Determining the Confidence Interval for the Center and Width of a Structure in Fitting Measured Data by the Regression Line

[1] G. Wimmer, [2]K. Karovič

[1]IMS and MI SAS, Bratislava, FNS UMB, Banská Bystrica, Slovakia
and FNS MU Brno, Czech Republic,
[2]Institute of Measurement Science SAS, Bratislava, Slovakia
Email: wimmer@mat.savba.sk

Abstract. *The two dimensional cross section of interest of a structure (e.g. a grating line) is schematically seen in Fig.2. Solid lines are structure's edges. The solid bold arrow at the horizontal axis is the structure's width w for given value F. Here the proper confidence intervals for width and center of the structure are derived.*

Keywords: Confidence Interval, Measurement Uncertainty

1. Introduction

In metrology in case of assumed linear dependence between two quantities x, y (e.g. length and electric signal, respectively) the functional dependence $y = f(x)$ is fitted using the regression line $y=a+bx$. It is assumed that measurement of the quantity y in (exact) point x is normally distributed, and the measurements are independent with equal standard deviations. The estimators of the regression coefficients $\hat{a}$, $\hat{b}$ and their standard deviations $s_{\hat{a}}, s_{\hat{b}}$ are determined from pairs of measured values $\{x_i, y_i\}_{i=1}^{n}$ using standard procedures (see e.g. [3]). A common estimator of the x value for assigned level F of the quantity y (e.g. electric signal) is

$$\hat{x} = \frac{\hat{F} - \hat{a}}{\hat{b}},$$

where $\hat{F}$ is a proper estimator of F, independent of $\hat{a}$, $\hat{b}$. Let f be the estimate of F and s_F be the estimate of the standard deviation of f. By using the Law of Propagation of Uncertainties (see [2]) we obtain the estimate of the standard deviation of the estimator $\hat{x}$ as

$$s_{\hat{x}} = \frac{1}{\hat{b}} \sqrt{\left[\frac{s^2}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2} \left(\sum_{i=1}^{n} x_i^2 - 2\frac{f - \hat{a}}{\hat{b}} \sum_{i=1}^{n} x_i + n\left(\frac{f - \hat{a}}{\hat{b}} \right)^2 \right) \right] + s_F^2},$$

where

$$s^2 = \frac{1}{n-2} \left(\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i \right)^2}{n} - \hat{b} \left(\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n} \right) \right). \qquad (1)$$

In deriving the (1-α)-confidence interval for x difficulties arise caused by the generally nonsymmetrical distribution of $\hat{x}$ and so this confidence interval cannot be determined in usual way as $\left(\hat{x} - ks_{\hat{x}}, \hat{x} + ks_{\hat{x}}\right)$.

The desired (1-α)-confidence interval for x can be determined by the below described procedure.

2. Subject and Methods

The (1-α)-confidence interval for x

It is obvious (see e.g. [3]) that for an arbitrary x the (1-α)-confidence interval for (nonrandom value) $a + bx$ (= y_x) is ($_l y_x$, $_u y_x$), where

$$_l y_x = \hat{a} + \hat{b} x - s d_x \left(t_{n-2} \left(1 - \frac{\alpha}{2} \right) \right), \qquad _u y_x = \hat{a} + \hat{b} x + s d_x \left(t_{n-2} \left(1 - \frac{\alpha}{2} \right) \right),$$

with

$$d_x = \sqrt{\frac{1}{n} \left(1 + \frac{\left(nx - \sum_{i=1}^{n} x_i \right)^2}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2} \right)},$$

and $t_{n-2}\left(1 - \frac{\alpha}{2}\right)$ is the $\left(1 - \frac{\alpha}{2}\right)$-quantile of the Student t-distribution with n-2 degrees of freedom, where s is given by (1), for illustration see Fig. 1. According to [1], pp. 509-512, in case of sufficiently steep edges of the structure and small value of s (what is assumed here), the (1-α)-confidence interval for x, dented by ($_l x$, $_u x$), can be constructed by the approach illustrated in Fig. 1. Given the errorless (nonrandom) value F, the following relations hold true for the boundaries $_l x$, $_u x$ of the (1-α)-confidence interval for x

$$F = \hat{a} + \hat{b}\, _l x - s d_x \left[t_{n-2} \left(1 - \frac{\alpha}{2} \right) \right],$$

$$F = \hat{a} + \hat{b}\, _u x + s d_x \left[t_{n-2} \left(1 - \frac{\alpha}{2} \right) \right].$$

