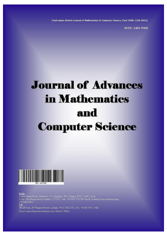

Journal of Advances in Mathematics and Computer Science

34(6): 1-12, 2019; Article no.JAMCS.53783
ISSN: 2456-9968
(Past name: British Journal of Mathematics & Computer Science, Past ISSN: 2231-0851)

Adopting Tolerance Regions in Environmental Economics

Christos P. Kitsos[1] **and Thomas L. Toulias**[2*]

[1]*Department of Informatics and Computer Engineering, University of West Attica, Athens, Greece.*
[2]*Department of Electrical and Electronics Engineering, University of West Attica, Athens, Greece.*

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JAMCS/2019/v34i630227
Editor(s):
(1) Dr. Metin Basarir, Professor, Department of Mathematics, Sakarya University, Turkey.
Reviewers:
(1) A. N. Chavan, Shivaji University, India.
(2) João Everthon da Silva Ribeiro, Universidade Federal da Paraíba, Brazil.
(3) Anton D. Murzin, Southern Federal University, Russian Federation.
Complete Peer review History: http://www.sdiarticle4.com/review-history/53783

Received: 28 October 2019
Accepted: 02 January 2020
Published: 11 January 2020

Original Research Article

Abstract

Uncertainty often lies when there is limited knowledge about the process one has to follow regarding the investigation of a real-world problem. In practice, uncertainty is related with the assumed estimation model of the physical problem, and mainly concerns the involved parameters. A typical example can be an Environmental Economics system. There are many model specifications that estimate the so-called Benefit Area of such system. For the evaluation of the optimal level of pollution, we can adopt the corresponding tolerance region, and hence we can refer to this optimal level via future observations rather than some parameters estimation. Tolerance regions can be either classical or expected tolerance regions. The associated (four) Benefit Areas can be evaluated through a proposed tolerance region procedure, and not through the usual confidence interval/region approach. Therefore, four possible optimal levels of pollution can be obtained, as well as the corresponding tolerance region for the reduction pollution point.

Keywords: Confidence interval/region; tolerance region; environmental economics; general linear model.

**Corresponding author: E-mail: t.toulias@teiath.gr, th.toulias@uniwa.gr, th.toulias@gmail.com;*

2010 Mathematics Subject Classification: 62F25, 62P12, 91B76.

1 Introduction

Uncertainty is a key element for the description of physical problems under investigation, and the easiest way to measure it is, through an information-theoretic approach, by the adoption and study of certain measures of information; see [1], [2] and [3] among others. A typical example of physical problems is the study of the Environment and, in particular, the Environmental Economics.

There are a number of model specifications that estimate, eventually, the Benefit Area, which is the area covered by the intersection between the marginal abatement function (MAD) and the marginal damage cost function (MD), restricted by the Cost or Benefits axis [4], [5]. As an example, research provides evidence that the Relative Risk (RR) differentiates under the gender factor in [6], [7]. In this survey, women seem to be more vulnerable to environmental disasters and climate change than men, mainly due to their social role and responsibilities. Therefore, different approaches are needed to analyze statistical parameters concerning the acquired environmental data; see [5] and [8] among others.

One question arising from the general study of these models is "what is the percentage of the future observations that lie within a predefined interval/region with a given probability?". Such a request gives in fact the definition of the so-called tolerance region (TR), as it was defined by [9]. We believe that the adoption and study of the TRs is essential in environmental problems, and in this research we shall try to implement the corresponding TRs in an Environmental Economics context.

Uncertainty is hidden in Environmental Economics, either in the choice of the model or in other factors, and have already been discussed in [4], [8], [10] and [11]. Initially, the aim is to investigate and to produce relationships among the real-world events that we study so that the involved variance and Relative Risk (RR) to be expressed and analyzed; see for example [12] among others. Environmental Economics is an important field that adopts such relationships and provides food for thought regarding Health and Economics.

In the theory of probability, the Borel Algebra on the set of real numbers, i.e. the algebra on which the Borel measure is defined, plays an important role on the foundation of the probabilistic aspect of a problem under investigation; the Environmental Pollution is that problem for this paper. That is why briefly we state the definition: Let $\mathscr{C}$ be the collection of real intervals in R. The smallest σ-field containing $\mathscr{C}$ is called the *Borel σ-field*. Any interval is a Borel set, and we can extent to regions (intervals) of $\mathbb{R}$. In principle, if $\mathscr{T}$ is a topological space and $\mathscr{B}$ is the smallest σ-field that contains all the open sets, then $\mathscr{B}$ is called the Borel σ-field. Moreover the Borel algebra on real numbers is the smallest σ-algebra on $\mathbb{R}$ that contains all the intervals. The probability space, defined through Measure Theory, is used in Appendix I for a theoretical approach regarding invariance.

Most of the work devoted on this subject is related to a confidence interval approach [8] which offers a solution towards the investigation of the involved uncertainty. The investigation of the future observations as well as the level of probability in order a future observation to be considered accepted or not, has its own importance which is equal if not greater than the important of some parameter estimations. This leads us to the concept of the tolerance interval/region (TR); roughly put, the interval/region in which a certain percentage of future observations lies with a given high probability. The above general idea of TR is discussed in Section 2.

