

**A VARIÂNCIA DOS PONTOS DE MÁXIMO OU DE MÍNIMO DE
EQUAÇÕES DE REGRESSÃO DE SEGUNDO GRAU**

MARLI DE BEM GOMES D'AULÍSIO

Orientador: DR. IZAIAS RANGEL NOGUEIRA

Dissertação apresentada à Escola Superior de Agricultura "Luiz de Queiroz", da Universidade de São Paulo, para obtenção do título de Mestre em Experimentação e Estatística.

P I R A C I C A B A
Estado de São Paulo - Brasil
Março - 1976

*A meus pais,
a minhas irmãs
e a minhas sobrinhas
Cibele, Cristiane e Isabella,*

D E D I C O .

A G R A D E C I M E N T O S

Ao Dr. Izaias Rangel Nogueira, pela orientação segu
ra e pelo incentivo a nós dispensado.

Ao Dr. F. Pimentel Gomes, pelas valiosas sugestões
apresentadas e pela revisão deste trabalho.

Aos Drs. Cássio Roberto de Melo Godoi e Vivaldo Franco
cisco da Cruz, pela valiosa ajuda prestada na programação para
o computador.

Aos bolsistas Gilberto Bertolino Rodrigues e Rosana
Clara Victória, pelo auxílio prestado.

Ao CNPq e à FAPESP, pelo auxílio financeiro prestado.

Aos colegas do Departamento de Matemática e Estatísti
ca da Escola Superior de Agricultura "Luiz de Queiroz", e
a todos aqueles que, de um modo ou de outro, concorreram para
o bom andamento deste trabalho.

Í N D I C E

	Página
1. INTRODUÇÃO	1
2. DESENVOLVIMENTO TEÓRICO E REVISÃO DA LITERATURA	4
2.1 - Generalidades	4
2.2 - Estimacão da Variância	7
2.3 - As Estimativas γ_1 e γ_2	7
2.4 - Intervalos de Confiança	8
3. MATERIAL E MÉTODOS	14
3.1 - Material	14
3.2 - Métodos	14
4. RESULTADOS E DISCUSSÃO	23
4.1 - Distribuicão de \underline{x}	23
4.2 - Histogramas de Frequências	27
4.3 - Intervalos de Confiança	28
5. CONCLUSÕES	53
6. RESUMO	55
7. ABSTRACT	58
8. BIBLIOGRAFIA CITADA	60

1. INTRODUÇÃO

Há atualmente no Brasil uma grande expansão no uso de adubos e no estudo de doses econômicas de adubação.

É de interesse do povo brasileiro e da humanidade em geral determinar as doses mais econômicas de fertilizantes a serem usadas. Durante muitos anos vem se utilizando a regressão quadrática para determinação dessa dose mais econômica, que seria o ponto de máximo do trinômio ajustado para o caso. Não se conhece, porém, a variância do ponto de máximo ou de mínimo da equação de regressão.

Admitindo-se para um dado fenômeno a equação de regressão quadrática

$$Y_i = a + bX_i + cX_i^2$$

o modelo matemático correspondente será:

$$Y_i = a + bX_i + cX_i^2 + e_i ,$$

onde se supõe $e_i \sim N(0, \sigma^2)$, isto é, com distribuição normal de média 0 e variância σ^2 .

Sabemos, porém, que se podem obter apenas as estimativas dos parâmetros \underline{a} , \underline{b} e \underline{c} , e a equação estimada

$$\hat{Y} = \hat{a} + \hat{b}X + \hat{c}X^2 .$$

Convém, pois, determinar o ponto de máximo ou de mínimo desta função.

A derivada primeira

$$\hat{Y}' = \hat{b} + 2\hat{c}X ,$$

deve anular-se, logo:

$$\hat{b} + 2\hat{c}X = 0 ,$$

donde,

$$(1) \quad X = - \frac{\hat{b}}{2 \hat{c}} ,$$

valor que é o ponto de máximo ou de mínimo da função.

Como vemos em (1), X é função das estimativas de \hat{b} e \hat{c} , sujeitas a erros experimentais e possui, portanto, uma distribuição probabilística.

O presente trabalho tem em mira a avaliação dos erros experimentais na determinação desses pontos de máximo

ou de mínimo, e estudar a distribuição probabilística a que eles pertencem.

2. DESENVOLVIMENTO TEÓRICO E REVISÃO DA LITERATURA

2.1 - Generalidades

O problema básico encarado neste trabalho, que é a determinação da variância dos pontos de máximo ou de mínimo de equações de regressão do segundo grau, tem estreita relação com o de saber qual a distribuição de um quociente de duas variáveis aleatórias de distribuição normal. Com efeito, de acordo com o que vimos no 1º capítulo, o ponto de máximo ou de mínimo relativo a uma equação de regressão

$$\hat{Y} = \hat{a} + \hat{b}X + \hat{c}X^2$$

é dada pela expressão

$$(1) \quad X = - \frac{\hat{b}}{2 \hat{c}} = \frac{1}{2} \cdot \frac{\hat{b}}{- \hat{c}} .$$

Se os valores observados de Y tiverem distribuição normal, sabe-se que as estimativas \hat{b} e \hat{c} , funções lineares dessas observações, serão também normalmente distribuídas. No entanto, \hat{b} e \hat{c} geralmente não serão independentes, se adotada essa forma para a equação de regressão, é isso que complica desnecessariamente o problema.

Podemos, porém, tomar a equação de regressão sob a forma

$$(2) \quad Y = a + b P_1(x) + c P_2(x) \quad ,$$

onde $P_1(x)$ e $P_2(x)$ são os polinômios ortogonais convenientes, de primeiro e de segundo grau, respectivamente. Neste caso, \hat{b} e \hat{c} têm ainda distribuição normal, mas são independentes, de sorte que seu coeficiente de correlação é nulo.

Sabe-se que podemos tomar

$$P_1(x) = x - \bar{X} \quad ,$$

$$P_2(x) = x^2 - k \quad ,$$

onde \bar{X} é a média dos valores de X , e k é um número real. Teremos, pois, neste caso, o ponto de máximo ou de mínimo

$$(3) \quad x = \frac{1}{2} \cdot \frac{\hat{b}}{-\hat{c}} \quad ,$$

onde, porém, \hat{b} e \hat{c} são variáveis normais independentes.

O problema da distribuição do quociente de duas variáveis aleatórias é discutido por KENDALL (1947) e por KENDALL e STUART (1963).

Consideremos um quociente

$$z = \frac{x_1}{x_2}$$

de variáveis normais, com $x_1 \sim N(m_1, \sigma_1^2)$, $x_2 \sim N(m_2, \sigma_2^2)$. Quando $m_1 = m_2 = 0$, e as variáveis aleatórias são independentes, demonstra-se que z tem uma distribuição de Cauchy (KENDALL e STUART, 1963). Mas este caso foge ao nosso problema objetivo, pois na prática da adubação temos $m_1 > 0$, $m_2 > 0$.

Por outro lado, admitida ainda a independência das duas variáveis, no caso de $m_2 > 0$ ser muito maior do que σ_2^2 (por exemplo $m_2 > 4\sigma_2$), de tal sorte que valores negativos de x_2 praticamente não ocorram, sabe-se que a variável

$$u = \frac{m_1 - m_2 z}{(\sigma_1^2 + \sigma_2^2 z^2)^{1/2}}$$

tem distribuição normal (aproximada) de média zero e variância um (KENDALL, 1947; KENDALL e STUART, 1963).

No caso geral, porém, que é o que realmente nos interessa, não se conhece a distribuição de z , nem se sa

be exatamente qual a sua variância.

2.2 - Estimação da Variância

Dada a equação (1)

$$X = - \frac{\hat{b}}{2 \hat{c}} \quad ,$$

obtemos por diferenciação

$$dX = - \frac{1}{2} \left[\frac{d\hat{b}}{\hat{c}} - \frac{\hat{b} d\hat{c}}{\hat{c}^2} \right]$$

de onde se deduz a fórmula aproximada

$$4V(X) = \frac{1}{\hat{c}^2} V(\hat{b}) + \frac{\hat{b}^2}{\hat{c}^4} V(\hat{c}) - 2 \frac{\hat{b}}{\hat{c}^3} \text{Cov}(\hat{b}, \hat{c}) \quad .$$

Se, porém, adotarmos a equação de regressão sob a forma (2), com polinômios ortogonais, \hat{b} e \hat{c} serão independentes. Neste caso obtemos, pelo mesmo método,

$$V(X) = \frac{1}{4} \left[\frac{1}{\hat{c}^2} V(\hat{b}) + \frac{\hat{b}^2}{\hat{c}^4} V(\hat{c}) \right] \quad .$$

2.3 - As Estatísticas γ_1 e γ_2

FISHER (1930) estudou, entre outras coisas, a distribuição de estimativas de

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \quad \text{e de} \quad \gamma_2 = \frac{\mu_4}{\sigma^4} - 3 \quad ,$$

obtendo seus momentos. Não conseguiu, porém, obter a função de densidade de nenhuma das distribuições, a não ser no caso particular de três amostras, em que deduziu a função de densidade para $\hat{\gamma}_1$. Sabe-se que γ_1 mede a assimetria, e γ_2 a curtose de uma distribuição. Demonstra-se que γ_1 e γ_2 valem zero quando estamos diante de uma distribuição normal. No caso de amostras, porém, podemos ter valores próximos de zero, daí a necessidade de sabermos se o valor obtido difere realmente de zero. Nesse trabalho FISHER deduziu, também, as fórmulas:

$$V(\hat{\gamma}_1) = \frac{6n(n-1)}{(n-2)(n+1)(n+3)}$$

$$V(\hat{\gamma}_2) = \frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}$$

que nos permitiram aplicar o teste t , para verificar se os valores estimados para γ_1 e γ_2 realmente diferem de zero.

2.4 - Intervalos de Confiança

Consideremos ainda o quociente

$$z = \frac{x_1}{x_2},$$

com $x_1 \sim N(m_1, \sigma_1^2)$, $x_2 \sim N(m_2, \sigma_2^2)$ e seja $\mu = \frac{m_1}{m_2}$. Um problema importante e relacionado com o assunto desta dis-

sertação é a determinação de intervalos de confiança de μ . Como a distribuição de \underline{z} não é conhecida, no caso geral, o problema tem sido abordado de maneira indireta, com base em um teorema de E.C. Fieller (KENDALL e STUART, 1963; ANDERSON e NELSON, 1971; BAUM et alii, 1956; FINNEY, 1975). O caminho seguido é o seguinte.

Sendo x_1 e x_2 variáveis normalmente distribuídas, a nova variável $u = x_1 - \mu x_2$ terá também distribuição normal, de média zero e variância $V(u) = \sigma_1^2 + \mu^2 \sigma_2^2$. Sendo s_1^2 e s_2^2 estimativas de σ_1^2 e σ_2^2 , respectivamente, conclui-se, pois, que o quociente

$$\frac{x_1 - \mu x_2}{\sqrt{s_1^2 + \mu^2 s_2^2}}$$

tem distribuição de \underline{t} . Podemos, pois, através das tabelas próprias, obter um intervalo $[-t; +t]$, para o qual tenhamos

$$P \left[-t \leq \frac{x_1 - \mu x_2}{\sqrt{s_1^2 + \mu^2 s_2^2}} \leq t \right] = 1 - \alpha.$$

Da desigualdade

$$(4) \quad \frac{x_1 - \mu x_2}{\sqrt{s_1^2 + \mu^2 s_2^2}} \leq t$$

podemos tentar deduzir intervalos (ou regiões) de confiança para μ . Com efeito, de (4) obtemos:

$$(5) \quad (x_2^2 - t^2 s_2^2)\mu^2 - 2 x_1 x_2 \mu + (x_1^2 - t^2 s_1^2) \leq 0 .$$

Devemos resolver esta inequação de segundo grau em μ .

Tomemos

$$p = x_2^2 - t^2 s_2^2 \quad , \quad q = x_1^2 - t^2 s_1^2$$

e as raízes da equação correspondente serão dadas pela expressão

$$a = \frac{x_1 x_2 \pm \sqrt{x_1^2 x_2^2 - p q}}{p} .$$

Para facilitar a discussão, tomaremos x_1 e x_2 como positivos, o que não traz perda real de generalidade.

Há vários casos a considerar:

$$1) \quad p > 0 \quad , \quad q > 0$$

Temos necessariamente $p < x_2^2$, $q < x_1^2$, logo

$$x_1^2 x_2^2 > x_1^2 x_2^2 - p q > 0 \quad ,$$

de onde se conclui que as duas raízes a_1 e a_2 são reais e positivas. A inequação estará satisfeita para

$$a_1 \leq \mu \leq a_2 .$$

$$2) p > 0 , q = 0$$

Neste caso as raízes são reais. Uma (a_1) é nula, e a outra,

$$a_2 = \frac{2 x_1 x_2}{p} ,$$

é positiva. A inequação estará satisfeita para

$$0 \leq \mu \leq a_2 .$$

$$3) p > 0 , q < 0$$

Neste caso temos

$$0 < x_1^2 x_2^2 < x_1^2 x_2^2 - p q .$$

As raízes são reais e de sinais contrários ($a_1 < 0$, $a_2 > 0$). A inequação estará satisfeita para

$$a_1 \leq \mu \leq a_2 .$$

$$4) p = 0$$

Neste caso obtemos

$$\mu \geq \frac{q}{2 x_1 x_2} .$$

$$5) p < 0, q > 0$$

Neste caso temos

$$0 < x_1^2 x_2^2 < x_1^2 x_2^2 - p q,$$

de sorte que as raízes (a_1 e a_2) serão reais e de sinais contrários. A inequação estará satisfeita para

$$\mu \leq a_1 < 0, \quad \mu \geq a_2 > 0.$$

$$6) p < 0, q = 0$$

Neste caso as raízes serão reais. Uma delas (a_2) é nula, e a outra é

$$a_1 = \frac{2 x_1 x_2}{p} < 0.$$

A inequação estará satisfeita para

$$\mu \leq a_1, \quad \mu \geq 0.$$

$$7) p < 0, q < 0$$

A natureza das raízes dependerá do valor do discriminante $x_1^2 x_2^2 - p q$.

$$7.1) x_1^2 x_2^2 - p q > 0$$

Temos então

$$x_1^2 x_2^2 > x_1^2 x_2^2 - p q > 0 ,$$

logo as raízes a_1 e a_2 serão reais e negativas. A inequação estará satisfeita para

$$\mu \leq a_1 , \quad \mu \geq a_2 .$$

$$7.2) x_1^2 x_2^2 - p q = 0$$

As raízes a_1 e a_2 são iguais e reais. A inequação estará satisfeita para qualquer valor real de μ .

$$7.3) x_1^2 x_2^2 - p q < 0$$

As raízes são complexas. A inequação estará satisfeita para qualquer valor real de μ .

