УДК 519.872.7:658.51

АНАЛИЗ ВЛИЯНИЯ ИНТЕРВАЛА СБРОСА СТАТИСТИКИ ИМИТАЦИОННОГО ЭКСПЕРИМЕНТА НА ТОЧНОСТЬ СРЕДНЕИНТЕГРАЛЬНОЙ ОЦЕНКИ

Солнцев Алексей Александрович, *кандидат технических наук, Московский автомобильно-дорожный государственный технический университет (МАДИ), 125319, Россия, г. Москва, Ленинградский проспект, 64, e-mail: kafedra@asu.madi.ru.*

Приходько Михаил Вячеславович, *кандидат технических наук, заместитель директора ЗАО «ЛонМАДИ», 125319, Россия, г. Москва, Ленинградский проспект, 64, e-mail: kafedra@asu.madi.ru.*

Кудрявцев Андрей Юрьевич, *кандидат технических наук, инженер ЗАО НПВФ «СВАРКА», 428003, Россия, Чувашская Республика, г. Чебоксары, Кабельный проезд, 9, e-mail: kafedra@asu.madi.ru.*

Котов Андрей Александрович, *кандидат технических наук, Московский автомобильно-дорожный государственный технический университет (МАДИ), 125319, Россия, г. Москва, Ленинградский проспект, 64, e-mail: kafedra@asu.madi.ru.*

Большинство современных методов, разработанных для получения оценок по результатам имитационных экспериментов, предполагают стационарность выходного процесса и в связи с этим отбрасывание переходного периода. В данной статье выполнен анализ точности оценки с учетом сброса статистики, накопленной на начальном периоде моделирования. Предполагается, что ковариационная функция условно нестационарного гауссовского процесса определяется на основании теоремы о нормальной корреляции для заданного стационарного гауссовского процесса, имеющего автокорреляционную функцию в виде смеси двух экспонент. Проведен анализ математического ожидания среднеинтегральной оценки условно нестационарного процесса при условии сброса статистики, накопленной на начальном интервале моделирования. Получены аналитические выражения математического ожидания и дисперсии среднеинтегральной оценки в зависимости от величины сброса статистики. Показано, что определяющим фактором является степень коррелированности процесса, а длительность интервала сброса статистики существенно отражается на дисперсии среднеинтегральной оценки. Сброс начальной статистики увеличивает дисперсию оценки и таким образом уменьшает ее точность. Смещение математического ожидания оценки тоже существенно зависит от длительности интервала сброса. Длительность интервала сброса статистики уменьшает систематическую погрешность оценки среднего значения. Таким образом, задача выбора интервала сброса начальной статистики в силу противоречивости критериев требует построения некоторой свертки исходных критериев. В качестве такой свертки предлагается использовать вероятность принадлежности оценки заданному интервалу. При малых значениях ковариации оптимум интервала сброса лежит в окрестности нуля, т.е. не учет любых значений переходного периода приводит к уменьшению вероятности принадлежности интервалу. Выполненное исследование позволяет разработать практические рекомендации по обработке выходного имитационного процесса с целью повышения точности и сокращения сроков моделирования.

Ключевые слова: *имитация, моделирование, гауссовские процессы, среднеинтегральные оценки, дисперсия, математическое ожидание, тренд, нестационарные процессы, автокорреляция, сходимость, оптимизация.*

ANALYSIS OF INFLUENCE OF INTERVAL RESET STATISTICS SIMULATION EXPERIMENT FOR EVALUATING THE ACCURACY OF AVERAGE INTEGRAL

Solntsev Alexey A., *Ph.D., Moscow Automobile and Road State Technical University (MADI), 64 Leningradsky av., Moscow, 125319, Russia, e-mail: kafedra@asu.madi.ru.*

Prikhodko Mikhail V., *Ph.D., Deputy Director of "LonMADI", 64 Leningradsky av., Moscow, 125319, Russia, e-mail: kafedra@asu.madi.ru.*

Kudryavtsev Andrey Yu., *Ph.D., Engineer, company NPVF "WELDING", 9 Cabelnyi passage, Cheboksary, 428003, Chuvash Republic, Russia, e-mail: kafedra@asu.madi.ru.*

Kotov Andrey A., *Ph.D., Moscow Automobile and Road State Technical University (MADI), 64 Leningradsky av., Moscow, 125319, Russia, e-mail: kafedra@asu.madi.ru.*

