

1. Prueba de bondad del ajuste

1.1. Ventas mensuales Y

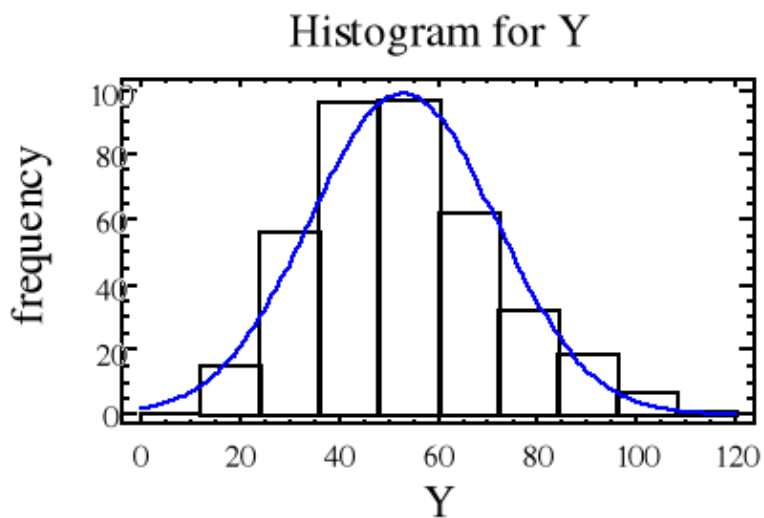
Sea X la variable aleatoria ventas mensuales (millones de dólares)

= X se distribuye $N(53.3127; \sigma^2=18.0734)$

= X no se distribuye $N(53.3127; \sigma^2=18.0734)$

mean = 52,9646

standard deviation = 18,6263



Prueba de bondad del ajuste para Y

Chi-Square Test

Lower Upper Observed Expected

Limit Limit Frequency Frequency Chi-Square

at or below 29,094 28 38,50 2,86

29,094 37,2883 52 38,50 4,73

37,2883 43,1969 49 38,50 2,86

43,1969 48,2457 41 38,50 0,16

48,2457 52,9646 50 38,50 3,43

52,9646 57,6836 32 38,50 1,10

57,6836 62,7323 27 38,50 3,43

62,7323 68,641 27 38,50 3,43

68,641 76,8353 31 38,50 1,46

above 76,8353 48 38,50 2,34

Chi-Square = 25,8309 with 7 d.f. P-Value = 0,000539938

Dado que la prueba de la chi-cuadrado no cumple con un valor necesario para que la hipótesis nula sobreviva, a pesar de que las frecuencias esperadas son mayores de 5, entonces se recurre a las pruebas no paramétricas.

Estimated Kolmogorov statistic DPLUS = 0,0756453

Estimated Kolmogorov statistic DMINUS = 0,0348178

Estimated overall statistic DN = 0,0756453

Approximate P-Value = 0,0244053

Dado que tanto el p-value que nos da la prueba de la chi-cuadrado = 0,000539938 con 7 y el p-value aproximado que nos brinda la Kolmogorov es = 0.0244053 son menores que = 0.04 (valor necesario para que dicha hipótesis () sobreviva); se puede decir que no hay evidencia suficiente para aceptar que la variable ventas mensuales se distribuye normalmente con 53.3127 y σ = 18.0734. Aceptándose entonces la hipótesis alternativa.

1.2. Inversión en capacitación al personal de ventas X1

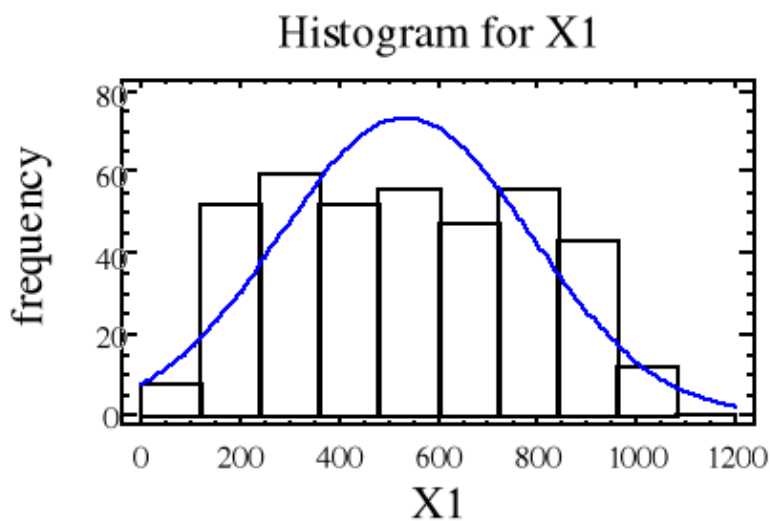
Sea X la variable aleatoria inversión en capacitación (dólares)

= X se distribuye N (544.6893; σ = 257.014)

= X no se distribuye N (544.6893; σ = 257.014)

mean = 534,701

standard deviation = 251,588



Prueba de bondad del ajuste para X1

Chi-Square Test

Lower Upper Observed Expected

Limit Limit Frequency Frequency Chi-Square

at or below 212,277 46 38,50 1,46

212,277 322,959 54 38,50 6,24

322,959 402,768 35 38,50 0,32

402,768 470,962 34 38,50 0,53

470,962 534,701 34 38,50 0,53

534,701 598,441 24 38,50 5,46

598,441 666,635 25 38,50 4,73

666,635 746,444 31 38,50 1,46

746,444 857,126 52 38,50 4,73

857,126 998,0 50 25,88 22,48

above 998,0 0 12,62 12,62

Chi-Square = 60,5559 with 8 d.f. P-Value = 3,62549E-10

Dado que la prueba de la chi-cuadrado no cumple con un valor necesario para que la hipótesis nula sobreviva, a pesar de que las frecuencias esperadas son mayores de 5, entonces se recurre a las pruebas no paramétricas.

Estimated Kolmogorov statistic DPLUS = 0,0703306

Estimated Kolmogorov statistic DMINUS = 0,0703477

Estimated overall statistic DN = 0,0703477

Approximate P-Value = 0,0442712

Dado que el p-value que nos da la prueba de la chi-cuadrado=3.62549E-10 con 8 es menor que nuestro alfa, pero al observar el p-value aproximado que nos brinda la Kolmogorov es = 0.0442712 nos dice que es mayor que = 0.04 (valor necesario para que dicha hipótesis () sobreviva); se puede decir que hay evidencia suficiente para aceptar que la variable inversión en capacitación se distribuye normalmente con 544.689 y $\sigma=538$. Aceptándose entonces la hipótesis nula.

