ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2025 Управление, вычислительная техника и информатика № 73
Tomsk State University Journal of Control and Computer Science

Original article
UDK 369:519.2
doi: 10.17223/19988605/73/8

Nonparametric evaluation of continuous *r*-year deferred *m*-year term life annuity using information on probabilistic characteristics of lifetime

Yury G. Dmitriev[1], Gennady M. Koshkin[2]

[1, 2] *National Research Tomsk State University, Tomsk, Russian Federation*
[1] *dmit@mail.tsu.ru*
[2] *kgm@mail.tsu.ru*

Abstract. The theory of pension annuities is closely related to the ideology of net premiums of the life insurance theory. The mathematical theory of insurance is widely used to solve many problems that are determined by the requirements of the market economy. The requirements of practice stimulate the development of insurance theory and the closely related theory of annuities and force researchers to turn to more complex mathematical models in this area. New methods of calculating annuities appear that reduce the time for making optimal decisions in the absence of sufficient information about the markets of new types of pension services. The article considers the problem of estimating continuous *r*-year deferred *m*-year term life annuity with making use of information on probabilistic characteristics of lifetime. Insurance companies often offer their clients to conclude contracts of *r*-year deferred *m*-year annuities. Nonparametric estimators of life annuities are constructed from data on the lifetimes of individuals. Found the principal terms of the asymptotic mean squared errors (MSEs) of the proposed estimators; their asymptotic normality is proved. It is shown that the use of auxiliary information often leads to a lower MSE of the modified estimator compared to the MSE of the traditional estimator. An adaptive estimator is proposed that can be applied in practice.

Keywords: *r*-year deferred *m*-year term life annuity; nonparametric evaluation; auxiliary information; mean squared error; asymptotic normality; adaptive estimator.

For citation: Dmitriev, Yu.G., Koshkin, G.M. (2025) Nonparametric evaluation of continuous *r*-year deferred *m*-year term life annuity using information on probabilistic characteristics of lifetime. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitelnaja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 73. pp. 65–74. doi: 10.17223/19988605/73/8

Научная статья
doi: 10.17223/19988605/73/8

Непараметрическое оценивание непрерывной *m*-летней временной ренты, отсроченной на *r* лет, с учетом информации о вероятностных характеристиках продолжительности жизни

Юрий Глебович Дмитриев[1], Геннадий Михайлович Кошкин[2]

[1, 2] *Национальный исследовательский Томский государственный университет, Томск, Россия*
[1] *dmit@mail.tsu.ru*
[2] *kgm@mail.tsu.ru*

Аннотация. Теория пенсионных рент тесно связана с идеологией нетто-премий теории страхования жизни. Математическая теория страхования широко используется при решении многих задач, которые определяются требованиями рыночной экономики. Требования практики стимулируют развитие теории страхования и тесно связанную с ней теорию рент и вынуждают исследователей обращаться к более сложным математическим моделям в указанной области. Появляются новые методы расчета рент, которые сокращают время принятия оптимальных решений в условиях отсутствия достаточной информации о рынках новых видов пенсионных услуг. В статье рассматривается проблема оценивания непрерывной *m*-летней временной ренты, отсроченной на *r* лет, с учетом информации о вероятностных характеристиках продолжительности жизни. Страховые компании часто предлагают клиентам заключать договоры именно *m*-летних рент, отсроченных на *r* лет. Непараметрические оценки рент строятся по данным продолжительностей жизни индивидуумов. Найдены главные части асимптотических среднеквадратичных ошибок (СКО) предложенных оценок, доказана их асимптотическая нормальность. Показано, что использование дополнительной информации часто приводит к меньшей СКО модифицированной оценки по сравнению с СКО традиционной оценки. Предлагается адаптивная оценка, которая может применяться на практике.

Ключевые слова: *m*-летняя рента, отсроченная на *r* лет; непараметрическое оценивание; дополнительная информация; среднеквадратическая ошибка; асимптотическая нормальность; адаптивная оценка.

***Для цитирования*:** Дмитриев Ю.Г., Кошкин Г.М. Непараметрическое оценивания непрерывной *m*-летней временной ренты, отсроченной на *r*-лет, с учетом информации о вероятностных характеристиках продолжительности жизни // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 73. С. 65–74. doi: 10.17223/19988605/73/8

Introduction

The current stage of development of public socio-economic relations requires non-trivial approaches to the ideology of calculating pension annuities [1. P. 13–46, 170–194], which is associated with:

– the impact on the insurance market of such unpredictable phenomena as epidemics, natural disasters, social cataclysms, etc. [2–4],

– the emergence of new types of insurance and pension services [5, 6].