Most modern methods developed to obtain estimates on the results of simulations suggest stationarity of the output process and, therefore, the rejection of the transition period. This article has evaluated the accuracy of the estimate, taking into account reset statistics accumulated during the initial period of simulation. It is assumed that the covariance function conditionally non-stationary Gaussian process defined on the basis of the theorem on normal correlation for a given stationary Gaussian process with autocorrelation function in the form of a mixture of two exponentials. The analysis of the expectation of mean integral evaluation of conditionally non-stationary process provided clear statistics accumulated during the initial interval of the simulation. The analytical expressions for the expectation and variance of mean integral evaluation depending on the discharge statistics are given. It is shown that the determining factor is the extent of the correlation process, and the duration of the interval reset statistics for the variance significantly reflects on dispersy of integral evaluation. Reset of the initial statistical variance of the estimate increases, and thus reduces its accuracy. The displacement of the expectation of evaluation also depends on the duration of the interval reset. Duration of reset interval statistics reduces the systematic error estimate of the mean value. Thus, the task of choosing the initial reset interval statistics by the contradictory criteria requires the construction of a convolution of the original criteria. As such convolutions are encouraged to use estimates of the probability to a given interval. For small values of the covariance reset interval is optimum in the neighborhood of zero, ie non-record of any values of the transition leads to a decrease in the probability of belonging to the interval. Performed study provides practical recommendations for handling the output of simulation process in order to improve accuracy and reduce the time of modeling.

Key words: *imitation, modeling, Gaussian processes, the mean integral evaluation, variance, expectation, the trend in non-stationary processes, autocorrelation, convergence, optimization.*

Большинство современных методов, разработанных для получения оценок по результатам имитационных экспериментов, предполагают стационарность выходного процесса и в связи с этим отбрасывание переходного периода. В данной статье выполнен анализ точности оценки с учетом сброса статистики, накопленной на начальном периоде моделирования. Предполагается, что ковариационная функция условно нестационарного гауссовского процесса (на основании теоремы о нормальной корреляции) определяется следующим выражением:

$$R(t,u) = r(|t-u|) - D_{\xi\theta}(t)\cdot D_{\theta\theta}^{-1}\cdot D_{\xi\theta}^{T}(u), \quad (t\ge t_1, u\ge t_1), \qquad (1)$$

где $D_{\xi\theta}(t) = (r(t-t_0), r(t-t_{-1}), \ldots, r(t-t_{-m}))$ – вектор-строка ковариаций; $D_{\theta\theta}=\| cov\,(\xi(t_i), \xi(t_j))\| = \| r\,(t_i - t_j)\|$ $i,j=0..-m$ – матрица ковариаций предыстории процесса в моменты t_i, t_j; $r(t)$ – автокорреляционная функция стационарного процесса.

$$r(t) = \sigma^2\left(\alpha_1 e^{-c_1 t} + \alpha_2 e^{-c_2 t}\right), \qquad (2)$$

где $c_1 > 0$ и $c_2 > 0$ параметры автоковариации, а α_1 и α_2 – некоторые функции параметров c_1 и c_2.

Характеристики среднеинтегральной оценки при наличии сброса статистики имитационного эксперимента. В случае проведения имитационного эксперимента с отбрасыванием статистики, накопленной за время переходного периода Δ, математическое ожидание среднеинтегральной оценки примет вид:

$$M\zeta(T,\Delta) = \frac{1}{T-\Delta}\int_{\Delta}^{T} M\xi(t \mid t_0, S_0)dt\,. \tag{3}$$

Функция ***W*** в этом случае имеет вид:

$$W(T,\Delta) = \frac{1}{T-\Delta}\int_{\Delta}^{T} D_{\xi\theta}(t)\cdot dt\,. \tag{4}$$

Стационарная составляющая дисперсии будет представлять следующий интеграл:

$$D_S\zeta(T,\Delta) = \frac{1}{(T-\Delta)^2}\int_{\Delta}^{T}\int_{\Delta}^{T} r(|\,t-u\,|)dtdu\,. \tag{5}$$

Выражение для нестационарной составляющей останется таким же, как и в предыдущем параграфе:

$$D_N\zeta(T,\Delta) = W(T,\Delta)\cdot D_{\theta\theta}^{-1}\cdot W^T(T,\Delta)\,. \tag{6}$$

В результате выражения для математического ожидания и дисперсии можно представить в виде:

$$\begin{aligned} M\zeta(T,\Delta) &= y + W(T,\Delta)\cdot D_{\theta\theta}^{-1}\cdot(S - y\cdot E),\\ D\zeta(T,\Delta) &= D_S\zeta(T,\Delta) - D_N\zeta(T,\Delta). \end{aligned} \tag{7}$$

Выполним анализ стационарной составляющей дисперсии.