1.3. Numero de automóviles registrados X2

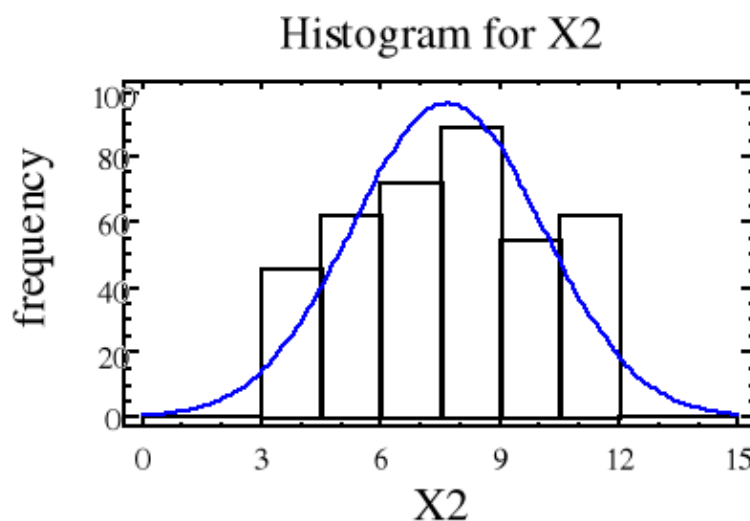
Sea X la variable aleatoria numero de automóviles registrados (millones)

= X se distribuye $N(7.6997; \sigma=2.4365)$

= X no se distribuye $N(7.6997; \sigma=2.4365)$

mean = 7,68447

standard deviation = 2,38604



Prueba de bondad del ajuste para X2

Chi-Square Test

Lower Upper Observed Expected

Limit Limit Frequency Frequency Chi-Square

at or below 4,62662 54 38,50 6,24

4,62662 5,67632 40 38,50 0,06

5,67632 6,43322 35 38,50 0,32

6,43322 7,07997 36 38,50 0,16

7,07997 7,68447 24 38,50 5,46

7,68447 8,28897 28 38,50 2,86

8,28897 8,93571 42 38,50 0,32

8,93571 9,69262 33 38,50 0,79

9,69262 10,7423 40 38,50 0,06

10,7423 12,99 53 33,46 11,41

above 12,99 0 5,04 5,04

Chi-Square = 32,7169 with 8 d.f. P-Value = 0,0000692569

Dado que la prueba de la chi-cuadrado no cumple con un valor necesario para que la hipótesis nula sobreviva, a pesar de que las frecuencias esperadas son mayores de 5, entonces se recurre a las pruebas no paramétricas.

Estimated Kolmogorov statistic DPLUS = 0,0545702

Estimated Kolmogorov statistic DMINUS = 0,0464019

Estimated overall statistic DN = 0,0545702

Approximate P-Value = 0,202032

Dado que el p-value que nos da la prueba de la chi-cuadrado=0.0000692569 con 8 es menor que nuestro alfa, pero al observar el p-value aproximado que nos brinda la Kolmogorov es = 0.202032 nos dice que es mayor que = 0.04 (valor necesario para que dicha hipótesis () sobreviva); se puede decir que hay evidencia suficiente para aceptar que la variable inversión en capacitación se distribuye normalmente con 7.69972 y $\sigma=2.4365$. Aceptándose entonces la hipótesis nula.

1.4. Ingreso personal promedio X3

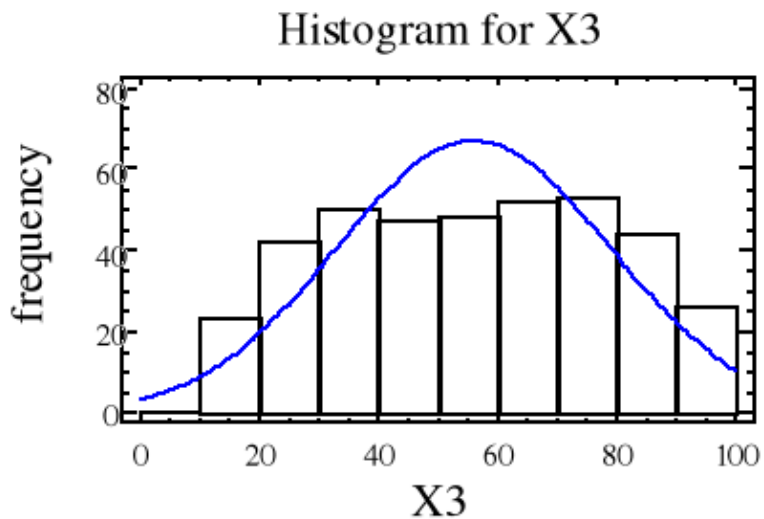
Sea X la variable aleatoria ingreso personal promedio (millones de dólares)

= X se distribuye $N(55.1265; \sigma^2=23.0394)$

= X no se distribuye $N(55.1265; \sigma^2=23.0394)$

mean = 55,974

standard deviation = 22,9769



Prueba de bondad del ajuste para X3

Chi-Square Test

Lower Upper Observed Expected

Limit Limit Frequency Frequency Chi-Square

at or below 26,5279 50 38,50 3,44

26,5279 36,6361 44 38,50 0,79

36,6361 43,9249 48 38,50 2,34

43,9249 50,1529 20 38,50 8,89

50,1529 55,974 29 38,50 2,34

55,974 61,7952 29 38,50 2,34

61,7952 68,0232 30 38,50 1,88

68,0232 75,3119 39 38,50 0,01

75,3119 85,4202 41 38,50 0,16

85,4202 95,8 55 22,51 46,87

above 95,8 0 15,99 15,99

Chi-Square = 85,0485 with 8 d.f. P-Value = 4,66294E-15

Dado que la prueba de la chi-cuadrado no cumple con un valor necesario para que la hipótesis nula sobreviva, a pesar de que las frecuencias esperadas son mayores de 5, entonces se recurre a las pruebas no paramétricas.

Estimated Kolmogorov statistic DPLUS = 0,0739703

Estimated Kolmogorov statistic DMINUS = 0,0667924

Estimated overall statistic DN = 0,0739703

Approximate P-Value = 0,0295997

Dado que tanto el p-value que nos da la prueba de la chi-cuadrado = 4,66294E-15 con 8 y el p-value aproximado que nos brinda la Kolmogorov es = 0.0295997 son menores que = 0.04 (valor necesario para que dicha hipótesis () sobreviva); se puede decir que no hay evidencia suficiente para aceptar que la variable ventas mensuales se distribuye normalmente con 55.1265 y $\sigma = 23.0394$. Aceptándose entonces la hipótesis alternativa.

2. Prueba de bondad del ajuste para ventas mensuales Y por región

2.1. Region 1

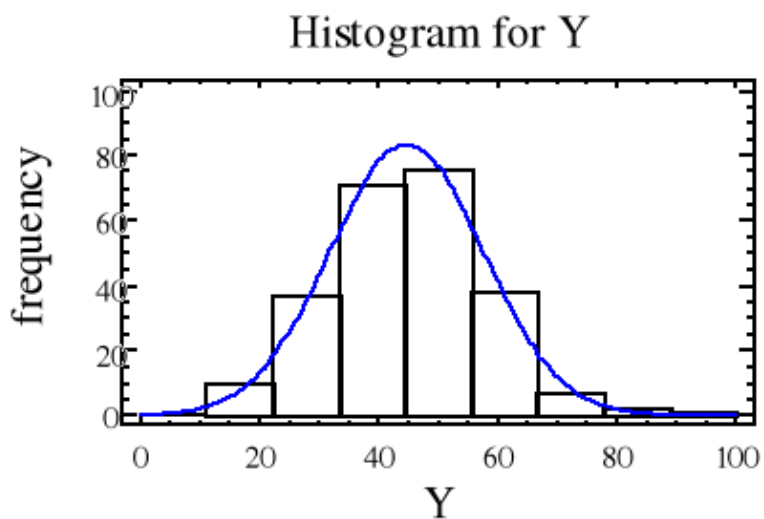
Sea X la variable aleatoria ventas mensuales (millones de dólares) para la región 1

