TOWARDS AN ACHIEVEMENT INDEX FOR COMPUTER SCIENCE STUDENTS

Sira Allende*[1], Carlos N. Bouza* Jose F. Garcia**, Jose M. Sautto*** and Hemant K.Verma****
*Facultad Matemática y Computación, Universidad de La Habana, Cuba.
bouza@matcom.uh.cu
**DACEA, Universidad Juárez Autónoma de Guerrero, México.
*** Campus Costa Chica, Universidad Autónoma de Guerrero, Mexico
**** Department of Statistics, Banaras Hindu University, Varanasi, India.

ABSTRACT.
We developed an index for measuring the achievement of students of Computer Sciences. The index was modeled by a regression equation. Different models were developed. Their behavior was evaluated for deciding which equation leaded to more accurate forecasting.

KEYWORDS: Least Squares, Least Absolute Deviation, Regression,

MSC: 97D60

RESUMEN
Hemos desarrollado un índex pare medir el aprovechamiento de los estudiantes de Ciencias de la Computación. Este es modelado mediante una regresión. Diferentes modelos se proponen. Su comportamiento fue evaluado para decidir que ecuación produce las predicciones mas exactas.

PALABRAS CLAVES: Mínimos Cuadrados, Mínima Desviación Media, Regresión.

1. INTRODUCTION

Education is an investment of the students. Hence they are interested in knowing the expected behavior of them when going into the labor market. The educational institutions are involved with different projects with the industry where students develop professional practice. It is of interest predicting the achievements of students using the information on their results up to the first years of college. We consider the development of an index for predicting the achievement of students. This index will be used by the students for considering their expectations, by the institutions for classifying the students in terms of their potentialities and by the industry for recruiting new workers.
We consider that when finishing the first years at college is possible to predict the degree of achievements of each student. The experts considered the variables which should be used, at a college devoted computer science, will be the notes obtained at High School, the classification tests and the notes obtained in the first years.
The degree of achievement is given by the criteria of a team of professors and employers. Professors consider the successfulness of the students considering their skills and employers give an evaluation after a training period. See recent research in the theme as Chen-Liao (2013), .Pratiwi et al. (2021), Lee (2021), Liem (2021), Davison (2015), Regenwetter-Ahmed (2022), Incekara et al., (2017) among others.
The relationships among achievement and the obtained notes suggest using a regression model. Using the information provided by students, that obtained their degrees in the course 2009-2012 we investigated the goodness of fitting a regression equation for forecasting the achievement. The normal based models are of common use, see Anderson (2003) and Johnson-Wichern. (2002). Due to the characteristic of the involved variables we considered not only the usual Normal based model but also the Laplace regression model. Considering that the residuals are Laplacians is not commonly studied. Section 2 is devoted to presenting the needed theoretical results on the Laplace distribution. Section 3 is concerned with details on multiple regression fitting and Section 4 presents the study developed for fitting the regressions. The results of the course 2012 were used for evaluating the behavior of the developed models.

[1] sira@matcom.uh.cu

2. SOME ASPECTS OF LAPLACE DISTRIBUTION

The Laplace or double exponential distribution arises commonly in problems supposedly normal in which the tails are heavier. That is the case in economics where the man intervenes in the process that generates random variables.

The classic Laplace density function (CLDF) is given by $L(\mu,\tau)=f(x|\mu,\tau)=\frac{e^{-|x-\mu|/\tau}}{2\tau}, \mu, x\in\Re, \tau\in\Re_+$

It is worth mentioning that $CLDF = \Psi_X(t)=\frac{e^{it\mu}}{1+\tau^2t^2}, t\in\Re$ is the characteristic function and $M_X(t)=(1-t^2)^{-1}$, is the moment generating function. Hence, is easily derived that $E(X)=\mu$, $V(X)=2\tau^2$ and the mean deviation is $MD(X)=E|X-\mu|=\tau$. Therefore $MD(X)/\sqrt{V(X)}=\tau/\tau\sqrt{2}\cong 0{,}71$.

