М. В. Михайлов, Н. В. Хованов, Л. А. Чудовская

МЕТОД РАНДОМИЗИРОВАННЫХ ТРАЕКТОРИЙ В ЗАДАЧАХ ПРОГНОЗИРОВАНИЯ ДИНАМИКИ ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ*

Введение

Анализ динамики финансово-экономических показателей базируется на исходной эмпирической информации, имеющей вид совокупности конечных временны́х рядов. Под *конечным временны́м рядом* $y(t)$ здесь понимается последовательность значений $y_0 = y(t_0),\ldots, y_m = y(t_m)$ исследуемого финансово-экономического показателя y, наблюдаемых в последовательные моменты времени $t_0,\ldots,t_m$, $t_0 < \ldots < t_m$ соответственно. Дальнейшая статистическая обработка такой исходной эмпирической информации обычно производится на основе теоретико-вероятностной модели, в рамках которой предполагается, что наблюдаемый временной ряд $y(t)$, $t = t_0,\ldots,t_m$ есть реализация («траектория») некоторого *случайного временного ряда* (стохастического процесса $\tilde{y}(t)$ с дискретным временем $t = t_0,\ldots,t_m$). Иными словами, вектор наблюдаемых значений $(y_0,\ldots,y_m)$ исследуемого показателя y интерпретируется как выборочное значение случайного вектора $(\tilde{y}_0,\ldots,\tilde{y}_m)$, каждая компонента $\tilde{y}_i = \tilde{y}(t_i)$ которого есть одномерное сечение случайного процесса $\tilde{y}(t)$. Этот гипотетический (напрямую не наблюдаемый) случайный временной ряд $\tilde{y}(t)$ описывается вероятностным пространством $(Y, \mathrm{A}_\sigma, P)$, где $Y = \{y(t;\theta), \theta \in \Theta\}$ есть множество всех возможных траекторий процесса $\tilde{y}(t)$, A_σ — сигма-алгебра подмножеств множества Y, а P определяет вероятность $P(A)$ любого события $A \in \mathrm{A}_\sigma$[1].

---

*Работа выполнена при поддержке Российского фонда фундаментальных исследований (Проект 09/06/00119).

Михаил Витальевич МИХАЙЛОВ — канд. экон. наук, доцент кафедры экономической кибернетики СПбГУ. В 1982 г. окончил Экономический факультет ЛГУ. В 1988 г. защитил кандидатскую диссертацию. Область научных интересов — многокритериальное оценивание экономических объектов в условиях неопределенности и дефицита информации, имитационное моделирование экономических процессов.

Николай Васильевич ХОВАНОВ — д-р физ.-мат. наук, профессор кафедры экономической кибернетики Экономического факультета СПбГУ. Область научных интересов — стохастические модели риска и неопределенности, теория и методы принятия решений в условиях дефицита информации.

Людмила Анатольевна ЧУДОВСКАЯ — старший преподаватель кафедры высшей математики С.-Петербургской государственной лесотехнической академии (СПбГЛТА), соискатель кафедры экономической кибернетики Экономического факультета СПбГУ. В 1982 г. окончила математико-механический факультет ЛГУ. Область научных интересов — стохастические модели оценки и прогнозирования временных рядов значений финансово-экономических показателей.

Однако описанный стандартный теоретико-вероятностный подход к интерпретации наблюдаемого временного ряда $y(t_0),\ldots,y(t_m)$ как реализация непосредственно ненаблюдаемого случайного временного ряда $\tilde{y}(t_0),\ldots,\tilde{y}(t_m)$, который описывается вероятностным пространством $(Y, \mathrm{A}_\sigma, P)$, вызывает ряд замечаний со стороны разных групп исследователей. Во-первых, некоторые специалисты отвергают саму идею описания эмпирически наблюдаемых временны́х рядов при помощи сложных и принципиально ненаблюдаемых математических схем[2]. Во-вторых, ряд исследователей указывают на возможность статистической обработки конечных временны́х рядов с использованием более простой и наглядной, чем теоретико-вероятностная конструкция, «геометрико-механической» интерпретации, согласно которой математическое ожидание случайной величины трактуется как центр распределенной вероятностной массы, а дисперсия — как момент инерции этой массы и т. д.[3] Наконец, в-третьих, многие специалисты, не отказываясь от использований концепции вероятностного пространства $(Y, \mathrm{A}_\sigma, P)$ для интерпретации наблюдаемых конечных временны́х рядов значений исследуемых финансово-экономических показателей, призывают максимально точно и осторожно применять абстрактные математические схемы, учитывая их неизбежные ограничения и приближенный характер описания реальных процессов[4]. Особенно важно учитывать ограниченность точности и достоверности теоретико-вероятностных объяснений в случае коротких временны́х рядов, сильно затрудняющих прогнозирование будущей динамики значений исследуемого показателя[5], а также при наличии неопределенности разных видов, возникающей при дефиците эмпирических данных и не сводящейся к неопределенности теоретико-вероятностного вида[6].

В настоящей статье разрабатываются теоретико-вероятностные модели, позволяющие (в определенных ситуациях и в определенной мере, разумеется) снизить влияние описанных выше проблем на точность и достоверность прогнозирования динамики значений финансово-экономических показателей на основе учета при построении соответствующих вероятностных пространств нечисловой, неточной и неполной экспертной информации.

В первом разделе статьи излагаются основы метода рандомизированных траекторий, в основе которого лежит байесовский подход к моделированию неопределенности выбора элемента из конечного множества однородных элементов при помощи задания вероятностной меры на этом конечном множестве. Второй раздел посвящен важному частному случаю рандомизации выбора элемента из конечного множества монотонных траекторий, связанному с построением соответствующего стохастического процесса с равновероятными дискретными монотонными траекториями. В третьем разделе подробно разобран пример использования построенного стохастического процесса с равновероятными монотонными траекториями для прогнозирования динамики цены облигации по нечисловой, неточной и неполной экспертной информации. В заключение сделан вывод о практической значимости разработанного метода рандомизированных траекторий в задачах прогнозирования динамики финансово-экономических показателей.

Метод рандомизированных траекторий

Уже в первой четверти прошлого века среди исследователей, занимающихся вопросами оценки и прогнозирования динамики финансово-экономических показателей, возникло представление о различных типах неопределенности такой оценки и прогнозирования. Четкое различение двух основных типов неопределенности проведено, например, в известной монографии Фрэнка Найта, вышедшей в 1921 г.[7]

Несколько модернизировав, обобщив и формализовав определения Ф. Найта, можно сказать, что выделенная им *неопределенность первого рода* связана с ситуацией, когда некоторый элемент y (например, значение финансово-экономического показателя или график функции, описывающей динамику этого показателя) известен исследователю «с точностью до множества Y». Иными словами, при неопределенности первого рода результатом оценки является указание некоторого множества Y, содержащего элемент y. Такую неопределенность, при которой исследователь знает только то, что оцениваемый (прогнозируемый) элемент y принадлежит некоторому множеству Y, будем далее называть *теоретико-множественной неопределенностью*.