Solving both preceding equations the bounds $_l x$ and $_u x$ are given by

$$_l x = \frac{-B}{2A} - \frac{\sqrt{B^2 - 4AC}}{2A}, \quad _u x = \frac{-B}{2A} + \frac{\sqrt{B^2 - 4AC}}{2A} \qquad (2)$$

for
$$A = \hat{b}^2 - \frac{ns^2 \left[t_{n-2}^2 \left(1 - \frac{\alpha}{2} \right) \right]}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2},$$

$$B = 2\left(\frac{s^2\left[t_{n-2}^2\left(1-\frac{\alpha}{2}\right)\right]\sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} - \hat{b}\left(F - \hat{a}\right) \right),$$

$$C = \left(F - \hat{a}\right)^2 - \frac{s^2}{n}\left[t_{n-2}^2\left(1-\frac{\alpha}{2}\right)\right]\cdot\left(1 + \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \right).$$

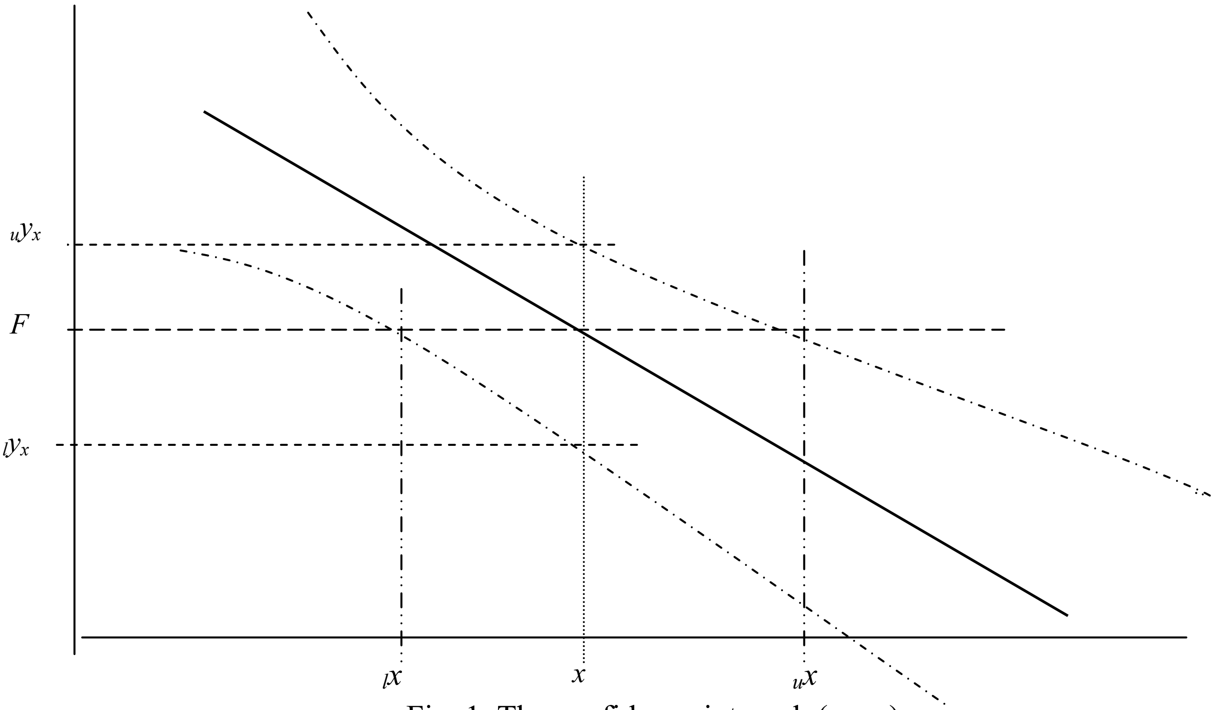

Fig. 1. The confidence interval $({}_l x,\ {}_u x)$

Determining some parameters of the structure

In order to determine the structure's center and width, for a given (errorless, nonrandom) F in analyzed cross section, it is necessary to use the x values from two confidence intervals $\left({}^{(1)}_{l}x, {}^{(1)}_{u}x\right)$ and $\left({}^{(2)}_{l}x, {}^{(2)}_{u}x\right)$ that correspond to the structure's edges (borders), see Fig. 2.