Interest is focused when the underlying model, with typical example being the General Linear Model (GLM) remains invariant to linear (affine) transformations, [13]. The δ-expected tolerance region is usually considered, which is the average TR, denoted with ATR(δ). The gain in information/knowledge is even more when we adopt the so-called expected TR which is asked to be at a certain probability level $\delta \in (0,\ 1)$; see Section 3 for a brief discussion. We comment here that, in bibliography, it is usually referred as the β-expectation tolerance region with the β notation omitted in this study (not to be confused with the β parameter vector of a GL model).

In this paper, our interest is focused on the intersection point $I(x_0, k_0)$ between the marginal abatement function (MAD) and the marginal damage cost (DC) function, known as the optimal level of pollution see [14]. The corresponding point x_0, in the Damage Reduction axis, is known as the optimal level of pollution reduction while the value k_0 on the Cost axis is known as optimal cost. The area covered from those curves (see Fig. 1) is known as the Benefit Area (BA). Regarding the optimal level of pollution, we can evaluate the corresponding tolerance region, either the classical or the expected (invariant) tolerance interval ATR(δ), and therefore we can obtain (from the intersection of the latter) four possible optimal levels of pollution and the corresponding tolerance interval for the reduction pollution point, as [5, 8] discussed for the confidence interval approach. The associated four Benefit Areas can be evaluated via to the adopted TR procedure, rather than a confidence interval approach.

But to what "amount" of pollution we are referring? And and what are the pollutants' future behavior? It is known that the atmosphere influences the climate and, therefore, the knowledge of the pollution is crucial, while the restrictions on the factors polluted the environment are also important. Typical examples are the CO_2 and CH_4 factors, while CFC's (not existed before 1938) with construction similar to CH_4, have a larger duration and destroy the ozon (O_3) layer, as it is known since 1985. Some researchers discussed different policies and taxation on SO_X, NO_X, CO_2 etc., [14] and [15] while others tackled the corresponding uncertainty problem, [11]. Moreover, the authors in [5] provided an extensive discussion on considering uncertainty, either through a mathematical or statistical point of view, by working on the theoretical identification of the Optimal Pollution Level, which was considered in [4] and extended in [10], with the adoption of different models describing the marginal abatement cost (MAC) and the marginal damage cost (MD), [14].

2 Tolerance Regions

Uncertainties in the functions of benefits and costs influence the policy design in a number of ways. In principle, are heavily depending on the real problem we face; see [16] and [17]. The Environmental investigation, even of the "local case", is depending on the international environment [18].

Let $\Omega = \mathbb{R}^n$ and $\mathscr{A} \subseteq \mathscr{B}$ with $\mathscr{B}$ being the Borel field $\mathscr{B} = \{[a,\ b) \in \mathbb{R}^2\}$. Consider the set function (Appendix I)

$$Q:\ \Omega \to \mathscr{A},\quad \mathbb{R}^n \ni \mathbf{y} = (y_1, y_2, \ldots, y_n) \overset{Q}{\longmapsto} Q(y_1, y_2, \ldots, y_n) \in \mathscr{A}. \tag{2.1}$$

We are restricted to statistical tolerance regions (TR) since $\Omega = \mathbb{R}^n$ and $\mathscr{A}$ is in $\mathscr{B}$. Hence, there exist two functions, say $L = L(\mathbf{y})$ and $U = U(\mathbf{y})$, such that

$$Q = [L,\ U),\quad L < U. \tag{2.2}$$

Wilks worked in [19] on a sample following a continuous distribution function (cdf) F, proved that (2.2), with

$$L = L(\mathbf{y}) = Y_{(k_1)},\quad U = U(\mathbf{y}) = Y_{(k_1+k_2)}, \tag{2.3}$$

defines indeed a tolerance region, where Y_i is the i-th ordered statistic. He also proved that

$$F(U) - F(L) = F\left(Y_{(k_1+k_2)}\right) - F\left(Y_{(k_1)}\right) \sim \text{Beta}(k_2, n - k_2 + 1). \tag{2.4}$$

The probability content of the tolerance region,denoted by $Q = Q(\mathbf{y})$, is based on independent observations with Pr_Y, and is called the *coverage* C of the TR Q, i.e. $C(Q) := \text{Pr}_Y\left(Q(\mathbf{y})\right)$, or in a function form,

$$C:\ \mathscr{B} \to [0,\ 1],\quad Q(\mathbf{y}) \overset{C}{\longmapsto} C\big(Q(\mathbf{y})\big) = \text{Pr}_Y\big(Q(\mathbf{y})\big). \tag{2.5}$$

The statistical tolerance region is a δ-content tolerance region, CTR(δ), with probability γ if

$$\Pr\left\{\Pr_Y\big(Q(\mathbf{y})\geq\delta\big)\right\}=\gamma\in[0,\,1]. \tag{2.6}$$

Recall the TR as considered above by [19]. Assuming $k_1=r$, $k_2=n-2r+1$ and $r<(n+1)/2$, it can then be proved, [20], that

$$\gamma=\Pr\left\{F\big(Y_{(n-r+1)}\big)-F\big(Y_{(r)}\big)\geq\delta\right\}=1-\mathrm{IBeta}_{\delta}(n-2r+1,2r), \tag{2.7}$$

where $\mathrm{IBeta}_{\delta}(p,q)$ denotes the incomplete beta distribution. We imposed one criterion to assure that, on average, the coverage would be δ and thus, the δ-expected TR, $\delta-$eTR, is then defined as,

$$\left[\Pr_Y\big(Q(y_1,y_2,\ldots,y_n)\big)\right]=\delta, \tag{2.8}$$

see the pioneering work of [21] and [22] among others. Therefore, we create a region, a two-sided tolerance interval. Notice that the TR is not unique; see the integral equation (2.9). It is then clear that TR's can be proved very practical in industry and not only [23].

