Exploratorio de Datos mediante BoxPlot para realizar un análisis comparativo del IMC de los alumnos versus alumnas de la cátedra de Probabilidad y Estadística del corriente año. Para esto tomamos de la población mencionada, una muestra para la población masculina y otra muestra para la población femenina, aplicando la fórmula correspondiente.
2. Utilizamos las muestras mencionadas con el siguiente procedimiento:
 - a. Analizamos como influye cada una de las variables regresoras, en el comportamiento de la variable dependiente “Materias Aprobadas de Primer Año”, mediante la aplicación del método de Regresión Lineal Simple.
 - b. Realizamos el análisis completo y en conjunto de la variable dependiente con las 3 variables independientes (regresoras) empleando Regresión Lineal Múltiple para ambos géneros.
 - c. Calculamos la ecuación de Regresión Lineal Múltiple para la muestra masculina y para la muestra femenina.
3. Conclusiones: Realizamos un análisis final comparando entre los dos géneros tiene mejor rendimiento académico y también viendo la influencia de la variable IMC en forma particular

DESARROLLO

Índice de masa corporal

El índice de masa corporal (IMC) se calculó dividiendo el peso en kilogramos por el cuadrado de la estatura en metros, siendo sus unidades (kg/m^2).

La OMS define la insuficiencia ponderal como un IMC igual o inferior a 18,50 y el sobrepeso como un IMC igual o superior a 25.

| Clasificación del IMC | |
|------------------------|-------------------|
| Insuficiencia Ponderal | $< 18,50$ |
| Intervalo Normal | $18,50$ a $24,90$ |
| Sobrepeso | $\geq 25,00$ |
| Preobesidad | $25,00$ a $29,90$ |
| Obesidad | $\geq 30,00$ |

Los siguientes son datos de la población:

| DESC FEMENINO | |
|--------------------------|------------|
| Media | 58,1166667 |
| Error típico | 1,02487935 |
| Mediana | 56,5 |
| Moda | 58 |
| Desviación estándar | 9,05148932 |
| Varianza de la muestra | 81,9294589 |
| Curtosis | 2,7565172 |
| Coeficiente de asimetría | 1,25498321 |
| Rango | 52 |
| Mínimo | 43 |
| Máximo | 95 |
| Suma | 4533,1 |
| Cuenta | 78 |

| DESC MASCULINO | |
|--------------------------|------------|
| Media | 72,3875817 |
| Error típico | 1,24223952 |
| Mediana | 70 |
| Moda | 65 |
| Desviación estándar | 15,3656543 |
| Varianza de la muestra | 236,103332 |
| Curtosis | 13,1273801 |
| Coeficiente de asimetría | 2,51743917 |
| Rango | 125 |
| Mínimo | 50 |
| Máximo | 175 |
| Suma | 11075,3 |
| Cuenta | 153 |

Usamos el cuadro comparativo de la variable Peso de los alumnos, porque de estos datos tomaremos las muestras que calcularemos a continuación:

Tamaño de la muestra para Población conocida

Calculamos el tamaño de la muestra para las mujeres (femenino) y para los hombres (masculino) para, a partir de ahora trabajar con datos muestrales.

$$n = \frac{N \cdot z^2 \cdot \sigma^2}{(N - 1) \cdot e^2 + z^2 \cdot \sigma^2}$$

Población Femenina:

N=78

z=1,96 para 95% de confianza

Varianza=81,93

e=Error máximo=2 Kg para variable Peso

Tamaño Muestra Mujeres: 39,4222441: 40

Población Masculina:

N=153

Z=1,96 para 95% de confianza

Varianza=236,10

e=Error máximo=2 Kg para variable Peso

Tamaño Muestra Varones: 91,5980911: 92

| IMC | M Reg | M Rendidas | M Aprobadas |
|-------|-------|------------|-------------|
| 21,38 | 2 | 2 | 0 |
| 20,55 | 3 | 0 | 0 |
| 22,61 | 4 | 3 | 2 |
| 22,96 | 3 | 0 | 0 |
| 20,50 | 4 | 5 | 3 |
| 21,91 | 3 | 5 | 3 |
| 22,43 | 4 | 2 | 2 |

