ACTA UNIVERSITATIS AGRICULTURAE ET SILVICULTURAE MENDELIANAE BRUNENSIS
SBORNÍK MENDELOVY ZEMĚDĚLSKÉ A LESNICKÉ UNIVERZITY V BRNĚ

Volume LVII 15 Number 3, 2009

ON NORMALITY ASSUMPTIONS FOR CLAIMS IN INSURANCE

M. Stehlík, L. Střelec

Received: December 15, 2008

Abstract

STEHLÍK, M., STŘELEC, L.: *On normality assumptions for claims in insurance.* Acta univ. agric. et silvic. Mendel. Brun., 2009, LVII, No. 3, pp. 141–146

The aim of this paper is to discuss effects of deviations from hypothetized normality. Two models are considered, one is the first pension pillar (and we consider here very small samples, which plays some role at start of some pension system or at early phases of it) and second one of modeling for IBNR (here we consider mid-samples). We will show that at early phases of 1st pension pillar in Slovakia the estimation of upper probability of oversizing of critical constant given by Potocký and Stehlík, 2005, fits well. For the case of IBNR reserves, the date given by Stelljes, 2006, are significantly more skewed and thus further research is needed for appropriate modelling of these reserves.

IBNR, pension pillar, classical tests of normality, robust tests of normality, modified tests of normality

Normality is one of the most common assumptions made in the development and use of statistical procedures. For example, in the majority of cases of relevant analysis is expected that returns derived from financial time series is Gaussian normal distributed random variable with constant expected value and constant variance. The problem has not suffered from lack of attention. There exists a vast amount of literature on tests for normality and their statistical properties, e.g. Anderson and Darling, 1952, Shapiro and Wilk, 1965, etc. Some robust modifications have been given by Gel and Gastwirth, 2008, and Brys et al. 2004. Střelec and Stehlík, 2008, and Stehlík and Střelec, 2009, provided a thorough study of general version of robust tests for normality.

The aim of this paper is to illustrate the impact of using an appropriate testing procedure for small (e.g. less then 10 observations) and mid-samples (around 40 observations). We illustrate this impact at a pay-as-you-go pillar studied by Potocký and Stehlík, 2005, and Potocký and Stehlík, 2007. Therein the probability of oversizing the limiting value of pillar is studied under normality, see Potocký and Stehlík, 2007. Potocký and Stehlík, 2007, considered the Cramer-Lundberg model in the case of a homogeneous portfolio with the attention focused on ruin probability for it under both light or heavy tails. They illustrated such situation in the setup of oversizing of the limiting value of the fund for the pay-as-yo-go pillar in Slovakia, see also Potocký and Stehlík, 2005. Potocký, 2007, gives further consequences for insurance. The second illustrating example is modelling of incurred but not reported (IBNR) losses. We continue in research in sequence of papers Stelljes, 2006, Adams, 2007, and Potocký and Stehlík, 2008.

MATERIALS AND METHODS

The topic of this paper is the problem of testing normality, i.e. problem of testing whether a sample of observations comes from a normal distribution. For normality testing we used these classical, robust and modified normality tests: the Anderson-Darling test (AD), the Cramer-von Mises test (CM), the D`Agostino test (DT), the Jarque-Bera test (JB), the Jarque-Bera-Urzua test (JBU), the robust Jarque-Bera test (RJB), the Lilliefors (Kolmogorov-Smirnov) test (L(KS)), the Pearson chi-square test (PT), the directed SJ test (dirSJ), the Shapiro-Wilk test (SW), the Shapiro-Francia test (SF), the standardized Geary test (GT), the standardized Uthoff test (UT), the directed Bonett-Seier test (dirBS), the skewness test (SKT), the kurtosis test (KT), the medcouple test (MC) and the modified JB, JBU, RJB and RJBU tests (RT tests).

