UNIVERSIDADE FEDERAL DO PARÁ

MÉTODOS QUANTITATIVOS APLICADOS À CONTABILIDADE
Prof. Héber Lavor Moreira
Plano de Aula

TEMA: REGRESSÃO MÚLTIPLA - Caso Multiplan S/A

INTRODUÇÃO

Em muitos casos uma variável pode estar relacionada com outras duas ou mais variáveis. Assim sendo, podemos ajustar um modelo, que permita descrever uma relação linear existente entre uma variável dependente Y e duas ou mais variáveis independentes, ou seja, diversas variáveis explicativas, podem ser utilizadas para prever o valor de uma variável dependente. Nosso interesse agora é o estudo do modelo de regressão com mais variáveis independentes, visando a melhor compreensão do comportamento da variável dependente.

Quando transformamos um modelo econômico com mais de uma variável explicativa em seu modelo estatístico correspondente, passamos a designá-lo como modelo de regressão múltipla.

Modelo Geral:

Para mais de duas variáveis independentes, em função de uma variável dependente.

$\hat{Y} = \beta_O + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + \ell$

Onde k

número de variáveis "independentes". (As variáveis independentes são chamadas também variáveis preditoras ou variáveis X).

$\hat{Y}$ = valor preditor da variável dependente Y (calculado a partir da equação de regressão múltipla).
X_1, X_2, ..., X_k, são as variáveis independentes.
β_O = intercepto Y, ou valor de Y quando todas as variáveis preditoras são zero.
b_1 = estimativa de β_O, baseada nos dados amostrais.
β_1 = inclinação de Y em relação à variável X_1, mantendo constantes as variáveis X_2, X_3, ..., X_k.
b_1 = estimativa de β_O, baseada nos dados amostrais.
β_2 = inclinação de Y em relação à variável X_2, mantendo constantes as variáveis X_1, X_3, ..., X_k.
.
.
.

β_k = inclinação de Y em relação à variável X_k, mantendo constantes as variáveis X_1, X_2, X_3, ..., X_k.
ε_i = erro aleatório em y, para a observação *i*.

Modelo de três Variáveis: Tal modelo envolve apenas três variáveis: uma dependente (Y) e duas independentes (X_1 e X_2), isto é: Y= $f(X_1, X_2)$.

$$Y = b_o + b_1x_{1i} + b_2x_{2i} + \varepsilon_i$$

Aplicação do Modelo

A aplicação será realizada para a análise do comportamento dos Custos Indiretos de Fabricação (CIF) em função de cada uma das Variáveis (HMOD e HM) isoladamente e em função destas duas variáveis simultaneamente.

Tab.1- Companhia Multiplast S/A- série de dados relativos aos 15 últimos meses

Meses	CIF	HM	LP
1	76.667	1.772	31
2	73.678	1.820	22
3	80.141	1.634	39
4	61.985	1.006	40
5	72.685	1.383	39
6	87.675	1.957	41
7	78.450	1.561	48
8	70.634	1.464	34
9	63.417	1.545	25
10	56.057	1.119	29
11	67.446	1.382	35
12	72.102	1.320	47
13	68.533	1.264	49
14	69.079	1.344	29
15	85.550	1.803	48
16	58.197	1.022	38
17	61.626	1.510	21
18	80.689	1.793	29
19	58.256	1.149	27
20	55.337	1.155	22
21	85.108	1.847	38
22	76.485	1.832	23
23	67.783	1.136	42
24	56.398	1.136	22
25	66.622	1.330	33
26	63.494	1.358	26
27	89.416	1.882	45
28	67.518	1.174	45
29	73.680	1.643	31
30	66.132	1.381	26

Com base nos cálculos apresentados no Excel, faremos a interpretação das estimativas, obedecendo o seguinte comando:

VARIÁVEIS: (HM)

Custos Indiretos de Fabricação (CIF) em função das Horas Máquina (HM)

ANÁLISE DOS RESULTADOS

Função: CIF = 27.901,23 + 29,13HM

$y = b_o + b_1 x_{1i}$

Podemos afirmar que:

O coeficiente angular (é a inclinação da reta quando Y é influenciado pelo valor de X) da variável independente HM, atende aos requisitos específicos no teste de hipótese da regressão. É diferente de zero e assim caracteriza a existência de regressão, mostrando-se a um nível de *significância* de 0,05.

O ***Coeficiente de determinação ou explicação (R^2)***, é de **74,12%**. Quer dizer que a Variável Independente – HM, é responsável por 74,12 % das variações ocorridas com o Custo Indireto de Fabricação (CIF), ou seja 74,12 % das variações ocorridas com o CIF, são explicadas pela variação ocorrida com a HM, enquanto que 25,88 % das variações ocorridas com o CIF, não são explicadas por HM. Isto é, são decorrentes de *causas aleatórias* não especificadas no modelo.

R – quadrado ajustado: é calculado para refletir tanto o número de variáveis explicativas no modelo, quanto o tamanho da amostra. No presente caso, ele demonstra que HM é responsável por 73,20% das variações ocorridas no CIF.

RESUMO DOS RESULTADOS

CIF = f (HM)

Estatística de regressão	
R múltiplo	0,860935717
R-Quadrado	0,74121031
R-quadrado ajustado	**0,731967821**
Erro padrão	5003,125148
Observações	30

ANALISANDO A SIGNIFICÂNCIA

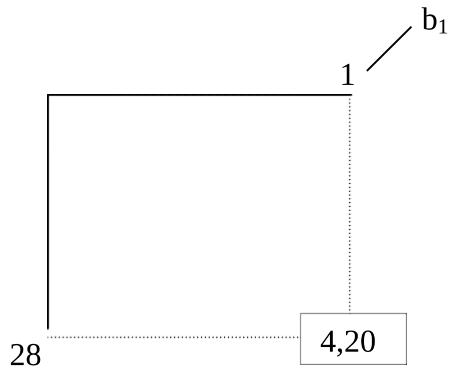
ANOVA

	gl	*SQ*	*MQ*	*F*	*F de significação*
Regressão	1	2007406122	2007406122	80,19596386	1,03716E-09
Resíduo	28	700875314,8	25031261,24		
Total	29	2708281437			

ANÁLISE DO F CRÍTICO

gl = n - P - 1
gl = 30 - 1 - 1
gl = 28

α 0,05
$H_o : F > F_{(1,28)}$ – rejeita
80,19596386 > 4,20



Vide Levine, pág. 741

Teste F

Testa a equação de regressão como um todo.