| IMC | M Reg | M Rendidas | M Aprobadas |
|-------|-------|------------|-------------|
| 32,77 | 5 | 3 | 2 |
| 24,22 | 2 | 0 | 2 |
| 27,10 | 3 | 2 | 3 |
| 20,90 | 0 | 0 | 0 |
| 29,38 | 2 | 5 | 3 |
| 25,56 | 6 | 2 | 2 |
| 24,77 | 6 | 5 | 4 |

| | | | |
|-------|---|---|---|
| 26,30 | 5 | 5 | 5 |
| 20,45 | 2 | 2 | 0 |
| 24,46 | 5 | 3 | 3 |
| 25,16 | 6 | 5 | 5 |
| 23,83 | 5 | 5 | 5 |
| 23,01 | 6 | 3 | 3 |
| 23,73 | 3 | 0 | 0 |

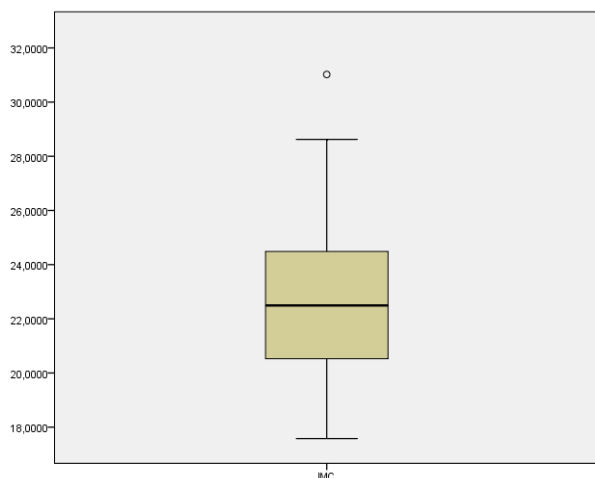
| | | | |
|-------|---|---|---|
| 25,51 | 3 | 7 | 5 |
| 23,67 | 4 | 2 | 3 |
| 28,36 | 6 | 3 | 3 |
| 24,09 | 2 | 1 | 1 |
| 21,83 | 4 | 3 | 2 |
| 22,99 | 3 | 6 | 2 |
| 21,93 | 4 | 3 | 0 |

Boxplots femenino y masculino: utilizando los datos de las respectivas muestras calculadas anteriormente.

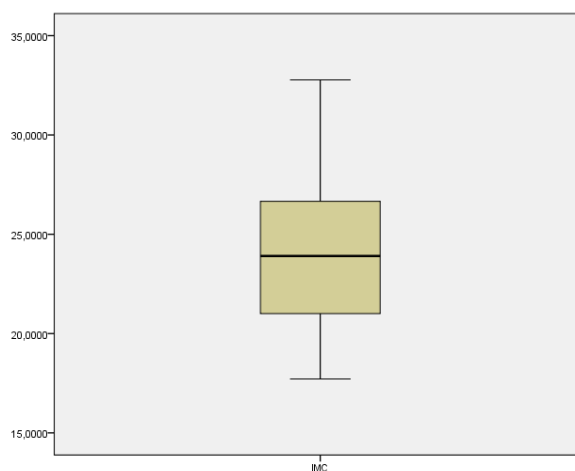
| <i>Femenino</i> | |
|-----------------------------------|------------|
| Media | 22,8363117 |
| Error típico | 0,44501439 |
| Mediana | 22,491756 |
| Moda | 20,4491374 |
| Desviación estándar | 2,81451813 |
| Varianza de la muestra | 7,92151229 |
| Curtosis | 0,72090877 |
| Coeficiente de asimetría | 0,71111435 |
| Rango | 13,4422832 |
| Mínimo | 17,578125 |
| Máximo | 31,0204082 |
| Suma | 913,45247 |
| Cuenta | 40 |
| Q1 | 20,525 |
| Q3 | 24,485 |
| DQ | 3,96 |
| Cotas Valores Alejados | |
| Q1-1,5 DQ | 14,585 |
| Q3+1,5 DQ | 30,425 |
| Cotas Valores Muy Alejados | |
| Q1-3 DQ | 8,645 |
| Q3+3 DQ | 36,365 |

