

МОДЕЛИРОВАНИЕ ДВИЖЕНИЯ ПОЕЗДА С УЧЕТОМ СТАТИСТИЧЕСКИХ СВЯЗЕЙ ЭЛЕМЕНТОВ ТРАЕКТОРИИ

Получено распределение случайного отклонения от графика при прибытии поезда на конечную станцию. При этом используется однопоездная стохастическая модель движения поезда, введенная авторами в статье 2020 г. Найденное распределение имеет интегральную форму, зависит от совместного распределения отклонения при отправлении с предпоследней станции и времени хода по последнему перегону.

Ключевые слова: стохастическая модель, отклонение от графика, распределение опоздания, критерий согласия

EDN: FKQESL



Б.И. Давыдов



В.И. Чеботарев

Нарушения графика движения возникают вследствие воздействия случайных факторов, в первую очередь, технических и технологических отказов. Оперативное диспетчерское управление корректирует траектории движения поездов с целью минимизации отклонений. Любое регулировочное решение основывается на прогностической модели, которую выстраивает диспетчер. В условиях интенсивного движения эти решения могут быть далекими от оптимальности, если модель недостаточно обоснована и имеется недостаток исходных данных.

Процесс текущего перепланирования обычно основывается на фиксированных координатных данных, состоянии инфраструктуры и на предположении об их неизменности в предстоящем периоде. Задача прогнозирования результатов диспетчерских действий при этом носит детерминированный характер. Отсутствие точного знания о будущих препятствиях движению поездов часто приводит к ошибкам в определении отклонений от графика и, соответственно, к малоэффективным регулировкам. Как результат, более 2% пассажирских и 43% грузовых поездов прибывают по назначению с нарушениями расписания (согласно

Давыдов Борис Исаевич, кандидат технических наук, доцент кафедры «Автоматика, телемеханика и связь» Дальневосточного государственного университета путей сообщения. Область научных интересов: проблемы управления в транспортной отрасли; моделирование транспортных объектов и процессов. Автор 109 научных работ, в том числе одной монографии. Имеет шесть патентов на изобретения.

Чеботарев Владимир Иванович, доктор физико-математических наук, доцент, главный научный сотрудник Вычислительного центра Дальневосточного отделения Российской академии наук – обособленное подразделение Хабаровского федерального исследовательского центра ДВО РАН. Область научных интересов: теория вероятностей и ее применения. Автор 76 научных работ, в том числе одной монографии и трех учебников.

Каблукова Ксения Сергеевна, младший научный сотрудник Вычислительного центра Дальневосточного отделения Российской академии наук – обособленное подразделение Хабаровского федерального исследовательского центра ДВО РАН. Область научных интересов: математическое моделирование, вычислительная математика. Автор 18 научных работ, в том числе одной монографии.

Пазойский Юрий Ошарович, доктор технических наук, профессор, заведующий кафедрой «Железнодорожные станции и транспортные узлы» Института управления и цифровых технологий Российского университета транспорта (ИУЦТ РУТ (МИИТ)), Почётный работник транспорта Российской Федерации, Почётный железнодорожник. Область научных интересов: пассажирские перевозки на железнодорожном транспорте, совершенствование работы станционных процессов. Автор 176 научных работ, в том числе четырех монографий и трех учебников.

официальной статистике ОАО «РЖД») — несмотря на резервы времени, заложенные в график и усилия диспетчерского персонала.

Показано [1], что при интенсивном движении подход, учитывающий фиксированные интервалы выполнения операций (равномерного хода, ускорений, остановок), дает ошибки прогнозных оценок моментов прибытия в горизонте трех часов, превышающие 5 мин. В реальной обстановке точность прогноза еще ниже, так как текущая информация о движении отягощена случайными погрешностями и, кроме того, нередко запаздывает. Обширный круг литературных источников (см., например, [2–4]) подтверждает, что использование вероятностного подхода при формировании траекторий движения поездов позволяет значительно (на порядок и более) ускорить процесс выработки прогноза и повысить его эффективность.

Предлагаемая в работе стохастическая модель представляет траекторию движения поезда в виде цепи последовательных операций, каждая из которых характеризуется случайным моментом ее начала и продолжительностью. Распределения указанных величин определяются на основе статистических данных, полученных в предшествующие периоды времени. Рассчитывается распределение моментов прибытия на заданную станцию, которое используется для фиксации факта опоздания поезда и оценки эффективности мер по предотвращению межпоездных конфликтов.

При проведении расчета вероятностного распределения моментов прибытия во многих работах (например, [2; 5–7]) используется предположение о статистической независимости интервалов реализации элементарных операций. Это допущение позволяет при нахождении распределения отклонения моментов прибытия от графика ограничиться использованием техники сверток распределений [8, стр. 272].