Now, let us consider a future response $\mathbf{z}=(z_1,z_2,\ldots,z_n)$ and its corresponding tolerance region $Q(\mathbf{z})$. Then, the affine tolerance region ATR is a statistic $Q(\cdot)$ on $\mathbb{R}^n$ over the space of future responses, based on the data such that

$$C\big(Q(\mathbf{y})\big):=\int_{Q(\mathbf{y})}H(\mathbf{z}|\mathbf{y})\,\mathrm{d}\mathbf{z}=\delta,\quad \delta\in[0,\,1]. \tag{2.9}$$

Moreover, we are asking the average of the probability coverage of the tolerance region to be δ for the future response, i.e.

$$_{\theta}\left[C\big(Q(\mathbf{y})\big)\right]=\delta,\quad \theta\in\Theta\subseteq\mathbb{R}^n, \tag{2.10}$$

and we are referring to it as ATR(δ). The density function $H(\mathbf{z}|\mathbf{y})$ in (2.9) is statistically well defined (see Appendix II and called the *prediction distribution* of the future response $\mathbf{z}$; see [24] and [22].

We emphasize here that in practice the “invariant”property, i.e. to remain invariant under affine transformations, might not be applicable in all cases; see Appendix I for the main points of invariance. Suppose, for example, that there is a source of pollution in a place M and suppose we transfer it in a distance d and rotate it by an angle ϑ, in a new position M'. Although, theoretically, the TR of M is equivalent to M' and is invariant under the affine transformation, the profile of the pollution (and hence the environmental analysis) might be completely different if M' is located near a river or a city, etc. This is an example where some mathematical concepts may not be always useful in practice.

The well-known 95% confidence interval $\hat{\mu}\pm1.96\hat{\sigma}$, with $\hat{\mu}$ being the sample mean and $\hat{\sigma}$ the sample standard deviation may not necessarily include the 95% of the population, as it depends on the variance of the estimates $\hat{\mu}$ and $\hat{\sigma}$. Therefore, a tolerance region is bounding this variance by requesting a certain percentage of the population (and not the parameters) to be included in tolerance interval.

The δ-expected TR for observations coming from the normal distribution can be evaluated due to the following Theorem 2.1. For the GLM, Theorem 2.2 is the appropriate one, while for the invariant case, Theorem 2.4 provides the corresponding ATR(δ). That is, we provide the tolerance regions Q_0 as in (2.11) for the classical TR, Q_1 as in (2.16) for the simple linear model, and Q_1^* as in (2.22) for the invariant case of the simple linear model.

In practice there are usually less than 33 observations with the t-distribution adopted at $(1-\delta)/2$ level of significance, where δ being the δ-content of the tolerance region.

Theorem 2.1. *Let us assume the normality of the error variable ε coming from the standard normal distribution $\mathcal{N}(0,1)$. Then, the $100(1-\delta)\%$ TR for a sample from the normal distribution, of the form*

$$Q_0 = \left[\bar{y} - k\,s,\ \bar{y} + k\,s\right). \tag{2.11}$$

is the δ-expected region, with

$$\bar{y} = \tfrac{1}{n}\sum y_i, \quad s^2 = \tfrac{1}{n-1}\sum\left(y_i - \bar{y}\right)^2, \quad k = \sqrt{1+n^{-1}}\; t_{n-1;(1-\delta)/2}, \tag{2.12}$$

where $t_{n-1;(1-\delta)/2}$ being, as usual, the t-distribution with $n-1$ degrees of freedom, exceeding with probability $(1-\delta)/2$.

In principle, two are the main families of models: Quantitative and Qualitative models. Typical examples for the former in Statistics are the General Linear Model (GLM) and the Regression Model, while for the latter is the Design Model. We shall focus on the GLM. Consider the matrix equation

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.13}$$

where $\mathbf{y} \in \mathbb{R}^{n\times 1}$ is an observable random vector, $\mathbf{X} \in \mathbb{R}^{n\times p}$ is a matrix of fixed observable non-random variables, $\boldsymbol{\beta} \in \mathbb{R}^{p\times 1}$ is a vector of unobservable parameters defined in a parameter space Θ and $\boldsymbol{\varepsilon} \in \mathbb{R}^{n\times 1}$ is an unobservable random vector with mean $[\boldsymbol{\varepsilon}] = 0$ and covariance matrix $\mathrm{Cov}(\boldsymbol{\varepsilon}) = \Sigma$. The only difference with the Regression Model is that the input variables $\mathbf{x}$ forming the matrix $\mathbf{X}$ are random. The normality assumption for the errors is imposed when inference is asked, and the well-known OLS (Ordinary Least Squares) procedure is performed.

For the General Linear Model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, the following holds [13, Th. 8.3.1-2]). In the following, Theorem 2.2 provides the classical TR for the GLM.

