ANALISIS DE RELACION Y CORRELACION

Los análisis se realizarán sobre las variables cuantitativas, ya que Stat Graphics no hace el análisis, estudio y graficas de las variables cualitativas. Por lo tanto se tomaran como variables dependientes e independientes, solo aquellas que se pueden ingresar en el sistema, como numéricas (las de character no pueden ser analizadas).
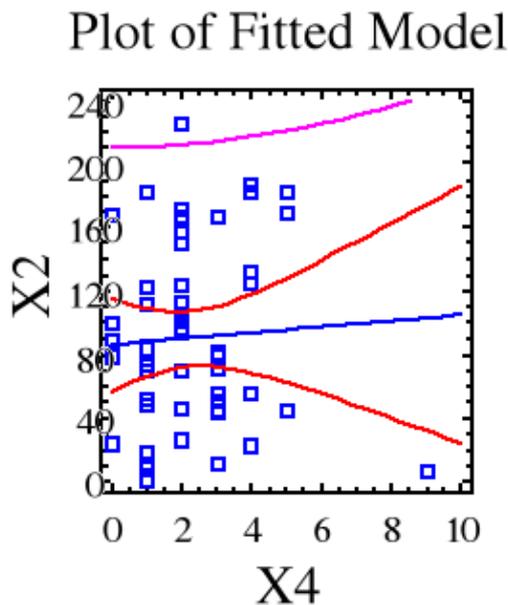
Regresión para pronóstico de corte transversal.

Tipo de Regresión: Simple

- **Variable dependiente: X4 (Número de cursos de educación continua terminados)**
- **Variable Independiente:X2 (Número de horas extra)**

Variable independiente X4 (Eje X)

Variable dependiente X2 (Eje Y)

Ecuación del modelo lineal: X2= 85.7147+1.90585X4

De la cual se deduce: Que el intercepto con el eje Y es 85.7147 y que cada vez que se incrementa en una unidad el número de cursos de educación continua terminados, aumenta en un 1.90585 las horas extras trabajadas.

1.COEFICIENTE DE CORRELACION: r ="R2 = 0.05370561

Como la pendiente es positiva, el coeficiente de correlación también es positivo.

Existe una relación de 0.0536071 entre las variables, la cual es una relación mas bien débil.

2.COEFICIENTE DE DETERMINACIÓN: $R^2 = 0.00288435$ O 0.288435%

El número de cursos de educación continua, determinan el número de horas extra

que trabajan los empleados en un 0.002874, por lo que se vé lo determina muy poco

3. ESTIMACIÓN DE LA VARIANZA DE ERRORES: $S_e^2 = SCE/n-2$ $S_e = 1.69363$

CONTRASTES– INTERVALOS DE CONFIANZA

" "

$Y = B_O + B_1 X_i$

La pendiente mide el grado de inclinación y la relación entre las dos variables.

$S^2 B_1 = S_e^2 / ("X_i^2 - (n * \emptyset X^2))$

$= 0.0000001$

$S_{B1} = 0.000380$

$1- = 95\%$ $= 5\%$

$t\, 0.025;\, 48 = 2.0106$

$LI = 1.90585 - (2.016 * 0.000380) = 1.1.905084$

$LS = 1.90585 + (2.016 * 0.000380) = 1.906616$

$1.905084 < B_1 < 1.906616$

PRUEBAS DE HIPÓTESIS

$H_0 : B = 0$

$H_1 : B " 0$

Rechaza si: $t_c < -t\, /2$ y $t_c > t\, /2$

"

$t_c = B_1 - B_1 / S_{b1}$

$t_c = 1.90585 - 0 / 0.000380$

$t_c = 5013.3787$

$t\, 0.025,\, 48 = 2.0106$

$5013.3787 > 2.0106$

Rechaza H0. Acepta H1

- **Variable independiente: X2 (Número de horas extra) (Eje X)**
- **Variable dependiente: X4 (Número de cursos de educación continua terminados) (Eje Y)**

Ecuación del modelo lineal : **X2**=85.7147+1.90585X4

De la cual se deduce: Que el intercepto con el eje Y es 85.7147 y que cada vez que se incrementa en una unidad las horas extra , aumenta en un 1.90585 el número de cursos de educación continua terminados.