Изучение статистических данных о реальном движении поездов показывает, что указанное предположение подтверждается далеко не всегда. Чтобы оценить масштаб распространения случаев статистической зависимости, в настоящей работе проводится исследование наличия корреляционной связи между операциями, реализуемыми в процессе движения. Предлагается метод расчета распределений моментов прибытия, который использует статистические данные о продолжительности операционных интервалов. Для проверки адекватности полученных распределений используется критерий Пирсона. Также обсуждается характер распределений времен хода и моментов прибытия, что полезно при качествен-

ной оценке дестабилизирующих влияний на процесс пропуска поездов. Результаты исследования могут быть использованы при решении задач по повышению уровня пунктуальности движения пассажирских поездов и эффективности пропуска потока по грузонапряженным направлениям железных дорог.

Обзор литературных источников

В [2] подчеркивается, что вид распределения моментов прибытия и его параметры зависят от отклонения от графика движения поезда в момент, когда вырабатывается прогноз и от текущей ситуации на предстоящих участках его маршрута. Вероятностный подход предполагает, что известен способ расчета распределения моментов прибытия на основе статистики, полученной путем анализа фактических отклонений от графика и прогноза будущих влияющих факторов. В основу методики расчета кладутся статистические связи выходного распределения с указанными факторами, которые должны быть известны до начала расчетной процедуры. Закономерности, которые используются при построении расчетной модели движения поездов, описываются в терминах теории марковских процессов [3] или байесовских сетей [4].

Отметим также работу [9], в которой рассматривается одна прогнозная задача, относящаяся к метрополитену. Более точно, в статье предлагается метод прогнозирования отклонения длительности остановки электропоезда метрополитена от плановой. Метод опирается на модель, которая, в свою очередь, основана на полиномах, являющихся наилучшими в смысле принципа наименьших квадратов приближениями к конечной последовательности фактических отклонений для поездов, которые уже проследовали через данную станцию.

Авторы также занимаются оценкой ошибки, возникающей при такой аппроксимации. В качестве погрешности они предлагают относительное среднеквадратическое отклонение разности прогнозного и фактического отклонений. Показывают, что погрешность аппроксимации выражается в терминах коэффициентов корреляции между указанными выше случайными отклонениями для предыдущих поездов. Однако для вычисления коэффициентов корреляции авторы не занимаются поиском распределений отклонений или совместных распределений пар отклонений. Вместо этого они берут два примера поведения коэффициентов корреляции и приводят таблицу соответствующих погрешностей при различных параметрах, от которых зависят примеры. Заметим, что в [9] не раскрывается, как указанные примеры связаны с реальностью.

Модель формирования распределения моментов прибытия поездов

Описание модели. Постановка задачи

Модель движения поезда представляется в виде последовательности операций движения и графических остановок на станциях, на которых производится высадка / посадка пассажиров или смена локомотива. Будем считать, что длительности операций имеют случайный характер. Предположим, что начальная станция отправления имеет номер 1, конечная станция – номер N . Мы будем пользоваться следующими обозначениями:

d_i – графический момент отправления со станции с номером i ;

a_i – графический момент прибытия на станцию i ;

δ_i – отклонение от графика при отправлении со станции i ;

ξ_i – отклонение от графика при прибытии на станцию i ;

$\rho_{i,i+1}$ – время хода от станции i до станции $i+1$;

τ_i – продолжительность стоянки на станции i .

Обращаем внимание читателя, что для обозначения случайных величин мы используем греческие буквы.

Рассматривается задача: найти распределение случайной величины ξ_N – величины отклонения от графика при прибытии на конечную станцию.

В основу настоящего исследования положен анализ статистики о фактическом движении контейнерных грузовых поездов в границах Восточного полигона РЖД, следующих по жесткому графику. Рассматривается процесс пропуска по пяти участкам, каждый протяженностью 250–400 км, ограниченными шестью техническими станциями (табл. 1). Соответственно здесь $1 \leq i \leq N=6$.

В рассматриваемом частном случае нас интересует распределение отклонения от графика при прибытии на станцию 6 (см. рис. 1), т.е. распределение случайной величины ξ_6 . Преодоление поездом участка между станциями 1 и 6 описывается следующей последовательностью случайных величин: $d_1 + \delta_1$ – момент отправления со станции 1; $\rho_{1,2}$ – время хода между станциями 1 и 2; τ_2 – время стоянки на станции 2; $\rho_{2,3}$ – время хода между станциями 2 и 3 и т.д.; $\rho_{5,6}$ – время хода между станциями 5 и 6; $a_6 + \xi_6$ – момент прибытия на станцию 6, т.е. последовательностью

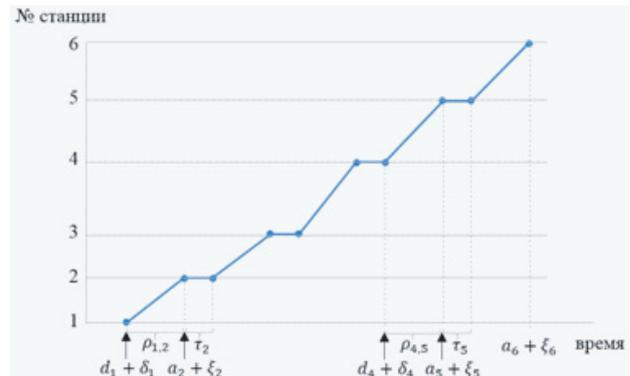


Рис. 1. Схема движения поезда между станциями 1 и 6

$$(d_1 + \delta_1, \{\rho_{i,i+1} + \tau_{i+1}\}_{i=1}^4, \rho_{5,6}, a_6 + \xi_6).$$

Схема движения поезда между станциями 1 и 6 изображена на рис. 1.

