SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2024
ORGANIZATION AND MANAGEMENT SERIES NO. 198

APPLICATION OF THE ECM ALGORITHM TO THE ESTIMATION OF THE LIKELIHOOD FUNCTION IN FINANCIAL AUDITING

Grzegorz SITEK

Uniwersytet Ekonomiczny w Katowicach, Katedra Statystyki, Ekonometrii i Matematyki;
grzegorz.sitek@ue.katowice.pl, ORCID: 0000-0002-7191-8631

Purpose: The book amounts are treated as values of a random variable whose distribution is a mixture of the distributions of the correct amount and the true amount contaminated by error. The mixing coefficient is equal to the proportion of the items with non-zero errors amounts. Below we consider a problem of testing appropriately formulated statistical hypotheses about admissibility of the total or the mean accounting errors. Hypotheses can be verified by the likelihood ratio test. In this paper, we show how to estimate parameters of the likelihood function.
Design/methodology/approach: The book amounts are treated as values of a random variable whose distribution is a mixture of the distributions of the correct amount and the true amount contaminated by error. The mixing coefficient is equal to the proportion of the items with non-zero errors amounts. Below we consider a problem of testing appropriately formulated statistical hypotheses about admissibility of the total or the mean accounting errors. Hypotheses can be verified by the likelihood ratio test. In this paper, we show how to estimate parameters of the likelihood function.
Findings: The work presents formulas for the parameters of the likelihood function. These parameters were obtained using the ECM algorithm.
Originality/value: The problem of estimating the average audit error is very common in economic research. A method for estimating the average audit error based on the likelihood function was proposed. The parameters of the likelihood function were estimated using the ECM algorithm.
Keywords: ECM algorithm, likelihood ratio test, mixture of probability distribution.
Category of the paper: Research paper.

1. Introduction

Letting X_i denote the book amount of the ith item in the account $X_U = \sum_{i=1}^{N} x_i$ called the population book amount, at regular periods, an auditor samples n line items from the account and compares them against correct amounts. Therefore, let Y_i denote the audited amounts for the ith line item and let $\tau_i = X_i - Y_i$ denote the error amount. Notice that the total book amount is known to the auditor. The fundamental problem is the problem of constructing confidence limits for mean or totals in finite populations, when the underlying distribution is highly skewed and contains a substantial proportion of zero values. This situation is often encountered in statistical applications such as statistical auditing, reliability and insurance. The most distinctive feature of accounting data is the large proportion of line items without error, while an audit sample may not yield any nonzero error amounts. For analyses of such data, which most observations are zero, the classical interval estimation of the total error amount based on the asymptotic normality of the sampling distribution is not reliable. In auditing practice, auditors are often more interested in obtaining lower or upper confidence limits than in obtaining two-sided confidence intervals. Independent public accountants are very often concerned in estimating the lower confidence bound for the total audited amount. An auditor wants to avoid overestimating this bound because of the potential legal liability that may follow from this. Stringer (1963) and Kaplan (1973) shows, that the accounting populations are highly positively skewed, and there is considerable diversity in the characteristics of error amounts in accounting populations across the accounting subsystem. There are several distributions that also exhibit the same form of the distribution observed in accounting populations. These include the Gamma, Log-normal, Weibull, and Beta distributions. The error rates are usually very low, which render many existing statistical procedures inappropriate for estimating and hypothesis testing of error rates and error amounts. There are two main types of audit tests for which the acquisition of information can profitably make use of statistical sampling. The first audit test, collecting data to determine the rate of procedural errors of a population of transactions is called a compliance test. The second, collecting data for evaluating the aggregate monetary error in the stated balance, is called a substantive test of details. Inference on the total error amount is usually based on confidence intervals. Of course, they are related to testing problems. The decision-making process in auditing is treated as a problem of testing statistical hypotheses about admissibility of the total or the mean accounting errors. This approach lets us control not only significance level (risk of incorrect rejection), but also probability of the type II error appearing (risk of incorrect acceptance). Substantive tests of details are concerned with the examination of the correctness of recorded monetary values in a financial statement. These tests provide direct evidence about the accuracy of total recorded monetary values. The auditor either applies substantive tests of detail extensively, or applies compliance tests to see if reliance on those controls are efficient and effective in reducing the tendency of material error in accounts. In compliance tests, the variable of interest is an error rate (proportion of transactions for which the internal control operates wrongly). Samples of transactions are used to make inferences about the error rate.

Wywiał (2018) proposed following model. Let U be the population of accounting documents of size N with given accounting totals. Some of the documents contain errors. In the population U there are given accounting totals (values) x_i for each element $i \in U$. Let $x^T = [x_1, \dots, x_N] \in R_+^N$ be the observation of a random vector $X^T = [X_1, \dots, X_N]$. We denote the true book values (without errors) by y_i, $i \in U$ and let $y^T = [y_1, \dots, y_N] \in R_+^N$ be the random vector observation $Y^T = [Y_1, \dots, Y_N]$. Vector of accounting values contaminated by errors $w^T = [w_1, \dots, w_N]$ will be an observation of the random vector $W^T = [W_1, \dots, W_N]$. Finally, let $Z^T = [Z_1, \dots, Z_N]$, where $Z_i = 0$ $(Z_i = 1)$ if $X_i = Y_i$ $(X_i \neq Y_i)$.

In practice, all values of X are known before auditing process. Observations x of X are treated as a specific auxiliary data. Auditing process leads to observation of values Z_i, Y_i and W_i,$i \in U$. Let $\overline{X} = \frac{1}{N}\sum_{i \in U} X_i$, $\overline{Y} = \frac{1}{N}\sum_{i \in U} Y_i$, $\overline{W} = \frac{1}{N}\sum_{i \in U} W_i$. Their values will be denoted $\overline{x}$, $\overline{y}$, $\overline{w}$.