Em resumo, intervalos de confiança propriamente ditos só são obtidos quando temos $p > 0$. Casos bem pouco satisfatórios e até bizarros ocorrem para $p \leq 0$. Ora, sendo $p = x_2^2 - t^2 s_2^2$, poderemos obter sempre valor nulo ou negativo de p , bastando, para isso, tomar t suficientemente grande. O mesmo ocorre, aliás, com $q = x_1^2 - t^2 s_1^2$. Tais fatos, bastante estranhos, justificam a pesquisa que fazemos neste trabalho, que pode também conduzir a intervalos de confiança, talvez mais adequados.

3. MATERIAL E MÉTODOS

3.1 - Material

Usou-se o computador IBM 1130 com 16 K palavras de memória, do Departamento de Matemática e Estatística da ESALQ, para os trabalhos de geração de variáveis aleatórias e para os cálculos mais complexos. Para operações mais simples, lançou-se mão de uma calculadora eletrônica Facit 1125.

3.2 - Métodos

No presente estudo o método fundamentalmente aplicado foi o da simulação de dados. Estimamos valores para $\hat{\beta}$ e para \hat{c} através do computador eletrônico 1130 da ESALQ, usando a sub-rotina CALL RANDU para geração de números alea

tórios. A sub-rotina baseia-se na fórmula:

$$(1) \quad Y = \frac{\sum_{i=1}^k X_i - \frac{k}{2}}{\sqrt{k/12}},$$

onde X_i é um número casual de distribuição retangular,

$$0 < X < 1,$$

k é o número de valores X_i que tomamos dentro da distribuição retangular. No nosso caso considerou-se $k = 60$.

A variância dessa distribuição retangular é:

$$V(X) = \int_0^1 \left(X - \frac{1}{2}\right)^2 dx = \frac{1}{12},$$

logo:

$$V\left(\sum_{i=1}^k X_i\right) = \frac{k}{12},$$

$$\sigma\left(\sum_{i=1}^k X_i\right) = \sqrt{k/12}.$$

Tomemos agora

$$Z = \sum_{i=1}^k X_i.$$

Esta variável terá distribuição aproximadamente normal (KENDALL e STUART, 1963), com média $\frac{k}{2}$ e variância

$\frac{k}{12}$. Segue-se que a variável

$$Y = \frac{Z - \frac{k}{2}}{\sqrt{k/12}} = \frac{\sum_{i=1}^k X_i - \frac{k}{2}}{\sqrt{k/12}} ,$$

terá, aproximadamente, distribuição normal reduzida, isto é, de média zero e variância um.

Assim se obtiveram 160 séries de 100 erros. Através de fatores apropriados foram eles multiplicados de modo a obter 20 séries de 100 erros relativos a cada uma das variâncias seguintes:

$$\begin{array}{ll} \sigma^2 = 0,015625 & \sigma^2 = 2,0000 \\ \sigma^2 = 0,0625 & \sigma^2 = 4,0000 \\ \sigma^2 = 0,2500 & \sigma^2 = 6,2500 \\ \sigma^2 = 1,0000 & \sigma^2 = 9,0000 \end{array}$$

Para $\sigma^2 = 0,015625$ ($\sigma = 0,125$), por exemplo, tivemos duas séries de 1000 erros (e_1 e e_2), com distribuição aproximadamente normal de média zero e com essa variância. Os valores de e_2 foram multiplicados por $\sqrt{3}$, assim obtendo-se, pois, 1000 erros $e_3 = \sqrt{3} e_2$, de variância $3\sigma^2 = 3 \times 0,015625$.

A seguir consideramos os valores exatos $b = 3$, $c = -1$ e tomamos as estimativas

$$\hat{b} = b + e_3 , \quad \hat{c} = c + e_1 .$$

Assim se obtiveram 1000 pares de valores de \hat{b} e de \hat{c} , em 10 grupos de 100. Para cada um desses grupos temos, evidentemente,

$$V(\hat{b}) = 3\sigma^2 \quad , \quad V(\hat{c}) = \sigma^2 \quad .$$

Os coeficientes de variação seriam:

$$C.V.(\hat{b}) = \frac{0,125 \sqrt{3} \times 100}{3} = 7,2\%$$

$$C.V.(\hat{c}) = \frac{0,125 \times 100}{1} = 12,5\% \quad ,$$

onde $0,125 = \sigma = \sqrt{0,015625}$.

Aí tomamos a média verdadeira $b = 3$, e consideramos o valor absoluto $|c| = 1$.

Com coeficientes de variação tão baixos valores negativos de \hat{b} e de $-\hat{c}$ só podem ocorrer com probabilidade extremamente reduzida. Para nos aproximarmos do que ocorre mais comumente na prática, basta considerar os outros valores de σ^2 , que nos levariam a completar o quadro a seguir.

Caso	σ^2	σ	C.V. (\hat{b})	C.V. (\hat{c})
1	0,015625	0,125	7,2%	12,5%
2	0,0625	0,250	14,4%	25,0%
3	0,2500	0,500	28,9%	50,0%
4	1,0000	1,000	57,7%	100,0%
5	2,0000	1,414	81,6%	141,4%
6	4,0000	2,000	115,5%	200,0%
7	6,2500	2,500	144,3%	250,0%
8	9,0000	3,000	173,2%	300,0%

É claro que no caso dos valores mais altos de σ há grande probabilidade de aparecerem valores negativos de \hat{b} , e maior ainda de \hat{c} . Com isto ocorrem frequentemente valores de \hat{c} próximos de zero, o que pode acarretar estimativas elevadas, em valor absoluto para x .

Poder-se-ia indagar o porquê dos valores fixados para \underline{b} e \underline{c} ($b = 3$, $c = -1$).

Em primeiro lugar, é fácil compreender que o importante não é o valor absoluto de cada parâmetro. Assim, tanto faz ter $b = 3$, $V(\hat{b}) = 0,7500$, como ter $b = 30$, $V(\hat{b}) = 75,000$. Raciocínio análogo vale para o parâmetro \underline{c} .

Por outro lado, poderíamos variar \underline{b} , em vez de $V(\hat{b})$, mas o resultado seria equivalente, o mesmo ocorrendo, "mutatis mutandis", para \underline{c} e $V(\hat{c})$.

Resta apenas explicar porque tomamos $b = -3c$, $V(\hat{b}) = 3V(\hat{c})$. Assim fizemos porque, em geral, nos experimentos de adubação, temos c mais próximo de zero e com variância menor. De qualquer forma, aliás, o parâmetro realmente mais importante na regressão quadrática é o coeficiente do termo de segundo grau (c), que nos indica, pelo seu sinal, se há máximo ou mínimo. Razões diferentes entre b e c , ou entre $V(\hat{b})$ e $V(\hat{c})$, poderiam ser tentadas, mas, dentro das limitações naturais de nosso trabalho, essas parecem adequadas.

Transformamos assim uma das séries de 100 erros casuais geradas em 100 valores de \hat{c} , e a outra em 100 valores de \hat{b} , com os quais calculamos

$$x = -\frac{\hat{b}}{2\hat{c}},$$

obtendo assim também 100 valores para x . Desses valores tiramos as médias \bar{c} , \bar{b} e \bar{x} . Com esses valores gerados da maneira exposta, foram feitos os cálculos abaixo:

3.2.1 - Estimamos $V(x)$, pela fórmula usual

$$V_1(x) = \frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N - 1},$$

onde tivemos sempre $N = 100$, de acordo com o que já foi explicado.

3.2.2 - Estimamos $V(x)$ pela fórmula deduzida por D'AULÍSIO (1970)

$$(4) \quad V_2(x) = \frac{1}{4} \left[\frac{1}{\bar{c}^2} V(\hat{b}) + \frac{\bar{b}^2}{\bar{c}^4} V(\hat{c}) \right] ,$$

onde obtivemos a $V(\hat{b})$ e $V(\hat{c})$ pelo método de estimativa da variância.

A fórmula (4) foi obtida a partir da diferenciação de $x = -\frac{\hat{b}}{2\hat{c}}$.

3.2.3 - Estimamos o terceiro (μ_3) e o quarto (μ_4) momentos de x ,

$$\hat{\mu}_3 = \frac{\sum (x - \bar{x})^3}{N - 1} ,$$

$$\hat{\mu}_4 = \frac{\sum (x - \bar{x})^4}{N - 1} .$$

3.2.4 - Obtivemos μ_4 (esperado), que é o quarto momento que teria a distribuição de x se fosse normal.

$$\mu_4(\text{esperado}) = 3 \sigma^4 .$$

3.2.5 - Estimamos γ_1 e γ_2 (FISHER, 1930) para x ,

$$\gamma_1 = \frac{\mu_3}{\sigma^3} , \quad \gamma_2 = \frac{\mu_4}{\sigma^4} - 3 .$$

Como se sabe, na distribuição normal temos

$$\gamma_1 = \gamma_2 = 0 .$$

3.2.6 - Obtivemos $V(\hat{\gamma}_1)$ e $V(\hat{\gamma}_2)$ (FISHER, 1930), para podermos aplicar a prova de t .

Uma vez que

$$V(\hat{\gamma}_1) = \frac{6N(N - 1)}{(N - 2)(N + 1)(N + 3)} ,$$

$$V(\hat{\gamma}_2) = \frac{24N(N - 1)^2}{(N - 3)(N - 2)(N + 3)(N + 5)} ,$$

segue-se que:

$$t = \frac{\hat{\gamma}_1 - 0}{\sqrt{V(\hat{\gamma}_1)}} ,$$

$$t = \frac{\hat{\gamma}_2 - 0}{\sqrt{V(\hat{\gamma}_2)}} .$$

3.2.7 - Foram feitos histogramas das frequências dos valores de x nos 8 casos estudados. Cada histograma representa os resultados de 1.000 dados, considerando nas classes próximas da média o valor de $\sigma/3$ como amplitude de classe, e de σ para as classes mais distantes, onde é pequena a frequência. Depois consideramos que os valores de x representam doses de adubação e tomamos $60 \text{ kg/ha} = 1$, isto é, uma dose. Considerando que se tivermos $x = 10$, a dose re-

comendada será 600 kg/ha, que é muito alta. Vamos eliminar todos os valores de $x \geq 10$ e todos os valores de $x < 0$, e fazemos novos histogramas com estas considerações, obtendo para isso novas médias.

3.2.8 - Foram calculados os intervalos de confiança para a média dos valores de \underline{x} , 1.000 dados para cada valor de σ^2 . Obtivemos os intervalos de confiança pelo método de Fieller, de acordo com o que foi explicado no item 2.4 do capítulo 2. Calculamos também, de uma maneira empírica, para 1.000 dados, o equivalente a um intervalo de confiança ao nível de 5% de probabilidade. Isto foi feito através de determinação do menor e do maior valor observado, após eliminação dos 25 maiores e dos 25 menores. Outros dois tipos de intervalos de confiança foram obtidos considerando

$$\bar{x} \pm t \sqrt{\hat{V}_1(x)} \quad \text{e} \quad \bar{x} \pm t \sqrt{\hat{V}_2(x)},$$

onde consideramos da mesma maneira a média \bar{x} de 1.000 dados, $t = 2$, e onde $\hat{V}_1(x)$ e $\hat{V}_2(x)$ são os valores médios dessas estimativas, obtidas para as 10 repetições feitas para cada valor de σ^2 , sendo $\hat{V}_1(x)$ calculados de acordo com o item 3.2.1 e $\hat{V}_2(x)$, de acordo com o item 3.2.2.

4. RESULTADOS E DISCUSSÃO

4.1 - Distribuição de \underline{x}

Os resultados apresentados, são provenientes de séries geradas de 100 dados. Assim sendo, em cada quadro há 10 linhas, cada uma delas relativa a uma amostra de 100 dados, de sorte que cada quadro resume estatísticas obtidas com 1.000 dados.

Nos Quadros de 1 a 10 temos:

Coluna 1: $\hat{V}(\hat{c})$ calculada pela fórmula do item 3.2.1 de Métodos, com substituição de \underline{x} por \hat{c} .

Coluna 2: Média de $-\hat{c}$.

Coluna 3: $\hat{V}(\hat{b})$ calculada pela fórmula do item 3.2.1 de Métodos, com substituição de \underline{x} por \hat{b} .

Coluna 4: Média de \hat{b} .

Coluna 5: Média de \underline{x} .

- Coluna 6: $\hat{V}_1(x)$, estimada pela fórmula do ítem 3.2.1 de Métodos.
- Coluna 7: $\hat{V}_2(x)$, estimada pela fórmula do ítem 3.2.2 de Métodos.
- Coluna 8: Valores de $\hat{\mu}_3$ pela fórmula citada no ítem 3.2.3 de Métodos.
- Coluna 9: Valores de $\hat{\mu}_4$ pela fórmula citada no ítem 3.2.3 de Métodos.
- Coluna 10: μ_4 (esperado), estimado pela fórmula citada no ítem 3.2.4 de Métodos, substituindo-se σ^2 por $\hat{V}_1(x)$.
- Coluna 11: $\hat{\gamma}_1$, estimado pela fórmula citada no ítem 3.2.5 de Métodos.
- Coluna 12: $\hat{\gamma}_2$, estimado pela fórmula citada no ítem 3.2.5 de Métodos.
- Coluna 13: Teste t aplicado a $\hat{\gamma}_1$, tomado em valor absoluto.
- Coluna 14: Teste t aplicado a $\hat{\gamma}_2$, tomado em valor absoluto.