Usada em regressão múltipla

F de significação = 0,00000000103716
Nível de confiança = 0,95
Nível de significância = 0,05

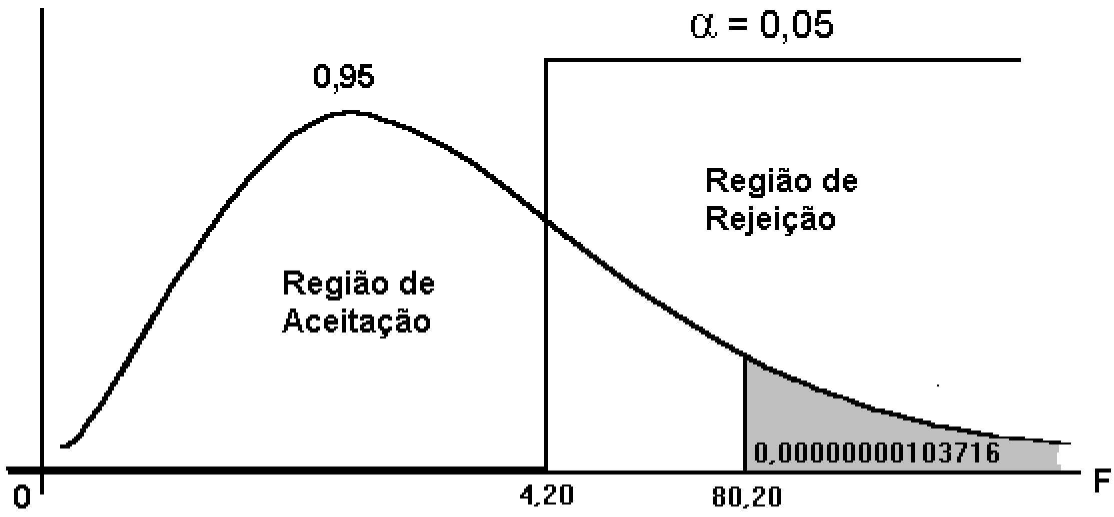
Regra Geral

Se F < 0,05 **α** (Nível de significância) : Regressão Significativa

Se 0,00000000103716 < 0,05: há Regressão Significativa
Se F ≥ 0,05 **α** (Nível de significância): Regressão Não Significativa

No exemplo, o valor 4,20 é o "F" crítico (F de Fischer)

Área vai usar 80,20 até final é 5%.



a) temos que pelo método do teste p que:
- Se o valor **p** for maior ou igual a **α**, a hipótese nula é aceita (H_O: **p ≥ α**). Não há regressão
- Se o valor p for menor do que a, a hipótese nula é rejeitada
- (H_1: **p < α**). **Há** regressão;

b) temos pela estatística F que:
- $(F>F_{s(p,n-p-1)})$ rejeita H_o
- $(F<F_{s(p,n-p-1)})$ aceita H_1

Nesta regressão, temos que o valor p é de 0,00000000103716 portanto < a, rejeita $H_{O,}$ e aceita $H_{1,}$ e configurando a existência de regressão.

Pela estatística **F: F = 80,19596386 > $F_{s\ (1,28)}$ = 4,20**, por conseguinte rejeita $H_{O,}$ e aceita H_1. Então inferimos com um nível de significância de 0,05 que, é provável que haja regressão linear, com uma forte correlação entre as variáveis, onde a variável explicativa relaciona-se aos gastos de CIF.

ANALISANDO A INCLINAÇÃO

	Coeficientes	*Erro padrão*	*Stat t*	*valor-P*	*95% inferiores*	*95% superiores*	*Inferior 95,O%*	*Superior 95,O%*
Interseção	27901,22754	4828,567607	5,778365307	3,32801E-06	18010,34407	37792,11102	18010,34407	37792,11102
HM	29,13414697	3,253314516	8,955219922	1,03716E-09	22,47002679	35,79826714	22,47002679	35,79826714

Com base nos dados acima, **rejeitou-se a hipótese da inclinação ser igual a zero**, o que significa admitirmos a existência de custo fixo.

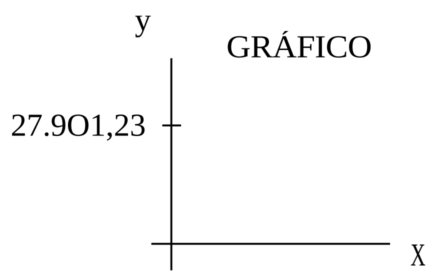
O valor de **b_1 = 29,13** (HM - Variável X_1), indica que, para cada crescimento correspondente a uma unidade em X, estima-se que o valor Y cresça a uma média de 29,13 unidades.

Isto é, para cada crescimento correspondente ao número de Horas Máquina, o modelo ajustado prevê a estimativa de que os custos indiretos de fabricação (CIF) cresçam em 29,13 unidades monetárias.

Depreende-se assim que, se a Hora Máquina aumenta em 10 unidades (de tempo - horas), espera-se que os custos indiretos de fabricação cresçam 291,34 unidades.

O **intercepto**, valor de **b_o = 27.901,23** representa o valor médio de **Y**, quando **X** é igual a (0) zero. Ou seja, indica que o custo indireto de fabricação é de 27.901,23 para uma referência (0) zero do número de Horas Máquina (X).

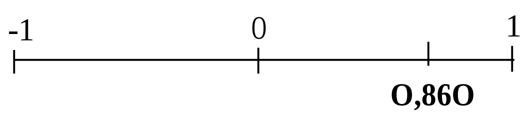
$y = b_0 + b_1 x_{1i}$
$y = 27.901,23 + 29,13 * 0$
$y = 27.901,23$



RESUMO DOS RESULTADOS
CIF = f (HM)

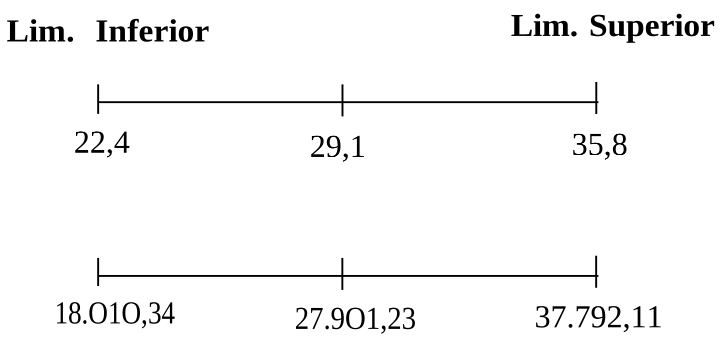
Estatística de regressão	
R múltiplo	**0,860935717**
R-Quadrado	0,74121031
R-quadrado ajustado	0,731967821
Erro padrão	5003,125148
Observações	30

O ***Coeficiente de Correlação (R),*** é de **86,09 %**, o que indica que há uma forte correlação (associação) linear, direta, entre o Custo Indireto de Fabricação e as Horas Máquina.



Como **CIF é *função* de HM**, com um valor de mão de obra direta elevado, teremos um elevado custo indireto de fabricação.

