DE GRUYTER OPEN

Folia Oeconomica Stetinensia
DOI: 10.1515/foli-2016-0016

WYDZIAŁ NAUK EKONOMICZNYCH I ZARZĄDZANIA
UNIWERSYTET SZCZECIŃSKI

CHARACTERISTICS AND PROPERTIES OF A SIMPLE LINEAR REGRESSION MODEL

Robert Kowal, Ph.D., Eng.

Kielce University of Technology
Faculty of Management and Computer Modelling
Al. Tysiąclecia Państwa Polskiego 7, 25-314 Kielce, Poland
e-mail: robertkk7@gmail.com

Received 21 August 2015, Accepted 27 June 2016

Abstract

A simple linear regression model is one of the pillars of classic econometrics. Despite the passage of time, it continues to raise interest both from the theoretical side as well as from the application side. One of the many fundamental questions in the model concerns determining derivative characteristics and studying the properties existing in their scope, referring to the first of these aspects. The literature of the subject provides several classic solutions in that regard. In the paper, a completely new design is proposed, based on the direct application of variance and its properties, resulting from the non-correlation of certain estimators with the mean, within the scope of which some fundamental dependencies of the model characteristics are obtained in a much more compact manner. The apparatus allows for a simple and uniform demonstration of multiple dependencies and fundamental properties in the model, and it does it in an intuitive manner. The results were obtained in a classic, traditional area, where everything, as it might seem, has already been thoroughly studied and discovered.

Keywords: simple linear regression model, OLS estimators, variance, unbiasedness

JEL classification: C01

Introduction

The surrounding economic reality is largely shaped by both economic and financial phenomena. For majority of them, identification, description or interpretation is a complex issue, frequently causing a multitude of problems for researches. The reason for that complexity is a large number of interacting factors, often of various nature and character, multiplicity of layers and directions of their influence, as well as other types of conditions or circumstances. That is why, the models and methods used for their description and analysis ought to be, on the one hand, as simple as possible in numerical and interpretational terms, and, on the other hand, provide possibly the most faithful reflection of reality in that regard. Therefore, unsurprisingly, linear models featuring the above-listed characteristics in various shapes constitute a fundamental tool for describing such types of phenomena and processes. Owing to their nature and character, a quantitative description of economic or financial phenomena must comprise two aspects: a statistical description and optimization. The underlining model for the description of such types of phenomena is a simple linear regression model. Despite the fact that it only links two factors, its role in various kinds of econometric studies is still relatively substantial. It is used, inter alia, in a sensitivity analysis, in accordance with the principle of *ceteris paribus*, an impact analysis, and the force of influence that variables exert on one another, or in direct or indirect form it serves as a basic component of many different special-purpose models. Thus, recognizing its properties and nature in the widest and most varied aspects is an important issue, both from the theoretical and practical points of view. One of them involves evaluating certain characteristics and examining their properties, within the scope of which a new approach is going to be presented to an issue that is crucial in this model. It will feature simplicity and originality of solutions, the use of only fundamental properties, substantial synthesis, and dissimilarity from the formulas found in literature.

1. Simple linear regression model

The model of simple linear regression is defined as follows:

$$y = \alpha + \beta x + \varepsilon,$$

where:

y – dependent variable,
x – independent variable,
α and β – structural parameters,
ε – random component.

In the design of n-observations performed on y and x it is written as follows:

$$y_i = \alpha + \beta x_i + \varepsilon_i, \qquad i = 1, 2, ..., n$$

From a formal point of view, we have an additive composite of a random component with a linear simple dependency of variables. Such a design reflects a direct impact or influence of one factor and the aggregation of all the other factors. It does not always entail a direct economic reinforcement, nevertheless, it is highly convenient in numerical terms. The random component gives the model a stochastic character (with suitable assumptions) and accumulates the influence of the variables omitted in the model. An adequate description of reality requires that suitable assumptions be taken in that regard. Typically, the so-called classic initial assumptions are made. The assumptions are as follows:

1. Invariance of an expected value, i.e. $E(\varepsilon_i) = 0$ for all i = 1, 2, ..., n.
2. Variance stability $D^2(\varepsilon_i) = \sigma^2$ for all i = 1, 2, ..., n.
3. Independence between one another, i.e. ε_i and ε_j are independent for all i and j.
4. Independence from x_i: ε_i and x_j are independent for all i and j (this is a natural assumption, since x_j are non-random, i.e. ε_i distribution does not depend on x_j).
5. Normality of distribution: ε_i have normal distributions for all i (a convenient assumption, though it may be weakened in the case of a practical need).

For the entire population the model equations assumes the following form:

$$E(y_i) = \alpha + \beta x_i, \quad i = 1, 2, ..., n.$$

It defines a function of regression in a population.