For instance $RT_{JB(U)}$ denotes the test where in all places of $\hat{\mu}_k$ of the Jarque-Bera (or the Jarque-Bera-Urzua) test statistic JB (JBU) the median is used instead of arithmetic mean. Similarly $RT_{RJB(U)}$ denotes the test where in all places of $\hat{\mu}_k$ of the robust Jarque-

Bera (or the robust Jarque-Bera-Urzua) test statistic RJB (RJBU) the median is used instead of arithmetic mean. RT_{JB}, RT_{JBU}, RT_{RJB} and RT_{RJBU} are special cases of general version of RT test, see Stehlík and Střelec, 2009.

RESULTS

Assumption of normality by 1st pension pillar

The problem that assets of a fund are not sufficient to cover its liabilities is of extreme importance. Such a situation may arise in some countries in connection with the so-called non-funded 1st pension pillar based on pay-as-you-go principle.

Here we consider the illustrative example of claims for the mandatory, non-funded 1st (pay-as-you-go) pillar given by Potocký and Stehlík, 2005. Therein is considered a closed group of Slovakian people, all aged 50 in the year 1998, and interest is in the estimation of the total claim amount for this group in the year 2010 when the members are supposed to retire. Table I contains the salaries given from Statistical Yearbook, 2004, Labour Market, III.3-10, Structure of average gross nominal monthly wage of employees in the economy of the SR. Potocký and Stehlík, 2005, they are interested in estimation of the probabilities $P(\Sigma_{k=1}^{N} X_k > C)$ where X_i are individual monthly claims of the members of the abovementioned group and C is a critical (limiting) value of the fund representing the amount the fund has gathered from the contributions of the active members or from other sources. It is possible to consider N as a constant or a random variable as it was treated in Potocký and Stehlík, 2005. In Potocký and Stehlík, 2007, the case that N is a random variable was considered. Then it is quite natural to choose a binomial model for N namely $N \sim bi(n, p)$ with $n = 130\,000$ and p representing the probability of surviving a 50-year person from the group to the age 62 years (such probabilities are regularly published by Slovak Statistical Office). Then one is looking for the largest C such that $P(\Sigma_{k=1}^{N} X_k > C) = p$ with p given in advance, e.g. 0.1 or 0.05.

I: *Salaries, 1998–2002*

year	salary
1998	24 233
1999	26 862
2000	30 021
2001	31 825
2002	34 041

Typically it is possible to model salaries as normal variables in short-terms and lognormal at long-terms. In Potocký and Stehlík, 2005, we have used the normal distribution which led to the following upper bound

$$\bar{p} = 1 - \Phi\left(\frac{C/(k_N t) - \mu}{\sigma}\right). \quad (1)$$

Here Φ is cdf of standardized normal distribution, C is a critical level as given above, μ and σ^2 are parameters of normal distribution of salaries, $\frac{P_t}{S_t}$ and N_t is the number of claims. In the case of Table I we have $\hat{\mu} = 29396.4$ and $\hat{\sigma} = 3903.5$.

As we can see at Table II, normality is not rejected by any of the given tests (classical, modified or robust). This has a conclusion, that at the early stage of the pension first pillar the approximation (1) derived in Potocký and Stehlík, 2005, fits very well.

II: *Test statistics and p-values for normality under the data from Table I*

test	statistic	p-value$_{emp}$
JB	0.3880	0.7380
JBU	0.4851	0.8293
LT	0.1636	0.9443
PT	0.6000	0.4651
RJB	0.2370	0.7807
SJdir	0.4558	0.7753
SW	0.9771	0.9117
SF	0.9865	0.9610
GTst	0.8854	0.7429
UTst	0.5086	0.4494
BSdir	–0.9944	0.6286
SKTst	–0.1681	0.7768
KTst	–0.5998	0.6507
MC	0.1300	0.4798
RT_{JB}	0.3911	0.6883
RT_{JBU}	0.4772	0.8498
RT_{RJB}	0.2746	0.7578
RT_{RJBU}	0.6086	0.8570

Now let us consider adjusted data from (SLOVSTAT on-line), see Table III.

III: *Salaries, 2000–2006, adjusted dataset*

year	salary
2000	29 737
2001	31 060
2002	34 262
2003	35 533
2004	34 490
2005	28 174
2006	30 077

It can be seen from Table IV one can find the tests which are close to rejection of normality at some size. The interpretation can be that for larger samples of wages normality is violated and we can consider light-tailed claims as given in Potocký and Stehlík, 2007.