Há 95 % de probabilidade de que b_1 (HM) esteja no intervalo de (22,47; 35,80) e de que b_o esteja no intervalo de (18.010,34; 37.792,11).



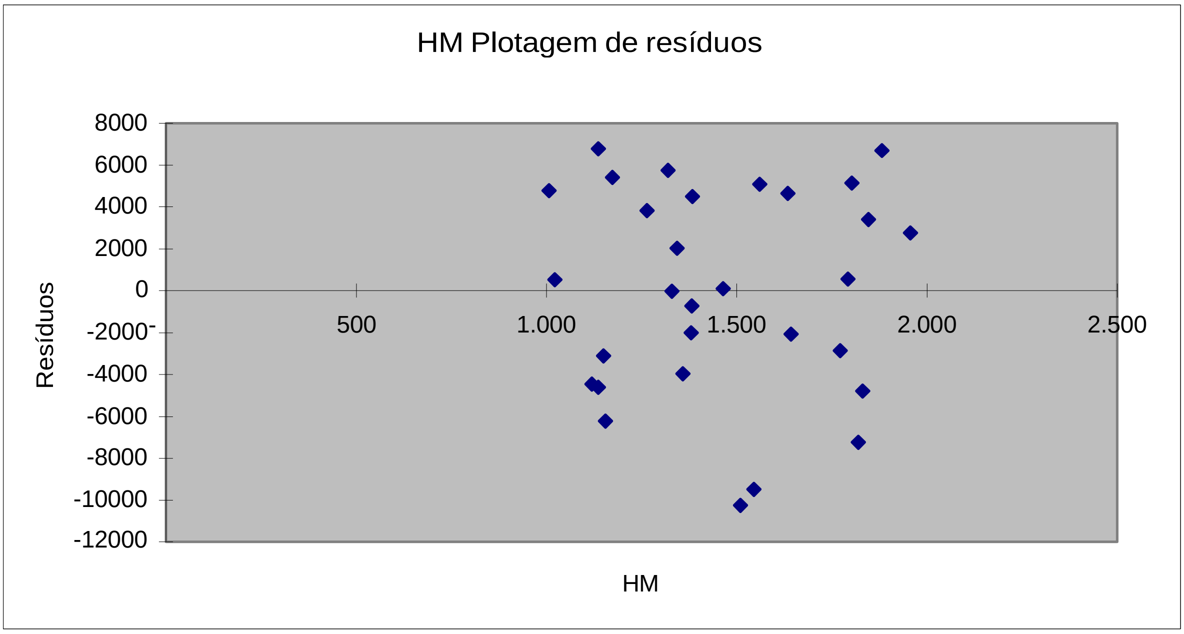
ANÁLISE DOS RESÍDUOS

Embora exista uma ampla dispersão no gráfico de resíduos, não há um padrão aparente. Os resíduos parecem estar uniformemente espalhados acima e abaixo de **0**, para os diferentes valores de **X** (HM). Portanto podemos concluir que em relação aos dados sobre o ClF e HM, que o modelo ajustado, parece apropriado.

RESULTADOS DE RESÍDUOS

Observação	*Previsto(a) CIF*	*Resíduos*	*Resíduos padrão*
1	79526,93597	-2859,935969	-0,581748037
2	80925,37502	-7247,375023	-1,474209997
3	75506,42369	4634,576312	0,94273288
4	57210,17939	4774,820607	0,971260387
5	68193,7528	4491,247201	0,913577882
6	84916,75316	2758,246842	0,561063151
7	73379,63096	5070,369041	1,031378769
8	70553,6187	80,38129669	0,016350597
9	72913,48461	-9496,484608	-1,931708032
10	60502,338	-4445,338	-0,904239355
11	68164,61865	-718,6186521	-0,146176346
12	66358,30154	5743,69846	1,168342698
13	64726,78931	3806,21069	0,774232578
14	67057,52107	2021,478933	0,411195011
15	80430,09452	5119,905475	1,041455121
16	57676,32574	520,6742559	0,105911891
17	71893,78946	-10267,78946	-2,088601435
18	80138,75306	550,2469447	0,111927359
19	61376,36241	-3120,362409	-0,63472215
20	61551,16729	-6214,167291	-1,264042155
21	81711,99699	3396,003009	0,69079102
22	81274,98479	-4789,984787	-0,974344978
23	60997,6185	6785,381502	1,380234527
24	60997,6185	-4599,618498	-0,935622008
25	66649,64301	-27,64300981	-0,005622946

26	67465,39912	-3971,399125	-0,807834046
27	82731,69214	6684,307865	1,35967484
28	62104,71608	5413,283917	1,101132098
29	75768,63101	-2088,63101	-0,424854614
30	68135,48451	-2003,484505	-0,407534711



"Após a análise, o controlador não ficou satisfeito com o poder Explicativo da Regressão, e procurou analisar melhor o processo de produção, concluindo que o número de Lotes de Produção (LP) afetam os custos indiretos de fabricação durante o mês. Assim resolveu efetuar uma nova análise tomando por base as duas variáveis independentes (HM e LP)".

Com base nos cálculos apresentados no Excel, e para efeito exploratório, faremos a interpretação das estimativas, obedecendo o seguinte comando:

VARIÁVEIS: (LP)

Custos Indiretos de Fabricação (CIF) em função dos Lotes de Produção (LP)

ANÁLISE DOS RESULTADOS

Função: $CIF = 52.675,31 + 518,15LP$
$y = b_o + b_1 x_{1i}$

Podemos afirmar que:

O coeficiente angular (é a inclinação da reta quando Y é influenciado pelo valor de X) da variável independente LP, atende aos requisitos específicos no teste de hipótese da regressão. É diferente de zero e assim caracteriza a existência de regressão, mostrando-se a um nível de *significância* de 0,05.

O ***Coeficiente de determinação ou explicação (R^2)***, é de **23,13%**. Quer dizer que a Variável Independente - LP, é responsável por apenas 23,13 % das variações ocorridas com o Custo Indireto de Fabricação (CIF), ou seja 23,13 % das variações ocorridas com o CIF, são explicadas pela variação ocorrida com a LP.

R - quadrado ajustado: é calculado para refletir tanto o número de variáveis explicativas no modelo, quanto o tamanho da amostra. No presente caso, ele demonstra que LP é responsável por 20,38 % das variações ocorridas no CIF.