Podemos ver, observando o Quadro 1, que há semelhança nos resultados obtidos nas colunas 6 e 7. Os dados da coluna 8 ($\hat{\mu}_3$) são próximos de zero. Os valores estimados para $\hat{\mu}_4$ (coluna 9) são bem semelhantes aos obtidos para μ_4 (esperado), que figura na coluna 10. Estudando os valores de $\hat{\gamma}_1$ e $\hat{\gamma}_2$ (colunas 11 e 12), vemos que só o $\hat{\gamma}_1$ da primeira e da sétima linha diferiram de zero ao nível de 5% de proba

bilidade; na quarta, na sexta e na décima linha diferiram ao nível de 1% de probabilidade. O $\hat{\gamma}_2$ diferiu de zero, ao nível de 1% de probabilidade, pelo teste de \underline{t} , apenas na 9.^a linha.

No Quadro 2 verificamos que os valores obtidos para a coluna 7 são sempre inferiores aos da coluna 6. O valor obtido para $\hat{\mu}_3$ só na 3.^a linha se aproxima de zero. O $\hat{\mu}_4$ da terceira linha se aproxima bastante do valor esperado, mas o das demais difere. Quanto a $\hat{\gamma}_1$ e $\hat{\gamma}_2$, só o valor de $\hat{\gamma}_2$ na terceira linha não difere significativamente de zero. O resultado obtido para $\hat{\gamma}_1$ na terceira linha difere de zero ao nível de 5% pela prova de \underline{t} ; os outros valores diferem ao nível de 1% de probabilidade pelo mesmo critério.

No estudo do Quadro 3, vemos que os resultados constantes da coluna 7 são sempre inferiores aos da coluna 6. Os valores calculados para $\hat{\mu}_3$ são bem diferentes de zero. Os resultados das colunas 9 e 10 diferem muito entre si. Os valores de $\hat{\gamma}_1$ e $\hat{\gamma}_2$ diferem todos significativamente de zero pelo teste \underline{t} , ao nível de 1% de probabilidade; conforme podemos verificar pelas colunas 13 e 14. Olhando agora os Quadros 4, 5 e 6, vemos que os resultados obtidos nos levam a observações semelhantes às feitas para o Quadro 3.

Vemos no Quadro 7, que os resultados da coluna 7 em geral subestimam os valores da coluna 6, com exceção do

valor obtido na 9.^a linha, que supera o da coluna 6. Os valores de $\hat{\mu}_3$ são bem diferentes de zero e os de $\hat{\mu}_4$ divergem bastante dos obtidos para $\hat{\mu}_4$ (esperado). Os resultados obtidos para $\hat{\gamma}_1$ e $\hat{\gamma}_2$ diferem bastante de zero. Com exceção dos valores obtidos na 3.^a e 9.^a linhas para $\hat{\gamma}_1$, todos os demais diferem de zero ao nível de 1% de probabilidade pelo teste t.

Examinemos agora o Quadro 8, cujos resultados nas colunas 6 e 7, que estimam o mesmo valor, estão sempre menores na coluna 7. O $\hat{\mu}_3$ apresenta estimativas bem distintas de zero, embora devessem ser nulos, se a distribuição fosse normal. Comparando os resultados constantes das colunas 9 e 10, vemos que diferem grandemente entre si. Sabemos pelo estudo da distribuição normal que o μ_3 é nulo e o $\mu_4 = 3\sigma^4$. Verifica-se, pois, que os pontos de máximo ou de mínimo calculados estão fugindo sensivelmente da distribuição normal. Olhando as colunas 11 e 12, vemos que temos valores diferentes de zero, confirmando a não normalidade dos dados. Indo mais além, vemos as colunas 13 e 14, onde podemos afirmar que os valores obtidos para $\hat{\gamma}_1$ e $\hat{\gamma}_2$ são significativamente diferentes de zero pelo teste t. Rejeita-se a hipótese da nulidade.

Salientemos que nos Quadros 1 e 2 todos os valores de $\hat{\gamma}_1$, são positivos, o que indica ser a distribuição viesada à direita. Nos demais Quadros aparecem sempre valo

res positivos e negativos para $\hat{\gamma}_1$, o que demonstra serem as curvas viezadas às vezes à direita, às vezes à esquerda.

Os valores de $\hat{\gamma}_2$ são negativos em 4 dos 10 casos do Quadro 1, e em apenas 1 caso do Quadro 2; nos demais, são sempre positivos. Isto indica que, à medida que cresce a variação do acaso (medida por σ) tende a distribuição de \underline{x} a se tornar leptocúrtica.

4.2 - Histogramas de Frequências

De acordo com o que foi visto no capítulo de Material e Métodos, Ítem 3.2.7, foram os 1.000 dados obtidos para cada caso levados a histogramas. Os gráficos de 1 a 8, foram feitos com todos os dados e cada um deles representa 1.000 dados, com exceção do Gráfico 3 ($\sigma^2 = 0,2500$) que foi feito com 999 dados, pois um dos resultados obtidos (7681,6601) era demasiadamente absurdo. Em todos os gráficos assinalamos a média. Observando o Gráfico 1, vemos que é ligeiramente viezado à direita, já que os Quadros 2, 3, 4 e 8 estão acentuadamente viezados à direita. Nesses cinco gráficos observamos que as classes de maior frequência estão sempre precedendo a média. No Gráfico 4 a média limita a classe de maior frequência. Nos Gráficos 5 e 7, observamos que a média está à esquerda das classes de maior frequência, portanto, observamos que a curva é viezada à es-

querda (SPIEGEL, 1971; LOBO, 1948), ambas, porém, são leptocúrticas. No Gráfico 6 a média limita a classe de maior frequência, ou classe modal, sendo, porém, uma curva leptocúrtica.

Os Gráficos de 9 a 14 foram feitos de acordo com o que foi visto no ítem 3.2.7 de Material e Métodos, eliminando os valores menores do que 0 e os iguais ou maiores do que 10. Calculou-se nova média para cada caso, de acordo com os valores que ficaram dentro do critério usado. Não houve necessidade de novos gráficos para os dados gerados com $\sigma^2 = 0,015625$ e com $\sigma^2 = 0,0625$, pois todos os valores obtidos já foram maiores que 0 e menores do que 10. Observando os Gráficos de 9 a 13, vemos que todos são viezados à direita e portanto as médias estão situadas depois das classes de maior frequência. Já no caso do Gráfico 14 a média limita a classe de maior frequência.

Observamos em quase todos os casos que a curva não se aproxima da distribuição normal. Com efeito, exceto para $\sigma^2 = 0,015625$, temos sempre assimetria acentuada, a tendência, quando σ^2 cresce, para distribuições leptocúrticas.

4.3 - Intervalos de Confiança

Obtidos pelos métodos indicados no ítem 3.2.8 do Capítulo 3, figuram eles no Quadro 9, em 8 linhas, uma para

cada um dos valores de σ^2 .

Para o nível mais baixo de $\sigma^2(0,015625, C.V.(\hat{c}) = 12,5\%)$ os 4 métodos dão resultados praticamente iguais.

Para $\sigma^2 = 0,0625, C.V.(\hat{c}) = 25,0\%$, o método de Fieller dá resultado bem concordante com o do intervalo de confiança empírico (na nossa denominação). Já os intervalos calculados a partir de $\hat{V}_1(x)$ e $\hat{V}_2(x)$, diferem bastante, mas não exageradamente.

Para $\sigma^2 \geq 0,2500$ o método de Fieller nos leva a resultados absolutamente insatisfatórios, uma vez que surgem extremos infinitos nos intervalos de confiança. Tomado como padrão o intervalo de confiança empírico (como parece recomendável), o intervalo dado por $\hat{V}_1(x)$ é excessivamente amplo, ao passo que o obtido a partir de $\hat{V}_2(x)$ é excessivamente restrito. Assim mesmo, porém, este último é o mais satisfatório em todos ou quase todos os casos.

QUADRO I - Resumo dos Resultados Obtidos para $\sigma^2 = 0.015625$

Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7	Col. 8	Col. 9	Col. 10	Col. 11	Col. 12	Col. 13	Col. 14
\bar{v} (c)	\bar{v} (d)	\bar{b}	\bar{x}	$\bar{v}_1(x)$	$\bar{v}_2(x)$	\bar{u}_3	\bar{u}_4	\bar{u}_4 (esperado)	Y_1	Y_2	Teste t $p/Y_1 \neq 0$	Teste t $p/Y_2 \neq 2$	
0,0213	1,0010	0,0428	2,9855	1,5240	0,0636	0,0581	0,0087	0,0116	0,0120	0,5438	- 0,1000	2,25*	0,209
0,0186	1,0062	0,0435	3,0013	1,5200	0,0571	0,0516	0,0035	0,0078	0,0099	0,2574	- 0,6364	1,07	1,33
0,0141	0,9896	0,0370	3,0161	1,5460	0,0451	0,0429	0,0024	0,0048	0,0060	0,2500	- 0,6000	1,04	1,25
0,0155	0,9993	0,0401	2,9913	1,5220	0,0539	0,0448	0,0083	0,0096	0,0087	0,6640	0,3103	2,75**	0,649
0,0141	0,9995	0,0518	3,0022	1,5222	0,0444	0,0448	0,0039	0,0053	0,0060	0,4149	- 0,3500	1,72	0,732
0,0187	1,0092	0,0465	2,9948	1,5115	0,0565	0,0519	0,0100	0,0111	0,0096	0,7463	0,4688	3,09**	0,980
0,0132	1,0043	0,0417	3,0048	1,5165	0,0446	0,0398	0,0054	0,0067	0,0060	0,5745	0,3500	2,38*	0,732
0,0117	0,9911	0,0523	3,0167	1,5404	0,0433	0,0409	0,0034	0,0066	0,0057	0,3778	0,4787	1,57	0,990
0,0141	0,9952	0,0522	2,9986	1,5268	0,0447	0,0455	0,0067	0,0086	0,0060	0,0094	1,3000	0,039	2,72**
0,0139	0,9937	0,0468	3,0018	1,5333	0,0514	0,0441	0,0078	0,0090	0,0078	0,6784	0,4615	2,81**	0,965

QUADRO 2 - Resumo dos Resultados Obtidos para $\sigma^2 = 0,062500$

Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7	Col. 8	Col. 9	Col. 10	Col. 11	Col. 12	Col. 13	Col. 14
$\varphi (c)$	$-\bar{x}$	$\varphi (b)$	\bar{b}	\bar{x}	$\varphi_1 (x)$	$\varphi_2 (x)$	\bar{u}_3	\bar{u}_4	\bar{u}_4 (esperado)	\bar{y}_1	\bar{y}_2	Teste t $P/Y_1 \neq 0$	Teste t $P/Y_2 \neq 0$
0,8570	0,8833	0,2234	3,0217	1,6434	0,3041	0,1971	0,2378	0,5265	0,2775	1,4180	2,6919	5,87**	5,63**
0,0580	1,0015	0,1696	2,9642	1,5912	0,3369	0,1689	0,4736	1,1960	0,3405	2,4225	7,5374	10,04**	15,76**
0,0492	1,0186	0,1863	2,9978	1,5458	0,1734	0,1477	0,0376	0,0687	0,0903	0,5208	- 0,0538	2,16*	0,111
0,0535	0,9899	0,1616	2,9984	1,6073	0,2500	0,1685	0,2685	0,6786	0,1875	2,1480	7,8576	8,90**	16,43**
0,0690	1,0026	0,2247	3,0319	1,6336	0,3186	0,2129	0,2325	0,4615	0,3045	1,2931	1,5468	5,36**	3,23**
0,0619	0,9940	0,1595	3,0151	1,6589	0,5474	0,1846	1,5541	7,0565	0,8988	3,8373	20,5531	15,90**	42,97**
0,0556	1,0055	0,1405	2,9692	1,5646	0,2046	0,1546	0,1067	0,1843	0,1241	1,1753	1,3162	4,71**	2,75**
0,0697	1,0193	0,1696	2,9814	1,6137	0,6815	0,1846	3,0254	18,8805	1,3932	5,3775	37,6557	22,28**	78,73**
0,0695	0,9984	0,1675	2,9732	1,6196	0,3303	0,1972	0,2469	0,4956	0,3273	1,3008	1,5426	5,39**	3,22**
0,0624	1,0021	0,2202	3,0001	1,6008	0,2647	0,1940	0,1728	0,3214	0,2103	1,2687	1,5849	5,26**	3,31**

QUADRO 3 - Resumo dos Resultados Obtidos para $\sigma^2 = 0,250000$

Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7	Col. 8	Col. 9	Col. 10	Col. 11	Col. 12	Col. 13	Col. 14
$\varphi(\xi)$	$-\bar{c}$	$\varphi(\xi)$	\bar{b}	\bar{x}	$v_1(\xi)$	$v_2(\xi)$	\bar{u}_3	\bar{u}_4	\bar{u}_4 (esperado)	y_1	y_2	Teste c $\rho/\gamma_1 \neq 0$	Teste t $\rho/\gamma_2 \neq 0$
0,2108	0,9787	0,7572	3,0141	1,9657	3,4498	0,7196	15,899	219	36	2,4813	15,4122	10,2888	32,2288
0,2358	1,0219	0,7006	3,0424	0,9572	38,9431	0,6680	- 2,011,190	116,658	4,550	-8,2758	78,9223	34,2888	154,5588
0,2453	1,0137	0,7591	3,0127	-0,8706	882,3087	0,7118	-355,106,688	74,882,256	2,333,406	-9,7340	93,1917	40,3288	194,8488
0,2705	1,0278	0,8422	2,9945	1,3138	34,7699	0,7427	- 1,498,056	82,866	3,627	-7,3067	65,5437	30,2788	137,0388
0,2483	1,0486	0,7410	3,0074	5,1934	738,9841	0,6328	153,781,719	34,603,904	1,638,292	7,6551	60,3657	31,7188	126,2188
0,2980	0,9667	0,6975	0,0396	1,3806	125,1870	0,9052	- 3,212,954	706,394	47,015	-2,2939	42,0743	9,5088	87,9788
0,2954	0,9762	0,6310	2,9982	2,8572	154,7383	0,8964	12,588,904	1,394,758	71,832	6,5402	55,2510	27,0988	115,5288
0,3071	0,9978	0,9007	2,9852	1,3207	162,5910	0,9164	- 5,979,981	747,278	79,707	-2,8844	25,7676	11,9588	52,8388
0,3002	1,0052	0,6336	2,9631	0,2152	219,5067	0,8020	-19,924,758	2,059,893	144,550	-6,1266	39,7513	25,3888	83,1188
0,2524	1,0081	0,7074	3,0379	-11,4817	21,451,8338	0,7378	-30,421,620,047	43,992,318,048	1,380,544,318	-9,6824	92,5000	40,1188	193,3988