IV: *Test statistics and p-values for normality under the data from Table III*

test	statistic	p-value$_{emp}$
JB	0.7139	0.3748
JBU	1.5048	0.3658
LT	0.2261	0.3511
PT	3.7143	0.0897
RJB	0.4681	0.4756
SJdir	–0.2972	0.9565
SW	0.9105	0.4010
SF	0.9286	0.4917
GTst	1.7015	0.0972
UTst	1.1276	0.0869
BSdir	–1.7784	0.9514
SKTst	0.0799	0.9039
KTst	–0.8411	0.1208
MC	1.0319	0.2452
RT$_{JB}$	0.7661	0.3353
RT$_{JBU}$	1.2624	0.5181
RT$_{RJB}$	0.5395	0.5094
RT$_{RJBU}$	1.2006	0.6012

Assumption of normality by IBNR reserves

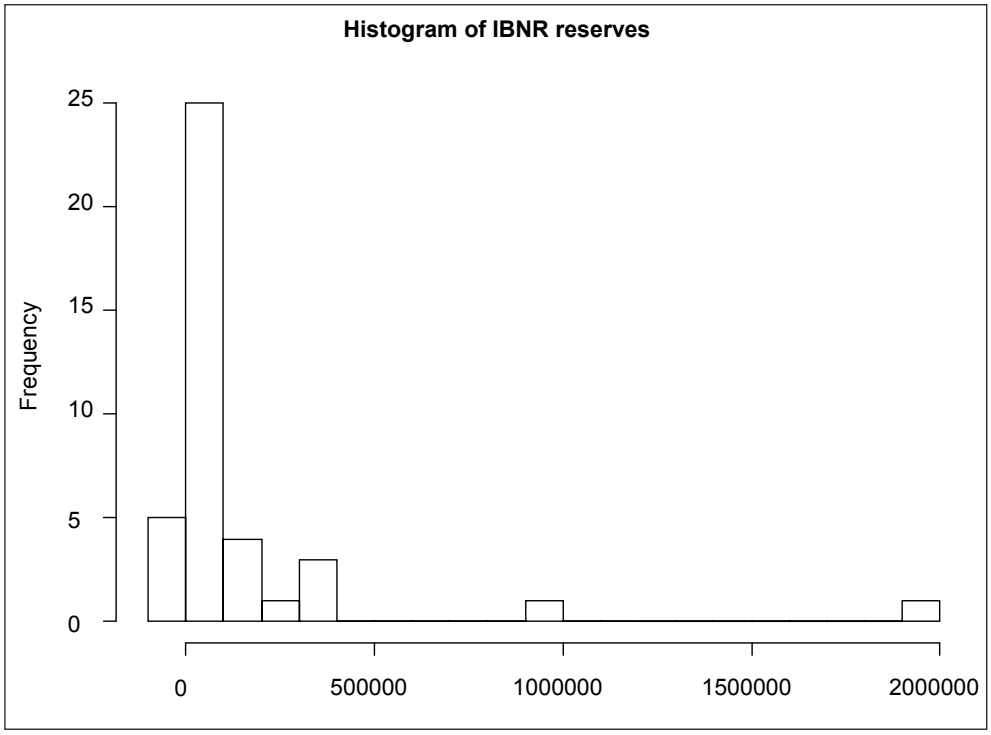
IBNR are insured losses that have occurred but have not been reported to a Primary Insurance company. These types of claims have a tremendous effect on a Reinsurance treaty, which may be showing a healthy profit when in reality it is losing money. Also, the reinsurer must establish an adequate reserve for IBNR claims to make a correct analysis of its business. As it is noted by Adams, 2007, and also by Potocký and Stehlík, 2008, one should take care about the fact that traditional nonlinear regression assumes the error terms to be normal which is a symmetric distribution with a range of whole real line. In Potocký and Stehlík, 2008, more accurate results for confidence intervals of parameters of IBNR nonlinear model are given. Therein a cubic approximation and usage of calculus given by Potocký and Van Ban, 1993, is employed. Incremental pure premium data may actually be skewed and can hardly ever be highly negative, therefore, using the normal distribution is approximation at best. The aim of this illustration is to show how far the distribution of IBNR claims from normality.