= X se distribuye N

= X no se distribuye N

mean = 44,748

standard deviation = 12,8387



Prueba de bondad del ajuste para Y

Chi-Square Test

Lower Upper Observed Expected

Limit Limit Frequency Frequency Chi-Square

at or below 29,0766 27 26,78 0,00

29,0766 34,9301 30 26,78 0,39

34,9301 39,218 24 26,78 0,29

39,218 42,9543 31 26,78 0,67

42,9543 46,5417 21 26,78 1,25

46,5417 50,278 28 26,78 0,06

50,278 54,5659 27 26,78 0,00

54,5659 60,4195 26 26,78 0,02

above 60,4195 27 26,78 0,00

Chi-Square = 2,67288 with 6 d.f. P-Value = 0,84864

Estimated Kolmogorov statistic DPLUS = 0,0341041

Estimated Kolmogorov statistic DMINUS = 0,0191459

Estimated overall statistic DN = 0,0341041

Approximate P-Value = 0,941947

Dado que tanto el p-value que nos da la prueba de la chi-cuadrado = 0.84864 con 6 y el p-value aproximado que nos brinda la Kolmogorov es = 0.941947 son mayores que = 0.04 (valor necesario para que dicha hipótesis () sobreviva); se puede decir que hay evidencia suficiente para aceptar que la variable ventas mensuales para la región 1 se distribuye normalmente. Esto nos lleva a concluir que las inferencias que realicemos tendrán una alta confiabilidad.

2.2. Region 2

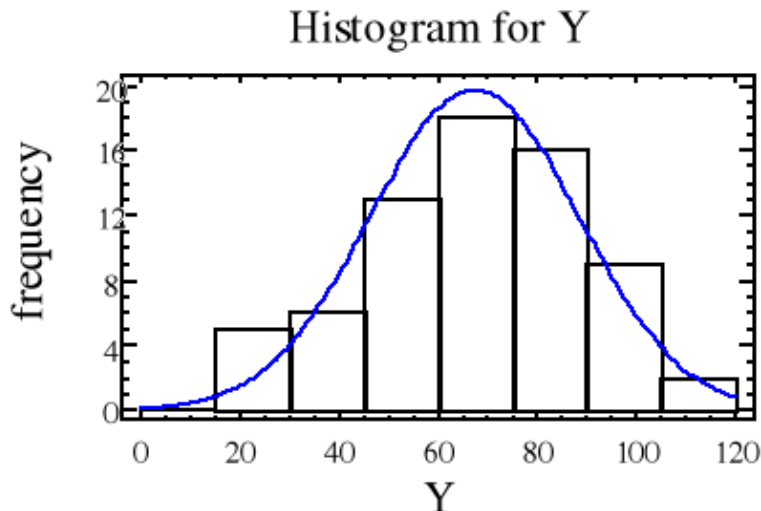
Sea X la variable aleatoria ventas mensuales (millones de dólares) para la región 2

= X se distribuye N

= X no se distribuye N

mean = 67,3334

standard deviation = 20,8828



Prueba de bondad del ajuste para Y

Chi-Square Test

	Lower	Upper	Observed	Expected
--	-------	-------	----------	----------

at or below 43,3108 10 8,62 0,22

43,3108 53,2481 10 8,62 0,22

53,2481 60,6793 4 8,62 2,48

60,6793 67,3334 6 8,63 0,80

67,3334 73,9875 9 8,63 0,02

73,9875 81,4187 14 8,62 3,35

81,4187 91,356 7 8,62 0,31

above 91,356 9 8,62 0,02

Chi-Square = 7,40594 with 5 d.f. P-Value = 0,192157

Estimated Kolmogorov statistic DPLUS = 0,0602213

Estimated Kolmogorov statistic DMINUS = 0,0846546

Estimated overall statistic DN = 0,0846546

Approximate P-Value = 0,705911

Dado que tanto el p-value que nos da la prueba de la chi-cuadrado = 0.192157 con 5 y el p-value aproximado que nos brinda la Kolmogorov es = 0.705911 son mayores que = 0.04 (valor necesario para que dicha hipótesis () sobreviva); se puede decir que hay evidencia suficiente para aceptar que la variable ventas mensuales para la región 2 se distribuye normalmente. Esto nos lleva a concluir que las inferencias que realicemos tendrán una alta confiabilidad.

2.3. Region 3

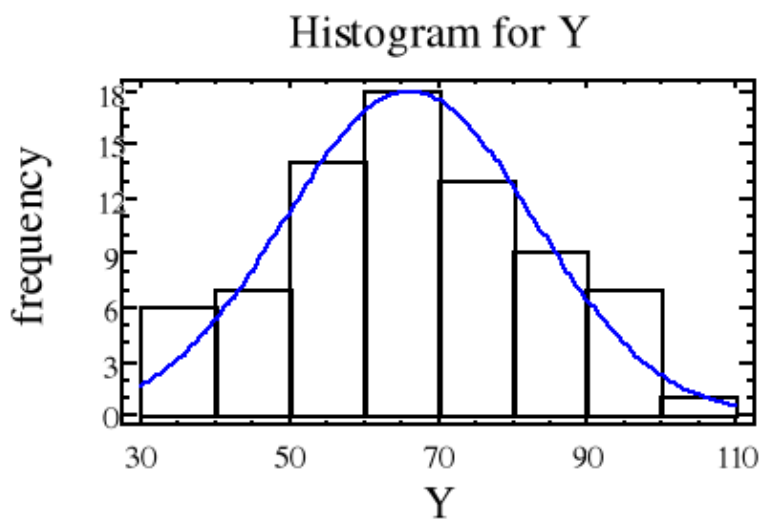
Sea X la variable aleatoria ventas mensuales (millones de dólares) para la región 3

= X se distribuye N

= X no se distribuye N

mean = 66,148

standard deviation = 16,6276



Prueba de bondad del ajuste para Y

Chi-Square Test

Lower Upper Observed Expected

Limit Limit Frequency Frequency Chi-Square

at or below 47,0204 11 9,37 0,28

47,0204 54,9328 9 9,37 0,01

54,9328 60,8498 7 9,37 0,60

60,8498 66,148 9 9,38 0,02

66,148 71,4463 11 9,38 0,28

71,4463 77,3632 10 9,37 0,04

77,3632 85,2757 8 9,37 0,20

above 85,2757 10 9,37 0,04

Chi-Square = 1,47994 with 5 d.f. P-Value = 0,915373

Estimated Kolmogorov statistic DPLUS = 0,036228

Estimated Kolmogorov statistic DMINUS = 0,0498449

Estimated overall statistic DN = 0,0498449

Approximate P-Value = 0,992263

Dado que tanto el p-value que nos da la prueba de la chi-cuadrado = 0.915373 con 5 y el p-value aproximado que nos brinda la Kolmogorov es = 0.992263 son mayores que = 0.04 (valor necesario para que dicha hipótesis () sobreviva); se puede decir que hay evidencia suficiente para aceptar que la variable ventas mensuales para la región 3 se distribuye normalmente. Esto nos lleva a concluir que las inferencias que realicemos tendrán una alta confiabilidad.