X2 = Horas extra (Variable independiente : X)

X4 = Cursos de educación continua terminados (Variable dependiente : Y)

1.COEFICIENTE DE CORRELACION: : r ="R2

=0.0537061

Como la pendiente es positiva, el coeficiente de correlación también es positivo.

Existe una relación de 0.0537061 entre las variables, la cual es una relación mas bien débil.

2.COEFICIENTE DE DETERMINACIÓN: R2 = 0.00288435

El número de horas extra que trabajan los empleados, determinan la cantidad de cursos de educación continua en un 0.002884, por lo que se vé lo determina muy poco

3. ESTIMACIÓN DE LA VARIANZA DE ERRORES: Se2 = SCE/n−2 Se= 60.1012

CONTRASTES− INTERVALOS DE CONFIANZA

" "

Y= BO+B1Xi

La pendiente mide el grado de inclinación y la relación entre las dos variables.

S2B1=Se2/ ("Xi2−(n*øX2))

=0.283698

SB1=0.532633

1− =95% = 5%

t 0.025; 48 = 2.016

LI= 1.90585 − (2.016*0.532633)= 0.832062

LS= 1.90585 + (2.016*0.532633)=2.979638

0.832062<B1<2.979638

PRUEBAS DE HIPÓTESIS

$H_0 : B = 0$

$H_1: B'' 0$

Rechaza si: $tc < -t /2$ y $tc > t /2$

" "

$tc = B_1 - B_1 / S_{b1}$

$tc = 1.90585 - 0 / 0.532633$

$tc = 3.578167$

$t_{0.025, 48} = 2.0106$

$3.578167 > 2.0106$

Rechaza H0. Acepta H1

ANÁLISIS DE UNA VARIABLE (INDEPENDIENTE)

Stat Graphics unicamente hace el análisis de variables cuantitativas, por eso tomaremos como variables independientes solo tomaremos a las variables cuantitativas.

1.Variable Independiente: X2 (Horas extra)

(Cursos de educación continua...)
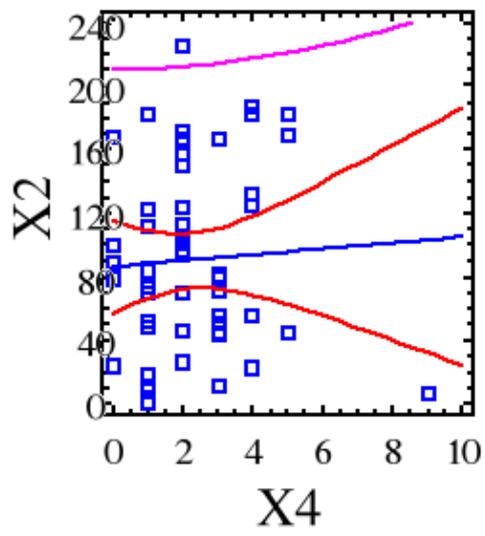
VARIABLE INDEPENDIENTE X4

TRABAJO DE ESTADÍSTICA

FECHA: NOVIEMBRE 3 DEL 2000
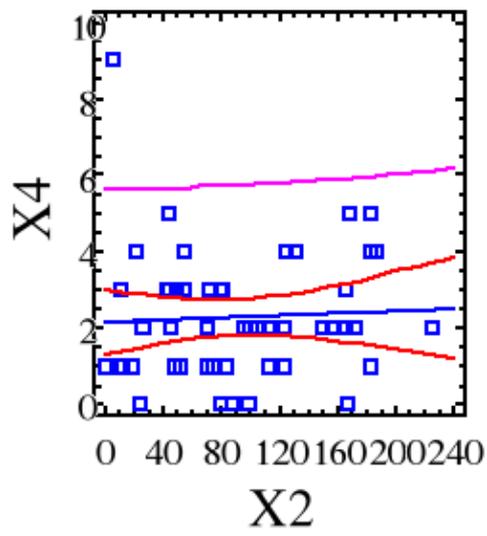
MATERIA: ESTADÍSTICA