О распределении случайной величины ξ_6

Учитывая рис. 1, легко понять, что

$$a_{i+1} + \xi_{i+1} - (d_i + \delta_i) = \rho_{i,i+1}$$

при $i = \overline{1,5}$, откуда

$$\xi_{i+1} = \delta_i + \rho_{i,i+1} + d_i - a_{i+1}, i = \overline{1,5}.$$

Таким образом,

$$\xi_6 = \delta_5 + \rho_{5,6} + d_5 - a_6. \quad (1)$$

Замечание 1. Формулу (1) можно получить иначе. Действительно, в силу нашей модели, изображенной на рис. 1,

$$a_6 + \xi_6 = d_1 + \delta_1 + \sum_{i=1}^4 (\rho_{i,i+1} + \tau_{i+1}) + \rho_{5,6}. \quad (1a)$$

Отсюда, учитывая очевидное равенство $\rho_{i,i+1} + \tau_{i+1} = d_{i+1} + \delta_{i+1} - (d_i + \delta_i)$, приходим к (1).

Формулы (1) и (1a) имеют место, когда номер N конечной станции равен 6. Аналогичные формулы справедливы при любом $N \geq 2$: вместо (1) –

$$\xi_N = \delta_{N-1} + \rho_{N-1,N} + d_{N-1} - a_N; \quad (1b)$$

вместо (1a) –

$$a_N + \xi_N = d_1 + \delta_1 + \sum_{i=1}^{N-2} (\rho_{i,i+1} + \tau_{i+1}) + \rho_{N-1,N}. \quad (1ab)$$

Таблица 1

Участок	1–2	2–3	3–4	4–5	5–6
Протяженность, км	246	245	399	343	318

Из (1b) и (1ab) при $N \geq 3$ вытекает равенство для условных вероятностей

$$P(\xi_N < t | \delta_i, \rho_{i,i+1} + \tau_{i+1}, 1 \leq i \leq N-2, \delta_{N-1}, \rho_{N-1,N}) = \\ = P(\xi_N < t | \delta_{N-1}, \rho_{N-1,N}),$$

напоминающее марковское свойство.

Функцию распределения случайной величины ξ_i обозначим $U_i(t) := P(\xi_i < t)$, $i=2,6$, а функцию плотности (мы считаем, что она существует) — $u_i(t)$. Функции распределения случайных величин δ_i и $\rho_{i,i+1}$ обозначим $F_i(x)$ и $L_{i,i+1}(y)$ соответственно. Их функции плотности обозначим теми же буквами, но строчными: $f_i(x)$ и $l_{i,i+1}(y)$. Введем также обозначение $g_i(x,y)$ для плотности совместного распределения этих случайных величин. Учитывая (1b), получаем

$$U_{i+1}(t) := P(\delta_i + \rho_{i,i+1} < t + a_{i+1} - d_i) = \\ = \iint_{x+y < t+a_{i+1}-d_i} g_i(x,y) dy dx = \\ = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{t+a_{i+1}-d_i-x} g_i(x,y) dy \right) dx \quad (2)$$

(см. также [10, формула (4)].) Из (2) вытекает равенство для плотности $u_{i+1}(t)$:

$$u_{i+1}(t) = \int_{-\infty}^{\infty} g_i(x, t + a_{i+1} - d_i - x) dx. \quad (3)$$

Предположим, что δ_i и $\rho_{i,i+1}$ независимы. В этом случае $g_i(x,y) = l_{i,i+1}(x) f_i(y)$. Тогда в силу (2) имеем

$$U_{i+1}(t) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{t+a_{i+1}-d_i-x} f_i(y) dy \right) l_{i,i+1}(x) dx = \\ = \int_{-\infty}^{\infty} F_i(t + a_{i+1} - d_i - x) l_{i,i+1}(x) dx \equiv \\ \equiv (F_i \cdot L_{i,i+1})(t + a_{i+1} - d_i),$$

а в силу (3) —

$$u_{i+1}(t) = \int_{-\infty}^{\infty} f_i(t + a_{i+1} - d_i - x) l_{i,i+1}(x) dx \equiv \\ \equiv (f_i * l_{i,i+1})(t + a_{i+1} - d_i). \quad (4)$$

Здесь $(F \cdot G)(x) = \int_{-\infty}^{\infty} F(x-y) dG(y)$ — свёртка функций распределения F и G ;

$(f \cdot g)(x) = \int_{-\infty}^{\infty} f(x-y) g(y) dy$ — свёртка плотностей f и g .

Далее формулу (4) будем использовать при $i=5$. В одном из следующих разделов проводится проверка

соответствия полученного модельного распределения реальным статистическим данным.