The ordinary least square (OLS) method is the most popular method for structural parameters estimation. The structural parameter estimators obtained in the model are expressed in the following formulas:

$$\begin{cases} \hat{\alpha} = \overline{y} - \hat{\beta}\overline{x} \\ \hat{\beta} = \dfrac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \end{cases}.$$

The so-called OLS estimators are the functions of the following variables y_i, x_i, i = 1, 2, ..., n.

2. Characteristics and parameters of a simple linear regression model

The best possible result of the structural parameters estimation of a simple linear regression model is obtaining unbiased and efficient estimators. Both of these questions are of significant importance from the point of view of subsequent practical applications of the model. The first one might not be defined or verified in the model in any complex manner, but merely intuitive one, nevertheless, it is not such, in fact. It seems that the unbiasedness still holds some 'secrets', and its inner nature, particularly in the interpretive layer, has not yet been completely identified and properly interpreted. In turn, the efficiency of structural parameter estimators is in itself a separate and independent problem. However, these questions, although interconnected in a sense, are not going to be the core subject of our focus, therefore, we are not going to elaborate on them in an extensive manner. The unbiasedness will be considered only to a limited, rather trace, and exclusively strictly context-defined scope, whereas the efficiency of structural parameter estimators will be the subject of a separate paper. We are going to focus mainly on other derivative characteristics with regard not only to $\hat{\alpha}$ and $\hat{\beta}$ OLS estimators of the structural parameters of the simple linear regression model, but also partly to the unbiased, linear ones in the design of the properties of basic and selected characteristics. Some of the obtained results and contents will be transferred to the multiple linear regression model, as such multiplication of regularities between the models occurs here.

The issue of unbiasedness of OLS structural parameter estimators of the simple linear regression model is relatively simple, therefore, in practice, in line with the above-stipulated suggestions, we are going to omit it. Within its scope, in order to signal the methodology and exemplify further solutions, we will only mention one of the themes, proof of unbiasedness of $\hat{\alpha}$ estimator derived by Maddala (2006):

$$E(\hat{\alpha}) = E(\overline{y} - \hat{\beta}\overline{x}) = E(\overline{y}) - \hat{\beta}\overline{x} = \alpha + \hat{\beta}\overline{x} - \hat{\beta}\overline{x} = \alpha.$$

The fact that it is exceptionally attractive, i.e. short and concise, is doubtless. The fact that it is also simple, can only be determined once we observe that it is underlined by a uniquely synthetic equation $\overline{y} = \overline{\hat{y}}$, applied here as well, which is neither direct nor easy. We will return once again to the unbiasedness in the same fundamental aspect as presented here at the end of this discussion in the context of its important general property. However, we are going to refer to it far more frequently in a secondary aspect, but it will be done successively, as the context will so require.

Such an ‘aggregated’ approach, as presented in the above proof, using the synthesis characteristic of the first normal equation, will be transferred and developed further onto all the significant characteristics and properties of $\hat{\alpha}$ and $\hat{\beta}$ OLS estimators of structural parameters as well as linear unbiased estimators. It will, however, require that at least one preliminary characteristic be defined as an initial value for further discussion. In principle, the main criterion for selecting such a characteristic is the ease of its determination, and so it is going to be in our case. Besides, we are not given much of a choice, since in this case the characteristic can only be $D^2(\hat{\beta})$ variance. Perhaps it is the only characteristic in this model that is normally determined in a direct and simple manner from a simplified form of $\hat{\beta}$ OLS estimator:

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})y_i}{\sum_{i=1}^{n}(x_i - \overline{x})^2},$$

by using the basic variance properties for that purpose. It is expressed in the following formula:

$$D^2(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}.$$

Nevertheless, within the scope of subsequent deliberations, this formula will also be derived in a contextually yet unknown, though simple, fashion, completely different from the methods typically adopted in standard literature.

The second of the basic characteristics, i.e. $D^2(\hat{\alpha})$ variance, is methodologically determined in the literature of the subject virtually in the same manner, however, from the numerical aspect it is slightly more complex. In this paper, it is going to be calculated differently, in a more synthetic fashion. We are going to use a certain, still implicit, secret property of the second estimator and conventionally the first normal equation, i.e.

$$\overline{y} = \hat{\alpha} + \hat{\beta}\overline{x},$$

hence, we successively get:

$$D^2(\hat{\alpha}) = D^2(\overline{y} - \hat{\beta}\overline{x}) = D^2(\overline{y}) + D^2(\hat{\beta}\overline{x}) - 2\operatorname{cov}(\overline{y}, \hat{\beta}\overline{x}) = D^2(\overline{y}) + \overline{x}^2 D^2(\hat{\beta}) =$$
$$= \frac{1}{n^2} D^2(y_1 + y_2 + \ldots + y_n) + \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}\overline{x}^2 = \frac{1}{n^2}\sum_{i=1}^{n} D^2(y_i) + \frac{\sigma^2\overline{x}^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{1}{n^2} n\sigma^2 + \frac{\sigma^2\overline{x}^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} =$$
$$= \sigma^2\left(\frac{1}{n} + \frac{\overline{x}^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}\right)$$

In fact, the above transformations can be shortened to just two equations (compare, e.g. Maddala (2006), Welfe (2009)). Furthermore, the applied methodology opens still other fairly wide-ranging generalizations and interpretations in that field. In particular, this proof, which is significant in this case, demonstrates at the same time an important general property of $\hat{\beta}$ OLS estimator, which shall be discussed further, and a complete variance structure of $\hat{\alpha}$ OLS estimator, i.e. which brings a certain contribution into this variance. Such a directly, defined variance structure of an absolute term is not present in the multiple linear regression model, because in it the estimators of the remaining structural parameters have already been correlated with one another.