In the special case $X\sim L(0,1)$ we deal with the standard (S)CLDF. Reparametrizing it accordingly with the usual normal notation we may write

$$L(\mu,\sigma)=f(x|\mu,\sigma)=\frac{e^{-\sqrt{2}|x-\mu|/\sigma}}{\sqrt{2}\sigma}, \mu, x\in\Re, \sigma\in\Re_+$$

$L(\mu,\sigma)$ is identified as "Laplace distribution" (LDF).

Calculating the moments of a CLDF we have that the coefficient of skewness

$$\gamma_1=\frac{E(X-E(X))^3}{\sqrt[3/2]{E(X-E(X))^2}}=0, \gamma_2=\frac{E(X-E(X))^4}{V(X)^2}-3=3.$$

Therefore, this distribution is symmetric and leptokurtic. See Kotz et al. (2001) for a detailed discussion of the Laplace distribution.

The CLDF arouses as the difference of two exponential variables W and W*. Hence the SCLDF is a result of the difference of two standard exponentially distributed rv`s. They are also generated by the product of a standard normal Z with a standard exponential W_0, a Rayleigh, or with $T^{-1}\sqrt{2}$, T . Then we have that the distribution is the brittle fracture density function $f_T(x)=\frac{2e^{\frac{1}{x^2}}}{x^3}$.

It is worth noting that exponential, Rayleigh and bi-fracture distributions are important in modeling the behavior of technical devices. As $W_0=\log P$, P a Pareto rv, $X=\log P/Q$ is SCLDF, if P and Q are Pareto. The same result follows when P and Q are uniform rv. The connection with economical issues is also derived from this fact. The CLDF is easily derived from any of these SCLDF`s.

It is interesting that if we have a geometric rv ν_p with $E(\nu_p)=1/p$, $p\in(0,1)$, X_i, $i=1,\ldots$, non-degenerate, symmetric and iid rv with common variance $\sigma^2>0$ and holds the distribution equality $p^{\frac{1}{2}}\sum_{i=1}^{\nu_p}X_i =^d X_0$

Then $X_0\sim L(0,\sigma^2)$. When ν_p is independent of the X_i`s and $\left(p^{\frac{1}{2}}\gamma+o(p^{1/2})\right)\sum_{i=1}^{\nu_p}X_i \to^d X_0$. We have that $X_0\sim L(0,\gamma^2\sigma^2)$.

The last two representations of Laplace rv`s will play an important role in the sequel. In the standard bivariate case the characteristic function is

$$\Psi_X(t_1,t_2)=\frac{1}{1+½(\sigma_1^2t_1^2+\sigma_2^2t_2^2)+\rho\sigma_1\ t_1\ \sigma_2\ t_2}, X\in\Re^2, \sigma_i\ >0, i=1,2;\ \rho\in[0,1]$$

Clearly

$$E(X)=\binom{0}{0}, V(X)=E(XX^T)=\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\ \sigma_2\\ \rho\sigma_1\ \sigma_2 & \sigma_2^2\end{bmatrix}$$

and in the case $X\in\Re^k$,

$$E(X)=\vec{0}\in\Re^k, V(X)=\vartheta_{k\times k}$$

$\vartheta_{k\times k}$ is the covariance matrix.

3. MODELING ACHIEVEMENT

The behavior of the achievements depends on a subjective evaluation made by a commission of potential employers and professors. They reached a consensus and an evaluation Y was given to each student. The control variables were the notes obtained in:

1. Classification exam in Hard Sciences
2. Classification exam in Soft Sciences
3. Overall High School in Hard Sciences

4. Overall High School in Soft Sciences
5. Average in Mathematics matters in the first two years in college
6. Average Computer Science matters in the first two years in college

The notes moved within (30, 100). The data provided the vector Y_i, $X_{i1},\dots,X_{i6}$. The relationship among them was supposed to be

$$Y_i = \varphi(X_{i1},\dots,X_{i6}) + \varepsilon_i, i = 1,\dots.,n$$

Definition: $\varphi(X_{i1},\dots,X_{i6})$ is an achievement index.□

We consider that the relation holds for some φ and ε_i follows a certain distribution. We look for an estimation of the index.

