7. NUMERICAL AND MONTE CARLO METHODS TO MAKE NORMAL RESIDUES IN REGRESSION

Daniel CIUIU*

Abstract

The Jarque-Bera normality test verifies if the residues $u_i = Y_i - A_0 - \sum_{j=1}^{k} A_j X_j^{(i)}$ *of the regression hyper-plane* $Y = A_0 + \sum_{i=1}^{k} A_i X_i$ *are normal random variables.In this paper we present some numerical and Monte Carlo methods to obtain normal residues if the Jarque-Bera test fails. We consider the case when we know the pdf, the cdf and the inverse of the cdf for the random variable Y (example: the exponential distribution), the case when we know only the first two elements (example: Erlang distribution) and the case when we know only the pdf (example: the gamma distribution). We consider also the case when we do not know even an analytical formula for the pdf. In this case we will estimate the pdf using some known kernels (see section 2).*

Keywords: Jarque-Bera test, linear regression, kernels, Monte Carlo, credits

JEL Classification: C12, C13, E51

1. Introduction

Consider *n* points in $\mathbb{R}^{k+1}$, $X^{(1)},...,X^{(n)}$, where $X^{(i)} = \left(X_1^{(i)}, X_2^{(i)},..., X_k^{(i)}, Y_i\right)$. The regression hyper-plane used in (Ciuiu, 2007) to classify patterns has the equation (Saporta, 1990)

$$H\ :\ Y = A_0 + \sum_{i=1}^{k} A_i X_i \text{ such that} \tag{1}$$

$$\sum_{i=1}^{n} u_i^2 \text{ is minimum,} \tag{1'}$$

* Technical University of Civil Engineering, Bucharest. Bd. Lacul Tei, nr. 124, Bucharest, Romania. E-mail: dciuiu@yahoo.com.

where the residues u_i have the formula

$$u_i = Y_i - A_0 - \sum_{j=1}^{k} A_j X_j^{(i)} \tag{1”}$$

For the computation of A_i from (1) we have to solve the system (Saporta, 1990)

$$\sum_{j=0}^{k} \overline{X_i \cdot X_j} \cdot A_j = \overline{X_i \cdot Y},\ i = \overline{0,k}\ , \tag{2}$$

where $\overline{X_0 \cdot X_i} = \overline{X_i}$ and $\overline{X_0^2} = 1$.

For the obtained estimators of A_i using (2) and of the residues u_i we have the following hypotheses (Jula 2003, Voineagu *et al.,* 2007):

1. The estimators of A_i are linear.
2. The estimators of u_i have the expectation 0 and the same variance (homoscedasticity).
3. The estimators of u_i are normal.
4. The random variables u_i are independent.

From the above hypotheses and from Gauss-Markov theorem we obtain the following properties (Jula 2003, Voineagu *et al*., 2007):

1. The estimators of A_i are consistent.
2. The estimators of A_i are unbiased.
3. The estimators of A_i have the minimum variance.
4. The estimators of A_i have the maximum likelihood.

In (Jula 2003) there is presented a test for the normality of u_i. First we compute the skewness:

$$S = \frac{\overline{u^3}}{\left(\overline{u^2}\right)^{\frac{3}{2}}} \tag{3}$$

and the kurtosis:

$$K = \frac{\overline{u^4}}{\left(\overline{u^2}\right)^2}\ . \tag{4}$$

The Jarque-Bera statistics is

$$JB = n\left(\frac{S^2}{6} + \frac{(K-3)^2}{24}\right). \tag{5}$$

For the normal distribution of u_i we have $S = 0$, $K = 3$ and $JB = 0$. JB has in fact the distribution χ_2^2 (χ^2 with two degrees of freedom), hence we accept the null hypothesis $H_0 : S = 0$ and $K = 3$ with the error ε if and only if $JB < \chi_{2;\varepsilon}^2$.

2. Making normal residues

In this section we consider the case when the Jarque-Bera test fails. We have the following proposition (Ciuiu, 2009).

Proposition 1. If Y_i have a normal distribution then the residues u_i have a normal distribution.

Proof: In the system (2) we have used $\frac{X^T X}{n}$ and $\frac{X^T Y}{n}$ instead of $X^T X$ and $X^T Y$ as in (Jula 2003) and (Voineagu *et al*., 2007).
In this proof we will use the last notations and we obtain

$$u = \left(I_n - X\left(X^T X\right)^{-1} X^T\right)Y \ .$$

Therefore each u_i is a linear combination of Y_j, and the proposition is proved.