Статистические данные о движении поездов

Проведен статистический анализ данных о движении контейнерных поездов на участке Восточного полигона РЖД, содержащем шесть технических станций (см. табл. 1). Исследуются вероятностные характеристики отклонений от графика при отправлении, прибытии, времен хода по участкам и продолжительности остановок. Последовательность случайных величин в порядке их возникновения состоит из следующих элементов: $\{\delta_i, \rho_{i,i+1}, \xi_{i+1}, \tau_{i+1}\}_{i=1}^4$, δ_5 , $\rho_{5,6}$, ξ_6 . Нас интересует корреляционная связь между перечисленными 19 случайными величинами.

Выборки из распределения каждой случайной величины последовательности приведены в табл. 2. (единица измерения здесь и далее — час).

Каждая строка табл. 2, начиная со второй, состоит из последовательности наблюдаемых значений случайных величин, которые указаны в первой строке (они описывают движение конкретного поезда). Столбцы таблицы представляют собой соответствующие выборки объемом $n=12$. Малость объема выборок связана со следующей трудностью, которая возникает при сборе статистических данных о временах выполнения цепи операций. Совокупность данных по всей цепи является конечномерным случайным вектором. Отсутствие хотя бы одной компоненты ведет к потере информации и, как следствие, к ошибкам в конкретных реализациях вектора в расчетах. Очевидно, чем большую размерность имеет вектор, тем более сложно получить полноценную выборку. Так, в проведенном исследовании анализируется 19-мерный случайный вектор, для которого удалось получить выборку с объемом, равным всего лишь 12.

Статистические связи в границах элементарных участков

Представим результаты вычислений выборочных парных коэффициентов корреляции между случайными величинами, характеризующими элементы траектории движения.

Напомним, что выборочным коэффициентом корреляции случайных величин X и Y называется величина

$$r^*(X,Y) = \frac{\text{cov}^*(X,Y)}{\sqrt{\mathbf{D}^*(X) \cdot \mathbf{D}^*(Y)}},$$

где $\mathbf{D}^*(Z)$ означает выборочную дисперсию случайной величины Z , а $\text{cov}^*(X,Y)$ — выборочную ковариацию случайных величин X и Y . Отсюда следует, что $r^*(X,Y)$ можно вычислять по формуле:

Таблица 2

δ_1	$\rho_{1,2}$	ξ_2	τ_2	δ_2	$\rho_{2,3}$	ξ_3	τ_3	δ_3	$\rho_{3,4}$	ξ_4	τ_4	δ_4	$\rho_{4,5}$	ξ_5	τ_5	δ_5	$\rho_{5,6}$	ξ_6
-0.1	7.1	-0.7	0.6	-0.1	4.6	0.5	0.1	-0.8	11.6	4.2	0.4	2.6	7.5	4.4	1.6	4.9	6.2	5.4
-0.1	7.1	-0.6	0.9	0.3	4.1	0.4	0.4	-0.5	10.5	3.5	1.4	2.8	10.6	7.8	0.7	7.4	6.7	8.4
0.0	12.0	4.4	0.2	4.5	4.1	4.7	0.8	4.2	8.2	5.7	0.3	4.1	7.6	6.0	0.5	5.5	10.0	9.7
-0.3	8.7	0.7	1.4	2.0	4.8	2.9	0.7	2.3	11.1	6.8	0.9	5.7	9.7	9.8	1.2	9.8	6.6	10.6
0.0	11.0	3.4	0.1	3.5	6.1	5.6	0.6	4.9	15.7	13.9	1.3	13.3	7.2	14.8	0.4	14.1	6.8	15.2
-0.1	12.6	4.9	0.6	5.5	6.4	7.9	0.6	7.1	7.9	8.4	1.4	7.8	7.4	9.5	0.6	9.0	6.6	9.8
0.2	7.6	0.1	0.1	0.2	4.9	1.2	0.5	0.4	9.7	3.6	0.6	2.1	6.7	3.1	0.8	2.8	6.7	3.7
-0.1	8.6	0.9	0.1	1.0	4.6	1.6	0.2	0.5	17.1	10.9	2.1	11.0	7.9	13.3	0.4	12.6	8.0	14.8
0.1	9.3	1.8	0.2	1.9	5.1	3.0	0.8	2.5	8.5	4.3	2.2	4.6	7.7	6.6	0.4	5.9	6.1	6.2
-0.1	12.0	4.2	0.2	4.4	4.0	4.5	0.9	4.0	9.2	6.7	0.9	5.4	7.8	7.6	0.3	6.8	6.3	7.3
-0.2	12.3	4.4	0.1	4.5	4.8	5.3	0.5	4.5	9.2	7.1	0.4	5.5	7.7	7.5	0.6	7.0	8.4	9.6
0.1	10.7	3.1	0.1	3.2	5.6	4.8	2.8	6.3	10.8	10.4	1.6	10.1	6.4	10.9	0.9	10.6	6.1	10.9

$$r^*(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

где x_i и y_i – выборочные значения; \bar{x} и \bar{y} – выборочные средние (см., например, [11, с. 83]). Авторы вычислили выборочные коэффициенты корреляции для всех возможных пар, состоящих из величин табл. 2. Из-за ограничений на объем публикации приведены результаты только для участков «1–2» (участок между станциями 1 и 2) и «2–3» (участок между станциями 2 и 3): табл. 3, 4.