RESUMO DOS RESULTADOS
CIF = f (LP)

Estatística de regressão	
R múltiplo	0,480956086
R-Quadrado	0,231318756
R-quadrado ajustado	**0,203865855**
Erro padrão	8622,655256
Observações	30

ANALISANDO A SIGNIFICÂNCIA

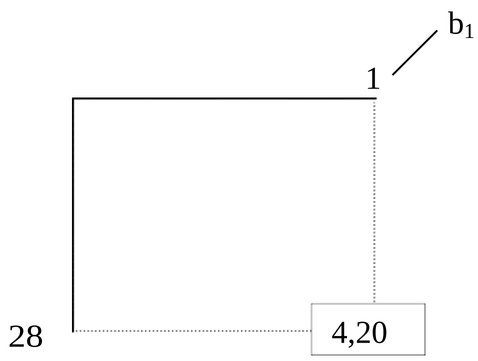
ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	1	626476294	626476294	8,426022145	0,007134689
Resíduo	28	2081805143	74350183,67		
Total	29	2708281437			

ANÁLISE DO F CRÍTICO

gl = n - P - 1
gl = 30 - 1 - 1
gl = 28

α 0,05
H_o : F > $F_{(1,28)}$ – rejeita
8,426022145 > 4,20



Vide Levine, pág. 741

Teste F

Testa a equação de regressão como um todo.

Usada em regressão múltipla

F de significação = 0,007134689
Nível de confiança = 0,95
Nível de significância = 0,05

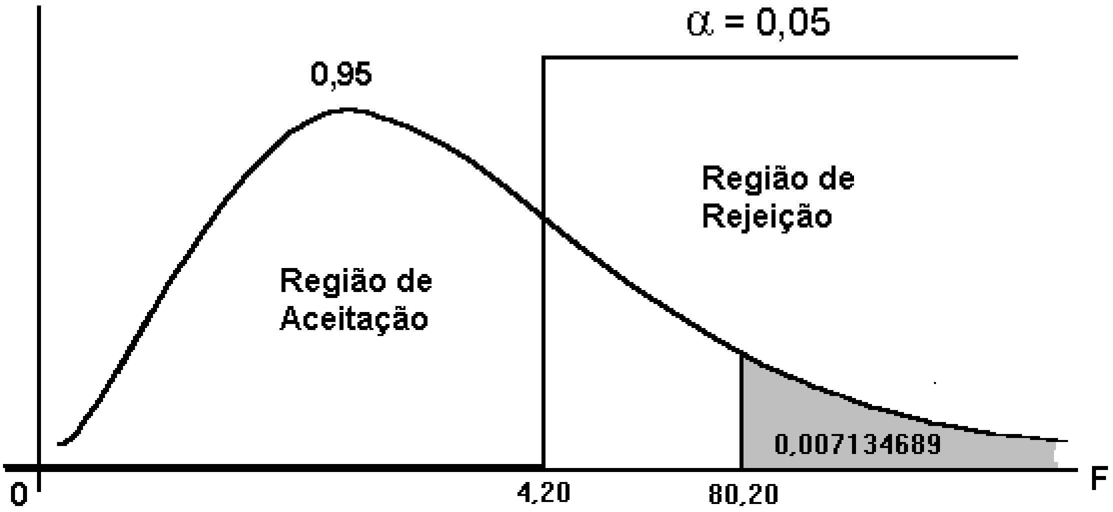
Regra Geral

Se F < 0,05 α (Nível de significância): Regressão Significativa

Se 0,007134689 < 0,05: há Regressão Significativa
Se F ≥ 0,05 α (Nível de significância): Regressão Não Significativa

No exemplo, o valor 4,20 é o "F" crítico (F de Fischer)

Área vai usar 80,20 até final é 5%.



c) temos que pelo método do teste p que:
- Se o valor **p** for maior ou igual a **α**, a hipótese nula é aceita (H_O: **p ≥ α**). Não há regressão
- Se o valor p for menor do que a, a hipótese nula é rejeitada
- (H_1: **p < α**). **Há** regressão;

d) temos pela estatística F que:
- ($F>F_{s(p,n-p-1)}$) rejeita H_o
- ($F<F_{s(p,n-p-1)}$) aceita H_1

Nesta regressão, temos que o **valor p** é de **0,007134689** portanto **< α, rejeita H_O**, e **aceita H_1**, e configurando a existência de regressão.

Pela estatística **F: F = 8,426022145 > $F_{s\ (1,28)}$ = 4,20**, por conseguinte rejeita H_O, e aceita H_1. Então inferimos com um nível de significância de 0,05 que, a regressão linear existe, com **baixa correlação**, embora exista, onde a variável explicativa relaciona-se aos gastos de ClF.

ANALISANDO A INCLINAÇÃO

	Coeficientes	*Erro padrão*	*Stat t*	*valor-P*	*95% inferiores*	*95% superiores*	*Inferior 95,O%*	*Superior 95,O%*
Interseção	52675,30884	6292,924081	8,370561627	4,17646E-09	39784,82373	65565,79395	39784,82373	65565,79395
LP	518,1452488	178,5008227	2,902761124	0,007134689	152,5024782	883,7880195	152,5024782	883,7880195

Com base nos dados acima, rejeitou-se a hipótese da inclinação ser igual a zero, o que significa admitirmos a existência de custo fixo.

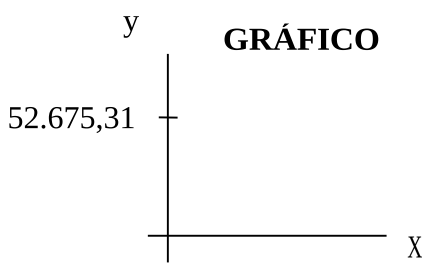
O valor de **b_1 = 518,15** (LP - Variável X_1), indica que, para cada crescimento correspondente a uma unidade em X, estima-se que o valor Y cresça a uma média de 518,15 unidades.

Isto é, para cada crescimento correspondente ao número de Lotes de Produção, o modelo ajustado prevê a estimativa de que os custos indiretos de fabricação (ClF) cresçam em 518,15 unidades monetárias.

Depreende-se assim que, se os Lotes de Produção aumentam em 10 unidades (de tempo - horas), espera-se que os custos indiretos de fabricação cresçam 5.181,45 unidades monetárias.

O **intercepto**, valor de **b_o = 52.675,30** representa o valor médio de **Y**, quando **X** é igual a (0) zero. Ou seja, indica que o custo indireto de fabricação é de 52.675,30 para uma referência (0) zero do número de Lotes de Produção (X).