Theorem 2.2. *For the GLM as in (2.13), the interval $Q_p = [L,\,U)$ with*

$$L = \hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}_0 - k_\delta\hat{\sigma} \text{ and } U = \hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}_0 + k'_\delta\hat{\sigma}, \tag{2.14}$$

is a δ-tolerance interval at the point $\mathbf{x}_0^{\mathrm{T}} = \left(x_0^1, x_0^2, \ldots, y_0^p\right)^{\mathrm{T}}$ with confidence coefficient $1-\gamma$, i.e. contains $100(1-\delta)\%$ of the observations with confidence $1-\gamma$, and

$$k_\delta = -K\,t_{1-\gamma;n-p;q}, \quad k'_\delta = K\,t_{1-\gamma;n-p;-q}, \quad K^2 = \mathbf{x}_0^{\mathrm{T}}\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{x}_0, \tag{2.15}$$

$k_\delta, k'_\delta \in \mathbb{R}$, where $t_{1-\gamma;n-p;q}$ denotes the upper $1-\gamma$ probability point of the non-central t-distribution with $n-p$ degrees of freedom (df) and non-centrality parameter $q = -Z_\delta/K$.

Notice that $-t_{1-\gamma;n-p;q} = t_{\gamma;n-p;-q}$. We also emphasize here that the evaluation of non-central t-distribution needs special care.

Corollary 2.3. *For the simple linear regression model $y = \beta_0 + \beta_1 x + \varepsilon = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x} + \varepsilon$, $\varepsilon \sim \mathcal{N}\left(0,\sigma^2\right)$, the δ-tolerance interval with confidence $1-\gamma$ is*

$$Q_1 = [l,\,u) = \left[\hat{\beta}_0 + \hat{\beta}_1 x_0 - k_\delta\,\hat{\sigma},\ \hat{\beta}_0 + \hat{\beta}_1 x_0 + k'_\delta\hat{\sigma}\right), \tag{2.16}$$

with $n-p = n-2$, $q = -Z_\delta/K$, k_δ, k'_δ as in (2.15) and

$$K^2 = \frac{1}{n} + \frac{(\bar{x}-x_0)^2}{\sum\limits_{i=1}^{n}(x_i-\bar{x})^2}. \tag{2.17}$$

Now, to construct a δ-expectation affine tolerance region ATR(δ) for the affine GLM, the prediction distribution is needed and the following holds; see [22] and Appendix II for a compact review and notation, especially (5.14).

Theorem 2.4. *Let the error variable ε follow the normal distribution with* 0 *mean and variance* 1*, i.e.*

$$f(\varepsilon_i)\,\mathrm{d}\varepsilon_i = \frac{1}{\sqrt{2\pi}}\exp\left\{-\tfrac{1}{2}\varepsilon_i^2\right\}\mathrm{d}\varepsilon_i, \quad i = 1,2,\dots,p. \tag{2.18}$$

We also assume that the matrix $\mathbf{X}_0$ of the GLM as in (2.13) *corresponds to the matrix of n' future responses. Then, for the central* $100(1-\delta)\%$ *normal distribution being sampled, the ellipsoidal region*

$$Q_p^* = \left\{\mathbf{y}\in\mathbb{R}^{n\times 1} \Big/ \left(\mathbf{y}-\mathbf{X}_0^{\mathrm{T}}\hat{\boldsymbol{\beta}}^{\mathrm{T}}\right)\mathbf{S}^{-1}\left(\mathbf{y}-\mathbf{X}_0^{\mathrm{T}}\hat{\boldsymbol{\beta}}^{\mathrm{T}}\right) \le \frac{p}{n-p}F_{n',n-p;\delta}\right\}, \tag{2.19}$$

is the δ-expectation ATR(δ)*, i.e.*

$$\mathrm{ATR}(\delta) = Q_p^*. \tag{2.20}$$

When $p = 1$, i.e. when we are referring to the simple linear model, the matrix $\mathbf{X}_0$ is reduced to the vector $\mathbf{x}_0$ and the following holds; see [25] for a detailed discussion.

Corollary 2.5. *For the simple linear model*

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}\left(0,\sigma^2\right), \quad i = 1,2,\dots,n, \tag{2.21}$$

the central δ-expectation invariant ATR(δ) *at point x_0 is $Q_1^* = \left[L^*,\, U^*\right)$ with*

$$L^* := \hat{\beta}_0 + \hat{\beta}_1 x_0 - \frac{k}{(S_1^{-1})^{1/2}}, \quad U^* := \hat{\beta}_0 + \hat{\beta}_1 x_0 + \frac{k}{(S_1^{-1})^{1/2}}, \tag{2.22}$$

where $\hat{\beta}_0 + \hat{\beta}_1 x_0 = (1, x_0)\cdot(\hat{\beta}_0, \hat{\beta}_1)^{\mathrm{T}} =: \mathbf{x}_0^{\mathrm{T}}\hat{\boldsymbol{\beta}}$, $S_1^{-1} := 1 - \mathbf{x}_0^{\mathrm{T}}\left(\mathbf{X}^{\mathrm{T}}\mathbf{X} + \mathbf{x}_0^{\mathrm{T}}\mathbf{x}_0\right)^{-1}\mathbf{x}_0$, and

$$k = \frac{t_{n-p;\delta/2}}{\sqrt{n-2}}(RSS)^{1/2}, \tag{2.23}$$

with $p = 2$ and RSS being the Residual Sum of Squares of the OLS model (2.21)*, while $\mathbf{X} := ((1,1,\dots,1)^{\mathrm{T}}, \mathbf{x}) \in \mathbb{R}^{n\times 2}$, $\mathbf{x} := (x_i)^{\mathrm{T}} \in \mathbb{R}^{n\times 1}$.*

Therefore, the interval $Q_1^* = \left[L^*,\, U^*\right)$ as in (2.22) is the ATR(δ). For the simple linear model, the corresponding TRs Q_1 and Q_1^*, as in (2.16) and (2.22) respectively, are the two candidates for the invariant TR-s in order to evaluate Q_1 or Q_1^*, so that a certain percentage of the future observations will lie, on average, within Q_1 or Q_1^*, with certain given probability.