As we can see from histograms and QQ plots for Stelljes, 2006 (see Fig. 1, Fig. 2 and Fig. 3), normality is very nonsignificant. We have conducted all tests from previous section and we have found that normality has been strictly rejected with $p = 0$ except of MC test where $p = 0.008$.

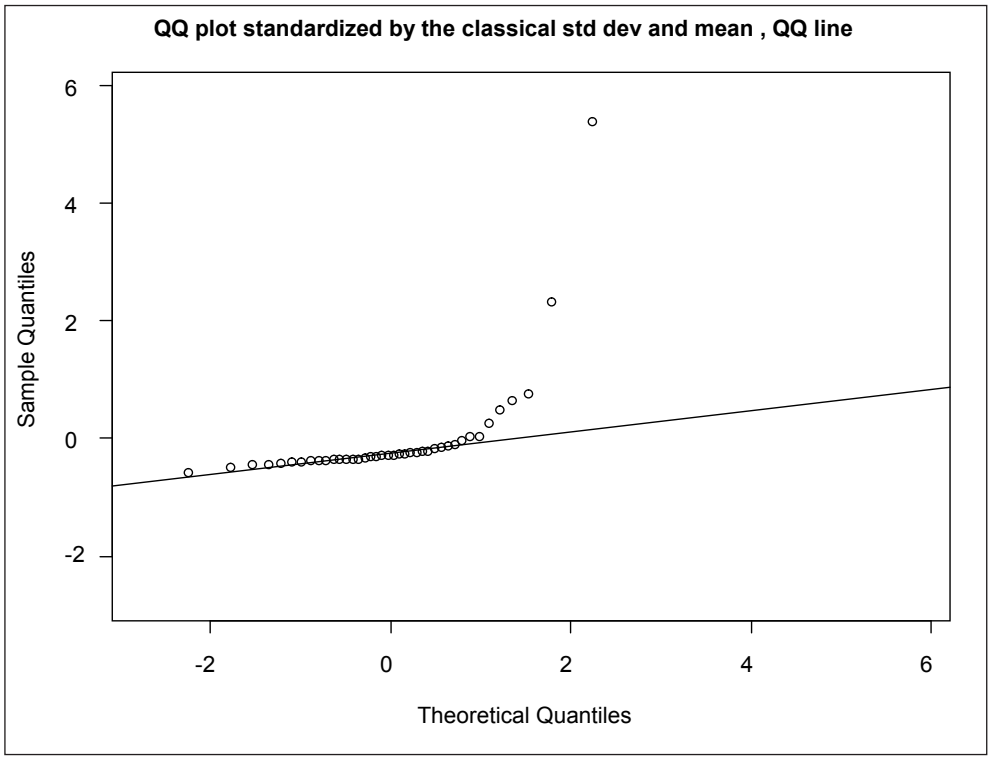
However, still approaches based on normality assumption can be a good approximation for the parameters of IBNR nonlinear model, etc. This is however the open problem for the next research.