QUADRO 4 - Resumo dos Resultados Obtidos para $\sigma^2 = 1,00$

q (G)	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7	Col. 8	Col. 9	Col. 10	Col. 11	Col. 12	Col. 13	Col. 14
	- z	q (S)	b	\bar{x}	v_1 (a)	v_2 (a)	u_3	u_4	\bar{u}_4 (esperado)	q_1	q_2	Teste t $p/N_1 \neq 0$	Teste t $p/N_2 \neq 0$	Teste t
1,0095	1,0029	2,6311	2,9134	4,2262	876,9301	2,7706	223,433,000	61,831,624	2,307,019	8,60	77,40	35,63**	161,82**	
1,0254	0,9641	2,7723	2,8539	0,9234	63,4796	3,1618	-1,124,872	100,314	12,089	-2,22	21,87	9,20**	45,72**	
1,0709	0,9476	3,4637	2,9380	0,9296	13,1138	3,9280	-	5,107	516	-3,20	26,70	13,26**	55,82**	
1,2041	1,0653	2,5724	2,9982	-0,3761	64,7230	2,6672	-2,686,780	1,4,928	12,567	-5,18	33,98	21,46**	71,04**	
0,9643	1,0265	2,6379	3,1478	-8,3784	10,765,2285	2,7818	-10,763,880,023	11,001,215,000	346,379,903	-9,66	92,28	40,02**	192,93**	
0,9237	0,9519	2,6033	3,0341	0,5283	80,4748	3,3074	1,490,841	145,025	19,428	-2,07	19,39	8,38**	40,54**	
0,9236	1,0785	2,6540	3,0056	1,1372	203,0328	2,1117	3,678,860	1,540,420	123,667	-1,27	34,37	5,26**	71,86**	
1,0716	1,0620	2,7747	3,0158	2,6657	357,9530	2,5299	25,672,609	5,564,388	384,391	3,79	40,43	15,70**	84,53**	
0,8041	1,0858	3,4458	2,9872	3,6417	1,028,5617	2,0162	152,005,532	55,096,296	3,173,818	4,61	49,08	19,10**	102,61**	
0,9177	1,0738	3,6586	3,0052	5,7930	3,352,1279	2,3513	1,789,525,753	1,013,635,329	33,710,284	9,22	87,20	38,19**	182,31**	

QUADRO 3 - Resumo dos Resultados Obtidos para $\sigma^2 = 2,00$

Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7	Col. 8	Col. 9	Col. 10	Col. 11	Col. 12	Col. 13	Col. 14
ϑ (C)	$-\bar{c}$	ϑ (S)	\bar{b}	\bar{x}	ϑ_1 (x)	ϑ_2 (x)	$\bar{\mu}_3$	$\bar{\mu}_4$	$\bar{\mu}_4$ (separado)	$\bar{\nu}_1$	$\bar{\nu}_2$	Teste t $p/Y_1 \neq 0$	Teste t $p/Y_2 \neq 0$
1,9971	1,0122	7,0469	3,0010	0,2801	38,9448	5,0022	- 3,154,841	243,479	23,734	-3,77	27,78	15,62**	56,08**
2,1545	0,8257	5,3545	3,2126	0,0188	120,1613	9,1929	-3,822,861	316,398	43,318	-2,90	18,91	12,01**	39,54**
2,0425	0,8277	7,3394	3,0408	0,0826	27,9409	8,5066	- 150,701	7,416	2,342	-1,02	6,50	4,23**	13,59**
2,1959	0,9992	5,4200	2,8142	0,8816	32,9177	5,7173	408,029	22,100	3,251	2,16	17,40	8,95**	36,38**
1,9541	1,0708	5,1838	2,8505	-2,7518	726,4195	4,1613	-154,155,537	36,529,584	1,583,056	-7,87	86,23	32,60**	138,47**
1,6739	1,0096	5,5053	3,0183	-2,6309	495,5569	5,0242	- 61,322,886	8,753,106	736,730	-5,56	32,64	23,03**	68,24**
1,8086	1,1064	6,9470	2,8027	-7,1719	4,870,2089	3,7881	-3,227,478,006	2.200.880,135	71.156,804	-9,58	89,79	39,35**	187,73**
1,7677	1,0895	6,5791	2,9595	0,8098	87,9855	4,1327	- 598,241	257,144	23,224	-0,72	30,22	2,98**	63,18**
2,2312	1,1104	5,5042	3,0178	-2,2705	1,388,8100	4,4564	-492,166,688	180,067,424	3,786,380	-9,51	98,36	39,40**	188,92*
2,2495	0,9184	5,1576	3,1746	1,0184	31,5812	9,5714	1,632,588	164,260	19,967	2,22	21,68	9,20**	45,33**

QUADRO 6 - Resumo dos Resultados Obtidos para $\sigma^2 = 4,00$

Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7	Col. 8	Col. 9	Col. 10	Col. 11	Col. 12	Col. 13	Col. 14
\bar{y} (C)	$-c$	\bar{y} (S)	b	\bar{x}	\bar{y}_1 (x)	\bar{y}_2 (x)	\bar{u}_3	\bar{u}_4	\bar{u}_4 (esperado)	\bar{y}_1	\bar{y}_2	Teste t $\frac{\bar{y}_1 - \bar{y}_2}{\sigma \sqrt{2}}$	Teste t $\frac{\bar{y}_1 - \bar{y}_2}{\sigma \sqrt{2}}$
3,5091	1,1320	13,9000	2,7200	-0,9383	284,1337	6,6627	-46,194,266	7,266,162	242,196	-9,2274	86,98	38,22**	181,83**
3,9161	0,8486	11,0280	2,7249	-0,3848	123,2472	17,8392	-6,992,120	552,228	45,570	-5,1103	33,35	21,17**	69,73**
3,6978	1,0035	13,6877	3,3139	0,0730	40,9853	13,5273	-430,100	29,204	5,039	-1,6392	14,38	6,79**	30,06**
4,6132	0,3616	12,9305	3,1684	-1,2340	164,6979	17,0315	-13,880,822	1,857,494	81,386	-7,5135	65,48	31,12**	136,90**
3,0048	1,1702	12,1102	3,1888	2,4651	124,5799	6,2838	7,721,986	577,983	46,560	5,5534	34,24	23,01**	71,55**
4,4540	1,0031	13,0698	3,1627	2,3210	403,6660	14,2474	68,945,322	13,241,322	488,839	8,5011	78,26	35,72**	163,62**
4,5310	0,9254	11,8723	3,9353	0,5466	21,6509	16,7703	108,880	6,626	1,406	1,0808	11,14	4,48**	23,29**
3,7138	0,8477	12,9238	3,2810	1,4207	189,4187	23,8431	19,215,109	2,427,164	107,638	7,3707	64,65	30,53**	135,17**
3,4296	0,9937	11,5733	2,8211	0,4740	34,4348	10,1810	527,473	32,740	3,557	2,6104	24,61	10,81**	51,45**
4,3318	0,9876	11,5603	2,7176	-0,1688	119,7670	11,3700	-4,274,158	514,014	43,032	-3,2610	37,83	13,51**	68,64**

QUADRO 7 - Resumo dos Resultados Obtidos para $\sigma^2 = 6,25$

Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7	Col. 8	Col. 9	Col. 10	Col. 11	Col. 12	Col. 13	Col. 14
\bar{y} (s)	\bar{c}	\bar{y} (s)	\bar{c}	\bar{x}	\bar{y}_1 (x)	\bar{y}_2 (x)	\bar{u}_3	\bar{u}_4	\bar{u}_4 (esperado)	\bar{y}_1	\bar{y}_2	Teste t	Teste t
												$\mu/\gamma_1 \neq 0$	$\mu/\gamma_2 \neq 0$
5,9221	1,0767	16,9388	2,6518	1,1408	24,85	11,40	294,04	10,311	1,853	2,3732	13,69	9,83**	28,62**
6,3830	0,9572	22,5175	3,0452	-0,8935	938,12	23,76	-161,325,69	51,675,840	2,640,189	-5,6146	55,72	23,26**	116,50**
5,2720	0,8937	18,9316	3,0707	0,1862	71,02	25,40	229,82	82,974	15,131	0,3340	13,45	1,59	28,12**
5,8955	1,1097	17,5151	3,2124	-55,3963	311,229,38	13,58	-1,701,280,515,50	9,396,736,258,047	290,591,180,926	-9,7900	94,00	40,56**	196,53**
6,1329	1,0685	18,9765	3,0636	0,2788	20,09	15,19	85,26	3,922	1,211	0,9465	6,71	3,92**	14,04**
6,7632	1,1391	21,0554	2,9729	1,5086	326,83	12,93	49,702,34	8,670,216	320,454	8,4119	78,17	34,85**	163,43**
6,2085	1,2433	18,5253	3,0371	4,0964	1,112,22	8,99	296,661,88	91,531,088	3,711,114	7,9979	70,99	33,13**	148,42**
7,4505	0,9337	17,4389	3,1982	1,7533	255,97	30,06	25,740,16	3,478,403	196,558	6,2854	50,09	26,04**	104,72**
7,3862	0,8814	15,7760	2,9913	-0,4428	24,85	32,45	42,82	14,489	1,853	0,3456	20,46	1,43	42,78**
6,1296	1,0079	19,4300	2,7650	6,7302	2,997,05	16,13	1,537,299,25	819,640,962	26,947,009	9,3695	88,25	38,81**	186,51**

QUADRO 8 - Resumo dos Resultados Obtidos para $\sigma^2 = 9,00$

Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7	Col. 8	Col. 9	Col. 10	Col. 11	Col. 12	Col. 13	Col. 14
\bar{y} (t)	$-\bar{z}$	\bar{y} (t)	\bar{b}	\bar{x}	$\bar{y}_1(x)$	$\bar{y}_2(x)$	$\bar{\mu}_3$	$\bar{\mu}_4$	$\bar{\mu}_4$ (esperado)	\bar{r}_1	\bar{r}_2	Teste t $\bar{r}_1 \neq 0$	Teste t $\bar{r}_2 \neq 0$
9,9433	1,0316	32,3583	3,3837	1,0849	117,39	32,73	10,344,246	1,082,372	41,340	8,1332	74,11	33,69**	154,94**
8,9258	0,9285	22,9715	3,1819	1,2823	231,48	37,06	24,761,897	3,724,009	160,748	7,0310	66,50	29,12**	139,03**
8,0070	1,0669	20,2267	2,6316	-3,3814	628,64	15,14	-117,815,984	24,267,596	1.185,572	-7,4748	58,41	30,96**	122,12**
10,0468	1,2321	24,4167	2,8016	3,6344	1,279,85	12,57	433,277,376	151,966,784	4,914,063	9,4630	89,77	39,20**	187,68**
10,0173	0,9815	24,1180	2,7392	-0,3771	49,29	26,51	41,370,787	63,934	7,289	-3,9611	23,31	16,41**	48,74**
8,9872	1,0260	31,7109	3,0082	2,6476	569,78	25,81	125,937,766	29,246,488	973,939	9,2597	87,09	38,36**	182,08**
9,6960	0,8424	24,0952	3,4511	5,3792	3,539,91	65,81	1,922,991,253	1.121,797,123	37.592,922	9,1304	86,52	37,82**	180,89**
9,1916	0,8466	33,0275	3,0866	0,8590	81,30	54,12	2,031,535	136,669	19,831	2,7711	17,67	18,48**	36,94**
9,8817	0,9984	24,3800	2,6060	1,2637	366,91	22,99	63,204,711	11,793,328	403,866	9,0074	84,60	37,31**	176,88**
8,8385	1,1503	23,2641	2,6830	1,3622	554,40	13,48	71,818,750	18,710,568	922,095	5,5017	57,87	22,79**	120,99**

QUADRO 9 - Intervalos de Confiança para Valores de \bar{x} .

