

ANÁLISE ESTATÍSTICA DOS ERROS DE CENTRAGEM E PONTARIA

Jair Mendes Marques
 Universidade Tuiuti do Paraná
 R. Marcelino Champagnat, 505 – CEP 80710-250
 e-mail: jairmm@utp.br

RESUMO

O objetivo deste trabalho consiste no desenvolvimento de uma análise estatística dos erros e centragem dos instrumentos e também dos erros de pontaria. O desenvolvimento da função densidade de probabilidade dos erros permite uma análise inferencial dos mesmos. São realizadas algumas simulações e a análises dos resultados obtidos.

ABSTRACT

This paper deals with the statistical analysis for the centring errors of the instrument and the target. The probability density function of errors allows an analysis of the same ones. Simulations and analyses of the obtained results are accomplished.

1. INTRODUÇÃO

Quando do uso de instrumentos que necessitam de centragem, envolvendo a obtenção de dados de alta precisão, o estudo do erro de centragem é fundamental para que se possa conhecer melhor o comportamento do mesmo, bem como sua possibilidade de correção. No caso do erro de pontaria, o estudo e a análise são semelhantes ao do erro de centragem.

A importância do conhecimento do comportamento estatístico dos erros de centragem e pontaria está relacionado às estimativas dos mesmos, tornando-os possíveis de controle.

2. A DISTRIBUIÇÃO DE RAYLEIGH

Seja a função densidade de probabilidade de uma distribuição normal bidimensional

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 \right]\right\} \quad (2.1)$$

onde $\mu_x = E(X)$, $\mu_y = E(Y)$, $\sigma_x^2 = V(X)$, $\sigma_y^2 = V(Y)$ e $\rho = \text{corr}(X, Y)$.

Bol. Ciênc. Geod., Curitiba, v. 5, p.15-26, 1999.

Considerando que o ponto de centragem corresponde a $P(\mu_x, \mu_y)$ onde $\mu_x = \mu_y = 0$ e que $\rho = 0$, ou seja, que o erro de fixação do ponto P independe das direções de X e Y, então, para garantir que qualquer par de eixos escolhidos fornece $\rho = 0$, deve-se ter $\sigma_x^2 = \sigma_y^2 = \sigma^2$.

Com as suposições propostas, a (2.1) fica

$$f(x, y) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2\sigma^2}(x^2 + y^2)\right], \quad -\infty < x < +\infty, \quad -\infty < y < +\infty \quad (2.2)$$

que passando para coordenadas polares, obtém-se

$$g(r, \theta) = f[G_1(r, \theta), G_2(r, \theta)] |J(r, \theta)| \quad (2.3)$$

onde

$$\begin{aligned} f[G_1(r, \theta), G_2(r, \theta)] &= \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2\sigma^2}(r^2 \cos^2 \theta + r^2 \sin^2 \theta)\right] \\ &= \frac{1}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad r \geq 0, \quad 0 \leq \theta \leq 2\pi. \end{aligned} \quad (2.4)$$

e

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \cdot \operatorname{sen} \theta \\ \operatorname{sen} \theta & r \cdot \cos \theta \end{vmatrix} = r \cdot \cos^2 \theta + r \cdot \operatorname{sen}^2 \theta = r. \quad ((2.5))$$

Substituindo a (2.4) e (2.5) em (2.3) resulta

$$g(r, \theta) = \frac{r}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad r \geq 0, \quad 0 \leq \theta \leq 2\pi \quad (2.6)$$

sendo a função densidade marginal de θ dada por

$$h_1(\theta) = \int_0^\infty g(r, \theta) dr = \int_0^\infty \frac{r}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) dr = \frac{1}{2\pi}, \quad 0 \leq \theta \leq 2\pi \quad (2.7)$$

ou seja, resulta numa função densidade uniforme. A função densidade marginal de r será

$$h_2(r) = \int_0^{2\pi} g(r, \theta) d\theta = \int_0^{2\pi} \frac{r}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) d\theta = \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad r \geq 0 \quad (2.8)$$

que é conhecida como **distribuição de Rayleigh**.