Let x be the age of an individual and at the moment *t* = 0 payments start. The idea of the *r*-year deferred *m*-year term life annuity in accordance with [1. P. 150] is this: from the moment *t*+*r* = *r*, an individual starts receiving a monetary amount once a year, but payments are made not more than *m* years. For ease of calculation, such a monetary amount is taken as equal to a conventional unit. So, payments are making in the time interval (*r, r*+*m*]. It is known that the life annuity is associated with the appropriate type of insurance. Thus, the average total cost of the present continuous *r*-year deferred *m*-year term life annuity gives the following formula (see [1. P. 184]):

$$_{r|}\bar{a}_{x:\overline{m}|}(\delta) = \frac{1 - {}_{r|}\bar{A}_{x:\overline{m}|}}{\delta},$$

where $_{r|}\bar{A}_{x:\overline{m}|} = \int_{x+r}^{x+r+m} e^{-\delta t} f_x(t)dt$ is a net premium (the expectation of the present value of an insured unit sum for an *r*-year deferred *m*-year term life insurance at age *x*), δ is a force of interest, $f_x(t) = \frac{f(x+t)}{S(x)}$ is a probability density of the future lifetime $T_x = X - x$ of an individual (*x*) [1. P. 62], $f(x)$ is a probability density of lifetime *X* of an individual (*x*), $S(x) = P(X > x)$ is a survival function. The essence of the present continuous *r*-year deferred *m*-year term life annuity is as follows: a client of age *x* who has entered into an agreement transfers to the company the sum of $_{r|}\bar{a}_{x:\overline{m}|}(\delta)$ conventional monetary units; then the company will pay one conventional monetary unit every year throughout the time interval (*r*, *r* + *m*]. It is clear that $_{r|}\bar{a}_{x:\overline{m}|}(\delta) > 1,$ and the value of the rent $_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta)$ increases with the growth of *m*.

Introduce the random variable

$$z(x) = \frac{1 - e^{-\delta T_x}}{\delta}, \ r + m \ge T_x > r. \tag{1}$$

Then, averaging the random variable *z*(*x*) (1) (see [7–9]), we get the formula of the *r*-year deferred *m*-year term life annuity:

$$_{r|}\bar{a}_{x:\overline{m}|}(\delta) = E(z) = \frac{1}{\delta}\left(1 - \frac{\Phi(x,r,m,\delta)}{S(x)}\right). \tag{2}$$

Here E is the symbol of the mathematical expectation,

$$\Phi(x,r,m,\delta)=e^{\delta x}\int_{x+r}^{x+r+m}e^{-\delta t}dF(t),$$

$F(x)=P\left(X\le x\right)=1-S(x)$ is a distribution function.

Note that the whole life annuity $\bar{a}_x(\delta)$ [7] is the special case of the annuity (2) at $r = 0$ and $m=\infty$.

Let us consider the problem of estimating continuous life annuities based on a sample of individuals' life expectancies [10–12]. The use of classical methods of statistical data processing often does not allow obtaining adequate models on the basis of which the insurance company's development strategy is built. When using classical parametric estimates and models, information about the phenomenon being studied is required with an accuracy of up to unknown parameters. In practice, problems often arise with the selection of suitable parametric estimates and models. Data processing using nonparametric statistical methods allows synthesizing simple and adequate (with known statistical properties) estimates and models in conditions where information about the phenomenon being studied is of a general nature [13].

1. Construction of the *r*-year deferred *m*-year term annuity estimator

Assume we have a random sample $X_1,\ldots,X_N$ of N individuals' lifetimes. Using the empirical survival function

$$S_N(x)=\frac{1}{N}\sum_{i=1}^{N}\mathrm{I}(X_i>x),$$

where I(A) is the indicator of an event A, we obtain the following estimator of (2):

$$ {}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta)=\frac{1}{\delta}\left(1-\frac{e^{\delta x}}{S_N(x)\cdot N}\sum_{i=1}^{N}\exp(-\delta X_i)\mathrm{I}(x+r+m\ge X_i>x+r)\right)=\frac{1}{\delta}\left(1-\frac{\Phi_N(x,r,m,\delta)}{S_N(x)}\right), \qquad (3)$$

$$\Phi_N(x,r,m,\delta)=\frac{e^{\delta x}}{N}\sum_{i=1}^{N}\exp(-\delta X_i)\mathrm{I}(x+r+m\ge X_i>x+r).$$

2. Bias and mean squared error of estimator (3)

In this section, we will obtain the principal term of the asymptotic MSE and the bias convergence rate of the estimator ${}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta)$.

Introduce the notation (see [14]): the function $H(t):R^s\to R^1$, $t=t(x)=\left(t_1(x),\ldots,t_s(x)\right)$, is an s-dimensional bounded function; $H_j(t)=\dfrac{\partial H(t)}{\partial t_j}$, $j=\overline{1,s}$, $\nabla H(t)=\left(H_1(t),\ldots,H_s(t)\right)$; the symbol T denotes the transpose; $t_N=\left(t_{1N},\ldots,t_{sN}\right)$ is an s-dimensional statistic, $t_{jN}=t_{jN}(x)=t_{jN}\left(x;X_1,\ldots,X_N\right)$, $j=\overline{1,s}$; $\|t_N\|=\sqrt{t_{1N}^2+\ldots+t_{sN}^2}$ is the Euclidean norm of t_N; $\Rightarrow N_s\{\mu,\sigma\}$ is the symbol of weak convergence of sequence of random variables to the s-dimensional normal random variable with a mean $\mu=(\mu_1,\ldots,\mu_s)$ and symmetric covariance matrix $\sigma=\|\sigma_{ij}\|$, $0<\sigma_{jj}=\sigma_{jj}(x)<\infty$, $j=\overline{1,s}$; $\Re$ is the set of the integers.