Still in the same style, i.e. relatively straightforwardly, synthetically, and through the application of the equation in a slightly self-conjugate manner, the third basic characteristic of the model will be determined, namely $\operatorname{cov}\left(\hat{\alpha}, \hat{\beta}\right)$ covariance of $\hat{\alpha}$ and $\hat{\beta}$ OLS estimators of the structural parameters in the simple linear regression model. We have here:

$$D^2(\overline{y}) = D^2(\hat{\alpha} + \hat{\beta}\overline{x}) = D^2(\hat{\alpha}) + D^2(\hat{\beta}\overline{x}) + 2\operatorname{cov}(\hat{\alpha}, \hat{\beta}\overline{x}) =$$
$$= D^2(\overline{y}) + \overline{x}^2 D^2(\hat{\beta}) + \overline{x}^2 D^2(\hat{\beta}) + 2\overline{x}\operatorname{cov}(\hat{\alpha}, \hat{\beta}),$$

hence:

$$0 = 2\overline{x}^2 D^2(\hat{\beta}) + 2\overline{x}\operatorname{cov}(\hat{\alpha}, \hat{\beta})$$

and eventually:

$$\operatorname{cov}(\hat{\alpha}, \hat{\beta}) = -\overline{x} D^2(\hat{\beta}) = -\frac{\overline{x}\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}.$$

Further on, the formula will be derived with other, simpler methods, which, however, are based on slightly different premises. Yet, the above-presented derivation contains a noteworthy, indirectly obtained dependency $0 = \overline{x}D^2(\hat{\beta}) + \text{cov}(\hat{\alpha}, \hat{\beta})$, which through the first normal equation leads to the following dependence $0 = \overline{x}\,\text{cov}(\hat{\beta}, \hat{\beta}) + \text{cov}(\hat{\alpha}, \hat{\beta})$, and consequently to an important equation $\text{cov}(\hat{\alpha} + \hat{\beta}\overline{x}, \hat{\beta}) = 0$, i.e. $\text{cov}(\overline{y}, \hat{\beta}) = 0$, which we have already implicitly signalled and which we soon shall discuss in detail.

Yet another noteworthy question is related to the equation. It seems that it ought to be formulated and interpreted completely differently from the way in which it is done in the literature of the subject. The proper form that reflects its actual nature, sense, and interpretation is a more synthetic equation of $\overline{y} = \overline{\hat{y}}$, on which, inter alia, the above-cited Maddala's (2006) proof of $\hat{\alpha}$ OLS estimator's unbiasedness was based. The right side of the equation is nothing more than just a theoretical mean of $\hat{y}$. What is more, the fact does not need to be proven, as is done in literature, but merely looking at properly should suffice. The notation in literature is merely its direct consequence. In any case, the equation directly results reversely from the first normal equation:

$$\sum_{i=1}^{n} e_i = 0.$$

Bearing in mind that:

$$e_i = y_i - \hat{y}_i, \qquad i = 1, 2, ..., n,$$

we have here:

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right) = 0.$$

Hence:

$$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \hat{y}_i,$$

i.e.:

$$\overline{y} = \overline{\hat{y}}\ .$$

On the basis of the first normal equation, certain general properties of the model are proven. Among those, one needs to distinguish and focus on the general property of the simple linear regression model of great practical import, namely the lack of correlation between $\overline{y}$ mean and $\hat{\beta}$ OLS estimator, which was already mentioned above and even used once to derive a formula for $D^2(\hat{\alpha})$: variance:

$$\operatorname{cov}(\bar{y},\hat{\beta}) = 0 \, .$$

The property needs to be classified in the group of the basic properties of the model. It can be proven into two ways. The following way is probably the simplest:

$$\operatorname{cov}(\bar{y},\hat{\beta}) = \operatorname{cov}(\hat{\alpha}+\hat{\beta}\bar{x},\hat{\beta}) = \operatorname{cov}(\hat{\alpha},\hat{\beta}) + \operatorname{cov}(\hat{\beta}\bar{x},\hat{\beta}) =$$

$$= \operatorname{cov}(\hat{\alpha},\hat{\beta}) + \bar{x}\operatorname{cov}(\hat{\beta},\hat{\beta}) = \frac{-\bar{x}\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} + \bar{x}D^2(\hat{\beta}) = -\frac{\bar{x}\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} + \frac{\bar{x}\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = 0,$$

however, one could conclude the proof with just the second equation (we get 0), taking into account the previously noted equation. Or a different, more direct, though less synthetic method than above, can be undertaken. We have here:

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})y_i}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})y_i}{S_{xx}}$$

$$\bar{y} = \sum_{i=1}^{n}\frac{y_i}{n}$$

hence:

$$\operatorname{cov}(\bar{y},\hat{\beta}) = \operatorname{cov}\Big(\sum_{i=1}^{n}\frac{y_i}{n},\sum_{i=1}^{n}\frac{(x_i-\bar{x})}{S_{xx}}y_i\Big) = \frac{1}{n}\sum_{i=1}^{n}\frac{(x_i-\bar{x})}{S_{xx}}\operatorname{cov}(y_i,y_i) = \frac{\sigma^2}{nS_{xx}}\sum_{i=1}^{n}(x_i-\bar{x}) = 0.$$

It is worth emphasizing that the second method, as opposed to the first one, is far more flexible and ‘more’ open to various generalizations. For instance, it directly demonstrates that not only $\hat{\beta}$ OLS estimator is uncorrelated to $\bar{y}$ mean, but neither is every $\tilde{\beta}$ linear unbiased estimator of β structural parameter. In fact, let $\tilde{\beta} = \sum_{i=1}^{n} d_i y_i$ be any linear unbiased estimator of β structural parameter. Then, from the unbiasedness condition:

$$E(\tilde{\beta}) = \sum_{i=1}^{n}E(d_i y_i) = \sum_{i=1}^{n}d_i E(y_i) = \sum_{i=1}^{n}d_i(\alpha+\beta x_i) = \alpha\sum_{i=1}^{n}d_i + \beta\sum_{i=1}^{n}d_i x_i = \beta\,,$$

results a secondary constraint:

$$\sum_{i=1}^{n}d_i = 0\,,$$

hence we obtain:

$$\mathrm{cov}(\overline{y},\tilde{\beta}) = \mathrm{cov}(\sum_{i=1}^{n}\frac{y_i}{n},\sum_{i=1}^{n}d_i y_i) = \frac{1}{n}\sum_{i=1}^{n}d_i\,\mathrm{cov}(y_i,y_i) = \frac{\sigma^2}{n}\sum_{i=1}^{n}d_i = 0.$$

Furthermore, as we are about to see, this property, which can be easily demonstrated, directly translates into a multidimensional case.

We will consistently calculate here the value of $\overline{y}$ mean covariance with $\hat{\alpha}$ and $\tilde{\alpha}$ estimators, in order to examine the entire problem, and, at the same time, in order to fully close the question of $\overline{y}$ mean correlation with all OLS's and linear unbiased estimators of structural parameters in the simple linear regression model. With regard to $\hat{\alpha}$ OLS estimator, it can be done by using, similarly as was done above, a mechanism of the first normal equation with a further assistance of, as in covariance, a certain type of its self-conjugate. Initially, we have here:

$$\overline{y}-\hat{\alpha}=\hat{\beta}\overline{x},$$

hence:

$$D^2(\overline{y}-\hat{\alpha}) = D^2(\hat{\beta}\overline{x}),$$

$$D^2\left(\overline{y}-\hat{\alpha}\right) = D^2\left(\hat{\alpha}\right) - D^2\left(\overline{y}\right)$$

and further:

$$D^2(\overline{\mathrm{y}}) + D^2(\hat{\alpha}) - 2\,\mathrm{cov}(\overline{\mathrm{y}},\hat{\alpha}) = \overline{x}^2 D^2(\hat{\beta}),$$

$$\frac{\sigma^2}{n} + \sigma^2\left(\frac{1}{n}+\frac{\overline{x}^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}\right) - 2\,\mathrm{cov}(\overline{\mathrm{y}},\hat{\alpha}) = \overline{x}^2\frac{\sigma^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2},$$

$$\mathrm{cov}(\overline{\mathrm{y}},\hat{\alpha}) = \frac{\sigma^2}{n} = \mathrm{cov}(\overline{\mathrm{y}},\overline{\mathrm{y}}) = D^2\left(\overline{\mathrm{y}}\right),$$

or much more succinctly, using direct non-correlation of $\overline{y}$ mean with $\hat{\beta}$ OLS estimator:

$$\mathrm{cov}(\overline{\mathrm{y}},\hat{\alpha}) = \mathrm{cov}(\overline{\mathrm{y}},\overline{y}-\hat{\beta}\overline{x}) = \mathrm{cov}(\overline{\mathrm{y}},\overline{y}) + \mathrm{cov}(\overline{\mathrm{y}},-\hat{\beta}\overline{x}) = D^2(\overline{\mathrm{y}}) = \frac{\sigma^2}{n}.$$