Remark 1. From the proof of the proposition 1 we can see that the reverse of the proposition is also true if $X\left(X^T X\right)^{-1} X^T$ has no eigenvalue equal to 1.

In (Jula, 2003) two methods to treat the case when the residues are not normal are presented. The first method consists in the identification of the distribution of u_i, and to apply a nonlinear regression. The second one is to transform Y_i and/or $X_j^{(i)}$ to eliminate the non-normality.

We will use the second method and the proposition 1: we transform Y_i to obtain normal variables. Suppose the pdf of Y_i is f, the cdf of Y_i is F and the inverse of the cdf is F^{-1}. Denote also by $\phi(x)$, $\Phi(x)$ and $\Phi^{-1}(x)$ the pdf and the cdf of the standard normal distribution, and respectively its inverse.
We make the substitution

$$Z_i = \Phi^{-1} \circ F(Y_i) \text{ ,`} \tag{6}$$

we compute the regression hyper-plane (1) with *Z* instead of *Y*, and we obtain the regression

$$Y = F^{-1} \circ \Phi\left(A_0 + \sum_{i=1}^{k} A_i X_i\right). \tag{7}$$

In the above formulae the main problem is to compute Φ and Φ^{-1}, for which we currently use the table of centils of the normal distribution. In (Ciuiu, 2009) we have presented some methods to compute the inverse of a cdf or the cdf. But these methods are only Monte Carlo ones and do not take into account the particularities of the considered distribution.

When we do not know an analytical formula for F^{-1} (as in the normal case $F = \Phi$) we estimate $F^{-1}(x)$ by numerical method as follows. If we know some x_0 and y_0 such that $F^{-1}(x_0) = y_0$ and the analytical formula for the pdf f we solve the Cauchy problem

$$\begin{cases} y' = \frac{1}{f(y)} \\ y(x_0) = y_0 \end{cases}. \tag{8}$$

For the normal or the Student distribution we take $x_0 = 0.5$ and $y_0 = 0$, but for instance in the case of the gamma distribution we can not take $x_0 = y_0 = 0$ because we can start with $f(y_0) = 0$ (if $\alpha > 1$). In this case we take some $y_0 \neq 0$ and $x_0 = F(y_0)$. The numerical method to solve the Cauchy problem (8) can be the Euler method, the modified Euler method or the Runge-Kutta method (Păltineanu *et al.*, 1998).

By Monte Carlo Methods we generate 10000 random variables $z_1, ..., z_{10000}$, we order these values in increasing order, and finally we take $F^{-1}(x) = z_{[10000 \cdot x]}$. For the random variables involving normal variables (normal, Student, chi square) we can generate the normal variables by the central limit method, by the Box-Muler method or by the Butcher 1 method. For those involving the exponential variables we generate the exponentials using the inverse method, the rejection method or the mixture method (Văduva, 2004). For the gamma distribution we use the rejection method or the mixture-rejection method.

When we know only an analytical formula for the pdf f we compute

$$F(x) = F(0) + \int_0^x f(t)dt \tag{9}$$

by the rectangles method, the trapezes method or by the Simpson method.

In the case of gamma distribution for computing $F(x)$ and $F^{-1}(x)$ we need to compute $\Gamma(\alpha)$ using the Gauss-Laguerre integration formula (Păltineanu *et al.*, 1998).

For Monte Carlo methods we generate 10000 random variables and $F(x) = F^*(x)$ (the empiric cdf). For the gamma distribution we estimate

$$F(x) = \frac{\int_0^{\frac{x}{\beta}} t^{\alpha-1} \cdot e^{-t} dt}{\Gamma(\alpha)}, \tag{10}$$

where the above integral and $\Gamma(\alpha)$ are computed by generating 100 random variables *Y* uniform on $\left[0, \frac{x}{\beta}\right]$, respectively $Y \sim \exp(1)$. The integral is estimated by the average of $\frac{x}{\beta} \cdot Y^{\alpha-1} \cdot e^{Y}$, and $\Gamma(\alpha)$ by the average of *Y* (Văduva, 2004).