Definition 1. The function $H(t):R^s\to R^1$ and the sequence $\left\{H(t_N)\right\}$ are said to belong to the class $\mathrm{N}_{\nu,s}(t;\gamma)$, provided that

1) there exists an ε-neighborhood $\left\{z:|z_i-t_i|<\varepsilon, i=\overline{1,s}\right\}$, in which the function $H(z)$ and all its partial derivatives $\dfrac{\partial H(z)}{\partial z_j}$ up to the order ν are continuous and bounded;

2) for any values of variables $X_1,...,X_N$ the sequence $\{H(t_N)\}$ is dominated by a numerical sequence $C_0 d_N^{\gamma}$, such that $d_N \uparrow \infty$, as $N \to \infty$, and $0 \le \gamma < \infty$.

We present Theorem 1 from [14].

Theorem 1. *Let the conditions* 1) $H(z)$, $\{H(t_N)\} \in \mathrm{N}_{2,s}(t;\gamma)$ and 2) $E\|t_N - t\|^i = O\left(d_N^{-i/2}\right)$ *hold for all* $i \in \Re$. *Then, for every* $k \in \Re$,

$$\left| E\left[H(t_N) - H(t)\right]^k - E\left[\nabla H(t)\cdot(t_N - t)^{\mathrm{T}}\right]^k \right| = O\left(d_N^{-(k+1)/2}\right). \quad (4)$$

If in formula (4) $k = 1$, *we obtain the principal term* $E\left[\nabla H(t)\cdot(t_N - t)^{\mathrm{T}}\right]$ *of the bias* $E\left[H(t_N) - H(t)\right]$ *for* $H(t_n)$, *and at k* = 2, *we have the principal term* $E\left[\nabla H(t)\cdot(t_N - t)^{\mathrm{T}}\right]^2$ *of the MSE* $E\left[H(t_N) - H(t)\right]^2$.

Theorem 2. *If the survival function* $S(x) > 0$ *and* $S(t)$ *is continuous at a point x, then*

1) *for the bias* $b\left({}_{r|}\bar{a}^N_{x:\bar{m}|}(\delta)\right)$ *of estimator (3) we have*

$$\left| b\left({}_{r|}\bar{a}^N_{x:\bar{m}|}(\delta)\right)\right| = \left| E\left({}_{r|}\bar{a}^N_{x:\bar{m}|}(\delta) - {}_{r|}\bar{a}_{x:\bar{m}|}(\delta)\right)\right| = O\left(N^{-1}\right);$$

2) *the MSE* $u^2\left({}_{r|}\bar{a}^N_{x:\bar{m}|}(\delta)\right)$ *is given by the formula*

$$u^2\left({}_{r|}\bar{a}^N_{x:\bar{m}|}(\delta)\right) = E\left({}_{r|}\bar{a}^N_{x:\bar{m}|}(\delta) - {}_{r|}\bar{a}_{x:\bar{m}|}(\delta)\right)^2 = \frac{\Phi(x,r,m,2\delta) - \Phi^2(x,r,m,\delta)/S(x)}{N\delta^2 S^2(x)} + O\left(\frac{1}{N^{3/2}}\right).$$

Proof. For the estimator ${}_{r|}\bar{a}^N_{x:\bar{m}|}(\delta)$ (3) in the notation of Theorem 1, we have: $s = 2$; $d_N = N$;

$$t_N = (t_{1N}, t_{2N}) = \left(\Phi_N(x,r,m,\delta),\ S_N(x)\right); \quad t = (t_1, t_2) = \left(\Phi(x,r,m,\delta),\ S(x)\right);$$

$$H(t_N) = \frac{1}{\delta}\left(1 - \frac{t_{1N}}{t_{2N}}\right) = \frac{1}{\delta}\left(1 - \frac{\Phi_N(x,r,m,\delta)}{S_N(x)}\right) = {}_{r|}\bar{a}^N_{x:\bar{m}|}(\delta);$$

$$H(t) = \frac{1}{\delta}\left(1 - \frac{t_1}{t_2}\right) = \frac{1}{\delta}\left(1 - \frac{\Phi(x,r,m,\delta)}{S(x)}\right) = {}_{r|}\bar{a}_{x:\bar{m}|}(\delta); \quad H_1(t) = -\frac{1}{\delta S(x)}, \quad H_2(t) = \frac{\Phi(x,r,m,\delta)}{\delta S^2(x)},$$

$$\nabla H(t) = \left(H_1(t), H_2(t)\right) = \left(-\frac{1}{\delta S(x)}, \frac{\Phi(x,r,m,\delta)}{\delta S^2(x)}\right) \ne 0.$$

The sequence $\{H(t_N)\}$ satisfies the condition 1) of Theorem 1 with $C_0 = \frac{1}{\delta}\left(1 + e^{-\delta r}\right)$ and $\gamma = 0$:

$$|H(t_N)| = \frac{1}{\delta}\left|1 - \frac{\Phi_N(x,r,m,\delta)}{S_N(x)}\right| \le \frac{1}{\delta}\left(1 + \frac{\Phi_N(x,r,m,\delta)}{S_N(x)}\right) \le \frac{1}{\delta}\left(1 + \frac{e^{\delta x}\sum_{i=1}^{N}\exp(-\delta X_i)\mathrm{I}(x+r+m \ge X_i > x+r)}{\sum_{i=1}^{N}\mathrm{I}(X_i > x)}\right) \le$$

$$\le \frac{1}{\delta}\left(1 + \frac{e^{\delta x}e^{-\delta(x+r)}\sum_{i=1}^{N}\mathrm{I}(x+r+m \ge X_i > x+r)}{\sum_{i=1}^{N}\mathrm{I}(X_i > x)}\right) \le \frac{1}{\delta}\left(1 + e^{-\delta r}\right).$$