σ^2	Intervalo de confiança pelo método de Fieller	\bar{x}	Intervalo de confiança empírico	Intervalo de confiança baseado em $\hat{\sigma}_1(x)$	Intervalo de confiança baseado em $\hat{\sigma}_2(x)$
0,015625	[1,14 ; 2,06]	1,53	[1,16 ; 2,02]	[1,08 ; 1,98]	[1,10 ; 1,96]
0,0625	[0,88 ; 3,12]	1,61	[0,91 ; 3,07]	[0,44 ; 2,78]	[0,76 ; 2,46]
0,2500	[0,50 ; + ∞]	1,75	[-3,00 ; 9,59]	[-95,8 ; 99,4]	[-0,01 ; 3,51]
1,0000	[- ∞ ; + ∞]	1,11	[-14,89 ; 19,27]	[-80,8 ; 83,0]	[-2,21 ; 4,43]
2,0000	(- ∞ ; + ∞)	1,60	[-14,61 ; 14,48]	[-54,6 ; 57,8]	[-3,3 ; 6,5]
4,0000	(- ∞ ; + ∞)	0,46	[-11,54 ; 14,46]	[-24,1 ; 25,0]	[-7,0 ; 7,9]
6,2500	(- ∞ ; + ∞)	-4,10	[-11,60 ; 17,20]	[-360,2 ; 352,0]	[-12,8 ; 4,6]
9,0000	(- ∞ ; + ∞)	1,38	[-12,48 ; 9,38]	[-53,1 ; 55,9]	[-9,7 ; 12,4]

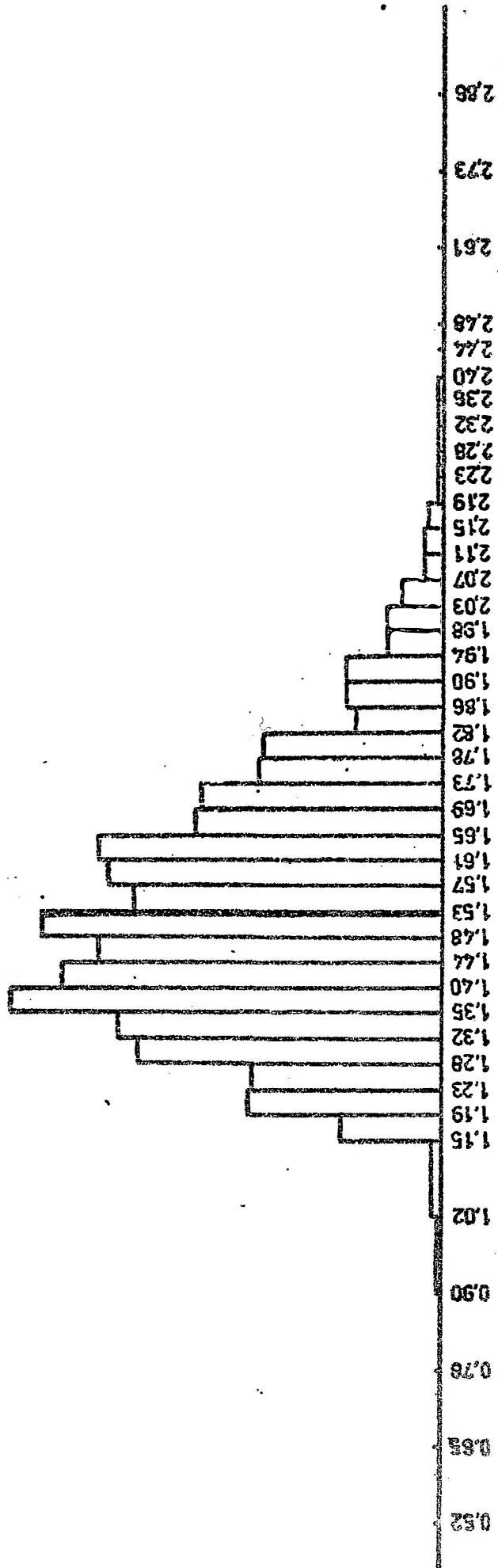


Gráfico 1 - Histograma de Frequências de 1.000 Valores de x Relativos ao Caso de $\sigma^2 = 0,015625$ e Média 1,5253.

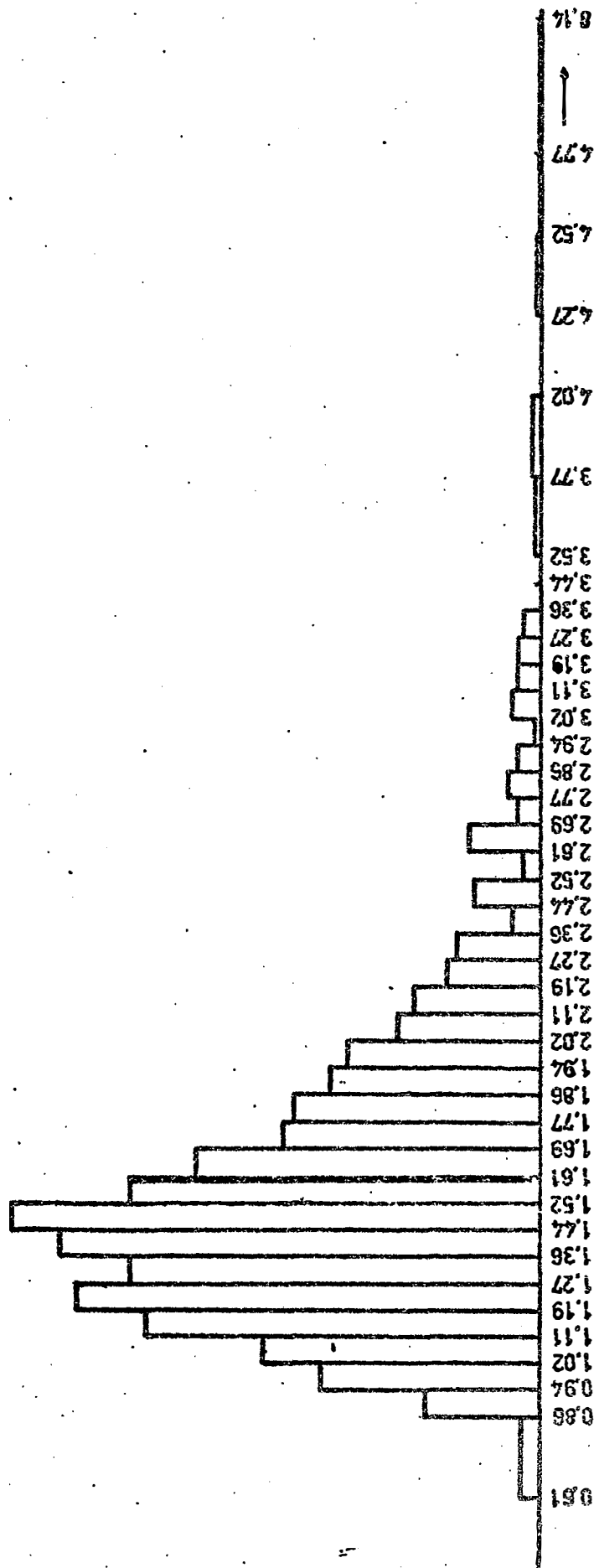
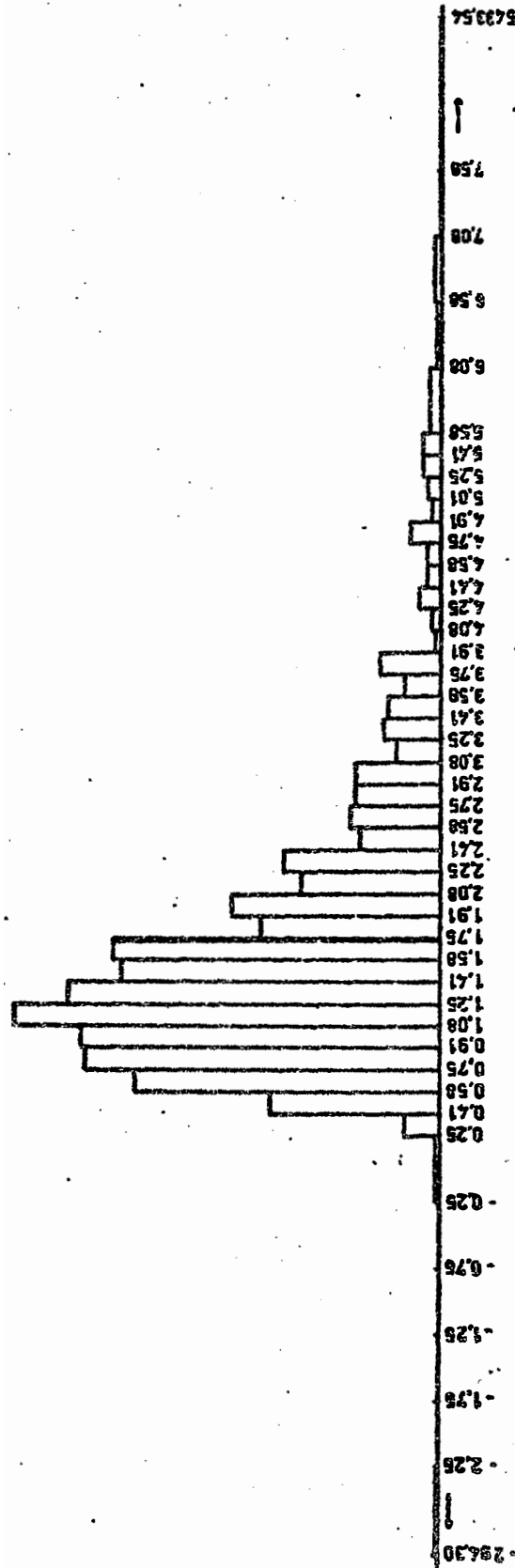


Gráfico 2 - Histograma de Frequências de 1.000 Valores de x Relativos ao Caso de $\sigma^2 = 0,0625$ e Média = 1,6079.



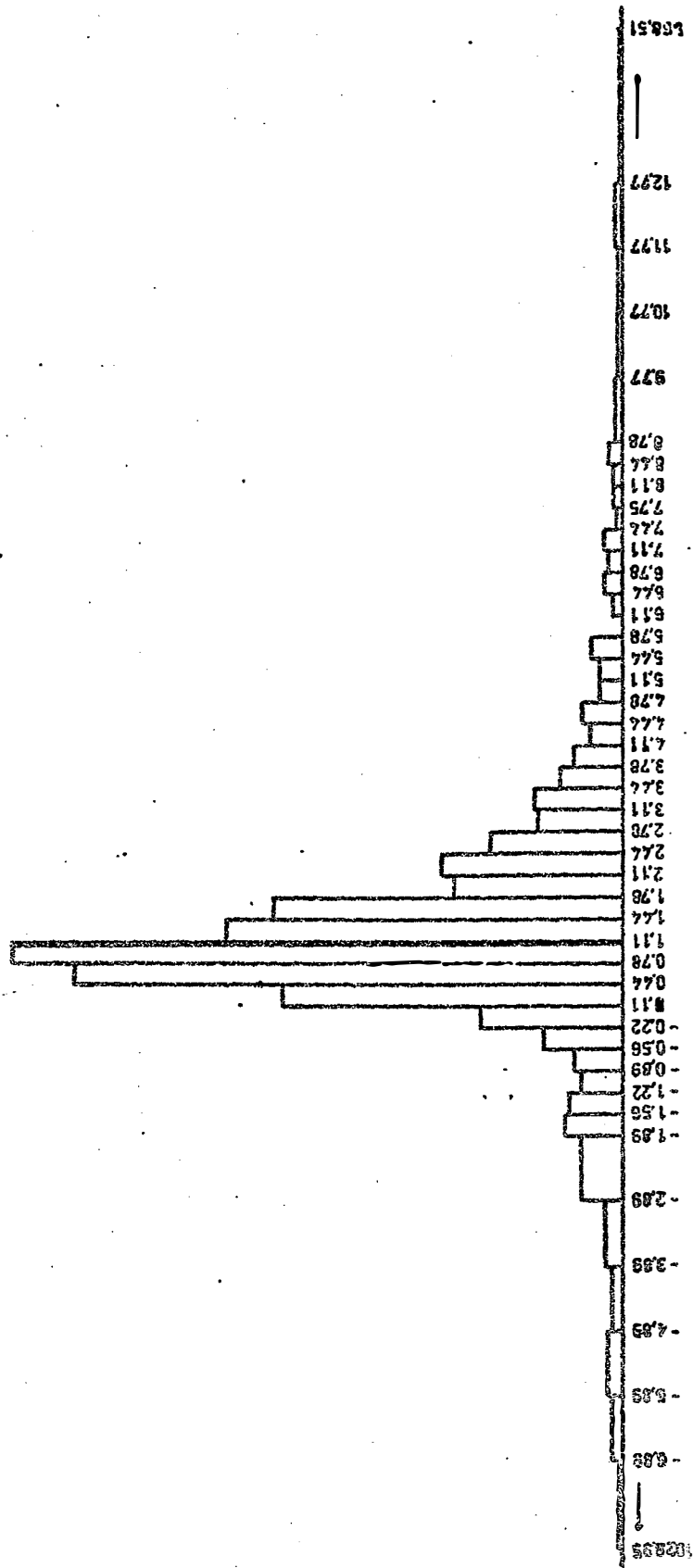


Gráfico 4 - Histograma de Frequências de 1.000 Valores de x Relativos ao Caso de $\sigma^2 = 1,00$ e Média = 1,1091.

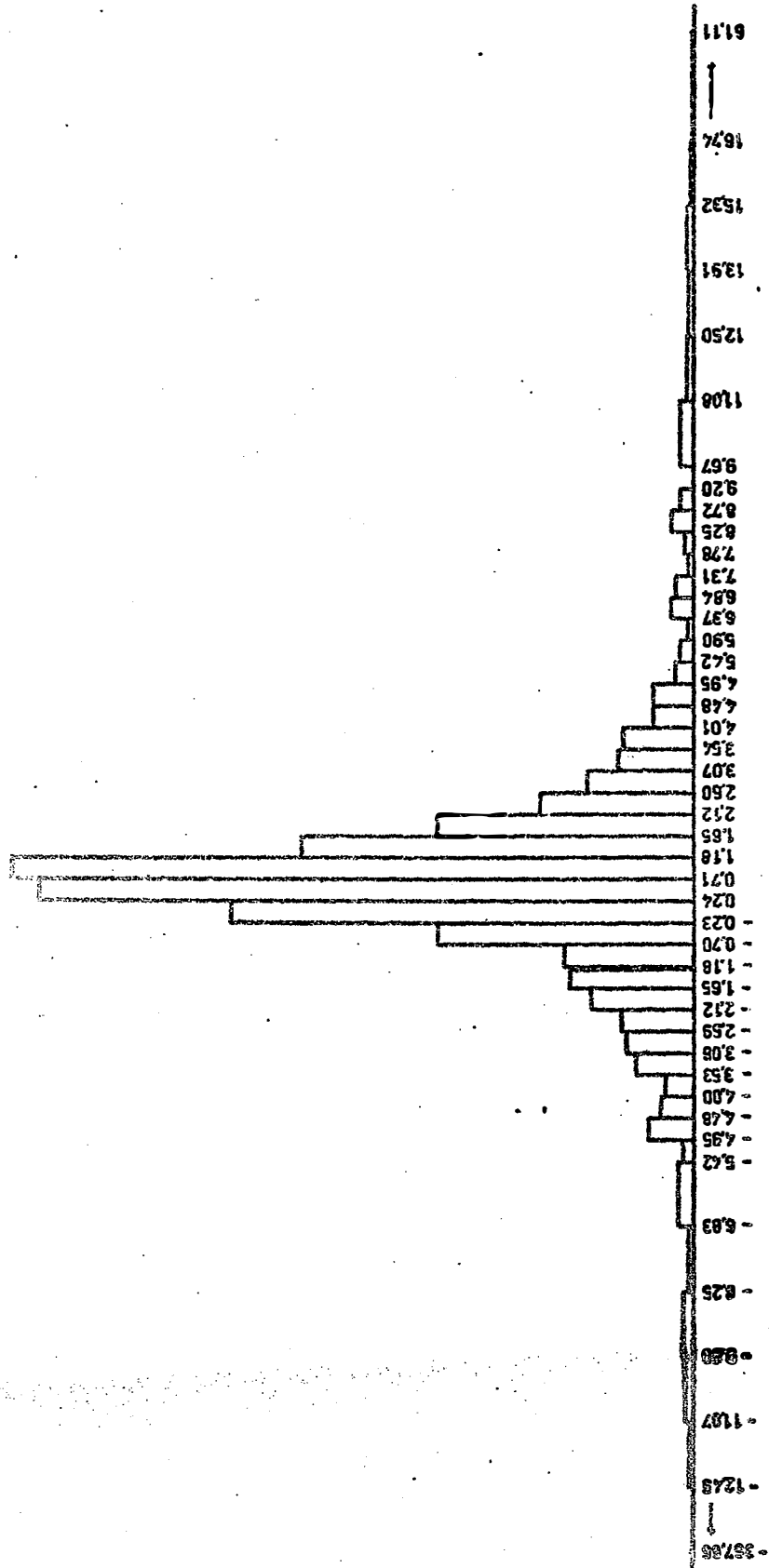


Gráfico 5 - Histograma de Frequências de 1.000 Valores de x Relativos ao Caso de $\sigma^2 = 2,00$ e Média = 1,1754.