Let an auditor arbitrary selects the sample s of the size n from U. Hence, X_s is the subvector of X, $n \leqslant N$. The random vector X_s is observed in s where the objects are controlled. After the auditing process the sample s is split into two disjoint sub-samples s_0 and s_1 where $s = s_0 \cup s_1$. The set s_1 is of size $n_1 = k$ and the set s_0 is of size $n_0 = n - k$. In the sub-sample s_0 here are observed accounting amounts without errors. Before auditing process, we have observations of the following data:

$$X = (X_i \colon i \in U) = (X_s, X_{U-s}),$$

where

$$X_s = (X_i \colon i \in s), X_{U-s} = (X_i \colon i \in U - s).$$

After the auditing process, we have observations of the following data:

$$T_U = (T_s, X_{U-s}), T_s = (X_i, Z_i) \colon i \in s) = (Y_{s_0}, W_{s_1}).$$

Values T, T_U, X, X_s, X_{U-s}, Y_{s_0} and W_{s_1} are denoted respectively by t, t_U, x, x_s, x_{U-s}, y_{s_0} and w_{s_1}. In the following work, we assume that $y_{s_0} = y_s$ and $w_{s_1} = w_s$.

Let $\tau = E(\overline{X} - \overline{Y})$ be the expected mean accounting error. Audit purpose is inference on τ or on the expected total accounting error $N\tau = E(\sum_{i \in U} X_i - \sum_{i \in U} Y_i)$. In particular, when we assume that τ_0 is the admissible mean accounting error then the inference reduces to testing the following hypothesis:

$$H_0 \colon \tau = \tau_0, H_1 \colon \tau > \tau_0, \tag{1}$$

where τ_1 is unadmissible level of the mean accounting error.

2. A mixture of three probability distributions as a model for generating accounting values

Let $F_0(y|\theta_0)$ be the probability distribution function of the random variable Y, whose values are true accounting and $\theta_0 \in \Theta_0$ where Θ_0 is the parameter space. The distribution function of W is denoted by $F_1(w|\theta_1)$, where $\theta_1 \in \Theta_1$. Moreover, let $\Theta = \Theta_0 \cup \Theta_1$. We assume that an accouting errors appears with probability p. We can write $Z = 1$ when an accouting error occurs $P(Z = 1) = p$ and $Z = 0$ when it does not occur $P(Z = 0) = 1 - p$. According to the well-known total probability thorem we have: $F(x) = F(x|Z=0)P(Z=0) + F(x|Z=1)P(Z=1)$ and finally

$$F(x|\theta) = (1-p)F_0(x|\theta_0) + pF_1(x|\theta_1), \quad (2)$$

where $\theta = \theta_0 \cup \theta_1$ and $\theta \in \Theta = \Theta_0 \cup \Theta_1$ is the parameter space.

Hance, the probability distribution of the observed accounting amounts is a mixture of the distribution function $F_0(x|\theta_0)$ of the true amounts and the distribution function $F_1(x|\theta_1)$ of the amounts contamineted by errors. When the random variables Y and W are continuous, by differentiating both sides of equation (2) we have

$$f(x|\theta) = (1-p)f_0(x|\theta_0) + pf_1(x|\theta_1). \quad (3)$$

Therefore, the probability density of the observed accouting amounts is a mixture of density $f_0(x|\theta_0)$ of the true amounts and density $f_1(x|\theta_1)$ of the amounts contaminated by errors. Let R and Y be independent and R is the accounting error. Hence $W = Y + R$, $X = Y + ZR$, $X = (1-Z)Y + ZW$.

The basic moments of the random variable X are:

$$E(X) = (1-p)E(X|Z=0) + pE(X|Z=0) = (1-p)E(Y) + pE(W). \quad (4)$$

$$V(X) = p(1-p)(E(W) - E(Y))^2 + pV(W) + (1-p)V(Y). \quad (5)$$

In the context of model approach our purpose is to test the hypothesis about the expected value of the following difference of the sum of observed in the population accounting amounts and the sum of the true values.

$$\tau = E(\overline{X} - \overline{Y}) = E(X) - E(Y) = p(E(W) - E(Y)), \quad (6)$$

or

$$\tau(\theta) = E(X|\theta) - E(Y|\theta_0) = p\big(E(W|\theta_1) - E(Y|\theta_0)\big). \quad (7)$$

The well-known gamma probability distribution we denote by $G(\alpha, \beta)$ where parameters $\alpha > 0$ and $\beta > 0$ are called scale and shape parameters. The shape of gamma density distribution does not depend on the scale parameter because its skewness and kurtosis coefficients are equal to $\frac{2}{\sqrt{\beta}}$ and $\frac{6}{\beta}$ respectively. Wywiał (2018) considered the model based on a mixture of gamma distributions. Let $Y \sim G(a, c)$ and $R \sim G(b, c)$ be independent random variables. The advantage of this model is that the density function for the sum of gamma distributions can be determined. Based on the above assumption, the random variable $W = Y + R \sim G(a + b, c)$. Using the previous considerations, we obtain

$$f(x|a,b,c) = (1-p)f_0(x|a,c) + pf_1(x|a,b,c), \quad (8)$$

where

$$f_1(x|a,b,c) = \frac{c^{a+b}}{\Gamma(a+b)} x^{a+b-1} e^{-cx}, x > 0, \quad (9)$$

$$f_0(x|a,c) = \frac{c^a}{\Gamma(a)} x^{a-1} e^{-cx}, x > 0. \quad (10)$$

From formulae (4) and (5) we obtain

$$E(X) = \frac{a+pb}{c}, V(X) = \frac{a+pb+p(1-p)b^2}{c^2}. \quad (11)$$

For more on the use of gamma decomposition to model accounting values, see the articles by Frost and Tamura (1986) or Tamura (1988). The book amounts are treated as values of a random variable which distribution is a mixture of the distributions of correct amount and the distribution of the true amount contaminated by error. Distributions of correct amount and true amount contaminated by error are right-skewed because small book amounts are more frequent than large book amounts. It is convenient to assume that the book values are additive function of true accounting amounts and accounting errors. Hence, we can expect that the above proposed quite simple model describes accounting data well.