DISCUSSION AND SUMMARY

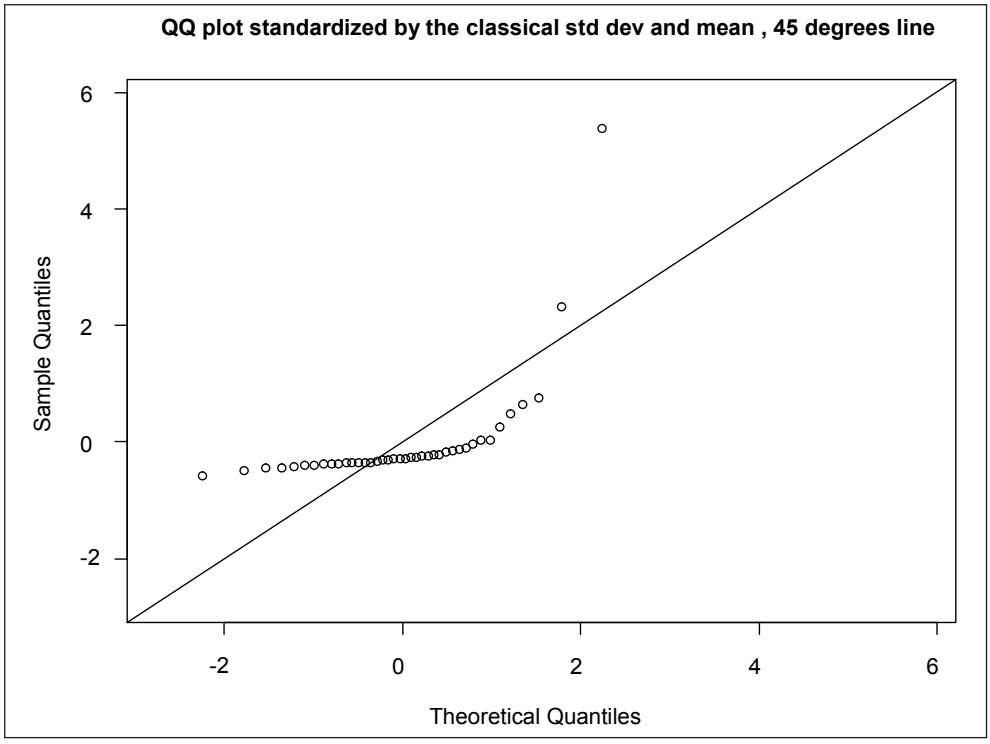
As we can see from Table II, normality is not rejected by any of the given tests. This has a conclusion, that at the early stage of the pension first pillar the approximation (1) derived in Potocký and Stehlík, 2005, fits very well. As it is noted by Adams, 2007, and also by Potocký and Stehlík, 2008, one should take care about the fact that traditional nonlinear regression assumes the error terms to be normal which is a symmetric distribution with a range of whole real line. Incremental pure premium data may actually be skewed and can hardly ever be highly negative, therefore, using the normal distribution is approximation at best. The normality of IBNR claim given by Stelljes, 2006, has been rejected by all considered tests. However, still approaches based on normality assumption can play as a good approximation for the parameters of IBNR nonlinear model.



1: *Histogram of IBNR reserves*



2: *QQ plot of IBNR reserves*



3: *QQ plot of IBNR reserves*

SOUHRN

K předpokladům normality pojistných událostí

Cílem tohoto článku je rozebrat efekty odchylek datového souboru od hypotetické normality. Uvažujeme dva modely, nejprve první penzijní pilíř pro velmi malý počet pozorování, který může být užitečný v případě startu penzijního systému nebo v jeho počátečních fázích. Druhým modelem je modelování IBNR rezerv pro středně velký počet pozorování.

K testování normality je využito následujících klasických, robustních a modifikovaných testů normality: Anderson-Darlingova testu (AD), Cramer-von Misessova testu (CM), D`Agostinova testu (DT), Jarque-Bera testu (JB), Jarque-Bera-Urzua testu (JBU), robustního Jarque-Bera testu (RJB), Lilliefors (Kolmogorov-Smirnov) testu (L(KS)), Pearsonova chí-kvadrát testu (PT), directed SJ testu (dirSJ), Shapiro-Wilkova testu (SW), Shapiro-Francia testu (SF), Gearyho testu (GT), Uthoffova testu (UT), directed Bonett-Seierova testu (dirBS), testu šikmosti (SKT), testu špičatosti (KT), medcouple testu (MC) a modifikovaných JB, JBU, RJB a RJBU testu (RT testy).

Například $RT_{JB(U)}$ označuje test, kde je ve všech případech Jarque-Bera testové statistiky JB, případně Jarque-Bera-Urzua testové statistiky JBU, využito mediánu místo aritmetického průměru. Obdobně $RT_{RJB(U)}$ značí test, kde je ve všech případech robustní Jarque-Bera testové statistiky RJB, případně robustní Jarque-Bera-Urzua testové statistiky RJB, využito mediánu místo aritmetického průměru. RT_{JB}, RT_{JBU}, RT_{RJB} a RT_{RJBU} jsou pak speciálními případy obecné verze RT testu (blíže viz Stehlík a Střelec, 2009).

