Polskie Towarzystwo Statystyczne
Oddział we Wrocławiu

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Silesian Statistical Review

Nr 8 (14)

Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2010

Redaktor Wydawnictwa

Joanna Szynal

Redaktor techniczny

Barbara Łopusiewicz

Korektor

Barbara Cibis

Skład i łamanie

Janusz Stanisławski

Projekt okładki

Beata Dębska

ADRES REDAKCJI

Katedra Statystyki
Uniwersytetu Ekonomicznego we Wrocławiu
ul. Komandorska 118/120, 53-345 Wrocław
tel. (71) 36-80-356, tel./fax (71) 36-80-357
e-mail: stanislaw.heilpern@ue.wroc.pl

Wrocław 2010

ISSN 1644-6739

Spis treści

Summaries

20. SCIENTIFIC STATISTICAL SEMINAR "MARBURG/KÖLN – WROCŁAW" WISŁA, SEPTEMBER 21-25, 2009

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 8(14)

EXTENDED SUMARIES OF THE PAPER

ISSN 1644-6739

SOME ASYMPTOTICS FOR THE DELAY TIME OF MOSUM CHANGE DETECTION PROCEDURES

Josef G. Steinebach (University of Cologne)

1. The model

In [Horváth et al. 2008] we discuss some "open-end" and "closed--end" monitoring procedures for detecting a "change in the mean" in the following location model:

$$X_i = \mu_i + \varepsilon_i, \quad i = 1, 2, ...,$$

where $\{\mu_i\}_{i=1,2}$ are unknown means and $\{\varepsilon_i\}_{i=1,2}$ are the unobservable, centered errors. It is assumed that there is no change in the mean of a "training sample" of size m, i.e., that $\mu_i = \mu,\ i = 1, ..., m.$ We are interested in constructing appropriate stopping rules for testing the null hypothesis:

$$H_0 : \mu_0 = \mu, \quad i = m, m+2, ...,$$

against the (two-sided) alternative

$$H_1 : \text{there is a } k_n^* \geq 1 \text{ such that } \mu_i = \mu,\ m < i < m + k^*,$$
$$\text{but } \mu_i = \mu + \Delta,\ i \geq m + k^*,\ \text{with some } \Delta \neq 0.$$

2. Stopping rules

Our rules for testing H_0 versus H_1 are based on "moving sum detectors" (MOSUM's), more precisely, on comparing:

ŚLĄSKI PRZEGLĄD STATYSTYCZNY Nr 8(14)

$$\overline{X}_m = \frac{1}{m}\sum_{i=1}^{m} X_i \quad \text{and} \quad \overline{X}_{m,k} = \frac{1}{h}\sum_{i=0}^{h-1} X_{m+k-i} \quad k = 1, 2, \ldots, \quad k = 1, 2, \ldots,$$

where $h = h(m)$ ($\leq m$) is a window size to be determined. For example, we study the (closed-end) stopping rule μ_i

$$\tau_{m,T} = \min\left\{k:\ 1 \leq k \leq mT,\ \frac{1}{\hat{\gamma}} |\ \overline{X}_{m,k} - \overline{X}_m\ | > ch^{-0.5} g\left(\frac{k}{m}\right)\right\}, \qquad (1)$$

where $\min \varnothing = +\infty$, $\hat{\gamma}^2$ is a certain variance estimator, and g is a weight function.

In Section 3 it is shown that the constant c in (1) can be chosen such that, under H_0, we have $\lim_{m\to\infty} P\{\tau_{m,T} < \infty\} = \alpha$, i.e. that the false alarm rate (asymptotically) equals α, where $0 < \alpha < 1$ is a prescribed level. In addition, some limiting distributions under H_1 are discussed in Section 4. Interestingly, it turns out that, under H_0, the asymptotics crucially depend on the relation between h and m, and, under H_1, also on the limits of k^*/h and k^*/m, respectively.

3. Null asymptotics

To obtain the null asymptotics, we assume that the errors $\{\varepsilon_i\}_{i=1,2}$ satisfy a functional central limit theorem (with asymptotic variance γ^2). Then, for example, if $\lim_{m\to\infty} h/m = b \in (0,1]$, we have

$$\lim_{m\to\infty} P\{\tau_{m,T} < \infty\} = P\left\{\sup_{0\leq t\leq T/b} \frac{1}{g(t)}\left|W\left(\frac{1}{b}+t\right) - W\left(\frac{1}{b}+t-1\right) - bW\left(\frac{1}{b}\right)\right| > c\right\}, \qquad (2)$$

where $\{W(t),\ t \geq 0\}$ denotes a standard Wiener process, i.e. the critical value in (1) can be determined via the weighted Gaussian process from (2). Similar results apply in case of $\lim_{m\to\infty} h/m = 0$, but require a more careful discussion (see [Horvath et al. 2008]).

4. Asymptotics under the alternative

For the limiting results under the alternative, it is assumed that the errors $\{\varepsilon_i\}_{i=1,2}$ satisfy a Hungarian (KMT) type of strong approxima-

tion. Various cases and stopping rules can be discussed then, depending on the orders of the ratios h/m, k^*/h and k^*/m, respectively. For example, if $h/m = b \in (0,1]$, $k^*/h \to a \geq 0$ and $T > ab$, then, for $y > 0$,

$$\lim_{m\to\infty} P\left\{\tau_{m,T} > k^* + \frac{\sqrt{h}}{|\Delta|} y \,\middle|\, \tau_{m,T} \geq k^*\right\} = P\left\{\sup_{0\leq t\leq y}\left|Z_1(a) + \frac{t}{\gamma g(a)}\right| \leq c \,\middle|\, \sup_{0\leq t\leq y}|Z_1(t)| \leq c\right\},$$

where $\{Z_1(t), t \geq 0\}$ is the weighted Gaussian process from (2) and c is the critical value therein. For other relations between h, k^* and m a number of similar asymptotics are available (cf. [Aue et al. 2009, p. 31]).

5. Some simulation results

In [Aue et al. 2009] we also present some simulation results concerning the behaviour of the suggested stopping rules under various alternatives and distributions (see Section 4 of [Aue et al. 2009, p. 31] for further details). In the table below we provide just one example showing empirical power values from simulation runs with:

- normal (0,1) errors ε_t,
- 2,500 repetitions,
- a training period of size $m = 100$,
- an observation period of size $Tm = 10m$,
- and a window of size $h = 0.1m$ for the MOSUM procedures.

We compare two MOSUM procedures:

- τ_1^{MS} based on the boundary function $g_1(t) = t_+^{1/v}$ (with $v = 10$),
- τ_2^{MS} based on the boundary function $g_2(t) = \sqrt{\log_+(1+t)}$,

and a CUSUM stopping rule τ^{CS}, which is known to perform quite well in case of "early changes" (cf., e.g., [Horváth et al. 2004]).

Across various change scenarios, e.g.

- $k^* = 0.1m$ ("early"), $4m$ ("intermediate"), $8m$ ("late"),

the table contains percentages:

- "fd" of changes "falsely detected" and
- "cd" of changes "correctly detected".

The critical values were chosen such that the

- nominal level equals $\alpha = 10\%$.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 8(14)

Table 1. The results of simulations

	τ_1^{MS}		τ_2^{MS}		τ^{CS}	
Δ	fd	cd	fd	cd	fd	cd
The "early change" scenario: $k^* = 0.1m$						
$m = 100$						
0.8	0.0000	1.0000	0.0004	0.9996	0.0116	0.9884
0.6	0.0004	0.9944	0.0004	0.9896	0.0168	0.9824
0.4	0.0000	0.8672	0.0004	0.8200	0.0140	0.9084
0.2	0.0000	0.3780	0.0008	0.3164	0.0160	0.3952
The "intermediate change" scenario: $k^* = 4m$						
$m = 100$						
0.8	0.0740	0.9260	0.0896	0.9044	0.0800	0.9068
0.6	0.0720	0.9072	0.0944	0.7488	0.0728	0.7600
0.4	0.0724	0.6416	0.0880	0.3064	0.0776	0.3768
0.2	0.0596	0.1568	0.0912	0.0248	0.0632	0.0736
The "late change" scenario: $k^* = 8m$						
$m = 100$						
0.8	0.0828	0.9108	0.0924	0.8148	0.0824	0.1236
0.6	0.0816	0.7664	0.0888	0.4332	0.0964	0.0636
0.4	0.0876	0.3128	0.1020	0.0884	0.0764	0.0260
0.2	0.0892	0.0416	0.0920	0.0020	0.0844	0.0044

Source: own calculations.

References

Aue A., Horváth L., Kühn M., Steinebach J., *On the reaction time of moving sum detectors*, Preprint, University of California, Davis, University of Utah, Salt Lake City, and University of Cologne 2009.

Horváth L., Hušková M., Kokoszka P., Steinebach J., *Monitoring changes in linear models*, "Journal of Statistical Planning and Inference", 126 (2004), pp. 225-251.

Horváth L., Kühn M., Steinebach J., *On the performance of the fluctuation test for structural change*,"Sequential Analysis", 27 (2008), pp. 126-140.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY
Nr 8(14)

USING STATISTICAL MODELS FOR SOCIAL STRESS ANALYSIS

Luisa Canal (Trient University)
Walenty Ostasiewicz (Wrocław University of Economics)

The paper addresses the critical review of statistical models that could be used in the social stress analysis. Such an analysis consists in the identification of the social stressors, and in the measurement of their potency to destroy social harmony. Four main groups of methods are discussed: item response models, factorial models, latent classification, and paired comparison.

Social stressor is defined here as any phenomenon, event, or condition which has a destructive impact on social life. For example legalized political corruption, cynicism of politicians, brutality in TV movies, immoral behaviour of higher officials, etc.

To discuss the problem, we assume the existence of some *common sense* or common feature characterizing a whole group of people. This characteristics which is not observed directly, will be denoted by symbol Z. It is assumed that it "drives", commands, or controls people's reaction to stressful phenomena. For the lack of established terminology, a latent variable Z will be called *susceptibility*, endurance, resistance or patience. To keep the discussion general enough, we admit a number of aspects of the susceptibility. Therefore, trait Z is considered as a d-dimensional variable $Z = (Z_1, Z_2, ..., Z_d)$.

As different people are endowed with different amounts of susceptibility, we will interpret trait Z as a random variable. The cumulative distribution of it is denoted by $H(z) = H(z_1, z_2, ..., z_d)$. All stressful phenomena will be denoted by symbols $Y_1, Y_2, ..., Y_p$.