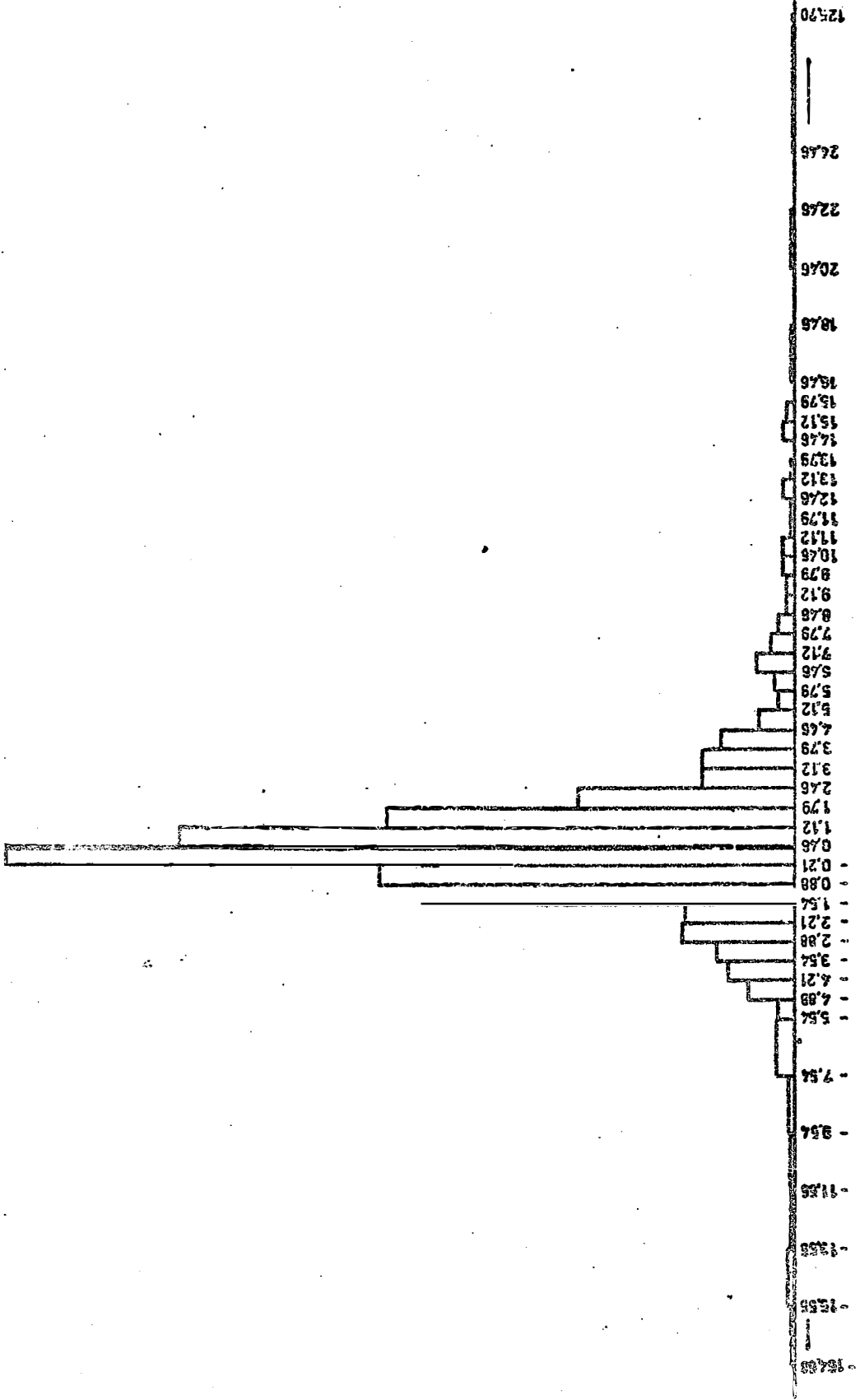


Gráfico 6 - Histograma de Frequências de 1.000 Valores de \bar{x} Relativos ao Caso de $\sigma^2 = 4,00$ e Média = $0,4554$.

Gráfico 7 - Histograma de Frequências de 1.000 Valores de \bar{x} Relativos ao Caso de $\sigma^2 = 6,25$ e Média = $-4,1039$.

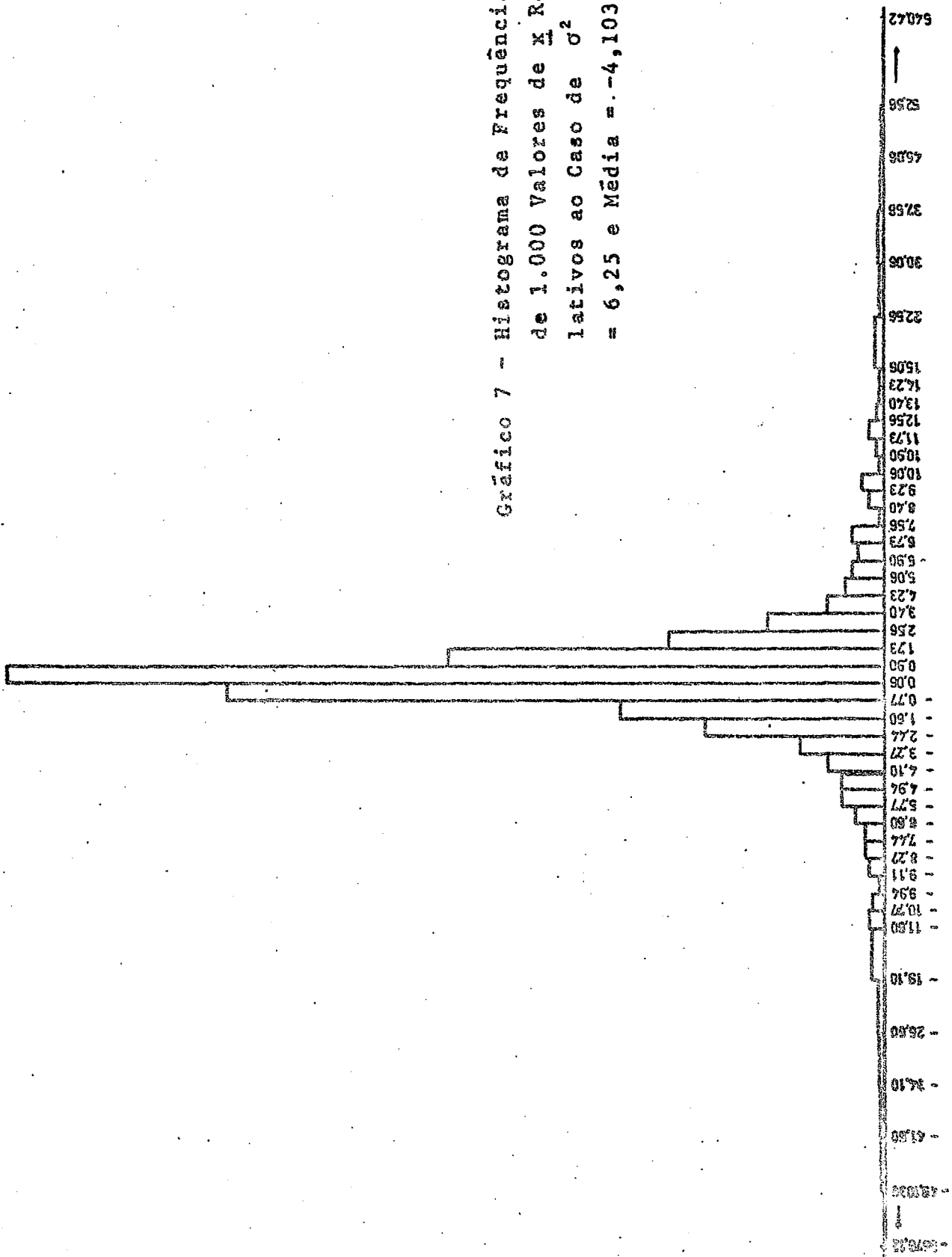
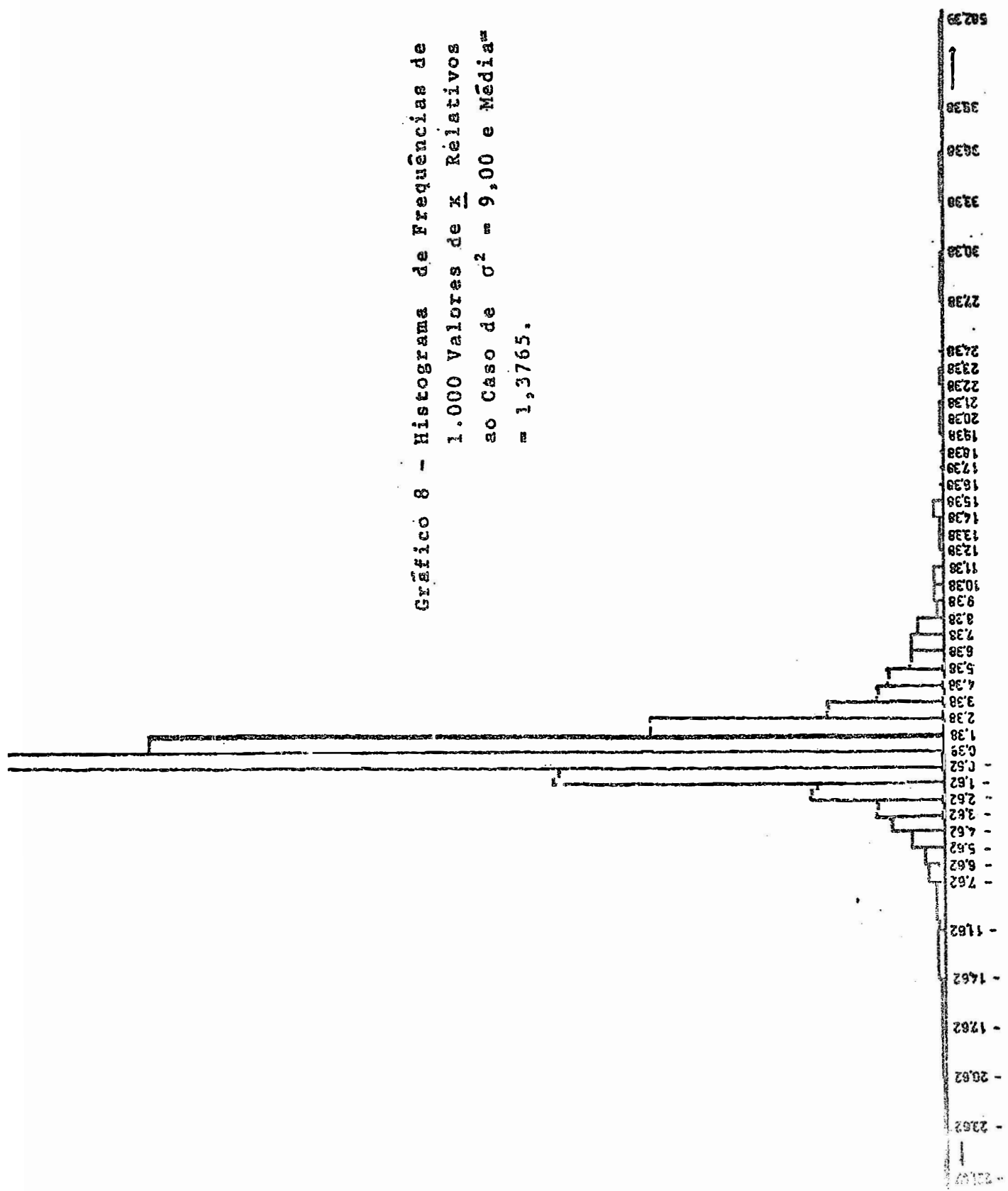


Gráfico 8 - Histograma de Frequências de
1.000 Valores de \bar{x} Relativos
ao Caso de $\sigma^2 = 9,00$ e Média
 $= 1,3765$.