Опираясь на известный факт теории вероятностей, что две случайные величины линейно зависимы тогда и только тогда, когда модуль их коэффициента корреляции равен 1, мы будем говорить, что случайные величины почти линейно зависимы, если модуль их выборочного коэффициента корреляции близок к 1.

В силу табл. 3 и 4 имеем: $r^*(\xi_2, \delta_2) = 0.98$, $r^*(\xi_3, \delta_3) = 0.97$. Закономерность сохраняется и для последующих участков «3–4», «4–5»: $r^*(\xi_4, \delta_4) = 0.98$, $r^*(\xi_5, \delta_5) = 0.99$. Поэтому можно считать, что между отклонением при прибытии ξ_i на станцию i и отклонением при отправлении с нее δ_i существует почти линейная зависимость.

Почти линейная связь в начале пути между временем хода $\rho_{i-1,i}$ и отклонением при прибытии ξ_i постепенно ослабевает к концу маршрута. Это можно наблюдать на участках «1–2», «2–3» (см. табл. 3, 4) и на последующих участках. Значения коэффициента корреляции снижаются по цепи: $r^*(\rho_{1,2}, \xi_2) = 0.99$, $r^*(\rho_{2,3}, \xi_3) = 0.61$, $r^*(\rho_{3,4}, \xi_4) = 0.64$, $r^*(\rho_{4,5}, \xi_5) = 0.02$, $r^*(\rho_{5,6}, \xi_6) = 0.29$.

Аналогичный вывод справедлив для отклонения δ_i при отправлении и предшествующего времени хода $\rho_{i-1,i}$: $r^*(\delta_2, \rho_{1,2}) = 0.99$, $r^*(\delta_3, \rho_{2,3}) = 0.61$, $r^*(\delta_4, \rho_{3,4}) = 0.65$, $r^*(\delta_5, \rho_{4,5}) = 0.044$.

Продолжительность стоянки τ_i почти не связана с отклонением прибытия ξ_i и с временем хода $\rho_{i-1,i}$ по предшествующему участку. Связь отклонения δ_i при отправлении и времени стоянки τ_i нестабильна.

Таблица 3

Участок «1–2»					
	δ_1	$\rho_{1,2}$	ξ_2	τ_2	δ_2
δ_1	1				
$\rho_{1,2}$	-0.13	1			
ξ_2	-0.07	0.993	1		
τ_2	-0.63	-0.37	-0.43	1	
δ_2	-0.21	0.993	0.98	-0.24	1

Таблица 4

Участок «2–3»					
	δ_2	$\rho_{2,3}$	ξ_3	τ_3	δ_3
δ_2	1				
$\rho_{2,3}$	0.35	1			
ξ_3	0.95	0.61	1		
τ_3	0.29	0.25	0.34	1	
δ_3	0.92	0.61	0.97	0.56	1

А связь отклонения ξ_i при прибытии с предшествующим отклонением δ_{i-1} при отправлении изменяется от почти некоррелированности к почти линейной зависимости: $r^*(\xi_2, \delta_1) = -0.07$, $r^*(\xi_3, \delta_2) = 0.95$, $r^*(\xi_4, \delta_3) = 0.55$, $r^*(\xi_5, \delta_4) = 0.94$, $r^*(\xi_6, \delta_5) = 0.94$.

О распределении отклонения при прибытии ξ_6

Поскольку выборочный коэффициент корреляции случайных величин δ_5 и $\rho_{5,6}$ равен $r^*(\delta_5, \rho_{5,6}) = -0.055$, т.е. близок к нулю, то можно считать, что указанные случайные величины некоррелированы. Это, в свою очередь, может служить основанием для предположения их независимости. Последнее упрощает вывод функции плотности распределения случайной величины ξ_6 . Для проверки того, насколько модельное распределение отличается от реального, будут применены статистические критерии согласия.

В соответствии с терминологией, случайные величины δ_5 , $\rho_{5,6}$ и соответствующие им плотности $f_5(\cdot)$, $l_{5,6}(\cdot)$ являются входными для формулы (4) при $i=5$. Эти плотности будем находить как аппроксимации для гистограмм, построенных по имеющимся статистическим данным.

Описание входных распределений

Приведем выборку из распределения случайной величины δ_5 . Представим ее в форме вариационного ряда (табл. 5). В табл. 5 первая строка отражает номер варианты j . Вторая – саму варианту $\delta_{5(j)}$. Третья – частоту ω_j (количество идентичных вариантов в выборке). Выборку из распределения $\rho_{5,6}$ приводить не будем из-за ограничений на объем публикации. Объемы выборки для δ_5 и $\rho_{5,6}$: $n_1=35$, $n_2=43$.