| <i>Masculino</i> | |
|-----------------------------------|------------|
| Media | 24,0716666 |
| Error típico | 0,3688088 |
| Mediana | 23,9072112 |
| Moda | 27,6816609 |
| Desviación estándar | 3,53748974 |
| Varianza de la muestra | 12,5138336 |
| Curtosis | -0,5241625 |
| Coeficiente de asimetría | 0,38386448 |
| Rango | 15,0550933 |
| Mínimo | 17,7154195 |
| Máximo | 32,7705128 |
| Suma | 2214,59333 |
| Cuenta | 92 |
| Q1 | 21,005 |
| Q3 | 26,655 |
| DQ | 5,65 |
| Cotas Valores Alejados | |
| Q1-1,5 DQ | 12,53 |
| Q3+1,5 DQ | 35,13 |
| Cotas Valores Muy Alejados | |
| Q1-3 DQ | 4,055 |
| Q3+3 DQ | 43,605 |

Boxplot IMC Femenino



Boxplot IMC Masculino



Si analizamos los boxplot vemos que el masculino esta desplazado hacia arriba respecto al de las mujeres. Podemos observar también analizando los coeficientes de curtosis y asimetría que la distribución de los datos femeninos tiene a ser Leptocurtica con asimetría positiva bien acentuada, mientras que el de los varones tiende a ser Platicurtica con distribución asimétrica hacia la derecha pero menos marcada.

La media es mayor que la mediana. Dentro de la Caja los valores de los varones se encuentran más dispersos (esto se observa en el diferencial DQ) y presenta un IMC promedio mayor que el de las mujeres.

En cuanto a los valores Atípicos, teniendo en cuenta las cotas calculadas vemos que en el caso Femenino hay un valor Alejado de 31,02; mientras que en los hombres no observamos ningún valor Alejado, aunque al máximo valor del IMC de los hombres es de 32,77; pero se encuentra dentro de las cotas de Valores Alejados.

Correlaciones

| | | M Aprob. | IMC | M Reg. | M Rend. |
|------------------------|----------|----------|-------|--------|---------|
| Correlación de Pearson | M Aprob. | 1,000 | -,018 | ,467 | ,562 |
| | IMC | -,018 | 1,000 | ,239 | ,019 |
| | M Reg. | ,467 | ,239 | 1,000 | ,276 |
| | M Rend. | ,562 | ,019 | ,276 | 1,000 |

| Sig. (unilateral) | M Aprob. | | ,457 | ,001 | ,000 |
|-------------------|----------|------|------|------|------|
| | IMC | ,457 | | ,069 | ,453 |
| | M Reg. | ,001 | ,069 | | ,042 |
| | M Rend. | ,000 | ,453 | ,042 | |

| | | M Aprob. | IMC | M Reg. | M Rend. |
|------------------------|----------|----------|-------|--------|---------|
| Correlación de Pearson | M Aprob. | 1,000 | ,162 | ,553 | ,693 |
| | IMC | ,162 | 1,000 | ,048 | ,211 |
| | M Reg. | ,553 | ,048 | 1,000 | ,432 |
| | M Rend. | ,693 | ,211 | ,432 | 1,000 |

| Sig. (unilateral) | M Aprob. | . | ,061 | ,000 | ,000 |
|-------------------|----------|------|------|------|------|
| | IMC | ,061 | . | ,325 | ,022 |
| | M Reg. | ,000 | ,325 | . | ,000 |
| | M Rend. | ,000 | ,022 | ,000 | . |

| | | | | | |
|---|----------|----|----|----|----|
| Z | M Aprob. | 40 | 40 | 40 | 40 |
| | IMC | 40 | 40 | 40 | 40 |
| | M Reg. | 40 | 40 | 40 | 40 |
| | M Rend. | 40 | 40 | 40 | 40 |

| | | | | | |
|---|----------|----|----|----|----|
| Z | M Aprob. | 92 | 92 | 92 | 92 |
| | IMC | 92 | 92 | 92 | 92 |
| | M Reg. | 92 | 92 | 92 | 92 |
| | M Rend. | 92 | 92 | 92 | 92 |

Variables introducidas/eliminadas^b

| Modelo | Variables introducidas | Variables eliminadas | Método |
|--------|---------------------------------|----------------------|------------|
| 1 | M Rend, IMC, M Reg ^a | | Introducir |

| Modelo | Variables introducidas | Variables eliminadas | Método |
|--------|---------------------------------|----------------------|------------|
| 1 | M Rend, IMC, M Reg ^a | | Introducir |