The usual assumption is that ε follows a Gaussian distribution. We considered that an alternative is considering a Laplacian. The motivation is given by the fact that a Laplacian can be viewed as generated by a function of a mixture of random variables, as quoted in the second section. Note that the residuals are Standard Laplacians one of the following probabilistic representation is valid that $\varepsilon \sim L(0,\sigma^2)$ *if*

$$\varepsilon = \begin{cases} Z\sqrt{2\epsilon}, \ Z \textit{ is a standard normal variable}, \epsilon \textit{ is a standard exponential} \\ ZR, \ Z \textit{ is a standard normal variable}, R \textit{ is a Rayleigh} \\ \frac{Z}{B}\sqrt{2}, \ Z \textit{ is a standard normal variable}, B \textit{ is a brittle fracture} \\ \epsilon - W, \quad \epsilon \textit{ and } W \textit{ are standard exponentials} \\ I\epsilon, \quad I \textit{ is equal to } + \textit{ or } - \textit{ with equal probalilities and } \epsilon \textit{ is a standard exponential} \\ log\left(\frac{P}{Q}\right), \ P \textit{ and } Q \textit{ are iid Pareto type I} \\ log\left(\frac{U}{Q}\right), \ U \textit{ and } Q \textit{ are iid Unifom in } (0,1) \\ Z_1Z_2 - Z_3Z_4, \ Z_i, i = 1,.2,3,4 \textit{ is a standard normal variable} \end{cases}$$

Note that in addition, if we take the geometric RV ν_p with $E(\nu_p)=1/p$, $p \in (0,1)$, some unknown symmetric non degenerate random effects Q_i, $t=1,\dots,$ with common variance $\sigma^2>0$, we may model the residuals as a sum of random effects and using $p^{\frac{1}{2}}\sum_{t=1}^{\nu_p} Q_t =^d \varepsilon$. We accept that $\varepsilon \sim L(0, \sigma^2)$. Under the independence of ν_p and the random effects

$$\left(p^{\frac{1}{2}}\gamma + o(p^{1/2})\right)\sum_{t=1}^{\nu_p} Q_t \to^d \varepsilon$$

and $\varepsilon \sim L(0,\gamma^2\sigma^2)$.

Therefore, the variety of mixtures of distributions generating the residuals is larger than considering simply that $\varepsilon \sim N(0, \sigma^2)$.

Clearly the evaluators considered the values of the notes and information on the results of the practice in institutions of the students during their studies. ε is considered as a noise. We considered that the set of indexes of achievements is a random vector. Its distribution may be considered as a multivariate normal. The justification that we deal with a multivariate Laplacian may come from the assumption that.

$$\underline{Y} =^d \underline{\mu}\epsilon + \underline{Z}\sqrt{\epsilon}, \qquad \epsilon \sim \exp(1), \underline{Z} \sim N_k\left(\underline{0}, \vartheta\right); \ \epsilon \textit{ and } \underline{Z} \textit{ independent}$$

Another justification is to consider that the stability property given by considering the existence of a geometrically distributed RV ν_p, independent of the $\underline{X}_i$`s such that

$$a_p \sum_{i=1}^{\nu_p}\left(\underline{X}_i + \underline{b}_p\right) \to^d \underline{Y} \textit{ as } p \to 0; \ a_p > 0, \underline{b}_p \in \Re^k$$

holds

The condition for accepting that the regression is linear is identical to those sustaining the model in the Gaussian case; see Kotz et al. (2001). The density of $\underline{Y}$ is

$$g\left(\underline{y}\right) = \frac{\exp\left(\underline{y}^T\vartheta^{-1}\underline{\mu}\right)\left(1 + ½\underline{\mu}^T\vartheta^{-1}\underline{\mu}\right)^{\frac{d}{2}-1}}{(2\pi)^{k/2}|\vartheta|^{1/2}}\int_0^\infty \exp\left(-\frac{a^2}{4z} - z\right) z^{-\frac{d-2}{2}-1} dz$$

$$a = \sqrt{2 + \left(\underline{\mu}^T\vartheta^{-1}\underline{\mu}\right)\left(\underline{y}^T\vartheta^{-1}\underline{y}\right)}$$

v_p is independent of the X_i`s .