UNIVERSIDAD EAFIT

MEDELLIN

Plot of Fitted Model

Plot of Fitted Model

Summary Statistics for X2

Count = 50
Average = 90,06
Variance = 3548,67
Standard deviation = 59,5707
Minimum = 0,0
Maximum = 225,0
Range = 225,0
Stnd. skewness = 1,02431
Stnd. kurtosis = -1,32

The StatAdvisor

    This table shows summary statistics for X2. It includes measures
of central tendency, measures of variability, and measures of shape.
Of particular interest here are the standardized skewness and
standardized kurtosis, which can be used to determine whether the
sample comes from a normal distribution. Values of these statistics
outside the range of -2 to +2 indicate significant departures from
normality, which would tend to invalidate any statistical test
regarding the standard deviation. In this case, the standardized
skewness value is within the range expected for data from a normal
distribution. The standardized kurtosis value is within the range
expected for data from a normal distribution.

Analysis Summary

Data variable: X2

50 values ranging from 0,0 to 225,0

The StatAdvisor

    This procedure is designed to summarize a single sample of data.
It will calculate various statistics and graphs. Also included in the
procedure are confidence intervals and hypothesis tests. Use the
Tabular Options and Graphical Options buttons on the analysis toolbar
to access these different procedures.

Scatterplot for X2

X2
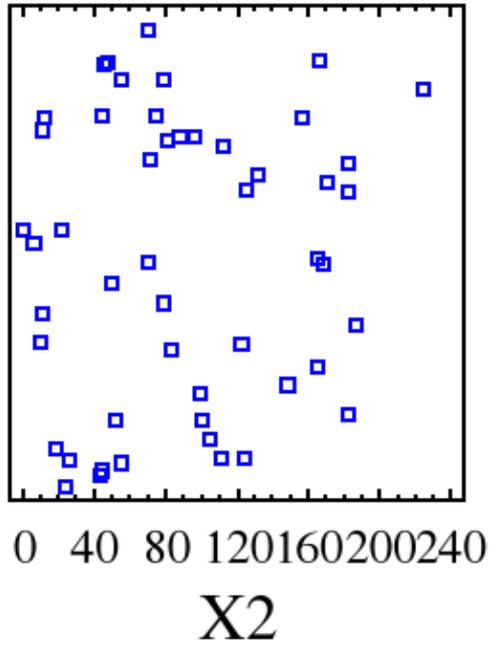
Scatterplot for X4

X4
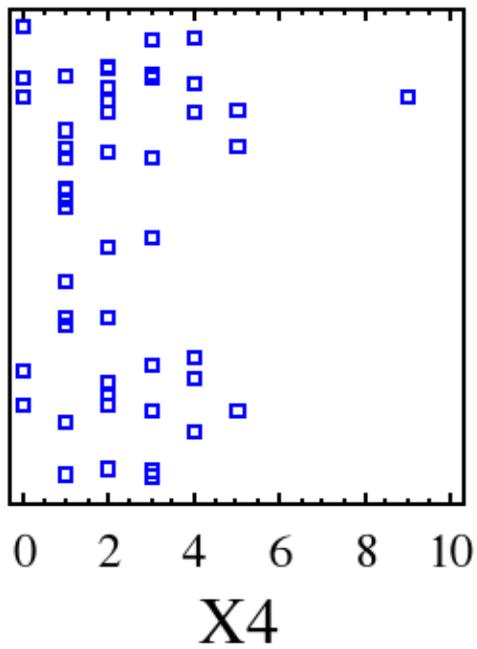
Analysis Summary

Data variable: X4

50 values ranging from 0,0 to 9,0

The StatAdvisor

    This procedure is designed to summarize a single sample of data.
It will calculate various statistics and graphs. Also included in the
procedure are confidence intervals and hypothesis tests. Use the
Tabular Options and Graphical Options buttons on the analysis toolbar
to access these different procedures.