Для введенной автокорреляции дисперсия среднеинтегральной оценки стационарного процесса будет равна:

$$D_S\zeta = \frac{1}{(T-\Delta)^2}\int_{\Delta}^{T}\int_{\Delta}^{T} r(|\,t_1-t_2\,|)\cdot dt_1 dt_2 = \frac{2}{(T-\Delta)^2}\int_{\Delta}^{T}\left(\int_{t_2}^{T} r(t_1-t_2)\cdot dt_1\right)dt_2\,. \tag{8}$$

Внутренний интеграл последнего выражения относительно первого слагаемого в выражении для $r(t_1\text{-}t_2)$ равен:

$$\begin{aligned} \int_{t_2}^{T}\alpha_1 e^{-c_1(t_1-t_2)}dt_1 &= \alpha_1 e^{c_1t_2}\int_{t_2}^{T} e^{-c_1t_1}dt_1 = -\frac{\alpha_1}{c_1}e^{c_1t_2}\left[e^{-c_1t_1}\right]_{t_2}^{T} = \\ &= -\frac{\alpha_1}{c_1}e^{c_1t_2}\left[e^{-c_1T} - e^{-c_1t_2}\right] \end{aligned} \tag{9}$$

И окончательно для одной компоненты автокорреляции получим:

$$D_S\zeta = \frac{2\sigma^2}{(T-\Delta)^2}\int_\Delta^T \frac{\alpha_1}{c_1}\left[1-e^{-c_1T}e^{c_1t_2}\right]dt_2 =$$

$$= \frac{2\sigma^2}{(T-\Delta)^2}\frac{\alpha_1}{c_1}\left[t+e^{-c_1T}\left(\frac{1}{c_1}e^{-c_1t_2}\right)\right]_\Delta^T = \tag{10}$$

$$= \frac{2}{(T-\Delta)}\sigma^2\frac{\alpha_1}{c_1} - \frac{2}{(T-\Delta)^2}\sigma^2\frac{\alpha_1}{c_1^2} + \frac{2}{(T-\Delta)^2}\sigma^2\frac{\alpha_1}{c_1^2}e^{-c_1(T-\Delta)}.$$

Выражая дисперсию через введенные функции r_1 и r_2

$$r_1(t) = \sigma^2\left(\frac{\alpha_1}{c_1}e^{-c_1t} - \frac{\alpha_2}{c_2}e^{-c_2t}\right) \qquad r_2(t) = \sigma^2\left(\frac{\alpha_1}{c_1^2}e^{-c_1t} - \frac{\alpha_2}{c_2^2}e^{-c_2t}\right) \tag{11}$$

получим:

$$D_S(\Delta,T) = \frac{2}{T-\Delta}r_1(0) - \frac{2}{(T-\Delta)^2}r_2(0) + \frac{2}{(T-\Delta)^2}r_2(T-\Delta). \tag{12}$$

Выполним анализ нестационарной составляющей дисперсии.

Значение функционала $\boldsymbol{W}$ для нестационарной составляющей равно:

$$W_i(T) = \frac{1}{T-\Delta}\int_\Delta^T r(t-u_i)\cdot dt\,. \tag{13}$$

После подстановки получим:

$$W_i(T) = \frac{1}{T-\Delta}\int_\Delta^T \sigma^2(a_1e^{c_1u_i}e^{-c_1t} + a_2e^{c_2u_i}e^{-c_2t})dt =$$

$$\frac{\sigma^2}{T-\Delta}\left[-\frac{\alpha_1}{c_1}e^{c_1u_i}e^{-c_1t} - \frac{\alpha_2}{c_2}e^{c_2u_i}e^{-c_2t}\right]_\Delta^T =$$

$$= \frac{\sigma^2}{T-\Delta}\left[\frac{\alpha_1}{c_1}e^{c_1u_i}(e^{-c_1\Delta}-e^{-c_1T}) + \frac{\alpha_2}{c_2}e^{c_2u_i}(e^{-c_2\Delta}-e^{-c_2T})\right]$$