A figura 2.1 mostra graficamente a distribuição de Rayleigh para $\sigma = 1$.

As variáveis r e θ são independentes, pois

$$g(r, \theta) = h_1(\theta) \cdot h_2(r) = \frac{r}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right).$$

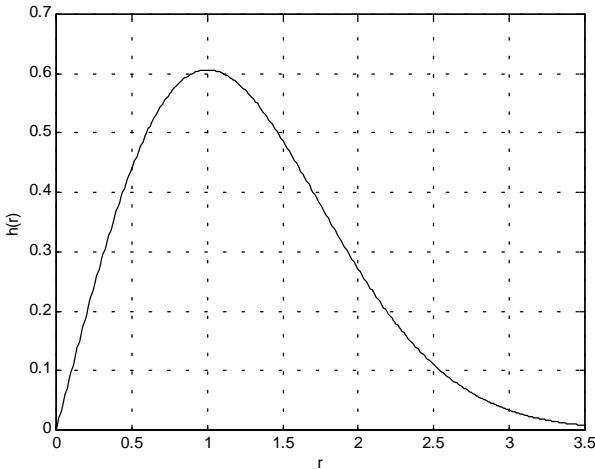


Figura 2.1 - Distribuição de Rayleigh

A esperança e variância de r são, respectivamente,

$$E(r) = \int r \cdot \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) dr = \sigma \sqrt{\frac{\pi}{2}} \quad (2.9)$$

e

$$V(r) = E(r^2) - [E(r)]^2 = 2\sigma^2 - \left(\sigma \sqrt{\frac{\pi}{2}}\right)^2 = \left(\frac{4-\pi}{2}\right)\sigma^2 = 0,43\sigma^2. \quad (2.10)$$

Genericamente, tem-se que

$$E(r^n) = 2^{\frac{n}{2}} \sigma^n \Gamma\left[\frac{n+2}{2}\right]. \quad (2.11)$$

Utilizando-se a transformação $v = \frac{r^2}{\sigma^2}$, resulta

$$h(v) = \frac{1}{2} \exp\left(-\frac{v}{2}\right), \quad v \geq 0 \quad (2.12)$$

ou seja, $v \sim \chi^2(2)$.

A função distribuição acumulada de Rayleigh será dada por

$$H(r) = P(R \leq r) = \int_0^r \frac{u}{\sigma^2} \exp\left(-\frac{u^2}{2\sigma^2}\right) du = 1 - \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad r \geq 0. \quad (2.13)$$

A figura 2.2 mostra graficamente a função distribuição acumulada de Rayleigh.

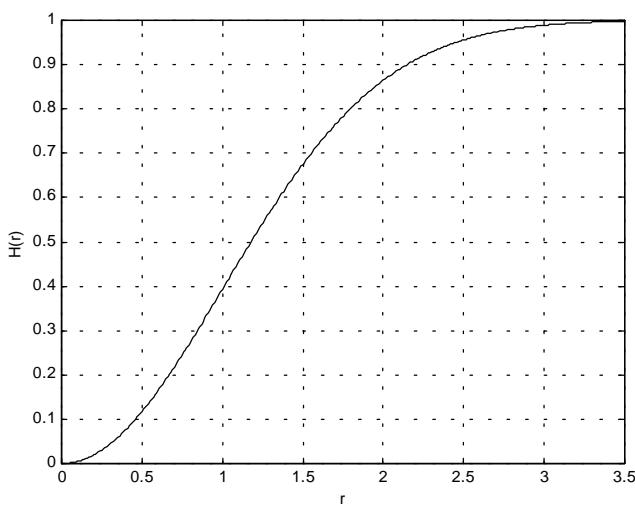


Figura 2.2 – Distribuição Acumulada de Rayleigh

3. RELAÇÃO ENTRE DIREÇÃO E DISTÂNCIA

Bol. Ciênc. Geod., Curitiba, v. 5, p.15-26, 1999.