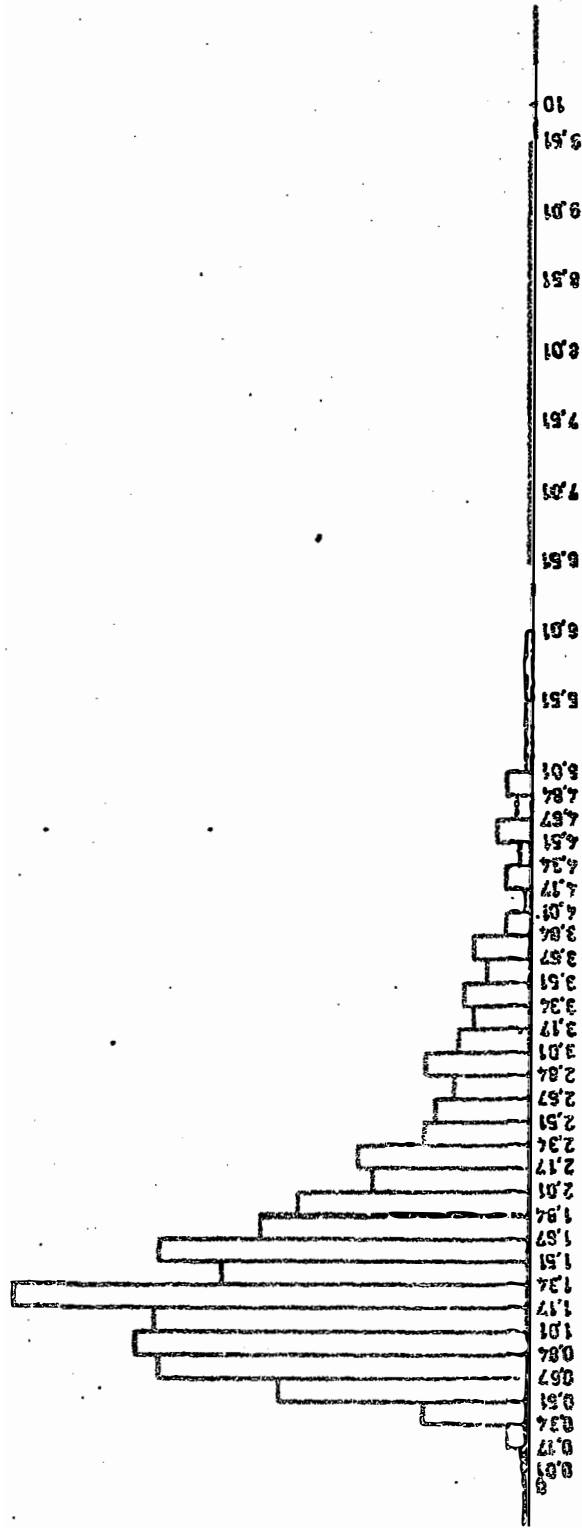


Gráfico 9 - Histograma de Frequências dos Valores de \bar{x} Maiores que Zero e Menores que Dez Relativos ao Caso $\sigma^2 = 0,2500$ e Média = 1,8404 ($0 < x < 10$).

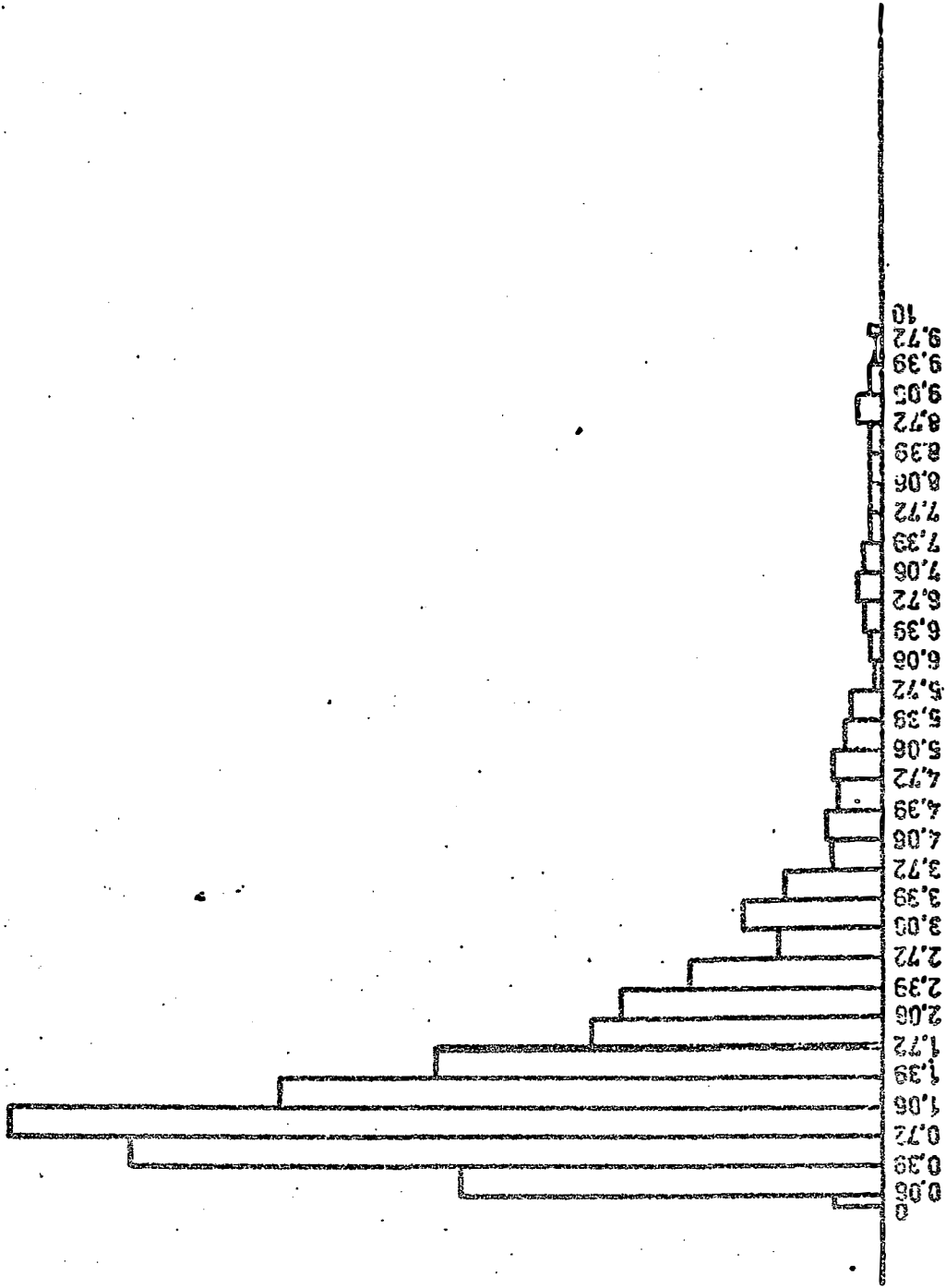


Gráfico 10 - Histograma de Frequências dos Valores de x Maiores que Zero e Menores que Dez Relativos ao Caso $\sigma^2 = 1,00$ e Média $\mu = 1,7230$ ($0 < x < 10$).

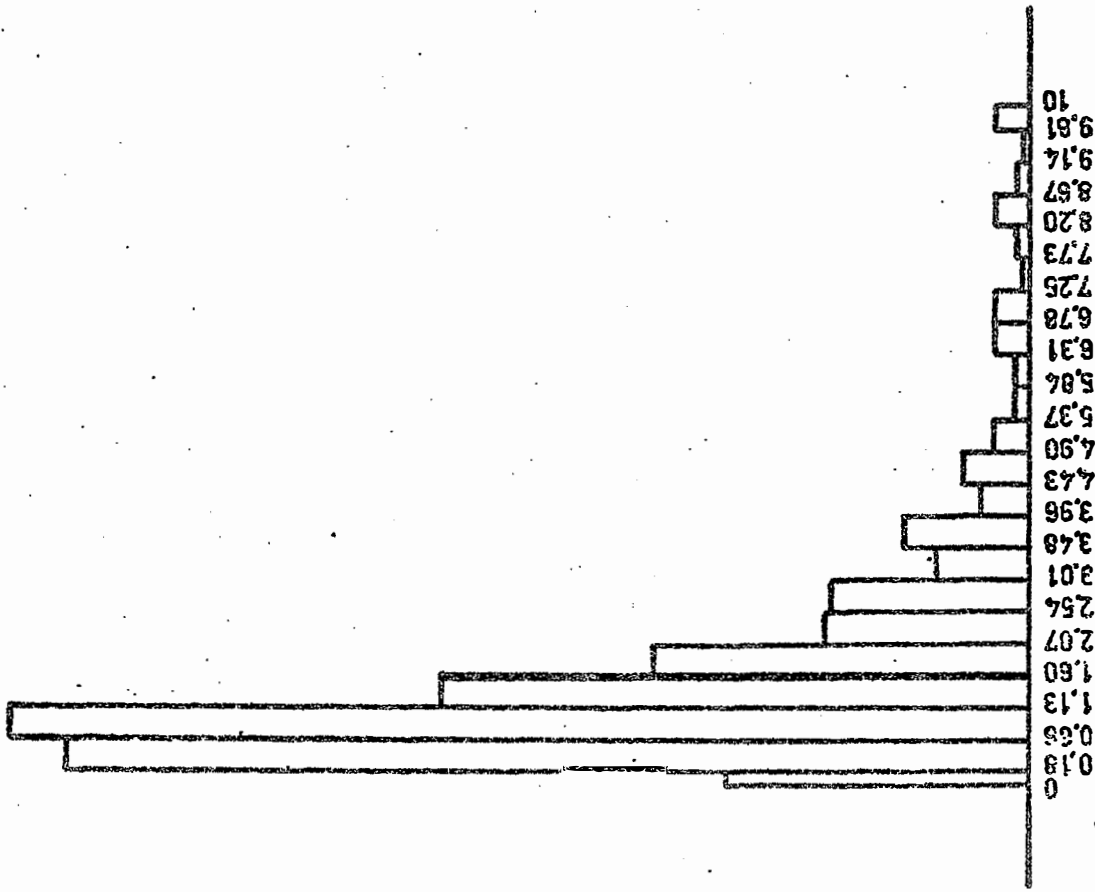


Gráfico 12 - Histograma de Frequências dos Valores de x Maiores que Zero e Menores que Dez Relativos ao Caso $\sigma^2 = 2,00$ e Média = 1,5981 ($0 < x < 10$).

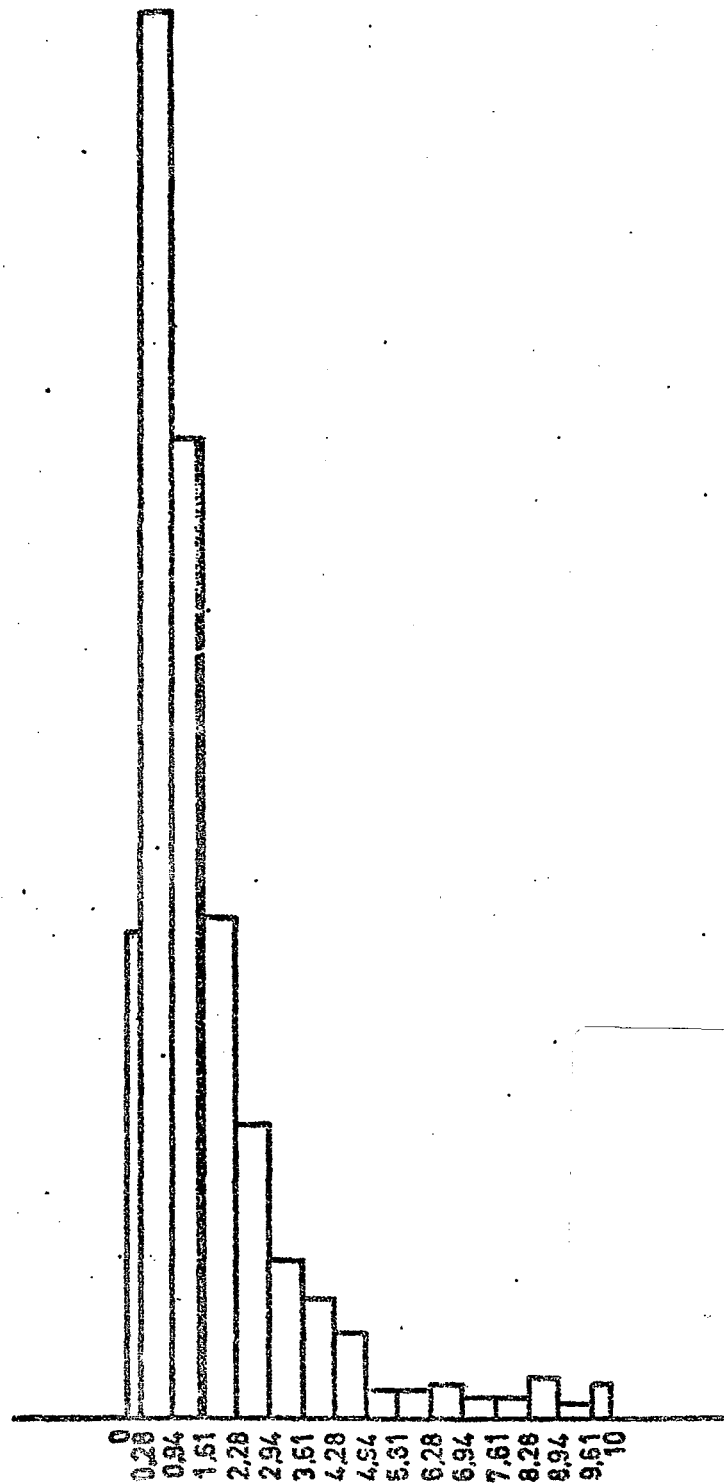


Gráfico 12 - Histograma de Frequências dos Valores de x Maiores que Zero e Menores que Dez Relativos ao Caso $\sigma^2 = 4,00$ e Média = 1,6089 ($0 < x < 10$).