The measurement of the strength of a stressor can be done by "observing" people's reaction. By a reaction we mean an answer to a question concerning undesired phenomena. Two kinds of questions and two broad approaches to the analysis of collected responses are being discussed: categorical responses and comparative responses.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY Nr 8(14)

In the first case we have the observation of the following kind:

$$y_{ij} = \begin{cases} 1, & \text{if item } Y \text{ is endorsed by ith respondent,} \\ 0, & \text{if item } Y \text{ is rejected by ith respondent.} \end{cases}$$

In the second case we have the observations n_{jk}, the number of respondents who asserted that Y_j is at least as dangerous as Y_k. For the convenience, we put $n_{jj} = n, \quad j = 1, 2, ..., p.$ The fundamental representation of the probability distribution of the observed data is following:

$$f(y) = \int f(y|z)dH(z).$$

The assumed three basic hypothesis:

(M) $P(Y = 1|Z = z)$ is a coordinatewise nondecreasing function in Z.

(LI) $$P(Y_1 = y_1, ..., Y_p = y_p \,|Z_1 = z_1, ..., Z_d = z_d) = \prod_{j=1}^{p} P(Y_j = y_j \,|Z = z)$$

(U) $$d = 1.$$

They are called correspondingly monotonicity, local independence and unidimensionality.

The most important consequences of these assumptions are following:

1. From condition (LI) and the lack of fit follows the evidence that $d \neq 1$.
2. Condition $d = 1$ and the lack of fit might be considered as the evidence of non-local independence.
3. The (LI) and (M) conditions imply that $Cov(g_1(Y), (g_2(Y)) \geq 0$, g_1 and g_2 nondecreasing.
4. If (LI) and (U) hold, then $Cov(Y_i, Y_j / Z = z) = 0$ for all z, and all pairs i and j.

Much more consequences could be drawn assuming some parametric form of the model.

In the simplest case it is so called logistic model which has the following form:

$$\pi_j(z_i) = P(Y_j = 1 / Z = z_i) = \frac{\exp(z_i - \alpha_j)}{1 + \exp((z_i - \alpha_j)}.$$

This model is called also as Rasch model. It depends on $n+p$ parameters:

$$\alpha_1, \alpha_2, ..., \alpha_p, z_1, z_2, ..., z_n.$$

Parameters determining susceptibility of the respondents $z_1, z_2, ..., z_n$ are treated as nuisance parameters. For the estimation there are used three approaches: joint maximum likelihood (JML), conditional maximum likelihood (CML) method, and the marginal maximum likelihood (MML) method.

Assuming that people's susceptibility to stressful phenomena is interpreted as a real valued random variable Z with a density distribution $h(z)$, we need additionally to estimate this function. Usually one assumes that $Z \sim N(\mu, \sigma^2)$. The problem is in the estimation of μ and σ^2. Parameters μ and σ^2 are estimated by the means of the so-called population likelihood function. In the simplest case, the society under the investigation (respondents) could be divided into two classes. These classes could be called, for example, "content" and "malcontent", or "sensible" and "insensible". In such a dichotomized situation one can assume that the latent trait Z is a binary random variable with distribution $\eta = P(Z = 1) = p$ (respondent is content), $1 - \eta = P(Z = 0) = p$ (respondent is malcontent).

The second big family of models which can be used for stressful phenomena analysis is known as the Factor analysis models:

$$Y_j = \alpha_j Z + \sigma_j \varepsilon_j.$$

This means that the individual's response is treated as a linear combination of susceptibility and random disturbances.

The third class of models discussed in the article is based on the principle of the paired comparisons. It is formulated in the form of the equation:

$$\pi_{ij} = P(Y_i \prec Y_j) = P(Y_i - Y_j > 0),$$

ŚLĄSKI PRZEGLĄD STATYSTYCZNY Nr 8(14)

where π_{ij} denotes the probability of the predominance of Y_i over Y_j.

After having analysed these three families of models we can conclude that the statistical methods developed in different fields of psychology, education and bioassay can be easily adopted for modelling of social phenomena. Particularly, the methods of item response theory can be directly used for social stressors analysis. Merely the little changes in the interpretation of parameters are needed.

References

Andersen E.B., *Discrete statistical models with social science applications*, North-Holland, Amsterdam 1980.

Andersen E.B., Madsen M., *Estimating the parameters of the latent population distribution*, "Psychometrika" 42 (1977), pp. 357-374.

Andrich D., *Rasch models for measurement*, Sage University Paper, 1988.

Bartholomew D.J., Knott M., *Latent variable models and factor analysis*, Arnold, London 1999.

Brunk H.D., *Mathematical models for ranking from paired comparisons*, "American Statistical Association Journal", 9 (1960), pp. 503-521.

David H.A., *The method of paired comparison*, Griffin, London 1969.

Everitt B.S., *An introduction to latent variable models*, Chapman& Hall, London 1984.

Everitt B.S., Hand D.J., *Finite mixture distributions*, Chapman& Hall, London 1981.

Fischer G.H., Molenaar I.W. (eds.), *Rasch models: foundations, recent developments and applications*, Springer-Verlag, New York 1995.

Hambleton R.K, Swaminathan H., Rogers H.J., *Fundamentals of item response theory*, Sage Publications, Newbury Park, CA 1991.

Holland P.W., Rosenbaum P.R., *Conditional association and unidimensionality in monotone latent variable models*, "Annals of Statistics", 14 (1986), pp. 1523-1543.

Junker B.W., Sijtsma K., *Nonparametric item response theory in action*, "Applied Psychological Measurement", 25 (2001), pp. 211-220.

Krauth J., *Testkonstruktion und Testtheorie*, BELTZ 1995.

Lazarsfeld P.F., Henry N.W., *Latent structure analysis*, Houghton-Mifflin, New York 1968.

Mosteller F., *Remarks on the method of paired comparisons. I*, "Psychometrika" 16 (1951), pp. 3-9.

Noether G., *Remarks about a paired comparison model*, "Psychometrika" 25 (1960), pp. 357-367.

Rasch G., *Probabilistic models for some intelligence and attainment tests*, Pœdagogiske Institut, Copenhagen 1960.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY
Nr 8(14)

DISCRETE PROCESS OF DEPENDENT RISKS

Stanisław Heilpern (Wrocław University of Economics)

1. General model

We will investigate the following discrete risk model:

$$U(t) = u + t - \sum_{i=1}^{t} Y_i,$$

where $t = 1, 2, \ldots$, $u \in N$ is an initial capital, $U(0) = u$ and $Y_i = I_i X_i$. We assume, that the discrete claims $X_i = 1, 2, \ldots$ are identically distributed and independent with the probability mass function $f(k)$, decumulative distribution function $F(n)$ and $m = E(X_i)$. The indicators

$$I_i = \begin{cases} 1 & \text{with probability } q \\ 0 & \text{with probability } p \end{cases}$$

are identically distributed and they may be dependent, but independent with the claims X_i. We will analyze the probability of ruin: $\psi(u) = P(U(t) < 0 \text{ for some } t \mid U(0) = u)$.

In the classical model, the independence between indicators I_j is assumed. We can compute the probability of ruin using the recurrence formulas [Shiu 1989]. We also have: $\psi_I(\infty) = 0$. We can compute the exact value of probability of ruin when the claims have the two-point or the exponential distribution.

In the next sections we will study the impact of the degree of dependence on the probability of ruin for different dependent structure of indicators I_j.

2. Strict dependence of I_j

For the strict dependent indicators the probability of ruin is equal

$$\psi_c(u) = \begin{cases} q & \text{for} \quad m > 1 \\ 0 & \text{for} \quad m = 1 \end{cases}.$$

We obtain the following relations between the probability of ruin for the independent $\psi_I(u)$ and strict dependent cases: $\psi_I(\infty) < \psi_c(\infty)$, $\psi_I(0)$

$> \psi_c(0)$, when $m + q > 2$, $\psi_I(0) = \psi_c(0)$ for $m + q = 2$ and $\psi_I(0) < \psi_c(0)$ else. We see, that there is not regularity when $m + q > 2$. For the smaller initial capital the probability of ruin when the indicators are independent is greater then in the strict dependence case. For the bigger initial capital we obtain reverse relation.

3. Archimedean copulas

Now, let us assume that the dependence structure of indicators is described by Archimedean copula C with the generator g. Then there exists the random variable $\Theta \sim F_\Theta$ [Frees 1998] with the Laplace transform $L_\Theta(s) = g^{-1}(s)$. The indicators are conditional independent for fixed $\theta \in \Theta$ in this case. We obtain the conditional risk process $U_\theta(u)$ and conditional indicators $I_{j|\theta}$ with the probability of claim $q(\theta) = \exp(-\theta g(q))$. The unconditional probability of ruin is equal

$$\psi(u) = \int_0^\infty \psi(u \mid \theta) dF_\Theta(\theta) = \int_{\theta_0}^\infty \psi(u \mid \theta) dF_\Theta(\theta) + F_\Theta(\theta_0),$$

where $\psi(u|\theta)$ is the conditional probability of ruin and $\theta_0 = \dfrac{\ln m}{g(q)}$. For the initial capital equals zero and infinity we obtain

$$\psi(0) = \int_{\theta_0}^\infty \frac{q(\theta)}{1-q(\theta)}(m-1) dF_\Theta(\theta) + F_\Theta(\theta_0), \qquad \psi(\infty) = F_\Theta(\theta_0).$$

When the claims X_i have the geometric distribution with β we have the exact formula for the probability of ruin

$$\psi(0) = \frac{\beta^{u+1}}{1-\beta} \int_{\theta_0}^\infty \frac{q(\theta)}{(1-q(\theta))^{u+1}} dF_\Theta(\theta) + F_\Theta(\theta_0),$$

when $\theta_0 = -\dfrac{\ln(1-\beta)}{g(q)}$.

In the case when the dependence structure is described by Clayton family

$$C_\alpha(u_1, ..., u_n) = (u_1^{-\alpha} + ... + u_n^{-\alpha})^{-1/\alpha}, \qquad \alpha > 0,$$

with generator $g(u) = (u^{-\alpha} - 1)/\alpha$, the induced random variable Θ has Gamma distribution $\mathrm{Ga}\left(\frac{1}{\alpha}, \alpha\right)$, $q_\alpha(\theta) = e^{\theta(1-q^{-\alpha})/\alpha}$ and the limit value of θ is equal $\theta_\alpha = \frac{\alpha q^\alpha \ln m}{1-q^\alpha}$. The parameter α reflects the degree of dependence. The Kendall coefficient of correlation takes the form $\tau = \alpha/(\alpha + 2)$ in this case.

Example. Let $q = 0.3$, claims X_i have the geometric distribution with $\beta = 0.5$ and dependence structure is described by Clayton copula with parameter α. There are graphs of the probability of ruin for the values of parameter $\alpha = 0, 0.1, 1, 2, 4, \infty$ on the figure 1.