3. Testing on the basis of the likelihood function

Wywiał (2018) considered the following likelihood function:

$$L(t|\theta) = L(t_s|\theta)L(x_{U-s}|\theta), \tag{12}$$

where

$$L(t_s|\theta) = \prod_{i\in s} p^{z_i} f_1^{z_i}(x_i|\theta_1)(1-p)^{1-z_i} f_0^{1-z_i}(x_i|\theta_0), \tag{13}$$

$$L(x_{U-s}|\theta) = \prod_{i\in U-s} f(x_i|\theta). \tag{14}$$

If $z_i = 0 (z_i = 1)$ then $x_i = y_i (x_i = w_i)$. The logarithm of the likelihood function (12) is given by the formula

$$\begin{aligned} l(t|\theta) = {} & k\ln(p) + (n-k)\ln(1-p) + \\ & \sum_{i\in s_1} \ln(f_1(x_i|\theta_1)) + \sum_{i\in s_0} \ln(f_0(x_i|\theta_0)) + \sum_{i\in U-s} \ln(f(x_i|\theta)). \end{aligned} \tag{15}$$

Hypotheses (1) can be verified by means of the well-known likelihood ratio test on the basis of the following statistic:

$$\lambda = \frac{sup_{\theta\in\Theta,\tau(\theta)=\tau_0} L(D|\theta)}{sup_{\theta\in\Theta} L(D|\theta)}. \tag{16}$$

We can expect that when hypothesis H_0 is true and $n, N, N-n, n_0$ and $n-n_0$ are sufficiently large then statistic $t = -2ln(\lambda)$ is well approximated by the chi-square distribution with one degree of freedom (Silvey 1959). Hypothesis H_0 is rejected if t is significantly large.

3.1. Mixture of two gamma distributions with the same scale parameter

We consider a mixture of two gamma distributions with the same scale parameter denoted by the symbol $\boldsymbol{c}$. For this mixture, the logarithm of the likelihood function (15) is written as

$$\begin{aligned} l = {} & k\ln(p) + (n-k)\ln(1-p) + Na\ln(c) + kb\ln(c) - \\ & k\ln\big(\Gamma(a+b)\big) - (n-k)\ln\big(\Gamma(a)\big) + (a-1)\sum_{j\in U} \ln\big(x_j\big) + b\sum_{j\in s_1} \ln\big(x_j\big) \\ & -c\sum_{j\in U} x_j + \sum_{j\in U-s} \ln\Big(\frac{1-p}{\Gamma(a)} + \frac{p(cx_j)^b}{\Gamma(a+b)}\Big). \end{aligned} \tag{17}$$

In order to determine the parameter estimators a, b, c and p we calculate the derivatives of equation (17). In financial auditing, the case $U = s$ is practically absent. The sample size n is usually a few percent of the population size N. These estimators can be obtained using the ECM algorithm. The ECM algorithm (Meng and Rubin (1993)) is an efficient combination of the CM and EM algorithms (McLachlan and Peel, (2000)). It replaces the maximization step of EM with a set of conditional maximization steps, and thus splits a difficult maximization problem into several easier ones.

3.2. ECM algorithm

Based on the previous expressions we write the following likelihood function

$$\boldsymbol{L = \prod_{i=1}^{k} p \frac{c^{(a+b)} w_{s_i}^{(a+b-1)} e^{-c w_{s_i}}}{\Gamma(a+b)} \prod_{i=k+1}^{n} (1-p) \frac{c^{a} y_{s_i}^{(a-1)} e^{-c y_{s_i}}}{\Gamma(a)}} \\ \boldsymbol{\prod_{i=n+1}^{N} [p \frac{c^{(a+b)} x_{(U-s)_i}^{(a+b-1)} e^{-c x_{(U-s)_i}}}{\Gamma(a+b)} + (1-p) \frac{c^{a} x_{(U-s)_i}^{(a-1)} e^{-c x_{(U-s)_i}}}{\Gamma(a)}]}. \tag{18}$$

We then determine the logarithm of the likelihood function (18). To simplify the notation, we introduce the variable $d = a + b$

$$l = k\ln(p) + (n-k)\ln(1-p) + kd\ln(c) - c\sum_{i=1}^{k} w_{s_i} \\ +(d-1)\sum_{i=1}^{k}\ln(w_{s_i}) - k\ln(\Gamma(d)) + (n-k)a\ln(c) - \\ c\sum_{i=k+1}^{n} y_{s_i} + (a-1)\sum_{i=k+1}^{n}\ln(y_{s_i}) - (n-k)\ln(\Gamma(a)) + \\ \sum_{i=n+1}^{N}\ln[p\frac{c^d}{\Gamma(d)} x_{(U-s)_i}^{d-1} e^{-c x_{(U-s)_i}} + (1-p)\frac{c^a}{\Gamma(a)} x_{(U-s)_i}^{a-1} e^{-c x_{(U-s)_i}}]. \tag{19}$$

To determine the parameter estimators a, d, c and p it is necessary to calculate the partial derivatives of the function (19) and solve the system of equations numerically. These estimators can be obtained using the ECM algorithm.