A figura 3.1 permite relacionar direção e distância, onde I representa a posição do instrumento, supostamente sem erro de centragem, e A representa o alvo onde ocorre erro de pontaria, sendo a visualização acontece no ponto E.

De acordo com a construção da figura 3.1, tem-se que $r \cdot \sin \theta$ causa variabilidade na **distância**, enquanto que $r \cdot \cos \theta$ causa variabilidade na **direção**.

Assim, as variâncias de interesse na determinação das variâncias das direções e distâncias são σ_X^2 e σ_Y^2 , e não a variância de r .

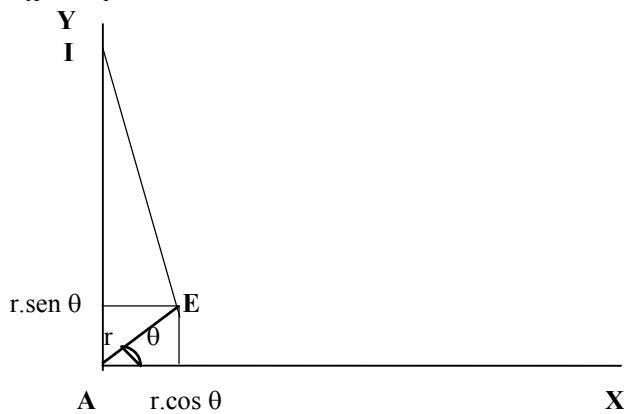


Figura 3.1 – Relação entre direção e distância

4. VARIÂNCIAS TOTAIS DE DIREÇÕES E DISTÂNCIAS

(I) Para as distâncias

Supondo que os erros de medidas sejam independentes dos erros de centragem do instrumento e que esses sejam independentes dos erros de pontaria, considere σ_I^2 = variância de centragem do instrumento, σ_A^2 = variância de centragem do alvo, σ_L^2 = variância da medida da distância e σ_D^2 = variância da medida da direção.

A variância total da distância será dada por

$$V(\text{Total}) = V(L - r_2 \cdot \sin \theta_2 + r_1 \cdot \sin \theta_1) = \sigma_L^2 + \sigma_I^2 + \sigma_A^2 . \quad (4.1)$$

Para $\sigma_I^2 = \sigma_A^2 = \sigma^2$, obtém-se

$$V(\text{total}) = \sigma_L^2 + 2\sigma^2 . \quad (4.2)$$

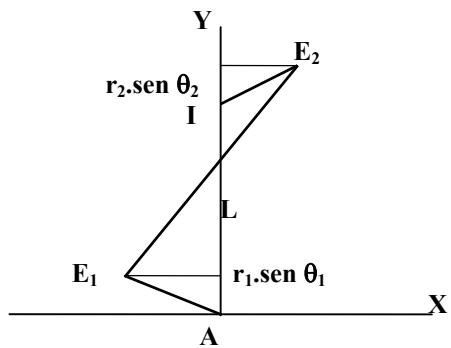


Figura 4.1 – Erros nas distâncias

(II) Para as direções

Considere as figuras 4.2-a onde supõe-se não haver erro na pontaria e 4.2-b onde supõe-se a inexistência de erro de centragem do instrumento. A variância total

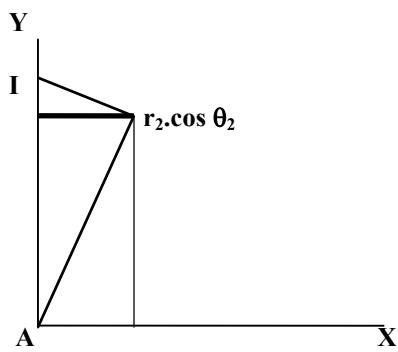


Figura 4.2-a – Sem erro de pontaria

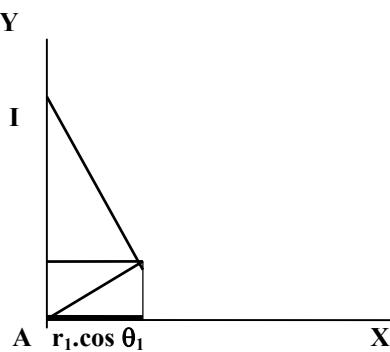


Figura 4.2-b – Sem erro de centragem

da direção será dada por

$$V(\text{Total}) = V(D) \pm V(r_1 \cos \theta_1) \pm V(r_2 \cos \theta_2)$$

$$= \sigma_D^2 + \left(\frac{\sigma_I \rho''}{L} \right)^2 + \left(\frac{\sigma_A \rho''}{L} \right)^2$$

