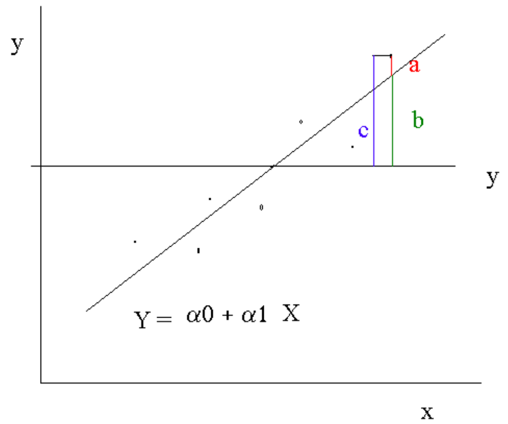
ANÁLISE DE VARIÂNCIA



SE NÃO HOUVER QUALQUER RELAÇÃO ENTRE AS VARIÁVEIS **X** E **Y**, AS MEDIÇÕES DE **Y** CONDUZIRIAM A UM VALOR MAIS PROVÁVEL QUE, À PARTE O ERRO ABSOLUTO, SERIA REPRESENTADO PELO **VALOR MÉDIO** $\overline{Y}$**.**

PODERÍAMOS, PORTANTO, DIZER QUE A VARIÂNCIA ASSOCIADA SERIA :

$$s_y^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(y_i - \overline{y}\right)^2 \qquad \Rightarrow \qquad c$$

c - DESVIOS DOS VALORES EXPERIMENTAIS EM RELAÇÃO AO VALOR MÉDIO

VERIFICA-SE QUE A SOMA $\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2$ É DECOMPONÍVEL EM DUAS PARCELAS:

$$c \quad \Leftrightarrow \quad \sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2 \left\langle \begin{array}{lll} \sum_{i=1}^{n}\left(y_i - Y_i\right)^2 & \Leftrightarrow & a \\ \sum_{i=1}^{n}\left(Y_i - \bar{y}\right)^2 & \Leftrightarrow & b \end{array} \right.$$

a - DESVIOS DOS VALORES EXPERIMENTAIS EM RELAÇÃO AOS VALORES DA RECTA.

b - ASSOCIADO À REGRESSÃO - DESVIOS DA RECTA EM RELAÇÃO AO VALOR MÉDIO

COM BASE NESTES DESVIOS É POSSÍVEL EFECTUAR A **ANÁLISE DE VARIÂNCIA** QUE NOS PERMITE DETERMINAR O **GRAU DE AJUSTE** DA RECTA AOS PONTOS EXPERIMENTAIS.

ANÁLISE DE VARIÂNCIA DA REGRESSÃO LINEAR SIMPLES

FONTE DE VARIAÇÃO	G. L.	QUADRADOS DOS DESVIOS (Q. D.)	VARIÂNCIAS OU MÉDIAS QUADRÁTICAS (M. Q.)	F
DESVIOS DA REGRESSÃO EM RELAÇÃO AO VALOR MÉDIO DE y ($\overline{y}$)	1*	$\sum\left(Y_i - \overline{y}\right)^2 = \dfrac{\sum\left[\left(x_i - \overline{x}\right)\left(y - \overline{y}\right)\right]^2}{\sum\left(x_i - \overline{x}\right)^2}$	MQR $= \dfrac{\sum\left(Y_i - \overline{y}\right)^2}{1}$	$F = \dfrac{MQR}{s_{y.x}^2}$
DESVIOS DOS VALORES EXPERIMENTAIS EM RELAÇÃO AOS RESPECTIVOS VALORES DA CORRELAÇÃO	n-2	$\sum\left(y_i - Y_i\right)^2 = \sum y_i^2 - \alpha_0 \sum y_i - \alpha_1 \sum x_i\, y_i$	$s_{y.x}^2 = \dfrac{\sum\left(y_i - Y_i\right)^2}{n - 2}$	
TOTAL - DESVIOS DOS VALORES EXPERIMENTAIS EM RELAÇÃO AO VALOR MÉDIO DE y ($\overline{y}$)	n-1	$\sum\left(y_i - \overline{y}\right)^2$	$s_y^2 = \dfrac{\sum\left(y_i - \overline{y}\right)^2}{n - 1}$	

* - DEPOIS DE DETERMINADA A EQUAÇÃO DA RECTA, OS VALORES DA RECTA APENAS TÊM **1 GL**, POIS PARA CADA x_i, HÁ APENAS UM VALOR DE Y_i E UM VALOR DE $\overline{y}$.