Using Bonferroni's inequality (see e.g. in [3]) for the structure's center x_s

$$P\left\{ x_S \in \left(\frac{{}^{(1)}_{l}x + {}^{(2)}_{l}x}{2}, \frac{{}^{(1)}_{u}x + {}^{(2)}_{u}x}{2} \right) \right\} \geq 1 - 2\alpha ,$$

i.e.

$$\left(x_{S_1} = \frac{{}^{(1)}_{l}x + {}^{(2)}_{l}x}{2}, \quad x_{S_2} = \frac{{}^{(1)}_{u}x + {}^{(2)}_{u}x}{2} \right) \qquad (3)$$

is at least (1-2α)-confidence interval for the structure's center.

Similarly, for the structure's width w ,

$$P\left\{w \in \left({}_{l}^{(2)}x - {}_{u}^{(1)}x,\ {}_{u}^{(2)}x - {}_{l}^{(1)}x\right)\right\} \geq 1 - 2\alpha \ ,$$

and from this,

$$\left(w_1 = {}_{l}^{(2)}x - {}_{u}^{(1)}x,\ w_2 = {}_{u}^{(2)}x - {}_{l}^{(1)}x\right) \tag{4}$$

is the at least (1-2α)-confidence interval for the structure's width w.

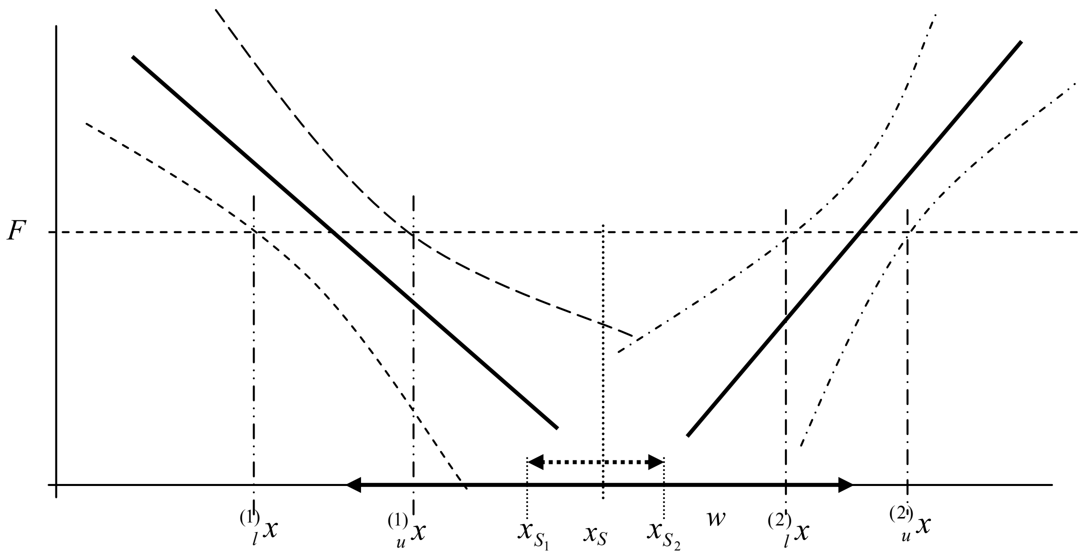

Fig. 2. Determining the sizes of the structure

3. Conclusion

The above achieved assertions are applicable to measurement of the geometry of two-dimensional structures (or cross sections of three-dimensional structures) in the following manner:

For a chosen (errorless) value (level) F of quantity y (e.g. electric signal of a length gauge) and $\alpha \in (0,1)$, the (1-α)-confidence interval for the structure's border (bound) is (${}_{l}x$, ${}_{u}x$) where ${}_{l}x$, ${}_{u}x$ are given in (2), the (1-2α)-confidence interval for the structure's center is given in (3), and the (1-2α)-confidence interval for the structure's width is given in (4).

Acknowledgment

Contribution was supported by Alexander von Humboldt foundation and grants APVV RPEU-0008-06, VEGA 1/0077/09, VEGA 2/7082/27 and MŠMT ČR LC06024.

References

[1] Cramér H. Mathematical Methods of Statistics. Princeton University Press, Princeton, 1946.

[2] Guide to the Expression of Uncertainty of Measurement (GUM) (1995), ISO, ISBN 91-67-10188-9, Geneve, Switzerland, 101 pages.

[3] Mood R.M., Graybill F.A., Boes D. Introduction to the Theory of Statistics. (3rd edn.). International Student Edition McGraw-Hill, Auckland, 1974.