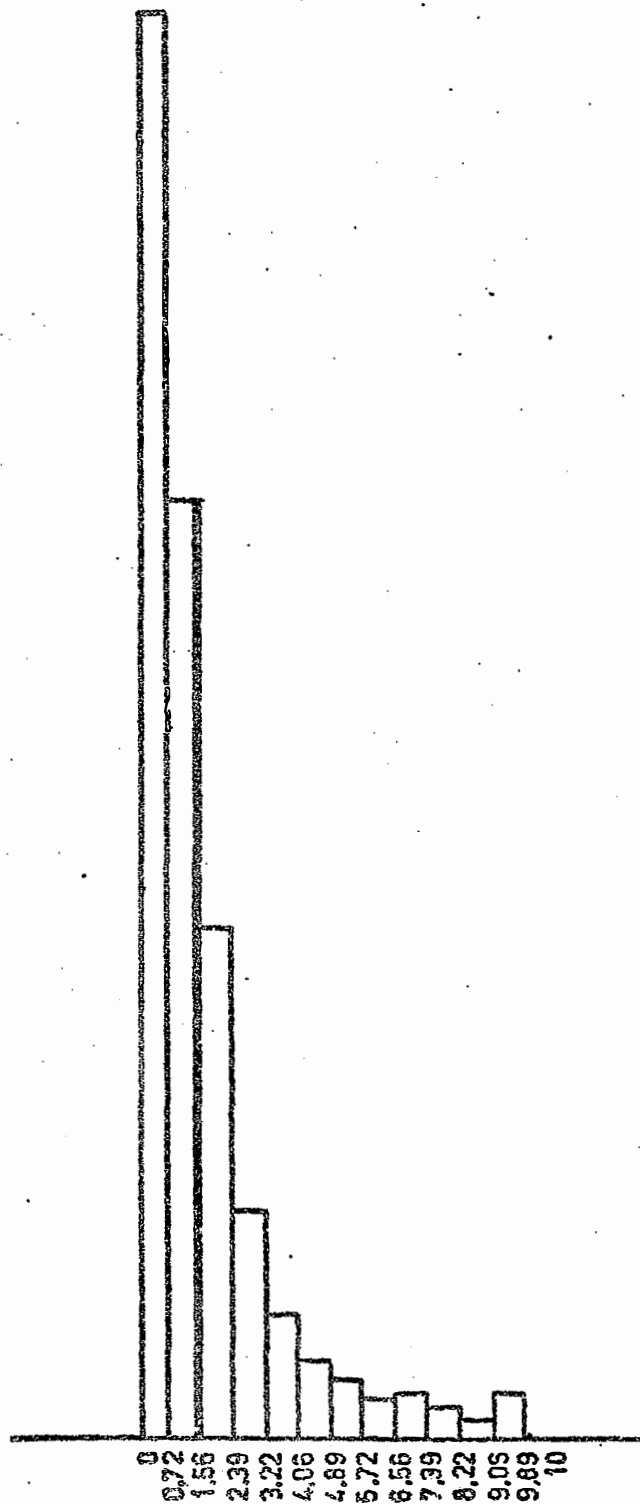


Gráfico 13 - Histograma de Frequências dos Valores de x Maiores que Zero e Menores que Dez Relativos ao Caso $\sigma^2 = 6,25$ e Média = 1,5558 ($0 < x \leq 10$).

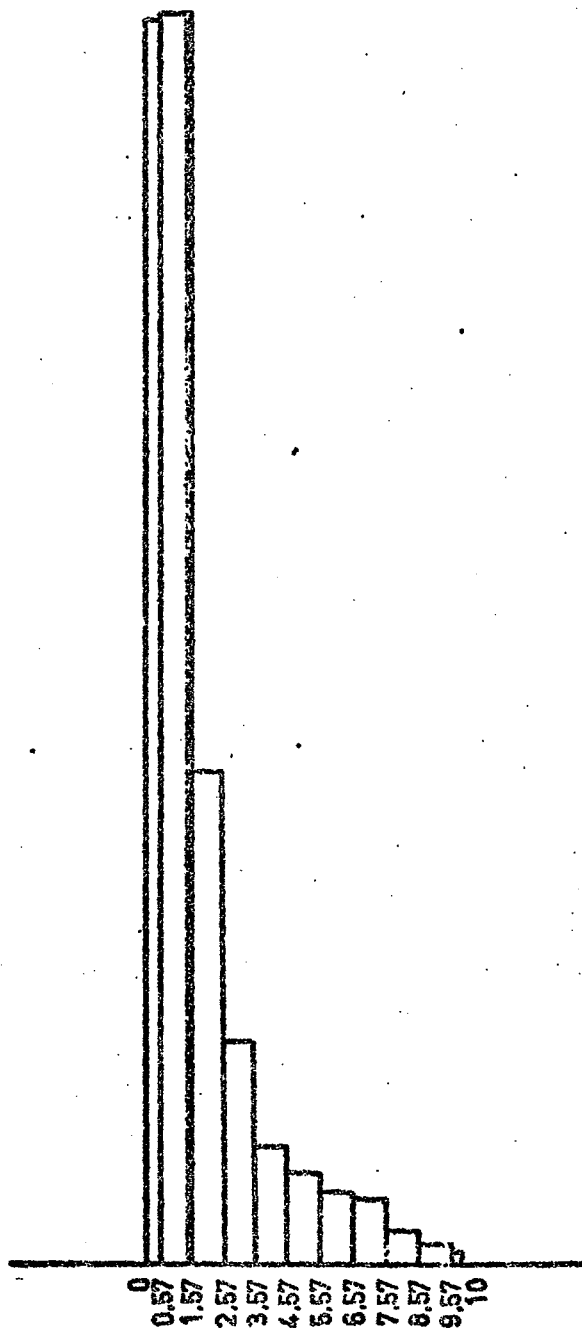


Gráfico 14 - Histograma de Frequências dos Valores de x Maiores que Zero e Menores que Dez Relativos ao Caso $\sigma^2 = 9,00$ e Média = $1,5736$ ($0 < x < 10$).

5. CONCLUSÕES

5.1) Só para o valor mais baixo de σ^2 ensaiado (0,015625, que dá um coeficiente de variação de 12,5% para \hat{c}), a normalidade aproximada da distribuição do quociente (\underline{x}) é aceitável. Nos demais casos \underline{x} foge completamente à distribuição normal.

5.2) Para os níveis mais baixos de σ^2 , são positivos os valores de $\hat{\gamma}_1$ relativos à distribuição de \underline{x} . À medida, porém, que cresce σ^2 , surgem valores positivos e também negativos para $\hat{\gamma}_1$.

5.3) Ao contrário, no caso de $\hat{\gamma}_2$ para os valores mais baixos de σ^2 , ocorrem estimativas positivas e negativas. No entanto, quando cresce σ^2 , passam a ocorrer somente valores positivos de $\hat{\gamma}_2$, o que caracteriza as distribuições leptocúrticas.

5.4) Para valores baixos de σ^2 , relativamente a \underline{c} (C.V. de 12,5% e de 25%) o método de Fieller e a simulação levam a intervalos de confiança semelhantes. Nos demais casos, porém (C.V. de 50% ou mais) o método de Fieller dá resultados estapafúrdios, ao passo que o obtido por simulação dos dados conduz a resultados aparentemente razoáveis.

5.5) Os valores de $V_2(x)$ foram, em praticamente todos os casos estudados, bem inferiores aos obtidos para $V_1(x)$, e conseqüentemente, conduziram a determinação de intervalos de confiança mais curtos.

5.6) Os intervalos de confiança mais recomendáveis são os calculados pelo método empírico e o obtido a partir de $V_2(x)$, embora este seja um tanto curto.

6. RESUMO

No presente trabalho procuraram-se avaliar os erros experimentais na determinação dos pontos de máximo ou de mínimo de equações de regressão de 2º grau e estudar a distribuição probabilística a que pertencem. Geraram-se, para isso, através da sub-rotina CALL RANDU, 16.000 dados de distribuição aproximadamente normal, que tiveram suas médias e variâncias ajustadas, para formarem os 8 casos estudados. Obtivemos, assim, 1.000 valores que chamamos de \hat{b} e 1.000 valores que chamamos de \hat{c} , para cada σ^2 considerado, valores esses relativos à equação

$$Y = \hat{a} + \hat{b} P_1(x) + \hat{c} P_2(x) ,$$

onde $P_1(x)$ e $P_2(x)$ são polinômios ortogonais. Portanto \hat{b} e \hat{c} são independentes. O ponto de máximo ou de mínimo (x) des

sa equação será dado por

$$x = - \frac{\hat{b}}{2 \hat{c}} .$$

Portanto, com os dados gerados obtivemos um total de 8.000 valores de \underline{x} relativos às diferentes variâncias estudadas: 0,015625; 0,0625; 0,2500; 1,0000; 2,0000; 4,0000; 6,2500; 9,0000, sendo 1.000 valores para cada variância. Com esses valores calculamos as estimativas de variância $\hat{V}_1(x)$ (fórmula comum de cálculo de variância) e $\hat{V}_2(x)$ (fórmula apresentada por D'AULÍSIO, 1970) e se pode observar que os valores de $\hat{V}_2(x)$ subestimam os de $\hat{V}_1(x)$. Calculou-se o terceiro e o quarto momento em relação à média, e também o quarto momento que deveríamos esperar se a distribuição fosse normal. Obtivemos ainda os valores de $\hat{\gamma}_1$ (que mede a assimetria) e de $\hat{\gamma}_2$ (que mede a curtose da distribuição), aos quais aplicamos o teste de \underline{t} . Concluimos que a distribuição de \underline{x} foge completamente da distribuição normal, exceto talvez para o valor mais baixo de σ^2 estudado.

Obtivemos quatro intervalos de confiança para \underline{x} , a saber: 1) um calculado pelo método de Fieller; 2) de uma maneira empírica, considerando o equivalente a um intervalo de confiança ao nível de 5% de probabilidade, tomando para cada 1.000 dados, o maior e o menor valor observados após a eliminação dos 25 maiores e dos 25 menores; 3 e 4) os outros

dois métodos usados aplicaram as fórmulas $\bar{x} \pm t \sqrt{V_1(x)}$, $\bar{x} \pm t \sqrt{V_2(x)}$. Concluimos que o método de Fieller nos levou a resultados absurdos para $\sigma^2 \geq 0,2500$. Em relação aos intervalos calculados pelo método empírico, tomados como padrão, foram excessivamente amplos os dados por $V_1(x)$, e excessivamente curtos os obtidos a partir de $V_2(x)$.

7. ABSTRACT

This paper has in view the evaluation of experimental errors in the estimation of the maximum or minimum point of second degree regression equations, and also the study of their distribution. For this purpose, 16,000 data were obtained, with their mean and variance adjusted to be included in the 8 cases studied. So, we had 1,000 values for \hat{b} and 1,000 for \hat{c} , for each of 8 values of σ^2 selected, referring to equation.

$$Y = \hat{a} + \hat{b} P_1(x) + \hat{c} P_2(x) ,$$

where $P_1(x)$ and $P_2(x)$ are orthogonal polynomials. Therefore, \hat{b} and \hat{c} are independent from each other. The maximum or minimum point of that equation is

$$x = - \frac{\hat{b}}{2 \hat{c}} .$$

So, with the data obtained by simulation, we obtained 8,000 values of \underline{x} , being 1,000 referring to each of the following values of $\sigma^2 = 0.015625; 0.0625; 0.2500; 1.0000; 2.0000; 4.0000; 6.2500$ and 9.0000 . With the \underline{x} values thus obtained, estimates of variance $\hat{V}_1(x)$, by the usual formula, as well as $\hat{V}_2(x)$, by a formula presented by D'AULÍSIO (1970), were calculated. It could be observed that in almost every case we had $\hat{V}_1(x) > \hat{V}_2(x)$.

The third and fourth moments, about the mean, of \underline{x} , were calculated, as well as the values of Fisher's $\hat{\gamma}_1$ and $\hat{\gamma}_2$, to which the t test was applied. Thus, it was shown that, except for the case of $\sigma^2 = 0.015625$, the distribution of \underline{x} was very far from normal.

Four types of confidence intervals for \underline{x} were obtained: 1) calculated by Fieller's method; 2) by taking the 950 values closer to the mean, of the 1,000 values of \underline{x} in each case; 3 and 4) by taking $\bar{x} \pm t \sqrt{\hat{V}_1(x)}$, $\bar{x} \pm t \sqrt{\hat{V}_2(x)}$. Fieller's method led to bizarre results for $\sigma^2 \geq 0.2500$. The intervals obtained with $\hat{V}_1(x)$ were too long, while those provided by $\hat{V}_2(x)$ were too short.

8. BIBLIOGRAFIA CITADA

ANDERSON, R.L. e L.A. Nelson - 1971 - Some Problems in the Estimation of Single Nutrient Response Functions. Reprint 244. Institute of Statistics, Raleigh.

BAUM, E.L.; E.O. Heady; J.T. Pesek e C.G. Hildreth - 1956 - Fertilizer Innovations and Resource Use. Iowa State College Press. Ames, Iowa.

D'AULÍSIO, M. de B.G. - 1970 - "A Variância dos Pontos de Máximo ou de Mínimo de Equações de Regressão de Segundo Grau". Resumos XXII Reunião Anual - SBPC (Nota prévia).

FINNEY, D.J. - 1975 - Bioassay. Trabalhos apresentados na Reunião Internacional Comemorativa do 20º Aniversário de Fundação da Região Brasileira da Sociedade Internacional de Biometria ("The Biometric Society"), realizada em Piracicaba, de 22 a 27 de setembro de 1975.

FISHER, R.A. - 1930 - The Moments of the Distribution for Normal Samples of Measures of Departure from Normality. Royal Society, Série A, 130:17-28.

KENDALL, Maurice G. - 1947 - The Advanced Theory of Statistics, vol. I. Charles Griffin & Company, Londres.

KENDALL, Maurice G. e Alan Stuart - 1963 - The Advanced Theory of Statistics, vols. 1 e 2, Charles Griffin & Company, Londres.

LOBO, Ary Maurell - 1948 - Tratado Teórico e Prático de Estatística das Grandes Massas de Dados e de Estatística das Mostras. Biblioteca Profissional Brasileira. Rio de Janeiro.

SPIEGEL, Murray R. - 1971 - Estatística, McGraw-Hill do Brasil, Rio de Janeiro.