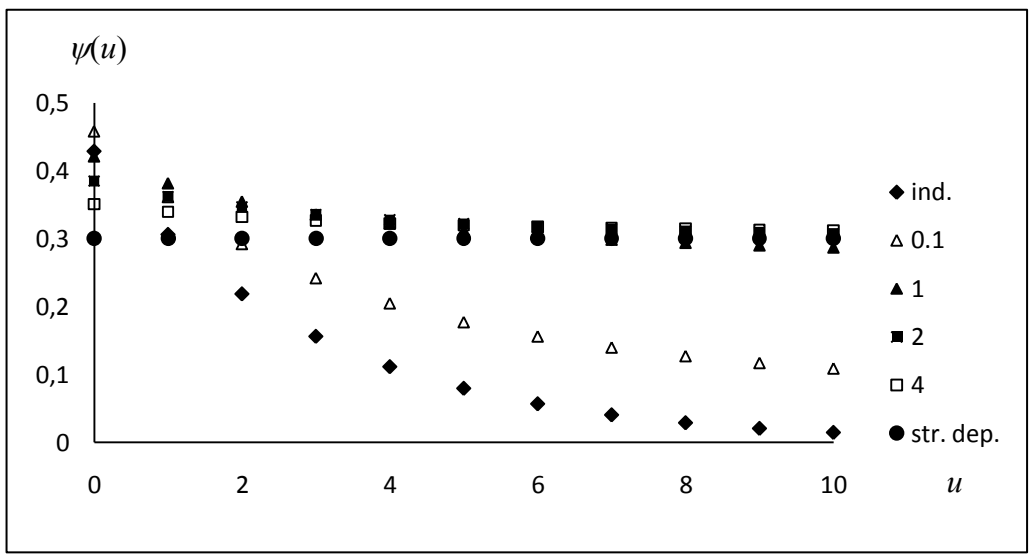

Fig. 1. Probability of ruin for different values of α

Source: own calculation.

We see that there is not regularity in the relation between the degree of dependence and the probability of ruin. For different values of initial capital u we obtain different order of the values of probability of ruin.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 8(14)

4. Markov binomial distribution

Let now assume that dependent structure of the indicators I_j is described by Markov stationary chain with state space {0, 1} and the following matrix of transition probabilities:

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = \begin{pmatrix} p + \pi q & p - \pi q \\ p - \pi q & q + \pi p \end{pmatrix},$$

where π is Pearson coefficient of correlation ($0 \leq \pi \leq 1$).
The probability of ruin takes the form

$$\psi(u) = (1 - q)\psi(u|0) + q\psi(u|1),$$

where $\psi(u|i)$ is conditional probability of ruin when $I_0 = i$, $i \in \{0, 1\}$ in this case. These conditional probabilities we can compute using the following recurrence equations (see [Cossete et al. 2004].

The greater value of the degree of dependence implies the greater probability of ruin [Cossete et al. 2003]:

$$\pi_1 < \pi_2 < 1 \quad \square \quad \psi_{\pi_1}(u) < \psi_{\pi_2}(u).$$

The limit value of probability of ruin is equal

$$\psi_g(u) = \lim_{\pi \to 1} \psi_\pi(u) = qm,$$

for any values of initial capital u. But, for $\pi = 1$ we have the strict dependence with the probability of ruin $\psi_c(u) = q$ for $m > 1$. We obtain the similar relation between $\psi_c(u)$ and $\psi_\pi(u)$ similar as in section 2.

$$\psi_c(u \mid 0) = 0 \qquad \psi_c(u \mid 1) = \begin{cases} 1 & \text{for} \quad m > 1 \\ 0 & \text{for} \quad m = 1 \end{cases},$$

$$\psi_c(u) = \begin{cases} q & \text{for} \quad m > 1 \\ 0 & \text{for} \quad m = 1 \end{cases}.$$

References

Cossete H., Landriault D., Marceau E., *Ruin probabilities in the compound Markov binomial model*, "Scandinavian Actuarial Journal", 4 (2003), pp. 301-323.

Cossete H., Landriault D., Marceau E., *Exact expressions and upper bound for ruin probabilities in the compound Markov binomial model*, "Insurance: Mathematics and Economics", 34 (2004), pp. 449-466.
Frees E.W., Valdez E. A., *Understanding relationships using copulas*, "North American Actuarial Journal", 2 (1998), pp. 1-25.
Shiu E., *The probability of eventual ruin in the compound binomial model*, "ASTIN Bulletin", 19 (1989), pp. 179-190.

RANDOMIZED RESPONSE VERSUS IMPUTATION – A COMPARISON REGARDING THE QUALITY OF DISTRIBUTION RECONSTRUCTION

Heiko Grönitz (University of Marburg)

1. Introduction

Surveys often contain sensitive questions like "How much do you earn?" or "Are you unemployed?" or "Have you ever evade taxes?". If such questions are asked, some interviewees will refuse responding, since they are afraid of resentments, small valuation or prosecution. In the missing data literature several imputation methods are suggested to repair the nonresponse. Each of them leads to a data set without missing values. This completed data set can be used to estimate the distribution of the considered variables.

A different approach is given by randomized response techniques. Thereby a sensitive question is not asked directly, but any respondent is requested to give a randomized response (RR), which does not provide his or her value of the sensitive variable. However, by the frequencies of the randomized replies the distribution of the underlying sensitive characteristic can be estimated.

We are going to compare the alternatives with respect to the quality of distribution reconstruction. We operate the comparison with the help of a simulation study. For that purpose we choose one special RR model and one special distribution in today's lecture. The RR model is the diagonal model and the distribution of interest is Germany's income-age distribution.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY Nr 8(14)

Before presenting the results of simulation we outline briefly some developments in the field of randomized responses.

2. Randomized response models

The randomized response theory started with the paper of Warner [1965]. Warner considered a dichotomous variable $K \in \{0, 1\}$. $K = 1$ indicates that the respondent has a sensitive characteristic. One could imagine that an interviewee has value $K = 1$ if he or she has ever evaded taxes and $K = 0$ otherwise. Every respondent has to operate a randomization device (RD). A RD is a random experiment. According to the result of the experiment and the value of the interesting variable the respondent gives a randomized response. For instance the respondent may be introduced to choose randomly one of the following two questions:

$Q = 1$: "Is your value of K equal to 1?"
$Q = 2$: "Is your value of K equal to 0?"

The question is selected for example by spinning a spinner, drawing cards or tossing a dice. The selection occurs hidden and the selected question is not revealed to the interviewer. The respondent replies either "yes" or "no", but the interviewer cannot identify the respondent's value of K. Hence one can assume truthful randomized responses.

Put $p := \mathrm{P}(Q = 1)$ and $\pi := \mathrm{P}(K = 1)$. Then the probability of an answer "yes" is

$$\mathrm{P}(\text{"}yes\text{"}) = p\pi + (1 - p)(1 - \pi).$$

Assume n respondents are asked and let $h := h_n$ be relative frequency of "yes" answers. Estimate π by $\hat{\pi} = (\mathrm{p} - 1 + h)/(2p - 1)$, where $p \neq 1/2$. The estimator is unbiased, but can attain values outside [0,1] for small n.

A large variety of extensions and versions of the Warner model has been discussed in the literature. For a review [Hedayat, Sinha 1991, chapter 11; Tan et al. 2009, section 2.2], can be recommended.

However, it must be mentioned that randomization devices have some disadvantages: the preparation and operation is extensive. An interviewer must always be present and it has to be ensured, that the

result of the RD is not revealed to the interviewer. Since the necessary presence of an interviewer, the methods are not useful for email surveys. These drawbacks motivated a newer development in the literature. Tian et al. [2007], Yu et al. [2008] and Tan et al. [2009] discussed some models without RD. Let us have a look at the crosswise model by Yu et al. [2008]. The authors considered a variable $X \in \{1, 2\}$ and chose an auxiliary variable $W \in \{1, 2\}$ with known distribution whereas X and W can be assumed as independent. As concrete W the period of birthday is suggested, e.g. $W = 1$ may indicate if a person is born between August and September. In this case the assumption $P(W = 1) = 5/12$ is reasonable. The interviewee gives a reply $A \in \in \{1, 2\}$ according to

$$\{A = 1\} = \{X = 1, \ \ W = 1\} \cup \{X = 2, \ \ W = 2\} \ \text{ and } \{A = 2\} = \overline{\{A = 1\}} \ .$$

Suppose $P(X = 2) = \pi$ and $P(W = 2) = p$. Then it is $P(A = 1) = p\pi + +(1 - \pi)(1 - p)$. Replacing $P(A = 1)$ by the relative frequency $h := h(A = 1)$ and solving the equation leads to an estimator for π:

$$\tilde{\pi} = (h - 1 + p) / (2p - 1).$$

To obtain an estimator with range [0, 1] modify $\tilde{\pi}$ to

$$\hat{\pi} = \min(1, \ \ \max(0, \ \tilde{\pi})).$$

The respondent's X-value is not identifiable by A. Hence no nonresponse and truthful answers are assumed.

3. Diagonal model

The crosswise model can only treat two-valued variables. So we thought about an extension for variables $X \in \{1, \ldots, k\}$, $k \geq 2$. As above choose an auxiliary variable W, but now with values 1, …, k. The knowledge of the distribution of W and the independence of X and W are supposed.

The respondent is requested to give the answer

$$A = [(W - X) \bmod k] + 1:$$

A describes the diagonal the respondent belongs to, e.g. for $k = 4$ we obtain responses according to the table

ŚLĄSKI PRZEGLĄD STATYSTYCZNY Nr 8(14)

Table 1. The answers of respondents.

X/W	$W=1$	$W=2$	$W=3$	$W=4$	$W=1$	$W=2$	$W=3$
$X=1$	1	2	3	4			
$X=2$		1	2	3	4		
$X=3$			1	2	3	4	
$X=4$				1	2	3	4
diagonal				**d1**	**d2**	**d3**	**d4**

Source: own calculations.

