

1.-INTRODUCCION

Hoy en día muchas empresas e instituciones manejan grandes cantidades de información, parte de la cual se desperdicia, ya que no se emplea para hacer un análisis completo del proceso y poder así mejorar la calidad del producto y/o servicio, recentrando y disminuyendo la variación de los procesos; razón por la cual la presente investigación contempla un análisis exhaustivo de la información disponible en los medios educativos.

La industria en la cual realizaré mi investigación es la educativa, esto es, la Facultad de Ciencias Físico Matemáticas de la Universidad Autónoma de N.L. Y para este proceso se tomarán las consideraciones pertinentes.

Y ante las dificultades de reproducir aspectos sociales en el laboratorio; el procedimiento para este tipo de investigación consistirá en la manipulación de diversas variables.

Ya que el proceso educativo a nivel superior es una responsabilidad compartida maestro-alumno, estos son dos grandes aspectos de los cuales partiremos , así algunas de las variables que intervienen en el proceso de educación son: hábitos de estudio, experiencia docente, apoyo audiovisual, reconocimiento de la labor docente, etc. las cuales presentan variación . Para su control, se hará una estimación de un modelo que explique la calidad educativa en función de las variables antes mencionadas.

La aportación principal de esta tesis consiste en construir un modelo de regresión en el cual intervengan de las variables mencionadas, aquellas que influyen en los factores que facilitan o favorecen al proceso educativo a la vez que se expresa dicho modelo como una función de optimización, en la cual las maximicemos para tender a la formación de estudiantes de calidad. Esto desde el enfoque "formativo" es decir dentro de la facultad.

Por lo cual, el título de esta tesis de maestría es:

APLICACION DE UNA FUNCION DE REGRESION COMO MODELO DE OPTIMIZACION PARA MAXIMIZAR LOS FACTORES QUE INCREMENTAN LA PRODUCTIVIDAD EDUCATIVA EN UNA INSTITUCION DE ENSEÑANZA SUPERIOR.

Cuyo objetivo es:

SE CONSTRUIRA UN PROCEDIMIENTO ESTANDAR QUE PERMITA UTILIZAR UN MODELO DE REGRESION LINEAL COMO CRITERIO DE OPTIMIZACION.

2.- PRESENTACION DE UN PROBLEMA DE CALIDAD EN LA ENSEÑANZA A NIVEL SUPERIOR.

Siendo uno de los objetivos básicos de la investigación científica descubrir las interrelaciones entre los fenómenos y variables, para hacer predicciones que permitan tanto estructurar políticas y estrategias de acción como contribuir al desarrollo del cuadro teórico de la ciencia social, este será uno de nuestros objetivos particulares. La investigación es un proceso que se inicia con el planteamiento de un problema que requiere solución y para encontrarla hay que construir un diseño básico que permita descubrir, explicar y si es posible predecir probabilísticamente determinadas situaciones, así como los efectos que en el proceso social va a tener la solución aportada. Por otra parte, es necesario estar conscientes de que no todos los problemas pueden ser analizados, debido a limitaciones teóricas, imposibilidad de establecer un método o por carecer de técnicas e instrumentos idóneos para llevar a cabo la investigación.

Ya que por ser un proceso humano no se cuenta con estructuras únicas ya establecidas, sino que sucede de acuerdo a circunstancias, partiremos del proceso para el cual ya tenemos organizados sus componentes ; esto es: el proceso de producción a partir del cual haremos una homologación con el sistema educativo. Así, en un proceso de producción tenemos lo que se conoce como las 5M del diagrama de Ishikawa, lo cual relacionaremos con el proceso educativo a continuación:

EN LA INDUSTRIA:

- *Maquinaria.
- *Materia prima.
- *Mano de obra.
- *Medio ambiente.
- *Método.

EN EL SISTEMA EDUCATIVO:

- *Material
didáctico, edificio, libros, etc.
- *Alumno.
- *Maestro.
- *Sociedad, familia.
- *Preparación Pedagógica y en
general del maestro.

Así el problema que trataremos es la calidad del egresado, siendo este un tema que nos atañe a todos los que de una manera u otra, participamos en este proceso.

Siendo este un tema tan amplio, y en el cual intervienen muchos factores diversos, delimitaremos nuestro marco de referencia, dentro del cual trabajaremos. Para esto, considerando el Reglamento Interno de la Facultad, en la parte de naturaleza y fines está la formación del estudiante y para ello atribuye ciertas cualidades. Por lo tanto, de acuerdo con lo anterior diremos que un egresado de calidad es aquel que posea las siguientes características:

- *Conocimientos de su área.
- *Formación humanística.
- *Responsable.
- *Auto-didacta.
- *Que participe en el desarrollo de la sociedad.
- *Con actitud de servicio.
- *Investigador.

Por lo tanto partiremos de estas características, para establecer las bases de nuestro análisis. Y tomando esto como referencia partiremos a la elaboración de encuestas.

Considerando que no se puede hablar del producto (calidad del egresado) sin hablar del proceso (todos los factores que intervienen); se tomarán en cuenta dos grandes aspectos para la elaboración de las encuestas: alumnos y maestros.

Y desglosando cada uno de ellos, tenemos la siguiente estructura:

- ALUMNO: Información personal: *Edad.
- *Sexo.
 - *Nacionalidad.
 - *Estado civil.
 - *Trabaja.
- Hábitos personales: *Cuánto estudia.
- *Cómo estudia.
 - *Dónde estudia.
- Familia: *Vive con ella.
- *Buenas relaciones.
 - *Apoyo.
 - *Nivel de estudios.
- Medio ambiente: *Vive lejos.
- *Hay transporte.
- Escuela *Condiciones del edificio.
- *Maestros.
 - *Planes de estudio.
 - *Biblioteca, laboratorios
cafetería .

- MAESTROS: Información personal: *Edad.
- *Antigüedad.
- Preparación profesional.
- Preparación Pedagógica.
- Relación maestro-alumno.
- Motivación.

Asignación de cursos.

Planes de estudio.

De lo anterior, podemos concluir que las características de un alumno de calidad están incluidas en estos aspectos, ya que: la formación humanística, la actitud responsable y la formación auto-didacta se logran mediante la participación del maestro y su actuación dentro y fuera del aula en la institución. Los conocimientos son logrados mediante el cumplimiento de los contenidos de los planes de estudio, apoyados con materiales y técnicas educativas. El desarrollo de la sociedad es impartido a través de la realización del servicio social, requisito para titularse. La investigación es fomentada a través de la actitud del maestro, realización de prácticas de laboratorio, participando en investigaciones cortas sugeridas por el instructor.

3.-PASOS DEL METODO CIENTIFICO EN LA SOLUCION DE UN PROBLEMA.

Los pasos del método científico en la solución de un problema son los siguientes:

1. Definir el problema.
2. Recolectar datos.
3. Formular hipótesis.
4. Probar hipótesis.
5. Evaluar resultados.
6. Obtener conclusiones.

1.- Definir el problema. Aquí estamos en el establecimiento de las fronteras de nuestro análisis. Este primer paso abarca lo que se refiere a: ubicar la investigación dentro de una perspectiva teórica que incluya algo del desarrollo histórico del problema. Formular el problema dentro de un marco teórico y conceptual elegido de antemano, para fundamentar adecuadamente el problema. Reducir el problema a términos concretos. Y formular objetivos claros y precisos que orienten la investigación.

2.-Recolectar datos. Elaborar instrumentos de recolección de datos adecuados y precisos a fin de obtener información empírica objetiva que sirva para la comprobación de las hipótesis. También se requiere de determinar los procedimientos para la prueba de las hipótesis. En nuestra investigación, utilizaremos la aplicación de encuestas, ya que son las que más se ajustan a nuestras circunstancias. Aplicaremos unas primeras encuestas, a partir de las cuales se elaborarán unas segundas las cuales consistirán solo de aquellas preguntas que presenten correlación ya que son las que necesitaremos para realizar nuestro análisis.

3.-Formulación de hipótesis. Debidamente fundamentadas en los marcos de la ciencia a fin de que las respuestas o explicaciones al problema o problemas

de investigación tengan mayor sustentación científica. A medida que las hipótesis se fundamenten adecuadamente con los elementos teóricos y empíricos disponibles habrá mayores posibilidades de que las hipótesis se comprueben en los términos planteados o que los ajustes sean menores. Las hipótesis deberán de tener las siguientes características:

1).- Términos claros y precisos.

2).- Con referencias empíricas.

3).- Deben de referirse a un ámbito espacio-temporal, para que sean corroboradas empíricamente.

4).- Deben de elaborarse en términos afirmativos.

4.- Pruébese hipótesis. Al probar las hipótesis, corregiremos nuestro curso, es decir, aquellas preguntas que en realidad no presenten correlación entre ellas, serán eliminadas.

5.- Evaluar resultados. En esta sección se presentará el análisis de regresión, de correlación, tablas de contingencia y nuestra función de regresión. Así como el análisis gráfico de nuestras variables.

6.- Obtener conclusiones. Concluiremos la función de regresión que nos describe la dependencia entre los factores que intervienen en la calidad del egresado, así como la función de optimización con sus restricciones respectivas.

4.-ALGUNOS CONCEPTOS IMPORTANTES.

A continuación se definen algunos términos empleados en esta investigación. Su aplicación en nuestros datos se puede observar en el apéndice H y en el H1.

ESTADISTICA: Tratar de definir a la Estadística desde un solo punto de vista resultaría algo incompleto, razón por la cual la definiremos desde algunos puntos de vista. Desde el punto de vista científico, la Estadística es la Ciencia Matemática que constituye la tecnología del método científico. También podemos decir que es un conjunto de herramientas sumamente útiles en investigación. Ofrece planeación, análisis e interpretación de resultados al investigador. Para el estadístico, el significado es relativo a la cantidad calculada apartir de la observación. Resumiendo, la ciencia de la Estadística trata con:

*Colección y compendio de datos.

*Diseño de experimentos y reconocimientos.

*Medición de la variación, tanto de datos experimentales como de reconocimiento.

*Estimación de parámetros de población y suministro de varias medidas de la exactitud y precisión de estas estimaciones.

*Ensayo de hipótesis respecto a poblaciones.

*Estudio de la relación entre dos o más variables.

PROBABILIDAD: La probabilidad es una medida de la frecuencia de ocurrencia de un evento casual. La definición clásica de probabilidad es: Si un evento puede ocurrir de N maneras mutuamente exclusivas e igualmente posibles y si n de ellas tienen una característica E , entonces la probabilidad de ocurrencia de E está dada por n/N . Lo cual se denota por $P(E) = n/N$.

TEOREMA DEL LIMITE CENTRAL: Si x es la media de una muestra aleatoria de tamaño n , que se toma de una población con media μ y varianza finita σ cuadrada, entonces la forma límite de la distribución de $Z = (\bar{x} - \mu) / \sigma / \sqrt{n}$ conforme n tiende a infinito es la distribución normal estándar $N(0,1)$.

MEDIDAS DE TENDENCIA CENTRAL: Son aquellas que nos dan aquel valor alrededor del cual se ubican los demás datos y son las siguientes: media aritmética, semirango, la mediana, la moda.

MEDIDAS DE DISPERSION O VARIABILIDAD: Son aquellas que muestran que tanta variación hay entre los datos y son las siguientes: rango, desviación estándar. Estas medidas y las anteriores se pueden apreciar en el apéndice H.

PROMEDIO: Está dado como la suma de todos los valores dividida entre el total.

MEDIA GEOMETRICA: La media geométrica de una serie de n números X_1, X_2, \dots, X_n es la raíz n -ésima del producto de los números: $(X_1 X_2 \dots X_n)^{1/n}$

ERROR ESTANDAR: Está dado como s/\sqrt{n} para un estimador puntual, es decir, coincide con la desviación estándar del estimador. Si se desconoce s y el muestreo es a partir de una distribución normal, σ reemplaza a s y tenemos σ/\sqrt{n} donde s es la desviación estándar.

MEDIA ARITMETICA: Se define como el promedio de todos los valores de la muestra.

SEMIRRANGO: Se define como: $(X_{\max} + X_{\min}) / 2$.

MEDIANA: Teniendo los datos ordenados de menor a mayor, la mediana es el $(n+1)/2$ ésimo dato.