Step E.

$$\hat{\gamma}_i = \frac{\hat{p}\frac{c^d}{\Gamma(d)} x_{(U-s)_i}^{d-1} e^{-c x_{(U-s)_i}}}{\hat{p}\frac{c^d}{\Gamma(d)} x_{(U-s)_i}^{d-1} e^{-c x_{(U-s)_i}} + (1-\hat{p})\frac{c^a}{\Gamma(a)} x_{(U-s)_i}^{a-1} e^{-c x_{(U-s)_i}}}. \tag{20}$$

Step CM.

The logarithm of the likelihood function (20) is written as follows

$$Q(a,d,c,p) = \sum_{i=n+1}^{N} ((1-\hat{\gamma}_i)(a\ln(c) + (a-1)\ln(x_{(U-s)_i}) - \ln(\Gamma(a)) - c x_{(U-s)_i}) + \\ \hat{\gamma}_i(d\ln(c) + (d-1)\ln(x_{(U-s)_i}) - \ln(\Gamma(d)) - c x_{(U-s)_i}) + \hat{\gamma}_i\ln(p) + (1-\hat{\gamma}_i)\ln(1-p)) \\ k\ln(p) + (n-k)\ln(1-p) + kd\ln(c) - c\sum_{i=1}^{k} w_{s_i} + (d-1)\sum_{i=1}^{k}\ln(w_{s_i}) - k\ln(\Gamma(d)) + \\ (n-k)a\ln(c) - c\sum_{i=k+1}^{n} y_{s_i} + (a-1)\sum_{i=k+1}^{n}\ln(y_{s_i}) - (n-k)\ln(\Gamma(a)). \tag{21}$$

Determining the partial derivatives of a function (21)

$$\frac{\partial Q}{\partial a} = \sum_{i=n+1}^{N} (1-\hat{\gamma}_i)(-\psi^{(0)}(a) + \ln(c) + \ln(x_i)) + \\ \sum_{i=k+1}^{n}\ln(y_i) - \psi^{(0)}(a)(n-k) + \ln(c)(n-k) = 0. \tag{22}$$

$$\frac{\partial Q}{\partial d} = \sum_{i=n+1}^{N} \hat{\gamma}_i(-\psi^{(0)}(d) + \ln(c) + \ln(x_{(U-s)_i})) + \sum_{i=1}^{k} \ln(w_{s_i}) - \psi^{(0)}(a)k + \ln(c)k = 0. \quad (23)$$

The previous two equations can be solved numerically using the Newton-Raphson method.

$$\frac{\partial Q}{\partial c} = \sum_{i=n+1}^{N} (1-\hat{\gamma}_i)\left(\frac{a}{c} - x_{(U-s)_i}\right) + \sum_{i=n+1}^{N} \hat{\gamma}_i\left(\frac{d}{c} - x_{(U-s)_i}\right) - \sum_{i=k+1}^{n} y_{s_i} - \sum_{i=1}^{k} w_{s_i} + \frac{a(n-k)}{c} + \frac{dk}{c} = 0. \quad (24)$$

After transformations, we obtain the estimator of the parameter c

$$\hat{c} = \frac{d(k+\sum_{i=n+1}^{N}\hat{\gamma}_i) + a(n-k+\sum_{i=n+1}^{N}(1-\hat{\gamma}_i))}{\sum_{i=k+1}^{n} y_{s_i} + \sum_{i=1}^{k} w_{s_i} + \sum_{i=n+1}^{N} x_{(U-s)_i}}. \quad (25)$$

$$\frac{\partial Q}{\partial p} = \sum_{i=n+1}^{N} -\frac{1-\hat{\gamma}_i}{1-p} + \sum_{i=n+1}^{N} \frac{\hat{\gamma}_i}{p} - \frac{n-k}{1-p} + \frac{k}{p} = 0, \quad (26)$$

$$\hat{p} = \frac{k+\sum_{i=n+1}^{N}\hat{\gamma}_i}{N}. \quad (27)$$

Let a_{y_s}- is the shape parameter obtained by the maximum likelihood method from the observation vector y_s, c_{y_s}- is the scale parameter obtained by the maximum likelihood method from the observation vector y_s, d_{w_s}- is the shape parameter obtained by the maximum likelihood method based on a vector of observations w_s, c_{w_s}-is a scale parameter derived by the maximum likelihood method from a vector of observations w_s. The estimators thus obtained constitute a vector of starting parameters for the ECM algorithm $(a_0, d_0, c_0, p_0) = (a_{y_s}, d_{w_s}, \frac{c_{y_s}+c_{w_s}}{2}, \frac{k}{n})$.

If the number of erroneous values k in the drawn sample is small, it may not be possible to use the maximum likelihood method for estimating the starting parameters. In this case, the method of moments can be used. The above-described ECM algorithm will not work if the $k = 0$.