3 Environmental Economics

In Environmental Economics, the marginal abatement cost (MAC) as well as the marginal damage cost (MD) play an important role as the optimal pollution level occurs at certain point for which $MAC = MD$. Although there is an uncertainty about the appropriate model choice for the approximation of MAC and MC, their typical presentation is given in Fig. 1; see [10] and [11].

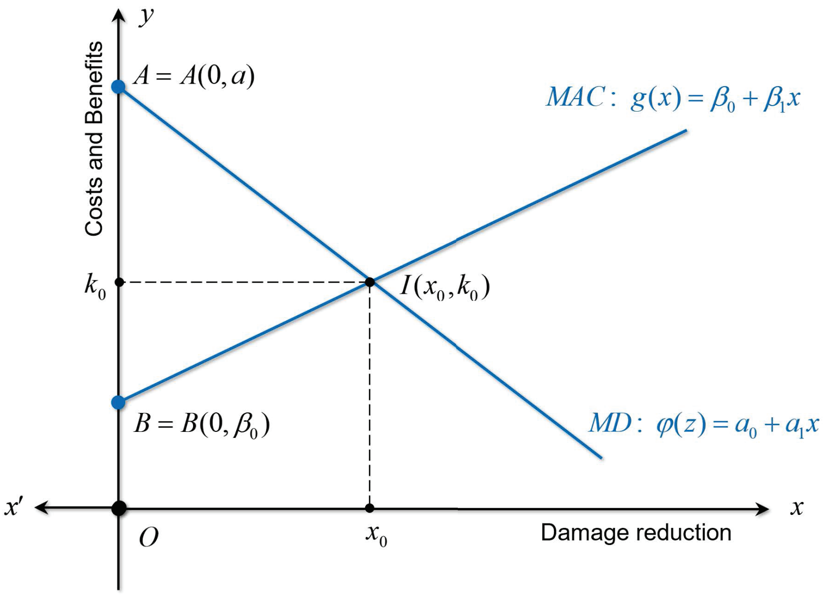

Fig. 1. Graphical presentation of the theoretical optimal level of pollution

Notice that the area of the triangle ABI corresponds to the Benefit Area (BA), where the curves $g = g(x)$ and $\varphi = \varphi(x)$ can be estimated with the OLS methodology; see [4]. For linear MAC = $g(x)$, and adopting regression technics, we can evaluate $\hat{g}(x) = \hat{\beta}_0 + \hat{\beta}_1 x$ as well as for MD = $\varphi(x)$: $\hat{\varphi}(x) = \hat{\alpha}_0 + \hat{\alpha}_1 x$. Therefore, the TR, either Q_1 or Q_1^*, can be evaluated for $g(x)$ and $\varphi(x)$ respectively, say Q_{1g} or Q_{1g}^* and $Q_{1\varphi}$ or $Q_{1\varphi}^*$. In Fig. 2 the TRs are presented together with the confidence intervals (CIs) for both MAC and MD. The benefit area BA = (AIB) has now different options, as the intersection I may vary, depending on the number of intersections of the associated TRs. In any case, however, it has to be of the form $I(x_0, k_0)$ with $x_0, k_0 > 0$. This presentation is another way to handle the BA and the existing underlying uncertainty, in Environmental Economics.

We believe that there is a real need, especially in the field of Environmental Economics, to work with the "future population" rather than the estimated measures of position or dispersion. The estimates of the mean, median, mode, and the percentiles, as well as the variance of the population, might offer information about the "center" and "scale" of the population, but does not provide information for the behavior of the future population coming from the source under study. Fig. 2 provides evidence that, as the TRs are larger than CIs, TRs might provide larger BAs and that could be a problem for the researcher who has then to decide what is the appropriate choice. We are working on this decision problem.

4 Discussion

We believe that, although a tolerance interval is less widely known than the confidence interval or the prediction interval, it is more useful in practical problems. While confidence intervals bounds the parameter estimates, such as mean, variance, proportion etc., tolerance interval bounds the range of the data that includes a specific proportion of the population.

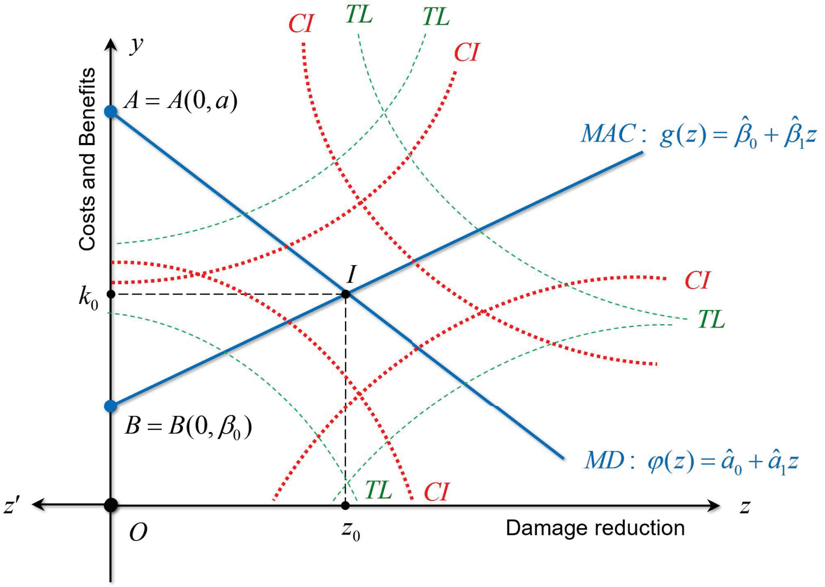

Fig. 2. Graphical presentation of the estimated level of pollution

Example 4.1. *Recall [4]. It has been evaluated , for quadratic MAC and linear MD, for the case of Greece that,*

$$\begin{aligned} MAC &= \hat{\beta}_0 + \hat{\beta}_i x + \hat{\beta}_2 x^2 = 2.29 + 0.026x + 0.0099x^2, \\ MD &= \hat{\alpha}_0 + \hat{\alpha}_1 x \qquad\quad = 3.13 + 0.0341x. \end{aligned}$$