Введенную Ф. Найтом *неопределенность второго рода* можно связать с ситуацией, когда кроме множества Y, содержащего элемент y, исследователю известно еще и распределение вероятности P, заданное на некоторой сигма-алгебре A_σ подмножеств множества Y. Распределение P определяет вероятность $P(Y')$ того, что элемент y содержится в подмножестве Y' множества Y, являющегося элементом сигма-алгебры A_σ. В простейшем случае конечного множества Y введенная неопределенность второго рода предусматривает прямое указание вероятности $p_i = P(\{y_i\})$ появления элемента $y_i \in Y$ при соответствующем случайном испытании. Далее будем называть описанную неопределенность второго рода *теоретико-вероятностной неопределенностью*.

Пусть перед исследователем стоит задача оценки (прогнозирования) значений $y(t_0),\dots,y(t_m)$, принимаемых изучаемым финансово-экономическим показателем в (будущие) моменты времени $t = t_0,\dots,t_m$, $t_0 < \dots < t_m$. Пусть задано конечное *множество* $Y = \{y(\theta) : \theta = 1,\dots,N\}$ *всех возможных траекторий* $y(\theta) = (y(t_0;\theta),\dots,y(t_m;\theta))$ временного ряда $y(t_0),\dots,y(t_m)$. Помимо значений $y(t_0;\theta),\dots,y(t_m;\theta)$ функции $y = y(t;\theta)$ будем рассматривать и приращения $d(t_i;\theta) = y(t_i;\theta) - y(t_{i-1};\theta)$, $i = 1,\dots,m$ этой функции. Очевидно, что значение функции $y = y(t;\theta)$ в точке $t = t_i$ определяется формулой $y(t_i;\theta) = y(t_0;\theta) + d(t_1;\theta) + \dots + d(t_i;\theta)$.

На основе экспертной информации I о значениях функций $y = y(t;\theta)$ и $d = d(t;\theta)$ возможна селекция элементов множества $Y = \{y(\theta) : \theta = 1,\dots,N\}$ всех допустимых траекторий $y(\theta) = (y(t_0;\theta),\dots,y(t_m;\theta))$ временно́го ряда $y(t_0),\dots,y(t_m)$. В результате селекции элементов множества Y, удовлетворяющих требованиям (ограничениям) информации I, формируется *множество* $Y(I)$ *всех допустимых* (*с точки зрения информации* I) функций $y = y(t;\theta)$, содержащее $N(I) \le N$ элементов. Если выполняется строгое неравенство $N(I) < N$, то можно говорить о *нетривиальной информации* I (о *нетривиальных ограничениях*, описываемых информацией I). Если же экспертная информация I *тривиальна* (множество ограничений, описываемых информацией I, является фактически *пустым* — $I = \varnothing$), то $N(I) = N(\varnothing) = N$.

Экспертная информация не носит, как правило, *числового характера* и может быть выражена лишь чисто сравнительными утверждениями типа «значение $y_i = y(t_i;\theta)$ функции $y = y(t;\theta)$ больше значения $y_j = y(t_j;\theta)$ этой же функции», «приращение $d(t;\theta)$ функции $y = y(t;\theta)$ в точке $t = t_i$ равно приращению этой функции в точке $t = t_j$» и т. п. Далее мы будем предполагать, что такая *нечисловая информация* может быть представлена в виде системы равенств и неравенств $IO = \{y(t_i;\theta) > y(t_j;\theta);\ d(t_i;\theta) = d(t_j;\theta); \dots\}$ для значений и приращений функций $y = y(t;\theta)$. Естественно назвать экспертную информацию IO, выражаемую указанной системой равенств и неравенств, *ординальной* (*порядковой*) *информацией*.

Помимо ординальной информации исследователь может также иметь и *неточную экспертную информацию* II о числовых значениях и приращениях функций $y = y(t;\theta)$, которую можно представить в виде системы $II = \{y_i^- \le y(t_i;\theta) \le y_i^+, d_j^- \le d(t_j;\theta) \le d_j^+, \ldots\}$ неравенств, указывающих возможные диапазоны варьирования значений и приращений функций $y = y(t;\theta)$. Естественно назвать экспертную информацию II, выражаемую системой указанных неравенств, *интервальной информацией*.

Объединяя системы неравенств IO и II, получаем *нечисловую и неточную информацию* $I = IO \cup II$ о значениях и приращениях функций $y = y(t;\theta)$. При этом возможно, что объединенная система равенств и неравенств I определяет функцию $y = y(t;\theta)$ не однозначно, а лишь с точностью до конечного *множества* $Y(I)$ *всех допустимых* (с точки зрения экспертной информации I) траекторий $y(\theta) = (y(t_0;\theta), \ldots, y(t_m;\theta))$, $\theta = 1, \ldots, N(I)$, $N(I) \le N$. Поэтому далее мы будем говорить о *нечисловой (ординальной), неточной (интервальной) и неполной информации* (ННН-*информации*) I о значениях и приращениях функций $y = y(t;\theta)$.

Измерить количество ННН-информации I можно, например, при помощи *коэффициента селекции* $SEL(I) = N/N(I)$. Для измерения количества ННН-информации в двоичных единицах (в битах) следует воспользоваться *логарифмической мерой информации* $INF(I)$, определяемой формулой $INF(I) = \log_2 SEL(I)$, где $\log_2 x$ есть двоичный логарифм числа x.

Итак, после построения с помощью экспертной ННН-информации I множества всех допустимых траекторий исследователь находится в условиях теоретико-множественной неопределенности, когда оцениваемая (прогнозируемая) траектория $y(\theta) = (y(t_0;\theta), \ldots, y(t_m;\theta))$ известна с точностью до конечного множества $Y(I)$, состоящего не менее чем из двух элементов $N(I) > 1$ [8]. Для моделирования неопределенности выбора конкретного элемента $y(\theta)$ из множества $Y(I)$ можно предложить подход, основанный на известной работе Т. Байеса[9] и состоящий в *рандомизации* такого выбора: траектория $y(\theta)$ случайно выбирается из множества $Y(I)$ с вероятностью $P(\theta) > 0$, $P(1) + \ldots + P(N(I)) = 1$. В результате такой рандомизации исследователь оказывается в ситуации, которая соответствует теоретико-вероятностной неопределенности, описываемой случайным вектором (стохастическим процессом, случайной функцией) $\tilde{y}(t) = y(t;\tilde{\theta}) = (y(t_0;\tilde{\theta}), \ldots, y(t_m;\tilde{\theta}))$[10].