3.3. Parameter estimation of a mixture of gamma distributions when the hypothesis H_0 is true

In this case, we estimate the parameters a_{H0}, c_{H0} i p_{H0}, because the parameter b then a function of the parameters τ_0, c and p ($b = \frac{c\tau_0}{p}$).

$$L = \prod_{i=1}^{k} p\frac{c^{(a+\frac{c\tau_0}{p})} w_{s_i}^{(a+\frac{c\tau_0}{p}-1)} e^{-cw_{s_i}}}{\Gamma(a+\frac{c\tau_0}{p})} \prod_{i=k+1}^{n} (1-p)\frac{c^a y_{s_i}^{(a-1)} e^{-cy_{s_i}}}{\Gamma(a)} \prod_{i=n+1}^{N} [p\frac{c^{(a+\frac{c\tau_0}{p})} x_{(U-s)_i}^{(a+\frac{c\tau_0}{p}-1)} e^{-cx_{(U-s)_i}}}{\Gamma(a+\frac{c\tau_0}{p})} + (1-p)\frac{c^a x_{(U-s)_i}^{(a-1)} e^{-cx_{(U-s)_i}}}{\Gamma(a)}]. \quad (28)$$

We then determine the logarithm of the likelihood function (28)

$$l = k\ln(p) + (n-k)\ln(1-p) + k(a + \frac{c\tau_0}{p})\ln(c) - c\sum_{i=1}^{k} w_{s_i}$$
$$+(a + \frac{c\tau_0}{p} - 1)\sum_{i=1}^{k} \ln(w_{s_i}) - k\ln(\Gamma(a + \frac{c\tau_0}{p})) + (n-k)a\ln(c) -$$
$$c\sum_{i=k+1}^{n} y_{s_i} + (a-1)\sum_{i=k+1}^{n} \ln(y_{s_i}) - (n-k)\ln(\Gamma(a)) + \quad (29)$$
$$\sum_{i=n+1}^{N} \ln[p\frac{c^{a+\frac{c\tau_0}{p}}}{\Gamma(a+\frac{c\tau_0}{p})} x_{(U-s)_i}^{a+\frac{c\tau_0}{p}-1} e^{-cx_{(U-s)_i}} + (1-p)\frac{c^a}{\Gamma(a)} x_{(U-s)_i}^{a-1} e^{-cx_{(U-s)_i}}].$$

Parameter estimators a, c and p can be determined by maximising the function (29) in the R environment with the nlm function. The starting parameters for the nlm function are chosen as follows $(a_0, c_0, p_0) = (a_{y_s}, \frac{c_{y_s}+c_{w_s}}{2}, \frac{k}{n})$.

4. Simulation study of the power of the likelihood ratio test

In the first step, the critical test values were determined by simulation. Then, in a second step, the test power was determined in a simulation manner for the previously determined critical values. For a mixture of gamma distributions, the following hypothesis was tested:

$$H_0 : \tau = \tau_0 = 50, H_1 : \tau = \tau_1 > \tau_0.$$

Table 1.
Power of the likelihood ratio test. A mixture of gamma distributions

τ_1	p	α	N					
			80	**120**	**160**	**200**	**400**	**600**
$1{,}01\tau_0$	0,1	0,05	0,047	0,05	0,05	0,05	0,051	0,052
		0,1	0,098	0,101	0,099	0,102	0,101	0,107
		0,2	0,203	0,199	0,190	0,203	0,206	0,206
	0,2	0,05	0,045	0,04	0,047	0,05	0,05	0,062
		0,1	0,092	0,09	0,102	0,095	0,10	0,108
		0,2	0,187	0,189	0,203	0,204	0,2	0,205
	0,3	0,05	0,047	0,049	0,05	0,051	0,055	0,064
		0,1	0,1	0,1	0,099	0,103	0,106	0,106
		0,2	0,22	0,22	0,209	0,205	0,207	0,206
$1{,}03\tau_0$	0,1	0,05	0,06	0,058	0,05	0,057	0,056	0,057
		0,1	0,102	0,111	0,102	0,11	0,109	0,11
		0,2	0,207	0,2	0,199	0,203	0,207	0,209
	0,2	0,05	0,05	0,04	0,048	0,057	0,05	0,053
		0,1	0,096	0,097	0,093	0,101	0,104	0,107
		0,2	0,187	0,191	0,2	0,197	0,206	0,211
	0,3	0,05	0,05	0,046	0,049	0,053	0,05	0,056
		0,1	0,097	0,094	0,099	0,107	0,105	0,113
		0,2	0,2	0,19	0,199	0,203	0,208	0,209

Cont. table 1.

$1{,}05\tau_0$	0,1	0,05	0,05	0,059	0,054	0,051	0,054	0,054
		0,1	0,101	0,112	0,114	0,119	0,116	0,119
		0,2	0,202	0,2	0,206	0,206	0,21	0,211
	0,2	0,05	0,045	0,05	0,05	0,049	0,05	0,053
		0,1	0,098	0,097	0,096	0,099	0,1	0,103
		0,2	0,203	0,199	0,194	0,195	0,199	0,209
	0,3	0,05	0,051	0,052	0,05	0,053	0,052	0,054
		0,1	0,1	0,102	0,102	0,101	0,1	0,102
		0,2	0,205	0,2	0,203	0,204	0,209	0,212
$1{,}075\tau_0$	0,1	0,05	0,059	0,058	0,054	0,62	0,07	0,072
		0,1	0,114	0,115	0,115	0,113	0,12	0,121
		0,2	0,207	0,2	0,209	0,205	0,21	0,21
	0,2	0,05	0,052	0,05	0,05	0,075	0,089	0,062
		0,1	0,103	0,1	0,096	0,112	0,116	0,115
		0,2	0,203	0,2	0,198	0,202	0,206	0,205
	0,3	0,05	0,056	0,055	0,051	0,05	0,05	0,06
		0,1	0,105	0,1	0,102	0,103	0,1	0,109
		0,2	0,197	0,2	0,202	0,204	0,204	0,206

Source: own calculations.

Parameters were used to generate a mixture of gamma distributions: $a = 2$, $c = 0{,}002$. For the mixing parameters $p = 0{,}1$, $p = 0{,}2$, $p = 0{,}3$ the corresponding parameters were determined $b = \frac{c\tau_1}{p}$. The results showing the power of the test are contained in Tables 1-2 and graphically interpreted in Figures 1-4.