When we do not know even the pdf f we can estimate this pdf using some kernels (Văduva, 1968, Văduva and Pascu, 2003). If we have the n-size sample $X_1, ..., X_n$ we estimate the pdf f in x by

$$f(x) = \frac{1}{n \cdot b_n} \cdot \sum_{i=1}^{n} K\left(\frac{x - X_i}{b_n}\right), \tag{11}$$

where K is the kernel and b_n is the window bandwidth.
There are some kernel functions $K = K_j$ used in literature:

$$\begin{cases} K_0(x) = \chi_{-\frac{1}{2},\frac{1}{2}}(x) = \begin{cases} 1 \text{ if } |x| \le \frac{1}{2} \\ 0 \text{ if } |x| > \frac{1}{2} \end{cases} \\ K_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\ K_2(x) = \frac{3}{4\sqrt{5}}\left(1 - \frac{x^2}{5}\right)\chi_{-\sqrt{5},\sqrt{5}}(x) \\ K_3(x) = (1 - |x|)\chi_{-1,1}(x) \\ K_4(x) = \frac{3}{4}\left(1 - x^2\right)\chi_{-1,1}(x) \\ K_5(x) = \frac{15}{16}\left(1 - x^2\right)^2 \chi_{-1,1}(x) \end{cases}, \tag{12}$$

where K_0 is the rectangular kernel, K_1 is the Gaussian kernel, K_2 is the Epanechnikoff kernel, K_3 is the triangular kernel, K_4 is the Bartlett-Priestley-Epanechnikoff kernel and K_5 is the biquadratic kernel.
The Epanechnikoff kernel K_2 minimizes the MISE (Min Integrated Square Error)

$$\mathrm{MISE}(\hat{h}) = \int_{-\infty}^{\infty} \mathrm{MSE}_x(\hat{h})dx, \tag{13}$$

where $\mathrm{MSE}_x(\hat{h})$ is the mean square error (Văduva 1968, Văduva and Pascu 2003):

$$\mathrm{MSE}_x(\hat{h}) = E\left\{(\hat{h}(x) - h(x))^2\right\}. \tag{13'}$$

The window bandwidth must be chosen such that $\lim_{n\to\infty} b_n = 0$ and $\lim_{n\to\infty} n \cdot b_n = \infty$ (Văduva, 1968, Văduva and Pascu, 2003). In our C++ program we take $b_n = \frac{1}{\sqrt{n}}$ and we obtain

$$f(x) = \frac{1}{\sqrt{n}} \cdot \sum_{i=1}^{n} K\left(\sqrt{n}(x - X_i)\right). \tag{11'}$$

Therefore for computing the cdf we use the formula

$$F(x) = \frac{1}{n}\sum_{i=1}^{n} \widetilde{K}\left(\sqrt{n}(x - X_i)\right), \tag{14}$$

where $\widetilde{K}$ is the cdf if the pdf is the kernel K. This formula is easy to use because, except the Gaussian kernel, we know the analytical formula for $\widetilde{K}$. In the case of

Gaussian kernel we compute $\widetilde{K}(x) = \Phi(x)$ by numerical or Monte Carlo methods as we have mentioned above.

The numerical method to compute the inverse of the cdf is the bisection method in the interval between the minimum of the values X_i and the maximum of them. If the kernel is defined on an interval (for instance $[-1,1]$ if the kernel is triangular) we subtract $\frac{1}{2\cdot\sqrt{n}}$ from the minimum value, and we add $\frac{1}{2\cdot\sqrt{n}}$ to the maximum one (*l* is the length of the interval). In the case of the Gaussian kernel we do the same operations tacking $l = 3$ (the $3 \cdot \sigma$ rule).

For the Monte Carlo method we use the same algorithm as for computing F^{-1} if we know f, and the random variables are generated using the mixture method as follows. First, we generate an integer uniform random variable $Y \in \{1,...,n\}$. Next, we generate a random variable *X* having the pdf *K*, and the desired generated value is $\frac{X}{\sqrt{n}} + X_Y$. For the Epanechnikoff, Bartlett-Priestley-Epanechnikoff and for the biquadratic kernel we generate the random variable having the pdf *K* by the inverse method (Văduva, 2004), the computation of $K^{-1}(U)$ being computed in the same way that it is computed $F^{-1}(x)$ in (Ciuiu, 2009): first we generate 1000 random variables uniform on $[-1,1]$. For each of the generated values *y* we take $F^{-1}(x) = \widetilde{y}$ such that $\Psi(y) - U \cdot y$ has its minimum in $\widetilde{y}$ ($\Psi(y)$ is a primitive of $F(y)$). For the Epanechnikoff kernel we multiply the result in the case of Bartlett-Priestley-Epanechnikoff kernel by $\sqrt{5}$.