The interviewer hears an answer $A \in \{1, \dots, k\}$, but it is not possible to identify the X-value with the help of the answer. Hence it is allowed to assume no nonresponse and truthful answers again. In the following define $\pi_i := P(X = i)$, $\pi := (\pi_1, \dots; \pi_k)^T$ and $c_i := P(W = i)$. It holds

$$(P(A = 1), \dots, P(A = k))^T = C_0\pi,$$

thereby C_0 is a $k \times k$ – Matrix where every row is a left-cyclic shift of the row above. The aim is to estimate the vector π. Therefore estimate the probabilities $P(A = j)$ by the corresponding relative frequencies $h_j := h(A = j)$. Define $h := (h_1; :::; h_k)^T$ and

$$\begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} := C_0^{-1}h \quad \text{and} \quad \hat{\pi} := \frac{1}{\sum_{i=1}^{k} \max(0, x_i)} \begin{pmatrix} \max(0, x_1) \\ \vdots \\ \max(0, x_k) \end{pmatrix}.$$

Obviously $\hat{\pi}$ is in the k-dimensional unit cube and the sum of its components is equal to one. Further $\hat{\pi}$ is a strongly consistent estimator for π and finally $\hat{\pi}$ is usually the unique maximum likelihood estimator (MLE) for π. More precisely the last property means: let $\pi_1, \dots, \pi_k > 0$, then $\hat{\pi}$ is with probability 1 for all sufficiently large sample sizes $n \geq N \in \mathbb{N}$ the unique MLE.

4. Simulation study

There are two possibilities to detect Germany's income-age distribution: on one hand request the interviewees for an answer according to the diagonal model whereas we suppose that no nonresponse occurs, since the respondents' privacy is protected, on the other ask directly

whereas missing values are removed by several imputation methods. We will check the suitability of both alternatives by simulations with MATLAB.

4.1. Data

Germany's income-age distribution is offered by the Federal Statistical Office in Germany[1]. In detail it divides income into 20 classes and age into 7 classes and provides the frequency of every combination. The population consists of the set of ca. 35 million taxpayers in 2004. We make some technical idealizations concerning the data, e.g. we assume age has range [16, 85], income has upper bound 10 million, no negative income and a uniform distribution within each income age combination.

4.2. Results of simulation

We processed following simulations:

1. Specify n_1 and n_2 income and age classes respectively ($k = n_1 n_2$ combinations). Further fix the vector $c = (c_1, \ldots, c_k)$, which describes the distribution of the auxiliary variable W.
2. Draw 50 samples of size n from the income-age distribution. For each sample let

h_{ij} : common relative frequency of i-th income class and j-th age class.

Then at first estimate the frequencies h_{ij} by DM estimator $\hat{h}_{ij}$ and calculate the reconstruction measure

$$A_{DM} = \sum_{j=1}^{n_2} \sum_{i=1}^{n_1} \left| h_{ij} - \hat{h}_{ij} \right| \in [0, 2].$$

Afterwards direct questioning is simulated whereas we consider 4 nonresponse mechanisms $N_1, \ldots, N_4$ and 4 imputation methods $I_1, \ldots, I_4$. For every pair (N_l, I_m) calculate $\hat{h}_{ij}^{N_i, I_m}$, i.e. the common relative

[1] Data are available on www.destatis.de (only in German): Fachserie 14, Reihe 7.1, Steuerfälle nach Alter und nach Grösenklasse der Summe der individuellen Einkünfte.

frequency of *i*-th income class and *j*-th age class after completing the data. Then determine

$$A_{N_l,I_m} = \sum_{j=1}^{n_2}\sum_{i=1}^{n_1}\left|h_{ij} - \hat{h}_{ij}^{N_l,I_m}\right| \quad (l,m = 1,...,4)$$

as the measure of reconstruction.

3. Average the 50 values for each of A_{DM} and A_{N_l,I_m} ($l,m = 1, \ldots, 4$).

We have studied four nonresponse mechanisms, two of the type missing completely at random (MCAR) and two of the type missing not at random (MNAR) – cf. Little, Rubin [2002]. The first MCAR mechanism causes a drop-out probability of each 30% for income and age while the second MCAR mechanism leads to about 50% nonresponse in each variable. Using the first and second MNAR mechanism each variable possesses a nonresponse rate of 30% and 50%, respectively, whereas the drop-out probabilities depend on the value of underlying variable. The considered imputation methods are mean replacement, regression imputation (with stochastic term of noise), hotdeck method and a model-based imputation. In detail for the last one we supposed a bivariate log-normal distribution, estimated parameters with the help of the observed data and removed missing data by drawing random numbers from the conditional distribution or the bivariate distribution.

For example for each two income and age classes we obtained Figure 1. Thereby "DM large std" and "DM small std" means the estimation by diagonal model with $c = c^{(1)}$ and $c = c^{(2)}$ respectively. Thereby we have empirical standard deviations std($c^{(1)}$) = 0:33 and std($c^{(2)}$) = 0:24, i.e. using the second one the distribution of the auxiliary variable is closer to a uniform distribution. Moreover, "Lmodel" is the abbreviation for the model-based imputation.

Figure 1 consists of four plots – one for each nonresponse mechanism. In each plot the reconstruction measure A (sum of absolute distances) is presented as a function of the sample size n (we have operated simulations for $n \in \{50, 100, 250, 500, 1000\}$).

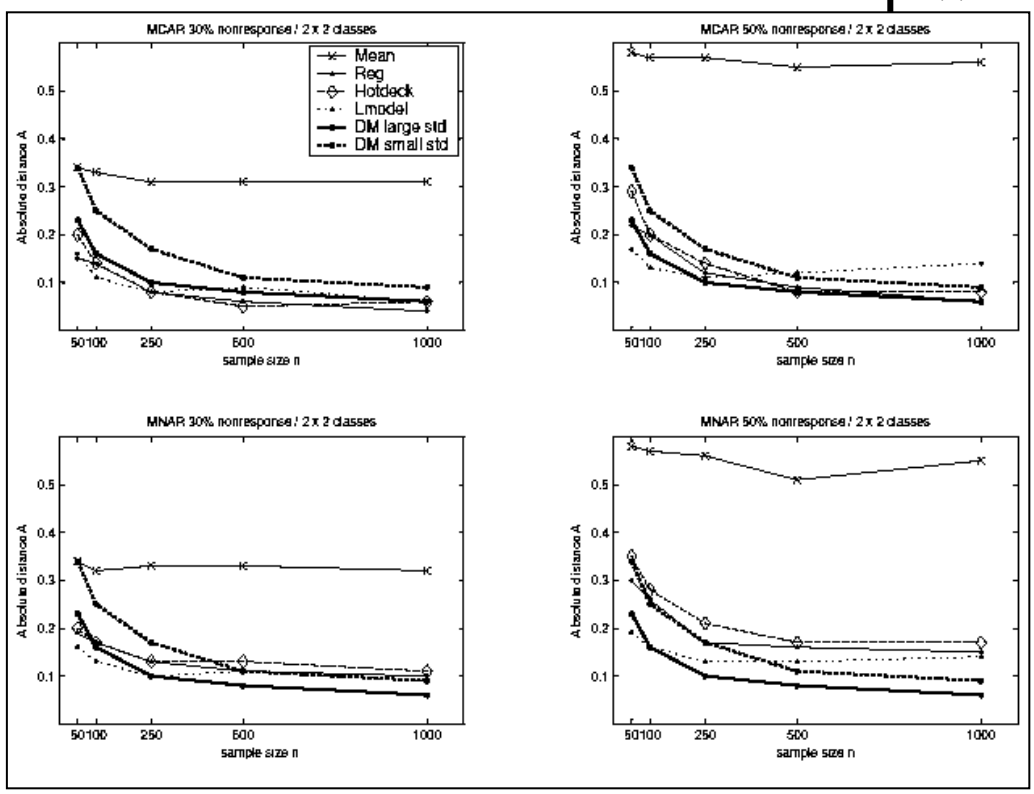

Fig. 1. Results of simulation

Source: own calculations.

In reality usually non-response of type MNAR with a high drop-out rate will occur if one asks for sensitive variables. Then we find a situation as in the lower right plot. Here the imputation methods produce quite bad results. However, a survey designed according to the diagonal model can be a helpful alternative.

References

Hedayat A.S., Sinha B.K., *Design and inference in finite population sampling*, Wiley, New York 1991.

Little R.J.A., Rubin D.B., *Statistical analysis with missing data*, Wiley, New York 2002.

Tan M.T., Tian G.L., Tang M.L., *Sample surveys with sensitive questions: a nonrandomized response approach*, "The American Statistician", 63 (2009), pp. 9-16.

Tian G.L., Yu J.W., Tang M.L., Geng Z., *A new non-randomized model for analysing sensitive questions with binary outcomes*, "Statistics in Medicine", 26 (2007), pp. 4238-4252.

Warner S.L., *Randomized response: A survey technique for eliminating evasive answer bias*, "Journal of the American Statistical Association", 60 (1965), pp. 63-69.

Yu J.W., Tian G.L., Tang M.L., *Two new models for survey sampling with sensitive characteristic: design and analysis*, "Metrika", 67 (2008), pp. 251-263.

AN APPROACH TO THE STUDY OF PROSPECTIVE RESERVES FOR MULTISTATE INSURANCE CONTRACTS

Joanna Dębicka (Wrocław University of Economics)

Irrespective of type, each insurance contract gives rise to two payment streams. The first one is a stream of premium payments which flows from the insured to the insurer. The second (in the opposite direction) is a stream of actuarial payment functions where fixed amounts under the annuity product and fixed insurance benefits are considered as a series of deterministic future cash flows. From the insurer's point of view, at the beginning net premiums are calculated in the way that the actuarial value of future benefits balances the actuarial value of future premiums since this balance is not preserved during insurance period. Thus with each insurance contract there is a special fund associated, called *insurance reserve*, which is the difference between actuarial value of future benefits and net premiums. This fund is used for the protection of solvency of the insurer.

The aim of the talk was to give a formula for prospective reserves for multistate insurance contracts, both for deterministic and stochastic rate of interest. In order to simplify the form of the derived expression matrix notation was used.

Multiple state modelling is a classical stochastic tool for designing and implementing insurance products. The multistate methodology is intensively used in the calculation of premiums and reserves of differ-

ent types of insurance, such as life, disability, sickness, marriage or unemployment insurance. The pair (S, T) is called a *multiple state model*, and describes all possible insured risk events as far as its evolution is concerned (usually up to the end of insurance). That is, at any time the insured risk is in one of a finite number of states belonging to the *state space* $\{S = \{1, 2, \dots, N\}$. Each state corresponds to an event which determines the cash flows (premiums and benefits). By T we denote the *set of direct transitions* between states of the state space.

We consider an insurance contract issued at time 0 (defined as the time of issue of the insurance contract) and, according to the plan, terminating at a later time n (n is the term of policy). Let $X(t)$ denote the state of an individual (the policy) at time t. Hence the evolution of the insured risk is given by a discrete-time stochastic process $\{X(t): t = 0, 1, 2, \dots\}$, with values in the finite set S. If we look at the evolution of the contract, then both the presence at a given state and the movement from state to another state may have some financial impact. We distinguish between the following types of cash flows related to multistate insurance:

- $b_j(k)$ – an annuity benefit at time k if $X(k) = j$,
- $d_j(k)$ – a lump sum at some fixed time k if $X(k) = j$,
- $c_{ij}(k)$ – a lump sum at time k if a transition occurs from state i to state j at that time,
- $\pi_j(k)$ – a premium amount at some fixed time k if $X(k) = j$,
- $p_j(k)$ – a period premium amount at time k if $X(k) = j$.