3. Prueba de hipótesis para la media

Sea X la variable aleatoria ventas mensuales $X \sim N(\mu, \sigma)$

Estadística descriptiva para Y

Count = 385

Average = 52,9646

Median = 50,4265

Variance = 346,94

Standard deviation = 18,6263

Coeff. of variation = 35,1675%

Analizando con unas ventas mensuales de por lo menos de 55 millones de pesos mensuales. Dado que y promedio = 52.9646 millones de pesos se plantea la hipótesis nula con mayor probabilidad de ser rechazada.

$H_0 : \mu \geq 5.5$

$H_1 : \mu < 5.5$

Prueba de hipótesis

Hypothesis Tests

Sample mean = 52,9646

Sample standard deviation = 18,6263

Sample size = 385

Null Hypothesis: mean = 55,0

Alternative: less than

Computed t statistic = -2,14414

P-Value = 0,0163244

Reject the null hypothesis for alpha = 0,0412.

El p-value es menor que el alfa por lo tanto existe suficiente evidencia para rechazar H .

Esto quiere decir que no hay suficiente evidencia para decir que el ingreso mensual es de por lo menos de 55 millones de pesos mensuales, esto para el sector metalmecánico no es muy recomendable puesto que no tiene bases para pronósticos de ventas, en años posteriores.

4. Prueba de hipótesis para la diferencia de medias y razón de varianzas

4.1. Prueba de hipótesis para la razón de varianzas

Sea Y: la variable aleatoria ventas mensuales de las agencias que tienen un supervisor.

Sea X: la variable aleatoria ventas mensuales de las agencias que tienen dos supervisores.

Ho : $X / Y = 1$

H1 : $X / Y \neq 1$

Prueba de hipótesis para la razón de varianzas

Hypothesis Tests

Sample standard deviations = 18,4547 and 18,7791

Sample sizes = 188 and 197

Null Hypothesis: ratio of variances = 1,0

Alternative: not equal

Computed F statistic = 0,965749

P-Value = 0,810651

Do not reject the null hypothesis for alpha = 0,0412.

ANÁLISIS:

Por medio de la estadística f (calculada) nos damos cuenta que esta es menor que la estadística f (tabulada), por lo tanto caemos en la región de aceptación de la hipótesis nula.

Otra forma de verificación es con el p-value el cual es mayor que el nivel de significancia permitiéndonos corroborar lo dicho, de que las varianzas son iguales.

4.2. Prueba de hipótesis para la diferencia de medias

Sea Y: la variable aleatoria ventas mensuales de las agencias que tienen un supervisor.

Sea X: la variable aleatoria ventas mensuales de las agencias que tienen dos supervisores.

$$H_0 : X - Y = 0$$

$$H_1 : X - Y \neq 0$$

Prueba de hipótesis para la diferencia de medias

Hypothesis Tests

Sample means = 51,8996 and 53,981

Sample standard deviations = 18,4547 and 18,7791

Sample sizes = 188 and 197

Null Hypothesis: difference between means = 0,0

Alternative: not equal

Computed t statistic = -1,09629

P-Value = 0,273642

Do not reject the null hypothesis for alpha = 0,0412.

(Equal variances assumed).

ANÁLISIS: Se observa que la estadística t (calculada) es menor que la estadística t (tabulada), lo que nos indica que el valor calculado cae en la región de aceptación de la hipótesis nula.

El p-value siendo mayor que el nivel de significancia, nos permite terminar de corroborar que las medias son iguales.

Con este resultado es importante que el sector entre a analizar porque las agencias que tienen dos supervisores, en promedio, tienen la misma utilidad que las que tienen uno. Ya que no es rentable tener agencias con mas de un supervisor puesto que no están aumentando el promedio de ventas mensuales.

5. Prueba de hipótesis para la diferencia de proporciones

Tabla de frecuencias cruzada

Row

1 2 3 Total

No Satisfecho| 18 | 16 | 11 | 45

| 7,47% | 23,19% | 14,67% | 11,69%

Satisfechos | 223 | 53 | 64 | 340

| 92,53% | 76,81% | 85,33% | 88,31%

Column 241 69 75 385

Total 62,60% 17,92% 19,48% 100,00%

Cell contents:

Observed frequency

Percentage of column

Prueba de hipotesis

Hypothesis Tests

Sample proportions = 0,9253 and 0,7681

Sample sizes = 241 and 69

Null Hypothesis: difference between proportions = 0,0

Alternative: not equal

Computed z statistic = 3,68427

P-Value = 0,000229424

Reject the null hypothesis for alpha = 0,0412.

Warning: normal approximation may not be appropriate for small sample sizes.

Dado que el p-value = 0.000229424 es menor que el nivel de significancia, la hipótesis nula cae en la región de rechazo. Por la tanto hay evidencia suficiente para rechazar la hipótesis nula y se acepta la hipótesis alternativa; de que las proporciones son diferentes.

Esto quiere decir que la proporción los supervisores que están satisfechos con las condiciones laborales en la región 1 es diferente a la proporción de los supervisores que esta satisfechos con las condiciones

laborales en la región 2; se tendría que entrar a analizar porque las regiones no tiene el mismo promedio de ventas mensuales.

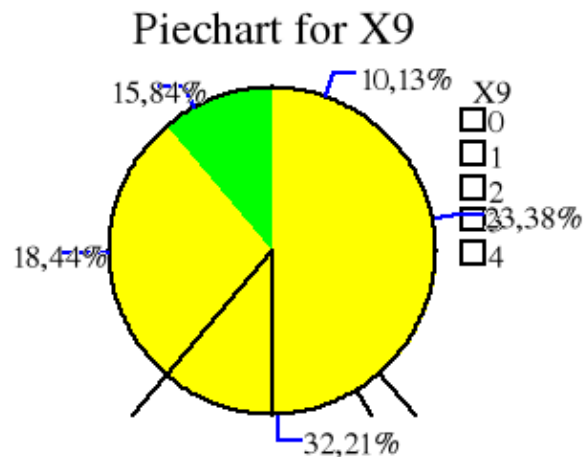
6. Prueba de hipótesis para proporción

Sea X la variable aleatoria: opinión sobre el estado de las vías.

X binomial. Como $n > 30$ podemos decir que $X \sim N$

$H_0 : p = 0.5$

$H_1 : p < 0.5$



Prueba de hipótesis

Hypothesis Tests

Sample proportion = 0,1584

Sample size = 385

Null Hypothesis: proportion = 0,5

Alternative: less than

P-Value = 0,0

Reject the null hypothesis for $\alpha = 0,0412$.

Si comparamos el p-value (0.0) con el nivel de significancia 0.04, vemos que el p-value es menor y llegamos a la conclusión de rechazar la hipótesis de que la calificación excelente del estado de las vías es menor que el 50% .

7. Prueba de independencia

Sea X: la variable aleatoria de cumplimiento del presupuesto

Sea Y: la variable aleatoria condiciones de las vías

Ho: Existe independencia entre el cumplimiento del presupuesto con respecto a las condiciones de las vías.

H1: Existe dependencia entre el cumplimiento del presupuesto con respecto a las condiciones de las vías.