MODA: Es el valor que ocurre más frecuentemente en la muestra. Una muestra puede tener mas de una moda y puede haber moda absoluta y moda relativa. La absoluta es en toda la muestra y la relativa se dá entre datos vecinos.

RANGO: Es una medida de dispersión y es la mas sencilla. Está dada por: $X_{\max} - X_{\min}$.

DESVIACION ESTANDAR: (s) Es la medida de variabilidad más conocida y más utilizada. Está dada por:

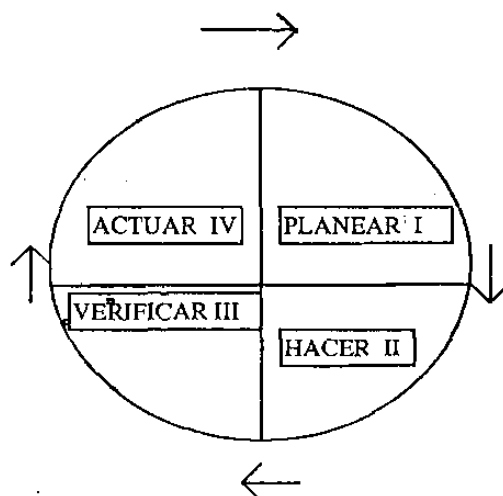
$$\sqrt{\{[\sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2 / n] / (n - 1)\}}$$

VARIANZA: Está dada como la raíz cuadrada de la desviación estándar. Es una medida de la variabilidad de los datos.

MATRIZ DE RANGO COMPETO: Una matriz de rango completo es aquella cuyo determinante es diferente de cero.

5.- COMPARACION DEL METODO CIENTIFICO CON EL DIAGRAMA DE DEMING PARA EL DIAGNOSTICO DE UN PROBLEMA DE CALIDAD.

A continuación presento el círculo de Deming, el cual es una herramienta muy valiosa y clave en el control total de calidad.



Los pasos principales del círculo de Deming son:

PLANEAR
HACER
VERIFICAR
ACTUAR

Y su importancia reside en que nunca termina su ejecución, y es aplicable a todo tipo de problemas.

Equiparándolo con el método científico podemos observar que:

Planear equivale a definir el problema, recolectar datos, formular hipótesis.

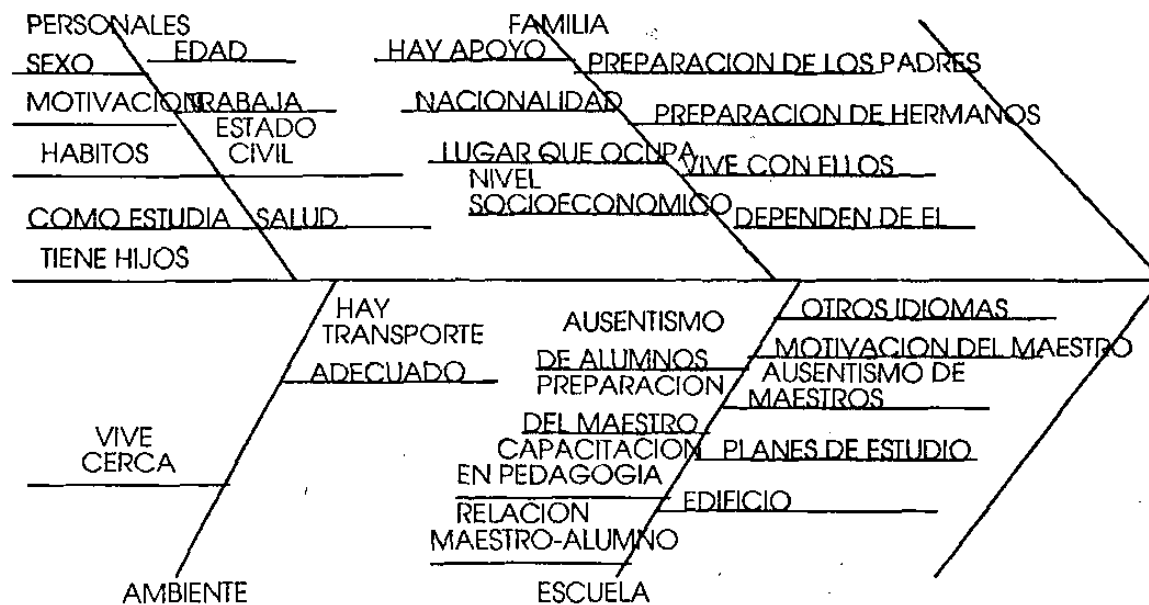
Hacer equivale a probar las hipótesis.

Verificar equivale a evaluar los resultados.

Y actuar equivale a obtener conclusiones y podríamos agregar además, el actuar de acuerdo a las conclusiones obtenidas.

En todo caso la diferencia con el método científico es que la aplicación de este círculo de Deming es interminable, de allí el sentido de las flechas.

6.- DIAGRAMA CAUSA EFECTO.



7.-ELABORACION DE HIPOTESIS

Así, nuestras primeras hipótesis, es decir, las relativas a las primeras encuestas son las siguientes:

- * Que la edad, influye en la forma de estudiar del alumno.
- * Que si el alumno es foráneo, o si es del mismo estado, sus hábitos de estudio pueden depender de esto, ya que el alumno foráneo tiene más responsabilidades con respecto a su alimentación, ropa, etc. y no debemos olvidar el factor económico.

* Así como sus hábitos de estudio. Yo diría que estos son los principales.

Y considero que estos están relacionados con la calificación que desea obtener así como también con el tiempo que dedica a prepararse para un examen parcial o un examen integrador.

Ahora las hipótesis referentes a las encuestas de maestros son las siguientes:

Su edad, categoría, experiencia, están correlacionadas entre sí.

Trabaja en otras partes, su preparación, su agrado con las materias que imparte, pueden presentar correlación.

Las instalaciones y el equipo que le proporciona la escuela para realizar su trabajo, influye en el agrado por sus cursos, y en como reaccionan los grupos en sus clases.

Sus hábitos, su vocación, su motivación, si se siente motivado económicamente son factores que también hay que tomar en cuenta.

8.- DISEÑO DE LAS ENCUESTAS.

Llevaremos a cabo la investigación directa, ya que la información será obtenida por medio de encuestas, por ser estas las que más se adaptan a las circunstancias de nuestro estudio. Algunas consideraciones importantes consideradas para la elaboración de las encuestas son las siguientes:

- * Los objetivos de estudio.
- * La disponibilidad de recursos.
- * El nivel de confianza para estimar los parámetros.
- * La homogeneidad de la población.
- * Tipo de preguntas: abiertas o cerradas.
- * El número de preguntas del cuestionario, que está en relación con el número de variables sujetas a investigación.
- * El plan de análisis estadístico.

Para el diseño de las encuestas hubo que considerar bibliografía referente al tema, la cual es mencionada, así como también todas las preguntas posibles de la lista de factores que intervienen como parte de nuestro estudio. Mostrándose la primera encuesta elaborada en el apéndice A. Esta encuesta se aplicó a una muestra piloto de 30 alumnos y 30 maestros.

La primer encuesta de alumnos consta de 60 preguntas y la de los maestros consta de 41 preguntas y siendo muchas las variables a considerar, se optó por la elaboración de una segunda encuesta que constará únicamente de aquellas preguntas correspondientes a variables que presenten correlación. Los datos obtenidos se muestran en el apéndice A.

Las variables a utilizar son las siguientes:

Para la encuesta de alumnos:

- X_1 : EDAD.
- X_2 : SEXO.
- X_3 : CARRERA.
- X_4 : NACIONALIDAD.
- X_5 : DE NUEVO LEON U OTRO ESTADO.
- X_6 : SI ES BECADO.
- X_7 : ESTADO CIVIL.
- X_8 : TIENE HIJOS.
- X_9 : TRABAJA.

- X₁₁ : APORTA DINERO A SUS PADRES.
- X₁₂ : PRACTICA ALGUN DEPORTE.
- X₁₃ : TIEMPO EN TRANSPORTARSE.
- X₁₄ : HAY TRANSPORTE ADECUADO.
- X₁₅ : TIENE AUTOMOVIL.
- X₁₆ : LLEGA A TIEMPO A CLASES.
- X₁₇ : LE IMPORTA LLEGAR A TIEMPO.
- X₁₈ : PIENSA QUE SI LE VA BIEN, DEPENDE DEL MAESTRO.
- X₁₉ : PREFIERE A UN MAESTRO "DURO".
- X₂₀ : PREFIERE A UN MAESTRO "FÁCIL".
- X₂₁ : PIENSA QUE SU DESEMPEÑO DEPENDE DEL CURSO.
- X₂₂ : LE AGRADA PRESENTAR MUCHOS EXAMENES.
- X₂₃ : SUFICIENTE MATERIAL DIDACTICO.
- X₂₄ : INFLUYE LA PREPARACION EN GENERAL DEL MAESTRO.
- X₂₅ : INFLUYE LA PREPARACION DEL MAESTRO EN EL CONTENIDO DEL CURSO.
- X₂₆ : INFLUYE LA PREPARACION DEL MAESTRO EN PEDAGOGIA.
- X₂₇ : QUE CALIFICACION DESEA OBTENER.
- X₂₈ : IMPORTANCIA DE LOS CURSOS PARA ESTUDIOS DE MAESTRIA.
- X₂₉ : LE INTERESA LA OPORTUNIDAD EN LA QUE PASE.
- X₃₀ : ESTUDIA EN CASA.
- X₃₁ : CUANTAS HORAS ESTUDIA EN CASA.
- X₃₂ : CUANTAS HORAS VE T.V.
- X₃₃ : ESTUDIA CON COMPAÑEROS.
- X₃₄ : CON CUANTO TIEMPO SE PREPERA PARA UN PARCIAL.
- X₃₅ : CON CUANTO TIEMPO SE PREPARA PARA UN INTEGRADOR.
- X₃₆ : TOMA APUNTES EN CLASE.
- X₃₇ : ESTUDIA DE LIBROS O DE APUNTES.
- X₃₈ : REALIZA ACTIVIDADES EXTRAESCOLARES.
- X₄₀ : DUERME EN EL DIA.
- X₄₂ : LE GUSTA SU CARRERA.
- X₄₃ : VE APLICACION DE LO QUE ESTUDIA EN LA SOCIEDAD.
- X₄₄ : LAS INSTALACIONES DEL SALON, SON APROPIADAS.
- X₄₅ : LAS INSTALACIONES DE LOS LABORATORIOS, SON ADECUADAS.

- X₄₆: LA BIBLIOTECA, ESTA BIEN EQUIPADA.
- X₄₇: LO APOYA SU FAMILIA.
- X₄₈: LUGAR QUE OCUPA ENTRE LOS HERMANOS.
- X₅₀: SU RELACION CON SUS PADRES.
- X₅₁: LOS PADRES, VIVEN JUNTOS.
- X₅₂: ESTUDIOS DEL PADRE.
- X₅₃: ESTUDIOS DE LA MADRE.
- X₅₄: ADMINISTRATIVO.
- X₅₅: ADMINISTRATIVO.

Obs: Las variables que no aparecen enlistadas, fueron eliminadas por tener las mismas respuestas, esto es solo ceros o solo unos; y son las siguientes:

- X_{10'}: CUANTAS HORAS AL DIA TRABAJA.
- X_{17'}: LE ES IMPORTANTE LLEGAR A TIEMPO A SUS CLASES.
- X_{39'}: CUANTAS HORAS REALIZA ACTIVIDADES EXTRAESCOLARES.
- X_{41'}: CUANTAS HORAS DUERME, SI DUERME.
- X_{49'}: OCUPACION Y ESTUDIOS DE LOS HERMANOS MAYORES.
- X_{53'}: EFICIENCIA DE LOS TRAMITES ADMINISTRATIVOS.

Y además, algunas de estas preguntas eran contestadas si contestaban sí a la anterior, no estaban completos los datos, por lo tanto no se incluyeron.

Y aparte las preguntas abiertas no se introdujeron ya que su procesamiento es diferente.