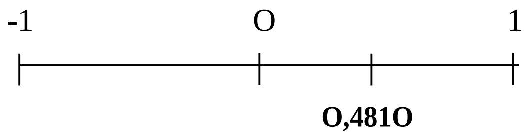
$y = b_0 + b_1 x_{1i}$
$y = 52.675,31 + 518,14 * 0$
$y = 52.675,31$



RESUMO DOS RESULTADOS
ClF = f (LP)

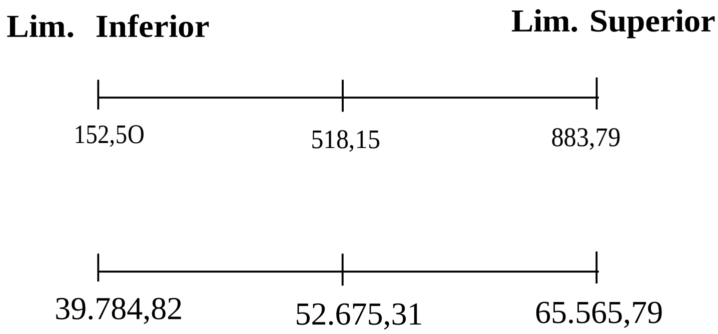
Estatística de regressão	
R múltiplo	0,480956086
R-Quadrado	0,231318756
R-quadrado ajustado	0,203865855
Erro padrão	8622,655256
Observações	30

O ***Coeficiente de Correlação (R)***, é de **48,10 %**, o que indica que há baixa correlação (associação) linear, direta, entre o Custo Indireto de Fabricação e os Lotes de Produção (LP).



Como **ClF é *função* de LP**, o custo indireto de fabricação será influenciado pelo aumento no volume de Lotes de Produção (LP).

Há 95 % de probabilidade de que b_1 (LP) esteja no intervalo de (152,50; 883,79) e de que b_o esteja no intervalo de (39.784,82; 65.565,79).



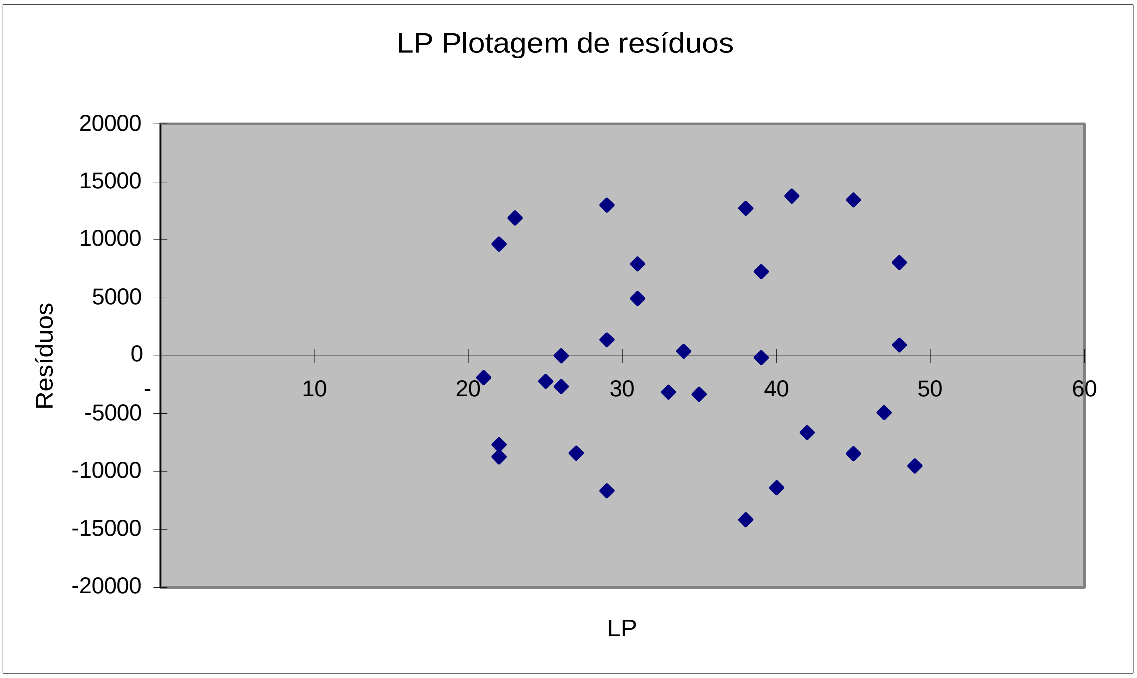
ANÁLISE DOS RESÍDUOS

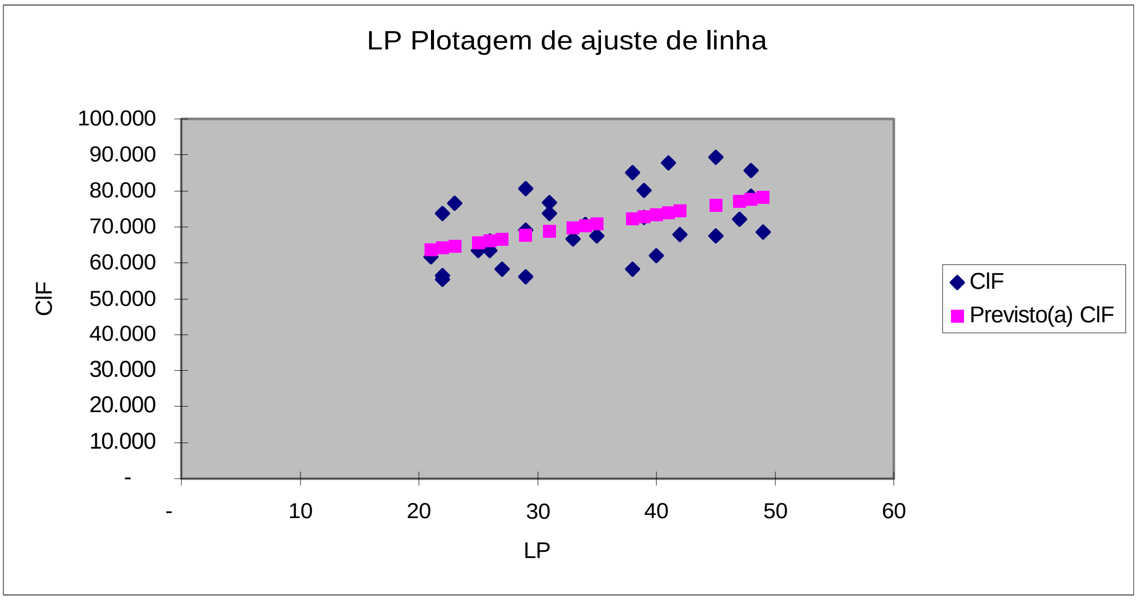
Embora exista uma ampla dispersão no gráfico de resíduos, não há um padrão aparente. Os resíduos parecem estar uniformemente espalhados acima e abaixo de **0**, para os diferentes valores de **X** (HM). Portanto podemos concluir que em relação aos dados sobre o CIF e HM, que o modelo ajustado, parece apropriado.