Tabla de frecuencias cruzadaFrequency

Row

0 1 2 3 4 Total

0 | 15 | 30 | 71 | 26 | 32 | 174

| 17,63 | 40,68 | 56,04 | 32,09 | 27,57 | 45,19%

| 0,39 | 2,80 | 3,99 | 1,16 | 0,71 |

1 | 24 | 60 | 53 | 45 | 29 | 211

| 21,37 | 49,32 | 67,96 | 38,91 | 33,43 | 54,81%

| 0,32 | 2,31 | 3,29 | 0,95 | 0,59 |

Column 39 90 124 71 61 385

Total 10,13% 23,38% 32,21% 18,44% 15,84% 100,00%

Cell contents:

Observed frequency

Expected frequency

Contribution to chi-squared

Prueba Chi Cuadrado

Chi-Square Df P-Value

16,52 4 0,0024

ANÁLISIS: Debido a que el p value de la Chi cuadrada nos da menor que nuestro alfa se rechaza la hipótesis nula, es decir que apoyados en la evidencia muestral se concluye que no hay evidencia suficiente para aceptar la hipótesis nula y con esto la idea de que las condiciones de las vías y el cumplimiento del presupuesto son variables dependientes.

Ósea que el cambio de una afectara a la otra ya que existe una relación.

Se aconsejaría a la empresa que a la hora de analizar el cumplimiento del presupuesto, para mejorarlo o para ver que factores influyen directamente en el aumento o la disminución de este, hacer relación con el estado de las vías ya que la prueba estadística nos muestra evidencia suficiente para rechazar su independencia.

8. Prueba Anova de un solo camino

Sea X la variable aleatoria ventas mensuales de acuerdo con la calificación del estado de las vías.

H1: Existe al menos uno diferente

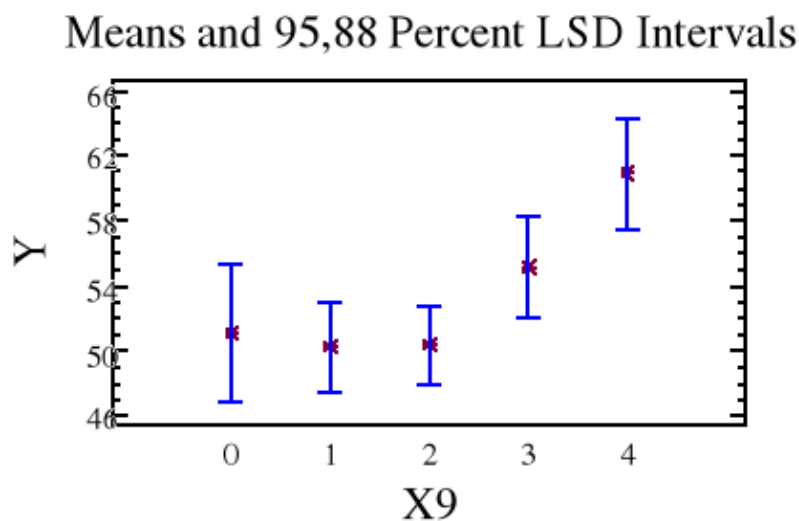


Tabla Anova Y by X9

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
--------	----------------	----	-------------	---------	---------

Between groups	5797,04	4	1449,26	4,32	0,0020
----------------	---------	---	---------	------	--------

Within groups 127428,0 380 335,336

Total (Corr.) 133225,0 384

Table of Means for Y by X9

with 95,88 percent LSD intervals

Std. error

X9 Count Mean (pooled s) Lower limit Upper limit

0 39 51,0709 2,9323 46,8235 55,3184

1 90 50,2203 1,93027 47,4243 53,0163

2 124 50,3925 1,64448 48,0105 52,7746

3 71 55,1872 2,17326 52,0392 58,3351

4 61 60,866 2,34464 57,4698 64,2622

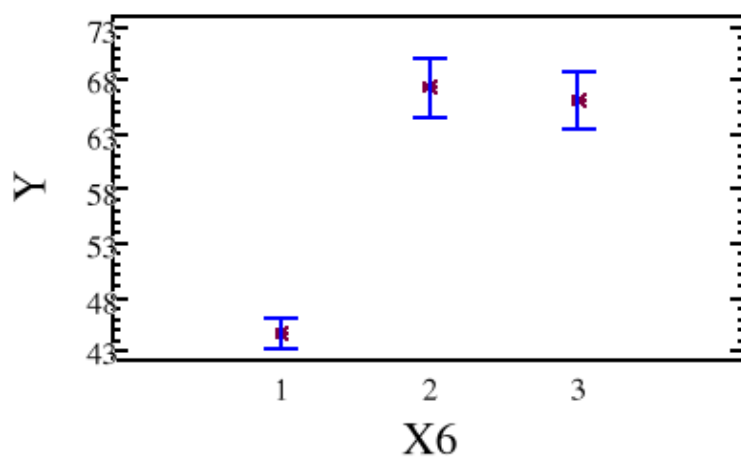
Total 385 52,9646

Como el p value de la prueba ANOVA es menor que el alfa, concluimos que hay al menos una media diferente entre las clases. En la tabla anterior se observa que las ventas mensuales son mayores cuando la opinión sobre el estado de las vías es la mejor, mientras que las mas bajas ventas mensuales se presentan cuando la opinión sobre las vías es la peor, de los resultados de la tabla también podemos concluir que las ventas mensuales no tienen una relación directa (lógica) con la opinión sobre el estado de las vías ya que se esperaría que mientras peor fuera el estado de las vías aumentarían mas las ventas mensuales. Cabe anotar que la desviación es mayor en aquellas opiniones que poseen un menor tamaño de muestra, por ejemplo cuando se observa la desviación de la respuesta no sabe/no responde que corresponde al valor 0, la desviación es la mayor de todas las respuestas sobre el estado de las vías. Por lo que puede que en algunos de los resultados obtenidos no se tenga suficiente certeza de que dan los resultados reales acerca de la prueba que se esta realizando. Por lo que una de las recomendaciones que se le podría dar a la compañía es que aumente el tamaño de la muestra con el fin de que la confianza en los resultados sea mayor.

9. Variables que estratifican a la variable ventas mensuales Y

9.1. Region X6

Means and 95,88 Percent LSD Intervals



ANOVA Table for Y by X6

Analysis of Variance

Source Sum of Squares Df Mean Square F-Ratio P-Value

Between groups 43551,5 2 21775,8 92,76 0,0000

Within groups 89673,3 382 234,747

Total (Corr.) 133225,0 384

Table of Means for Y by X6

with 95,88 percent LSD intervals

Std. Error

X6 Count Mean (pooled s) Lower limit Upper limit

1 241 44,748 0,986942 43,3185 46,1776

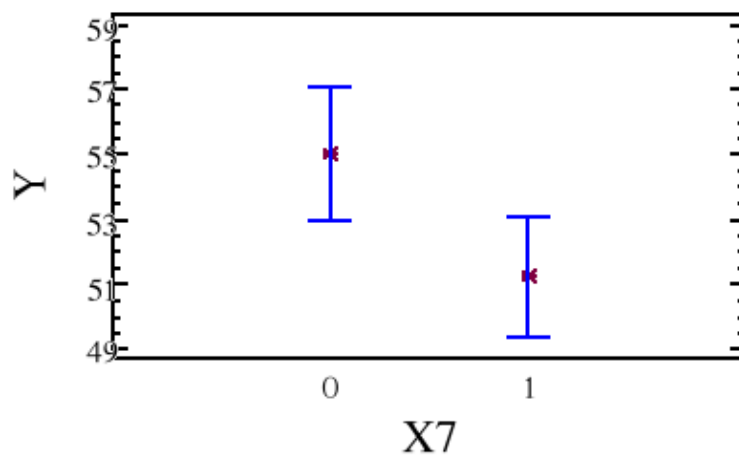
2 69 67,3334 1,84449 64,6617 70,0051

3 75 66,148 1,76917 63,5855 68,7106

Total 385 52,9646

9.2. Cumplimiento presupuesto de ventas X7

Means and 95,88 Percent LSD Intervals



ANOVA Table for Y by X7

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
--------	----------------	----	-------------	---------	---------