Because we focus on discrete-time model, it means that insurance payments are made at the ends of time intervals. Practically it means, that annuity and insurance benefits are paid immediately before the end of the unit time (for example: year or month). Premiums are paid immediately after the beginning of the unit time.

In view of financial mathematics, future cash flows, which are realized at time k, are discounted to the present (to time t) by some interest rate. This produces the cash value of future payment stream $\Upsilon_t^{\wp,j}(k)$, where $\wp$ denotes one of the types of cash flows ($\wp \in \{p, \pi, b, d, c_1, c_2, \dots, c_N\}$ and c_i is the benefit paid if process $\{X(t)\}$ leaves state i). If $\wp \in \{p, \pi, b, d\}$, then cash value of cash flow is given by

ŚLĄSKI PRZEGLĄD STATYSTYCZNY Nr 8(14)

$$\Upsilon_t^{\wp,j}(k) = \upsilon(t,k)\mathbf{1}_{\{X(k)=j\}}\wp_j(k),$$

while for $\wp \in \{c_1, c_2, \ldots, c_N\}$ we have

$$\Upsilon_t^{c_i,j}(k) = \begin{cases} \upsilon(t,k)\mathbf{1}_{\{X(k-1)=i \wedge X(k)=j\}} c_{ij}(k) & \text{for} \quad i \neq j \\ 0 & \text{for} \quad i = j \end{cases},$$

where for stochastic interest rate $Y(t)$, the discount function is in the following form $v(t, k) = e^{-(Y(k)-Y(t))}$ (for $0 \leq t \leq k$ and $Y(t)$ denotes the rate of interest in time interval $[0,t]$).

At moment t the sum of cash value of future payment stream is called prospective loss ${}_tL$ of the insurer at time t, so

$${}_tL = \sum_{\wp\in\{b,d,c_1,\ldots,c_N\}} \sum_{j\in S} \sum_{k=t+1}^{n} \Upsilon_t^{\wp,j}(k) - \sum_{\wp\in\{p,\pi\}} \sum_{j\in S} \sum_{k=t}^{n-1} \Upsilon_t^{\wp,j}(k).$$

Benefits are an inflow representing an income to loss fund. Premiums represent an outgo from a loss fund of the insurer. Let us observe that ${}_tL$ is a random variable and its distribution depends on the probabilistic structure of multistate model and the stochastic interest rate. Moreover, at moment t insurance contract may be at any state, then for a given moment t we need to count the prospective reserves for all states separately. Then prospective reserve is a conditional expectation of prospective loss under the condition that at time t the insurance contract is at state i

$$V_i(t) = \mathrm{E}\left({}_tL \mid X(t) = i\right)$$

$$= \sum_{\wp\in\{b,d,c_1,\ldots,c_N\}} \sum_{j\in S} \sum_{k=t+1}^{n} \mathrm{E}\left(\Upsilon_t^{\wp,j}(k) \mid X(t) = i\right) - \sum_{\wp\in\{p,\pi\}} \sum_{j\in S} \sum_{k=t}^{n-1} \mathrm{E}\left(\Upsilon_t^{\wp,j}(k) \mid X(t) = i\right).$$

Note that this formula looks quite complex. Namely, we look at it and we do not see the structure of the analyzed insurance contract. Fortunately, it appears that it is possible to simplify it using matrix notation. To do this we have to introduce the extended multistate model and describe matrices related to: multistate model and its probabilistic structure ($\mathbf{P}(0)$ – vector of initial distribution and $\{\mathbf{Q}(k)\}_{k=0,1,2,\ldots}$ – sequence of matrices transition of process $\{X(t)\}$),

cash flows ($\mathbf{C}_m$ consists only of an income to a particular fund, $\mathbf{C}_{out}$ consists only of an outgo from a fund and $\mathbf{C}_m + \mathbf{C}_{out} = \mathbf{C}$) and discount function ($\mathbf{\Lambda}$ consists of discount and accumulated functions). Now the following theorem is proven.

Theorem

For the insurance contract described by extended multistate model (S, T) vector $\mathbf{V}(t) = (V_1(t), V_2(t), \ldots, V_N(t))^T$, of prospective reserves at moment t is in the following form

$$\mathbf{V}(t) = \left(\mathbf{C}_{out}^T + \sum_{k=t+1}^{n} \prod_{u=t}^{k-1} \mathbf{Q}(u) \mathbf{C}^T \mathbf{I}_{k+1} \mathbf{I}_{k+1}^T \Lambda \right) \mathbf{I}_{t+1},$$

where $\mathbf{I}_{t+1}$ is a vector which consists of zeros except for 1 at $t + 1$ coordinate.

Matrix approach enables us to give a flexible tool not only for numerical calculations but also for the analysis of gross reserves, emerging costs and profit testing and helps in analysing both a single policy and a portfolio of policies.

As a numerical illustration, a health insurance contract was considered, for which prospective reserves in the whole insurance period were calculated, using the above introduced theorem.

MONITORING CHANGES IN LINEAR MODELS WITHOUT INTERCEPT

Alexander Schmitz (University of Cologne)

1. Introduction

This note contains a further discussion of a sequential change-point test proposed by Horváth et al. [2004] and Hušková, Koubková [2005]. They designed a test to detect a change in the parameter (α_i, β_i) of a linear model

$$y_i = \alpha_i + x_i \beta_i + \varepsilon_i, \qquad i = 1, 2, \ldots$$

ŚLĄSKI PRZEGLĄD STATYSTYCZNY Nr 8(14)

We focus on the detection of a change in the regression parameter β_i solely. Thus, we consider a simple linear model without intercept α_i, i.e.

$$y_i = x_i \beta_i + \varepsilon_i, \quad i = 1, 2, ..., \tag{1}$$

where $\{x_i\}_{1 \le i < \infty}$ is the real-valued regressor sequence and $\{\varepsilon_i\}_{1 \le i < \infty}$ denotes the error process. Our common approach rests upon a monitoring scheme by Chu et al. [1996]. They assumed a historical period of length m with a constant but unknown regression parameter β_0, i.e.

$$\beta_i = \beta_0, \qquad i = 1, ..., m. \tag{2}$$

Since an infinite monitoring period starts subsequently to the historical period, their change-point test is designed as a sequential analysis. The parameter stability null hypothesis

$$H_0 : \beta_i = \beta_0, \ i = m+1, ...,$$

is checked after each arrival of a new data against a certain change alternative H_A.

2. Two regressor sequences

Another feature of our model is the consideration of two regressor sequences: $\{x_{i,0}\}_{1 \le i \le m}$ on the historical period and $\{x_{i,1}\}_{1 \le i < \infty}$ on the monitoring period. For the ease of notation we set

$$x_i = \begin{cases} x_{i,0}, & 1 \le i \le m \\ x_{i-m,1}, & i = m+1, m+2, ... \end{cases}. \tag{3}$$

This reflects the following situation. After the historical period there are no longer observations for the first regressor sequence available. But it is possible to use data from a second source and the historical regression parameter remains. The historical regression parameter is suitable for the new model until the detection of a parameter shift from β_0 to a different value β_* (say). Therefore, it seems appropriate to detect the change-point k^* (say) via a sequential analysis. Follow-

ing Chu et al. [1996], the testing procedure stops at time $\tau(m)$, according to the first excess of a detector $\hat{Q}_m(\cdot)$ over a boundary function $g_m^*(\cdot)$, i.e.

$$\tau(m) = \inf\left\{k : \left|\hat{Q}_m(k)\right| > \sqrt{d}\sigma c(\alpha) g_m^*(k)\right\},$$

where σ and $\sqrt{d}$ are positive constants and $c(\alpha)$ is a critical constant. Moreover, we set $\inf \varnothing = \infty$, if the path of the detector never exits the boundary. For the purpose of an asymptotically controlled level α, the critical constant $c(\alpha)$ can be determined via a limit distribution. Moreover, the test is shown to be consistent against a large class of change-point alternatives.

3. Detector and model assumptions

With a view to gain consistency, the residual based cumulative sum detector (CUSUM) includes regression weights, i.e.

$$\hat{Q}_m(k) = \sum_{i=m+1}^{m+k} x_i \hat{\varepsilon}_i \quad , \qquad k = 1, 2, \ldots \tag{4}$$

The empirical residuals $\hat{\varepsilon}_i = y_i - x_i \hat{\beta}_m$ are computed via the least squares estimator:

$$\hat{\beta}_m = \left(\sum_{i=1}^{m} x_i x_i\right)^{-1} \sum_{i=1}^{m} x_i y_i . \tag{5}$$

The least squares estimator relies only on the historical period. Next, assume that the error sequence $\{\varepsilon_i\}_{1 \le i < \infty}$ is a strictly stationary process satisfying:

$$E\varepsilon_1 = 0, \quad E\varepsilon_1^2 = \sigma^2 \text{ and } \quad E\varepsilon_1 \varepsilon_i = 0 \quad \forall i > 1. \tag{6}$$

We allow for an M-dependence among the error variables, i.e.

$$\varepsilon_i \text{ and } \varepsilon_j \text{ are independent, if } |i - j| > M . \tag{7}$$

ŚLĄSKI PRZEGLĄD STATYSTYCZNY Nr 8(14)

This dependence should reflect a certain correlation between the two regressor sequences involved. We need a further moment condition:

$$E\,|\,\varepsilon_1\,|^{2+\delta} < \infty, \text{ for some } \delta > 0. \tag{8}$$

Although we observe the regressor data, we need a condition on the data generating process, which in turn yields a convenient large sample behaviour of the realisations. We assume that the squared regressors obey a strong law of large numbers with a certain rate, i.e. there are positive constants d and $0 < \tau < 1/2$, such that

$$\frac{1}{n^{1-\tau}} \sum_{i=1}^{n} \left(x_{i,0}^2 - d \right) \xrightarrow{a.s.} 0 \tag{9}$$

holds almost surely, as $n \to \infty$. And similar for the second regressor sequence:

$$\frac{1}{n^{1-\tau}} \sum_{i=1}^{n} \left(x_{i,1}^2 - d \right) \xrightarrow{a.s.} 0 \tag{10}$$

holds almost surely, as $n \to \infty$. As a consequence of (9) and (10), the variance of each regressor sequence is asymptotically equal to d. Horváth et al. [2004] introduced a class of boundary functions being analytically convenient for the CUSUM monitoring:

$$g_m^*(k) = m^{1/2} \left(1 + \frac{k}{m} \right) \left(\frac{k}{m+k} \right)^{\gamma}, \qquad 0 \le \gamma < 1/2. \tag{11}$$

The parameter γ is the so-called tuning constant influencing the detection ability.