Для моделирования распределений входных случайных величин будем использовать плотность гамма-распределения (см. [12, с. 121], [13]):

$$g_1(x) = I(x > b) \frac{\exp\{-(x-b)/\beta\} (x-b)^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)}, \quad (5)$$

где $\alpha > 0$, $\beta > 0$, $b \in \mathbb{R}$; $I(x > b)$ означает индикатор множества $\{x: x > b\}$, т.е. $I(x > b) = 1$ если $x > b$ и $I(x > b) = 0$ если

$x \leq b$. Оформи́м результаты, полученные для входных случайных величин, в форме табл. 6.

Используя критерий согласия Пирсона, проверим, согласуются ли выборочные данные с гипотетическими плотностями из табл. 6. Выпишем эти плотности в явном виде:

$$f_5(t) = I(t > -1.5) \frac{\exp\left\{-\frac{(t+1.5)}{1.59}\right\}(t+1.5)^{2.93}}{(1.59)^{3.93} \Gamma(3.93)} \quad (6)$$

— гипотеза 1 для δ_5 ,

$$l_{5,6}(t) = I(t > 5.42) \frac{\exp\left\{-\frac{(t-5.42)}{0.77}\right\}(t-5.42)^{1.12}}{(0.77)^{2.12} \Gamma(2.12)} \quad (7)$$

— гипотеза 2 для $\rho_{5,6}$.

Критерий Пирсона опирается на так называемую «статистику хи-квадрат»:

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i}, \quad (8)$$

где k — число интервалов в подправленном интервальном вариационном ряду (каждый интервал содержит не менее пяти выборочных значений); i — номер интервала; n_i — число выборочных значений, попавших в i -й интервал; p_i — гипотетическая вероятность попадания в этот интервал [8, с. 152, 153].

Выборка из распределения случайной величины δ_5 представлена в табл. 5. Приведем соответствующий

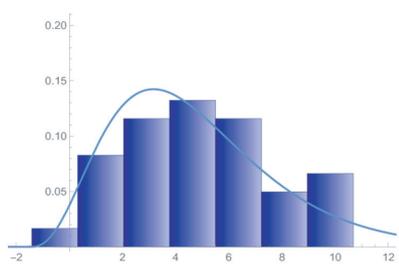
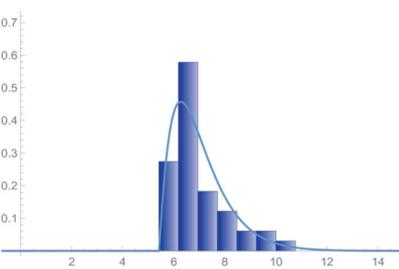
Таблица 5

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\delta_{5(j)}$	-0.55	0.95	1.13	1.32	1.82	1.97	2.38	2.52	2.77	3.25	3.77	4.20	4.40	4.45	4.87	5.17
ω_j	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	2

Окончание табл. 5

j	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
$\delta_{5(j)}$	5.42	5.45	5.92	6.10	6.17	6.27	6.77	7.00	7.38	7.93	8.80	8.97	9.33	9.38	9.82
ω_j	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1

Таблица 6

Случайная величина	Гистограмма и гипотетическая функция плотности	Гипотетическая функция плотности
Отклонение при отправлении δ_5		Плотность гамма-распределения из (5) с параметрами $\alpha=3.93$, $\beta=1.59$ и $b=-1.5$
Время хода $\rho_{5,6}$		Плотность гамма-распределения из (5) с параметрами $\alpha=2.12$, $\beta=0.77$ и $b=5.42$

ей подправленный интервальный вариационный ряд в форме табл. 7, где $(a_j, a_{j+1}]$ обозначает полуоткрытый интервал с номером j и концами a_j, a_{j+1} . Выборка из распределения случайной величины $\rho_{5,6}$ ранее не приводилась; соответствующий ей подправленный интервальный вариационный ряд представлен в форме табл. 8.

Согласно теореме Пирсона, если выдвинутая статистическая гипотеза верна, то при увеличении n распределение статистики (8) приближается к распределению хи-квадрат с r степенями свободы, где $r=k-1-k_H$, а k_H – число параметров, оцениваемых по выборке (k_H совпадает с так называемым числом связей между количествами $n_i, i=1, k$) [8, с. 152, 153].

Задается уровень значимости критерия α , близкий к нулю, и определяется критическое значение χ^2_α , удовлетворяющее равенству

$$P(\chi^2(r) > \chi^2_\alpha) = \alpha,$$

где $\chi^2(r)$ – случайная величина, имеющая распределение хи-квадрат с r степенями свободы. Если найденное значение статистики (8) удовлетворяет условию $\chi^{2*} < \chi^2_\alpha$, то гипотеза считается не противоречащей опытным данным по критерию Пирсона с уровнем значимости α . В табл. 9 представлены результаты применения критерия Пирсона с уровнем значимости $\alpha=0.05$.