Para las encuestas de maestros, las variables son las siguientes:

- Y₁: EDAD.
- Y₂: CATEGORIA.
- Y₃: ANTIGUEDAD EN LA FACULTAD.
- Y₅: TRABAJA FUERA DE LA U.A.N.L.
- Y₇: NIVEL DE ESTUDIOS.
- Y₈: CUENTA LA ESCUELA CON INSTALACIONES ADECUADAS.
- Y₉: CUENTA CON MATERIAL DIDACTICO.
- Y₁₀: LOS LABORATORIOS Y LA BIBLIOTECA ESTAN BIEN EQUIPADOS.
- Y₁₁: PLAN DE ESTUDIOS.
- Y₁₂: SEGUIMIENTO DEL PLAN DE ESTUDIOS.
- Y₁₄: ASIGNACION DE MATERIAS.
- Y₁₅: SISTEMA ROTATIVO DE ASIGNACION DE CURSOS.
- Y₁₆: PERIODO MAYOR DE ASIGNACION DE UN CURSO.
- Y₁₇: SISTEMA DE EVALUACION.
- Y₁₈: INDICE DE REPROBACION.
- Y₁₉: INGLES.
- Y₂₀: PUNTUALIDAD.
- Y₂₁: EXIGE PUNTUALIDAD.

Y_{22} : CUENTA LA ASISTENCIA EN LA CALIFICACION.
 Y_{23} : GUSTO POR LAS MATERIAS QUE IMPARTE.
 Y_{25} : CONSIDERA CLARAS SUS EXPLICACIONES EN CLASE.
 Y_{27} : LE AGRADA ACLARAR DUDAS DE SUS CLASES.
 Y_{30} : CURSOS DE PEDAGOGIA.
 Y_{31} : EL AMBIENTE DE SUS CLASES.
 Y_{32} : EFECTUA EL TRABAJO EN EQUIPOS.
 Y_{33} : CONOCE EL TRABAJO EN EQUIPOS.
 Y_{34} : CURSOS DE CAPACITACION.
 Y_{35} : PREPARACION DE CLASES.
 Y_{36} : FOMENTA EN ALUMNOS LA INVESTIGACION EN LA BIBLIOTECA.
 Y_{37} : COMO SE CONSIDERA COMO MAESTRO.
 Y_{41} : AMBIENTE DE TRABAJO.

Similarmente para los maestros, las preguntas que no se incluyeron son:

$Y_{4'}$ $Y_{6'}$ $Y_{26'}$ $Y_{28'}$ $Y_{29'}$ $Y_{38'}$ $Y_{39'}$ $Y_{40'}$

La relación de las preguntas para las encuestas de los alumnos es la siguiente:

- 1.- Factores personales:preg:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14.
- 2.- Medio ambiente:preg:15, 16, 17.
- 3.- Hábitos:puntualidad:preg:18, 18.
de estudio: preg:33, 33, 34, 35, 36, 37, 38, 39.
otros: 42, 43.
- 4.- Maestros:preg:20, 21, 22, 23, 24, 25, 26, 27, 28.
- 5.- Su preparación preg:29, 30, 31, 45, 46, 47, 48.
- 6.- Actividades extra escolares:preg:40, 41, 44.
- 7.- Instalaciones de la escuela:preg:49, 50, 51.
- 8.- Familia: preg: 52, 53, 54, 55, 56, 57, 58.
- 9.- Administrativo preg:59, 60.

Y para las encuestas de maestros:

- 1.- Edad, antigüedad, y categoría: 1, 2, 3, 4.
- 2.- Trabaja en otras partes: 5, 6.
- 3.- Preparación: 7, 30.
- 4.- Instalaciones de la escuela: 8, 9, 10.
- 5.- Planes de estudio: 11, 12.
- 6.- Asignación de cursos: 13, 14, 15, 16, 23.
- 7.-Desempeño en el aula: 17, 18, 19, 22, 24, 25, 26, 27, 28, 29, 31, 32, 33, 34, 35, 36.
- 8.- Hábitos en el aula: 20, 21.
- 9.- Motivación: 37, 38, 39, 40, 41.

9.- TAMAÑO DE LA MUESTRA.

Aquí analizaremos a quienes y a cuantas personas será aplicada la encuesta elaborada. Es importante recalcar que la muestra que tomemos, debe ser significativa, es decir, que conserve las mismas características de la población. La muestra será elegida al azar, y dentro de este tenemos los siguientes tipos de muestreo probabilístico:

- * Aleatorio simple.
- * Estratificado.
- * Por racimos.
- * Sistemático.

De los anteriores, aplicaremos el muestreo estratificado ya que nuestra población está formada por estratos, de los cuales nos interesa obtener una representación. Así, de cada uno se selecciona una muestra, cuya suma representa la muestra total.

Si p es la proporción de éxito o de afirmaciones y $q = 1 - p$. Considerando que en las respuestas de los alumnos hay una variación aproximada de $p = 0.5$ y $q = 0.5$, con una confianza del $(1 - \alpha)100\%$, si $\alpha = 0.05$ esto indica que con una confianza del 95% el error será menor que la cantidad que se especifique como e , entonces aplicando el Teorema del Límite Central el tamaño de la muestra es aproximadamente

$$n = (Z_{\alpha/2}^2 pq) / e^2$$

Por lo tanto, para las encuestas de alumnos si $p = 0.5$, $q = 0.5$, $Z = 1.96$, esto es con un 95 % de confianza de que el error sea menor que 0.6 tenemos:

$$n = (1.96)^2(0.5)(0.5) / (0.6)^2 = 267.$$

Y para la muestra de los maestros, considerando que las opiniones no difieren tanto, es decir dando a p un valor de 0.9 y $q = 0.1$, con una confianza del 95 % de que el error será menor que 0.31 tenemos :

$$n = (1.96)^2(0.9)(0.1) / (0.31)^2 = 35$$

Por lo tanto se consideraron muestras de tamaño 350 para los alumnos y de 35 para los maestros.

Ahora, para formar los estratos consideramos lo siguiente: debido a que los estratos con los cuales trabajaremos son de tamaños muy diferentes (alumnos de Física, Matemáticas y Computación) y para que nuestro estudio tenga validez consideraremos todos los alumnos de Matemáticas y de Física y el resto de Computación quedando así nuestros estratos:

Alumnos de la carrera de Física : 44.

Alumnos de la carrera de Matemáticas : 30

Siendo aproximadamente estas las cantidades de alumnos en estas carreras.

Y el resto de computación, esto es 276 alumnos.

Se efectuó un sorteo para aplicar las encuestas a los grupos de Computación, y se aplicaron de manera directa dándose las indicaciones pertinentes y aclarando dudas que se presentaran.

Ya que entre los maestros hay algunos que imparten materias a alumnos de las tres carreras, se aplicó la encuesta aleatoriamente en forma general.

10.-ANALISIS ESTADISTICO.

10.1.- ANALISIS DE LA MATRIZ DE CORRELACION SIMPLE.

Algunas de las conclusiones que podemos presentar son las siguientes:

Analizando las matrices de correlacion simple (apéndice B), tenemos que las variables que presentan correlación con una significancia menor o igual que 0.005 son las siguientes (para la elaboración de las segundas encuestas).

Las parejas de variables correlacionadas mediante correlación simple son :

Para las primeras encuestas:

VARIABLES: CORRELACION: SIGNIFICANCIA:

X_1 CON X_3	0.3344	0.0002
X_3 CON X_{40}	-0.3940	0.0000
X_{24} CON X_{25}	0.4845	0.0000
X_{29} CON X_{40}	0.3353	0.0002
X_{13} CON X_{15}	-0.5211	0.0000
X_{31} CON X_{32}	-0.4259	0.0000
X_{54} CON X_{55}	0.3675	0.0000
X_{30} CON X_{31}	-0.5006	0.0000
X_{34} CON X_{35}	0.5196	0.0000
X_1 CON X_{48}	-0.5398	0.0000
X_5 CON X_{33}	0.5218	0.0000
X_{32} CON X_{40}	0.4583	0.0000
X_{12} CON X_{27}	-0.3511	0.0001
X_{22} CON X_{29}	-0.4472	0.0000
X_{28} CON X_{47}	0.3563	0.0001

Así, las variables que formarán parte de la encuesta definitiva son:

(Las variables que aparecen entre paréntesis son aquellas cuya numeración corresponde a las encuestas originales. El número de variable que aparece a la izquierda, corresponde a la renumeración, después de quitar las preguntas no consideradas).

- X_1 : Edad.
 X_3 : Carrera.
 X_5 : Si es de otro Estado.(X_4)
 X_9 : Si trabaja.(X_{10}).
 X_{12} : Practica algún deporte.(X_{14})
 X_{13} : Cuánto tiempo hace de su casa a la escuela.(X_{15}).
 X_{15} : Tiene automóvil.(X_{17}).
 X_{22} : Su agrado por los exámenes.(X_{24}).
 X_{24} : Preparación en general del maestro.(X_{26}).
 X_{25} : Preparación del maestro en su curso.(X_{27}).
 X_{26} : Preparación en Pedagogía.(X_{28}).
 X_{27} : Con que calificación le interesa acreditar.(X_{29}).
 X_{28} : Visión para estudios de maestría.(X_{30}).
 X_{29} : Si estudia para acreditar en primera oportunidad.(X_{31}).
 X_{30} : Si estudia en casa.(X_{32}).
 X_{31} : Cuántas horas estudia por semana.(X_{33}).
 X_{32} : Cuántas horas de T.V. ve al día. (X_{34}).
 X_{33} : Estudia con sus compañeros.(X_{35}).
 X_{34} : Tiempo de preparación para un examen parcial.(X_{36}).
 X_{35} : Tiempo de preparación para un examen integrador.(X_{37}).
 X_{40} : Si duerme durante el día.(X_{42}).
 X_{47} : Apoyo por parte de la familia.(X_{52}).
 X_{48} : Lugar que ocupa en su familia, entre sus hermanos.(X_{53}).
 X_{54} : Tiempo de tramitación en la escuela.(X_{60}).
 X_{55} : Estudios de la madre.(X_{58}).
 Total : 25.

Algunas sugerencias dadas por los alumnos fueron referentes al edificio de la escuela, uso de pintarrones, abanicos, mayor apoyo económico a laboratorios, más libros y más actualizados. Además, hicieron incapié en la preparación del maestro en Pedagogía.

La nueva encuesta elaborada se puede apreciar en el apéndice A.

De las correlaciones de las encuestas de maestros podemos apreciar lo siguiente:

VARIABLES:	CORRELACION :	SIGNIFICANCIA:
Y_1 Y_{47}	0.5834	0.0062
Y_{25} Y_{31}	0.4404	0.0389
Y_{18} Y_{17}	-0.4366	0.0406
Y_{41} Y_{20}	0.4115	0.0536
Y_{11} Y_{12}	-0.5161	0.0155
Y_{17} Y_{30}	-0.5305	0.0128
Y_{17} Y_{18}	-0.4366	0.0406
Y_2 Y_{27}	-0.5112	0.0165
Y_3 Y_{16}	0.4763	0.0255

Por lo tanto, las variables que aparecerán en la encuesta definitiva de los maestros son:

- Y₁: Edad.
 - Y₂: Categoría.
 - Y₃: Antigüedad como docente en la facultad.
 - Y₇: Nivel de estudios.
 - Y₁₁: Estructuración del plan de estudios.
 - Y₁₂: Seguimiento del plan de estudios.
 - Y₁₆: Período con la asignación de una materia.
 - Y₁₇: Sistema de evaluación.
 - Y₁₈: Considera natural que en un grupo repruebe la mayoría.
 - Y₂₀: Es puntual en sus clases.
 - Y₂₅: Considera que sus alumnos le entienden a sus explicaciones.
 - Y₂₇: Le agrada que sus alumnos lo busquen, para preguntar dudas.
 - Y₃₀: Considera necesario tomar cursos de Pedagogía.
 - Y₃₁: Cómo es el ambiente que hay en sus clases.
 - Y₄₁: Le agrada el ambiente de su trabajo.
- Total : 15.

Así, las encuestas definitivas pueden apreciarse en el apéndice A.