Теперь можно предложить в качестве оценки (прогноза) случайного значения $\tilde{y}_i = \tilde{y}(t_i) = y(t_i;\tilde{\theta})$ исследуемого финансово-экономического показателя математическое ожидание $\bar{y}_i = \bar{y}(t_i) = y(t_i;1)P(1) + \ldots + y(t_i;N(I))P(N(I))$ случайной величины $\tilde{y}_i = \tilde{y}(t_i)$ (одномерного сечения стохастического временного ряда $\tilde{y}(t)$, $t = t_0, \ldots, t_m$). Точность полученной оценки $\bar{y}_i$ естественно измерять величиной стандартного отклонения $\sigma_i = \sqrt{D\tilde{y}_i}$, где $D\tilde{y}_i$ есть дисперсия случайной величины $\tilde{y}_i = \tilde{y}(t_i)$, вычисляемая по формуле $D\tilde{y}_i = [y(t_i;1) - \bar{y}_i]^2 P(1) + \ldots + [y(t_i;N(I)) - \bar{y}_i]^2 P(N(I))$.

Обычно при рандомизации теоретико-множественной неопределенности в качестве распределения вероятностей $P(\theta)$ выбирается равномерное распределение $P(\theta) = 1/N(I)$, соответствующее «максимальному дефициту информации», который имеется у исследователя[11]. В этом случае для искомых оценок $\bar{y}_i = \bar{y}(t_i)$ и для определения мер $\sigma_i = \sqrt{D\tilde{y}_i}$ их точности получаются наиболее простые вычислительные формулы $\bar{y}_i = 1/N(I) \cdot [y(t_i;1) + \ldots + y(t_i;N(I))]$ и $D\tilde{y}_i = 1/N(I) \cdot \{[y(t_i;1) - \bar{y}_i]^2 + \ldots + [y(t_i;N(I)) - \bar{y}_i]^2\}$ соответственно.

Стохастические процессы с дискретными монотонными траекториями

Рассмотрим важный частный случай применения метода рандомизированных траекторий, который был описан выше, для оценки (прогнозирования) временны́х рядов с монотонными (неубывающими или невозрастающими) реализациями.

Пусть опять перед исследователем стоит задача оценки (прогнозирования) значений $y(t_0),\ldots,y(t_m)$, принимаемых изучаемым финансово-экономическим показателем в моменты времени $t=t_0,\ldots,t_m$, $t_0<\ldots<t_m$. Пусть задано конечное *множество* $Y(m,n)=$ $=\{y(t;m,n;\theta):\theta=1,\ldots,N(m,n)$ *всех возможных траекторий* $y(t;m,n;\theta)$, $t\in\{t_0,\ldots,t_m\}$, $y(t;\theta)=y(t;m,n;\theta)\in\{y_0,\ldots,y_n\}$, $y_0<\ldots<y_n$, временного ряда $y(t_0),\ldots,y(t_m)$, удовлетворяющих *условию монотонности* ($y(t_{i-1};\theta)\le y(t_i;\theta)$) и двум *краевым условиям* ($y(t_0;\theta)=y_0$, $y(t_m;\theta)=y_n$). Помимо значений $y_j=y(t_i;\theta)$ функции $y=y(t;\theta)$ будем рассматривать и неотрицательные приращения $d(t_i;\theta)=y(t_i;\theta)-y(t_{i-1};\theta)\ge 0$, $i=1,\ldots,m$ этой функции. Очевидно, что значение функции $y=y(t;\theta)$ в точке $t=t_i$ определяется формулой $y(t_i;\theta)=y(t_0;\theta)+d(t_1;\theta)+\ldots+d(t_i;\theta)$. Поэтому каждой траектории $y(t;m,n;\theta)$ из множества $Y(m,n)$ сопоставляется соответствующий набор приращений $d(t_i;m,n;\theta)=d(t_i;\theta)$, $i=1,\ldots,m$ из множества $D(m,n)=\{d(t;m,n;\theta):\theta=1,\ldots,N(m,n)\}$.

Далее будем рассматривать простейший вариант, когда имеются равноотстоящие моменты времени $t_i=t_{i-1}+h=t_0+ih$ ($h=[(t_m-t_0)/m]>0$ — шаг отсчета времени), $i=1,\ldots,m$ и равноотстоящие возможные значения $y_j=y_{j-1}+u=y_0+ju$ ($u=[(y_n-y_0)/n]>0$ — шаг отсчета показателя), $j=1,\ldots,n$ исследуемого финансово-экономического показателя. В этом случае можно установить взаимно однозначное соответствие между множествами $Y(m,n)=\{y(t;m,n;\theta):\theta=1,\ldots,N(m,n)\}$ и $J(m,n)=\{j(i;m,n;\theta):\theta=1,\ldots,N(m,n)\}$, где $J(m,n)$ есть множество *всех возможных траекторий* $j(i;\theta)=j(i;m,n;\theta)$, $i\in\{0,1,\ldots,m\}$, $j(i;\theta)\in\{0,1,\ldots,n\}$, заданных на плоской целочисленной решетке $[0,m]\times[0,n]=\{(i,j):i=$ $=0,1,\ldots,m,\ j=0,1,\ldots,n\}$ и удовлетворяющих *условию монотонности* ($j(i-1;\theta)\le j(i;\theta)$), а также двум *краевым условиям* ($j(0;\theta)=0$, $j(m;\theta)=n$). Указанное взаимно однозначное соответствие множеств $Y(m,n)$ и $J(m,n)$ устанавливается формулой $y_j=y(t_i;\theta)=$ $=y(t_0;\theta)+j(i;\theta)u=y(t_0;\theta)+j(i;\theta)[\{y_n-y_0\}/n]$, $i=0,1,\ldots,m$.

Помимо функции $j=j(i;\theta)$ будем рассматривать и неотрицательные приращения $\delta(i;\theta)=j(i;\theta)-j(i-1;\theta)\ge 0$, $i=1,\ldots,m$ этой функции. Очевидно, что значение функции $j=j(i;\theta)$ определяется формулой $j(i;\theta)=\delta(1;\theta)+\ldots+\delta(i;\theta)$. Взаимно однозначное соответствие между множеством $D(m,n)$ приращений $d(t_i;\theta)$ и множеством $\Delta(m,n)=\{\delta(i;m,n;\theta):\theta=1,\ldots,N(m,n)\}$ приращений $\delta(i;\theta)=\delta(i;m,n;\theta)$ устанавливается формулой $d(t_i;\theta)=\delta(i;\theta)u=\delta(i;\theta)[(y_n-y_0)/n]$, $i=0,1,\ldots,m$.

Равновероятная рандомизация теоретико-множественной неопределенности выбора конкретной траектории $j(i;\theta)=j(i;m,n;\theta)$ из множества $J(m,n)$ дает стохастический процесс $\tilde{j}(i)=\tilde{j}(i;m,n)=j(i;m,n;\tilde{\theta})$, порожденный равномерно распределенным случайным параметром $\tilde{\theta}:P(\{\tilde{\theta}=\theta\})=1/N(m,n)$, где число $N(m,n)$ элементов множества $J(m,n)$ определяется известной формулой $N(m,n)=(n+m-1)!/[n!(m-1)!]$[12].