4. Results

Under the null hypothesis H_0, suppose (1)-(11) hold, then we have:

$$\lim_{m \to \infty} P\left(\frac{1}{\sqrt{d}\,\sigma} \sup_{1 \le k \le \infty} \frac{\left| \hat{Q}_m(k) \right|}{g_m^*(k)} > c \right) = P\left(\sup_{0 < t \le 1} \frac{|W(t)|}{t^{\gamma}} > c \right).$$

The limit distribution is a functional of a standard Wiener process $\{W(t)\}_{0\le t<\infty}$. Selected quantiles are given in [Horváth et al. 2004]. An application of the monitoring procedure in practice requires a consistent estimation of the unknown error deviation σ. Estimators for the parameter σ are available using a "non-overlapping blocking" approach, cf. [Schmitz, Steinebach 2008]. Now we discuss the consistency of the test under several change alternatives. Therefore, we allow the change-point k^* and the parameter shift $\Delta_m = \beta_* - \beta_0$ to vary with m. We assume that the change-point does not occur too late, relatively to the size of the historical period:

$$k^* = O(m/\log m) \qquad m \to \infty \ . \tag{12}$$

Under the "fixed-change alternative", i.e. $\Delta_m = \Delta$, suppose (1)-(12) hold, then we have:

$$\frac{1}{\sqrt{m}} \frac{\left|\hat{Q}_m(m+k^*)\right|}{g_m^*(m+k^*)} \xrightarrow{P} \frac{d\Delta}{2^{1-\gamma}} \qquad m \to \infty.$$

And under the "shrinking-change alternative", i.e. $\lim_m \Delta_m m^\tau = \Theta$, suppose (1)-(11) hold, then we have:

$$\frac{1}{m^{(1/2)-\tau}} \frac{\left|\hat{Q}_m(m+k^*)\right|}{g_m^*(m+k^*)} \xrightarrow{P} \frac{d\Theta}{2^{1-\gamma}} \qquad m \to \infty .$$

From these two stochastic limits we derive that a fixed change can be detected earlier than a shrinking change. Moreover, if there are two constants, such that $C_1 m^{-\tau} \le \Delta_m \le C_2$ holds, the testing procedure has asymptotical power one.

5. Remarks

In the regression weighted CUSUM monitoring by Hušková and Koubková [2005] an independent error sequence is assumed. The present note shows that the monitoring procedure permits for an M--dependence among the error variables. In [Horváth et al. 2004] an

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 8(14)

additional parameter constraint, i.e. $0 \le \gamma < \min\left\{1/2, \tau\right\}$, is assumed. Since this constraint is due to the intercept, we do not need this constraint here. Assumption (9) and (10), that the squared regressor sequences obey a strong law of large numbers with a certain rate, hold for a large class of stochastic processes. Some extension of the so-called Marcinkiewicz-Zygmund law of large numbers to dependent processes will be presented elsewhere.

References

Chu C.S.J., Stinchcombe M., White H., *Monitoring structural change*, "Econometrica", 64 (1996), pp. 1045-1065.

Horváth L., Hušková M., Kokoszka P., Steinebach J.G., *Monitoring changes in linear models*, "Journal of Statistical Planning and Inference", 126 (2004), pp. 225-251.

Hušková M., Koubková A., *Monitoring jump changes in linear models*, "Journal of Statistical Research", 39 (2005), pp. 51-70.

Schmitz A., Steinebach J.G., *A note on the monitoring of changes in linear models with dependent errors*, Preprint University of Cologne (2008), pp. 1-14.

RUIN PROBABILITY IN INFINITE TIME

Aleksandra Iwanicka (Wrocław University of Economics)

We consider a risk model for three classes of insurance business as an example of a multiclasses risk model, i.e. a risk model for several classes of insurance business. The classes of business are correlated. The correlation between classes can be the effect of some outside risk factors like natural disasters that causes various kinds of insurance claims. The main aim is to investigate the impact of some outside risk factors which causes additional claims in each class of insurance business on ruin probability in infinite time.

We consider a risk model involving a book of three dependent classes of insurance business. Let $\left\{X_{ij}\right\}_{i=1}^{\infty}$ be a sequence of independent claim size random variables for i-th class of business with com-

mon probability function f_{X_i} and mean μ_i. Then the aggregate claim sizes process for a book of three classes of business is given by:

$$S(t) = \sum_{i=1}^{3} \sum_{j=1}^{N_i(t)} X_{ij},$$

where $\{N_i(t)\}_{t\geq 0}$ is the claim number process for i-th class. It is assumed that all claim sizes are independent and that they are independent of all claim counting processes. The claim number processes are correlated in the way:

$$N_1(t) = M_2(t) + M_4(t), \quad N_2(t) = M_2(t) + M_4(t) + M_5(t) \text{ and}$$
$$N_3(t) = M_3(t) + M_5(t)$$

with $\{M_1(t)\}, \{M_2(t)\}, \{M_3(t)\}, \{M_4(t)\}$ and $\{M_5(t)\}$ being independent Poisson processes with intensities respectively $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and λ_5. In the i-th class of business the underlying risks of this class cause claim numbers according the process $\{M_i(t)\}$. Additionally, in the first class and in the second class some outside risks common for both of these classes cause claim numbers according to the process $\{M_4(t)\}$. Also in the second and in the third class some other outside risks common for both of these classes cause claim numbers according to the process $\{M_5(t)\}$. The situation of the impact of all risks factors on three classes of business are shown in auxiliary figure 1. Then the risk process for a book of these classes is given by:

$$R(t) = u + ct - S(t), \tag{1}$$

where u is the amount of initial surplus and c is the constant rate of premium per unit time. To satisfy the net profit for the insurance company we assume that $c = (1+\theta)E(S(1))$, where $\theta > 0$ is called the relative safety loading. A risk model for one class of insurance business with claim counting process being Poisson process is called a classical risk model. The infinite time ruin probability is:

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 8(14)

$$\Psi(u) = P(R(t) < 0 \text{ for some } t \geq 0).$$

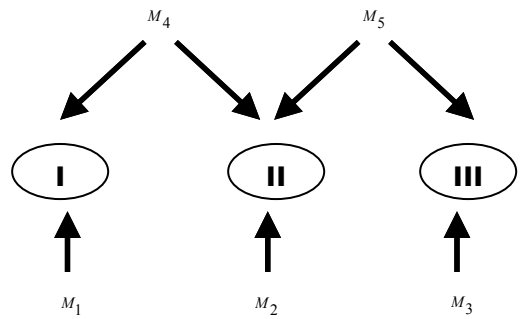

Fig. 1. Impact of risks factors on three classes of insurance business.

Source: own elaboration

The risk model (1) can be converted to a classical risk model [Ambagaspitiya 1998]:

$$R'(t) = u + ct - \sum_{i=1}^{N(t)} X_i',$$

where $\{N(t)\}_{t\geq 0}$ is a Poisson process with intensity $\lambda = \sum_{i=1}^{5} \lambda_i$ and $\{X_i'\}_{i=1}^{\infty}$ is a sequence of independent new claim sizes with the probability function given by:

$$f_{X'}(x) = \frac{1}{\lambda} \sum_{i=1}^{5} \lambda_i (f_{X_1}^{*(a_{1i})} * f_{X_2}^{*(a_{2i})} * f_{X_3}^{*(a_{3i})})(x) \quad \left(f_{X_i}^{*(0)} \equiv 1\right).$$

In a case of a classical risk model there are known a lot of methods of calculation or approximation of the infinite time ruin probability [Asmussen 2000; Rolski et al. 1998]. In further analysis we use De Vylder's approximation, which is given by [Rolski et al. 1998]:

$$\Psi(u) \approx \frac{1}{1+\bar{\theta}} \exp\left(-\frac{\bar{\theta}\beta u}{1+\bar{\theta}}\right),$$

where $\overline{\beta} = \frac{3m_2}{m_3}, \overline{\lambda} = \frac{9\lambda m_2^3}{2m_3^2}, \overline{\theta} = \frac{2mm_3}{3m_2^2}\theta$ and $m_k = EX^k$.

It is worth noticing that in actuarial literature we distinguish between light- and heavy-tailed claim sizes distributions [Rolski et al. 1998]. Light-tailed distribution with c.d.f. $F_X(x)$ means that there exist constants $a > 0, b > 0$ such that the tail $\overline{F_X}(x) = 1 - F_X(x) \le a\exp(-bx)$ or equivalently there exists $z > 0$ such that the moment generating function $M_X(z) < \infty$. If any distribution is not light-tailed, it is said to be heavy-tailed.

We consider four following cases of an impact of outside risk factors in risk model (1) on the infinite time ruin probability:

1) $\lambda_1 = 30, \lambda_2 = 60, \lambda_3 = 30$ and assume only in this case that there are no outside risk factors affecting the classes of business;

2) $\lambda_1 = 20, \lambda_2 = 40, \lambda_3 = 20, \lambda_4 = 10, \lambda_5 = 10$;

3) $\lambda_1 = 10, \lambda_2 = 20, \lambda_3 = 10, \lambda_4 = 20, \lambda_5 = 20$;

4) $\lambda_4 = 30, \lambda_5 = 30$ and assume in this case that there is no affect of underlying risk factors in each class of business.

We analyse an impact of outside risk factor considering four above cases and different types of claim sizes distributions. In figures 2 and 3 there are diagrams of the ruin probability as a function of initial capital u. In figure 2 there are diagrams in the situation where in each class there are light-tailed distributed claim sizes, i.e. $X_{1j} \sim \text{Gamma}(0.9, 1.1)$, $X_{2j} \sim \text{Gamma}(0.75, 0.8)$ and $X_{3j} \sim \text{Gamma}(0.5, 0.75)$. In figure 3 there are diagrams in the situation where all of claim sizes are heavy-tailed distributed, i.e. $X_{1j} \sim \text{Weibull}(1.1, 0.8)$, $X_{2j} \sim \text{Pareto}(5.1, 3.1)$ and $X_{3j} \sim \text{Burr}(6.5, 9.2, 0.9)$. In each considered situation we observe that with the growing strength of outside risks factors affecting three classes of business in considered cases 1-4 the ruin probability is increasing. We can also notice that for the fixed value of initial capital u increase of ruin probability in each case is almost proportional to the increase of strength of outside risk factors. A similar situation is in case of ruin probability in finite time, which was investigated earlier by us.