Further, in view of $t_2 = S(x) > 0$, the function $H(t)$ satisfies the condition 1) of Theorem 1. Also, this function satisfies the condition 2) of Theorem 1 due to Lemma 3.1 [15], as for all $i \in \Re$ such inequalities hold:

$$E\left[e^{i\delta x}e^{-i\delta X}\mathrm{I}^i(x+r+m \ge X > x+r)\right] \le e^{i\delta x}e^{-i\delta(x+r)}\left[S(x+r) - S(x+r+m)\right] =$$

$$= e^{-i\delta r}\left[S(x+r) - S(x+r+m)\right] \le 1, \qquad E\left[\mathrm{I}^i(X > x)\right] = S(x) \le 1.$$

It is well known that $S_N(x)$ is an unbiased and consistent estimator of $S(x)$ Show that $\Phi_N(x,r,m,\delta)$ is an unbiased estimator of $\Phi(x,r,m,\delta)$:

$$E\Phi_N(x,r,m,\delta)=\frac{e^{\delta x}}{N}E\left[\sum_{i=1}^{N}\exp(-\delta X_i)\mathrm{I}(x+r+m\geq X_i>x+r)\right]=\Phi(x,r,m,\delta).$$

The ratio of two unbiased estimators can have a bias. All conditions of Theorem 1 are satisfied and $E\left[\nabla H(t)(t_N-t)^{\mathrm{T}}\right]=0,$ therefore, in accordance with (4), we obtain the order of the bias $b\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta)\right)$:

$$\left|E\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta)-{}_{r|}\bar{a}_{x:\overline{m}|}(\delta)\right)-E\left[\nabla H(t)(t_N-t)\right]\right|=\left|E\,{}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta)-{}_{r|}\bar{a}_{x:\overline{m}|}(\delta)\right|=O\left(N^{-1}\right).$$

Now, calculate the variance of estimator $\Phi_N(x,r,m,\delta)$:

$$D\Phi_N(x,r,m,\delta)=D\left[\frac{e^{\delta x}}{N}\sum_{i=1}^{N}\exp(-\delta X_i)\mathrm{I}(x+r+m\geq X_i>x+r)\right]=$$

$$=\frac{e^{2\delta x}}{N^2}\sum_{i=1}^{N}D\left[\exp(-\delta X_i)\mathrm{I}(x+r+m\geq X_i>x)\right]=\frac{1}{N}\left(\Phi(x,r,m,2\delta)-\Phi^2(x,r,m,\delta)\right).$$

Similarly, we find the covariance matrix $\Omega_{2\times2}\left(x,r,m,\delta\right)=\begin{bmatrix}\sigma_{11}\sigma_{12}\\ \sigma_{21}\sigma_{22}\end{bmatrix}$ for statistics $S_N(x)$ and $\Phi_N(x,r,m,\delta)$:

$$\sigma_{11}=ND\left[\Phi_N(x,r,m,\delta)\right]=\Phi(x,r,m,2\delta)-\Phi^2(x,r,m,\delta);\quad \sigma_{22}=ND\left[S_N(x)\right]=S(x)\left(1-S(x)\right);$$

$$\sigma_{12}=\sigma_{21}=N\operatorname{cov}\left(S_N(x),\,\Phi_N(x,\delta)\right)=$$

$$=N\left[E\left[S_N(x)\Phi_N(x,r,m,\delta)\right]-E\left[S_N(x)\right]E\left[\Phi_N(x,r,m,\delta)\right]\right]=\left(1-S(x)\right)\Phi(x,r,m,\delta).$$

Using the previous results on the bias and the covariance matrix, we obtain

$$u^2\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta)\right)=E\left[\nabla H(t)(t_N-t)^{\mathrm{T}}\right]^2+O\left(\frac{1}{N^{3/2}}\right)=$$

$$=H_1^2(t)\sigma_{11}+H_2^2(t)\sigma_{22}+2H_1(t)H_2(t)\sigma_{12}+O\left(\frac{1}{N^{3/2}}\right)=\frac{W\left(x,r,m,\delta\right)}{N}+O\left(\frac{1}{N^{3/2}}\right),\tag{5}$$

$$W\left(x,r,m,\delta\right)=\frac{\Phi(x,r,m,2\delta)-\Phi^2(x,r,m,\delta)\,/\,S(x)}{\delta^2S^2(x)}.\tag{6}$$

The proof of Theorem 2 is completed.

3. Asymptotic normality of estimator (3)

We need Theorem 3 (the central limit theorem from [16]) and Theorem 4 from [14].