Sometimes we have a shortcut for obtaining Z_i in (6). For instance, if the distribution of *Y* is log-normal we can do the substitution

$$Z_i = \ln Y_i, \tag{6'}$$

and (7) becomes

$$Y = \exp\left(A_0 + \sum_{i=1}^{k} A_i X_i\right). \tag{7'}$$

The centil of the χ^2 distribution with two degrees of freedom is computed by the formula $\chi^2_{2;\varepsilon} = -2\ln\varepsilon$, because the χ^2_2 distribution is identical to the $\exp\left(\frac{1}{2}\right)$ distribution (Ciuiu, 2005, Ciuiu, 2006).

3. Applications

Example 1. *Consider the same monthly series as in the previous examples, but the table is about the other goal credits. The results are in the following table (Statistical Section of the NBR Monthly Bulletin):*

X_1	358.7	381.6	428.5	448	497.4	519.4	571.1	613.3	661.1
X_2	323.7	304.1	307.8	327.4	423.5	402	478.8	549.2	604.4
Y	10	13.2	12.9	13.4	38.4	85.3	126.7	174.6	226.3

X_1	710	761.4	840.3	901	1006.6	1121.1	1186.1	1248.8	1309.3
X_2	655.6	767.4	901.3	980.5	1069.5	1231	1400.3	1571.5	1791.5
Y	281.7	373.4	464.8	544.7	638.2	745	806.3	860.5	968.1

X_1	1365.9	1402.3	1432.1	1437.8
X_2	1916.1	2129.4	2213.5	1821.9
Y	1000.1	1094.9	1270.7	1402

The regression plane is $Y = -445.22095 + 0.92707X_1 + 0.14184X_2$, the skewness is 2.08023 and the kurtosis 8.31283. Because the Jarque-Bera statistics is 41.7409 and the χ_2 centil of $0.05 = 5.99146$ we have $41.7409 > 5.99146$, hence we reject that the residues are normal.

Even we can compute also the integrals by the rectangles or by the trapezes method, in this example we will refer only to the Simpson method, because is the most precise (Păltineanu et al., 1998). For solving the Cauchy problem we refer only to Runge-Kuta method, by the same reasons. For the Monte Carlo method we will refer only to the Box-Muler method to generate the normal variables, because the method is the most rapid (Văduva, 2004).

First, we consider the distribution $\exp(\lambda)$. We estimate $\lambda = \frac{1}{\bar{X}} = 0.00197$. We are in the case when we know analytical formulae for F and F^{-1}.

If we generalize and we consider the distribution Erlang we estimate the parameters $n = 1$ and $\lambda = 0.00197$, hence it is the previous case. Therefore we obtain the same results, the only exception being the last value of Y when we give values to X_1 and X_2, because we compute the inverse of the Erlang distribution by numerical or Monte Carlo methods. The values obtained by the Monte Carlo Method are not the same, because of the randomness.

More general, if Y has a Gamma distribution with we estimate the parameters by the moments method and we obtain $\alpha = 1.32812$ and $\beta = 381.64604$. In fact from these values we estimate the Erlang parameters choosing $n = 1$ because it is the closest integer number to α.

The obtained regressions, the new skewness, the new kurtosis and the new Jarque-Bera statistics for the above distributions of *Y* and for the unknown distribution using the kernels presented in this paper are listed in table 1. The corresponding applications of the regressions from table 1 are listed in table 2.