- a. Todas las variables solicitadas introducidas.
b. Variable dependiente: M Aprobadas

Resumen del modelo^b

| Modelo | R | R ² | R ² corregida | Error típ. de la estimación | Durbin-Watson |
|--------------|-------------------|----------------|--------------------------|-----------------------------|---------------|
| dimension0 1 | ,658 ^a | ,433 | ,386 | 1,516 | 2,109 |

| Modelo | R | R ² | R ² corregida | Error típ. de la estimación | Durbin-Watson |
|--------------|-------------------|----------------|--------------------------|-----------------------------|---------------|
| dimension0 1 | ,749 ^a | ,561 | ,546 | 1,315 | 1,948 |

- a. Variables predictoras: (Constante), M Rendidas, IMC, M Regularizadas
b. Variable dependiente: M Aprobadas

ANOVA^b

| Modelo | Suma de cuadrados | gl | Media cuadrática | F | Sig. |
|-------------|-------------------|----|------------------|-------|-------------------|
| 1 Regresión | 63,264 | 3 | 21,088 | 9,179 | ,000 ^a |
| Residual | 82,711 | 36 | 2,298 | | |
| Total | 145,975 | 39 | | | |

| Modelo | Suma de cuadrados | gl | Media cuadrática | F | Sig. |
|-------------|-------------------|----|------------------|--------|-------------------|
| 1 Regresión | 194,203 | 3 | 64,734 | 37,410 | ,000 ^a |
| Residual | 152,276 | 88 | 1,730 | | |
| Total | 346,478 | 91 | | | |

- a. Variables predictoras: (Constante), M Rendidas, IMC, M Regularizadas
b. Variable dependiente: M Aprobadas

Tanto en cuadro Femenino como en el Masculino, el F-Test, muestra en la columna Sig el valor 0,000 en la fila de Regresión, este resultado indica que es un buen modelo para la Regresión Lineal Múltiple.

Es decir, al ser el valor de Sig =0,000, menor que 0,05 se concluye que el modelo es estadísticamente significativo y por tanto las variables independientes explican el comportamiento de la variable dependiente. Cuanta es la incidencia, se analizará más adelante con el coeficiente de determinación de los modelos.

Coeficientes^a

| Modelo | | Coeficientes no estandarizados | | Coeficientes tipificados | t | Sig. |
|--------|-------------|--------------------------------|------------|--------------------------|-------|------|
| | | B | Error tip. | Beta | | |
| 1 | (Constante) | ,472 | 2,009 | | ,235 | ,815 |
| | IMC | -,079 | ,089 | -,114 | -,883 | ,383 |
| | M Reg. | ,441 | ,162 | ,367 | 2,724 | ,010 |
| | M Rend. | ,431 | ,122 | ,463 | 3,541 | ,001 |

| Modelo | | Coeficientes no estandarizados | | Coeficientes tipificados | t | Sig. |
|--------|-------------|--------------------------------|------------|--------------------------|--------|------|
| | | B | Error tip. | Beta | | |
| 1 | (Constante) | -1,087 | ,991 | | -1,097 | ,276 |
| | IMC | ,017 | ,040 | ,031 | ,426 | ,671 |
| | M Reg. | ,352 | ,088 | ,313 | 3,994 | ,000 |
| | M Rend. | ,509 | ,074 | ,551 | 6,878 | ,000 |

a. Variable dependiente: M Aprobadas

Analizando en forma individual la influencia de las variables empleando el T-Test que se encuentra en la última columna, un valor de Sig menor que 0,05 indica que las variables influyen en forma particular y significativa sobre el modelo, mientras que si es mayor a 0,05, significa que las variables influyen con menor intensidad.

Ahora analizando el coeficiente Beta vemos que en las mujeres el valor de Beta = - 0,114 significa que a mayor masa corporal presentan menor rendimiento académico pero no tan marcado, mientras que en caso de los hombres un valor de Beta= 0,031, indica que prácticamente no tiene relevancia el IMC en el rendimiento académico.

Mientras que para materias rendidas y regularizadas para ambos casos los valores de Beta superan los 0,300 demostrando que ambas variables son las más determinantes en el modelo.