Between groups	1362,81	1	1362,81	3,96	0,0473
----------------	---------	---	---------	------	--------

Within groups	131862,0	383	344,287		
---------------	----------	-----	---------	--	--

Total (Corr.)	133225,0	384			
---------------	----------	-----	--	--	--

Table of Means for Y by X7

with 95,88 percent LSD intervals

Std. error

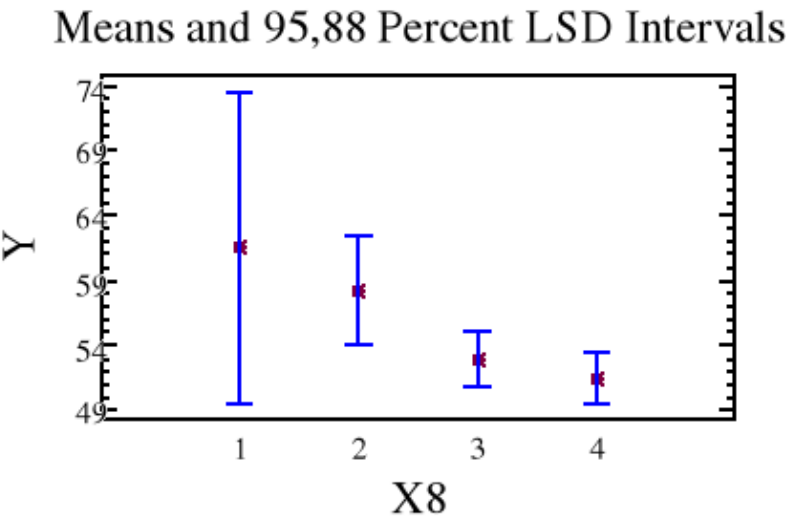
X7	Count	Mean (pooled s)	Lower limit	Upper limit
----	-------	-----------------	-------------	-------------

0	174	55,0365	1,40665	52,999 57,0739
---	-----	---------	---------	----------------

1	211	51,2561	1,27738	49,4059 53,1063
---	-----	---------	---------	-----------------

Total 385 52,9646

9.3. Calificacion de condiciones laborales X8



ANOVA Table for Y by X8

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
--------	----------------	----	-------------	---------	---------

Between groups	1834,25	3	611,417	1,77	0,1518
----------------	---------	---	---------	------	--------

Within groups	131391,0	381	344,857		
---------------	----------	-----	---------	--	--

Total (Corr.)	133225,0	384			
---------------	----------	-----	--	--	--

Table of Means for Y by X8

with 95,88 percent LSD intervals

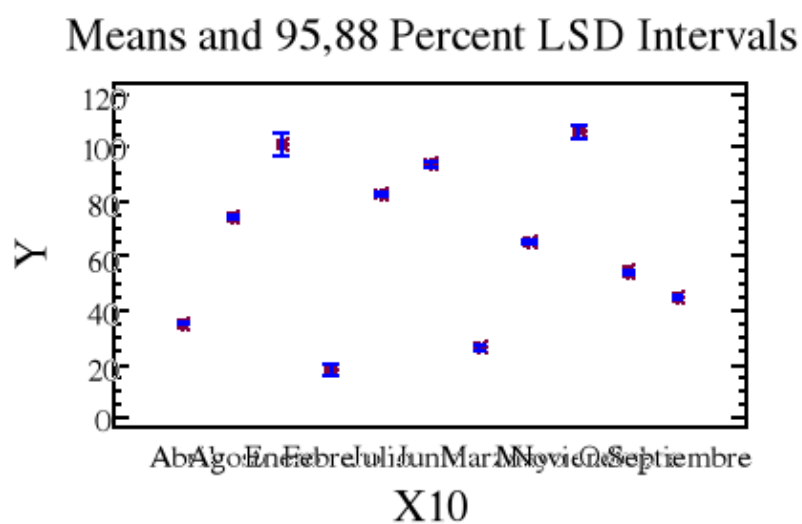
Std. error

X8	Count	Mean (pooled s)	Lower limit	Upper limit
----	-------	-----------------	-------------	-------------

1	5	61,4849	8,3049	49,4554	73,5144
2	40	58,2191	2,93623	53,966	62,4721
3	165	52,9684	1,4457	50,8743	55,0624
4	175	51,5166	1,40379	49,4833	53,55

Total 385 52,9646

9.4. Mes X10



ANOVA Table for Y by X10

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
--------	----------------	----	-------------	---------	---------

Between groups	129883,0	10	12988,3	1453,71	0,0000
----------------	----------	----	---------	---------	--------

Within groups	3341,54	374	8,93461		
---------------	---------	-----	---------	--	--

Total (Corr.)	133225,0	384			
---------------	----------	-----	--	--	--

Table of Means for Y by X10

with 95,88 percent LSD intervals

Std. error

X10 Count Mean (pooled s) Lower limit Upper limit

Abril 63 35,0354 0,376589 34,4899 35,5809

Agosto 33 74,1304 0,520332 73,3767 74,8842

Enero 1 101,145 2,98908 96,8148 105,475

Febrero 5 18,1994 1,33676 16,263 20,1358

Julio 19 83,1014 0,685742 82,108 84,0947

Junio 16 93,5908 0,74727 92,5084 94,6733

Marzo 29 25,866 0,555058 25,0619 26,67

Mayo 49 64,9688 0,427012 64,3503 65,5874

Noviembre 3 105,835 1,72575 103,336 108,335

Octubre 77 54,316 0,340637 53,8226 54,8095

Septiembre 90 44,8433 0,315077 44,3869 45,2997

Total 385 52,9646

De los anteriores resultados se puede decir que las variables que menos se traslapan al ser comparadas con las ventas mensuales son la region y el cumplimiento del presupuesto. de la prueba anova de las ventas mensuales por región podemos observar que existe un traslape entre las regiones 2 y 3, al mostrar que existe homogeneidad entre estos grupos, lo que quiere decir que no hay diferencia significativa entre las medias de las ventas mensuales con respecto a estas dos regiones, entre la región 1 y las regiones 2 y 3 no se presenta ningún tipo de traslape, al mostrar que la homogeneidad entre estos grupos es muy baja, por lo tanto podemos decir que existe una diferencia significativa entre la media de las ventas mensuales de la región 1 y las medias de las ventas mensuales de la región 2 y 3. Partiendo de esta prueba podemos decir que la variable región tiene ciertas posibilidades en comparación con otras variables de estratificar a la variable ventas mensuales debido a que existe poco traslape entre las curvas.

Pero al observar la prueba Anova para las ventas mensuales por cumplimiento del presupuesto podemos observar que las curvas no se traslapan, al mostrar que no existe prácticamente homogeneidad entre estos grupos, lo que nos permite decir que existe una diferencia significativa entre las medias de las ventas

mensuales con respecto al cumplimiento del presupuesto de ventas. Partiendo de esta prueba se puede decir que la variable que mejor estratifica las ventas mensuales es el cumplimiento del presupuesto de ventas, debido a que entre las curvas no existe traslape alguno.