4. PREDICTION OF THE ACHIEVEMENT INDEX

Consider a linear regression model

$$\varphi(X_{i1},\dots,X_{i6}) = \sum_{t=0}^{T} B_t\, \varphi_t(X_{i1},\dots,X_{i6}), i = 1,\dots.,n.$$

We consider adjusting the parameters of the model function to best fit a data set. A common assumption is that the distribution of the errors is a Normal distribution with zero-expectation, conditional on the independent variables, uncorrelated and having equal variances. In such cases the best linear unbiased estimator of any linear combination of the observations is obtained solving the L_2 optimization problem

$$Min_{B_1,\dots,B_T}\{\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n}(Y_i - \sum_{t=0}^{T} B_t\, \varphi_t(X_{i1},\dots,X_{i6}))^2\} = \left\|\underline{Y} - \underline{X}\ \underline{B}\right\|^2,.$$

The solution of this problem leads to the well known least-squares estimators

$$\underline{\hat{b}}_{lsN} = \left(\underline{X}^T\underline{X}\right)^{-1}\underline{X}^T\underline{Y}$$

See Kotz et al. (2001).

Under the assumption of normality they are also maximum likelihood estimators. The residual vector is

$$\underline{\hat{\varepsilon}}_N = \underline{Y} - \underline{X}\ \underline{\hat{b}}_{lsN}$$

When we deal with Laplace distributed errors the LS problem is essentially the same and

$$\underline{\hat{b}}_{lsL} = \left(\underline{X}^T\underline{X}\right)^{-1}\underline{X}^T\underline{Y}$$

is the LS-estimator.

The hypotheses of Gauss-Markov Theorem are valid in both cases, but the regression parameters estimator and the residuals are not independent if they have a Laplace distribution.

An alternative model is obtained by considering a L_1 optimization problem

$$Min_{B_1,\dots,B_T}\{\sum_{i=1}^{n}|\varepsilon_i| = \sum_{i=1}^{n} |Y_i - \sum_{t=0}^{T} B_t\, \varphi_t(X_{i1},\dots,X_{i6})|\},.$$

This is known as least absolute deviations (LAD) or Least Absolute Errors (LAE), Least Absolute Value (LAV). Though it is structurally similar to the L_2 problem it poses a different computation problem. Its solution leads to the maximum likelihood estimate if the errors have a Laplace distribution. The computation of the regression parameters has been studied in detail for decades. The robustness of LAD under certein circumstances has been proved, see Thanoon (2015) an Chen-Derezi´nski (2021). Barrodale (1973) has pointed out how LAD does not have an analytical solving method and that iterative methods are needed. He proposed an algorithm. Osborne (1987) developed another algorithm which was thoroughly used. Nowadays, some usual techniques available for solving LAD are:

- Simplex-based methods
- Iteratively re-weighted least squares
- Wesolowsky's direct descent method
- Li-Arce's maximum likelihood approach, Li-Arce (2004).

Simplex-based methods are commonly used as they are available in standard optimization software packages. For LAD there are not inferential results as those available for the LS problem where we may derive confidence ellipsoid $\underline{\hat{b}}_{lsL}$for using the F-distribution. In the LS case we have for a confidence coefficient α the confidence region

$$(\underline{b} - \underline{\hat{b}}_{lsL})^T\underline{X}^T\underline{X}(\underline{b} - \underline{\hat{b}}_{lsL}) \quad \le k\,\frac{\underline{\hat{\varepsilon}^T}_L\,\underline{\hat{\varepsilon}}_L}{n-k}F_{k,n-k}(1-\alpha)$$