ou, então

$$V(\text{Total}) = \sigma_D^2 + \frac{2(\sigma\rho')^2}{L^2}. \quad (4.3)$$

5. ESTIMAÇÃO DE σ^2 USANDO MEDIDAS REPETIDAS DE r

Para o desenvolvimento dos diversos métodos que serão considerados, suponha que:

- (a) seja possível obter r experimentalmente;
- (b) cada r_i tem distribuição de Rayleigh, sendo essas distribuições idênticas e os r_i independentemente distribuídos;
- (c) sejam efetuadas n medidas.

(I) Estimador de máxima verossimilhança de σ^2

A função de verossimilhança é dada por

$$L(\sigma^2) = \frac{\prod_{i=1}^n r_i}{\sigma^{2n}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n r_i^2\right)$$

ou

$$\ln L(\sigma^2) = \sum_{i=1}^n \ln r_i - n \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n r_i^2. \quad (5.1)$$

Assim, a derivada de (5.1), $\frac{\partial \ln L(\sigma^2)}{\partial \sigma^2} = 0$, resulta no EMV de σ^2 , ou seja,

$$\hat{\sigma}^2 = \frac{1}{2n} \sum_{i=1}^n r_i^2. \quad (5.2)$$

Pode-se verificar que $\hat{\sigma}^2$ como estimador do parâmetro σ^2 é um estimador não-tendencioso, consistente, de variância mínima e suficiente. Portanto possui todas as qualidades desejáveis de um estimador.

(II) Suponha que sejam medidas as componentes x_i e y_i ao invés de r_i

A função de verossimilhança será dada por

$$L(\sigma^2) = \frac{1}{(2\pi\sigma^2)^n} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i^2 + y_i^2)\right]$$

Bol. Ciênc. Geod., Curitiba, v. 5, p.15-26, 1999.

ou

$$\ln L(\sigma^2) = -n \cdot \ln 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i^2 + y_i^2), \quad (5.3)$$

resultando no EMV

$$\hat{\sigma}^2 = \frac{1}{2n} \sum_{i=1}^n (x_i^2 + y_i^2) \quad (5.4)$$

que é idêntico ao obtido para r_i , apenas exigindo duas vezes mais observações.

(III) Método dos momentos

O primeiro momento centrado na origem corresponde a

$$\mu = E(r) = \sigma \sqrt{\frac{\pi}{2}}, \quad (5.5)$$

enquanto que

$$\bar{r} = \frac{\sum_{i=1}^n r_i}{n}. \quad (5.6)$$

Portanto, de acordo com o método dos momentos tem-se que

$$\left(* \hat{\sigma} \sqrt{\frac{\pi}{2}} \right)^2 = \left(\frac{\sum_{i=1}^n r_i}{n} \right)^2,$$

resultando no estimador

$$* \hat{\sigma}^2 = \frac{2}{\pi n^2} \left(\sum_{i=1}^n r_i \right)^2. \quad (5.7)$$

O estimador pelo método dos momentos é tendencioso, pois

$$E(* \hat{\sigma}^2) = \sigma^2 + \frac{\sigma^2}{n} \left(\frac{4}{\pi} - 1 \right), \quad (5.8)$$

sendo porém consistente.