Table 1

The new skewness, the new kurtosis and the new Jarque-Bera statistics

Distribution of *Y*, method and possible kernel	Regression	Skewness	Kurtosis	Jarque-Bera
$\exp(0.00197)$, numerical methods	$Y = F^{-1} \circ \Phi(-3.46804 + 0.00575X_1 - 0.00164X_2)$	–0.13519	1.84367	1.29268
$\exp(0.00197)$, Monte Carlo method	$Y = F^{-1} \circ \Phi(-3.49338 + 0.00578X_1 - 0.00165X_2)$	–0.12949	1.88158	1.2081
$E_1(0.00197)$, numerical methods	$Y = F^{-1} \circ \Phi(-3.46804 + 0.00575X_1 - 0.00164X_2)$	–0.13519	1.84367	1.29268
$E_1(0.00197)$, Monte Carlo method	$\mathrm{Y} = \mathrm{F}^{-1} \circ \mathit{\Phi}(-3.44675 + 0.00571\mathrm{X}_1 - 0.00163\mathrm{X}_2)$	–0.15207	1.8774	1.24
$\Gamma(1.32812,\ 381.64604)$, numerical methods	$Y = F^{-1} \circ \Phi(-4.09842 + 0.00664X_1 - 0.0019X_2)$	–0.13973	1.84982	1.28427
$\Gamma(1.32812,\ 381.64604)$, Monte Carlo method	$Y = F^{-1} \circ \Phi(-4.10605 + 0.00679X_1 - 0.00207X_2)$	–0.10649	1.90272	1.14527
Unknown, rectangular kernel, numerical methods	$Y = F^{-1} \circ \Phi(-2.49728 + 0.00368X_1 - 0.00071X_2)$	–0.45465	2.43823	1.04721
Unknown, rectangular kernel, Monte Carlo method	$Y = F^{-1} \circ \Phi(-2.48813 + 0.00369X_1 - 0.00073X_2)$	–0.59871	2.66449	1.41753
Unknown, Gaussian kernel, numerical methods	$Y = F^{-1} \circ \Phi(-2.39127 + 0.00342X_1 - 0.00058X_2)$	–0.32775	2.41367	0.70901
Unknown, Gaussian kernel, Monte Carlo method	$Y = F^{-1} \circ \Phi(-2.48644 + 0.00368X_1 - 0.00072X_2)$	–0.32886	2.43095	0.69339
Unknown, Epanechnikoff kernel, numerical methods	$Y = F^{-1} \circ \Phi(-2.48903 + 0.00366X_1 - 0.0007X_2)$	–0.4451	2.4966	0.95872
Unknown, Epanechnikoff kernel, Monte Carlo method	$Y = F^{-1} \circ \Phi(-2.46621 + 0.00362X_1 - 0.00069X_2)$	–0.48837	2.4334	1.16882
Unknown, triangular kernel, numerical methods	$Y = F^{-1} \circ \Phi(-2.49719 + 0.00368X_1 - 0.00071X_2)$	–0.45467	2.43848	1.04703

Distribution of *Y*, method and possible kernel	Regression	Skewness	Kurtosis	Jarque-Bera
Unknown, triangular kernel, Monte Carlo method	$Y = F^{-1} \circ \Phi\left(-2.47615 + 0.00362 X_1 - 0.00068 X_2\right)$	–0.42246	2.48799	0.89472
Unknown, Bartlett-Priestley-Epanechnikoff kernel, numerical methods	$Y = F^{-1} \circ \Phi\left(-2.49715 + 0.00368 X_1 - 0.00071 X_2\right)$	–0.45468	2.4386	1.04694
Unknown, Bartlett-Priestley-Epanechnikoff kernel, Monte Carlo method	$Y = F^{-1} \circ \Phi\left(-2.48045 + 0.00364 X_1 - 0.00069 X_2\right)$	–0.51071	2.45505	1.22856
Unknown, biquadratic kernel, numerical methods	$Y = F^{-1} \circ \Phi\left(-2.49726 + 0.00368 X_1 - 0.00071 X_2\right)$	–0.45465	2.43827	1.04718
Unknown, biquadratic kernel, Monte Carlo method	$Y = F^{-1} \circ \Phi\left(-2.46174 + 0.00360 X_1 - 0.00067 X_2\right)$	–0.4819	2.42036	1.15949

For all the above case we accept the normality because the new Jarque-Bera statistics is less than 5.99146. If we consider $X_1 = 1500.5$ and $X_2 = 1675.67$ we obtain the following results.