Y verificando aquellas parejas de variables para las cuales la significancia es menor o igual que 0.05 en la aplicación de las segundas encuestas:

	COEFICIENTE	SIGNIFICANCIA
AX ₁ CON AX ₃	-0.2039	0.0001
AX ₆ CON AX ₇	-0.2006	0.0002
AX ₁₃ CON AX ₁₄	0.1290	0.0157
AX ₈ CON AX ₉	0.1104	0.0389
AX ₁₄ CON AX ₁₅	0.1745	0.0011
AX ₁₉ CON AX ₂₀	0.5592	0.0000
AX ₁₈ CON AX ₁₉	0.1395	0.0090
AX ₄ CON AX ₁₇	-0.2103	0.0001
AX ₁₀ CON AX ₁₉	0.0493	0.3582
AX ₁₃ CON AX ₂₆	-0.1390	0.0092
AX ₁₃ CON AX ₂₃	-0.1116	0.0368
AX ₁₃ CON AX ₈	0.1048	0.0501
AX ₁₄ CON AX ₁₈	-0.0923	0.0547
AX ₁₂ CON AX ₁₆	0.1225	0.0219
AX ₂ CON AX ₁₁	0.1027	0.0550
AX ₁ CON AX ₄	0.2015	0.0001
Para los maestros :		
AY ₁ CON AY ₃	0.7525	0.0000
AY ₁ CON AY ₇	0.4048	0.0159
AY ₆ CON AY ₈	-0.3246	0.0571

AY ₂ CON AY ₃	0.3741	0.0268
AY ₃ CON AY ₂	0.4213	0.0117

10.2.-ANALISIS DE REGRESION PARA LAS SEGUNDAS ENCUESTAS.

El análisis para la regresión lineal donde se muestran los valores para b_0 (intercepción con el eje y) y para b_1 (pendiente de la recta), asumiendo que la significancia es menor o igual que 0.05 se muestran en el apéndice I. En estas tablas también aparece el coeficiente de correlación, el cual como ya se hizo mención, nos proporciona una medida de la correlación entre las variables (entre más se acerque a 1) y además R^2 , que es el porcentaje de variación total que nos es explicado por la gráfica. Además podemos observar la tabla del ANOVA para cada par de variables correlacionadas. Por lo tanto R^2 varía de 0 a 100% y entre mayor sea, es mejor. En estas tablas podemos observar que las considerando la significancia menor o igual que 0.05 los valores de R^2 son mucho muy pequeños.

10.3.-DIAGRAMAS DE PASTEL.

Una de las herramientas gráficas de la Estadística son los diagramas de pastel. Su ventaja consiste en que nos permiten visualizar, sin esforzarnos los porcentajes para cada nivel de cada variable. Los diagramas de pastel para las variables de las segundas encuestas se muestran en el apéndice F.

10.4. - ANALISIS DE LAS TABLAS DE CONTINGENCIA.

En el apéndice D podemos observar las tablas cruzadas y los valores de X^2 para cada pareja de variables cuyos coeficientes de correlación simples indican que tienen correlación. Estas tablas se obtuvieron de las encuestas definitivas (AX₁, AY₁).

Las conclusiones para estas tablas son las siguientes:

- 1.- Tablas cruzadas para AX₁; edad con AX₂; es de otro Estado.
Ho: No hay relación entre estas dos variables.

$X^2 = 19.8698$. Con una significancia de 0.001 y 4 grados de libertad, X^2 en las tablas es 18.465. Por lo tanto se rechaza H_0 .

El 0.3 % tiene menos de 16 años y es de N.L.. El 62.3 % tiene entre 16 y 20 años de los cuales el 49.7 % es de N.L. y el resto es de otros Estados, hay un 0.3% que es extranjero. El 35.4 tiene entre 21 y 25 años de los cuales el 24.9 es de N.L. y el 10.6 % es de otros Estados. El 1.4 % tiene entre 25 y 30 años de edad de los cuales el 1.1 % es de otro Estado. Y por último hay un 0.6 % de alumnos que tienen entre 30 y 35 años y son de otros Estados.

2.- Tabla cruzada para AX_4 : trabaja con AX_1 : edad.

H_0 : No existe relación entre si trabaja y la edad.

$X^2 = 14.7504$. Con una significancia de 0.005 y 4 grados de libertad $X^2 = 14.860$. Con lo cual se acepta H_0 . Es decir que no hay relación.

3.- Tabla cruzada de AX_2 : carrera con AX_3 : es de otro Estado.

H_0 : No existe relación entre la carrera con el Estado de donde viene.

X^2 obtenida = 4.2536 . X^2 de las tablas con $a = 0.30$ y $n = 2$ es de 3.665 . Como X^2 obtenida es mayor que la de las tablas, se rechaza H_0 . Esto es, si hay relación de la carrera con la edad.

En general el 75 % de los alumnos son de N.L. y el resto son de otro Estado. El 61% corresponde a alumnos de Computación que son de N.L. y un 17.4 que son de otros Estados. Para los matemáticos tenemos que el 5.4% son de N.L. y el 3.1% son de otros Estados. Y para los físicos el 8.9 son de N.L. y el 4 % son de otro Estado.

4.- Tabla cruzada de AX_2 : carrera con AX_5 : practica algún deporte.

H_0 : No hay relación entre la carrera y si practica algún deporte.

X^2 obtenida es 2.0255 . X^2 de las tablas, con una significancia de 0.3 y 2 grados de libertad es 2.408. Y de lo anterior concluimos que se acepta H_0 , esto es, no hay relación.

5.- Tabla cruzada para AX_2 :carrera con AX_{11} : piensa que es importante una preparación en Pedagogía para el maestro.

H_0 :No hay relación entre estas dos variables.

X^2 obtenida es 3.69678. X^2 de las tablas, con una significancia de 0.1 y dos grados de libertad es: 4.605. Por lo tanto concluimos que se acepta H_0 , esto es no hay relación.

6.- Tabla cruzada para AX_3 : si es de N.L. con AX_6 : tiempo que hace en transportarse de su casa a la escuela.

H_0 : No hay relación entre estas dos variables.

X^2 obtenida es 8.06256. X^2 de las tablas, con una significancia de 0.1 y 4 grados de libertad, es 7.779. Por lo tanto se rechaza H_0 .

En la tabla cruzada podemos observar que del 75% que son alumnos de N.L. un 36% hacen entre 30 y 60 min. para llegar a la escuela. El 26% hace de 10 a 30 min., el 6% hace entre 5 y 10 min y el 5% hace de una a dos horas.

7.- Tabla cruzada para AX_{17} : horas al día en las que ve T.V. con AX_4 : trabaja.

Ho: las horas que ve T.V. al día no se relacionan con el hecho de que trabaje o no.

X^2 obtenida = 17.2794. La de las tablas con 3 grados de libertad y $\alpha = 0.001$ es 16.268. Por lo tanto se rechaza Ho.

En general un 75.4 de alumnos no trabajan, de los cuales un 35.4 % ven T.V. de 5 a 10 horas por semana, un 20.3 % ven T.V. de 10 a 15 horas por semana un 12.3% no ve T.V. y un 7.4% ve T.V. 15 o más horas. Un 24.6 % son alumnos que trabajan, de los cuales un 12.6 % ve de 5 a 10 h. de T.V. a la semana, un 8% no ve T.V. un 2.9% ve de 10 a 15 horas y un 1.1% ve 15 o mas horas. En general podemos decir que las personas que no trabajan son las que ven más T.V. lo cual era de esperarse.

8.- Tablas cruzadas para AX_6 : tiempo que hace para llegar de su casa a la escuela con AX_7 : tiene automóvil.

Ho: No existe relación entre estas variables.

$X^2 = 27.1557$. Con una significancia de 0.005 y 4 grados de libertad, $X^2 = 14.860$.

Con lo cual se rechaza Ho.

El 12.6 % tardan en llegar entre 5 y 10 min. de los cuales el 8.6 % no tiene automóvil y el 3.7 si. El 35.7 % hace entre 10 y 30 min. de los cuales el 31.7 no tiene automóvil. El 45.1 % hace entre 30 y 60 min. de los cuales, el 42.3 % no tiene automóvil. Y el 6.3 % hace entre 1 y dos horas. Y el .3 % hace mas de 2 horas.

9.-Tabla cruzada para AX_8 : le agrada que el maestro aplique muchos exámenes, con AX_{18} : se reúne a estudiar con sus compañeros.

Ho: No existe relación entre estas variables.

X^2 obtenida es 3.61343. La de las tablas con 2 grados de libertad y una significancia de 0.2 es de 3.219. Por lo cual se rechaza Ho.

Observando la tabla tenemos que al 78% de los alumnos les agrada que los maestros apliquen muchos exámenes, de los cuales el 60% algunas veces se reúne a estudiar con sus compañeros, el 10% siempre estudian con sus compañeros y un 9% nunca estudia con compañeros. Un 22% no tienen agrado por muchos exámenes, de los cuales el 17% algunas veces estudia con compañeros, un 4% nunca y un 1% siempre.

10.- Tabla cruzada para AX_8 : agrado por que el maestro aplique muchos exámenes con AX_{26} : tiempo para efectuar trámites administrativos.

H_0 : No existe relación entre las variables.

X^2 obtenida es 4.398. Con una significancia de 0.1 y 2 grados de libertad, en las tablas nos encontramos el valor de 4.605. Por lo cual se acepta H_0 .

11.-Tabla cruzada para AX_{10} : piensa que si el maestro no está bien preparado no aprende con AX_{11} : piensa que es importante la preparación en Pedagogía del maestro.

H_0 : No existe relación entre estas variables.

X^2 obtenida es 11.3810. La de las tablas, con 3 grados de libertad y una significancia de 0.01 es de 11.345. Por lo cual se rechaza H_0 .

Observando las tablas de contingencia, tenemos que un 93% piensa que es importante la preparación en Pedagogía del maestro de los cuales un 32.9% piensan que si el maestro no está bien preparado, entonces no aprenden, un 48.6% piensan que depende de la materia y un 11% piensan que su aprendizaje no depende de si el maestro está bien preparado.

12.-Tabla cruzada para AX_{10} : piensa que si el maestro no está bien preparado, no aprende

con AX_{19} : tiempo de preparación para un examen parcial.

H_0 : No existe relación entre estas dos variables.

X^2 obtenida es 20.7303. La de las tablas, con 12 grados de libertad y una significancia de 0.05 es de 21.026; por lo cual se acepta H_0 .

13.- Tabla cruzada para AX_{11} : piensa que es importante la preparación en Pedagogía del maestro con AX_{13} : tiene visión de realizar estudios de maestría.

H_0 : No existe relación entre estas dos variables.

X^2 obtenida es 4.49. La de las tablas, con 1 grado de libertad y una significancia de 0.025 es de 5.024; razón por la cual se acepta H_0 .

14.-Tabla cruzada para AX_{11} : piensa que es importante la preparación en Pedagogía del maestro con AX_{17} : cuántas horas al día ve T.V..

H_0 : No existe relación entre estas dos variables.

X^2 obtenida es 6.97186. La de las tablas, con 3 grados de libertad y una significancia de 0.10 es de 6.251. Por lo tanto se rechaza H_0 .

De las tablas cruzadas, podemos observar que del 93% son los alumnos que piensan que es importante la preparación en Pedagogía del maestro, de los

cuales un 45% ve T.V. entre una y dos horas al día, un 22% ve T.V. entre dos y tres horas, un 17% no ve T.V. y un 8% ve mas de 3 horas al día.

15.- Tabla cruzada de AX_{16} : Cuántas horas estudia por semana con AX_{12} : calificación que desea obtener.

Ho: no hay relación entre las horas que dedica a estudiar por semana con la calificación que desea obtener.

X^2 obtenida = 28.0752. X^2 de tablas con 12 grados de libertad y $\alpha = 0.02$ es de 24.054. Por lo tanto se rechaza Ho.

En general, el 45% de los alumnos desean pasar con 100. El 22 % con 90.

El 16% con 80, el 7.7 % con 70 y el 7.4 % no están interesados en alguna calificación en particular.

Cruzando las variables tenemos lo siguiente: el 21.1% desea pasar con 100 y estudia de 0 a 5 horas por sem., el 13.1 % estudia de 5 a 10 horas por semana y desea pasar con 90, el 12.6 % estudia de 0 a 5 horas por sem. y desea pasar con 90. El 10% estudia de 0 a 5 horas por semana y desea pasar con 80, el 6.6 % estudia de 10 a 15 horas por semana y desea pasar con 100, el 5.4 % estudia de 5 a 10 horas por semana y desea pasar con 90, el 5.1 % estudia de 0 a 5 horas y desea pasar con 70, el 4.9 % estudia de 0 a 5 horas y no desea alguna calificación en particular.