The case of $\tilde{\alpha}$ linear unbiased estimator of α structural parameter is considered similarly to the previous case of $\tilde{\beta}$ linear unbiased estimator of β structural parameter. Let $\tilde{\alpha} = \sum_{i=1}^{n} f_i y_i$ be any linear unbiased estimator of α structural parameter.
From the unbiasedness condition:

$$E(\tilde{\alpha}) = E\left(\sum_{i=1}^{n} f_i y_i\right) = \sum_{i=1}^{n} E(f_i y_i) = \sum_{i=1}^{n} f_i E(y_i) = \sum_{i=1}^{n} f_i(\alpha + \beta x_i) = \alpha\sum_{i=1}^{n} f_i + \beta\sum_{i=1}^{n} f_i x_i = \alpha,$$

a secondary constraint arises:

$$\sum_{i=1}^{n} f_i = 1.$$

It appears that each $\tilde{\alpha}$ linear unbiased estimator of α structural parameter satisfies the dependence such as $\hat{\alpha}$ OLS estimator does, since we have here:

$$\mathrm{cov}(\bar{y}, \tilde{\alpha}) = \mathrm{cov}(\sum_{i=1}^{n} \frac{y_i}{n}, \sum_{i=1}^{n} f_i y_i) = \frac{1}{n}\sum_{i=1}^{n} f_i \,\mathrm{cov}(y_i, y_i) = \frac{\sigma^2}{n}\sum_{i=1}^{n} f_i = \frac{\sigma^2}{n}.$$

Thus, we observe that in the simple linear regression model in the case of $\hat{\alpha}$ OLS estimator and $\tilde{\alpha}$ linear unbiased estimator of α structural parameter, the situation of correlating with $\bar{y}$ mean is slightly different than that of the second of the structural parameters. At this point, we are only able to possibly guess the reason of that what the subsequently conducted deliberations are going to be helpful with. In the context of not just further discussion, but also in a general sense, we receive the following dependency as the resultant synthetic consequence of the above: $\mathrm{cov}(\bar{y}, \hat{y}_i - \bar{y}) = 0$ for any $i = 1, 2, \ldots, n$. Since we have:

$$\begin{aligned} \mathrm{cov}(\bar{y}, \hat{y}_i - \bar{y}) &= \mathrm{cov}(\bar{y}, \hat{y}_i) - \mathrm{cov}(\bar{y}, \bar{y}) = \mathrm{cov}(\bar{y}, \hat{\alpha} + \hat{\beta}x_i) - \mathrm{cov}(\bar{y}, \bar{y}) = \\ &= \mathrm{cov}(\bar{y}, \hat{\alpha}) - \mathrm{cov}(\bar{y}, \bar{y}) = 0, \qquad i = 1, 2, \ldots, n. \end{aligned}$$

Several derivative dependencies of a general character typical for simple linear regression result from the above proven property of non-correlation of $\bar{y}$ mean with $\hat{\beta}$ OLS estimator, and $\tilde{\beta}$ linear unbiased estimators of β structural parameter. We will refer to and demonstrate some of them in a typically synthetic design on account of their significance, as well as their subsequent use in this paper. We shall start from the non-correlation of $\hat{}$ and $\hat{\beta}$ OLS estimators of the structural parameters in the simple linear regression model with the residuals of e_i, $i = 1, 2, \ldots, n$. These particular dependencies are known, however, the manner presented in this paper is different from the one found in the literature of the subject, which is laborious, not entirely clear, and completely failing to reflect the heart of the matter. Owing to the conjugation of proofs only to one side occurring here, which is discussed below, and the related sequentiality, as well as deriving various complexity of the proof process, we shall commence with an easier case and demonstrate it first for $\hat{\beta}$ OLS estimator. We have here for any $i = 1, 2, ..., n$:

$$\mathrm{cov}(\hat{\beta}, e_i) = \mathrm{cov}(\hat{\beta}, \mathrm{y}_i - \hat{\mathrm{y}}_i) = \mathrm{cov}(\hat{\beta}, \mathrm{y}_i) - \mathrm{cov}(\hat{\beta}, \hat{\mathrm{y}}_i) = \mathrm{cov}(\hat{\beta}, \mathrm{y}_i) - \mathrm{cov}(\hat{\beta}, \hat{\alpha} + \hat{\beta} x_i) =$$
$$= \mathrm{cov}(\hat{\beta}, \mathrm{y}_i) - \mathrm{cov}(\hat{\beta}, \hat{\alpha}) - \mathrm{cov}(\hat{\beta}, \hat{\beta}\,\mathrm{x}_i) = \mathrm{cov}(\hat{\beta}, \mathrm{y}_i) - \mathrm{cov}(\hat{\beta}, \hat{\alpha}) - \mathrm{x}_i\,\mathrm{cov}(\hat{\beta}, \hat{\beta}) =$$
$$= \frac{\left(x_i - \overline{x}\right)}{\sum_{i=1}^{n}(x_i - \overline{x})^2}\mathrm{cov}(\mathrm{y}_i, \mathrm{y}_i) + \frac{\overline{x}\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} - x_i\frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = 0.$$