RESULTADOS DE RESÍDUOS

Observação	*Previsto(a) CIF*	*Resíduos*	*Resíduos padrão*
1	68737,81155	7929,188446	0,935853136
2	64074,50431	9603,495686	1,133465501
3	72882,97354	7258,026456	0,856638339
4	73401,11879	-11416,11879	-1,347402782
5	72882,97354	-197,9735444	-0,023366094
6	73919,26404	13755,73596	1,623539246
7	77546,28078	903,719216	0,106662677
8	70292,2473	341,7526998	0,040335822
9	65628,94006	-2211,940061	-0,261067202
10	67701,52106	-11644,52106	-1,374360266
11	70810,39255	-3364,392549	-0,397086958
12	77028,13554	-4926,135535	-0,581413775
13	78064,42603	-9531,426033	-1,124959383
14	67701,52106	1377,478944	0,162578806
15	77546,28078	8003,719216	0,94464973
16	72364,8283	-14167,8283	-1,672176999
17	63556,35907	-1930,359065	-0,227833226
18	67701,52106	12987,47894	1,532864678
19	66665,23056	-8409,230558	-0,992510752
20	64074,50431	-8737,504314	-1,031255704
21	72364,8283	12743,1717	1,504029988
22	64592,64956	11892,35044	1,403610663
23	74437,40929	-6654,409291	-0,785395611
24	64074,50431	-7676,504314	-0,906029751

25	69774,10205	-3152,102051	-0,372031086
26	66147,08531	-2653,085309	-0,313133964
27	75991,84504	13424,15496	1,584403953
28	75991,84504	-8473,845037	-1,000136963
29	68737,81155	4942,188446	0,583308442
30	66147,08531	-15,08530941	-0,001780464





Com base nos cálculos apresentados no Excel, faremos a interpretação das estimativas, obedecendo o seguinte comando:

VARIÁVEIS: (HM, LP)

⇨ **Custos Indiretos de Fabricação (CIF)** em função das **Horas Máquina (HM)** e **Lotes de Produção (LP)**

ANÁLISE DOS RESULTADOS

⇨ Função: CIF = **f** (HM, LP)
⇨ Função: CIF = 11.620,93 + 28,73HM + 494,27LP

$$y = b_o + b_1x_{1i} + b_2X_2$$

Podemos afirmar que:

O coeficiente angular (é a inclinação da reta quando **Y** é influenciado pelo valor de **X**) das variáveis independentes HM e LP, atende aos requisitos específicos no teste de hipótese da regressão. É diferente de zero e assim caracteriza a existência de regressão, mostrando-se a um nível de *significância* de 0,05.

O ***Coeficiente de determinação ou explicação (R^2)***, é de **95,16 %.** Quer dizer que as Variáveis Independentes - HM e LP, são responsáveis por 95,16 % das variações ocorridas com o Custo Indireto de Fabricação (CIF), ou seja 95,16% das variações ocorridas com o CIF, são explicadas pelas variações ocorridas com a HM e LP, se tratadas conjuntamente, enquanto que 4,84 % das variações ocorridas com o CIF, não são explicadas por HM e LP. Isto é, são decorrentes de *causas aleatórias* não especificadas no modelo.

R - quadrado ajustado: é calculado para refletir tanto o número de variáveis explicativas no modelo, quanto o tamanho da amostra. No presente caso, ele demonstra que HM e LP, conjuntamente, são responsáveis por 94,80,16% das variações ocorridas no CIF.

RESUMO DOS RESULTADOS

CIF = f (HM,LP)

Estatística de regressão	
R múltiplo	0,975478099
R-Quadrado	0,951557521
R-quadrado ajustado	0,947969189
Erro padrão	2204,338035
Observações	30

ANALISANDO A SIGNIFICÂNCIA

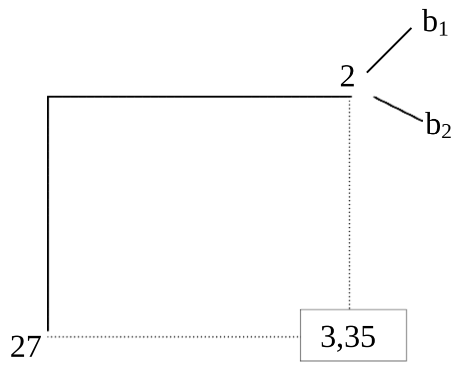
ANOVA

	gl	SQ	MQ	F	F de significação
Regressão	2	2577085570	1288542785	265,1810312	1,78056E-18
Resíduo	27	131195866,6	4859106,171		
Total	29	2708281437			

ANÁLISE DO F CRÍTICO

gl = n - P - 1
gl = 30 - 2 - 1
gl = 27

α 0,05
$H_o : F > F_{(2,27)}$ – rejeita
265,1810312 > 3,35



Vide Levine, pág. 741

Teste F

Testa a equação de regressão como um todo.

Usada em regressão múltipla

F de significação = 0,00000000000000000178056
Nível de confiança = 0,95
Nível de significância = 0,05

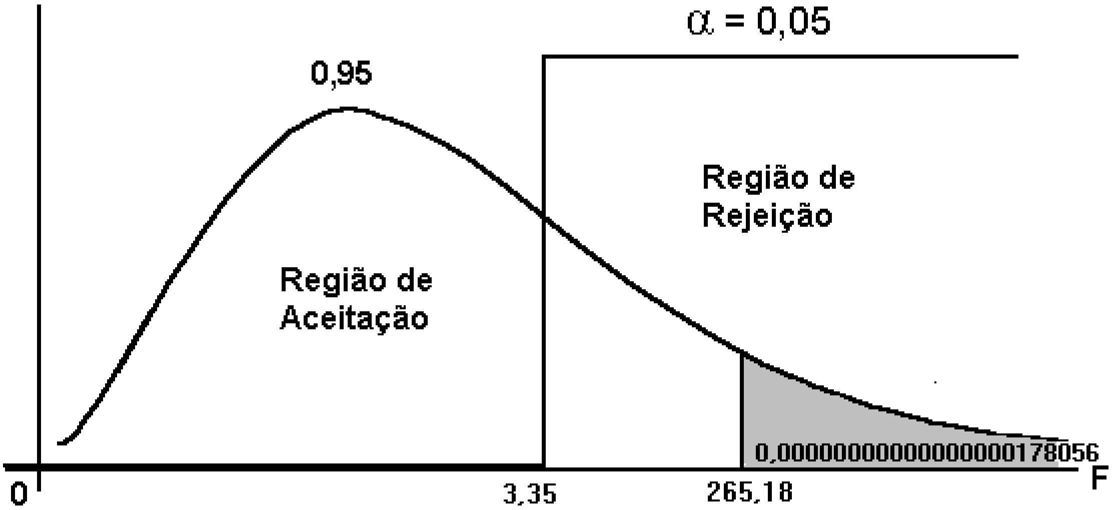
Regra Geral

Se F < 0,05 α (Nível de significância): Regressão Significativa

Se 0,00000000000000000178056 < 0,05: há Regressão Significativa
Se F ≥ 0,05 α (Nível de significância): Regressão Não Significativa

No exemplo, o valor 3,35 é o "F" crítico (F de Fischer)

Área vai usar 265,18 até final é 5%.



e) temos que pelo método do teste p que:

- Se o valor **p** for maior ou igual a **α**, a hipótese nula é aceita (H_O: **p ≥ α**). Não há regressão
- Se o valor p for menor do que a, a hipótese nula é rejeitada
- (H_1: **p < α**). Há regressão;

f) temos pela **estatística F** que:

- $(F > F_{s(p,n-p-1)})$ rejeita H_o
- $(F < F_{s(p,n-p-1)})$ aceita H_1

Nesta regressão, temos que o valor p é < a, portanto rejeita $H_{O,}$ e aceita $H_{1,}$ e configurando a existência de regressão.