Theorem 3. *If* $\xi_1,\xi_2,...,\xi_N,...$ *is a sequence of independent and identically distributed s-dimensional vectors,* $E\{\xi_k\}=0,$ $\sigma=E\{\xi_k^{\mathrm{T}}\xi_k\},$ *and* $t_N=\frac{1}{N}\sum_{k=1}^{N}\xi_k$*, then, as* $N\to\infty$*,*

$$\sqrt{N}t_N\Rightarrow N_s\left\{0,\sigma\right\}.$$

Theorem 4. *If* $q_N\left(t_N-t\right)\Rightarrow N_S\left\{\mu,\sigma\right\}$ *for some number sequence* $q_N\uparrow\infty,$ *the function H(z) is differentiable at a point* μ, $\nabla H(\mu)\neq0,$ *then*

$$q_N\left(H(t_N)-H(\mu)\right)\Rightarrow N_1\left\{\mu\nabla H(\mu)\mu^{\mathrm{T}},\ \nabla H(\mu)\sigma\nabla H^{\mathrm{T}}(\mu)\right\}.$$

Theorem 5. *Under the conditions of Theorem 2*

$$\sqrt{N}\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta)-{}_{r|}\bar{a}_{x:\overline{m}|}(\delta)\right)\Rightarrow N_1\left\{0,\frac{\Phi(x,r,m,2\delta)-\Phi^2(x,r,m,\delta)/S(x)}{\delta^2S^2(x)}\right\}=N_1\left\{0,W\left(x,r,m,\delta\right)\right\}.$$

Proof. In the notation of Theorem 4, we have $s = 2$, $\sigma = \Omega_{2\times2}\left(x, r, m, \delta\right)$ in accordance with Section 2, and $q_N = \sqrt{N}$. Now, applying Theorem 3, we get

$$\sqrt{N}\left[\left[\Phi_N(x,r,m,\delta), S_N(x)\right] - \left[\Phi(x,r,m,\delta), S(x)\right]\right] \Rightarrow N_2\left\{(0,0), \Omega_{2\times2}\left(x,r,m,\delta\right)\right\}.$$

The function *H*(*z*) is differentiable at the point $t = \left(\Phi(x,r,m,\delta), S(x)\right)$, $\nabla H(t) \neq 0$, Consequently, all conditions of Theorem 4 hold, and using (5) and (6), we obtain the desired result.

The proof of Theorem 5 is completed.

4. Construction of estimators using information on probabilistic characteristics of lifetime

Let ω be the limiting age, i.e. $S(x) > 0$ for $x < \omega$, and $S(x) = 0$ for $x \geq \omega$. Suppose we know the average of the lifetime functional

$$Eg(X) = \int_0^{\omega} g(x)dF(x) = J, \tag{7}$$

where $g(x)$, $x \in [0, \omega]$, is a known function. If $g(x)$ is the indicator of the set $\left\{\omega : C_1 \leq g(x) \leq C_2\right\}$, then one gets the probability $P\left(C_1 \leq X \leq C_2\right)$; for $g(x) = x^r$ one comes to the initial moment of order *r*, and for $g(x) = \left(x - EX\right)^r$ – to the central moment of order *r*, and so on. The estimator using such information can be taken in the following from [17, 18]:

$${}_{r|}\bar{a}^N_{x:\overline{m}|}(\delta,\lambda) = \frac{1}{\delta}\left(1 - \frac{\Phi_N(x,r,m,\delta)}{S_N(x)} - \lambda\left(J_N - J\right)\right), \tag{8}$$

where $J_N = \frac{1}{N}\sum_{i=1}^{N} g\left(X_i\right)$ is an estimator of *J*, the parameter λ we will find minimizing the principal term of the asymptotic MSE of ${}_{r|}\bar{a}^N_{x:\overline{m}|}(\delta,\lambda)$ (8). Estimator (8) combines the available empirical information containing in (3) and prior information (7).

For estimator ${}_{r|}\bar{a}^N_{x:\overline{m}|}(\delta,\lambda)$ in the notation of Theorem 1, we have: $s = 3$, $d_N = N$;

$$t_N = \left(t_{1N}, t_{2N}, t_{3N}\right) = \left(\Phi_N(x,r,m,\delta), S_N(x), J_N\right); \quad t = \left(t_1, t_2, t_3\right) = \left(\Phi(x,r,m,\delta), S(x), J\right);$$

$$H(t) = H(t_1,t_2,t_3) = \frac{1}{\delta}\left(1 - \frac{t_1}{t_2} - \lambda\left(t_3 - J\right)\right) = \frac{1}{\delta}\left(1 - \frac{\Phi(x,r,m,\delta)}{S(x)} - \lambda\left(J - J\right)\right) = {}_{r|}\bar{a}_{x:\overline{m}|}(\delta,\lambda);$$

$$H(t_N) = \frac{1}{\delta}\left(1 - \frac{\Phi_N(x,r,m,\delta)}{S_N(x)} - \lambda\left(J_N - J\right)\right) = {}_{r|}\bar{a}^N_{x:\overline{m}|}(\delta,\lambda);$$

$$\nabla H(t) = \left(H_1(t), H_2(t), H_3(t)\right) = \left(-\frac{1}{\delta S(x)}, \frac{\Phi(x,r,m,\delta)}{\delta S^2(x)}, -\frac{\lambda}{\delta}\right) \neq 0. \tag{9}$$

5. Bias and MSE of estimators using auxiliary information

Arguing as in the proof of Theorem 2, it is easy to show that the sequence $\left\{H(t_N)\right\}$ in (9) satisfies the condition 1) of Theorem 1 with $C_0 = \frac{1}{\delta}\left(1 + e^{-\delta r} + |\lambda|\left(K + |J|\right)\right)$, where $\sup_{x\in[0,\omega]} |g(x)| = K < \infty$, and $\gamma = 0$. Also, the statistic t_N satisfies the condition 2) due to Lemma 3.1 [15], provided that $Eg^i(X) \leq K^i < \infty$ for all $i \in \Re$. Hence, given that $E\left(t_N - t\right) = 0$, for the bias of (7) we obtain

$$\left|E\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda)-{}_{r|}\bar{a}_{x:\overline{m}|}(\delta,\lambda)\right)-E\left[\nabla H(t)(t_N-t)\right]\right|=$$
$$=\left|E\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda)-{}_{r|}\bar{a}_{x:\overline{m}|}(\delta,\lambda)\right)\right|=\left|b\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda)\right)\right|=O\left(N^{-1}\right). \quad (10)$$