Jak je z článku patrné, tak v případě prvního penzijního pilíře se nedá normalita zamítnout, což umožňuje použít vzorec pro pravděpodobnost přesáhnutí kritické hranice z článku Potocký a Stehlík, 2005. Naopak v případě IBNR rezerv se nedá předpoklad normality přijmout, jelikož rozdělení IBNR rezerv je signifikantně zešikmené, což otevírá potřebu dalšího výzkumu použitelnosti normálních aproximací pro modelování parametrů nelineární regrese s odezvou IBNR.

IBNR, penzijní pilíř, klasické testy normality, robustní testy normality, modifikované testy normality

Research was supported by Project 51p7 "AKTION Czech Republic-Austria": Robust testing for the normality and its applications for the economy, sports and Basel II.

REFERENCES

ADAMS, J. H., 2007: A Nonlinear Regression Model of Incurred But Not Reported Losses, discussion. *CASE Forum Summer 2007*. 1–14.

ANDERSON, T. W., DARLING, D. A., 1952: Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23. 193–212.

BRYS, G., HUBERT, M., STRUYF, A., 2004: A Robustification of the Jarque-Bera Test of Normality. *COMPSTAT 2004, Proceedings in Computational Statistics, ed. J. Antoch, Springer, Physica Verlag*. 753–760.

GEL, Y. R., GASTWIRTH, J. L., 2008: A robust modification of the Jarque-Bera test of normality. *Economics Letters*, 99. 30–32.

POTOCKÝ, R., 2007: Modelling of the claims for non-life insurance. *Forum Statisticum Slovacum*, 3 (4). 122–126.

POTOCKÝ R., STEHLÍK M., 2005: Analysis of pensions in the 1st pilar under the expected demographic development. *Proceedings of 10th Slovak Conference on Demography, Smolenice*.

POTOCKÝ R., STEHLÍK M., 2007: Stochastic models in insurance and finance with respect to Basel II. *JAMSI*, 3 (2). 237–245.

POTOCKÝ R., STEHLÍK M., 2008: Improved nonlinear regression modeling of parameters of IBNR reserves. *Communications in Dependability and Quality Management*. 11(2008)4. 54–60.

POTOCKÝ R., VAN BAN, T., 1993: On parameter-effects arrays in non-linear regression models. *Appl. Math*. 38. 123–132.

SHAPIRO, S. S., WILK, M. B., 1965: An analysis of variance test for normality (complete samples). *Biometrika*. 52 (3 and 4). 591–611.

SLOVSTAT online. Statistical database of indicators of economic and social-economic development in the Slovak Republic.

STATISTICAL YEARBOOK OF SLOVAKIA, 1998–2003.

STELLJES, S., 2006: Nonlinear regression model of incurred but not reported losses. *Casualty Actuarial Society Forum Casualty Actuarial Society – Arlington*, Virginia. 353–377.

STŘELEC, L., STEHLÍK, M., 2008: Some properties of robust tests of normality and their applications. *Firma a konkurenční prostředí 2008 – 1. část*. Brno: MSD, spol. s r. o., ISBN 978-80-7392-020-3. (In Czech)

STEHLÍK, M., STŘELEC, L., 2009: On robust testing for normality. Accepted to *Proceedings of the 6th St. Petersburg Workshop on Simulation*.

Address

RNDr. Milan Stehlík, Ph.D., Ústav aplikované statistiky, JKU Linz, Freistädter Straße 315, A-4040 Linz a. D., Rakousko, e-mail: Milan.Stehlik@jku.at, Ing. Luboš Střelec, Ústav statistiky a operačního výzkumu, Mendelova zemědělská a lesnická univerzita v Brně, Zemědělská 1, 613 00 Brno, Česká republika, e-mail: xstrelec@node.mendelu.cz