Table 2

Application of the above-mentioned regressions

Distribution of *Y*, method and possible kernel	Normal value of *Y*	Uniform value of *Y*	Initial value of *Y*
$\exp(0.00197)$, numerical methods	2.41305	0.99209	2453.09113
$\exp(0.00197)$, Monte Carlo method	2.42354	0.9926	2486.85713
$E_1(0.00197)$, numerical methods	2.41305	0.99209	2452.80966
$E_1(0.00197)$, Monte Carlo method	2.40228	0.9922	2460.17346
$\Gamma(1.32812, 381.64604)$, numerical methods	2 . 67312	0.99624	3452.53477
$\Gamma(1.32812, 381.64604)$, Monte Carlo method	2.61345	0.9959	2455.24388
Unknown, rectangular kernel, numerical methods	1.8362	0.96684	1401.95142
Unknown, rectangular kernel, Monte Carlo method	1.8362	0.96684	1401.95672
Unknown, Gaussian kernel, numerical methods	1.76291	0.96104	1401.83521
Unknown, Gaussian kernel, Monte Carlo method	1.90436	0.973	1401.94525
Unknown, Epanechnikoff kernel, numerical methods	1.83108	0.96646	1401.84323
Unknown, Epanechnikoff kernel, Monte Carlo method	1.8128	0.9651	1401.52327
Unknown, triangular kernel, numerical methods	1.83613	0.96683	1401.9434
Unknown, triangular kernel, Monte Carlo method	1.81451	0.9669	1401.93914
Unknown, Bartlett-Priestley-Epanechnikoff kernel, numerical methods	1.83611	0.96683	1401.93238
Unknown, Bartlett-Priestley-Epane-chnikoff kernel, Monte Carlo method	1.82547	0.9626	1401.7868
Unknown, biquadratic kernel, numerical methods	1.83619	0.96683	1401.94541
Unknown, biquadratic kernel, Monte Carlo method	1.81337	0.9648	1401.7868

If we consider the random variable Y lognormal we obtain the normal variables Y : $(2.30259, 2.58022, 2.55723, 2.59525, 3.64806, 4.44617, 4.84182, 5.16250, 5.42186, 5.64084, 5.92265, 6.14161, 6.30024, 6.45865, 6.61338, 6.69246, 6.75751, 6.87534, 6.90786, 6.99842, 7.14732, 7.24566)$, and finally the regression $Y = exp(0.4192 + 0.00968 X_1 - 0.00342 X_2)$. The new skewness is -0.26974, the new kurtosis is 1.79648, and the new Jarque-Bera statistics is 1.59454. Because we have $1.59454 < 5.99146$ we accept that the residues are normal. If $X_1 = 1500.5$, $X_2 = 1675.67$ we obtain the normal variable $Y = 9.21196$ and the initial variable $Y = 10016.21484$.

4. Conclusions

When the Jarque-Bera test fails we have to transform the random variable Y in a normal one using (6) and proposition 1. If we want to compute a new value of Y when we know new values for the explanatory variables we do the inverse transformation.

If we know analytical formulae for the pdf f, the cdf F and the inverse of the cdf F^{-1} we use the formula (6). If some of the above elements are not known we compute them by numerical or Monte Carlo methods, as in the second section.

The obtained normal variables Z are standard ones, except the lognormal case, when we obtain the normal variables using logarithm. An open problem is to use some other normal random variables (with other expectation and other variance).

When we go from exponential random variable to the Erlang one the estimation of the parameters is done using the moments' method to estimate the gamma distribution parameters, and we choose n being the closest integer to α. After we choose n we re-compute $\beta = \frac{\overline{X}}{n}$, and finally $\lambda = \frac{1}{\beta}$. Even we obtain in example 3 $n = 1$ the computation of F^{-1} is done by numerical or Monte Carlo methods. The differences between the regressions in the cases of the exponential and Erlang distributions (opposite the numerical methods, when we obtain the same regressions) are due to the fact that in this case we obtain as solution a random variable with the expectation being the true solution. Even in the case of the same distribution we obtain different (but closer) results in different moments of running the program.

We can see that when we go to the most general case of Gamma distribution we obtain $\alpha = 1.32812$ and $\beta = 381.64604$, close to the exponential (Erlang with $n = 1$) distribution. The differences are first from the approximations of $\alpha = 1 \neq 1.32812$ and $\beta = \frac{1}{\lambda} = \frac{1}{0.00197} = 507.61421 \neq 381.64604$. Secondly the differences come also from the numerical computation of $\Gamma(\alpha)$, respectively from the use of the formula (10) and from the simulations of Gamma random variables in the case of the final value of Y.