Diagnósticos de colinealidad^a

| Modelo | Dimensión | | Autovalores | Índice de condición | Proporciones de la varianza | | | |
|------------|------------|---|-------------|---------------------|-----------------------------|-----|--------|---------|
| | | | | | (Constante) | IMC | M Reg. | M Rend. |
| dimension0 | dimension1 | 1 | 3,666 | 1,000 | ,00 | ,00 | ,01 | ,02 |
| | | 2 | ,238 | 3,926 | ,01 | ,01 | ,02 | ,95 |
| | | 3 | ,089 | 6,432 | ,03 | ,02 | ,96 | ,02 |
| | | 4 | ,007 | 22,584 | ,97 | ,97 | ,01 | ,01 |

| Modelo | Dimensión | | Autovalores | Índice de condición | Proporciones de la varianza | | | |
|------------|------------|---|-------------|---------------------|-----------------------------|-----|--------|---------|
| | | | | | (Constante) | IMC | M Reg. | M Rend. |
| dimension0 | dimension1 | 1 | 3,653 | 1,000 | ,00 | ,00 | ,01 | ,02 |
| | | 2 | ,229 | 3,990 | ,01 | ,01 | ,00 | ,78 |
| | | 3 | ,108 | 5,824 | ,01 | ,02 | ,95 | ,17 |
| | | 4 | ,010 | 19,077 | ,97 | ,97 | ,04 | ,03 |

a. Variable dependiente: M Aprobadas

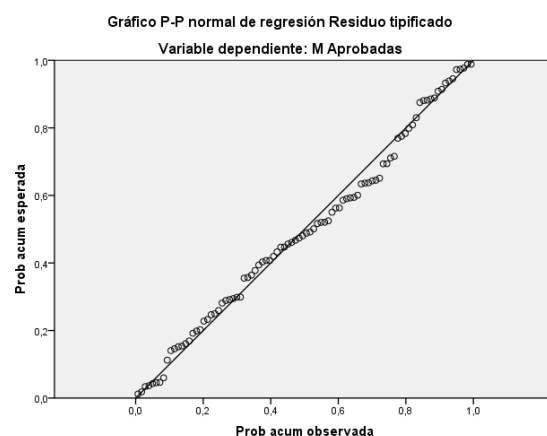
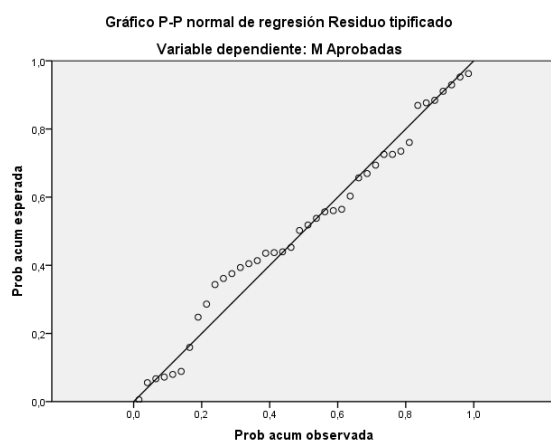
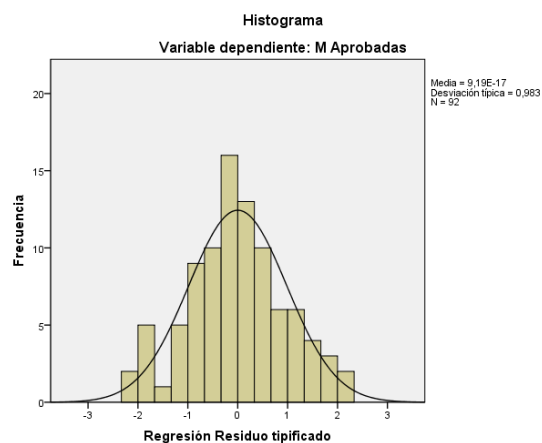
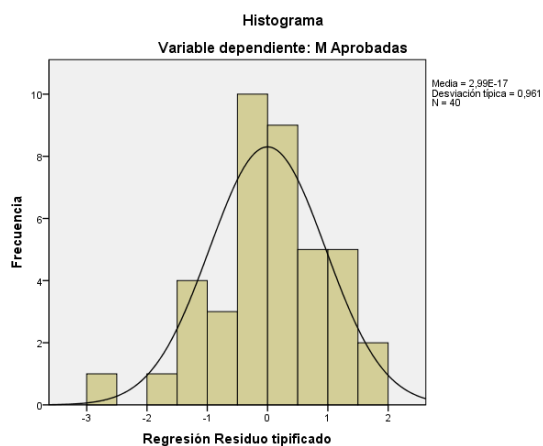
Estadísticos sobre los residuos^a