10. Regresión Lineal Simple

Matriz de correlación

Y X1 X2 X3 X4

Y -0,0403 0,1997 0,5323 0,0407

(385) (385) (385) (385)

0,4301 0,0001 0,0000 0,4258

X1 -0,0403 -0,0637 -0,0539 -0,0988

(385) (385) (385) (385)

0,4301 0,2124 0,2911 0,0529

X2 0,1997 -0,0637 -0,0045 0,0510

(385) (385) (385) (385)

0,0001 0,2124 0,9295 0,3185

X3 0,5323 -0,0539 -0,0045 -0,0333

(385) (385) (385) (385)

0,0000 0,2911 0,9295 0,5153

X4 0,0407 -0,0988 0,0510 -0,0333

(385) (385) (385) (385)

0,4258 0,0529 0,3185 0,5153

X5 0,0559 -0,0587 0,0568 0,0555 -0,0554

(385) (385) (385) (385) (385)

0,2736 0,2508 0,2664 0,2777 0,2782

X5

Y 0,0559

(385)

0,2736

X1 -0,0587

(385)

0,2508

X2 0,0568

(385)

0,2664

X3 0,0555

(385)

0,2777

X4 -0,0554

(385)

0,2782

X5

Correlation

(Sample Size)

P-Value

Al analizar la matriz de correlación, se concluye que la variable que mayor predice es la variable ingreso personal promedio X3, puesto que posee el menor p value y la mayor correlación. Con esta seguiremos realizando los demás puntos.

Regresión Simple– Y vs. X3

Regression Analysis – Linear model: $Y = a + b \cdot X$

Dependent variable: Y

Independent variable: X3

Standard T

Parameter Estimate Error Statistic P-Value

Intercept 28,8129 2,12141 13,582 0,0000

Slope 0,43148 0,0350674 12,3043 0,0000

Correlation Coefficient = 0,532262

R-squared = 28,3303 percent

R-squared (adjusted for d.f.) = 28,1432 percent

Standard Error of Est. = 15,7892

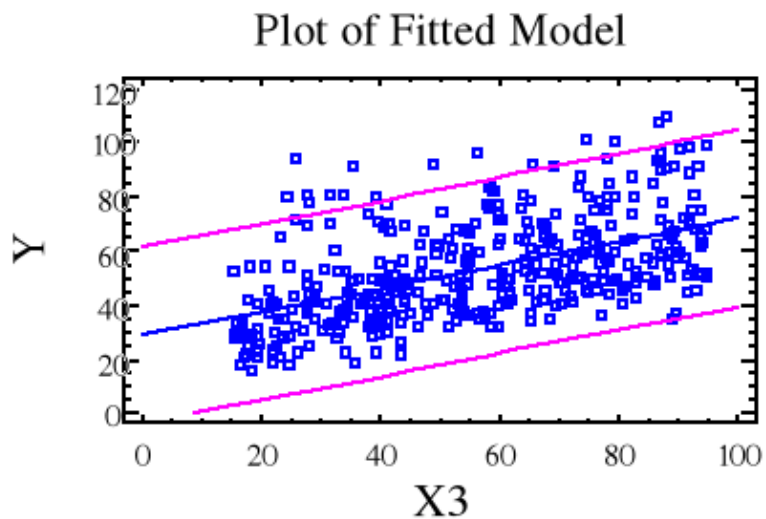
Mean absolute error = 12,5493

Durbin-Watson statistic = 1,16673 (P=0,0000)

Lag 1 residual autocorrelation = 0,404786

Al analizar la relación existente entre las variables ingreso personal y ventas promedio, el coeficiente de correlación que ha sido calculado por Statgraphics para estas dos variables $r = 0.532262$. Cuando r se encuentra entre 0.30 y 0.60 se dice que existe una mala linealidad entre las dos variables.

En la grafica que se presenta a continuación también nos podemos dar cuenta que los puntos en ella no siguen ningún patrón. Una razón mas para afirmar que la linealidad es mala. Puede verse utilizando el modelo lineal la predicción se va a corregir en 28,3303% o se deja de cometer un error del 28,3303 cuando se utiliza este modelo.



Prueba de bondad del ajuste para los errores

Prueba de bondad del ajuste para los residuals de la regression lineal

Prueba Chi-cuadrado

Limite Limite Frecuencia Frecuencia

inferior superior observada esperada Chi-cuadrado

at or below -20.2324 13 38.40 16.80

-20.2324 -13.287 63 38.40 15.76

-13.287 -8.27895 71 38.40 27.68

-8.27895 -3.99972 39 38.40 0.01

-3.99972 1.22399E-10 37 38.40 0.05

1.22399E-10 3.99972 31 38.40 1.43

3.99972 8.27895 34 38.40 0.50

8.27895 13.287 29 38.40 2.30

13.287 20.2324 19 38.40 9.80

above 20.2324 48 38.40 2.40

Chi-cuadrado = 76.7302 con 7 g.l. P-Value = 6.38378E-14

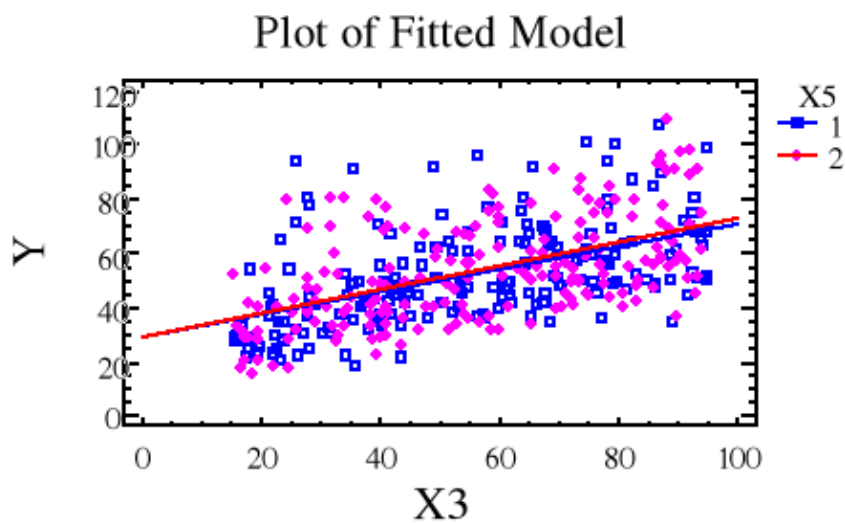
Kolmogorov estadístico estimado DPLUS = 0.0951313

Kolmogorov estadístico estimado DMINUS = 0.0691041

Estimated overall statistic DN = 0.0951313

P-Value aproximado= 0.00191655

11. Comparacion de modelos lineales



Regression Simple – Y vs. X3 (X5=1)