16.- Tabla cruzada para AX_{13} : estudia en casa con AX_{14} : desea pasar en primera oportunidad.

Ho: No hay relación entre si estudia en casa, con pasar en primera oportunidad.

$X^2 = 10.62$. X^2 con una significancia de 0.001 y 1 grado de libertad es 10.827

por lo tanto se rechaza Ho.

En general, el 73.4 % de los alumnos desea pasar en primera oportunidad de los cuales el 60.2% estudia en casa. Del 26.6 % que no estudian en casa un 17.5 % desea pasar en primera oportunidad.

17.- Tablas cruzadas para AX_{16} : cuántas horas estudia por semana con AX_{19} : tiempo para prepararse para un examen parcial.

Ho: No existe relación entre estas dos variables.

$X^2 = 33.1580$. Con una significancia de 0.005, y 16 grados de libertad, $X^2 = 34.267$. Por lo tanto se acepta Ho.

18.- Tablas cruzadas para AX_{19} : tiempo para prepararse para un examen parcial con AX_{18} : estudia con sus compañeros.

Ho: No existe relación entre estas variables.

$X^2 = 27.0780$. Con una significancia de 0.001 y 8 grados de libertad, $X^2 = 26.125$. Por lo tanto se rechaza H_0 .

El 76.6 % de los alumnos estudian con sus compañeros algunas veces de los cuales el 40.9% lo hace con menos de una semana para un parcial, el 17.4 % con una semana, el 4.3 % con mas de una semana y el 13.7 % un día antes. El 0.3 % no se prepara. Un 12.3% nunca se reúne a estudiar con sus compañeros. Y el 11.1 % siempre se reúne a estudiar con compañeros.

19.- Tabla cruzada para AX_{20} : preparación para un examen integrador con AX_{19} : preparación para un examen parcial.

H_0 : No hay relación entre estas dos variables.

$X^2 = 203.874$. Con 16 grados de libertad y una significancia de 0.001, $X^2 = 39.252$. Por lo tanto se rechaza H_0 .

En la preparación para un examen parcial el porcentaje mas alto corresponde a los alumnos que se preparan con menos de una semana (52.6 %) de los cuales el 22.6 % se preparan con una semana para presentar un examen integrador. Para presentar un examen integrador el mayor porcentaje corresponde a los que estudian con una semana de anticipación (37.4 %), siguiéndolo un 28.9 % de aquellos que toman menos de una semana y un 27.1 % de aquellos que toman mas de una semana. Los que no se preparan ni para un examen parcial ni para un integrador son un 0.3%. Los que no se preparan para un parcial son un 1.4 % y los que no se preparan para un integrador son 0.6 %.

Los que estudian un día antes para un parcial son un 18.3 % y para un integrador un 5.7 %.

20.- Tabla cruzada para AX_{23} : lugar que ocupa de entre sus hermanos con AX_{25} : nivel de estudios de su papá.

H_0 : No existe relación entre estas dos variables.

$X^2 = 30.0739$. La de las tablas, con 20 grados de libertad y una significancia de 0.05 es de 31.410. Por lo tanto se acepta H_0 .

21.- Tabla cruzada para AX_{24} : nivel de estudios de la madre con AX_{25} : nivel de estudios del padre.

H_0 : No hay relación entre estas variables.

$X^2 = 235.641$. Con una significancia de 0.005 y 80 grados de libertad, en las tablas encontramos el valor de 51.2. Por lo cual se rechaza H_0 . Esto es, si hay relación entre estas dos variables.

Para las encuestas de los maestros, el análisis de las tablas de contingencia es el siguiente:

1- Tablas cruzadas para AY_3 : antigüedad como docente en la facultad con AY_1 : edad.

Ho: No existe relación entre estas variables.

$X^2 = 110.992$. Con una significancia de 0.001 y 30 grados de libertad, $X^2 = 59.703$. Por lo tanto se rechaza Ho.

El 28 % de los maestros tiene entre 35 y 40 años, de los cuales 18% tienen entre 5 y 15 años de antigüedad en la facultad un 6% tienen entre 2 y 5 años de antigüedad y un 3% tienen entre 1 y 2 años. Un 14% tienen entre 45 y 50 años, de los cuales un 9% tienen entre 20 y 25 años de antigüedad y un 6% tienen entre 15 y 20 años.

2.- Tablas cruzadas para AY_1 : edad con AY_7 : mayor período de tiempo con la asignación de una misma materia.

Ho: No existe relación entre estas variables.

$X^2 = 28.1750$. La de las tablas, con 30 grados de libertad y una significancia de 0.7 es de 25.508. Por lo cual se rechaza Ho.

Observando las tablas de contingencia tenemos que a mayor antigüedad, mayor tiempo con la asignación de una misma materia y este corresponde a entre 10 y 15 años.

3.- Tabla cruzada para AY_2 : categoría con AY_5 : considera que hay seguimiento de las materias en el plan de estudios.

Ho: No existe relación entre estas dos variables.

$X^2 = 12.1859$. La de las tablas, con 6 grados de libertad y una significancia de 0.05, es de 12.592. Por lo cual se acepta Ho.

4.- Tabla cruzada para AY_2 : categoría con AY_6 : considera que es normal que en un grupo repruebe la mayoría.

Ho: No existe relación entre estas dos variables.

$X^2 = 7.23349$. Con una significancia de 0.3 y 6 grados de libertad, la de las tablas es: 7.23 por lo tanto se rechaza Ho.

Observando las tablas cruzadas, tenemos que un 66% de maestros consideran que no es normal que en un grupo repruebe la mayoría tienen tiempo completo, un 14% son de medio tiempo y un 3% son maestros por horas. Un 20% considera que depende de la materia de los cuales un 17% son maestros de tiempo completo y un 3% son maestros exclusivos.

5.- Tablas cruzadas para AY_{10} : es puntual en sus clases con AY_{12} : le agrada que sus alumnos le busquen para preguntar dudas de clase.

Ho: No existe relación entre estas variables.

$X^2 = 5.30303$. Con una significancia de 0.07 y 2 grados de libertad en las tablas encontramos 5.991. Por lo tanto se rechaza Ho.

Observando las tablas cruzadas tenemos que un 69% de los maestros no son puntuales en sus clases y les agrada que sus alumnos le busquen para aclarar dudas de clase. Un 3% de maestros son regularmente puntuales y les agrada

que los busquen para aclarar dudas. Un 29% es puntual en sus clases de los cuales a un 6% no le agrada que lo busquen sus alumnos para aclarar dudas.

10.5.-GRAFICAS DE DISPERSION

Podemos observar que a incrementos en una variable hay cambios en otra variable. Al graficar estas variables podemos observar el tipo de relación que guardan o si no hay relación. Es importante considerar que estamos trabajando con parejas de variables y que una de ellas será la absisa y la otra la ordenada. Los tipos de correlación que se pueden presentar son los siguientes:

1)CORRELACION POSITIVA: Un incremento en Y produce un incremento en X. Si X es encontrada Y será controlada (al graficar los puntos, estos tienden a formar una recta con pendiente positiva).

2)POSIBLE CORRELACION POSITIVA: Si X aumenta, Y se incrementará un poco, aunque parece haber otras causas diferentes (al graficar los puntos, se aprecia cierta tendencia a acumularse sobre una recta con pendiente positiva).

3)CORRELACION NEGATIVA:Un aumento en X causa una disminución en Y. Por lo tanto X puede ser controlada en lugar de Y (al graficar los puntos podemos observar que se acumulan sobre una recta con pendiente negativa).

4)POSIBLE CORRELACION NEGATIVA: Un aumento en X causará una tendencia a disminuir Y, aunque puede haber otras causas (al graficar los puntos, se aprecia cierta tendencia a acumularse sobre una recta con pendiente negativa).

5)NO CORRELACION: No se aprecia ninguna relación.

Como se podrá observar en los datos, hay varias variables que solo contienen ceros y unos, por lo que al obtener su gráfica de dispersión no nos es de mucha ayuda.

Razón por la cual se muestran las gráficas de dispersión para las últimas listas de datos, las cuales se van a obtener a continuación.

11.- VARIABLES BINOMIALES, PROPORCIONES Y PROMEDIOS.

A las variables cuyos datos constan de ceros y unos solamente les llamaremos binomiales, al resto no binomiales o multinomiales. Donde el uno indica presencia de algo y el cero la ausencia para las binomiales. El análisis para estas variables y el resto se efectuará usando un método propuesto por el Dr. Rolando Peña en su tesis doctoral, el cual se describirá más adelante.

En nuestra investigación hay varias variables binomiales, como lo son: para la encuesta de alumnos:

AX_3 : Nació en N.L.

AX_4 : Trabaja.

AX_5 : Practica algún deporte.

- AX₇: Cuenta con automóvil.
 AX₈: Le agradan muchos exámenes.
 AX₉: Le es importante la preparación en general del maestro.
 AX₁₁: Le es importante la preparación en Pedagogía del maestro.
 AX₁₃: Tiene una visión de que las materias de la licenciatura le pueden ser útiles para estudios de maestría.
 AX₁₄: Le es significativo aprobar las materias en primera oportunidad.
 AX₁₅: Estudia en casa.
 AX₂₂: A la familia, le gusta lo que está estudiando.

Y para los maestros :

- AY₁₂: Le agrada que sus alumnos le busquen para aclarar dudas.
 AY₁₃: Considera que es necesario tomar cursos de Pedagogía.

Nuestro interés ahora consiste en saber como trabajar con ellas al formar nuestra función de regresión. Algo importante a considerar es que al trabajar con variables binomiales los coeficientes de la función de regresión con este tipo de variables, son los mismos que se obtienen con la regresión para las variables no binomiales.

Las variables binomiales nos van a servir de pivote para llevar a cabo el cruzamiento de datos.

Así procederé a explicar el método propuesto por el doctor Rolando Peña.

Haré la explicación para la encuesta de alumnos, ya que se realizó lo mismo con la de maestros. Ya que entre las variables con categorías en sus respuestas está la edad, considerando que la edad hasta cierto punto, marca la madurez de la persona y de acuerdo con esto la persona adquiere mayor responsabilidad, criterio, metas, etc., esta será nuestra variable de agrupación o variable de cruce. Para los maestros también se tomó la edad como la variable de cruce. Las categorías que presenta la edad (en los alumnos) son las siguientes:

- 1.- menos de 16 años.
- 2.- de 16 a 20 años.
- 3.-de 21 a 25 años.
- 4.-de 26 a 30 años.
- 5.-de 31 a 35 años.
- 6.-de 36 a 40 años.

Así, en nuestros nuevos datos tendremos seis niveles o categorías.

Consideraremos ahora las variables binomiales y contaremos ahora cuantos ceros(o unos) aparecen en cada categoría de la edad y esta cantidad la dividiremos entre el total de elementos pertenecientes a este nivel de la edad

(ceros y unos) y esto lo repetiremos para todos los niveles y para todas las variables binomiales, y a estos nuevos datos los llamaremos proporciones y las denotaremos por PX_j , j = número correspondiente a la variable original.

Así AX_k será la edad (variable de cruce).

k es el número correspondiente a la variable de cruce, en nuestro caso es 1.

$AX_k = \{ 1, 2, \dots, 6 \}$ X_k tiene 6 categorías las cuales corresponden a nuestras nuevas variables.

Ahora procederemos a trabajar con el resto de las variables, es decir, con las que no son binomiales. Cruzaremos cada variable no binomial con una de las binomiales, de preferencia alguna que se relacione, los cruces se efectuaron de la siguiente manera:

variable binomial	con	variable no binomial
PX_{11}	con	PX_{12}
PX_{15}	con	PX_{16}
PX_{15}	con	PX_{17}
PX_8	con	PX_{19}

Una variable de la forma PX_j coincide con una variable AX_j , solo que las PX_j corresponden a proporciones de las originales (AX_j).

Haciendo la aclaración de que se tomaron solo una parte del resto de las variables para mostrar el método, para cuyas repuestas son continuas, se decidió obtener como medida el promedio, ya que es la medida de tendencia central que minimiza la varianza de los datos.

Elijamos una categoría. Obtendremos la suma de los datos en la variable no binomial para los cuales en la binomial hay solo ceros (o unos según lo que se haya considerado desde un principio). Se saca el promedio con esta suma y se repite para todas las categorías. Así nuestras nuevas variables son ahora promedios y las denotaremos por PMX_j .