If we consider in example 3 being Student with bias we estimate the number of degrees of freedom $n = 2$ and bias 506.87273. This is in fact the expectation of Y. Because the k-moments of the Student distribution can be computed only for $k < n$, we have theoretical skewness and kurtosis only for $n \geq 5$. For these values of n we obtain the theoretical skewness 0 and the theoretical kurtosis $3 + \frac{6}{n-4}$ decreasing from 9 to 3. These values are quite different to those obtained in the example. Of course, we take into account the sum form of cumulants for independent random variables, and that $\frac{\sum_{i=1}^{n} a_i^t}{\left(\sum_{i=1}^{n} a_i\right)^t} \leq 1$ for $t = \frac{3}{2}$ or $t = 2$. When we run the program we obtain for instance by Monte Carlo method all the uniform random variables Y tacking the values 0 or 1 (or very close to these values), and the obtained regression coefficients have the very high absolute value. Therefore we obtain multicoliniarity. If we consider Y distributed χ_n^2 we obtain n = 507 and the regression coefficients of X_1 and X_2 being 0. This error comes from the fact that the expectation of Y is 506.87173 and the variance is $193445.96649 \neq 2 \cdot 506.87173 = 1013.74546$ (for χ^2 distribution the variance is twice the expectation, and $E(X) = n$). For the case of Snedecor-Fisher distribution we estimate the degrees of freedom by the moments' method, and we obtain $n_1 = 0$ and $n_2 = 2$, which is an error. This comes from the fact that the expectation of this distribution is equal to the variance of the Student distribution with n_2 degrees of freedom, and n_1 is computed using the ratio $\frac{E(X^2)}{E(X)}$.

The non-normality in example 3, opposite the other two examples, can be explained first by the high variance of the resulting variable Y: it starts from 10 (January 2007) to 1402 (October 2008). We can notice also that in the case of the household credits (example 2) the values of Y are generally small in function of the credits amounts in Euro (X_2). This can be explained by the use of prices in Euro in household market, and it can also be an explanation of the normality of residues.

The same situation is in example 1 with X_1 (credits in lei). In the last example we have another situation: in January 2007 X_1 and X_2 had comparable values and the resulting variable Y was small in function of the other two, and, even the explanatory variables remain comparable until the end of the period, Y becomes also in October 2008 comparable to X_1 and X_2. This can also explain the non-normality of Y.

References

Ciuiu, D., (2005), "Une modalité d'eviter les centiles aux regions de confiance et aux tests statistiques", *Analele Universității din Craiova*, Vol. 32, No. 1, 2005, pp. 98-105.

Ciuiu, D., (2006), "Statistical Inference for Queuing Systems", *Buletinul Ştiinţific al UTCB*, No 4, 2006, pp. 2-8.

Ciuiu, D., (2007), "Pattern Classification using Polynomial and Linear Regression", *Proceedings of the International Conference Trends and Challenges in Applied Mathematics, 20-23 June 2007*, Technical University of Civil Engineering, Bucharest, Romania, pp. 153-156.

Ciuiu, D., (2009), "On the Jarque-Bera normality test", *Working Papers of Macroeconomic Modeling Seminar*, No. 081802, 2008, pp. 16-24.

Jula, D., (2003), *Introducere în Econometrie*, Editura Professional Consulting, Bucureşti.

Păltineanu, G., Matei, P. and Trandafir, R., (1998), *Analiză numerică*, Editura Conspress, Bucureşti.

Saporta, G., (1990), *Probabilités, Analyse des Donées et Statistique*, Editions Technip, Paris.

Saporta, G. and Ben Ammou S., (1998), "Sur la normalité asymptotique des valeurs propres en ACM sous l'hypothèse d'indépendence des variables", *Revue de Statistique Appliquée*, Vol. 46, No. 3, pp. 21-35.

Văduva, I. (1968), "Contribuții la teoria estimațiilor statistice ale densităților de repartiție şi aplicații", *Studii şi Cercetări Matematice*, No. 8, pp. 1027-1076.

Văduva, I. and Pascu M. (2003), "Nonparametric Estimate of the Hazard Rate: A Survey", *Revue Roumaine Math. Pures Appl.*, Vol. 58, No. 2, pp. 173-191.

Văduva, I. (2004), *Modele de simulare*, Editura Universității Bucureşti.

Voineagu, V. *et al*, (2007), *Teorie şi practică econometrică*, Meteor Press, Bucureşti.

Secțiunea Statistică a Buletinului Lunar BNR, decembrie 2007, www.bnr.ro

Secțiunea Statistică a Buletinului Lunar BNR, octombrie 2008, www.bnr.ro