Математическое ожидание $\mu_{\tilde{j}}(i)=E\tilde{j}(i)$ и дисперсия $\sigma^2_{\tilde{j}}(i)=D\tilde{j}(i)$ стохастического процесса $\tilde{j}(i)=\tilde{j}(i;m,n)=j(i;m,n;\tilde{\theta})$ определяются по формулам $\mu_{\tilde{j}}(i)=n[i/m]$ и $\sigma^2_{\tilde{j}}(i)=n^2[i(m-i)]/[m^2(m+1)]+n[i(m-i)]/[m(m+1)]$ соответственно. Отсюда находим математическое ожидание $\mu_{\tilde{y}}(t_i)=E\tilde{y}(t_i)$ и дисперсию $\sigma^2_{\tilde{y}}(t_i)=D\tilde{y}(t_i)$ стохастического процесса $\tilde{y}(t_i)=y(t_0)+\tilde{j}(i)u$ по формулам $\mu_{\tilde{y}}(t_i)=y(t_0)+[y(t_n)-y(t_0)][i/m]$ и $\sigma^2_{\tilde{y}}(t_i)=[i(m-i)]/[m^2(m+1)]+(1/n)[i(m-i)]/[m(m+1)]$.

Найденный тренд $\mu_{\tilde{y}}(t_i)$ стохастического процесса $\tilde{y}(t_i)$ может использоваться как искомая оценка (прогноз) значений исследуемого финансово-экономического показателя на моменты времени $t_0,\ldots,t_m$. Наглядное представление о точности полученных оценок $\mu_{\tilde{y}}(t_i)$, $i=0,1,\ldots,m$ дает область вокруг тренда $\mu_{\tilde{y}}(t_i)$, ограниченная графиками функций $\mu^-_{\tilde{y}}(t_i)=\mu_{\tilde{y}}(t_i)-\sigma_{\tilde{y}}(t_i)$, $\mu^+_{\tilde{y}}(t_i)=\mu_{\tilde{y}}(t_i)+\sigma_{\tilde{y}}(t_i)$, где $\sigma_{\tilde{y}}(t_i)=\sqrt{D\tilde{y}(t_i)}$ есть стандартное отклонение случайного процесса $\tilde{y}(t_i)$, $i=0,1,\ldots,m$.

Пусть теперь исследователь располагает нечисловой, неточной и неполной экспертной информацией I_y о значениях $y(t_1;\theta),\ldots,y(t_{m-1};\theta)$ траекторий из множества $Y(m,n)$, а также аналогичной ННН-информацией I_d о соответствующих наборах $d(t_1;\theta),\ldots,d(t_m;\theta)$ приращений из множества $D(m,n)$. Объединенная ННН-информация $I = I_y \cup I_d$ позволяет построить множество траекторий $Y(m,n;I) \subseteq Y(m,n)$, содержащее число элементов $N(m,n;I) \le N(m,n)$: $Y(m,n;I) = \{y(t;m,n;\tau) : \tau = 1,\ldots,N(m,n;I)\}$.

Умея генерировать траектории из множества $Y(m,n;I)$ [13], можно сосчитать математическое ожидание $\mu_{\tilde{y}}(t_i;I) = E\,\tilde{y}(t_i;I)$ и дисперсию $\sigma^2_{\tilde{y}}(t_i;I) = D\,\tilde{y}(t_i;I)$ стохастического процесса $\tilde{y}(t_i;I)$ с равновероятными реализациями из множества $Y(m,n;I)$ по формулам

$$\mu_{\tilde{y}}(t_i;I) = \frac{1}{N(m,n;I)} \sum_{\tau=1}^{N(m,n;I)} y(t_i;m,n;\tau),$$

$$\sigma^2_{\tilde{y}}(t_i;I) = \frac{1}{N(m,n;I)} \sum_{\tau=1}^{N(m,n;I)} [\,y(t_i;m,n;\tau) - \mu_{\tilde{y}}(t_i;I)\,]^2$$

соответственно.

Найденный тренд $\mu_{\tilde{y}}(t_i;I)$ стохастического процесса $\tilde{y}(t_i;I)$ может использоваться как искомая оценка (прогноз) значений исследуемого финансово-экономического показателя на моменты времени $t_0,\ldots,t_m$. Наглядное представление о точности полученных оценок $\mu_{\tilde{y}}(t_i;I)$, $i = 0,1,\ldots,m$ дает область вокруг тренда $\mu_{\tilde{y}}(t_i;I)$, ограниченная графиками функций $\mu^-_{\tilde{y}}(t_i;I) = \mu_{\tilde{y}}(t_i;I) - \sigma_{\tilde{y}}(t_i;I)$, $\mu^+_{\tilde{y}}(t_i;I) = \mu_{\tilde{y}}(t_i;I) + \sigma_{\tilde{y}}(t_i;I)$, где $\sigma_{\tilde{y}}(t_i;I) = \sqrt{D\,\tilde{y}(t_i;I)}$ есть стандартное отклонение случайного процесса $\tilde{y}(t_i;I)$, $i = 0,1,\ldots,m$.

Прогнозирование цены облигации методом рандомизированных траекторий

Применим изложенную в предыдущем параграфе модификацию метода рандомизированных траекторий для прогнозирования временного ряда значений цены облигации по нечисловой, неточной и неполной экспертной информации. Рассмотрим динамику цены некоторой облигации, имеющей номинал 100 руб. и продающейся в момент эмиссии по цене 80 руб. Облигация погашается через шесть месяцев по номинальной цене 100 руб. Требуется оценить значение цены облигации на конец каждого из шести месяцев.

Построим сначала множество $Y = \{y(t;\theta), \theta \in \Theta\}$ возможных траекторий $y^{(\theta)}(t)$ значения цены облигации, предполагая, что эмиссия соответствует моменту времени $t = 0$, а единицей измерения времени является один месяц. Дополнительно укажем, что цена акции измеряется с точностью до рубля — $y(t;\theta) \in \{80,81,\ldots,99,100\}$. Тогда множество $Y = Y(6,20) = \{y(t;\theta),\ t = 0,1,\ldots,6;\ \theta = 1,\ldots, N(6,20)\}$ всех возможных траекторий конечно ($\Theta = \{1,\ldots,N(6,20)\}$, $N(6,20) = 53130$), а каждая траектория цены является монотонно неубывающей дискретной функцией $y(t;\theta) \in \{80,81,\ldots,99,100\}$ дискретного аргумента $t \in \{0,1,\ldots,6\}$ и представляет собой набор точек $(0,y(0,\theta)) = (0,80), (1,y(1;\theta)),\ldots,(5,y(5;\theta)), (6,y(6;\theta)) = (6,100)$, где $y(t-1;\theta) \le y(t;\theta)$, $t = 1,\ldots,6$. Приращения $d(t;\theta)$, $t = 1,\ldots,6$, траектории $y(t;\theta)$, задаваемые соотношением $d(t;\theta) = y(t;\theta) - y(t-1;\theta)$ позволяют определить значение функции $y(t;\theta)$ в точке t как сумму $y(t;\theta) = d(1;\theta) + \ldots + d(t;\theta)$.