A estas listas de proporciones y promedios se aplicaron las tablas de contingencia, análisis de regresión y gráficas de dispersión, mostrándose esto en el apéndice E.

12.-ESTABLECIMIENTO DE LAS FUNCIONES DE REGRESION.

Para la encuesta de alumnos, las variables correlacionadas, con una confianza de 95% o más son las siguientes:

$$\text{ECUACION: } PX_{14} = 0.100 + 0.8012PX_{15}$$

$R^2 = 83.08\%$

$N = 350, k = 6$, NIVEL DE SIGNIFICANCIA: 0.03118

PX_{14} : PROPORCION DE ALUMNOS QUE ^{NO} LES INTERESA ACREDITAR EN PRIMERA OPORTUNIDAD.

PX_{15} : PROPORCION DE ALUMNOS QUE ^{NO} ESTUDIA EN CASA.

ECUACION: $PMX_{16} = -0.02557 + 2.237PMX_{17}$

$R^2: 98.46\%$.

$N = 350, K = 6$, NIVEL DE SIGNIFICANCIA: 0.00081

PMX_{16} : PROMEDIO QUE CORRESPONDE A LAS HORAS QUE ESTUDIA POR SEMANA.

PMX_{17} : PROMEDIO QUE CORRESPONDE A LAS HORAS DE T.V. QUE VE AL DIA.

ECUACION: $PX_{15} = 0.061 + 0.93PX_{22}$

$R^2: 88.34\%$

$N = 350, K = 6$, NIVEL DE SIGNIFICANCIA: 0.01752

PX_{15} : PROPORCION QUE CORRESPONDE A SI ^{NO} ESTUDIA EN CASA.

PX_{22} : PROPORCION QUE CORRESPONDE A SI A LA FAMILIA ^{NO} LE GUSTA LO QUE ESTA ESTUDIANDO.

ECUACION: $PX_{13} = 0.0124 + 4.81PX_9$

$R^2: 88.44\%$

$N = 350, K = 6$, NIVEL DE SIGNIFICANCIA: 0.01731

PX_{13} : PROPORCION QUE CORRESPONDE A QUE ^{NO} PIENSA QUE LAS MATERIAS DE LA LICENCIATURA LE PUEDEN SERVIR POSTERIORMENTE AL REALIZAR ESTUDIOS DE MAESTRIA.

PX_9 : PROPORCION QUE CORRESPONDE A QUE ^{NO} LE ES IMPORTANTE LA PREPARACION DEL MAESTRO.

Y PARA LOS MAESTROS:

ECUACION: $PMAY_7 = 0.0024 + 1.47PMAY_3$

$R^2: 91.46\%$

$N = 35, K = 9$, NIVEL DE SIGNIFICANCIA: 0.00005

$PMAY_7$: PROMEDIO QUE CORRESPONDE AL MAYOR TIEMPO CON LA ASIGNACION DE UNA MISMA MATERIA.

$PMAY_3$: PROMEDIO QUE CORRESPONDE A LA ^{NO} ANTIGÜEDAD COMO DOCENTE.

ECUACION: $PMAY_7 = 0.08 + 0.15PAY_{13}$

R²: 43.58 %

N = 35, K = 9, NIVEL DE SIGNIFICANCIA: 0.05298

PMAY₇: PROMEDIO QUE CORRESPONDE AL MAYOR TIEMPO CON LA ASIGNACION DE UNA MISMA MATERIA.

PAY₁₃: PROPORCION QUE CORRESPONDE A SI PIENSA QUE ^{NO} SON NECESARIOS CURSOS DE PEDAGOGIA.

ECUACION: $PMAY_3 = 0.06 + 0.11PAY_{13}$

R²: 54.65 %

N = 35, K = 9, NIVEL DE SIGNIFICANCIA: 0.02284

PMAY₃: PROMEDIO QUE CORRESPONDE A LA ANTIGÜEDAD COMO DOCENTE.

PAY₁₃: PROPORCION QUE CORRESPONDE A QUE PIENSA ^{NO} QUE SON NECESARIOS CURSOS DE PEDAGOGIA.

13.-OBTENCION DE LAS FUNCIONES DE OPTIMIZACION.

Realizando el análisis de regresión múltiple obtenemos las siguientes funciones las cuales establecemos como nuestras funciones de optimización.

Primeramente presentamos nuestra función de regresión:

$$PX_{13} = 0.0008 + 4.38PX_7 + 0.0085PX_3$$

R²: 91.28%, NIVEL DE SIGNIFICANCIA = 0.0871

Con lo cual podemos concluir que con una confianza ^{no} del 92% la proporción de alumnos piensan que las materias de la licenciatura les pueden servir para realizar estudios de maestría está en función de la proporción de alumnos para los cuales ^{no} es importante la preparación del maestro y de la proporción de alumnos que ^{no} son de N.L. Y esto último en el sentido de que al no ser de N.L. implica algunas situaciones económicas.

A continuación presentamos la función de optimización correspondiente:

$$\min Z = \min PX_{13} = 0.67PX_7 + 4.38PX_3 + 0.0008$$

sujeta a :

$$0 \leq PX_7 \leq 1$$

$$0 \leq PX_3 \leq 1$$

Donde al minimizar Z se favorecerían los siguientes aspectos considerados como parte de la calidad del alumno:

- *Que tenga conocimientos de su área.
- *Que sea investigador.
- *Que sea responsable.
- *Que tenga actitud de servicio a la sociedad.

La siguiente función de regresión está dada por:

$$PX_{15} = -0.008 + 0.5PX_{14} + 0.6PX_{22}$$

R²: 92.93%, NIVEL DE SIGNIFICANCIA: 0.0707

Con una confianza del 93%, la proporción de alumnos que estudian en casa está en función de la proporción de alumnos que les interesa acreditar en primera oportunidad y la proporción de alumnos a cuya familia le gusta lo que está estudiando. Pasando a nuestra función de optimización:

$$\min Z = \min PX_{15} = 0.5PX_{14} + 0.6PX_{22} - 0.008$$

sujeta a :

$$0 \leq PX_{14} \leq 1$$

$$0 \leq PX_{22} \leq 1$$

Donde al minimizar Z se favorecerá a los siguientes aspectos considerados como parte de la calidad del alumno:

- * Que sea auto didacta.
- * Que tenga conocimientos de su área.
- * Favorecer a su formación humanística.
- * Que sea responsable.

Para los maestros tenemos las siguientes funciones de regresión:

$$PMAY_7 = 0.15 + 0.67PMAY_3 - 0.45PAY_{13}$$

R²= 91.93%, NIVEL DE SIGNIFICANCIA:0.0005.

Por lo tanto, con una confianza mayor a 95% el promedio de mayores tiempos con la asignación de una misma materia está en función del promedio de la antigüedad como docente del maestro y de la proporción de maestros que piensan que son necesarios cursos de Pedagogía. Así nuestra función de optimización está dada por:

$$\min Z = PMAY_7 = 0.67PMAY_3 - 0.45PAY_{13} + 0.15$$

sujeta a:

$$0 \leq PMAY_3 < \text{INFINITO}$$

$$0 \leq PAY_{13} \leq 1$$

Y por último tenemos la función de regresión:

$$PMAY_3 = -0.099 + 1.28PMAY_7 + 1.29PAY_{13}$$

$R^2 = 93.52\%$, NIVEL DE SIGNIFICANCIA: 0.0003.

Con una confianza mayor al 95% podemos decir que el promedio correspondiente a la antigüedad como docente es una función del promedio de el mayor tiempo que ha tenido con la asignación de una misma materia y de la proporción de la necesidad de cursos de Pedagogía.

Así, la función de optimización quedaría de la siguiente manera:

$$\min Z = \min PMAY_3 = 1.28PMAY_7 + 1.29PAY_{13} - 0.099$$

sujeta a:

$$0 \leq PMAY_7 < \text{INFINITO}$$

$$0 \leq PAY_{13} < 1$$

14.-CONCLUSIONES Y COMENTARIOS.

Las hipótesis establecidas originalmente se pueden corroborar mediante los resultados obtenidos. De tal manera que las variables que están correlacionadas y favorecen a los aspectos que intervienen en la calidad del egresado son:

Si nació en N.L.

Es importante la preparación del maestro.

Piensa que las materias de licenciatura le pueden servir para realizar estudios de maestría.

Si al alumno le es importante acreditar en primera oportunidad.

Estudia en casa.

A su familia le gusta lo que está estudiando.

Poniendo atención en estas variables podemos eficientar nuestro proceso educativo.

Con respecto a los maestros, como ya observamos, los alumnos los toman mucho en cuenta dentro de su aprendizaje, así a lo que tenemos que poner atención en lo que se refiere a los maestros es a:

Antigüedad como docente.

Mayor tiempo con la asignación de una misma materia.

Piensa que son necesarios cursos de Pedagogía.

De lo anterior destacamos la importancia de impartir cursos de Pedagogía

para el personal docente lo cual se puede considerar como un factor que interviene en la calidad del maestro que favorece a la calidad del alumno.

COMENTARIOS:

- * Se sugiere considerar otros aspectos al realizar estudios futuros, como son: coeficiente intelectual del alumno y actitudes psicométricas.
- * Realizar un seguimiento de los egresados para hacer un diagnóstico de la calidad académica de dichos egresados.
- * Una sugerencia es asignar a los maestros jóvenes a los últimos semestres y los maestros antiguos a los primeros, de tal manera que los primeros semestres se vean beneficiados con la experiencia y los maestros más jóvenes se esfuercen más por prepararse.

15.-APORTACIONES

El pasar de una función de regresión en la cual posteriormente se consideren unicamente las variables que son significativas en nuestro proceso a una de optimización en la cual se maximice la calidad del egresado ofrece los siguientes beneficios a la sociedad:

Permite controlar el proceso ya que este análisis nos da a conocer cuales son las variables que realmente son significativas, y por lo tanto las que requieren de mayor atención para llevar a cabo el control.

Disminución de la variabilidad, y centrar el proceso. En el proceso educativo con esto nos referimos a tratar de disminuir los altos índices de reprobación, de ausentismo por parte de los alumnos, como por parte de los maestros y al centrar el proceso trataremos de que disminuyan los altos índices de reprobación, tratando de que las calificaciones se mantengan mas cercanas al promedio.

El poder adaptar planes de estudio en la facultad, concordantes con las características que se desea que tenga el egresado, y con las que espera la sociedad en ese tiempo o contexto coincidiendo mas con sus necesidades.

Una comparación del método científico con el diagrama de Deming lo cual nos ofrece un panorama que antes no se había propuesto.

Contar con un modelo que describa el proceso, para el cual solo se requiera actualizar cada año.

Al efectuar este análisis en la industria, llevar a cabo una capacitación eficiente y de buena calidad que conlleve al logro de sus objetivos.

Se provee de una estructura completa para llevar a cabo investigación social.

Identificar aquellas variables que afectan al proceso educativo y que su control aportaría beneficios al mismo. Tanto al alumno como a la institución y a la sociedad misma.

16.- FUNDAMENTACION MATEMATICA.

16.1.- MATRIZ DE CORRELACION SIMPLE.

Bajo las suposiciones fundamentales al modelo de regresión:

i) Homocedasticidad: Las varianzas son iguales para cada X y puede ser asignada por σ^2

ii) Normalidad: El error ϵ se distribuye normalmente y con media cero.

iii) Independencia: El error es independiente.

Planteamos el siguiente modelo:

$$Y = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \epsilon$$

Centrando los datos, es decir, restando la media Z a los datos esto se transforma en:

$$Y - \bar{Y} = \beta_1(Z_1 - \bar{Z}_1) + \beta_2(Z_2 - \bar{Z}_2) + \epsilon$$

cuando el modelo es escrito de esta forma, la matriz del problema toma la forma:

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{13} & S_{14} \end{bmatrix}$$

donde $S_{ji} = \sum_{i=1}^n (Z_{ji} - \bar{Z}_j)(Z_{li} - \bar{Z}_l)$, $i = 1, 2, \dots, n$; $j, l = 1, 2$.