The power diagrams are approximated by a broken line. The highest power is obtained for the mixing parameter $p = 0{,}1$ which corresponds to the value of the parameter $b = \frac{50 \cdot 0{,}002}{0{,}1} = 1$. The test power for $\tau_1 = 1{,}3\tau_0 = 65$ and $\alpha = 0{,}2$ is 0,477. In this case, the difference in the mean value of the documents with errors and the mean value of the documents without errors is $\overline{w} - \overline{y} = 500$. The lowest power is obtained for the mixing parameter $p = 0{,}3$, which corresponds to the value of the parameter $b = \frac{50 \cdot 0{,}002}{0{,}3} = \frac{1}{3}$. In this case, the difference of the average value of the documents with errors and the average value of the documents without errors is $\overline{w} - \overline{y} = 150$.

Table 2.
Power of the likelihood ratio test. A mixture of gamma distributions

τ_1	p	α	n					
			80	120	160	200	400	600
$1{,}1\tau_0$	0,1	0,05	0,065	0,064	0,065	0,066	0,069	0,078
		0,1	0,12	0,119	0,124	0,135	0,139	0,143
		0,2	0,2	0,22	0,22	0,215	0,219	0,23
	0,2	0,05	0,05	0,054	0,055	0,05	0,052	0,071
		0,1	0,106	0,114	0,111	0,117	0,106	0,132
		0,2	0,212	0,209	0,213	0,211	0,212	0,228
	0,3	0,05	0,06	0,055	0,054	0,051	0,051	0,065
		0,1	0,112	0,111	0,128	0,122	0,12	0,134
		0,2	0,227	0,218	0,217	0,224	0,227	0,229

Cont. table 2.

$1{,}15\tau_0$	0,1	0,05	0,07	0,075	0,082	0,098	0,1	0,109
		0,1	0,12	0,122	0,137	0,134	0,157	0,182
		0,2	0,214	0,218	0,231	0,256	0,254	0,294
	0,2	0,05	0,066	0,063	0,062	0,07	0,079	0,109
		0,1	0,112	0,113	0,111	0,128	0,138	0,162
		0,2	0,207	0,204	0,203	0,226	0,239	0,267
	0,3	0,05	0,057	0,059	0,053	0,064	0,089	0,101
		0,1	0,109	0,109	0,108	0,123	0,156	0,161
		0,2	0,229	0,218	0,223	0,237	0,257	0,264
$1{,}2\tau_0$	0,1	0,05	0,406	0,429	0,46	0,077	0,09	0,151
		0,1	0,131	0,139	0,139	0,144	0,157	0,23
		0,2	0,224	0,222	0,231	0,229	0,275	0,335
	0,2	0,05	0,056	0,054	0,053	0,066	0,077	0,136
		0,1	0,124	0,12	0,13	0,127	0,149	0,187
		0,2	0,219	0,216	0,211	0,246	0,272	0,333
	0,3	0,05	0,057	0,054	0,058	0,053	0,074	0,099
		0,11	0,104	0,1	0,111	0,116	0,145	0,175
		0,2	0,23	0,218	0,251	0,234	0,269	0,293
$1{,}3\tau_0$	0,1	0,05	0,09	0,123	0,127	0,133	0,163	0,235
		0,1	0,154	0,195	0,218	0,219	0,261	0,326
		0,2	0,258	0,281	0,327	0,341	0,394	0,477
	0,2	0,05	0,083	0,08	0,08	0,1	0,12	0,204
		0,1	0,144	0,142	0,145	0,162	0,225	0,309
		0,2	0,249	0,247	0,247	0,282	0,376	0,443
	0,3	0,05	0,054	0,055	0,055	0,07	0,098	0,177
		0,1	0,11	0,104	0,133	0,16	0,181	0,299
		0,2	0,231	0,225	0,249	0,261	0,314	0,434

Source: own calculations.

Figures 1-2 show that the test is unconstrained because the power of the test is not less than the significance level. The Power increases in an approximately linear manner. Figure 4 shows that the increment in test power as a function of sample size is slow. As the sample size increases by 100, the power of the test increases by 0.05.

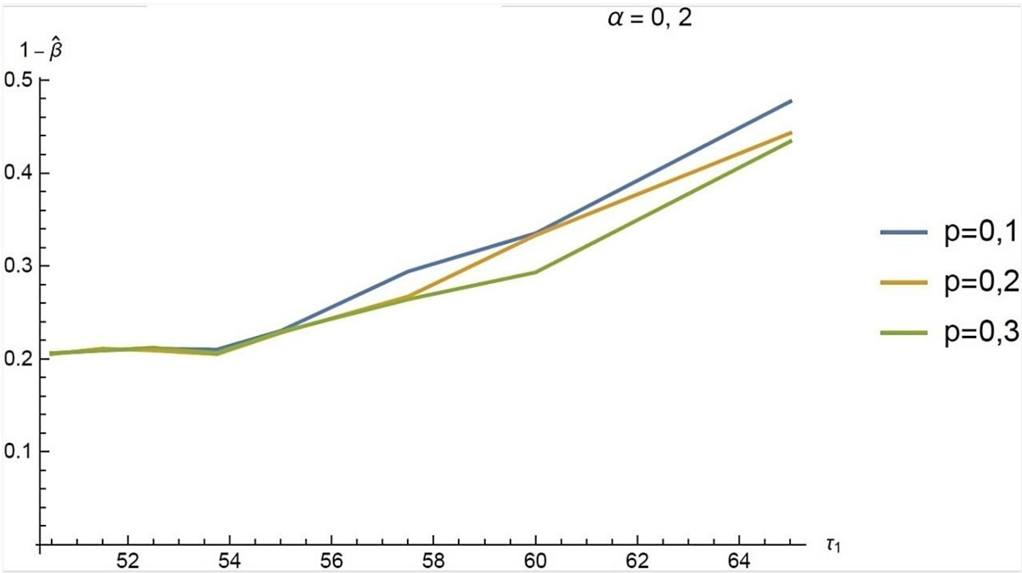

Figure 1. Test power depending on $\tau_0$and mixing parameters for a mixture of gamma distributions, α = 0,2, τ_0= 50.