Окончательно, используя выражения для r_1 и r_2, получим:

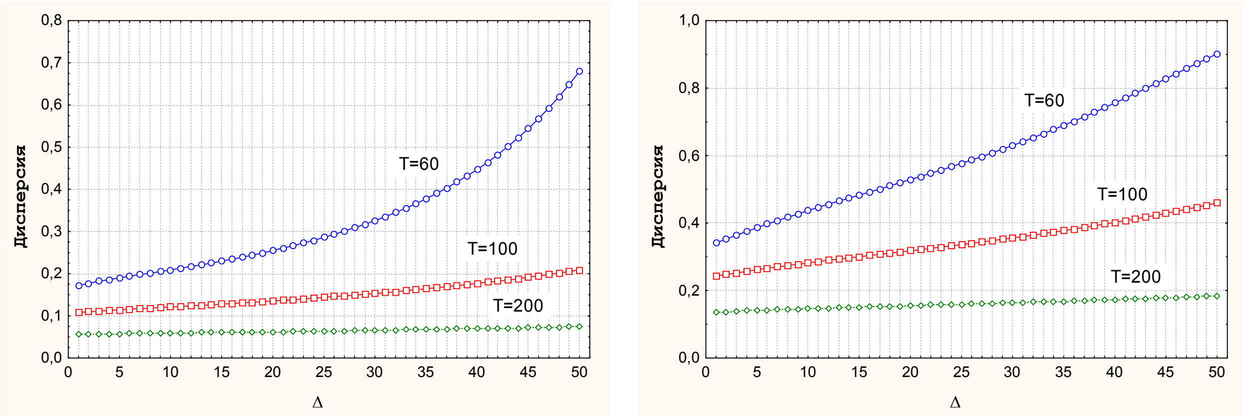
$$W_i(\Delta,T) = \frac{1}{T-\Delta}\left[r_1(\Delta-u_i) - r_1(T-u_i)\right]. \tag{14}$$

Полученные соотношения для дисперсии и математического ожидания позволяют провести анализ влияния сброса на эффективность оценки.

Математическое ожидание и дисперсию оценки следует рассматривать как функции переменных Δ, St, S, Cv, T:

$$M\xi(\Delta) = M_\xi(\Delta \mid St,S,Cv,T),$$
$$D\xi(\Delta) = D_\xi(\Delta \mid St,Cv,T). \tag{15}$$

Поскольку значение дисперсии не зависит от начальных условий моделирования, будем оценивать влияние коррелированности Cv^1 = (0,2; 0,1), Cv^2 = (0,4; 0,3) и общего времени моделирования. На рисунке 1 приведены графики дисперсии среднеинтегральной оценки условно нестационарного процесса в зависимости от длины интервала сброса для различных значений интервалов моделирования $T \in \{60, 100, 200\}$.

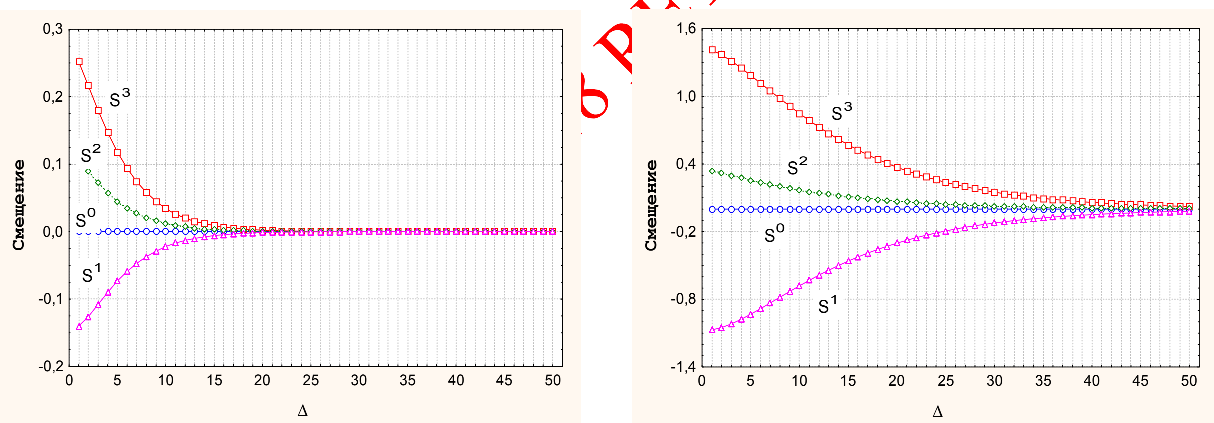


а) б)