The corresponding optimal pollution level is

$$x_0 = -\frac{\hat{\beta}_1 - \hat{\alpha}_1}{2\hat{\beta}_2} = 42.12,$$

and the Benefit Area BA = 42.6*. Based on the confidence interval (CI) for the coefficients of MAC and MD, two more curves are obtained, depending on the lower and upper values of the coefficients' CIs, say* MAC_1, MAC_2, *and* MD_1, MD_2*; see the related curves of CIs in Fig. 2. In our case*

$$\begin{aligned} MAC_1 &= 2.29 + 0.0235x + 0.0022x^2, \\ MAC_2 &= 2.33 + 0.295x + 0.0169x^2, \\ MD_1 &= 3.16 + 0.0159x, \\ MD_2 &= 4.33 + 0.841x. \end{aligned}$$

It is clear that there are different intersections $I(x_0, k_0)$*, new* x_0 *and another BA evaluation. For this particular case, a more detailed statistical analysis is needed since the CIs of* $\hat{\alpha}_1$ *includes zero. Now, for the future observations, the estimates of the values of MAC and MD are obtained; see the corresponding tolerance region (TR) curves in Fig. 2. The TRs depend on the given number of future observations,* n'*, and on a non-central t-distribution, which makes the evaluation more complicated than the one for the corresponding CIs. Moreover, the choice of the proportion* δ *for the* δ*-content, we usually chose* $\delta = 0.90$ *or* 0.95*. We shall present soon the related TR evaluations.*

5 Conclusions

Tolerance intervals are related to prediction intervals, [26], while other researchers worked to create bounds for the variance of the future response; see [22]. In practice, the future response is of great importance as it is mainly a method of "building" a model to explore the future. Roughly speaking, a prediction interval is an estimate confidence interval since future observations of some given data (analyzed in principle with the regression analysis) will fall in that interval with an assigned probability level. We emphasize that a prediction interval bounds only a single future sample. The superiority of the tolerance interval lies in the fact that it concerns the entire population while ATR(δ) works on the average of future possible samples. This is the main reason that tolerance intervals are, we believe, more appropriate in practice. Notice also that due to this simple relation,

$$1 < \sqrt{n+1} \Rightarrow \frac{1}{\sqrt{n}} < \sqrt{1+\tfrac{1}{n}} \Rightarrow \frac{s}{\sqrt{n}} < s\sqrt{1+\tfrac{1}{n}} \Rightarrow \frac{ks}{\sqrt{n}} < k\sqrt{1+\tfrac{1}{n}}, \tag{5.1}$$

with $k > 0$, the provided tolerance region Q_0 is wider than the corresponding confidence interval. This is a simple heuristic proof of the fact that TR > CI, either as length, area, or volume. Keep in mind that TRs may be wider than CIs, however they are refer directly to the population rather than the sample parameters; see also [27].

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Kitsos CP, Toulias TL. Hellinger distance between generalized normal distributions. British Journal of Mathematics and Computer Science. 2017;21(2):1-16.

[2] Toulias TL, Kitsos CP. Entropy and information extensions through the generalized normal distribution. In: Bozeman JR et al. (Eds). Stochastic and Data Analysis Methods and Applications in Statistics and Demography. ISAST. 2016;141-156.

[3] Toulias TL. Entropy and information measures for the generalized normal distribution. In: Filus L et al. (Eds). Stochastic Modeling, Data Analysis & Statistical Applications. ISAST. 2015;3-20.

[4] Halkos G, Kitsos CP. Optimal pollution level: a theoretical identification. Applied Economics. 2005;37(2):1475-1483.

[5] Halkos G, Kitsos CP. Mathematics vs. statistics in tackling environmental economics uncertainty. MRBA. 2018;85280.

[6] Pan-American Health Organization. Gender and natural disasters: women, health and development program. Fact Sheet of the Program on Women, Health and Development. Washington; 1998.

[7] Mitchell T, Tanner T, Lussier K. We know what we need: South Asian women speak out in climate change application. Action Aid International, Johannesburg, London; 2007.

[8] Halkos G, Kitsos CP. Uncertainty in environment economics: the problem of entropy and model choice. Economic Analysis and Policy. 2018;60:127-140.

[9] Wilks S. Determination of sample sizes for setting tolerance limits. Ann of Mat Stat. 1941;12:91-96.

[10] Halkos G, Kitsou D. Uncertainty in optimal pollution levels modelling and evaluating the benefit area. Journal of Environmental Planning and Management. 2018;55(4):678-700.

[11] Kitsou D. Estimating Damage and Abatement Cost Functions to Define Appropriate Environmental Policies. PhD thesis, Univ. of Thessaly; Volos, Greece; 2015.