(IV) Estimador de mínimos quadrados

O estimador não-viciado da variância de r é dado por

$$\hat{\sigma}_r^2 = \frac{1}{n-1} \sum_{i=1}^n (r_i - \bar{r})^2, \quad (5.9)$$

onde

$$\bar{r} = \frac{1}{n} \sum_{i=1}^n r_i. \quad (5.10)$$

Como foi visto $V(r)$ em (2.10), tem-se então que minimizando a função $Q = [\hat{\sigma}_r^2 - V(r)]^2$ resulta no estimador de mínimos quadrados

$$\bar{\sigma}_r^2 = \frac{1}{n-1} \left(\frac{2}{4-\pi} \right) \sum_{i=1}^n (r_i - \bar{r})^2. \quad (5.11)$$

Pode-se mostrar que esse estimador é não-tendencioso, pois $E(\hat{\sigma}_r^2) = \sigma^2$.

Quadro 5.1 – Resumo dos estimadores desenvolvidos

Tipo do Estimador	Observações	Estimador
(I) E.M.V.	r_1, r_2, \dots, r_n	$\hat{\sigma}^2 = \frac{1}{2n} \sum_{i=1}^n r_i^2$
(II) E.M.V.	x_1, x_2, \dots, x_n y_1, y_2, \dots, y_n	$\hat{\sigma}^2 = \frac{1}{2n} \sum_{i=1}^n (x_i^2 + y_i^2)$
(III) Momentos	r_1, r_2, \dots, r_n	$*\hat{\sigma}^2 = \frac{2}{\pi n^2} \left(\sum_{i=1}^n r_i \right)^2$
(IV) Mínimos Quadrados	r_1, r_2, \dots, r_n	$\bar{\sigma}^2 = \frac{1}{n-1} \left(\frac{2}{4-\pi} \right) \sum_{i=1}^n (r_i - \bar{r})^2$

6. SIMULAÇÕES

Considere uma amostra aleatória de $n=100$ pontos gerados a partir de uma distribuição de Rayleigh, com desvio padrão para a distribuição normal bidimensional igual a $\sigma = 1$. A figura 6.1 mostra a distribuição desses pontos, com ponto de centragem na origem.

A distribuição de Rayleigh nesse caso tem como média

$$E(r) = \sigma \sqrt{\frac{\pi}{2}} = 1 \sqrt{\frac{\pi}{2}} \cong 1,25 \quad (6.1)$$

e variância

$$V(r) = 0,43\sigma^2 = 0,43 \cdot 1 = 0,43.$$

Para a amostra gerada o valor estimado de $E(r)$ é de aproximadamente 1,27, enquanto que os valores estimados de $V(r)$, de acordo com o método de estimação utilizado, são:

$$\hat{\sigma}^2 = 1,0122 \text{ (E.M.V.)}, \quad * \hat{\sigma}^2 = 1,0239 \text{ (Momentos)} \quad \text{e} \quad \bar{\sigma}^2 = 0,979 \text{ (M.M.Q.)}.$$

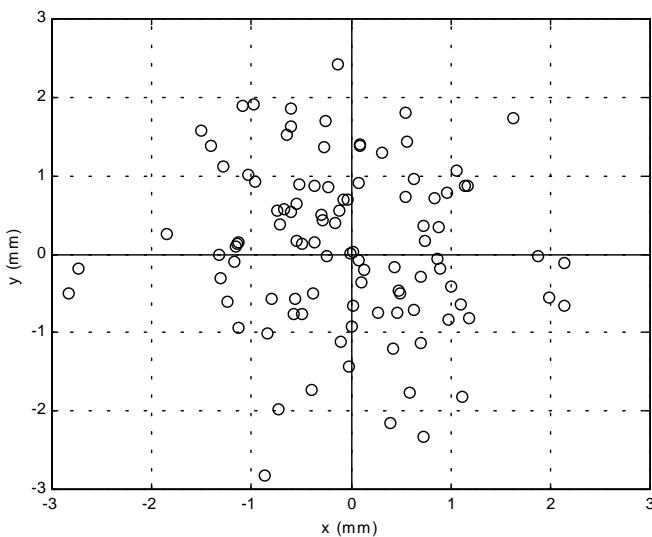


Figura 6.1 – Dispersão de $n = 100$ pontos – Distribuição de Rayleigh

Para a simulação realizada, considerando apenas o erro de centragem, e uma distância verdadeira de 20 m, obtiveram-se os resultados mostrados no quadro 6.1

Quadro 6.1 – Resultados da simulação para $n = 100$

Distância (metros)	Erro (milímetros)
Máxima = 20,00287276	2,877276
Mínima = 19,99743409	2,566591
Média = 19,99997366	0,026336

Para uma amostra de 10 pontos gerados nas mesmas condições da simulação anterior, resultou no diagrama de dispersão dado pela figura 6.2.