The non-correlation of $\hat{\alpha}$ OLS estimator with e_i, $i = 1, 2, ..., n$ is proven by the use of the above fact in the first normal equation conjugating both OLS estimators:

$$\mathrm{cov}(\hat{\alpha}, e_i) = \mathrm{cov}(\overline{\mathrm{y}} - \hat{\beta}\,\overline{\mathrm{x}}, \mathrm{e}_i) = \mathrm{cov}(\overline{\mathrm{y}}, \mathrm{e}_i) + \mathrm{cov}(-\hat{\beta}\,\overline{\mathrm{x}}, \mathrm{e}_i) =$$
$$= \mathrm{cov}(\overline{\mathrm{y}}, \mathrm{y}_i - \hat{\mathrm{y}}_i) = \mathrm{cov}(\overline{\mathrm{y}}, \mathrm{y}_i) - \mathrm{cov}(\overline{\mathrm{y}}, \hat{\mathrm{y}}_i) = \frac{1}{n}\mathrm{cov}(\mathrm{y}_i, \mathrm{y}_i) - \mathrm{cov}(\overline{\mathrm{y}}, \hat{\alpha} + \hat{\beta}\,\mathrm{x}_i) =$$
$$= \frac{\sigma^2}{n} - \mathrm{cov}(\overline{\mathrm{y}}, \hat{\alpha}) - \mathrm{cov}(\overline{\mathrm{y}}, \hat{\beta}) = \frac{\sigma^2}{n} - \frac{\sigma^2}{n} - 0 = 0.$$

These properties of OLS estimators of the structural parameters in the simple linear regression model and the residuals further entail the non-correlation of theoretical values with residuals, i.e.

$$\mathrm{cov}(\hat{\mathrm{y}}_j, \mathrm{e}_i) = 0, \quad i, j = 1, 2, ..., n,$$

and in particular:

$$\mathrm{cov}(\hat{\mathrm{y}}_i, \mathrm{e}_i) = 0, \quad i = 1, 2, ..., n,$$

and $\overline{y}$ mean, since also through the first normal equation, $\mathrm{cov}(\overline{\mathrm{y}}, \mathrm{e}_i) = 0$ for $i = 1, 2, ..., n$.

The subsequent value will be the value of a variance $D^2(\hat{\mathrm{y}}_i - \overline{y})$, $i = 1, 2, ..., n$ displaced by $\overline{y}$ mean, necessary further on in the aspect of its calculation. The property of the non-correlation of $\overline{y}$ mean with $\hat{\beta}$ OLS estimator simplifies the process to a significant degree. Namely, we have here:

$$D^2(\hat{y}_i - \overline{y}) = D^2(\hat{y}_i) + D^2(\overline{y}) - 2\,\mathrm{cov}(\hat{y}_i, \overline{y}),$$
$$D^2(\hat{y}_i - \overline{y}) = D^2(\hat{y}_i) + D^2(\overline{y}) - 2\,\mathrm{cov}(\hat{\alpha}, \overline{y}),$$

$$D^2(\hat{y}_i - \overline{y}) = D^2(\hat{y}_i) + D^2(\overline{y}) - 2\,\mathrm{cov}(\overline{y} - \beta\overline{x}, \overline{y}),$$
$$D^2(\hat{y}_i - \overline{y}) = D^2(\hat{y}_i) + D^2(\overline{y}) - 2\,\mathrm{cov}(\overline{y}, \overline{y}),$$
$$D^2(\hat{y}_i - \overline{y}) = D^2(\hat{y}_i) + D^2(\overline{y}) - 2D^2(\overline{y}),$$
$$D^2(\hat{y}_i - \overline{y}) = D^2(\hat{y}_i) - D^2(\overline{y}),$$

or much more succinctly, starting from the equation:

$$\hat{y}_i = \left(\hat{y}_i - \overline{y}\right) + \overline{y},$$

hence:

$$D^2(\hat{y}_i) = D^2(\hat{y}_i - \overline{y} + \overline{y}) = D^2(\hat{y}_i - \overline{y}) + D^2(\overline{y}) - 2\,\mathrm{cov}(\hat{y}_i - \overline{y}, \overline{y}),$$

that is:

$$D^2(\hat{y}_i) = D^2(\hat{y}_i - \overline{y}) + D^2(\overline{y}).$$

The last consideration in this aspect, i.e. numerical one, to be analysed is the reference of $\hat{\alpha}$ and $\hat{\beta}$ OLS estimators of the model structural parameters to $\mathrm{cov}\left(\hat{\alpha}, \hat{\beta}\right)$ covariance, which can be determined differently than was done previously, more simply, while still remaining in the complete methodological trend of the presented approach. Here, in turn, we have:

$$\mathrm{cov}(\hat{\alpha}, \hat{\beta}) = \mathrm{cov}(\overline{\mathrm{y}} - \hat{\beta}\overline{x}, \hat{\beta}) = \mathrm{cov}(-\hat{\beta}\overline{x}, \hat{\beta}) = -\overline{x}\,\mathrm{D}^2(\hat{\beta}) = -\frac{\overline{x}\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}.$$