Построенному множеству $Y(6,20)=\{y(t;\theta), t=0,1,...,6, \theta=1,...,N(6,20)\}$ можно взаимно однозначно сопоставить множество $J(6,20)=\{j(i;\theta), i=0,1,...,6, \theta=1,...,N(6,20)\}$ всех возможных дискретных монотонных путей (траекторий) $j(i;\theta)$ на целочисленной решетке $[0,6]\times[0,20]=\{(i,j): i=0,1,...,6; j=0,1,...,20\}$, принимающих дискретные значения из множества $\{0,1, ..., 20\}$ и удовлетворяющих условиям $j(i-1;\theta)\le j(i;\theta)$, $j(0;\theta)=0$, $j(6;\theta)=20$. Более того, в рассматриваемом случае траектории $j(i;\theta)$ и $y(t;\theta)$ связаны в каждой точке $t=i$ простым соотношением $y(t;\theta)=80+j(t;\theta)$, $t=0,1,...,6$. Поэтому далее мы будем изучать, в основном, непосредственно траектории $y(t;\theta)$, опуская соответствующие описания, связанные с траекториями $j(i;\theta)$.

Будем моделировать неопределенность выбора траектории $y(i;\theta)$ из множества всех возможных траекторий $Y(6,20)$ при помощи стохастического процесса $\tilde{y}(i)=$ $=\tilde{y}(t;m,n)=\tilde{y}(t;6,20)$, реализациями (траекториями) которого служат дискретные функции $y(t;\theta)=80+j(t;\theta)$, $\theta=1,...,N(m,n)=53130$, дискретного аргумента $t=1,....,m=6$. В случае, когда отсутствует дополнительная экспертная информация I о вероятностях появления траекторий $y(t;\theta)$ ($I=\varnothing$), стохастический процесс $\tilde{y}(t;6,20)$ может быть задан, как это уже было отмечено выше, при помощи равномерно распределенного на множестве $\Theta=\{1,...,N(6,20)\}$ случайного параметра $\tilde{\theta}$: $P(\{\tilde{y}(t;6,20)=y(t;6,20;\theta), t=0,1,...,6\})=P(\{\tilde{\theta}=\theta\})=1/N(6,20)=1/53130$.

Тогда для математического ожидания $\mu_{\tilde{y}}(t)=n[t/m]$ стохастического процесса $\tilde{y}(i;m,n)$ получаем искомую оценку $\mu_{\tilde{y}}(t)=80+t\,[10/3]\approx 80+3.3333t$. Аналогично, для дисперсии $\sigma^2_{\tilde{y}}(t)=[t(m-t)]/[m^2(m+1)]+(1/n)[t(m-t)]/[m(m+1)]$ и стандартного отклонения $\sigma_{\tilde{y}}(t)=\sqrt{D\tilde{y}(t)}$ стохастического процесса $\tilde{y}(i;m,n)$ получаем искомые оценки $\sigma^2_{\tilde{y}}(t)=t(6-t)[20\cdot 26]/[7\cdot 36]\approx 2.0635t(6-t)$ и $\sigma_{\tilde{y}}(t)\approx 1.4365\sqrt{t(6-t)}$ соответственно.

Теперь помимо тренда $\mu_{\tilde{y}}(t)$ стохастического процесса $\tilde{y}(t;6,20)$ можно ввести ожидаемые нижнюю $\mu^-_{\tilde{y}}(t)=\mu_{\tilde{y}}(t)-\sigma_{\tilde{y}}(t)$ и верхнюю $\mu^+_{\tilde{y}}(t)=\mu_{\tilde{y}}(t)+\sigma_{\tilde{y}}(t)$ границы для возможных траекторий этого процесса. Значения функций $\mu(t)=\mu_{\tilde{y}}(t)$, $\sigma_{\tilde{y}}(t)$, $\mu_-(t)=\mu^-_{\tilde{y}}(t)$, $\mu_+(t)=\mu^+_{\tilde{y}}(t)$ приведены в табл. 1, а соответствующие графики изображены на рис. 1.

Таблица 1

Значения функций

$\mu(t)=\mu_{\tilde{y}}(t)$, $\sigma_{\tilde{y}}(t)$, $\mu_-(t)=\mu^-_{\tilde{y}}(t)$, $\mu_+(t)=\mu^+_{\tilde{y}}(t)$

t	$\mu_{\tilde{y}}(t)$	$\sigma_{\tilde{y}}(t)$	$\mu^-_{\tilde{y}}(t)$	$\mu^+_{\tilde{y}}(t)$
0	0	0	0	0
1	83.33	3.21	80.12	86.55
2	86.67	4.06	82.60	90.73
3	90.00	4.31	85.69	94.31
4	93.33	4.06	89.27	17.4
5	96.67	3.21	93.45	99.88
6	100	0	100	100

Рассмотрим теперь другую информационную ситуацию, в которой исследователь обладает определенным количеством нечисловой, неточной и неполной информации $I\ne\varnothing$. Пусть, например, эксперт описывает свои представления о скорости роста цены облигации, указывая систему неравенств $ID=\{d(1)<d(2)<d(3)=d(4)>d(5)>d6)\}$ для приращений

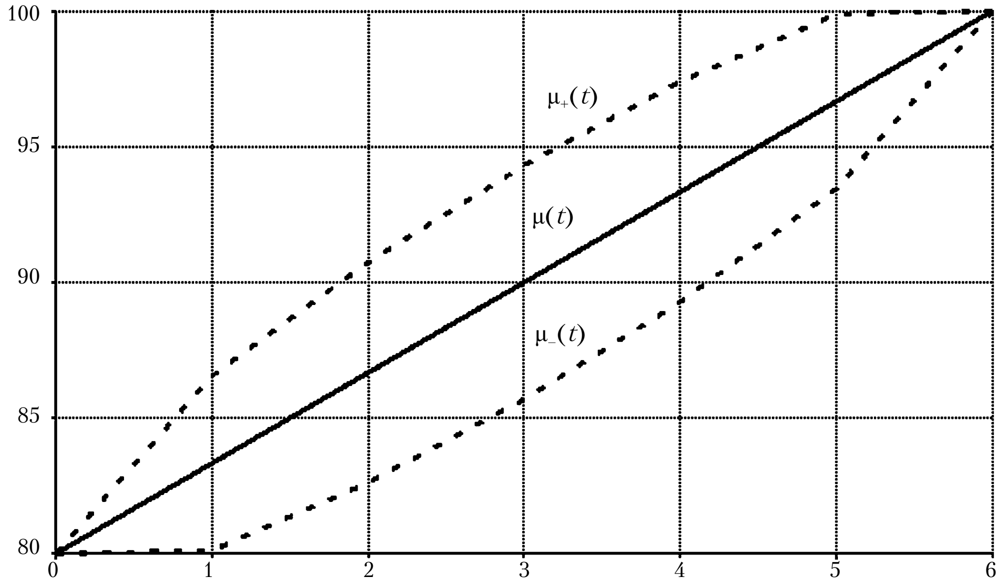

Рис. 1. Графики функций $\mu(t) = \mu_{\tilde{y}}(t)$, $\mu_{-}(t) = \mu^{-}_{\tilde{y}}(t)$, $\mu_{+}(t) = \mu^{+}_{\tilde{y}}(t)$.