O QUOCIENTE $F = \dfrac{MQR}{s_{y.x}^2}$ PERMITE-NOS DETERMINAR COM QUE PROBABILIDADE A CORRELAÇÃO **SE AJUSTA** AOS PONTOS EXPERIMENTAIS, COMPARANDO O SEU VALOR COM VALORES TABELADOS QUE DEPENDEM **DA PROBABILIDADE DE AJUSTE E DO NÚMERO DE GRAUS DE LIBERDADE DO NUMERADOR E DO DENOMINADOR**.

MEDIÇÕES DE Y REPETIDAS PARA UM DADO VALOR DE X

QUANDO PARA CADA VALOR DE X, SE EFECTUAM VÁRIAS MEDIÇÕES DE Y, A ANÁLISE DE VARIÂNCIA É DIFERENTE:

- DIVIDEM-SE OS DESVIOS DOS VALORES EXPERIMENTAIS EM RELAÇÃO AOS VALORES DA RECTA, EM DUAS PARTES:

$$a \Leftrightarrow \sum_{i=1}^{n} \left(y_i - Y_i\right)^2 \left\langle \begin{array}{ll} \displaystyle\sum_{i=1}^{M}\sum_{j=1}^{p} \left(y_{ij} - \bar{y}_i\right)^2 & \Leftrightarrow a_1 \\ \displaystyle\sum_{i=1}^{n} \left(\bar{y}_i - Y_i\right)^2 & \Leftrightarrow a_2 \end{array} \right.$$

a_1 - DESVIOS DE CADA VALOR REPETIDO DE y, EM RELAÇÃO À MÉDIA DE y, PARA CADA VALOR DE x.

a_2 - DESVIOS DAS MÉDIAS ($\bar{y}_i$) PARA CADA VALOR DE x, EM RELAÇÃO À CORRELAÇÃO.

ANÁLISE DE VARIÂNCIA DA REGRESSÃO LINEAR SIMPLES PARA MEDIÇÕES REPETIDAS

FONTE DE VARIAÇÃO	G. L.	QUADRADOS DOS DESVIOS (Q. D.)	VARIÂNCIAS OU MÉDIAS QUADRÁTICAS (M. Q.)	F
DESVIOS DA REGRESSÃO EM RELAÇÃO AO VALOR MÉDIO DE y ($\bar{y}$)	1	$\sum\left(Y_i - \bar{y}\right)^2$	$MQR = \dfrac{\sum\left(Y_i - \bar{y}\right)^2}{1}$	$F = \dfrac{MQM}{MQD}$
DESVIOS DE CADA VALOR REPETIDO DE y, EM RELAÇÃO AO VALOR MÉDIO DE y, PARA CADA VALOR DE x.	SOMA DOS G.L. DE CADA CONJUNTO $\sum_i\left(\sum_j n_{ij} - 1\right)$	$\sum_i \sum_j\left(y_{ij} - \bar{y}_i\right)^2$	$MQD = \dfrac{\sum_i \sum_j\left(y_{ij} - \bar{y}_i\right)^2}{\sum_i\left(\sum_j n_{ij} - 1\right)}$	
DESVIOS DAS MÉDIAS DE y, PARA CADA x, EM RELAÇÃO AOS RESPECTIVOS VALORES DA CORRELAÇÃO	Nº DE MÉDIAS - 2 M-2	$\sum\left(\bar{y}_i - Y_i\right)^2$	$MQM = \dfrac{\sum\left(\bar{y}_i - Y_i\right)^2}{M - 2}$	
TOTAL	n-1	$\sum \sum\left(y_{ij} - \bar{y}\right)^2$	$s_y^2 = \dfrac{\sum\left(y_i - \bar{y}\right)^2}{n - 1}$	

NESTE CASO É O VALOR DO QUOCIENTE $\left(\dfrac{MQM}{MQD}\right)$ QUE VAMOS COMPARAR COM OS VALORES DAS TABELAS DE **F** QUE NOS PERMITE DETERMINAR A PROBABILIDADE DO **AJUSTE** OU **FALTA DE AJUSTE** DA CORRELAÇÃO AOS PONTOS EXPERIMENTAIS.