We are able to forecast the index of achievement of a student with notes $\underline{X}_0 = (X_{01},\dots,X_{06})^T$ using the regression equation fitted:

$$\hat{Y}_0 = \underline{X}_0^T\ \underline{\hat{b}}$$

As both LS- estimators satisfy the Gauss-Markov hypothesis, for the Normal and the Laplace cases, this prediction is the Best Linear Unbiased Estimator of $\underline{Y_0}$ and

$$t_0\,\frac{\sqrt{(n-k)}\left(Y_0 - \underline{X}_0^T\ \underline{\hat{b}}\ \ \right)}{\sqrt{\left(\underline{\hat{\varepsilon}}^T\quad \underline{\hat{\varepsilon}}\ \ \right)\left(1+\underline{X}_0^T(\underline{X}^TX)^{-1}\underline{X}_0^T\right)}}$$

has a T-Student distribution with n-k degrees of freedom.

In some situations is needed to establish some constraints. A particular study was developed by of Shi-Lukas (2002).

Definition. A prediction of the achievement index of a student with note $\underline{X}_0 = (X_{01}, \dots, X_{06})^T$ is $\hat{Y}_0 = \underline{X}_0^T \, \underline{\hat{b}}$.□

There are 3 ways of assessing to our college:

1. Public School
2. Private School
3. Grant form the government

We considered a model for each source of High School graduates and for the whole group considered of 319 students. The structure is given in the following table

Table 1. Number of graduate students

Source	2008	2010	2011	total
Public School	32	27	44	103
Private School	29	21	32	82
Grants	42	48	44	134
total	103	96	120	319

The initial model considered used as controlled variables $\underline{X} = (X_1, \dots, X_6)^T$ and the cross products of degree up to 4. The regression equations obtained were:

1. LS under the Normal model

$$Public\ School: \hat{Y} = 25{,}1 - 0{,}94X_1 - 0{,}99X_2 - 1{,}02X_3 + 1{,}64X_4 + 2{,}64\sqrt{X_1X_3} - 1{,}05\frac{X_5}{X_6}$$

$$Private\ School: \hat{Y} = 38{,}2 + 0{,}18X_3 + 0{,}27X_4 + 3{,}72\sqrt{X_4/X_3}$$

$$Grants: \hat{Y} = 7{,}88 + 0{,}29X_1 + 0{,}17X_2 + 0{,}12X_3 + 0{,}09X_4 + 0{,}11\sqrt{X_4X_5/X_1}$$

$$Total: \hat{Y} = 16{,}3 + 0{,}26X_4 + 0{,}07X_5 + 0{,}48\sqrt{X_6}$$

Note that Classification exam in Hard Sciences is unimportant for private Schools and in the population model. It works negatively in public schools. A similar behavior is present in Classification exam in Soft Sciences. The meaning of the cross variables must motivate further pedagogical research.

2. LS under the Laplace model

$$Public\ School: \hat{Y} = 36{,}6 - 1{,}02X_1 - 1{,}12X_2 + 0{,}87X_3 + 0{,}23X_4 + 1{,}62\sqrt{X_4/X_3}$$

$$Private\ School: \hat{Y} = 59{,}9 + 0{,}07X_3 + 0{,}12X_4 + 0{,}29\sqrt{X_6X_5}$$

$$Grants: \hat{Y} = 11{,}74 + 0{,}10X_1 + 0{,}08X_2 + 0{,}11X_3 + 0{,}06X_4 + 0{,}54\sqrt{X_4X_5/X_1}$$

$$Total: \hat{Y} = 9{,}23 + 0{,}33X_3 + 0{,}23X_4 + 3{,}34\sqrt{X_5}$$

It is interesting that Classification exam in Hard Sciences and in Soft Sciences have a behavior similar in public and private schools: they have an unimportant role. The notes in Overall High School in Hard Sciences and in Soft Sciences play a very important role in predicting the achievement. This fact is particularly important in public schools. The interactions between the averages in Mathematics and in Computer Science matters, in the first two years in college, have a considerable weigh in the prediction in private schools. The direct interaction between Overall High School in Soft Sciences and the Average in Mathematics in the first two years in college is modified inversely by the Classification exam in Hard Sciences for the grants.