Los S_{ji} son de diferentes tamaños. Una transformación para centrar los datos es:

$$X_{ji} = (Z_{ji} - \bar{Z}_j) / \sqrt{S_{jj}} \quad j = 1, 2. \quad Y_i = (Y_i - \bar{Y}) / \sqrt{S_{yy}}$$

Una nueva forma del modelo ya centrado, es la siguiente:

$$\begin{aligned} y\sqrt{S_{yy}} &= \beta_1(S_{11}) x_1 + \beta_2(S_{22}) x_2 + \varepsilon \\ \text{ó} \\ y &= \alpha_1 x_1 + \alpha_2 x_2 + \varepsilon' \end{aligned}$$

donde $\alpha_1 = \beta_1 \sqrt{(S_{11} / S_{yy})}$ y $\alpha_2 = \beta_2 \sqrt{(S_{22} / S_{yy})}$ son nuevos coeficientes que serán estimados por la transformación de datos (y_i, x_{1i}, x_{2i} , $i = 1, 2, \dots, n$) y representan escalas de los coeficientes originales β_1 y β_2 . Cuando el modelo es escrito en forma matricial, toma la nueva forma:

$$\begin{bmatrix} 1 & r_{12} \\ r_{21} & 1 \end{bmatrix}$$

llamada matriz de correlación de Z 's, para la cual:

$$r_{12} = S_{12} / \sqrt{(S_{11} S_{22})} = r_{21}$$

r_{12} es la correlación entre Z_1 y Z_2 , y generalizando, podemos escribir:

$$r_{jy} = S_{jy} / \sqrt{(S_{jj} S_{yy})}$$

$$\text{donde } S_{jy} = \sum_{i=1}^n (Z_{ji} - \bar{Z}_j)(Y_i - \bar{Y})$$

esto es, la correlación entre Z_j y Y ($j = 1$ ó 2).

Las ecuaciones normales para este nuevo modelo son:

$$\begin{bmatrix} 1 & r_{12} \\ r_{21} & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} r_{1y} \\ r_{2y} \end{bmatrix}$$

donde a_1 y a_2 son estimadores de α_1 y α_2

Calculando el determinante de la matriz de correlación: $D = 1 - r_{12}^2$

Si r_{12} se acerca mucho a 1 o -1 podemos decir que Z_1 y Z_2 están correlacionadas o que hay dependencia entre ellas.

En general, en un problema de correlación podemos tener la matriz:

$$\begin{bmatrix} 1 & r_{12} & r_{13} & \dots & r_{1p} \\ r_{21} & 1 & r_{23} & \dots & r_{2p} \\ r_{31} & r_{32} & 1 & \dots & r_{3p} \\ \dots & \dots & \dots & \dots & \dots \\ r_{p1} & r_{p2} & r_{p3} & \dots & 1 \end{bmatrix}$$

16.2.-REGRESION LINEAL.

En muchos trabajos experimentales se trata de estudiar cómo el cambio en una variable afecta a otra variable, esto es, dadas dos o más variables, nos interesa saber su relación. Esta variación puede deberse a dos motivos: errores en la medición o variación entre los individuos. Si hay dependencia, podemos encontrar la ecuación que nos especifique esta dependencia.

Si se alcanza a trazar una curva o una recta, a lo que se foirme se le llama curva de regresión. Si es una recta, se tratará de regresión lineal, si se trata de una curva, de regresión no lineal. A la ecuación que relaciona a las variables graficadas (X, Y) se le llama ecuación de regresión. Si dadas dos variables, asumimos que el modelo es lineal, y bajo las condiciones de homocedasticidad, normalidad, e independencia anteriormente mencionadas, la ecuación de regresión está dada por :

$$Y = \beta_0 + \beta_1 X + \epsilon$$

donde β_0 y β_1 son llamados los parámetros del modelo.

Para nuestro análisis requeriremos solo del modelo lineal, el cual explicaremos a continuación.

Las β 's son fáciles de encontrar, ϵ depende de las condiciones para cada observación Y. Si usamos sus estimadores, esto es b_0 y b_1 en lugar de β_0 y β_1 , podemos escribir $\hat{Y} = b_0 + b_1 X$ donde \hat{Y} denota la predicción de Y para cada X cuando se han determinado b_0 y b_1 . b_0 y b_1 denotan estimaciones de los parámetros dados por las letras β_0 y β_1 .

La estimación de b_0 y b_1 es por medio del método de los mínimos cuadrados.

Este método fué descubierto independientemente por Carl Friedrich Gauss (1777-1855) y Adrien Marie Legendre (1752 - 1833). Gauss lo utilizó después de 1803 y Legendre lo publicó en 1805. El descubrimiento del método de mínimos cuadrados trajo gran controversia en aquella época.

Tomemos el modelo lineal. Procedamos a calcular los parámetros de regresión para

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

donde tenemos n conjuntos de observaciones $(X_1, Y_1), \dots, (X_n, Y_n)$. Para $i = 1, \dots, n$ la suma de cuadrados para la línea está dada por:

$$S = \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2$$

Estimaremos los valores para b_0 y b_1 y los sustituiremos por β_0 y β_1 para calcular S.

Diferenciando la ecuación anterior respecto a β_0 y luego respecto a β_1 :

$$\frac{dS}{d\beta_0} = -2 \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)$$

I

$$\frac{dS}{d\beta_1} = -2 \sum_{i=1}^n X_i (Y_i - \beta_0 - \beta_1 X_i)$$

Las estimaciones de b_0 y b_1 están dadas por :

$$\sum_{i=1}^n (Y_i - b_0 - b_1 X_i) = 0$$

$$\sum_{i=1}^n X_i (Y_i - b_0 - b_1 X_i) = 0$$

II

Así tenemos :

$$\sum_{i=1}^n Y_i - nb_0 - b_1 \sum_{i=1}^n X_i = 0$$

$$\sum_{i=1}^n X_i Y_i - b_0 \sum_{i=1}^n X_i - b_1 \sum_{i=1}^n X_i^2 = 0$$

III

De otra manera :

$$nb_0 + b_1 \sum_{i=1}^n X_i = \sum_{i=1}^n Y_i$$

$$b_0 \sum_{i=1}^n X_i + b_1 \sum_{i=1}^n X_i^2 = \sum_{i=1}^n X_i Y_i \quad \text{IIIa.}$$

A las ecuaciones anteriores se les llama normales.

Despejando b_1 de IIIa :

$$b_1 = \{ \sum_{i=1}^n X_i Y_i - [(\sum_{i=1}^n X_i)(\sum_{i=1}^n Y_i)]/n \} / [\sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2 / n]$$

$$b_1 = [\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})] / [\sum_{i=1}^n (X_i - \bar{X})^2]$$

Donde $\bar{Y} = (y_1 + y_2 + \dots + y_n) / n$

$$\bar{X} = (x_1 + x_2 + \dots + x_n) / n$$

$$\begin{aligned} \text{y } \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) &= \sum_{i=1}^n X_i Y_i - \bar{X} \sum_{i=1}^n Y_i - \bar{Y} \sum_{i=1}^n X_i + n\bar{X}\bar{Y} \\ &= \sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y} \\ &= \sum_{i=1}^n X_i Y_i - (\sum_{i=1}^n X_i)(\sum_{i=1}^n Y_i) / n \end{aligned}$$

$\sum_{i=1}^n X_i^2$ es llamada la suma de cuadrados corregida de las X's o S.C. de las desviaciones respecto a la media.

$(\sum_{i=1}^n X_i)^2 / n$ es llamada la corrección por la media de las X's o SC debido a la media.

Similarmente $\sum_{i=1}^n X_i Y_i$ es llamada la suma corregida de productos y $(\sum_{i=1}^n X_i)(\sum_{i=1}^n Y_i) / n$ es llamada la corrección por las medias.

Una notación conveniente y corta es la siguiente:

$$\begin{aligned} S_{xy} &= \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \\ &= \sum_{i=1}^n (X_i - \bar{X})Y_i \\ &= \sum_{i=1}^n X_i(Y_i - \bar{Y}) \\ &= \sum_{i=1}^n X_i Y_i - (\sum_{i=1}^n X_i)(\sum_{i=1}^n Y_i) / n \\ &= \sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y} \end{aligned}$$

Donde todas las formas anteriores son equivalentes. Similarmente podemos escribir :

$$\begin{aligned} S_{xx} &= \sum_{i=1}^n (X_i - \bar{X})^2 \\ &= \sum_{i=1}^n (X_i - \bar{X})X_i \\ &= \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2 / n \\ &= \sum_{i=1}^n X_i^2 - n\bar{X}^2. \end{aligned}$$

$$\begin{aligned} S_{yy} &= \sum_{i=1}^n (Y_i - \bar{Y})^2 \\ &= \sum_{i=1}^n Y_i(Y_i - \bar{Y}) \end{aligned}$$

$$= \sum_{i=1}^n n Y_i^2 - (\sum_{i=1}^n n Y_i)^2 / n$$

$$= \sum_{i=1}^n n Y_i^2 - n \bar{Y}^2.$$

Y finalmente : $b_1 = S_{xy} / S_{xx}$.

La solución para b_0 , al interceptar con $X = 0$ está dada por $b_0 = \bar{Y} - b_1 \bar{X}$.

Sustituyendo en la ecuación anterior:

$$Y^{\wedge} = \bar{Y} + b_1(X - \bar{X}).$$

R^2 está definida como: $\sum_{i=1}^n n (Y_i^{\wedge} - Y)^2 / \sum_{i=1}^n n (Y_i - \bar{Y})^2$.

R^2 es el coeficiente de determinación múltiple.

R^2 adj. = $1 - [\sum_{i=1}^n n (Y_i - \bar{Y}_i)^2 / (n - k - 1)] / [\sum_{i=1}^n n (Y_i - \bar{Y})^2 / (n - 1)]$

R^2 es el coeficiente de determinación múltiple ajustado.

16.3.-REGRESION MULTIPLE.

Describiremos matricialmente el modelo bajo las suposiciones para el modelo de regresión lineal citadas anteriormente.

$$Y = \beta_0 X_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

Aquí Y es nuestra variable predecible o dependiente de X_0, X_1, X_2

Formemos la matriz Y de una columna y n renglones donde n es el número de datos. La matriz X formada por tres columnas: X_0, X_1, X_2 y n renglones. La matriz β formada por $\beta_0, \beta_1, \beta_2$. Y ϵ formada por n renglones y una columna.

La estimación de $\beta_0, \beta_1, \beta_2$ está dada por

$$\beta = (X'X)^{-1} X'Y$$

donde β es un vector de estimaciones de los elementos de la matriz b , siempre que $X'X$ sea no singular. Con lo cual al conocer los coeficientes, completamos el modelo. X' es la matriz transpuesta de X . Mientras que $X'X$ es una matriz no singular, es decir de rango completo.

16.4.- TABLAS DE CONTINGENCIA.

En toda investigación social es utilizado el análisis de contingencia. Principiaremos por decir que una tabla de contingencia es una prueba donde se aplica la X^2 para probar la hipótesis de relación entre dos variables. Las tablas de contingencia nos permiten establecer relaciones entre escalas nominales con

cualquier número de categorías. Aquí H_0 : no existe relación alguna entre las variables cruzadas mediante la tabla.

$$X^2 = \sum (f_o - f_e)^2 / f_e$$

donde f_o es el número de frecuencias observadas y f_e es el número de frecuencias esperadas.

Comparando X^2 obtenida con la de las tablas podremos rechazar H_0 si X^2 calculada es mayor que la de las tablas. Los grados de libertad para consultar X^2 en las tablas se obtienen multiplicando el número de columnas menos uno por el número de renglones menos uno. Los grados de libertad nos indican la libertad aleatoria que se tiene de acomodar los datos respetándolos. En el apéndice D podemos observar las tablas cruzadas y los valores de X^2 para cada pareja de variables cuyos coeficientes de correlación simples indican que tienen correlación. Estas tablas se obtuvieron de las encuestas definitivas (A_x , A_y).

No debe de utilizarse el criterio descrito anteriormente a no ser de que cada una de las frecuencias esperadas sea al menos igual que 5. Esta restricción puede requerir de la combinación de celdas adyacentes, lo cual resulta en una reducción de los grados de libertad.

En una tabla de contingencia 2×2 donde se tiene solo un grado de libertad, se aplica una corrección que recibe el nombre de corrección de Yates para continuidad.