Exemplo:

Na determinação da massa específica de um óleo a diferentes temperaturas, obtiveram-se os seguintes valores:

ρ / g cm^{-3}	ϕ / ºC
0,906 0,905 0,907	14,0
0,894 0,892 0,893	23,0
0,884 0,886	32,9
0,876 0,877 0,875	41,9
0,866 0,868	52,4

a) Determinar a melhor aproximação, sob o ponto de vista estatístico, da correlação linear $\rho = \rho(\phi)$.

b) Determinar o valor da massa específica a 25 ºC, e o respectivo erro.

c) Efectuar a análise de variância.

R:

a) $\rho = (0{,}918 \pm 0{,}003) + (-1{,}00 \pm 0{,}09) \times 10^3\ \phi$

b) $Y_o = \hat{Y}_o \pm 4{,}184\mathrm{x}10^{-3}\sqrt{1 + \frac{1}{13} + \frac{(x_o - 31{,}28)^2}{2324{,}54}}$; $\rho_{25ºC} = (0{,}893 \pm 0{,}004)\ \mathrm{gcm}^{-3}$

c) Ajuste superior a 99 %.

ANÁLISE DE VARIÂNCIA NA REGRESSÃO LINEAR MÚLTIPLA

FONTE DE VARIAÇÃO	G. L.	QUADRADOS DOS DESVIOS (Q. D.)	VARIÂNCIAS OU MÉDIAS QUADRÁTICAS (M. Q.)	F
DESVIOS DA REGRESSÃO EM RELAÇÃO AO VALOR MÉDIO DE y ($\bar{y}$)	k	$SSR = \sum\left(Y_i - \bar{y}\right)^2$	$MQR = \dfrac{\sum\left(Y_i - \bar{y}\right)^2}{k}$	
DESVIOS DOS VALORES EXPERIMENTAIS EM RELAÇÃO AOS RESPECTIVOS VALORES DA CORRELAÇÃO	n-2	$SRM = \sum_{i=1}^{n}\left[y_i - \sum_{j=1}^{k}\beta_j x_{ij}\right]^2$	$s^2_{(y.x)m} = \dfrac{\sum\left(y_i - Y_i\right)^2}{n - (k+1)}$	$F = \dfrac{MQR}{s^2_{(y.x)m}}$
TOTAL	n-1	$\sum\left(y_i - \bar{y}\right)^2$	$s^2_y = \dfrac{\sum\left(y_i - \bar{y}\right)^2}{n - 1}$	

ESCOLHA DE ESCALAS

A ESCALA DAS COORDENADAS DE UM GRÁFICO DEVE SER TAL QUE OS ERROS DAS VARIÁVEIS REPRESENTADAS TENHAM SIGNIFICADO QUANDO COMPARADOS COM A MENOR DIVISÃO DA ESCALA.

NO PAPEL MILIMÉTRICO, A IMPRECISÃO DE UM NÚMERO DEVE ESTAR CONTIDA NO MILÍMETRO DO PAPEL.

QUANDO, PARA SATISFAZER ESTE REQUISITO, SÃO NECESSÁRIAS ESCALAS MUITO GRANDES, DEVEM ESCOLHER-SE ESCALAS TAIS QUE, REPRESENTANDO O MESMO FENÓMENO COM **ERRO RELATIVO SEMELHANTE**, REDUZAM O TAMANHO DO GRÁFICO.

PODEM USAR-SE:
- ESCALAS LOGARÍTMICAS
- ESCALAS SEMI-LOGARÍTMICAS
- ESCALAS DE INVERSOS
- ...

MUDANÇA DE VARIÁVEIS

NORMALMENTE PROCURA-SE QUE A REPRESENTAÇÃO DE UM FENÓMENO SEJA EFECTUADO ATRAVÉS DE UM CORRELAÇÃO LINEAR SIMPLES, PORQUE É MAIS FÁCIL EFECTUAR INTERPOLAÇÕES E EXTRAPOLAÇÕES.

MUITAS VEZES ISSO É POSSÍVEL EFECTUANDO UMA MUDANÇA DE VARIÁVEL DE TAL MODO QUE A RELAÇÃO PASSE A SER UMA RECTA.