$d\ (t)$, $t = 1,\ldots,6$. Пусть, далее, эксперт дает интервальную информацию $II = \{81 \le y(1) \le 82;$ $II = \{81 \le y(1) \le 82; 83 \le y(2) \le 84; 88 \le y(3) \le 90; 93 \le y(4) \le 99; 98 \le y(5) \le 99\}$ для цены $y\ (t)$, $t = 1,\ldots,6$, рассматриваемой облигации.

Таким образом, исследователь обладает нечисловой (ординальной) и неточной (интервальной) экспертной информацией $I = ID \cup II$, определяемой объединенной системой неравенств, входящих в системы неравенств D и I. Теперь сформируем множество $Y(I) = Y(m,n;I) = Y(6,20;I)$ всех допустимых (с точки зрения ННН-информации I) траекторий $y(t;\theta)$, $\theta = 1,\ldots,N(m,n;I) = N(6,20;I) < N(6,20)$. Для генерации всех возможных траекторий с последующей селекцией допустимых (т. е. удовлетворяющих неравенствам, входящим в объединенную систему неравенств I) используем систему поддержки принятия решений (СППР) APIS[14].

СППР APIS по ННН-информации I сформировала множество $Y\ (6,20;I)$ допустимых траекторий, состоящее всего из пяти траекторий $y(t;\theta)$, $\theta = 1,\ldots,N(6;20;I) = 5$, перечисленных в табл. 2. «Пучок» этих пяти траекторий изображен на рис. 2.

Таблица 2

Допустимые (с точки зрения ННН-информации I) траектории $y(t;\theta)$ стохастического процесса $\tilde{y}\,(t)$

θ	$y(0;\theta)$	$y(1;\theta)$	$y(2;\theta)$	$y(3;\theta)$	$y(4;\theta)$	$y(5;\theta)$	$y(6;\theta)$
1	80	81	83	89	95	98	100
2	80	81	83	89	95	99	100
3	80	81	84	89	97	99	100
4	80	81	84	90	94	98	100
5	80	81	84	90	96	99	100

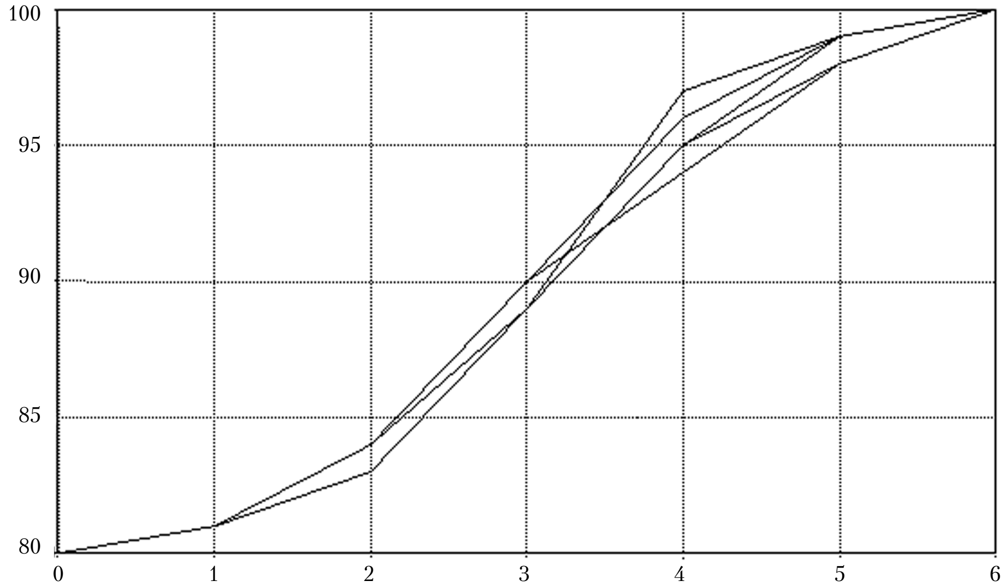

Рис. 2. «Пучок» допустимых траекторий $y(t;\theta)$, $\theta = 1,\ldots,5$, цены облигации.

Данные табл. 2 позволяют сосчитать математическое ожидание $\mu_{\tilde{y}}\ (t;I)$ стохастического процесса $\tilde{y}(t;I)$ с равновероятными траекториями из множества $Y(6,20;I)$ по формуле $\mu_{\tilde{y}}(t;I) = [y(t;1)+\ldots+y(t;5)]/5$, а дисперсию $\sigma^2_{\tilde{y}}(t;I)$ этого процесса — по формуле $\sigma^2_{\tilde{y}}(t;I) = \{[y(t;1)-\mu_{\tilde{y}}(t;1)]^2+\ldots+[y(t;5)-\mu_{\tilde{y}}(t;5)]^2\}/5$.

Теперь помимо тренда $\mu(t;I) = \mu_{\tilde{y}}(t;I)$ стохастического процесса $\tilde{y}\ (t;I)$ можно ввести ожидаемые нижнюю $\mu_-(t;I) = \mu(t;I) - \sigma(t;I)$ и верхнюю $\mu_+(t;I) = \mu(t;I) + \sigma(t;I)$ границы для допустимых траекторий этого процесса. Значения функций $\mu(t;I)$, $\sigma(t;I) = \sqrt{D\,\tilde{y}(t;I)}$, $\mu_-(t;I)$, $\mu_+\ (t;I)$ приведены в табл. 3, а соответствующие графики изображены на рис. 3.

Таблица 3

Значения функций $\mu(t;I)$, $\sigma(t;I)$, $\mu_-(t;I)$, $\mu_+(t;I)$

t	$\mu(t;I)$	$\sigma(t;I)$	$\mu_-(t;I)$	$\mu_+(t;I)$
0	80	0	80	80
1	81	0	81	81
2	83.6	0.49	83.11	84.09
3	89.4	0.49	88.91	89.89
4	95.4	1.02	94.38	96.42
5	98.6	0.49	98.11	99.09
6	100.0	0	100.00	100.00

До сих пор предполагалось, что все допустимые траектории $y(t;\theta) = y(t;6,20;\theta)$, $\theta = 1,\ldots,N(6;20;I) = 5$ равновероятны. Однако эксперт зачастую обладает дополнительной ННН-информацией P о вероятностях $p(\theta)$ появления допустимых траекторий $y(t;\theta)$, $\theta = 1,\ldots,N(m,n;I)$, $p(\theta) \geq 0$, $p(1)+\ldots+p(N(m,n;I)) = 1$. Пусть, например, эксперт задает такую ННН-информацию о вероятностях системой неравенств $IP = \{p(3) > p(5) > p(2) > p(1) > p(4)\}$. По этой информации СППР APIS строит оценки