Now, find the covariance matrix $\Omega_{3\times3}\left(x,r,m,\delta\right)=\begin{bmatrix}\sigma_{11}\sigma_{12}\sigma_{13}\\ \sigma_{21}\sigma_{22}\sigma_{23}\\ \sigma_{31}\sigma_{32}\sigma_{33}\end{bmatrix}$ for statistics $\Phi_N(x,r,m,\delta)$, $S_N(x)$, and J_N : $\sigma_{33}=NDJ_N=Dg(X)$; $\sigma_{13}=\sigma_{31}=N\operatorname{cov}\left(J_N,\Phi_N(x,r,m,\delta)\right)=C_1(x,r,m,\delta)-J\Phi(x,r,m,\delta)$; $\sigma_{23}=\sigma_{32}=N\operatorname{cov}\left(S_N(x),J_N\right)=C_2(x,r,m,)-JS(x)$, where $C_1(x,r,m,\delta)=e^{\delta x}\int\limits_{x+r}^{x+r+m}e^{-\delta u}g(u)dF(u)$, $C_2(x,r,m)=\int\limits_{x+r}^{x+r+m}g(u)dF(u)$, and σ_{11}, σ_{12}, σ_{21}, σ_{22} are defined in Section 2. Using (4) at k = 2, (5), (6), (9), (10), and covariance matrix $\Omega_{3\times3}\left(x,r,m,\delta\right)$, we obtain:

$$u^2\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda)\right)=E\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda)-{}_{r|}\bar{a}_{x:\overline{m}|}(\delta)\right)^2=$$
$$=E\left[\nabla H(t)(t_N-t)\right]^2+O\left(\frac{1}{N^{3/2}}\right)=\frac{W\left(x,r,m,\delta,\lambda\right)}{N}+O\left(\frac{1}{N^{3/2}}\right), \quad (11)$$

$$W\left(x,r,m,\delta,\lambda\right)=\sum_{p=1}^{3}\sum_{j=1}^{3}H_j(t)\sigma_{jp}H_p(t)=H_1^2(t)\sigma_{11}+H_2^2(t)\sigma_{22}+H_3^2(t)\sigma_{33}+2H_1(t)H_2(t)\sigma_{12}+$$
$$+2H_1(t)H_3(t)\sigma_{13}+2H_2(t)H_3(t)\sigma_{23}=W\left(x,r,m,\delta\right)+\frac{\lambda^2\sigma_{33}}{\delta^2}-\frac{2\lambda H_1\sigma_{13}}{\delta}-\frac{2\lambda H_2\sigma_{23}}{\delta}= \quad (12)$$
$$=W\left(x,r,m,\delta\right)+\lambda^2Q_1-2\lambda Q_2,$$

where $Q_1=\frac{\sigma_{33}}{\delta^2}>0$, $Q_2=\frac{H_1\sigma_{13}+H_2\sigma_{23}}{\delta}$.

Thus, the derived formulas (10)–(12) determine the bias and MSE of the estimate (8) and allow us to formulate the following theorem.

Theorem 6. *If the survival function* $S(x)>0$, $S(t)$ *is continuous at a point* x, $\sup\limits_{x\in[0,\omega]}\left|g(x)\right|=K<\infty$, *then*

1) *for the bias of estimator (8) the following relation holds*: $\left|b\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda)\right)\right|=O\left(N^{-1}\right)$;

2) *the MSE of estimator (8) is given by the formula (11).*

The minimum of $W\left(x,r,m,\delta,\lambda\right)$ (12) with respect to λ is achived at $\lambda_0=Q_2/Q_1$. Such λ_0 minimizes the principal term of MSE $u^2\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda)\right)$, and this minimum is as follows:

$$\frac{W\left(x,r,m,\delta,\lambda_0\right)}{N}=\frac{1}{N}\left(W\left(x,r,m,\delta\right)-\frac{Q_2^2}{Q_1}\right)\le\frac{W\left(x,r,m,\delta\right)}{N}. \quad (13)$$

So, the principal term of MSE (11) at λ_0 is not more than the principal term of MSE (5), and, in accordance with (13), the estimator

$${}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda_0)=\frac{1}{\delta}\left(1-\frac{\Phi_N(x,r,m,\delta)}{S_N(x)}-\lambda_0\left(J_N-J\right)\right) \quad (14)$$

will be called the optimal estimator in the mean square sense. The non-negative quantity Q_2^2/Q_1 in (13) determines the decrease of the principal term of MSE for the optimal estimator ${}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda_0)$ by making use of auxiliary information (7).

Theorem 7. *If the conditions of Theorem 6 hold, then*

1) $\left|b\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda_0)\right)\right| = O\left(N^{-1}\right);$

2) $u^2\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}\left(\delta,\lambda_0\right)\right) = \dfrac{W\left(x,r,m,\delta,\lambda_0\right)}{N} + O\left(\dfrac{1}{N^{3/2}}\right),$

where $W\left(x,r,m,\delta,\lambda_0\right)$ *is defined by the formula (13).*

Note that Theorem 7 is a simple consequence of Theorem 6.