| | Mínimo | Máximo | Media | Desviación típica | N |
|-------------------------|--------|--------|-------|-------------------|----|
| Valor pronosticado | -,91 | 4,24 | 1,73 | 1,274 | 40 |
| Residual | -3,865 | 2,701 | ,000 | 1,456 | 40 |
| Valor pronosticado tip. | -2,068 | 1,978 | ,000 | 1,000 | 40 |
| Residuo típ. | -2,550 | 1,782 | ,000 | ,961 | 40 |

| | Mínimo | Máximo | Media | Desviación típica | N |
|-------------------------|--------|--------|-------|-------------------|----|
| Valor pronosticado | -,73 | 6,31 | 2,20 | 1,461 | 92 |
| Residual | -2,974 | 2,989 | ,000 | 1,294 | 92 |
| Valor pronosticado tip. | -2,004 | 2,815 | ,000 | 1,000 | 92 |
| Residuo típ. | -2,261 | 2,272 | ,000 | ,983 | 92 |

a. Variable dependiente: M Aprobadas

GRAFICOS



El modelo de regresión lineal múltiple asume que la distribución de los errores es Normal. Para estudiar si se verifica esta hipótesis realizamos estos gráficos que nos indican que se cumple esta condición. La falta de normalidad influye en el modelo en que los estimadores no son eficientes y los intervalos de confianza de los parámetros del modelo y los contrastes de significación no son exactos.

INFLUENCIA DE CADA VARIABLE REGRESORA EN EL MODELO

Femenino en forma individual cada variable regresora X

| Estadísticas de la regresión | IMC | Regularizadas | Rendidas |
|---|--------------|---------------|-------------|
| Coeficiente de correlación múltiple | 0,017767479 | 0,46722743 | 0,561938502 |
| Coeficiente de determinación R ² | 0,000315683 | 0,218301472 | 0,31577488 |
| R ² ajustado | -0,025991799 | 0,197730458 | 0,297768956 |
| Error típico | 1,959651675 | 1,73287442 | 1,621238658 |
| Observaciones | 40 | 40 | 40 |

En este cuadro podemos ver claramente mediante comparación, cuales son las variables regresoras que tienen mayor influencia en el modelo. Si analizamos el coeficiente de determinación R², empleando el método de Regresión Lineal Simple para cada variable independiente relacionada a la variable dependiente por separado.

Masculino en forma individual cada variable regresora X

| Estadísticas de la regresión | IMC | Regularizadas | Rendidas |
|---|-------------|---------------|-------------|
| Coeficiente de correlación múltiple | 0,16221514 | 0,552942416 | 0,693221998 |
| Coeficiente de determinación R ² | 0,026313752 | 0,305745316 | 0,480556739 |
| R ² ajustado | 0,015495015 | 0,298031375 | 0,474785147 |
| Error típico | 1,936093197 | 1,634843369 | 1,414118484 |
| Observaciones | 92 | 92 | 92 |

Haciendo la comparación con el cuadro Femenino vemos que tanto el IMC como las otras dos variables regresoras son notablemente mayores en su coeficiente de determinación, lo que nos estaría anticipando que con el Análisis de Regresión Lineal Múltiple, también se ajustara mucho mejor el modelo para el caso de los hombres.

ECUACIÓN DEL MODELO DE REGRESIÓN LINEAL MULTIPLE FEMENINO

FEMININO $Y = 0,47234749 - 0,0785446 X1 + 0,44077075 X2 + 0,43141897 X3$

Estadísticas de la regresión

| | |
|---|-----------|
| Coeficiente de correlación múltiple | 0,6583228 |
| Coeficiente de determinación R ² | 0,4333889 |
| R ² ajustado | 0,3861713 |
| Error típico | 1,5157603 |
| Observaciones | 40 |

X1=IMC

X2=Materias regularizadas de primer año

X3=Materias rendidas de primer año

Y=Materias aprobadas de primer año

Realizando el análisis de Regresión Lineal Múltiple y observando el valor del coeficiente de correlación de 0,6583228 se concluye que es un indicador muy fuerte, y asociado a su Coeficiente de determinación 0,4333889, que indica que el 43,34% de los datos del modelo esta explicado con esta ecuación que obtuvimos.