La fórmula corregida se convierte en:

$$X^2 \text{ corregida} = \sum \{(|f_o - f_e| - 0.05)^2\} / f_e$$

Si las frecuencias esperadas de celdas son grandes, los resultados corregidos y sin corregir son casi los mismos. Cuando las frecuencias esperadas están entre 5 y 10 debe de aplicarse la corrección de Yates.

16.5.- TRANSFORMACION DE UNA VARIABLE BINOMIAL A UNA CONTINUA PARA K CATEGORIAS.

Uno de los conceptos mas importantes dentro de las matemáticas es el de función o aplicación de un conjunto en otro. Así podemos definir una función de la siguiente manera: Sean S y T conjuntos; una función o aplicación f de S en T es una regla que asigna a cada elemento $s \in S$ un elemento único $t \in T$. Y explicaremos más a fondo lo que esto significa. Si s es un elemento dado de S , entonces hay un elemento t de T que es asociado a s mediante la aplicación. Cuando s varía en S , t varía en T . Y además a todo elemento $s \in S$ le corresponde un elemento $t \in T$. De otra manera:

$$f : S \rightarrow T$$

$$f(s) = t$$

Si además S y T son conjuntos tales que los podemos considerar como planos, esta aplicación es llamada mapeo o transformación. La imagen de un punto s en el dominio de definición S es el punto $f(s) = t$; el conjunto de imágenes de todos los puntos de un conjunto T, que está contenido en S, recibe el nombre de imagen de T. La imagen del dominio completo de definición S recibe el nombre de rango de f.

En nuestro estudio contamos con dos hiperplanos, a saber: A (de seis dimensiones) y B de j dimensiones y estableceremos un mapeo entre ellos.

Sea el conjunto de categorías de la variable de cruce elegida llamado A.

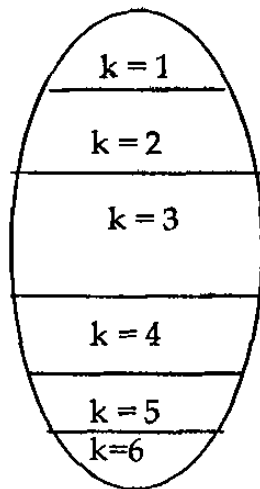
Sea X_j la variable de cruce elegida la cual tiene k categorías.

Sea B el conjunto de variables binomiales (con n variables) con la cantidad de datos respectivos (350 en esta ocasión).

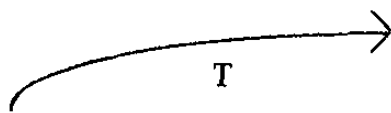
Sea T una aplicación tal que toma una categoría y le asigna a todos los elementos de B su categoría correspondiente, reorganizando a B y formando subconjuntos.

$$T : A \rightarrow B$$

$$T(k_j) = b_{jk}$$



CONJUNTO A
VARIABLE DE CRUCE X_j



X_1	X_2	X_j
0	1		0
0	0		1
.	.		.
.	.		.
.	.		.

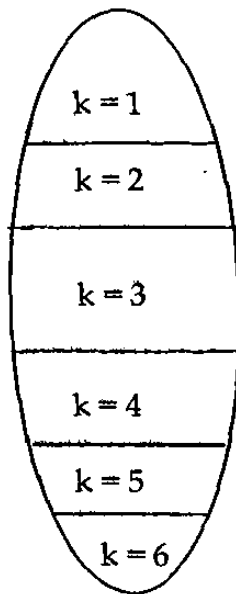
CONJUNTO B
VARIABLES
BINOMIALES

Ahora dividiremos a B en subconjuntos b_{jk} donde j indica el subíndice de la variable y k la categoría a la que pertenece de tal manera que para cada

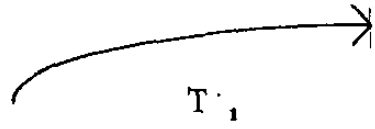
categoría del primer conjunto le va a corresponder un subconjunto de cada variable del conjunto B.

$$T_1: A \rightarrow B$$

$$T_1(k_i) = b_{ij}$$



Conjunto A de las categorías de la variable de cruce.



PX_1	PX_2	$PX_3 \dots$	PX_j
b_{11}	b_{21}	$b_{31} \dots$	b_{j1}
b_{12}	b_{22}	$b_{32} \dots$	b_{j2}
b_{13}	b_{23}	$b_{33} \dots$	b_{j3}
b_{14}	b_{24}	$b_{34} \dots$	b_{j4}
b_{15}	b_{25}	$b_{35} \dots$	b_{j5}
b_{16}	b_{26}	$b_{36} \dots$	b_{j6}

Conjunto B de las variables binomiales.

Ahora contaremos cuantos ceros (o unos) hay para cada subconjunto b_{ij} y se dividirá entre n_{ij} donde n_{ij} es el número de elementos de los subconjuntos b_{ij} obteniendo así lo que llamamos proporciones, y con estas formaremos las nuevas listas de datos continuos.

Sí las variables son no binomiales, requeriremos de una variable de cruce que sea binomial para obtener las nuevas listas de datos, el proceso será similar solo que en lugar de contar los ceros haremos una suma de los elementos correspondientes a cada categoría que coinciden con los ceros (o unos, según lo que se haya elegido) de la variable de cruce en tal categoría, la cual puede ser diferente para cada variable no binomial, y de preferencia debe de estar relacionada y dividiremos entre n_{jk} donde n_{jk} es el total de elementos de X_j (variable no binomial) pertenecientes a la categoría k , obteniendo así el promedio.

Como se hizo mención anteriormente el promedio minimiza la varianza de los datos lo cual se mostrará a continuación:

$$\text{Sea } Q = (1 / (n - 1)) \sum_{i=1}^n (Y_i - k)^2, \quad k = f(Y)$$

$$\delta Q / \delta k = Q' = 0,$$

entonces $k = (\sum_{i=1}^n Y_i) / n = Y$. Donde Y es la media incondicional.

Ahora definiremos una aplicación o mapeo G del conjunto A (de categorías) al conjunto P (de la variable binomial de cruce) de tal manera que a cada $k \in A$ le corresponde un b_j elemento del conjunto formado por la variable binomial elegida como de cruce y también definiremos otra aplicación o mapeo F de este conjunto P al conjunto C el cual está formado por las variables no binomiales de las cuales ya se obtuvo su promedio en la forma en la que se indicó anteriormente.

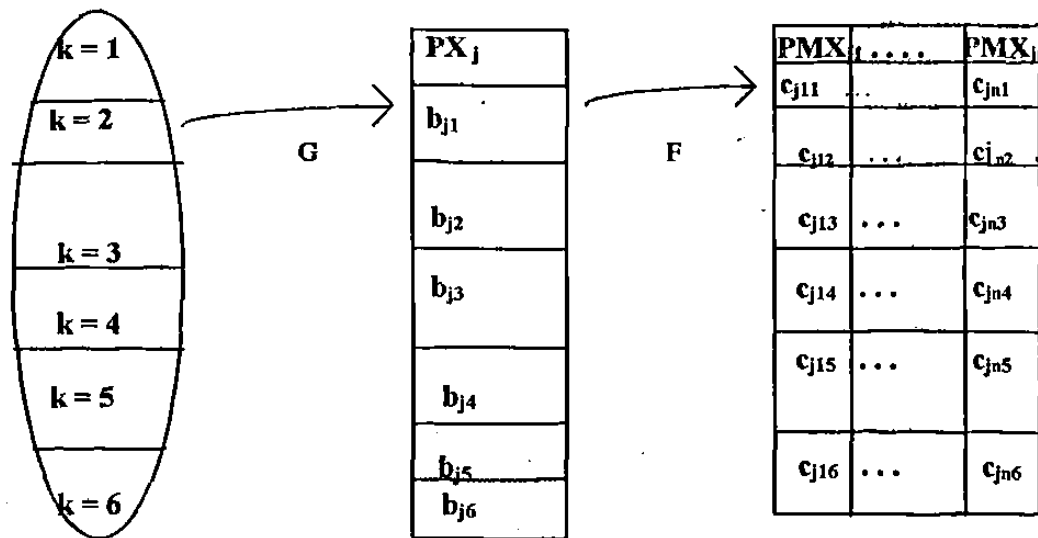
En la siguiente gráfica podemos apreciar las correspondencias ya establecidas.

$$G: A \rightarrow P$$

$$G(k_j) = b_j$$

$$F: P \rightarrow C$$

$$f(b_j) = c_{jnk}$$



Conjunto A de las categorías de la variable de cruce.

Conjunto P
Variable binomial de cruce.

Conjunto C
de variables no binomiales

Así nuestras nuevas variables no binomiales son renombradas como PMX_j , donde j corresponde a la numeración original de la variable. Cada c_{jnk} donde j corresponde a la numeración de la variable original, n corresponde a la reenumeración de las variables no binomiales y k indica la categoría a la que pertenece, forma parte de las nuevas tablas o listas.

Estas nuevas listas se pueden apreciar en el apéndice E.

17.- CONSTRUCCION DE UN MODELO MATEMATICO.

Una vez que se ha identificado un problema para el cual haya que tomar decisiones se procede a formular el modelo matemático que represente la esencia del problema.

Podemos decir que un modelo matemático es una representación idealizada en términos de símbolos y expresiones matemáticas. A las decisiones que se puedan tomar relacionadas unas con otras, las llamaremos variables de decisión (ejem: x_1, x_2, \dots) para las que se deben de determinar los valores respectivos. La medida de efectividad compuesta se expresa entonces como una función matemática de estas variables de decisión y a esta función se le llama función objetivo ($AX_1+BX_2+\dots+ZX_n$). También se representan matemáticamente todas las limitaciones que se puedan imponer sobre las variables de decisión, casi siempre en forma de ecuaciones o desigualdades. Tales expresiones matemáticas de las limitaciones reciben el nombre de restricciones ($0 < X_j < 1$ por ejemplo). Los coeficientes de las variables en las restricciones y en la función objetivo se llaman parámetros del modelo. El modelo matemático puede presentarse entonces como el problema de elegir los valores de las variables de decisión de manera que se maximice la función objetivo, sujeta a las restricciones dadas.

Un modelo matemático presenta algunas ventajas, como son: una descripción del problema en una forma más consisa, así la estructura del problema se vuelve más comprensible y ayuda a revelar las relaciones importantes entre causa y efecto. De esta manera se indica con más claridad que datos son más importantes para el análisis, también facilita el manejo del problema en su totalidad y el estudio de todas sus interrelaciones. Y por último un modelo matemático es un puente para poder emplear técnicas matemáticas poderosas en el análisis del problema.

Sí en el análisis del problema se contemplan más de una función objetivo es necesario transformar y combinar las medidas respectivas en una medida compuesta de efectividad llamada medida global de efectividad.

18.-BIBLIOGRAFIA

N. R. Draper, H. Smith
APPLIED REGRESSION ANALISIS
Second Edición.
Wiley Series in Probability and Mathematical Statics.

Walpole , Myers.
PROBABILIDAD Y ESTADISTICA
Cuarta edición.
Ed. Mc Graw Hill.

Rojas Serrano, Raul
GUIA PARA REALIZAR INVESTIGACIONES SOCIALES
Edición revisada y actualizada.
Plaza y Versalles.

Zorrilla Arerna, Santiago.
INTRODUCCION A LA METODOLOGIA DE LA
INVESTIGACION.
Aguilar Leon y Cal editores.

James R. Newman.
SIGMA Σ EL MUNDO DE LAS MATEMATICAS. TOMO 3.
Editorial grijalbo.

Norman h. Nie, C.Haldai Hull, J.G. Jenkins,
K.Steinbrenner,D.H.Bent.
STATICAL PACKAGE FOR THE SOCIAL SCIENCES.
Segunda edición.
Ed. Mc. Graw-Hill.

Hillier and Liberman .
INTRODUCCION A LA INVESTIGACION DE OPERACIONES.
Quinta edición.
Editorial McGraw-Hill

Bernard Ostle
ESTADISTICA APLICADA
Ed. Limusa Noriega.

I. N. Herstein.
ALGEBRA ABSTRACTA
Grupo Editorial Iberoamerica.

R.V. churchill, J.W.Brown, R.F.Verhey.
VARIABLE COMPLEJA Y SUS APLICACIONES.
Ed. Mc Graw Hill. 2da. edición.