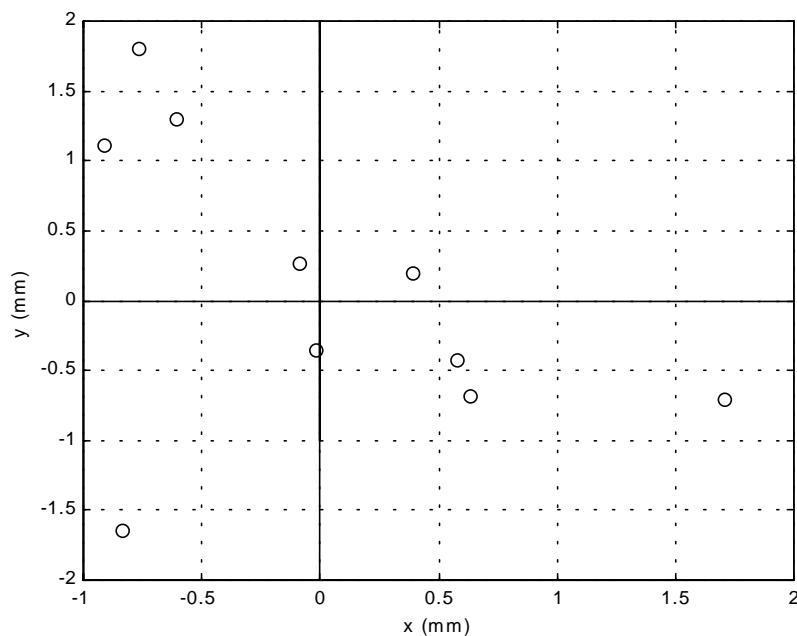


Figura 6.2 - Dispersão de $n = 10$ pontos – Distribuição de Rayleigh

O quadro 6.2 resume os resultados para a amostra de $n=10$ pontos, considerando-se apenas o erro de centragem.

Quadro 6.2 - Resultados da simulação para $n = 10$

Distância (metros)	Erro (milímetros)
Máxima = 20,00144621	1,44621
Mínima = 19,99875402	1,24598
Média = 20,00030940	0,30940

7. CONCLUSÕES E SUGESTÕES

Analizando os resultados das simulações realizadas, verifica-se que os erros cometidos na centragem de instrumentos poderão ser minimizados com o aumento da amostra de medidas, sendo essas amostras obtidas com a repetição das centragens.

Nos casos analisados, em 100 centragens o erro máximo foi de 2,87276 mm, com erro médio de 0,02634 mm, enquanto que no caso de 10 centragens o erro máximo foi de 1,44621 mm com um erro médio de 0,30940 mm. Se considerarmos também o erro de pontaria, esses erros poderiam ser o dobro dos resultados obtidos.

Como o número de centragem influencia na obtenção dos erros, recomenda-se a repetição das medidas com um número de centragem maior, visando à minimização dos mesmos.

Outra sugestão seria a realização de um estudo que permitisse dimensionar o tamanho da amostra mínima para a obtenção das medidas dentro de uma precisão e nível de confiança definidos pelo interessado.

Bibliografia

- BICKEL, P.J.; DOCKSUN, K. A. **Mathematical statistics**. San Francisco: Holden-Day, 1977.
- PATTERSON, R. Estimation of the variance of centring errors. **The Australian Surveyor**, v.29, n. 6, p. 365-374, June 1979.
- WHITE, L. A. Distribution of centring. **The Australian Surveyor**, v. 28, n. 8, December 1977.

(Recebido em 13/05/99. Aceito para publicação em 14/07/99.)