6. Asymptotic normality of estimators using auxiliary information

Theorem 8. *Under the conditions of Theorem 6*

$$\sqrt{N}\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda) - {}_{r|}\bar{a}_{x:\overline{m}|}(\delta)\right) \Rightarrow N_1\left\{0, W\left(x,r,m,\delta,\lambda\right)\right\}.$$

Proof. In the notation of Theorem 4, we have $s=3$, $\sigma = \Omega_{3\times 3}\left(x,r,m,\delta\right)$ in accordance with Section 5, and $q_N = \sqrt{N}$. Now, applying Theorem 3, we get

$$\sqrt{N}\left[\left[\Phi_N(x,r,m,\delta), S_N(x), J_N\right] - \left[\Phi(x,r,m,\delta), S(x), J\right]\right] \Rightarrow N_3\left\{(0,0,0), \Omega_{3\times 3}\left(x,r,m,\delta\right)\right\}.$$

The function $H(z)$ is differentiable at the point $t = \left(\Phi(x,r,m,\delta), S(x), J\right)$, $\nabla H(t) \neq 0$. Consequently, all conditions of Theorem 4 hold, and using (15) and (16), we obtain the desired result.

The proof of Theorem 8 is completed.

The asymptotic normality of the optimal estimator ${}_{r|}\bar{a}^{N}_{x:\overline{m}|}\left(\delta,\lambda_0\right)$ is determined by Theorem 9.

Theorem 9. *Under the conditions of Theorem 6*

$$\sqrt{N}\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda_0) - {}_{r|}\bar{a}_{x:\overline{m}|}(\delta)\right) \Rightarrow N_1\left\{0, W\left(x,r,m,\delta,\lambda_0\right)\right\}.$$

Theorem 9 is a simple consequence of Theorem 8.

7. Adaptive Estimator

The statistic ${}_{r|}\bar{a}^{N}_{x:\overline{m}|}\left(\delta,\lambda_0\right)$ can be used as an estimator for ${}_{r|}\bar{a}_{x:\overline{m}|}(\delta)$ if we know λ_0; otherwise, it is required to construct an adaptive estimator. We need a more detailed formula for λ_0:

$$\lambda_0 = \frac{1}{S(x)Dg(X)}\left[\frac{\Phi(x,r,m,\delta)}{S(x)}\left(C_2(x,r,m) - JS(x)\right) - C_1(x,r,m,\delta) + J\Phi(x,r,m,\delta)\right]. \tag{15}$$

Using (15), we consider the following adaptive estimator:

$${}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\hat{\lambda}_0) = \frac{1}{\delta}\left(1 - \frac{\Phi_N(x,r,m,\delta)}{S_N(x)} - \hat{\lambda}_0\left(J_N - J\right)\right) \tag{16}$$

with

$$\hat{\lambda}_0 = \frac{1}{s^2 S_N(x)}\left[\frac{\Phi_N(x,r,m,\delta)}{S_N(x)}\left(\hat{C}_2(x,r,m) - JS_N(x)\right) - \hat{C}_1(x,r,m,\delta) + J\Phi_N(x,r,m,\delta)\right], \tag{17}$$

where $s^2 = \dfrac{1}{N-1}\sum\limits_{i=1}^{N}\left(g(X_i) - J_N\right)^2$ is an unbiased estimator of the variance $Dg(X)$,

$\hat{C}_2(x,r,m) = N^{-1}\sum\limits_{i=1}^{N} g\left(X_i\right)\mathrm{I}\left(x+r+m \geq X_i > x+r\right)$, $\hat{C}_1(x,r,m,\delta) = N^{-1}\sum\limits_{i=1}^{N} e^{-\delta X_i} g(X_i)\mathrm{I}\left(x+r+m \geq X_i > x+r\right).$

Theorem 10. *Under the conditions of Theorem 6*

$$\sqrt{N}\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\hat{\lambda}_0) - {}_{r|}\bar{a}_{x:\overline{m}|}(\delta)\right) \Rightarrow N_1\left\{0, W\left(x,r,m,\delta,\lambda_0\right)\right\}.$$

Proof. The following equality holds:

$$\sqrt{N}\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\hat{\lambda}_0) - {}_{r|}\bar{a}_{x:\overline{m}|}(\delta) \right) = \sqrt{N}\left({}_{r|}\bar{a}^{N}_{x:\overline{m}|}(\delta,\lambda_0) - {}_{r|}\bar{a}_{x:\overline{m}|}(\delta) \right) + R_N ,$$

where $R_N = \delta^{-1}\left(\lambda_0 - \hat{\lambda}_0\right)\sqrt{N}\left(J_N - J\right).$ All the estimators, used in (17), converge almost surely to their true values according to the strong law of large numbers (the Second Theorem of Kolmogorov [19]). Thus, from the First Continuity Theorem of Borovkov [16], estimator $\hat{\lambda}_0$ converges almost surely to λ_0. Based on the central limit theorem $\sqrt{N}(J_N - J) \Rightarrow N_1\{0, Dg(X)\},$ we retrieve $R_N \Rightarrow 0.$ Now, the statement of Theorem 10 is proved by using Theorem 9.