Pela estatística **F: F = 265,1810312 > $F_{s\ (2,27)}$ = 3,35**, por conseguinte rejeita $H_{O,}$ e aceita H_1. Então inferimos com um nível de significância de 0,05 que, é provável que haja regressão linear, com uma forte correlação entre as variáveis, onde as variáveis explicativas relacionam-se aos gastos de CIF.

ANALISANDO A INCLINAÇÃO

	Coeficientes	*Erro padrão*	*Stat t*	*valor-P*	*95% inferiores*	*95% superiores*	*Inferior 95,O%*	*Superior 95,O%*
Interseção	11620,92705	2605,128761	4,460787975	0,000129622	6275,647939	16966,20616	6275,647939	16966,20616
HM	28,7288258	1,433873794	20,03581202	9,73536E-18	25,78676176	31,67088984	25,78676176	31,67088984
LP	494,2680341	45,64838764	10,82772163	2,50973E-11	400,605342	587,9307261	400,605342	587,9307261

Com base nos dados acima, **rejeitou-se a hipótese da inclinação ser igual a zero**, o que significa admitirmos a existência de custo fixo.

O valor de $\mathbf{b_2 = 494,27}$ (LP - Variável X_2), indica que, para cada crescimento correspondente a uma unidade de X_2, mantendo-se inalterada HM, estima-se que o valor do CIF cresça a uma média de 494,27 unidade monetárias.

O valor de $\mathbf{b_1 = 28,73}$ (HM - Variável X_1), indica que, para cada crescimento correspondente a uma unidade em X_1, mantendo-se inalterada LP, estima-se que o valor de **Y** cresça a uma média de 28,73 unidades monetárias.

Isto é, para cada crescimento correspondente ao número de Horas Máquina, o modelo ajustado prevê a estimativa de que os custos indiretos de fabricação (CIF) cresçam em 28,73 unidades monetárias.

Depreende-se assim que, se a Hora Máquina aumenta em 10 unidades (de tempo - horas), espera-se que os custos indiretos de fabricação cresçam 287,29 unidades monetárias.

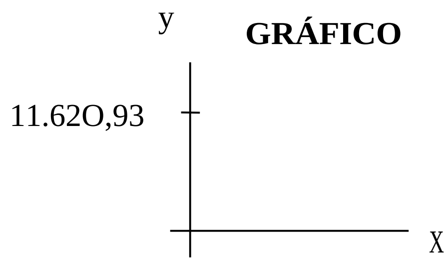
Igual raciocínio se aplica a LP. Desde que mantido inalterado HM.

O **intercepto**, valor de $\mathbf{b_o = 11.620,93}$ representa o valor médio de **Y**, quando $\mathbf{X_1}$ **e** $\mathbf{X_2}$ são iguais a (0) zero. Ou seja, indica que o custo indireto de fabricação é de 11.620,93 para uma referência (0) zero do número de Horas Máquina (X_1) e Lotes de Produção (X_2).

$y = b_0 + b_1x_{1i} + b_2X_{2i}$

$y = 11.620,93 + 28,72HM + 494,27LP$

$\mathbf{y = 11.620,93}$

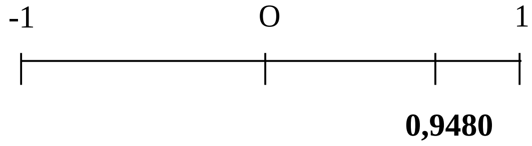


RESUMO DOS RESULTADOS

CIF = f (HM,LP)

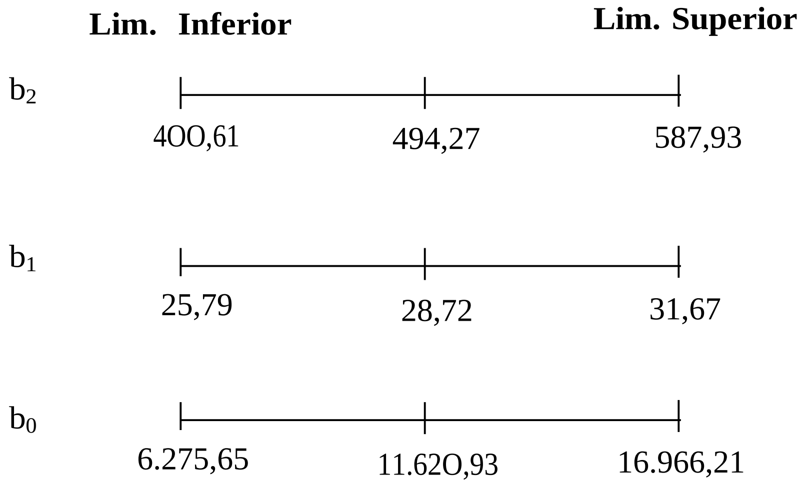
Estatística de regressão	
R múltiplo	0,975478099
R-Quadrado	0,951557521
R-quadrado ajustado	**0,947969189**
Erro padrão	2204,338035
Observações	30

O ***R - Quadrado Ajustado****,* é de **94,80 %**, o que indica que há uma forte correlação (associação) linear, direta, entre o Custo Indireto de Fabricação e as Variáveis Horas Máquina (HM) e Lotes de Produção (LP).