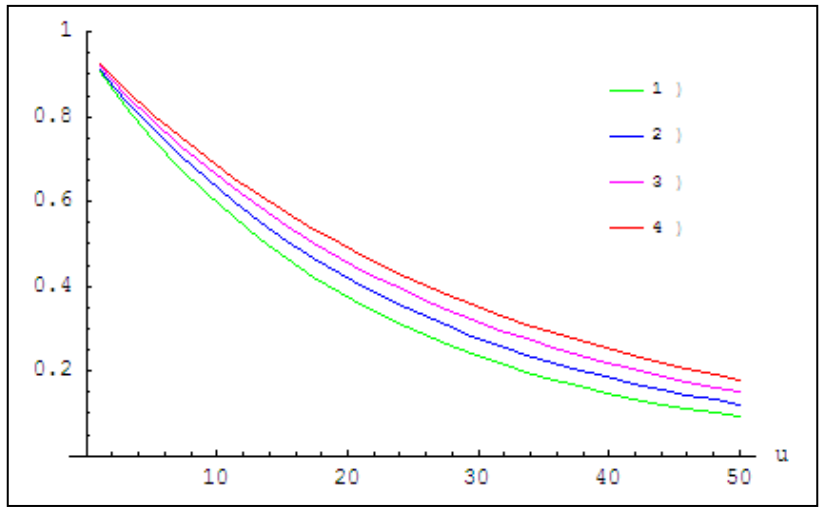

Fig. 2. Ruin probability in case of light-tailed claim sizes distributions ($\theta = 0.05$)

Source: own elaboration.

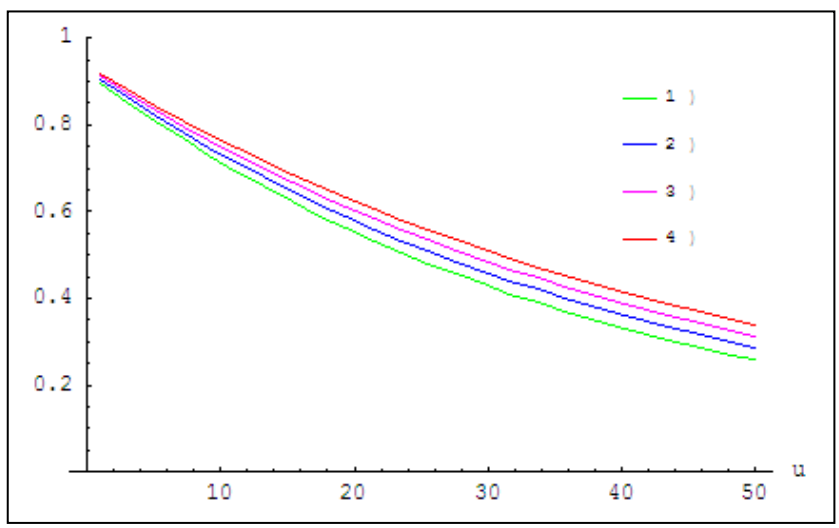

Fig. 3. Ruin probability in case of heavy-tailed claim sizes distributions ($\theta = 0.05$)

Source: own elaboration.

References

Ambagaspitiya R.S., *On the distribution of a sum of correlated aggregate claims*, "Insurance: Mathematics and Economics", 23 (1998), pp. 15-19.

Asmussen S., *Ruin probabilities*, Advanced Series on Statistical Science & Applied Probability, 2000.

Rolski T., Schmidli H., Schmidt V., Teugles J., *Stochastic processes for insurance and finance*, Wiley, New York 1998.

OBTAINING MISSING NOT AT RANDOM DISTRIBUTION'S PARAMETERS FROM MICROECONOMIC SURVEYS

Christian Westphal (University of Marburg)

1. Motivation and model

Today a remaining problem when dealing with missing data is the problem of missing not at random data (MNAR). A variable Y that is missing not as random is defined as $\Pr(R = 1|Y,\cdot) \neq \Pr(R = 1|\cdot)$, where R is indicating response ($R = 1$) or nonresponse ($R = 0$) and the dot stands for everything else besides Y[2].

Dealing with MNAR data depends on modelling the missingness[3], and therefore has not received much attention in the general statistical analysis of missing data problems. As many of these problems are from the field of microeconomics[4], I will give a general model for all of these problems. The problem will be illustrated by the example of income surveys where income is the MNAR variable. This example has proven to be a reliable point in any discussion and there exists

[2] See [Rubin 1976; Little, Rubin 2002, p. 12].

[3] [Rubin 1976, p. 589; Little, Rubin 2002, chapter 15].

[4] For a recent summary see [Simmons, Wilmot 2004]. Philipson [2001] is quite different from the general conclusion of the former article in that Philipson gets very clear results from a postpaid incentive albeit from a very specific population.

prior information about the distribution(s) of income[5] and the behaviour of survey participants when asked about their income[6].

The model is as in [Westphal 2009]. In that model I required a postpaid incentive[7]. This requirement shall be abandoned for a more generalized approach[8]. Instead, incentives shall be considered as "shocks" to an individual's utility. It does not matter whether the incentive is postpaid or prepaid. All that matter is that the incentive influences the survey participant towards response regardless of its time of payment[9].

2. Findings in simulation

I will illustrate my findings by a simple simulation where there is a missing not at random variable Y ("income") and an artificial variable uncorrelated with the income Z ("incentive to respond").

Exactly "income" respective "incentive" are $Y, Z \overset{\text{i.i.d.}}{\sim} N(0, 1)$. The response decision depends, in accordance with the random utility binary choice model[10], on the value of the random utility function $U = Z - Y + E$ where E once again is a standard normal random variable. Response is given if $U > 0$.

This represents the following behaviour observed in the Philipson [2001] experiment: with rising income, the willingness to share information about this income decreases, and the willingness to give up the income information can be increased by paying an incentive. The results are illustrated in figure 1. The first column depicts the distribution of the (simulated and therefore known) response probabilities of

[5] Especially surveys; for summary see [Pinkovskiy, Sala-i-Martin 2009] as well as literature concerned with optimal income taxation where the income distribution is of crucial importance for the results.

[6] [Philipson 2001].

[7] Inspired by [Philipson 2001], incentive theory [Laffon, Martimort 2002] and rational choice theory.

[8] Abandoning the requirement for postpaid incentives does not contradict incentive theory or rational choice theory – we simply do not know what "rational" behaviour looks like, reacting to a prepaid incentive may be considered rational.

[9] This accomodates the findings in many prepaid incentive experiments, e.g. [Mehlkop, Becker 2007, section 3; Becker et al. 2007]; also see [Singer 2002] in: [Groves et al. 2002].

[10] As in [Greene 2008, p. 777].

the simulated agents who have been given an incentive within different ranges. The second column shows how with increasing incentive the response probabilities increase among the higher incomes.

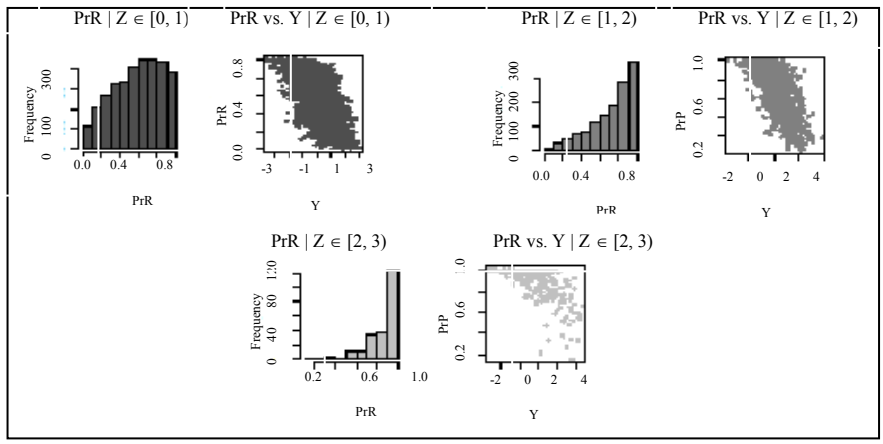

Fig. 1. Simulation results: response probabilities

Source: own elaboration.

The relationship between these two findings results in increasing incomes reported with increasing incentives[11]. This rather clearly shows there can be information about the reaction of the observed distribution of a MNAR variable in a variable uncorrelated to the MNAR variable of interest.

3. Extracting information

In section 2 we concluded there clearly lies information about the (simulated) income variable in the (simulated) incentive variable[12].

By the set up of the experiment we know there is no correlation between the incentive and the income. This directly leads us to the need of explaining the observations from section 2. This explanation is simple. By design there cannot be information about the income variable in the incentive variable. However, there is information about the observed incomes given the incentive variable in the incentive

[11] As in [Philipson 2001].
[12] The same seems to be true in the Philipson [2001] real-life experiment.

variable. This is due to the correlation of the response probability and the incentive variable and the correlation of the response probability and the income variable. This is illustrated by the distributional relations in figure 1.

Two independent random variables Y and Z which are arguments of a utility function $U(y, z; \cdot)$ do have a joint distribution $f_{Y,Z}(y, z)$ which is simply the product of the marginal distributions $f_Y(y) \cdot f_Z(z)$.[13] Via the transformation $(Y, Z) \rightarrow U$ a distribution of the random variable U, $f_U(u)$ is obtained where we add white noise ($N(0, \sigma^2)$) via the function $U(y, z; \cdot)$ to account for different response decisions given the same income and incentive level. Doing this transformation yields a joint density

$$f_{U,Z}(u, z) \tag{1}$$

as well.

From $f_Z(z)$ and (1) we can now find conditional densities

$$f_{U|Z}(u \mid Z \in \underline{\bar{z}}) = \frac{\int_{\underline{z}}^{\bar{z}} f_{U,Z}(u,z)\mathrm{d}z}{\int_{\underline{z}}^{\bar{z}} f_Z(z)\mathrm{d}z} \tag{2}$$

for any $\underline{\bar{z}} = [\underline{z}, \bar{z})$.

Therefore the integral from 0 to ∞ of u over (2) does give the expected percentage of respondents of all individuals receiving an incentive in the interval of $\underline{\bar{z}}$:

$$f_{U|Z}(u \mid Z \in \underline{\bar{z}}) = \int_0^{\infty} \frac{\int_{\underline{z}}^{\bar{z}} f_{U,Z}(u,z)\mathrm{d}z}{\int_{\underline{z}}^{\bar{z}} f_Z(z)\mathrm{d}z}\mathrm{d}u \tag{3}$$

The expected percentage of respondents cannot be obtained as the utility function's parameters and the parameters of $f_Y(y)$ are not known. However, we would expect these percentagesto roughly correspond to the observed response percentage of all individuals receiving an incentive in the interval $\underline{\bar{z}}$.