Рис. 1. Зависимость дисперсии от длительности интервала сброса: а) $Cv = (0{,}4; 0{,}3)$; б) $Cv = (0{,}2; 0{,}1)$

Из графиков видно, что длительность интервала сброса статистики существенно отражается на дисперсии оценки. В любом случае сброс начальной статистики *увеличивает дисперсию оценки*, т.е. уменьшает точность оценки.

На рисунке 2 приведены графики математического ожидания среднеинтегральной оценки процесса в зависимости от интервала сброса. Значение длительности интервала моделирования равно $T = 200$, и в качестве начальных состояний были выбраны: $St = (0; -1)$; $S^0 = (0;0)$, $S^1 = (5;0)$, $S^2 = (5;5)$, $S^3 = (0;5)$.



а) б)

Рис. 2. Зависимость математического ожидания оценки от интервала сброса:
а) Cv = (0,4; 0,3); б) Cv = (0,2; 0,1)

Из графиков видно, что смещение математического ожидания оценки также существенно зависит от длительности интервала сброса. В любом случае сброс *уменьшает систематическую погрешность* оценки среднего значения.

Таким образом, задача выбора интервала сброса начальной статистики в силу противоречивости критериев требует построения некоторой свертки исходных критериев. В качестве такой свертки предлагается использовать вероятность попадания оценки в заданный интервал погрешности.

Вероятность принадлежности оценки интервалу погрешности. Определим вероятность того, что значение оценки в результате моделирования будет лежать в δ-окрестности среднего значения функции *y*.

Эту вероятность можно рассматривать как критерий выбора длительности интервала сброса статистики. Чем больше вероятность, тем более точной является оценка среднего значения.

Вероятность принадлежности оценки δ-окрестности равна:

$$P(\delta) = \frac{1}{\sqrt{2 \cdot \pi \cdot \mathrm{D}\zeta(\Delta)}} \cdot \int_{y-\delta}^{y+\delta} \exp\left[\frac{(t - \mathrm{M}\zeta(\Delta))^2}{2\mathrm{D}\zeta(\Delta)}\right] dt \, . \tag{16}$$

Очевидно, что

$$P(\delta) = P(\delta \mid T, St, S, Cv, \Delta) \, . \tag{17}$$

Проведем полный 4^4 факторный эксперимент со значениями всех параметров на 4-х уровнях. В таблице 1 приведены значения варьируемых факторов.

Таблица 1

Значения факторов

N	Фактор 1 T	Фактор 2 Δ(D)	δ(G) *100	Фактор 3 S*10 s(0)	s(-1)	Фактор 4 Cv*10 c1	c2
1	300	0	0,2	0	0	0,2	0,1
2	350	50	0,25	0	5	0,3	0,2
3	400	100	0,3	5	5	0,4	0,3
4	450	150	0,35	5	0	0,5	0,4

Значения эффектов взаимодействия факторов приведены в таблице 2.

Таблица 2

Эффекты взаимодействия факторов

Однофакторное		Двухфакторное		Трехфакторное	
1	.159723	12	.001759	123	.000000
2	.288725	13	.000000	124	.000250
3	.000000	23	.000000	134	.000000
4	1.908926	14	.001808	234	.000000
		24	.241754		
		34	.000000		

Из таблицы 2 видно, что самым существенным фактором является коррелированность процесса. Начальные условия практически не влияют на значение критерия. Из двухфакторных взаимодействий существенным является сочетание коррелированности и длительности интервала сброса.

Диаграммы Бокса-Кокса однофакторных влияний для δ = 0,3 приведены на рисунке 3.