We can obtain it equally straightforwardly by starting from this property directly, and not as above, by using it indirectly:

$$\mathrm{cov}(\overline{\mathrm{y}}, \hat{\beta}) = \mathrm{cov}(\hat{\alpha} + \hat{\beta}\overline{x}, \hat{\beta}) = \mathrm{cov}(\hat{\alpha}, \hat{\beta}) + \overline{x}D^2(\hat{\beta}) = 0.$$

As already mentioned, some of the dependencies for the simple linear regression model proven in this paper are also typical to the multiple linear regression model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon.$$

Since in that model $\overline{y}$ average is also uncorrelated with all OLS estimators and linear unbiased estimators of the structural parameters, except for the absolute term, with independent variables:

$$\mathrm{cov}(\overline{\mathrm{y}}, \hat{\beta}_j) = 0, \qquad j = 1, 2, \ldots, k.$$

The proof of this parameter is conducted immediately for any linear unbiased estimator of the structural parameter β_j, $j=1, 2, ..., k$. We deal here with a general form of the model:

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+...+\beta_k x_k+\varepsilon,$$

and the form of the model in the design comprising all observations:

$$y_i=\beta_0+\beta_1 x_{1i}+\beta_2 x_{2i}+...+\beta_k x_{ki}+\varepsilon_i, \qquad i=1, 2, ..., n.$$

Let $\tilde{\beta}_j=\sum_{i=1}^{n} d_{ji} y_i$, $j=1, 2, ..., k$ be any linear unbiased estimator of the structural parameter β_j, $j=1, 2, ..., k$. From the condition of unbiasedness:

$$E(\tilde{\beta}_j)=\sum_{i=1}^{n} d_{ji} E(y_i)=\sum_{i=1}^{n} d_{ji} E(\beta_0+\beta_1 x_{1i}+\beta_2 x_{2i}+...+\beta_k x_{ki}+\varepsilon_i)=$$

$$=\beta_0\sum_{i=1}^{n} d_{ji}+\beta_1\sum_{i=1}^{n} d_{ji} x_{1i}+\beta_2\sum_{i=1}^{n} d_{ji} x_{2i}+...+\beta_k\sum_{i=1}^{n} d_{ji} x_{ki}=\beta_j, \qquad j=1, 2, ..., k,$$

results the following secondary constraint:

$$\sum_{i=1}^{n} d_{ji}=0,$$

whence we obtain:

$$\operatorname{cov}\left(\bar{y},\tilde{\beta}_j\right)=\operatorname{cov}\left(\sum_{i=1}^{n}\frac{y_i}{n},\sum_{i=1}^{n} d_{ji} y_i\right)=\frac{1}{n}\sum_{i=1}^{n} d_{ji}\operatorname{cov}\left(y_i, y_i\right)=\frac{\sigma^2}{n}\sum_{i=1}^{n} d_{ji}=0, \qquad j=1, 2, ..., k.$$

As a consequence of the above, the dependency for $\hat{\beta}_0$, OSL estimator is automatically repeated, since:

$$\operatorname{cov}\left(\bar{y},\hat{\beta}_0\right)=\operatorname{cov}\left(\bar{y},\bar{y}-\sum_{j=1}^{k}\hat{\beta}_j\bar{x}_j\right)=\operatorname{cov}\left(\bar{y},\bar{y}\right)=D^2\left(\bar{y}\right)=\frac{\sigma^2}{n}$$

and the same holds true for any linear estimator of unbiased $\tilde{\beta}_0$ of β_0, structural parameter, because we have as follows. Let $\tilde{\beta}_0=\sum_{i=1}^{n} f_i y_i$ be any linear unbiased estimator of β_0, structural parameter. From the unbiasedness condition:

$$E(\tilde{\beta}_0)=\sum_{i=1}^{n} f_i E(y_i)=\sum_{i=1}^{n} f_i E(\beta_0+\beta_1 x_{1i}+\beta_2 x_{2i}+...+\beta_k x_{ki}+\varepsilon_i)=$$

$$=\beta_0\sum_{i=1}^{n} f_i+\beta_1\sum_{i=1}^{n} f_i x_{1i}+\beta_2\sum_{i=1}^{n} f_i x_{2i}+...+\beta_k\sum_{i=1}^{n} f_i x_{ki}=\beta_0,$$

results a secondary constraint:

$$\sum_{i=1}^{n} f_i = 1,$$

from which we obtain:

$$\operatorname{cov}(\overline{y},\tilde{\beta}_0) = \operatorname{cov}(\sum_{i=1}^{n}\frac{y_i}{n},\sum_{i=1}^{n} f_i y_i) = \frac{1}{n}\sum_{i=1}^{n} f_i \operatorname{cov}(y_i, y_i) = \frac{\sigma^2}{n}\sum_{i=1}^{n} f_i = \frac{\sigma^2}{n}.$$

Consequently, a resultant dependency is also repeated for this model:

$$\operatorname{cov}(\overline{\mathrm{y}},\hat{y}_i - \overline{\mathrm{y}}) = 0.$$

The latter one, and indeed the fact of the non-correlation of $\overline{y}$ mean, further results in $D^2(\hat{y}_i - \overline{y}) = D^2(\hat{y}_i) - D^2(\overline{y}),$ in the multiple linear regression model, which is directly derived in the same manner.