Now a system of equations using the observed response percentages as left handed side and the right handed side of equation (3) as

[13] Note that in the example given above $f_Z(z)$ and all its parameters are known.

right handed side can be set up. With assumptions about the utility function's form and an assumed/known parametric distribution family for Y the estimation of the unknown parameters in $f_Y(y)$ and $f_U(u)$ and thereby $U(\cdot)$ becomes feasible by employing a target function and – by guess – in most cases numerical methods.

This is in line with simpler suggestions[14] for the unbiased estimation of the distribution's parameters. I trade in [Pinkovskiy, Sala-i-Martin's 2009] suggested assumption of MNARness only at the top and the bottom of the distribution for the assumption of an underlying utility mechanism.

4. Conclusion

We see that with two rather general assumptions, namely (a) the decision to respond is based on a random utility binary choice and (b) the missing not at random variable can be modelled by a parametric distribution, information can be obtained about the missing not at random variable's distribution parameters.

Combined with good prior information about the distribution's and the utility function's form, good survey design and testing as proposed by Yuan [2009] might help solve the problem of MNAR data in microeconomic surveys in which the response or participation decision is at the discretion of the survey participant.

Furthermore, the incentive variable may be crucial for Yuan's testing method as the random incentive satisfies all requirements for the testing method proposed by Yuan and can be easily generated in a survey.

References

Becker R., Imhof R., Mehlkop G., *Effects of prepaid monetary incentives on the return of mail survey and self reporting about delinquency*, "Methoden – Daten – Analysen", 1(2) (2007), pp. 131-159.

Greene W.H., *Econometric analysis*, 6th ed., Pearson, New York 2008.

Groves R.M., Dillman D.A., Eltinge J.L., Little R.J.A. (eds.), *Survey nonresponse*, Wiley, New York 2002.

Laffont J.-J., Martimort D., *The theory of incentives*, Princeton, New York 2002.

[14] [Pinkovskiy, Sala-i-Martin 2009, p. 15].

Little R.J.A., Rubin D. B., *Statistical Analysis with Missing Data*, 2nd ed., Wiley, New York 2002.

Mehlkop G., Becker R., *The effects of monetary incentives on the response rate in mail surveys on self-reported criminal behavior*, "Methoden –Daten – Analysen", 1(1) (2007), pp. 5-24.

Philipson T., *Data markets, missing data, and incentive pay*, "Econometrica", 69(4) (2001), pp. 1099-1111.

Pinkovskiy M., Sala-i-Martin X., *Parametric estimations of the world distribution of income*, NBER Working Paper No. 15433 (2009, October).

Rubin D. B., *Inference and missing data*, "Biometrika", 63(3) (1976), pp. 581-592.

Simmons E., Wilmot A., *Incentive payments on social surveys: a literature review*, "Survey Methodology Bulletin", 53 (2004), pp. 1-11.

Singer E., *The use of incentives to reduce nonresponse in household surveys*, [in:] Groves et al., *Survey nonresponse*, Wiley, New York 2002.

Westphal C., *How economic modeling may help with MNAR data*, "Silesian Statistical Review", 13(7) (2009), pp. 96-101.

Yuan K.-H., *Identifying variables responsible for data not missing at random*, "Psychometrika", 74(2) (2009), 233-256.

APPROXIMATION OF THE STOP-LOSS PREMIUM ANALYSIS AND COMPARISON

Anna Nikodem (Wrocław University of Economics)

In order to protect oneself against large individual claims or against the fluctuation in the number of claims, the insurer takes out reinsurance cover for his insurance portfolio. The expected cost of this insurance is called the stop-loss premium and is defined as the expected value of the excess of an agreed retention d of the aggregate claim amount S accumulated during a certain time period. For the discrete and continuous case the premium can be obtained by the formulae respectively

$$\pi(d) = \begin{cases} \sum_{s>d} (s-d) f_S(s), \\ \int_{s>d} (s-d) f_S(s) ds, \end{cases} = \begin{cases} \sum_{s>d} [1 - F_S(s)], \\ \int_{d}^{\infty} [1 - F_S(s)] ds. \end{cases}$$

where $f_S(s)$ is the probability density or the probability function of aggregate claim amount and $F_S(s)$ is the distribution function of *S*. The aggregate claim amount *S* of the insurer portfolio is defined by a random sum of the random variables X_i

$$S = \begin{cases} \sum_{i=1}^{N} X_i & \text{for} \quad N \geq 1, \\ 0 & \text{for} \quad N = 0, \end{cases}$$

where *N* is a random variable, X_i is independent and identically distributed and *N* and X_i are independent.

In order to compute the stop-loss premium we have to determine the distribution of the aggregate claims amount. If the number of claims distribution is in the (a, b, 0) class, i.e. when the number of claims has Poisson, Binomial, Geometric, Negative binomial distribution and the individual claims distribution is discrete, we can use the Panjer's recursion (see i.e. [Klugman et al. 1998]). This method can be used also after discretization of the probability density function of individual claim size. Appling the recursive method, the stop-loss premium can be calculated recursively. For integer retention *d* we have (see i.e. [Klugman et al. 1998])

$$\pi(d) = E[(S-d)_+] = \pi(d-1) - [1 - F_S(d-1)],$$

where $\pi(0) = E(S)$. The stop-loss premium can be calculated recursively for the arithmetic individual claims distribution. For some fixed *h* and for retention *jh* we have

$$\pi(jh) = \pi\big((j-1)h\big) - h[1 - F_S\big((j-1)h\big)],$$

where $\pi(0) = E(S)$.

Unfortunately, in a lot of situations the distribution of aggregate claim amount is intractable. Then we need approximation. We can approximate the density by a function that uses the mean, variance and skewness of the aggregate claim amount, which are relatively easily obtained. In literature various approximations of the stop-loss premium are described (see [Kaas et al. 2001]). The compound distribution of *S* can be approximated by a normal distribution, when the

ŚLĄSKI PRZEGLĄD STATYSTYCZNY Nr 8(14)

skewness is equal to zero. The next approximation is NP-approximation, which can be used when the skewness of the aggregate claim amount is small, i.e. from zero to one. Since most aggregate claim distributions have roughly the same shape as the gamma distribution, the cumulative distribution function $F_S(s)$ can be also approximated by the gamma cumulative distribution function $G(s-x_0;\alpha,\beta)$. Comparing this approximation in examples we can observe that the translated gamma approximation is the best one, but this approximation cannot be used when the mean of number of claims is very small and the mean of the individual claim size is large. Besides some values of parameters of the compound distribution of S we cannot calculate the stop-loss premium using the translated gamma approximation. Hence the Haldane approximation is also considered (see [Pentikainen 1987]). For this approximation the stop-loss premium has a form of

$$\pi(d)=\int_{w(d)}^{\infty}(q(u)-d)\cdot\phi(u)du=\int_{w(d)}^{\infty}q(u)\cdot\phi(u)du-d\left[1-\Phi(w(d))\right],$$

where $q(u)=E(S)\cdot\left[u\cdot\sigma_Y+E(Y)\right]^{1/h}$ and

$$h=1-\frac{\gamma_S E(S)}{3D(S)},$$

$$E(Y)=1-\frac{1}{2}h(1-h)\cdot\left[1-\frac{1}{4}(2-h)(1-3h)v^2\right]\cdot v^2,$$

$$\sigma_Y=h\cdot v\cdot\sqrt{\left[1-\frac{1}{2}(1-h)(1-3h)\cdot v^2\right]}\quad\text{for}\quad v=\frac{D(S)}{E(S)}.$$

Comparing the Haldane approximation and the translated gamma approximation we have that both approximations give the same results. The Haldane approximation is better for smaller skewness. This approximation can be used for those values of parameters of compound distribution of S for which we cannot use the translated gamma approximation, in a condition where for those parameters σ_Y is positive.

Using the translated gamma approximation and the Haldane approximation to compute the stop-loss premium, when the aggregate claim amount has the compound Poisson distribution with parameter

λ, we have that for a larger parameter λ those approximations are better. In practice the mean of the number of claims is rather small. In this situation we can use the Gaussian exponential approximation (see [Hurlimann 2003]). For this approximation the stop-loss premium has the following form

$$\pi(s)=\begin{cases} E(S)\cdot\exp\left(-\alpha\cdot\dfrac{s}{E(S)}-\dfrac{1}{2}\left(\dfrac{\alpha}{\eta}\cdot\dfrac{s}{E(S)}\right)^2\right), & 0\le s\le s_0,\\ \pi(s_0)\cdot\exp\left(-\left(\dfrac{s-s_0}{m(s_0)}\right)\right), & s\ge s_0, \end{cases}$$

where the unknown parameter and threshold can be found by using the following formulas: first we have to find the parameter η from equation $2\dfrac{\eta_0}{\alpha}\cdot\dfrac{\Phi(\eta_0)}{\phi(\eta_0)}=1+v^2$ and next we obtain the threshold from $s_0\cdot E(S)=\dfrac{1}{2}\left(1+\sqrt{1+4\dfrac{\eta_0}{\alpha}}\right)$, where $\alpha=1-e^{-\lambda}$. In contrast to previous approximations this approximation is better when the Poisson parameter is small. The approximation gives good results for a small λ parameter, even though the skewness is large.

References

Daykin C.D., Pentikainen T., Pesonen M., *Practical risk theory for actuaries*, Chapman & Hall, London 1994.

Haldane J.B.S., *The approximate normalization of a class of frequency distributions*, "Biometrica", 29 (1938), pp. 392-404.

Hurlimann W., *A Gaussian exponential approximation to some compound Poisson distributions*, "Astin Bulletin", 33(1) (2003), pp. 41-55.

Kaas R., Goovaerts M., Dhaene J., Denuit M., *Modern actuarial risk theory*, Kluwer Academic Publishers, Boston 2001.

Klugman S.A., Panjer H.H., Willmot G.E., *Loss models. From data to decision*, John Wiley & Sons, New York 1998.

Ostasiewicz W. (red.), *Modele aktuarialne*, AE, Wrocław 2000.

Panjer H.H., Willmot G.E., *Insurance risk models*, Society of Actuaries, Schaumburg 1992.

Pentikainen T., *Approximative evaluation of the distribution function of aggregate claims*. "Astin Bulletin", 17(1) (1987), pp. 15-39.

Reijnen R., Alberts W., Kallenberg W.C.M., *Approximation for stop-loss reinsurance premiums*, "Insurance: Mathematics and Economics", 36 (2005), pp. 237-250.