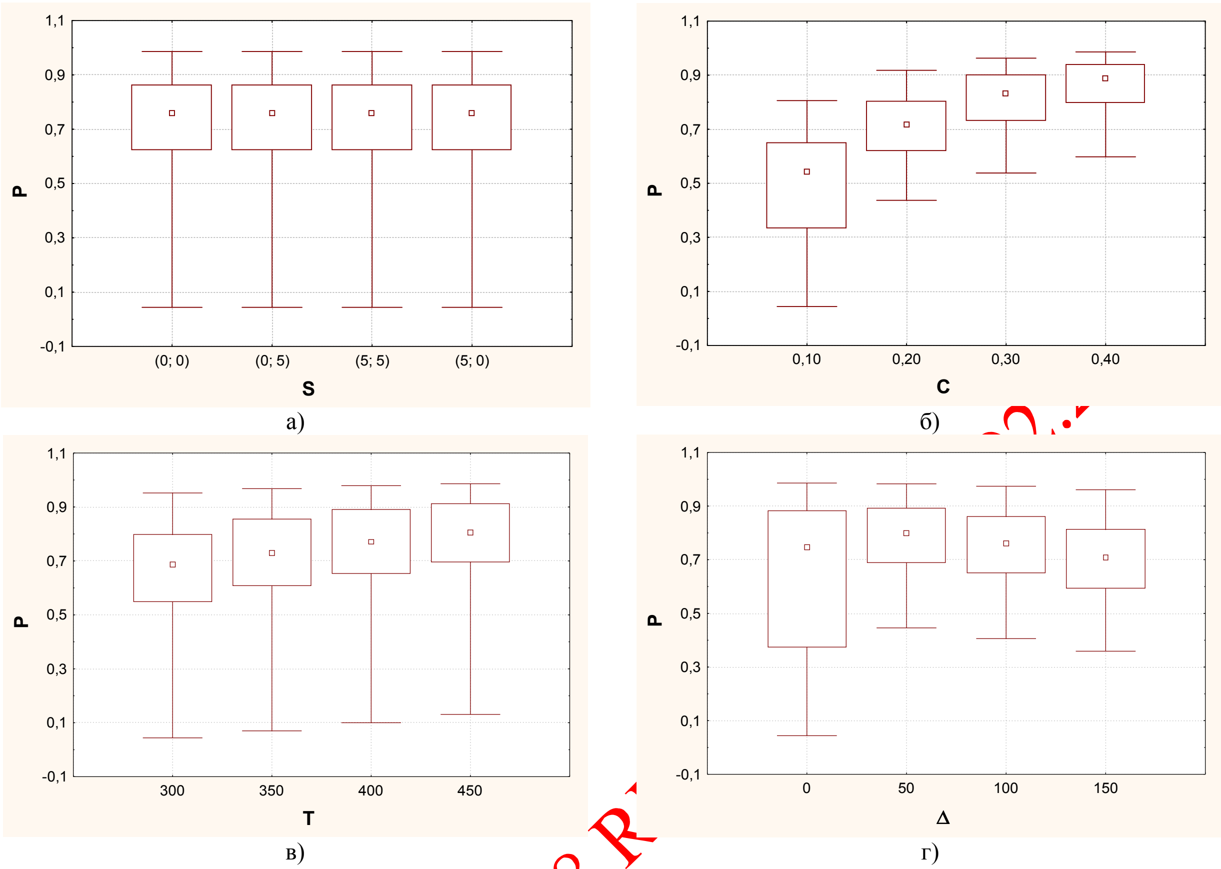


Рис. 3. Графики однофакторного влияния на вероятность:
а) влияние S; б) влияние Cv; в) влияние T; г) влияние Δ

Из диаграмм (рис. 3г) видно, что существует некоторый оптимум длительности интервала сброса. Более наглядно существование оптимума представлено на рисунке 4. Здесь приведены графики усредненной по всем значениям коррелированности вероятности попадания в заданный интервал для каждого начального условия и для каждой длительности интервала моделирования.

На рисунке 5 приведены графики зависимостей вероятности, усредненной по начальным условиям для каждого значения коррелированности и каждой длительности интервала моделирования. Из графиков видно, что при малых значениях ковариации оптимум интервала сброса лежит в окрестности нуля, т.е. не учет любых значений переходного периода приводит к уменьшению вероятности. Более того, малые длительности интервала моделирования делают оптимум более отчетливым.