Finally, let us consider the issue of unbiasedness with regard to linear estimators of both structural parameters of the simple linear regression model in a basic design. They demonstrate one curious, continuous property, significant from the cognitive standpoint. We shall prove it first for such estimators of α structural parameter. Let us suppose that $\tilde{\alpha}_1, \tilde{\alpha}_2$ are any two linear unbiased estimators of α structural parameter of a simple linear regression model. Then $\lambda\tilde{\alpha}_1 + (1-\lambda)\tilde{\alpha}_2,$ estimator, where λ – any real number, is also a linear unbiased estimator of α structural parameter. The proof is very simple here. Let $\tilde{\alpha}_1 = \sum_{i=1}^{n} f_{1i} y_i$ and $\tilde{\alpha}_2 = \sum_{i=1}^{n} f_{2i} y_i$ be any two linear unbiased estimators of α structural parameter. From this condition the following secondary constraints arise:

$$\begin{cases} \sum_{i=1}^{n} f_{1i} = 1 \\ \sum_{i=1}^{n} f_{1i} x_i = 0 \end{cases}$$

for the distribution weights of $\tilde{\alpha}_1$ estimator and secondary constraints:

$$\begin{cases} \sum_{i=1}^{n} f_{2i} = 1 \\ \sum_{i=1}^{n} f_{2i} x_i = 0 \end{cases}$$

for the distribution weights of $\tilde{\alpha}_2$ estimator. Multiplying the first design by λ and the second one by $(1-\lambda)$, then adding the sides of corresponding equations, we get:

$$\begin{cases} \sum_{i=1}^{n} \left(\lambda f_{1i} + (1-\lambda) f_{2i} \right) = 1 \\ \sum_{i=1}^{n} \left(\lambda f_{1i} + (1-\lambda) f_{2i} \right) x_i = 0. \end{cases}$$

These equations already directly determine the unbiasedness of $\lambda\tilde{\alpha}_1 + (1-\lambda)\tilde{\alpha}_2$ estimator. Interestingly, this property can also be proven in a much simpler manner, by calculating the expected value of $\lambda\tilde{\alpha}_1 + (1-\lambda)\tilde{\alpha}_2$ estimator. We then have the following:

$$E\left(\lambda\tilde{\alpha}_1 + (1-\lambda)\tilde{\alpha}_2\right) = E\left(\lambda\tilde{\alpha}_1\right) + E\left((1-\lambda)\tilde{\alpha}_2\right) =$$
$$\lambda E\left(\tilde{\alpha}_1\right) + (1-\lambda) E\left(\tilde{\alpha}_2\right) = \lambda\alpha + (1-\lambda)\alpha = \alpha.$$

It is identically proven for a linear unbiased estimator of the second structural parameter, i.e. β Generally, however, as was previously mentioned, the nature of the very unbiasedness has not been fully investigated yet. It seems to be hiding and opening certain areas for further research exploration in that regard, particularly in the interpretational layer.

Conclusions

The synthetic approach proposed in the paper, based on variance and using certain classic dependencies, simplifies, at least to a certain degree, the deliberations related to the simple linear regression model and brings some new and fresh ideas into it. Simultaneously, it further shows that new approaches or solutions can still be found in the models that at present are considered to be classic. And such endeavours ought to be undertaken. Every different look at a question contributes to developing new, ever more effective solutions in a given area.

References

Johnston, J. (1987). *Econometric Methods*. New York: McGraw-Hill.

Kmenta, J. (1986). *Elements of Econometrics*. New York: Macmillan.

Maddala, G.S. (2006). *Ekonometria*. Warszawa: Wydawnictwo Naukowe PWN.

Magnus, J.R., Katyszew, P.K., Pieriesieckij, A.A. (1997). *Ekonometrika*. Moskwa: Naczalnyj kurs, Delo.

Spanos, A. (1986). *Statistical Foundations of Econometric Modelling*. Cambridge: Cambridge University Press.

Weintraub, E.R. (1982). *Mathematics for Economics*. An Integrated Approach. Cambridge: Cambridge University Press.

Welfe, A. (2009). *Ekonometria. Metody i zastosowania*. Warszawa: Polskie Wydawnictwo Ekonomiczne.

Wooldridge, J.M. (2009). *Introductory Econometrics. A Modern Approach*. Fourth Edition. Canada: South-Western.