Source: Based on Tables 1 and 2.

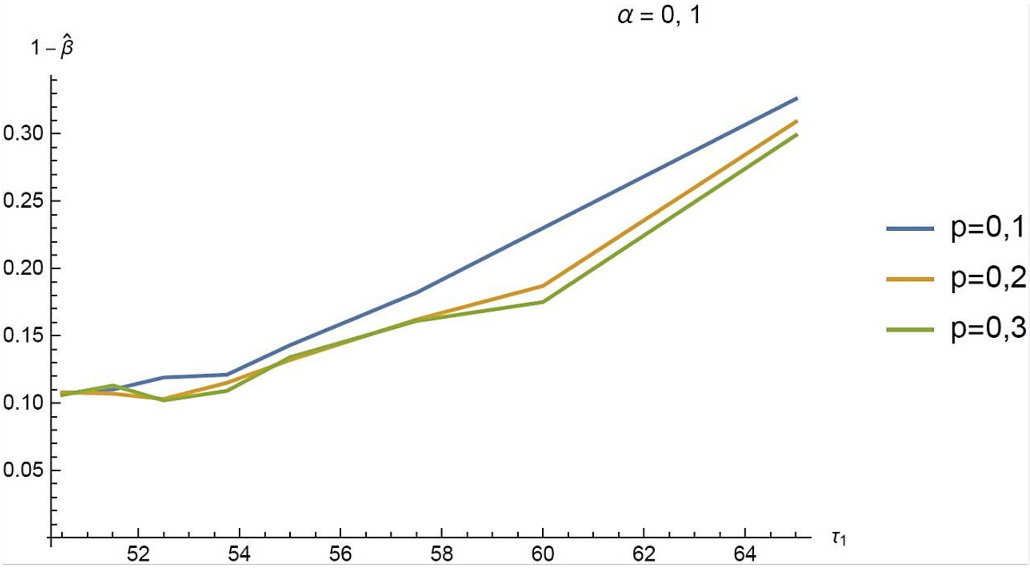

Figure 2. Test power depending on τ_0 and mixing parameters for a mixture of gamma distributions, α = 0,1, τ_0 = 50.

Source: Based on Tables 1 and 2.

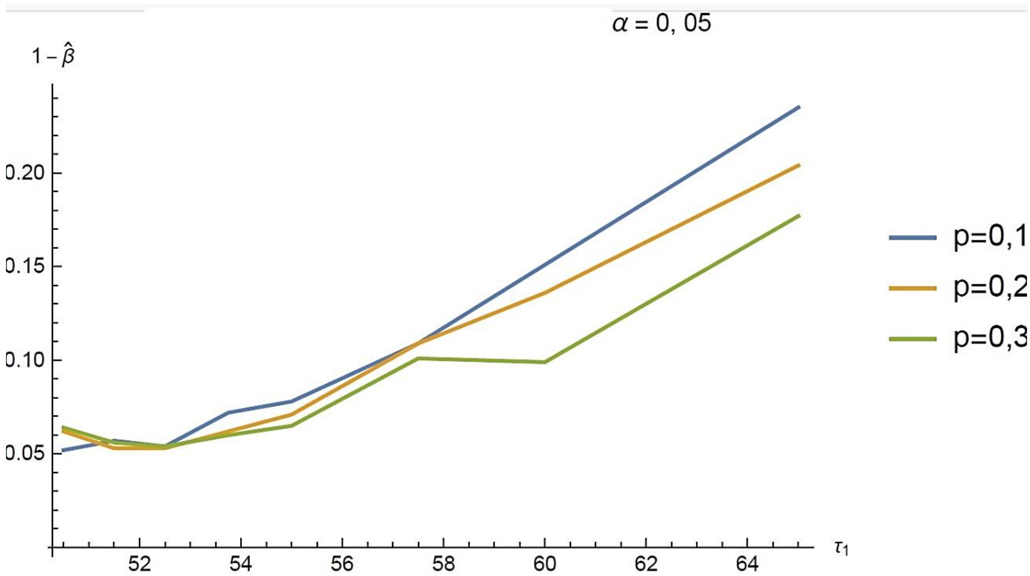

Figure 3. Test power depending on τ_0 and mixing parameters for a mixture of gamma distributions, α = 0,05, τ_0 = 50.

Source: Based on Tables 1 and 2.