Для случая необходимости получения грубых оценок средних значений были проведены расчеты в условиях:

- $\sigma = 1$,
- короткий интервал моделирования ($T = 100$),
- широкий доверительный интервал ($\delta = 0{,}5$),
- слабокоррелированный процесс $Cv = (0{,}4;\ 0{,}5)$.

Графики вероятности попадания оценки в δ-окрестность для различных начальных состояний $S^0 = (1;\ 1)$, $S^1 = (1;\ 3)$, $S^2 = (3;\ 1)$, $S^3 = (5;\ 5)$, приведены на рисунке 6.

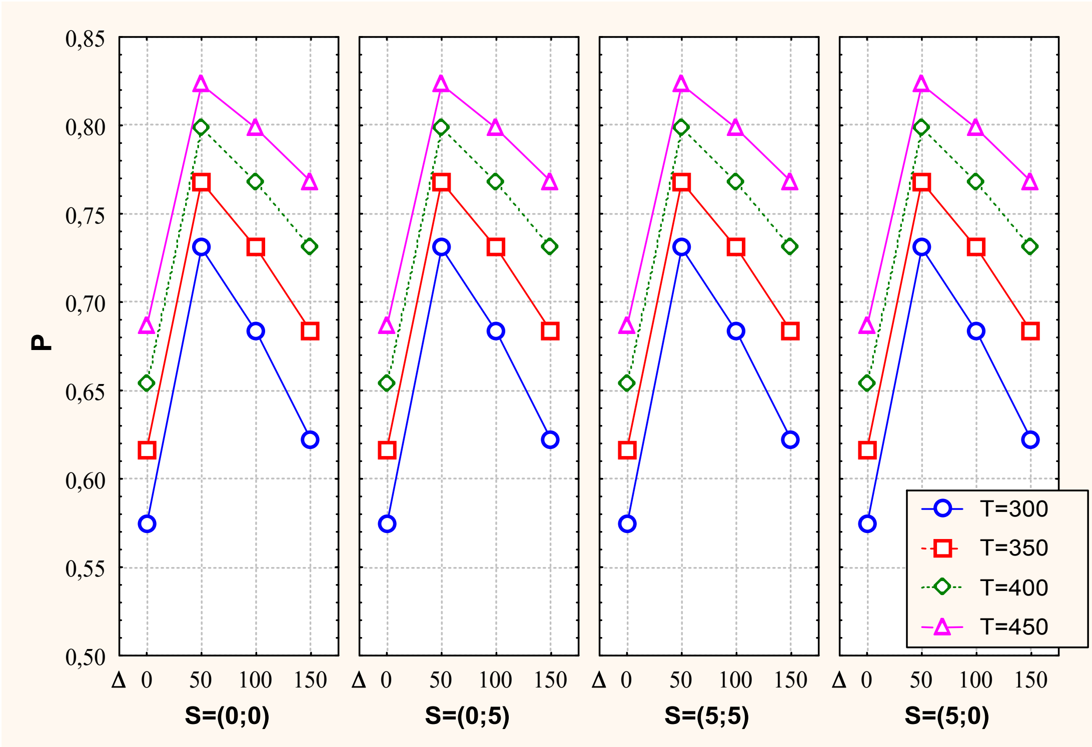


Рис. 4. Усредненные по коррелированности оценки влияния сброса

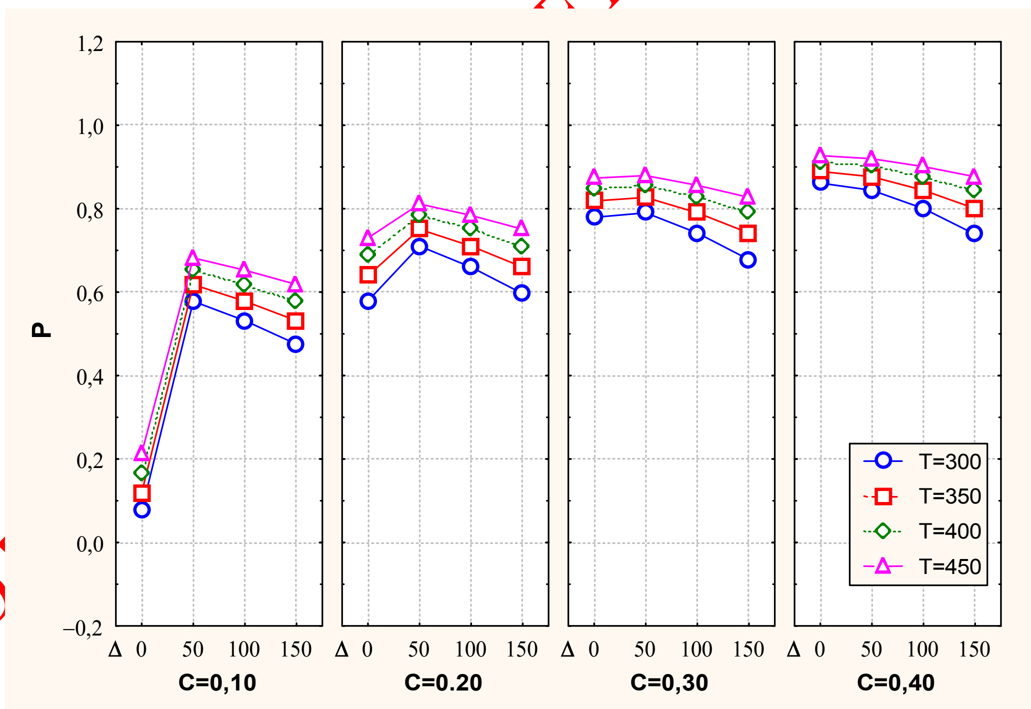


Рис. 5. Усредненные по начальному состоянию оценки влияния сброса

Как видно из графиков, для состояния S^0 нужно учитывать все данные, для S_3 (достаточно далекие начальные условия) оптимальный интервал сброса равен всего 4–5 единицам из 100.

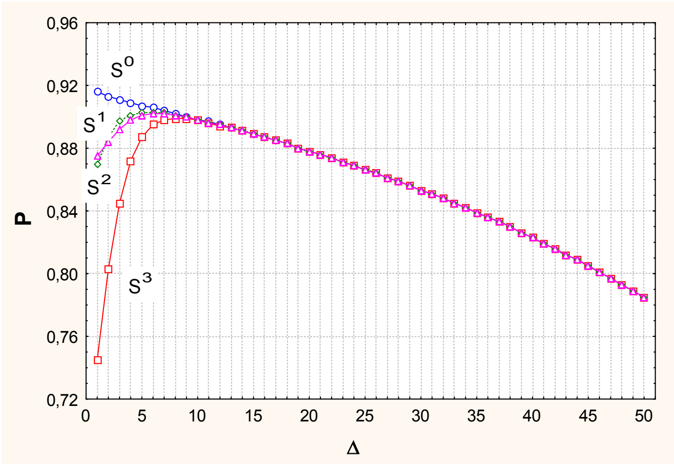


Рис. 6. Влияние сброса на точность для грубых оценок

Таким образом, проведен анализ математического ожидания среднеинтегральной оценки условно нестационарного процесса при условии сброса статистики, накопленной на начальном интервале моделирования, равном Δ. Получены аналитические выражения математического ожидания и дисперсии среднеинтегральной оценки в зависимости от величины Δ. В результате дальнейших исследований показано, что увеличение Δ уменьшает систематическую погрешность, но увеличивает дисперсию оценки. В силу этого предложено выполнить свертку обоих показателей посредством определения вероятности того, что значение оценки в результате моделирования и сброса начальной статистики будет лежать в δ-окрестности стационарного значения. Проведена серия экспериментов с целью выявления факторов, влияющих на этот показатель. Показано, что определяющим фактором является степень коррелированности процесса: при сильной коррелированности сброс необходим, однако при широком доверительном интервале и слабой коррелированности нет необходимости производить сброс статистики. При малых значениях корреляции оптимум интервала сброса начальной статистики лежит в окрестности нуля.

Таким образом, выполненное исследование позволяет оценить влияние автокорреляционных свойств выходных процессов, начальных значений, длительности моделирования, длительности интервала сброса начальной статистики на точность результатов, получаемых в процессе моделирования, и разработать практические рекомендации по обработке выходного имитационного процесса с целью повышения точности и сокращения сроков моделирования.

Предложенные статистические и аналитические модели выходных имитационных процессов позволяют существенно расширить сферу исследований различных процедур организации имитационного моделирования и построения управляющих и оптимизирующих алгоритмов.