[12] Halkos G, Kitsos CP. Relative risk and innovation activities: the case of Greece. Innovation Management, Policy & Practice. 2012;14(1):156-159.

[13] Graybill AF. Theory and Applications of the Linear Model. Duxbury Press, Massachussets; 1976.

[14] Halkos G. Economy and Environment (in Greek). Liberal Books Publ., Athens, Greece; 2013.

[15] Halkos G. Optimal abatement of sulphure missions in Europe. Environmental & Resource Economics. 1994;4(2):127-150.

[16] Newbery D. Acid rain. Economic Policy. 1990;11:288-346.

[17] Newbery D. The impact of EC environmental policy on British coal. Oxford Review of Economic Policy. 1993;9(4):66-95.

[18] Mäller KG. International environmental problems. Oxford Review of Economic Policy. 1990;11:80-108.

[19] Wilks S. Mathematical Statistics. John Wiley, New York; 1962.

[20] Kendall MG, Stuart A. The Advanced Theory of Statistics. C. Griffin Ltd, London, UK. 1968;2.

[21] Guttman I. Construction of β-content tolerance regions at confidence level γ for large samples for k-variate normal distribution. Ann Met Stat. 1970;41:376-400.

[22] Muller CH, Kitsos CP Optimal design criteria based on tolerance regions. In: di Bucchianno A et al. (Eds.) Advances in Model-Oriented Design and Analysis (MODA7). 2004;107-115.

[23] Zarikas V, Kitsos CP. Risk analysis with reference class forecasting adopting tolerance regions. In: Kitsos CP et al. (Eds.) Theory and Practice of Risk Assessment. Springer. 2015;235-247.

[24] Fraser D, Haq MS. Structural probability and prediction for the multivariate model. J Royal Stat Soc B. 1969;31:317-332.

[25] Ellerton RRW, Kitsos CP, Rinco S. Choosing the optimal order of a response polynomial-structural approach with minimax criterion. Comm Stat Theory and Methods. 1986;15(1):129-136.

[26] Schervish IM. Theory of statistics. Springer Series in Statistics, Springer; 1995.

[27] Kitsos CP, Toulias TL. Confidence and tolerance regions for the signal process. Recent Patents on Signal Processing. 2012;2(2):149-155.

[28] Halkos G. Econometrics: Theory and practice: Instructions in using Eviews, Minitab, SPSS and Excel. Gutenberg, Athens, Greece; 2011.

Appendix I (Invariance)

Let $\{(\Omega,\mathscr{A},P),\ P\in\mathscr{P}\}$ be a probability space associated with points $\mathbf{y}\in\mathbb{R}^n$. Let g be any one-to-one transformation of Ω onto itself. The collection of all sets gA with $A\in\mathscr{A}$, is a sigma-field $\mathscr{A}$, and $g\mathscr{P}$ is the probability measure on $g\mathscr{A}$ induced by g such that

$$gP(gA)=P(A),\quad A\in\mathscr{A}. \tag{5.2}$$

Any function φ on Ω generates a new function $g\varphi$ such that

$$\varphi(x)=g\varphi(gx). \tag{5.3}$$

Then $(\Omega,\mathscr{A},P)\cong(g\Omega,g\mathscr{A},gP)$, i.e. g is an isomorphism in the sense that if $\mathscr{B}\subseteq\mathscr{A}$ then $g\mathscr{B}\subseteq g\mathscr{A}$, and if φ is any $\mathscr{A}$-$\mathscr{P}$-integrable function on Ω, the $g\varphi$ is $g\mathscr{A}$-$g\mathscr{P}$-integrable on $g\Omega$ and

$$g_P(g\varphi\mid g\mathscr{A})\ =\ g_P(\varphi\mid\mathscr{B}), \tag{5.4}$$

(i.e. except on a $g\mathscr{P}$-null set). If we now let G be a group with element g such that

$$g\mathscr{A}=\mathscr{A},\quad gP\in\mathscr{P},\quad \text{with } P\in\mathscr{P}, \tag{5.5}$$

then the family $\{(\Omega,\mathscr{A},P),\ P\in\mathscr{P}\}$ is said to be invariant under G. □

The following lemma is essential for Section 3 and the analysis therein.

Lemma 5.1. [Kitsos (2011)] *The set*

$$G=\left\{g=\begin{pmatrix}\mathbf{I}_p & \mathbf{0}\\ \mathbf{h}^{\mathrm{T}} & \lambda\end{pmatrix},\quad \mathbf{I}_p\in\mathbb{R}^{p\times p},\quad \lambda>0,\quad \mathbf{h}^{\mathrm{T}}=(h_i)\in\mathbb{R}^p\right\}, \tag{5.6}$$

is a group of transformations.