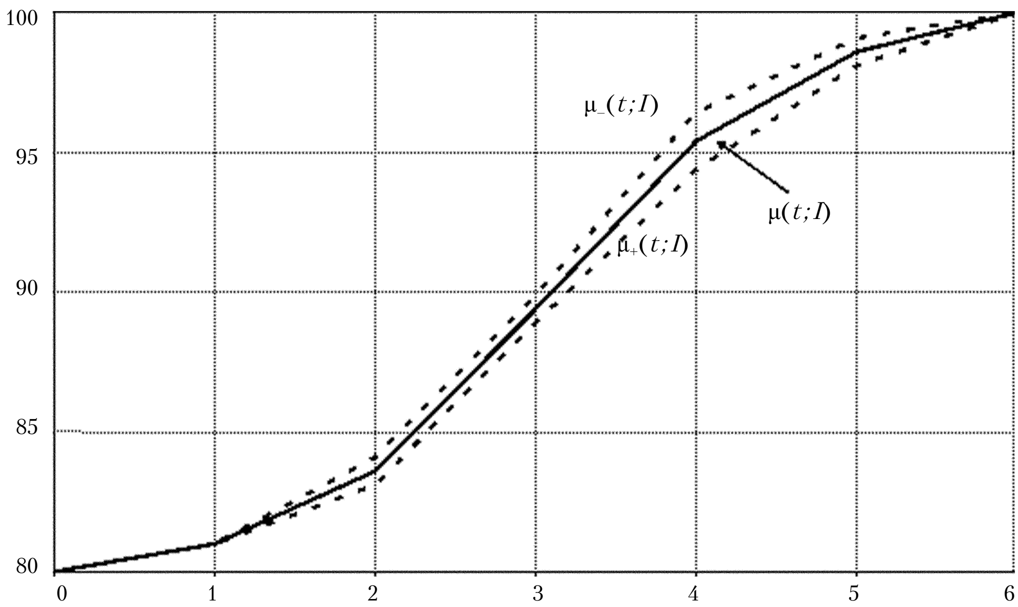

Рис. 3. Графики функций $\mu(t;I)$, $\mu_{-}(t;I)$, $\mu_{+}(t;I)$.

$\overline{p}(1;IP)=0.09$, $\overline{p}(2;IP)=0.16$, $\overline{p}(3;IP)=0.46$, $\overline{p}(4;IP)=0.03$, $\overline{p}(5;IP)=0.26$ вероятностей $p(1),..., p(5)$.

Полученные оценки вероятностей позволяют сосчитать математическое ожидание $\mu(t;I,IP)=\mu_{\tilde{y}}(t;I,IP)=E\,\tilde{y}(t;I,IP)$ стохастического процесса $\tilde{y}(t;I,IP)=\tilde{y}(t;6,20;I,IP)$ с равновероятными траекториями из множества $Y(6,20;I)$ по следующей формуле: $\mu(t;I,IP)=y(t;1)\,\overline{p}(1;IP)+...y(t;5)\,\overline{p}(5;IP)$. Дисперсия $\sigma^{2}(t;I,IP)=D\,\tilde{y}(t;I,IP)$ стохастического процесса $\tilde{y}(t;I,IP)=\tilde{y}(t;m,n;I,IP)$ вычисляется по формуле

$$\sigma^{2}(t;I,IP)=[y(t;1)-\mu(t;I,IP)]^{2}\,\overline{p}(1;IP)+...+[y(t;5)-\mu(t;I,IP)]^{2}\,\overline{p}(5;IP).$$

Теперь помимо тренда $\mu(t;I,IP)$ стохастического процесса $\tilde{y}(t;6,20;I.IP)$ можно ввести ожидаемые нижнюю $\mu_{-}(t;I,IP)=\mu(t;I,IP)-\sigma(t;I,IP)$ и верхнюю $\mu_{+}(t;I,IP)=$ $=\mu(t;I,IP)+\sigma(t;I,IP)$ границы для допустимых траекторий этого процесса. Значения функций $\mu(t;I,IP)$, $\sigma(t;I,IP)=\sqrt{D\,\tilde{y}(t;I,IP)}$, $\mu_{-}(t;I,IP)$, $\mu_{+}(t;I,IP)$ приведены в табл. 4.

Таблица 4

Значения функций
$\mu(t;I,P)$, $\sigma(t;I,P)$, $\mu_{-}(t;I,P)$, $\mu_{+}(t;I,P)$

t	$\mu(t)$	$\sigma(t)$	$\mu_{-}(t)$	$\mu_{+}(t)$
0	80	0	80	80
1	81	0	81	81
2	83.75	0.43	83.32	84.18
3	89.29	0.45	88.84	89.74
4	96.15	0.9	95.25	97.05
5	98.88	0.32	98.56	99.2
6	100	0	100	100

Сравнение таблиц 3 и 4 показывает, что учет вероятностей появления траекторий несколько уменьшает стандартное отклонение, что позволяет увеличить точность и достоверность экспертных оценок будущей динамики цен на облигацию.

Приведенный пример прогнозирования методом рандомизированных траекторий динамики цены облигации показывает, что привлечение нечисловой (порядковой, ординальной), неточной (интервальной) и неполной экспертной информации позволяет существенно повысить точность и надежность получаемых оценок. Наиболее перспективным представляется развитие метода рандомизированных траекторий в направлении привлечения вероятностных оценок экспертной информации и соединения (в рамках байесовской схемы оценивания) нечисловой экспертной информации с эмпирическими числовыми данными[15].

---

[1] См., напр., неоднократно переиздававшиеся учебники: *Вентцель Е. С.* Теория вероятностей: 9-е изд. М., 2003; *Гнеденко Б. В.* Курс теории вероятностей: 8-е изд. М., 2005.

[2] См., напр., довольно резкую критику использования абстрактных теоретико-вероятностных схем для объяснения эмпирически наблюдаемой стабилизации частот событий в работах: *Алимов Ю. И.* 1) Еще раз о реализме и фантастике в приложениях теории вероятностей // Автоматика. 1979. № 4. С. 83–90; 2) Альтернатива методу математической статистики. М., 1980; *Алимов Ю. И., Кравцов Ю. А.* Является ли вероятность «нормальной» физической величиной? // Успехи физических наук. 1992. Т. 162. № 7. С. 150–182.

[3] См., напр.: *Айвазян С. А., Енюков И. С., Мешалкин Л. Д.* Прикладная статистика. Основы моделирования и первичная обработка данных. М., 1983; *Вишняков И. В.* Экономико-математические модели оценки деятельности коммерческих банков. СПб., 1999; *Вишняков И. В., Хованов Н. В.* Система нормативов надежности коммерческих банков. СПб., 1998; *Хованов Н. В.* Феноменологическая теория стабильных метаденег // Финансы и бизнес. 2005. № 4. С. 18–21; *Brunk H., Gref L.* A geometrical approach to probability // Mathematics Magazine. 1964. Vol. 37. N 5. P. 287–296.