ANÁLISIS DE VARIANZA

| | Grados de libertad | Suma de cuadrados | Promedio de los cuadrados | F | Valor crítico de F |
|-----------|--------------------|-------------------|---------------------------|-----------|--------------------|
| Regresión | 3 | 63,263949 | 21,087983 | 9,1785485 | 0,0001204 |
| Residuos | 36 | 82,711051 | 2,2975292 | | |
| Total | 39 | 145,975 | | | |

| | Coefficientes | Error típico | Estadísticos | Probabilidad | Inferior 95% | Superior 95% | Inferior 95,0% | Superior 95,0% |
|--------------|---------------|--------------|--------------|--------------|--------------|--------------|----------------|----------------|
| Intercepción | 0,4723475 | 2,0089361 | 0,2351232 | 0,8154454 | -3,601964 | 4,54665873 | -3,6019638 | 4,5466587 |
| Variable X1 | -0,0785446 | 0,0889178 | -0,883339 | 0,3829142 | -0,258878 | 0,1017891 | -0,2588782 | 0,1017891 |
| Variable X2 | 0,4407708 | 0,1618059 | 2,7240704 | 0,009886 | 0,1126131 | 0,7689284 | 0,1126131 | 0,7689284 |
| Variable X3 | 0,431419 | 0,1218364 | 3,5409697 | 0,0011226 | 0,1843233 | 0,67851463 | 0,1843233 | 0,6785146 |

ECUACIÓN DEL MODELO DE REGRESIÓN LINEAL MULTIPLE MASCULINO

MASCULINO $Y = -1,1674185 + 0,01909761 X1 + 0,35545146 X2 + 0,51088036 X3$

Estadísticas de la regresión

Coefficiente de correlación múltiple

0,7486684

Coefficiente de determinación R²

0,56050437

R² ajustado

0,54552156

Error típico

1,31544877

Observaciones

92

X1=IMC

X2=Materias regularizadas de primer año

X3=Materias rendidas de primer año

Y=Materias aprobadas de primer año

ANÁLISIS DE VARIANZA

| | Grados de libertad | Suma de cuadrados | Promedio de los cuadrados | F | Valor crítico de F |
|-----------|--------------------|-------------------|---------------------------|------------|--------------------|
| Regresión | 3 | 194,202579 | 64,7341932 | 37,4098408 | 1,11143E-15 |
| Residuos | 88 | 152,275681 | 1,73040547 | | |
| Total | 91 | 346,478261 | | | |

| | Coefficientes | Error típico | Estadístico t | Probabilidad | Inferior 95% | Superior 95% | Inferior 95,0% | Superior 95,0% |
|--------------|---------------|--------------|---------------|--------------|--------------|--------------|----------------|----------------|
| Intercepción | -1,0869713 | 0,99104638 | -1,0967915 | 0,27572485 | -3,056467698 | 0,88252517 | -3,0564677 | 0,88252517 |
| Variable X 1 | 0,01702566 | 0,03992759 | 0,42641336 | 0,67084796 | -0,062322043 | 0,09637336 | -0,062322 | 0,09637336 |
| Variable X 2 | 0,35245126 | 0,08823786 | 3,99433157 | 0,0001343 | 0,177097059 | 0,52780546 | 0,17709706 | 0,52780546 |
| Variable X 3 | 0,50936976 | 0,07405402 | 6,87835399 | 8,4442E-10 | 0,362202955 | 0,65653656 | 0,36220296 | 0,65653656 |

Y como era de esperar los datos de la muestra Masculina en el análisis de Regresión Lineal Múltiple y observando el coeficiente de correlación de 0,74916816 es un indicador que muestra que las variables tienen una asociación fuerte positiva, y su Coeficiente de determinación 0,5612529 indica que el 56,13% de los datos del modelo esta explicado con esta ecuación que obtenido, además de ver que es sensiblemente superior al de la muestra Femenina.