Let the GLM to be of the form (2.13). Notice that the transpose is then

$$\mathbf{y}^{\mathrm{T}}=\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}+\sigma\boldsymbol{\varepsilon}^{\mathrm{T}}, \tag{5.7}$$

and therefore, it is easy to prove that it holds

$$\begin{pmatrix}\mathbf{X}\\ \mathbf{y}^{\mathrm{T}}\end{pmatrix}=\begin{pmatrix}\mathbf{I}_p & \mathbf{0}\\ \boldsymbol{\beta}^{\mathrm{T}} & \sigma\end{pmatrix}\begin{pmatrix}\mathbf{X}^{\mathrm{T}}\\ \boldsymbol{\varepsilon}^{\mathrm{T}}\end{pmatrix}, \tag{5.8}$$

with $\mathbf{0}^{\mathrm{T}}=(0,0,\ldots,0)\in\mathbb{R}^p$, $\mathbf{I}_p\in\mathbb{R}^{p\times p}$ denotes the identity matrix, i.e. $\mathbf{I}_p:=\operatorname{diag}(1,1,\ldots,1)$. If we let

$$\mathbf{Y}:=\begin{pmatrix}\mathbf{X}^{\mathrm{T}}\\ \mathbf{y}\end{pmatrix},\quad \mathbf{g}:=\begin{pmatrix}\mathbf{I}_p & \mathbf{0}\\ \boldsymbol{\beta}^{\mathrm{T}} & \sigma\end{pmatrix},\quad \mathbf{E}:=\begin{pmatrix}\mathbf{X}^{\mathrm{T}}\\ \boldsymbol{\varepsilon}^{\mathrm{T}}\end{pmatrix}, \tag{5.9}$$

then (5.8) forms an affine transformation with matrix $\mathbf{g}$ being an element g of a group of transformations, say G, and hence (5.8) is of the form $\mathbf{Y}=g\mathbf{E}$. So we can have an affine transformation for $\mathbf{E}$.

Appendix II

Let us consider the response $\mathbf{y}^{\mathrm{T}}=(y_1,y_2,\ldots,y_n)^{\mathrm{T}}$ and its corresponding tolerance region $Q(\mathbf{y})$. Then, the affine invariant tolerance region is a statistic $Q(\cdot)$ on $\mathbb{R}^{n\times 1}$, and with the space of future responses based on the data such that

$$\Pr\big(Q(y_1,y_2,\ldots,y_n)\big)=\int_{Q(.)}\Pr(\mathbf{y}\mid\boldsymbol{\theta})\,\mathrm{d}\mathbf{y}, \tag{5.10}$$

with $\boldsymbol{\theta}$ being an element or the parameter space $\Theta \subseteq \mathbb{R}^{n\times 1}$. For the δ-expectation affine invariant equivalent tolerance region, it can be proved that is of the form

$$\mathrm{ATR}(\delta) =_{\boldsymbol{\theta}} \left[\Pr_{\boldsymbol{\theta}}(Q(\mathbf{y}))\right] = \int\limits_{\Theta}\left(\,\int\limits_{Q(\cdot)} \Pr(\mathbf{y}|\boldsymbol{\theta})\,\mathrm{d}\mathbf{y}\right) h^{*}\left(\boldsymbol{\theta} \mid \mathrm{data}\right)\mathrm{d}\boldsymbol{\theta} = \delta, \tag{5.11}$$

with $h^{*}(\boldsymbol{\theta} \mid \mathrm{data})$ being the structural distribution of parameters; see [24] and [22]. Under Fubini's theorem, and denoting

$$H(\mathbf{z}|\mathbf{y}) = \int\limits_{\Theta} \Pr_{\Theta}(\mathbf{z} \mid \boldsymbol{\theta})\, h^{*}(\boldsymbol{\theta} \mid \mathbf{z})\,\mathrm{d}\boldsymbol{\theta}, \quad \boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^{n\times 1}, \tag{5.12}$$

relation (5.11) is reduced to

$$\mathrm{ATR}(\delta) := \int\limits_{Q} H(\mathbf{z} \mid \mathbf{y})\,\mathrm{d}\mathbf{y} = \delta. \tag{5.13}$$

The density function $H(\mathbf{z}|\mathbf{y})$ has been defined as the prediction distribution of the future response $\mathbf{z}$ [24].

The prediction distribution (5.12) for the affine GLM is

$$H(\mathbf{z}|\mathbf{y}) = \frac{|\mathbf{S}|^{-1/2}\,\Gamma\left(\frac{n+n'-p}{2}\right)}{\pi^{n'/2}\,\Gamma\left(\frac{n-p}{2}\right)} \left|\mathbf{I}_n + \left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)^{\mathrm{T}}\mathbf{S}^{-1}\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)\right|^{-\frac{n+n'-p}{2}} \mathrm{d}\mathbf{y}, \tag{5.14}$$

with $\hat{\boldsymbol{\beta}}$ being the usual OLS estimators for the parameter vector $\boldsymbol{\beta}$, while the variance s^2 and covariance $\mathbf{S}$ are given below

$$\mathbf{S}^{-1} = s^{-2}(\mathbf{y})\mathbf{S}_1^{-1}, \text{ with } \left|\mathbf{S}_1^{-1}\right| = \frac{|\mathbf{X}^{\mathrm{T}}|}{\left|\mathbf{X}^{\mathrm{T}} + \mathbf{X}_0^{\mathrm{T}}\mathbf{X}_0\right|}, \tag{5.15}$$

$$s^2 = s^2(\mathbf{y}) = \left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)^{\mathrm{T}}\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right), \quad \hat{\boldsymbol{\beta}} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}, \tag{5.16}$$

$$\mathbf{S}_1^{-1} = \mathbf{I}_n - \mathbf{X}_0\left(\mathbf{X}^{\mathrm{T}}\mathbf{X} + \mathbf{X}_0^{\mathrm{T}}\mathbf{X}_0\right)^{-1}\mathbf{X}_0^{\mathrm{T}}, \quad \mathbf{X}_0 \in \mathbb{R}^{n'\times p}, \tag{5.17}$$

for given n' future responses; see for details [25] and [22]. To fit a linear model see [28] among others.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sdiarticle4.com/review-history/53783