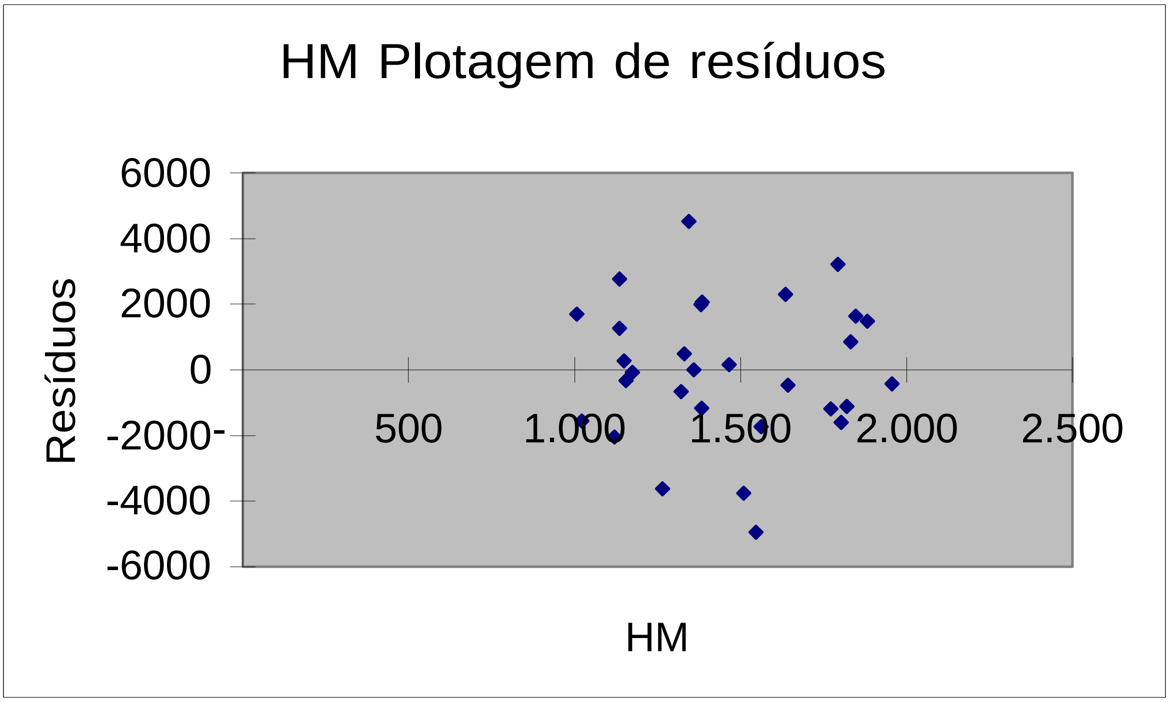
Como **ClF é *função* de HM e LP**, com um valor de Horas Máquina e/ou Lotes de Produção, elevados, teremos um efeito similar, repercutindo em custo indireto de fabricação.

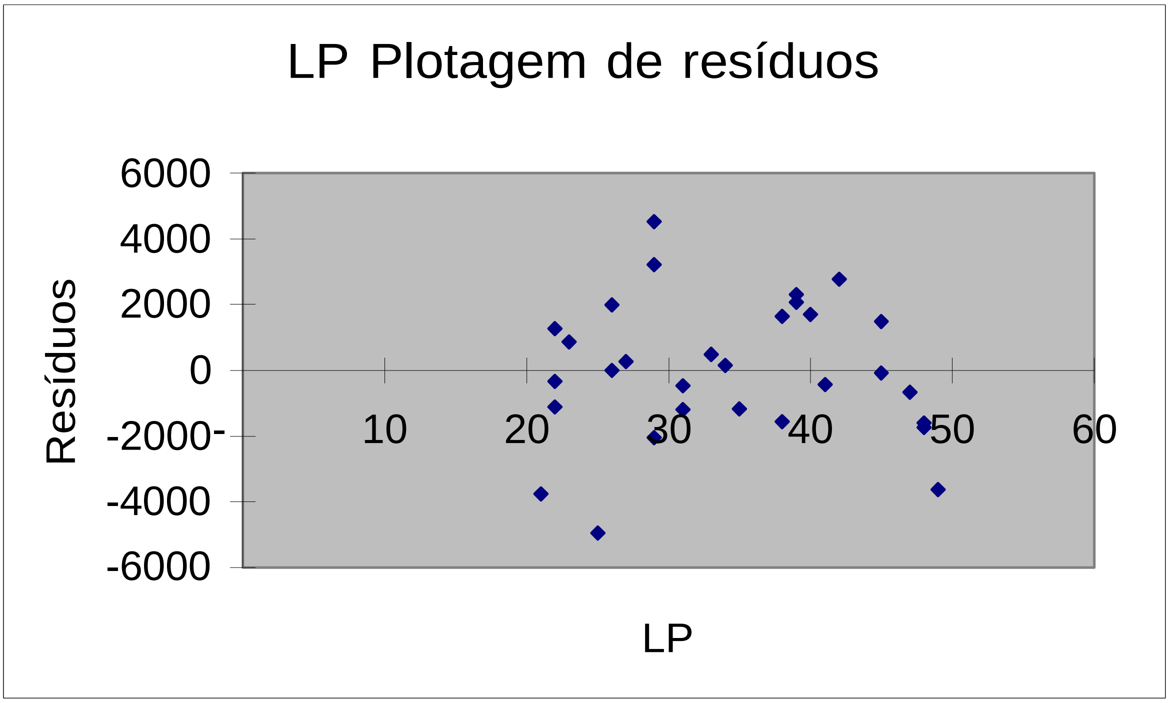
Há 95 % de probabilidade de que b_2 (LP) esteja no intervalo de (400,61; 587,93) de que b_1 (HM) esteja no intervalo de (25,79; 31,67) e de que b_o esteja no intervalo de (6.275,65; 16.966,21).



ANÁLISE DOS RESÍDUOS

Embora exista uma ampla dispersão no gráfico de resíduos de HM e LP, não há um padrão aparente. Os resíduos parecem estar uniformemente espalhados acima e abaixo de 0, para os diferentes valores de **X_1 (HM) e X_2 (LP)**. Portanto podemos concluir que em relação aos dados sobre o ClF, HM e LP, que o modelo ajustado, parece apropriado.


HM Plotagem de resíduos
6000
4000
2000
0
-2000
-4000
-6000
Resíduos
500
1.000
1.500
2.000
2.500
HM


LP Plotagem de resíduos
6000
4000
2000
0
-2000
-4000
-6000
Resíduos
10
20
30
40
50
60
LP

VERIFICANDO A MULTICOLINEARIDADE

Como já vimos, a Multicolinearidade se dá quando as variáveis explicativas são fortemente relacionadas entre si.

"Nessas situações, variáveis colineares não fornecem novas informações, e torna-se difícil separar o efeito dessas variáveis na variável dependente ou variável resposta".
Levine - "Estatística: Teoria e Aplicações" - LTC Editora.

"Um método de mensuração da colinearidade utiliza o **fator inflacionário da variância (FIV)** para cada variável explicativa".

"Assim sendo, o FIV seria calculado de uma fórmula no formato de:
= 1/(1-valor de r^2)".

"Se um conjunto de variáveis explicativas não for correlacionado FIV_j será igual a 1" (o grifo é nosso).

Em nosso caso os dados revelam o seguinte:

CORRELAÇÃO

	CIF	*HM*	*LP*
CIF	1		
HM	0,860935717	1	
LP	0,480956086	0,026106658	1

FIV = 1

Não há Multicolinearidade entre as variáveis explicativas. Concluímos portanto que os coeficientes são confiáveis.