Summary Statistics for X4

Count = 50
Average = 2,28
Variance = 2,81796
Standard deviation = 1,67868
Minimum = 0,0
Maximum = 9,0
Range = 9,0
Stnd. skewness = 4,0335
Stnd. kurtosis = 5,57409

The StatAdvisor

    This table shows summary statistics for X4. It includes measures
of central tendency, measures of variability, and measures of shape.
Of particular interest here are the standardized skewness and
standardized kurtosis, which can be used to determine whether the
sample comes from a normal distribution. Values of these statistics
outside the range of -2 to +2 indicate significant departures from
normality, which would tend to invalidate any statistical test
regarding the standard deviation. In this case, the standardized
skewness value is not within the range expected for data from a normal
distribution. The standardized kurtosis value is not within the range
expected for data from a normal distribution.

```
Regression Analysis - Linear model: Y = a + b*X
----------------------------------------------------------------------
Dependent variable: X2
Independent variable: X4
----------------------------------------------------------------------
                            Standard          T
Parameter       Estimate      Error       Statistic       P-Value
----------------------------------------------------------------------
Intercept       85,7147      14,4303       5,93993         0,0000
Slope           1,90585       5,11467      0,372625        0,7111
----------------------------------------------------------------------


                        Analysis of Variance
----------------------------------------------------------------------
Source         Sum of Squares    Df  Mean Square   F-Ratio    P-Value
----------------------------------------------------------------------
Model              501,544        1    501,544      0,14       0,7111
Residual         173383,0        48   3612,15
----------------------------------------------------------------------
Total (Corr.)    173885,0        49

Correlation Coefficient = 0,0537061
R-squared = 0,288435 percent
Standard Error of Est. = 60,1012


The StatAdvisor
---------------
   The output shows the results of fitting a linear model to describe
the relationship between X2 and X4.  The equation of the fitted model
is

   X2 = 85,7147 + 1,90585*X4

Since the P-value in the ANOVA table is greater or equal to 0.10,
there is not a statistically significant relationship between X2 and
X4 at the 90% or higher confidence level.

   The R-Squared statistic indicates that the model as fitted explains
0,288435% of the variability in X2.  The correlation coefficient
equals 0,0537061, indicating a relatively weak relationship between
the variables.  The standard error of the estimate shows the standard
deviation of the residuals to be 60,1012.  This value can be used to
construct prediction limits for new observations by selecting the
Forecasts option from the text menu.
```

```
Regression Analysis - Linear model: Y = a + b*X
-------------------------------------------------------------------------------
Dependent variable: X4
Independent variable: X2
-------------------------------------------------------------------------------
                               Standard          T
Parameter        Estimate       Error        Statistic        P-Value
-------------------------------------------------------------------------------
Intercept        2,1437        0,43722        4,90302          0,0000
Slope            0,00151342    0,0040615      0,372625         0,7111
-------------------------------------------------------------------------------


                         Analysis of Variance
-------------------------------------------------------------------------------
Source          Sum of Squares   Df  Mean Square   F-Ratio       P-Value
-------------------------------------------------------------------------------
Model            0,39827          1    0,39827       0,14         0,7111
Residual         137,682          48   2,86837
-------------------------------------------------------------------------------
Total (Corr.)    138,08           49

Correlation Coefficient = 0,0537061
R-squared = 0,288435 percent
Standard Error of Est. = 1,69363


The StatAdvisor
---------------
   The output shows the results of fitting a linear model to describe
the relationship between X4 and X2.  The equation of the fitted model
is

   X4 = 2,1437 + 0,00151342*X2

Since the P-value in the ANOVA table is greater or equal to 0.10,
there is not a statistically significant relationship between X4 and
X2 at the 90% or higher confidence level.

   The R-Squared statistic indicates that the model as fitted explains
0,288435% of the variability in X4.  The correlation coefficient
equals 0,0537061, indicating a relatively weak relationship between
the variables.  The standard error of the estimate shows the standard
deviation of the residuals to be 1,69363.  This value can be used to
construct prediction limits for new observations by selecting the
Forecasts option from the text menu.
```