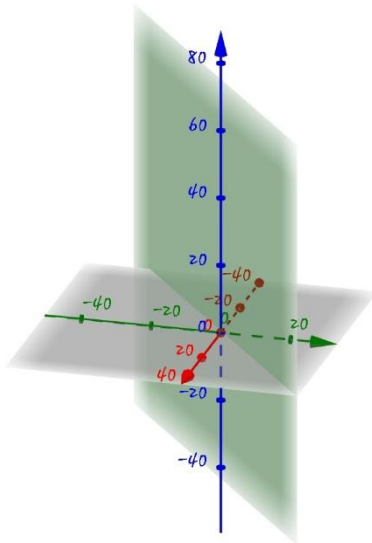


Gráfico de la ecuación estimada por el método de RLM para la muestra femenina eliminando la variable de menor incidencia IMC

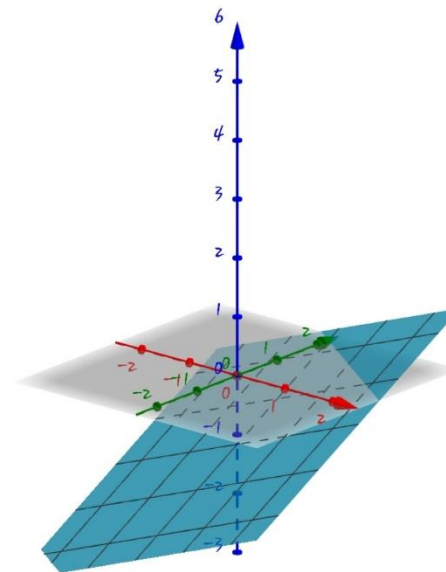


Gráfico de la ecuación estimada por el método de RLM para la muestra masculina eliminando la variable de menor incidencia IMC

CONCLUSIONES

Cuando se inició este trabajo fue pensando en determinar cómo afecta en el Rendimiento Académico el IMC, tomando como Variable dependiente las Materias Aprobadas de primer año, ya que los alumnos de probabilidad, con los que hicimos las encuestas son de segundo año. Durante el transcurso de este estudio vimos que las variables seleccionadas influyen en este modelo de regresión lineal múltiple pero en el siguiente orden de importancia de mayor a menor:

1. Materias Rendidas
2. Materias Regularizadas
3. IMC

Contrario a lo que pensábamos que el IMC iba a ser el que más influía en el rendimiento de los alumnos. Pero de todas maneras y comparando entre los varones y las mujeres con todas las interpretaciones parciales que fuimos haciendo durante este estudio, logramos ver que en el caso de las mujeres a mayor IMC se produce un menor rendimiento académico, mientras que en los hombres prácticamente no influye en su rendimiento.

Tenemos también calculadas las formulas con las que se estima la recta de regresión lineal múltiple para probar y jugar dándole valores a las variables regresoras y pronosticar la incidencia sobre el la cantidad de materias aprobadas.

Como conclusión general observamos que tanto para varones como para mujeres, analizando sus respectivos coeficientes de correlación y determinación tanto en forma individual y conjunta de las variables, además de los F-Test y T-Test, realizados entre otras interpretaciones se estableció que el modelo obtenido de Regresión Lineal Múltiple, para evaluar, pronosticar y relacionar las variables mencionadas en este trabajo, permite realizar inferencias con respecto a la población, minimizando errores sobre el comportamiento de dichas variables.

BIBLIOGRAFIA

2008 - Anderson, David R., Dennis J. Sweeney y Thomas A. Williams. Estadística para administración y economía, Ed. Cengage Learning. 10a. Edición.

2004 - Douglas C Montgomery. Diseño y Análisis de Experimentos. Ed. Limusa. 2da Edición

1996 - Douglas C. Montgomery y George C. Runger - Probabilidad y Estadística aplicadas a la Ingeniería. Ed. Mc Graw-Hill.

2008 - Jay L. Devore Probabilidad y Estadística para Ingeniería y Ciencias. Séptima Edición.

2004 – Levin Richard, Rubin David. Estadística para Administración y Economía. Séptima edición. Ed. Pearson.

2009 - Ronald E. Walpole, Raymond H. Myers, Sharon I. Myers y Keying Ye. Probabilidad y estadística para ingeniería y ciencias, 9ª Edición. Ed. Pearson educación
Estadística. Novena edición. Triola, Mario F. Ed. Pearson educación