Таблица 7

j	1	2	3	4	5
$(a_j, a_{j+1}]$	(-1.414, 2.042]	(2.042, 3.77]	(3.77, 5.498]	(5.498, 7.226]	(7.226, 10.68]
n_j	6	7	8	7	7

Таблица 8

j	1	2	3	4
$(a_j, a_{j+1}]$	(5.418, 6.182]	(6.182, 6.946]	(6.946, 7.71]	(7.71, 10.766]
n_j	9	19	6	9

Таблица 9

Случайная величина	Значение статистики χ^{2*} , вывод о согласованности
δ_5	Так как $\chi^{2*} \approx 1.669 < \chi^2_{0.05} \approx 5.991$, то по критерию Пирсона гипотеза 1 не противоречит опытным данным с уровнем значимости $\alpha=0.05$. Пояснение. Два параметра: $\alpha=3.93, \beta=1.59$ найдены по выборке (методом моментов). Поэтому $k_H=2$. Учитывая, что $k=5$, получаем $r=5-1-2=2$. По таблице распределения хи-квадрат [14] находим $\chi^2_{0.05} \approx 5.991$. Параметр $b=-1.5$ подобран эмпирически
$\rho_{5,6}$	Так как $\chi^{2*} \approx 3.135 < \chi^2_{0.05} \approx 3.841$, то по критерию Пирсона гипотеза 2 не противоречит опытным данным с уровнем значимости $\alpha=0.05$. Пояснение. Два параметра: $\alpha=2.12, \beta=0.77$ найдены по выборке. Поэтому $k_H=2$. Учитывая, что $k=4$, получаем $r=4-1-2=1$. По таблице распределения хи-квадрат [14] находим $\chi^2_{0.05} \approx 3.841$. Параметр $b=5.42$ подобран эмпирически

Таким образом, входные распределения, согласующиеся с данными реального движения, заданы. Теперь можно приступить к нахождению плотности $u_6(t)$.

О согласованности модельного распределения со статистическими данными

График функции $u_6(t)$ из (4), изображенный вместе с гистограммой (см. рис. 2), построен с помощью пакета Wolfram Mathematica. Выборка из распределения ξ_6 , по которой построена гистограмма, представлена табл. 10.

Проверим статистическую гипотезу: выборка, которой соответствует гистограмма, изображенная на рис. 2, взята из распределения с плотностью $u_6(t)$, определенной равенством (4), где $d_5=31.98$ час., $a_6=37.3$ час., а плотности $f_5(t)$ и $l_{5,6}(t)$ взяты из (6) и (7) соответственно. Проверку будем проводить, применяя критерий Пирсона с уровнем значимости $\alpha=0.05$. Подправленный интервальный вариационный ряд приведен в форме табл. 11.

По формуле $p_j = \int_{a_j}^{a_{j+1}} u_6(t) dt$ находим гипотетические вероятности: $p_1=0.202079$, $p_2=0.172208$, $p_3=0.173423$, $p_4=0.332924$. Используя (8), получаем $\chi^{2*} \approx 0.9$.

Так как плотность $u_6(t)$ полностью определена и не зависит от выборки из распределения случайной величины ξ_6 , то $k_H=0$. Поэтому $r=4-1-0=3$. По таблице

распределения хи-квадрат [14] находим $\chi^2_{0.05} \approx 7.815$. Таким образом, убеждаемся, что $\chi^{2*} < \chi^2_{0.05}$.

Отсюда следует, что выдвинутая статистическая гипотеза не противоречит опытным данным по критерию Пирсона с уровнем значимости $\alpha=0.05$.

Другими словами, модельная функция плотности, определенная равенством (4), где плотности $f_5(t)$ и $l_{5,6}(t)$ взяты из (6) и (7) соответственно, согласуется со статистическими данными из табл. 10.

Некоторые простые выводы из построенной модели

Учитывая равенство (1), рассмотрим вопрос о влиянии исходных распределений на итоговое в терминах математических ожиданий и дисперсий.

Из (1) и свойств математического ожидания следует, что

$$E\xi_6 = E\delta_5 + E\rho_{5,6} + d_5 - a_6.$$

Таким образом, сдвиг δ_5 на величину c_1 , а $\rho_{5,6}$ — на величину c_2 , приводит к сдвигу $E\xi_6$ на сумму $c_1 + c_2$.

Предположим, что случайные величины δ_5 и $\rho_{5,6}$ некоррелированы. Тогда по свойствам дисперсии

$$D\xi_6 = D\delta_5 + D\rho_{5,6}.$$

Следовательно, если в (1) случайные величины δ_5 и $\rho_{5,6}$ уменьшить в k раз, то $D\xi_6$ уменьшится в k^2 раз. Это значит, что разброс значений случайной величины ξ_6 (измеряемый посредством среднеквадратического отклонения $\sqrt{D\xi_6}$) уменьшится тоже в k раз.