[4] См., напр.: *Селезнева Т. В., Тутубалин В. Н., Угер Е. Г.* Исследование прикладных возможностей некоторых моделей стохастической финансовой математики // Обозрение прикладной и промышленной математики. 2000. Т. 7. Вып. 2. С. 210–238; *Тутубалин В. Н.* 1) Статистическая обработка рядов наблюдений. М., 1973; 2) Границы применимости: вероятностно-статистические методы и их возможности. М., 1977.

[5] См., напр.: *Айзинова И. М.* Система показателей краткосрочных процессов в народном хозяйстве // Методологические проблемы анализа и прогноза краткосрочных процессов. М., 1979. С. 9–27; *Кипнис В. М.* Проблема прогнозирования временных рядов в условиях малых выборок // Методологические проблемы анализа и прогноза краткосрочных процессов. М., 1979. С. 107–134.

[6] Подробный разбор различных моделей неопределенности приведен в работе: *Хованов Н. В.* Три типа математических моделей неопределенности // Измерительная техника. 2005. № 9. С. 39–44.

[7] *Knight F.* Risk, Uncertainty, and Profit. Boston (MA, USA): Houghton Mifflin Co., 1921.

[8] Ср. с методом формирования множества допустимых траекторий, предложенным в работе: *Головченко В. Б.* Прогнозирование дискретных в пространстве состояний и времени процессов. Иркутск, 1988.

[9] *Bayes Th.* An essay towards solving a problem in the doctrine of chances // Biometrika. 1958. Vol. 5. Part 3–4. P. 296–315 (Reproduced from Philosophical Transactions of London Royal Society. 1763. Vol. 53).

[10] Подробнее о рандомизации теоретико-множественной неопределенности см. в работах: *Маркова Е. В., Маслак А. А.* Рандомизация и статистический вывод. М., 1986; *Хованов Н. В.* 1) Анализ и синтез показателей при информационном дефиците. СПб., 1996; 2) Математические модели риска и неопределенности. СПб., 1998.

[11] Об аргументах в пользу выбора именно равномерного распределения в качестве распределения, моделирующего отсутствие информации, см., работы: *Зельнер А.* Байесовские методы в эконометрии. М., 1980; *Barmish B., Lagoa C.* The uniform distribution: a rigorous justification for its use in robustness analysis // Mathematical Control, Signals, Systems. 1997. Vol. 10. P. 203–222; *Evans R.* The principle of minimal information // IEEE Transactions on Reliability. 1969. Vol. 18. P. 87–89; *Jaynes E.* Foundations of Probability Theory and Statistical Mechanics. New York, Springer, 1967; *Janes E.* Where do we stand on maximum entropy? // The Maximum Entropy Formalism / Ed by R. Levin. Cambridge: Cambridge University Press, 1979. P. 15–118; *Kan Yu., Kibzun A.* Sensitivity analysis of worst-case distribution for probability optimization problems // Probabilistic Constrained Optimizations / Ed by S. Uryasev. New York, 2000. P. 31–46; *Shimony A.* The status of the principle of maximum entropy // Synthese. 1985.

Vol. 63. P. 35–53; *Villegas C.* On the representation of ignorance // Journal of American Statistical Association. 1977. Vol. 72. N 359. P. 651–654.

[12] Все используемые далее результаты, связанные с изучением стохастических процессов, получены и подробно исследованы в работах: *Буре В. М., Колесникова О. Н., Корников В. В.* Простой статистический метод выявления монотонной зависимости среди наблюдаемых траекторий стохастического процесса. Л., 1983; *Хованов Н. В.* Стохастические процессы и поля с равновероятными монотонными дискретными реализациями // Управление, надежность и навигация. Вып. 5. Саранск, 1979. С. 136–139; *Хованов Н. В., Колесникова О. Н.* Прямой байесовский метод оценки распределений и параметров. Л., 1981; *Хованов Н. В., Рожков Н. Н.* Статистическая оценка показателя надежности с помощью нестационарных марковских случайных процессов // Управление, надежность и навигация. Вып. 6. Саранск, 1979. С. 127–131; *Чудовская Л. А.* Стохастические процессы с монотонными линейно-ограниченными дискретными реализациями // Вестн. Ленингр. гос. ун-та. Сер. 1: Математика, механика, астрономия (Депонировано ВИНИТИ 27.01.88 № 1330-В88. М., 1988).

[13] Простейший алгоритм генерации всех монотонных траекторий из множества $Y(m,n) \supseteq Y(m,n;I)$ приведен в работе: *Хованов Н. В.* Стохастические модели теории квалиметрических шкал. Л., 1986.

[14] Генерация траекторий из множества $Y(m,n;I)$ производилась при помощи одной из бета-версий системы поддержки принятия решений (СППР) APIS (Aggregated Preference Indices System), созданной фирмой Polyidea Ltd. (London, G.B.) в виде модификации сертифицированной СППР АСПИД-3W (см.: *Хованов К. Н., Хованов Н. В.* Система поддержки принятия решений АСПИД-3W (Анализ и синтез показателей при информационном дефиците). Свидетельство об официальной регистрации программы для ЭВМ № 960087 от 22.03.1996. Российское агентство по правовой охране программ для ЭВМ, баз данных и топологии интегральных микросхем (РосАПО). М., 1996).

[15] См., например, следующие работы по экспертной и/или байесовской оценке вероятностей альтернатив динамики финансово-экономических показателей: *Колесников Г. И., Федотов Ю. В., Хованов Н. В.* Оценка вероятностей альтернатив развития фондового рынка в условиях дефицита числовой информации // Вестн. С.-Петерб. ун-та. 2005. Серия 10: Прикладная математика, информатика, процессы управления. Вып. 2. С. 151–160; *Макаров А. В., Федотов Ю. В., Хованов Н. В.* Байесовская модель оценки вероятностей альтернативных состояний финансово-экономической среды реализации инвестиционных проектов // Материалы международной научной конференции «Экономическая наука: проблемы теории и методологии». Санкт-Петербург, 16–18 мая 2002 г. Секции 5–10. СПб., 2002. С. 141–142; *Хованов Н. В., Федотов Ю. В.* Рациональная оценка вероятностей альтернатив состояния среды осуществления проектов — основа эффективного стратегического менеджмента // Материалы конференции «Концепции и инструменты эффективного менеджмента». Санкт-Петербург, 28 октября 2005 г. СПб., 2005. С. 31–32; *Hovanov N. V., Yudaeva M. S., Kotov N. V.* Event-Tree with randomized transition probabilities as a new tool for alternatives probabilities estimation under uncertainty // Proceedings of the Sixth International Scientific School “Modeling and Analysis of Safety and Risk in Complex Systems”. St. Petersburg, 2006. July 4–8. SPb., RAS, 2006. P. 118–125.

Статья поступила в редакцию 24 декабря 2008 г.