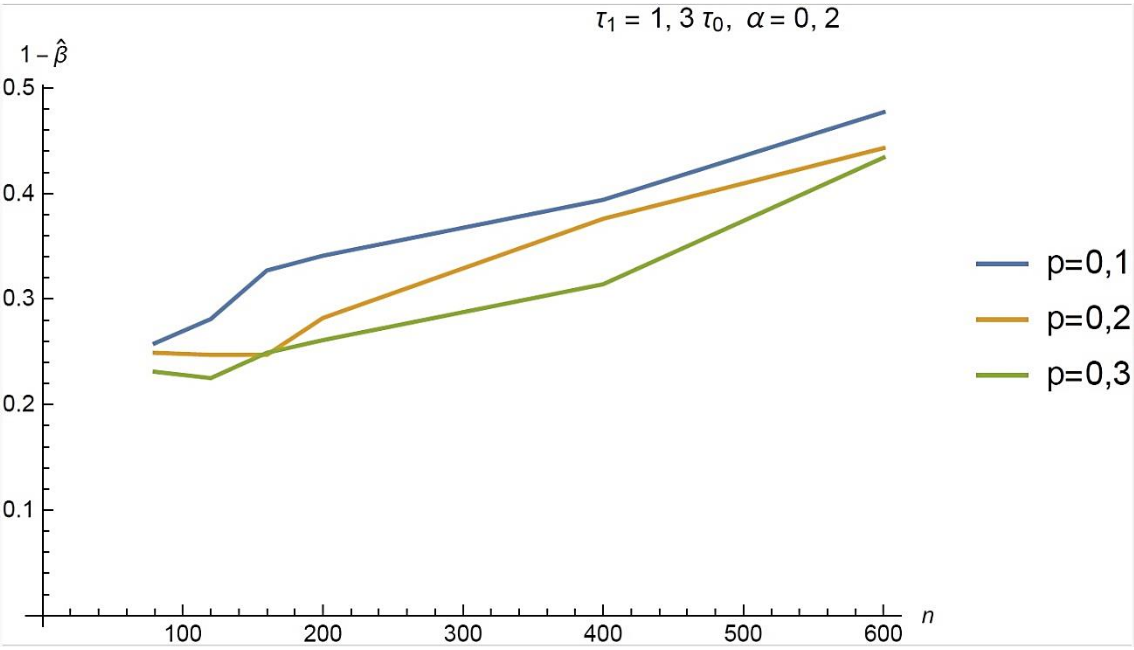

Figure 4. Test power as a function of sample size and mixing parameters for a mixture of gamma distributions, α = 0,2, τ_1= 1,3τ_0.

Source: Based on Tables 1 and 2.

5. Conclusions

In this paper, we show how to estimate parameters of the likelihood function. Particularly, the results let us construct confidence intervals or the statistics to testing hypotheses on the mean or the total amount error with assumed risk of incorrect rejection H_0 (significant level).

The considered maximum likelihood estimators are usually the solutions of the systems of the non-linear equations. In order to calculate those solutions, some numerical methods have to be used.

In the first step, the critical test values were determined by simulation. Then, in a second step, the test power was determined in a simulation manner for the previously determined critical values. For a mixture of gamma distributions, the following hypothesis was tested: $H_0: \tau = \tau_0 = 50, H_1: \tau = \tau_1 > \tau_0$.

Parameters were used to generate a mixture of gamma distributions: $a = 2$, $c = 0{,}002$. For the mixing parameters $p = 0{,}1$, $p = 0{,}2$, $p = 0{,}3$ the corresponding parameters were determined $b = \frac{c\tau_1}{p}$. The results showing the power of the test are contained in Tables 1-2 and graphically interpreted in Figures 1-4.

The power diagrams are approximated by a broken line. The highest power is obtained for the mixing parameter $p = 0{,}1$ which corresponds to the value of the parameter $b = \frac{50 \cdot 0{,}002}{0{,}1} = 1$. The test power for $\tau_1 = 1{,}3\tau_0 = 65$ and $\alpha = 0{,}2$ is 0,477. In this case, the difference in the mean value of the documents with errors and the mean value of the documents without errors is $\overline{w} - \overline{y} = 500$. The lowest power is obtained for the mixing parameter $p = 0{,}3$, which corresponds to the value of the parameter $b = \frac{50 \cdot 0{,}002}{0{,}3} = \frac{1}{3}$. In this case, the difference of the average value of the documents with errors and the average value of the documents without errors is $\overline{w} - \overline{y} = 150$. Figures 1-2 show that the test is unconstrained because the power of the test is not less than the significance level. The Power increases in an approximately linear manner. Figure 4 shows that the increment in test power as a function of sample size is slow. As the sample size increases by 100, the power of the test increases by 0.05.

The presented model has some limitations in practical application. If the number of erroneous values k in the drawn sample is small, it may not be possible to use the maximum likelihood method for estimating the starting parameters. In this case, it may not be possible to estimate the average audit error or the estimate may be incorrect. This model can be used in other cases where the population is heterogeneous and a mixture of different distributions needs to be used to describe it.

References

1. Frost, P.A., Tamura, H. (1986). Accuracy of auxiliary information interval estimation in statistical auditing. *Journal of Accounting Research*, *24*, 57-75.
2. Kaplan, R.S. (1973). Statistical Sampling in Auditing with Auxiliary Information Estimators. *Journal of Accounting Research*, 238-258.
3. McLachlan, G., Peel, D. (2000). *Finite Mixture Models.* New York: Wiley.
4. Meng, X.L., Rubin, D.B. (1993). Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika, 80,* 267-278.
5. Silvey, S.D. (1959). The Lagrangian multiplier test. *The Annals of Mathematical Statistics, 30, 2*, 389-407.
6. Stringer, K.W. (1963). *Practical Aspects of Statistical Auditing.* Preceeding of Business and Economic Statistics Section of the American Statistical Association, 405-411.
7. Tamura, H. (1988). Estimation of rare errors using judgement. *Biometrika*, *75*, 1-9.
8. Wywiał, J.L. (2016). *Contributions to Testing Statistical Hypotheses in Auditing*. Warsaw: PWN, 91-95.
9. Wywiał, J.L. (2018). Application of two gamma distributions mixture to financial auditing. *Sankhyã B: The Indian Journal of Statistics.*