3. LAD under the Laplace model

$$Public\ School: \hat{Y} = 5{,}17 + 0{,}31X_1 - 0{,}08X_2 + 0{,}05X_4 + 0{,}49X_5 + 1{,}62\sqrt{X_6}$$

$$Private\ School: \hat{Y} = 48{,}7 + 0{,}01X_3 + 0{,}28\frac{X_3}{X_4} + 0{,}24\frac{X_3}{X_5}$$

$$Grants: \hat{Y} = 22{,}47 + 0{,}07X_1 + 0{,}01X_3 + 0{,}26X_4 + 0{,}42\sqrt{X_4X_5/X_1}$$

$$Total: \hat{Y} = 12{,}35 + 0{,}48X_3 + 0{,}14X_4 + 0{,}27\sqrt{X_5/X_6}$$

Classification exam in Hard Sciences and in Soft Sciences play jointly a role only in public schools. Classification exam in Hard Sciences is important in the forecasts in public schools and is very low in the grants. The Overall High School note in Hard Sciences is the leading variable in the prediction of achievement in private schools, where it appears with a small weigh as a direct variable, but is reduced by the values of Overall notes in High School in Soft Sciences and the .Average in Mathematics matters in the first two years in college. For the grants the Overall High School in Soft Sciences appears as an important factor both linearly and weighted by the ratio of the Average in Mathematics matters in the first two years in college and the Classification exam in Hard Sciences. For the total population Overall High School in Hard Sciences and Overall High School in Soft Sciences appear with considerably high weights as well as the ratio of the Average in Mathematics matters in the first two years in college with respect to the Classification exam in Hard Sciences.

The graduate students of the course 2012 were evaluated and we analyzed the behavior of the predicted achievement index $\hat{Y}_{0i}$ comparing it with the value Y_0 assigned by the experts.

The accuracy was evaluated by computing

$$A(M) = \frac{\sum_{i=1}^{r(M)} \left|\hat{Y}_{0i} - Y_0\right|}{r(M)}, \qquad M = LSNormal, LSLaplace, LAD$$

The results obtained are given in the following table

Table. 2. Mean Absolute error of the predictions

Source	A(LSNormal)	A(LSLaplace)	A(LAD)l
Public School	23,07	12,9o	12,16
Private School	25,15	14,44	14,28
Grants	26,96	11,52	11,31
total	43,27	21,10	23,58

Note that the use of Laplace model is a better choice than the Normal one. LS and LAD had a similar behavior.

The number of graduate students were larger than 25 in all the cases. Then we can use the T-Student tests. We considered the hypothesis H$_0$: $E(\hat{Y}_0) = Y_0$ and defined the Bernoulli RV

$$I(Y_{0i}) = \begin{cases} 1 \ if \ E(\hat{Y}_{0i}) = Y_{0i} \ is \ accepted \\ \quad 0 \ otherwise \end{cases}$$

Fixing $P=1-\alpha=0{,}95$ and evaluating a sample of size $r(M)$ we performed test of hypothesis that H_0: $P=0{,}95$ for each M.

We computed

$$p(M) = \frac{\sum_{i=1}^{r(M)} I(Y_{0i})}{r(M)}$$

And used the fact that $\sum_{i=1}^{r(M)} I(Y_{0i})$

The results are given in the next table is a Binomial RV with parameters $r(M)$ and $P(M)=0{,}95$.

Table. 3. Estimated $P(M)$ and p-value of the hypothesis $P(M)\geq 0{,}95$

Source	A(LSNormal)		A(LSLaplace)		A(LAD)	
	p(M)	p-value	*p(M)*	p-value	*p(M)*	p-value
Public School	0,667	0,048	0,697	0,054	0,737	0,066
Private School	0,583	0,031	0,546	0,018	0,674	0,056
Grants	0,734	0,065	0,519	0,012	0,852	0,577
total	0,747	0,003	0,694	0,001	0,661	0,001

From the result in the above table we have that:

1. LS-Laplace and LAD models behave adequately for the prediction in Public Schools.
2. LAD is the only model behaving adequately for Private Schools.
3. LS-Normal and LAD behave adequately for Grants.
4. There is a bad behavior for the prediction of the whole population.
5. LAD has the best behavior as a whole.