Regression Analysis – Linear model: $Y = a + b \cdot X$

Dependent variable: Y

Independent variable: X3

Selection variable: X5=1

Standard T

Parameter Estimate Error Statistic P-Value

Intercept 28,9444 2,99587 9,66146 0,0000

Slope 0,419875 0,0505785 8,30146 0,0000

Correlation Coefficient = 0,519945

R-squared = 27,0343 percent

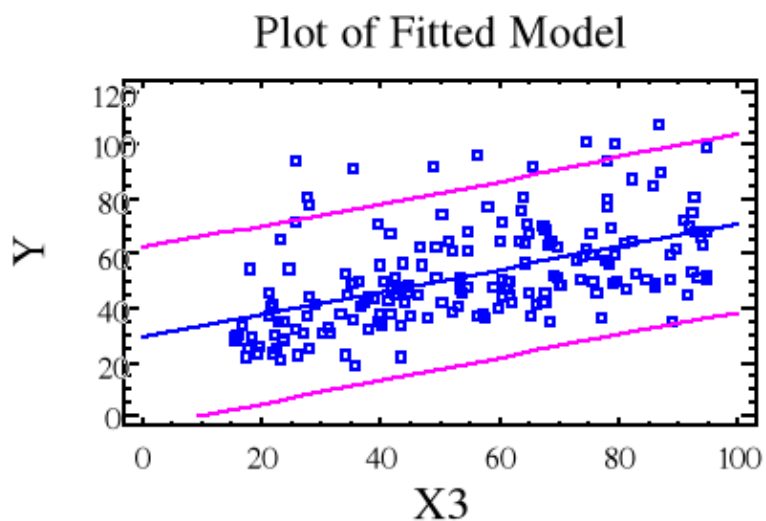
R-squared (adjusted for d.f.) = 26,642 percent

Standard Error of Est. = 15,8063

Mean absolute error = 12,073

Durbin-Watson statistic = 1,22583 (P=0,0000)

Lag 1 residual autocorrelation = 0,362988



Regrsion simple – Y vs. X3 (X5=2)

Regression Analysis – Linear model: $Y = a + b \cdot X$

Dependent variable: Y

Independent variable: X3

Selection variable: X5=2

Standard T

Parameter Estimate Error Statistic P-Value

Intercept 28,8039 3,02211 9,53107 0,0000

Slope 0,440027 0,0489994 8,98025 0,0000

Correlation Coefficient = 0,540896

R-squared = 29,2568 percent

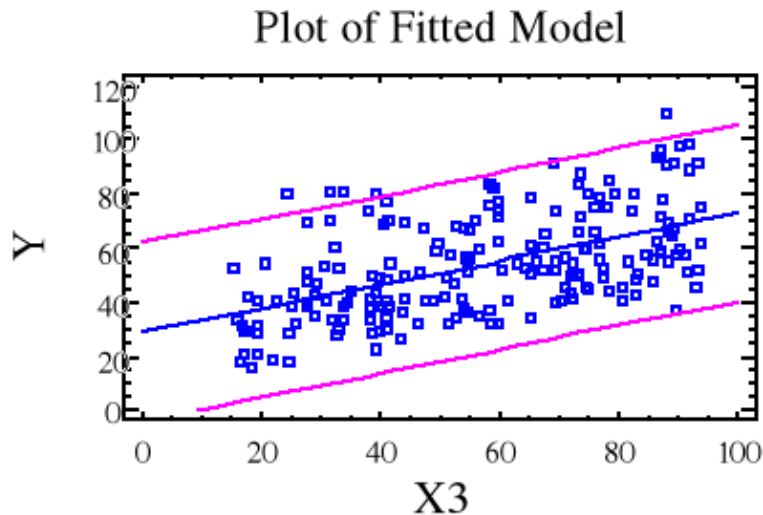
R-squared (adjusted for d.f.) = 28,894 percent

Standard Error of Est. = 15,8354

Mean absolute error = 12,9793

Durbin-Watson statistic = 1,33716 (P=0,0000)

Lag 1 residual autocorrelation = 0,322907



Con los resultados obtenidos podemos determinar que el modelo de predicción de las ventas mensuales promedio basado en el numero de supervisores es mejor para el caso cuando son dos puesto que tiene un mayor porcentaje de corrección de errores. Además en la gráfica podemos observar que las observaciones en el caso cuando es un supervisor, están más centradas alrededor de la recta que en el caso cuando son uno.

Debemos realizar las siguientes pruebas de hipótesis para verificar que tan buen modelo es este:

Prueba de hipótesis para el intercepto

Como el valor de esta prueba (0) es menor que el con el que venimos trabajando, la hipótesis nula es rechazada. Lo que nos dice que el intercepto es diferente a cero (0) y que probablemente hay otras variables que deban estar incluidas dentro del modelo.

Prueba de hipótesis para la pendiente

Como el valor de esta prueba es igual a cero (0) y menor que el , esta hipótesis debe ser rechazada, por lo tanto la pendiente es diferente de 0. Esto quiere decir que la variable ingreso personal promedio de las agencias que tienen dos supervisores si debe estar incluida en este modelo y además que por cada unidad de ingreso personal promedio cuando $X_5=2$ que se incremente, el valor de las ventas mensuales promedio aumenta en 0,440027 unidades.

12. Regresion Multiple

Regresion multiple – Y

Multiple Regression Analysis

Dependent variable: Y

Standard T

Parameter Estimate Error Statistic P-Value

CONSTANT 10,5693 6,25538 1,68964 0,0919

X1 0,000530975 0,00315158 0,168479 0,8663

X2 1,55338 0,330533 4,69962 0,0000

X3 0,433054 0,0342848 12,6311 0,0000

X4 1,09181 0,927429 1,17724 0,2398

X5 0,677283 1,57859 0,429043 0,6681

R-squared = 32,684 percent

R-squared (adjusted for d.f.) = 31,796 percent

Standard Error of Est. = 15,3827

Mean absolute error = 11,8448

Durbin-Watson statistic = 1,11166 (P=0,0000)

Lag 1 residual autocorrelation = 0,431718

Multiple Regression – Y

Multiple Regression Analysis

Dependent variable: Y

Standard T

Parameter Estimate Error Statistic P-Value

CONSTANT 16,6472 3,26579 5,09746 0,0000

X2 1,57776 0,328357 4,80501 0,0000

X3 0,432221 0,0340983 12,6757 0,0000

R-squared = 32,4151 percent

R-squared (adjusted for d.f.) = 32,0613 percent

Standard Error of Est. = 15,3527

Mean absolute error = 11,928

Durbin-Watson statistic = 1,10934 (P=0,0000)

Lag 1 residual autocorrelation = 0,433462

Para realizar la regresión múltiple, primero se deben buscar las variable que mejor predican las ventas mensuales; este proceso se realiza sacando las variables con un p value mayor que el alfa del 4%. Al final quedaron que las mejores variables que predican a Y son el ingreso personal promedio y el numero de automóviles registrados, aunque cabe anotar que nos es un muy buen modelo puesto que no corrige un buen porcentaje errores. Se deberían buscar otras variables que mejoren el modelo.

Prueba de hipótesis para el intercepto

Como el valor de esta prueba (0) es menor que el con el que venimos trabajando, la hipótesis nula es rechazada. Lo que nos dice que el intercepto es diferente a cero (0) y que probablemente hay otras

variables que deban estar incluidas dentro del modelo.

Prueba de hipótesis para la pendiente

Como el valor de esta prueba es igual a cero (0) y menor que el , esta hipótesis debe ser rechazada, por lo tanto la pendiente es diferente de 0. Esto quiere decir que la variable ingreso personal promedio y número de automóviles registrados si deben estar incluida en este modelo