Conclusion

The paper deals with the problem of estimating the present values of the continuous whole life annuity using auxiliary information about the expectation of life. It is shown that the usage of such auxiliary information can often provide the MSE not more then that of standard estimators. We proved the results on asymptotic properties of the proposed estimators: unbiasedness, consistency and normality. Also, the main parts of the asymptotic MSEs of the estimators were found. An adaptive estimator is constructed; such estimator is equivalent (in the sense of asymptotic distribution) to the estimator with the optimal weight coefficient λ_0. Note that the improved estimators of life annuities (8) and (12) can be obtained by substituting of empirical survival functions by the smooth empirical survival functions [20].

References

1. Falin, G.I. (2002) *Mathematical foundations of the theory of life insurance and pension schemes*. Moscow: Ankil.
2. Al-Nator, M.S., Al-Nator, S.V. & Soloviev, A.K. (2018) Actuarial and statistical substantiation of the retirement age in the context of socio-economic uncertainty. *Modern Mathematics and Concepts of Innovative Mathematical Education.* 5(1). pp. 144–148.
3. Ovanesyan, N.M. & Midler, E.A. (2016) Long-term life insurance as a factor in the sustainable development of the insurance services market in an inflationary environment. *Financial Research.* 1(50). pp. 128–134.
4. Roik, V.D. (2013) Ways to improve insurance of occupational disease risks. *Occupational Safety and Economics*. 2(23). pp. 4–14.
5. Soloviev, A.K. (2018) Risks of forecasting pension reform in the context of digitalization. *Bulletin of the Department of Statistics of the Plekhanov Russian University of Economics: Statistical studies of the socio-economic development of Russia and prospects for sustainable growth: materials and reports.* Moscow. pp. 257–260.
6. Winter, P. & Planchet, F. (2022) Modern tontines as a pension solution: a practical overview premium. *European Actuarial Journal.* 12. pp. 3–32.
7. Gubina, O.V. & Koshkin, G.M. (2015) Estimation of actuarial present value of the whole continuous life annuity *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitelnaya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science.* 30(1). pp. 38–43.
8. Bowers, N., Gerber, H., Hickman, J., Jones, D. & Nesbitt C. (1986) *Actuarial Mathematics.* Itasca: Society of Actuaries.
9. Gerber, H. (1997) *Life Insurance Mathematics*. 3rd ed. New York: Springer-Verlag.
10. Kokina, E.P. & Tregubova A.A. (2015) Formation of insurance tariff classes: application of statistical methods. *Financial Research.* 2(47). pp. 115–122.
11. Awondo, S., Ramirez, O., Datta, G.S., Colson, G. & Fonsah, E.G. (2018) Estimation of grop yields and insurance premiums using a shrinkage estimator. *North American Actuarial Journal.* 22(2). pp. 289–308.
12. Gubina, O.V. & Koshkin, G.M. (2020) Estimation of the present value of *n*-year life annuity for endowment insurance. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitelnaya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science.* 50. pp. 39–46.
13. Hu, J. & Hong, L (2022) A nonparametric sequential learning procedure for estimating the pure premium. *European Actuarial Journal*. 2. pp. 485–502.
14. Koshkin, G.M. (1999) Deviation moments of the substitution estimator and its piecewise smooth approximations. *Siberian Mathematical Journal.* 40(3). pp. 515–527.
15. Ibragimov, I.A. & Khasminskii, R.Z. (1981) *Statistical Estimation: Asymptotic Theory*. Berlin – New York: Springer.
16. Borovkov, A.A. (1998) *Mathematical Statistics*. New York: Gordon & Breach.
17. Dmitriev, Yu.G. & Koshkin, G.M. (1987) Using additional information in nonparametric estimation of density functionals. *Automation and Remote Control*. 48(10). pp. 1307–1316.

18. Dmitriev, Y.G. & Koshkin, G.M. (2018) Nonparametric estimators of probability characteristics using unbiased prior conditions. *Statistical Papers*. 59(4), pp. 1559–1575. doi: 10.1007/s00362-018-1044-7
19. Rao, C.R. (1965) *Linear Statistical Methods and Its Applications*. New York: Willey.
20. Koshkin, G.M. (2014) Smooth estimators of the reliability functions for non-restorable elements. *Russian Physics Journal*. 57(5). pp. 672–681.

***Information about the authors*:**
Dmitriev Yury G. (Associate Professor, Doctor of Physical and Mathematical Sciences, Head of the Department of the System Analysis and Mathematical Modeling of the National Research Tomsk State University, Tomsk, Russian Federation). E-mail: dmit@mail.tsu.ru
Koshkin Gennady M. (Professor, Doctor of Physical and Mathematical Sciences, Professor of the Department of the System Analysis and Mathematical Modeling of the National Research Tomsk State University, Tomsk, Russian Federation). E-mail: kgm@mail.tsu.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

***Информация об авторах*:**

Дмитриев Юрий Глебович – доцент, доктор физико-математических наук, заведующий кафедрой системного анализа и математического моделирования Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: dmit70@mail.ru
Кошкин Геннадий Михайлович – профессор, доктор физико-математических наук, профессор кафедры системного анализа и математического моделирования Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: kgm@mail.tsu.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Received 24.08.2025; accepted for publication 02.12.2025

Поступила в редакцию 24.08.2025; принята к публикации 02.12.2025