5. CONCLUSIONS

The use of the proposed procedure seems to be good for predicting the achievement of students.

The importance of the variables varies considerably in dependence of the model and the kind of institution of the students. The prediction must be developed taking into account the type of students.

ACKNOWLEDGEMENTS: This research was supported by the project PN223LH010-005 Desarrollo de nuevos modelos y métodos matemáticos para la toma de decisiones

RECEIVED: MARCH 2021.
REVISED: MARCH 2022.

REFERENCES

[1] ANDERSON, T. W. (2003): **An Introduction to Multivariate Statistical Analysis,** 3rd edn. John Wiley and Sons, New York.

[2] BARRODALE, F.D.K. R. (1973): An improved algorithm for discrete L1 linear approximation. **SIAM J. Numer. Anal**., 10, 839-848.

[3] CHEN X. and M. DEREZI´NSKI (2021): **Query Complexity of Least Absolute Deviation Regression via Robust Uniform Convergence,** arXiv:2102.02322v2.

[4] CHEN, M.H, LIAO, J.L. (2013). Correlations among Learning Motivation, Life Stress, Learning Satisfaction, and Self-Efficacy for Ph.D Students. **The Journal of International Management Studies**, 8, 157 – 162.

[5] DAVISON, C.B. (2015). Assessing IT Student Performance Using Virtual Machines. **Tech Directions**, 74, 23-25.

[6] INCEKARA, A., T. GUZ and G. SENGUN, (2017): Measuring the technology achievement index: comparison and ranking of countries. **Journal of Economics, Finance and Accounting** (JEFA), 4, 164-174

[7] KOTZ, S., T. J. KOZUBOWSKI and K. PODGÓRSKI (2001): **The Laplace Distribution and Generalizations. A Revisit with Applications to Communications, Economics, Engineering and Finance.** Birkhäuser, Boston.

[8] JOHNSON, R:: and WICHERN, D. W. (2002): **Applied Multivariate Statistical Analysis.**. Prentice Hall, N. Jersey.

[9] LEE, E. J. (2021): Biased self-estimation of maths competence and subsequent motivation and achievement: Differential effects for high- and low-achieving students. **Educational Psychology,** 41, 446–466.

[10] LI, Y. and G. R. ARCE (2004): A Maximum Likelihood Approach to Least Absolute Deviation Regression". **EURASIP Journal on Applied Signal Processing,** 12, 1762–1769.

[11] LIEM, G. A. D. (2021): Achievement and motivation, **Educational Psychology . An International Journal of Experimental Educational Psychology,** 4 , 379-382

[12] OSBORNE, M.R. (1987): The reduced gradient algorithm. In Y. Dodge (ed.) **Statistical data analysis based on the L1 norm and related methods,** 95-108. North-Holland. Amsterdam.

[13] PRATIWI L- E., N. LAILY , I. FITRI S., and I. F. SHOLICHAH. (2021): The relationship between self-efficacy and academic achievement , **Journal Universitas Muhammadiyah Gresik Engineering, Social Science, and Health International Conference (UMGESHIC).**

[14] REGENWETTEr L. and F. AHMED (2022): Design target achievement index: a differentiable metric to enhance deep generative models in multi-objective inverse design. **Proceedings of the ASME 2022 International Design Engineering Technical Conferences and Computers and Information in Engineering Conferences IDETC/CIE2022** August 14-17, 2022, St. Louis, MO, USA

[15] SHI, M. and M. A. LUKAS (2002): An L_1 estimation algorithm with degeneracy and linear constraints. **Computational Statistics & Data Analysis** 39, 35–55.

[16] SAS INSTITUTE (2004):. **SAS 9.1.3 Procedures Guide Volume 4.** SAS Institute, Cary, North Caroline.

[17] THANOON, F.T. (2015): **Robust Regression by Least Absolute Deviations Method**, International Journal of Statistics and Applications 5, 109-112

[18] WOLBERG, J. (2005). **Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments.** Springer, Berlin.