Таблица 10

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\xi_6(j)$	0.23	1.88	2.38	2.52	2.77	2.88	3.52	3.72	3.75	3.93	4.72	5.38	5.55	5.93	5.97	6.05
w_j	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1	1

Окончание табл. 10

j	17	18	19	20	21	22	23	24	25	26	27
$\xi_6(j)$	6.23	6.43	6.72	6.88	7.30	8.02	8.37	8.50	9.63	9.72	9.80
w_j	1	1	1	1	1	1	2	1	1	1	1

Таблица 11

j	1	2	3	4
$(a_j, a_{j+1}]$	(-0.45, 3.651]	(3.651, 5.018]	(5.018, 6.385]	(6.385, 10.486]
n_j	7	5	7	11

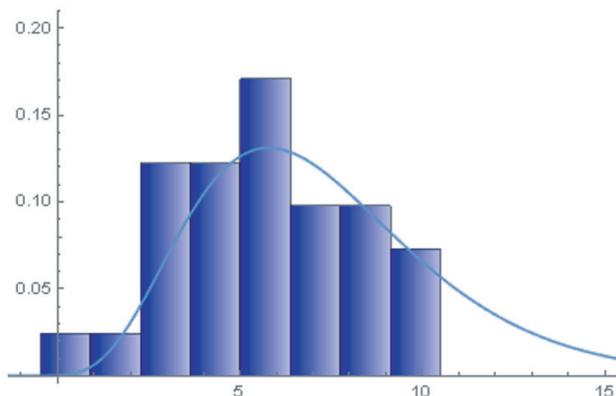


Рис. 2. Гистограмма, построенная по выборке из распределения случайной величины ξ_6 , и функция $u_6(t)$ из (4)

Заключение

Использование стохастического моделирования позволяет повысить качество прогнозирования развития поездной ситуации и оперативность принятия решений в условиях интенсивного движения. В результате улучшается пунктуальность движения пассажирских, контейнерных поездов и степень использования пропускной способности железнодорожной линии.

Качество прогноза улучшается, если при расчете вероятностного распределения моментов прибытия

учитывать статистические связи между величинами, характеризующими элементы маршрута, — временами хода и остановок. Проведен корреляционный анализ реальных данных по движению поездов на магистральной железнодорожной линии, который позволяет оценить закономерности в области межэлементной связи. Определено, что интенсивная связь присутствует между отклонениями при отправлении на участок и прибытии на следующую техническую станцию. Что касается времен хода, то их статистическая связь с моментами прибытия по мере движения по маршруту ослабевает.

С другой стороны, оказалось, что для решения задачи нахождения распределения отклонения от расписания при прибытии на конечную станцию не требуется знать корреляционные связи между всеми случайными величинами или их распределения, а достаточно знать только совместное распределение двух следующих случайных величин: отклонения от расписания при отправлении с предпоследней станции и времени хода на последнем участке.

Показано, что теоретическое распределение, полученное в результате предложенной модели, в случае независимости двух указанных случайных величин, согласуется со статистическими данными реального движения. В дальнейшем предполагается рассмотреть более общий случай, когда условие независимости не выполняется. 

Литература

1. Cacchiani, V. An overview of recovery models and algorithms for real-time railway rescheduling / V. Cacchiani, D. Huisman, M. Kidd, L. Kroon, P. Toth, L.P. Veelenturf, J. Wagenaar // *Transportation Research, Part B*. —2014. —vol. 63. —P. 15–37.
2. Buker, T. Stochastic modelling of delay propagation in large networks / T. Buker, B. Seybold // *Journal of Rail Transport Planning & Management*. —2012. —vol. 2. —P. 34–50.
3. Kecman, P. Train delay evolution as a stochastic process / P. Kecman, F. Corman, L. Meng // *The 6th International Conference on Railway Operations Modelling and Analysis (RailTokyo2015)*. —2015. —P. 1–19 (007).
4. Corman, F. Stochastic prediction of train delays in real-time using Bayesian networks / F. Corman, P. Kecman // *Transportation Research, Part C: Emerging Technologies*. —2018. —vol. 95. —P. 599–615.
5. Berger, A. Stochastic Delay Prediction in Large Train Networks / A. Berger, A. Gebhardt, M. Müller-Hannemann, M. Ostrowski // *The 11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS'11)*, Saarbrücken, Germany. —2011. —P. 100–111.
6. Carey, M. Stochastic approximation to the effects of headways in knock-on delays of trains / M. Carey, A. Kwiecieński // *Transportation Research, Part B*. —1994. —vol. 28. —P. 251–267.
7. Вентцель, Е.С. Теория вероятностей : учебник для студентов вузов / Е.С. Вентцель. —10-е изд., стер. —Москва: Academia, 2005. —572 с. —ISBN 5-7695-2311-5. —Текст: непосредственный.

8. Carey, M. Properties of expected costs and performance measures in stochastic models of scheduled transport / M. Carey, A. Kwieceński // European Journal of Operational Research. – 1995. – vol. 83. – P. 182–199.
9. Централизованное управление движением поездов городских железных дорог современного мегаполиса / Л.А. Баранов, Е.П. Балакина, С.Е. Иконников, Д.А. Антонов. – Текст: непосредственный // Наука и техника транспорта. – 2020. – № 1. – С. 30–38.
10. Davydov, V. On the distribution of the random arrival time / V. Davydov, V. Chebotarev, K. Kablukova // WIT Transactions on The Built Environment. – 2020. – vol. 199. – P. 193–205. DOI 10.2495/CR200181
11. Теория вероятностей и математическая статистика / В.С. Мхитарян, Л.И. Трошин, Е.В. Адамова, К.К. Шевченко. – Москва: Московский международный институт эконометрики, информатики, финансов и права, 2003. – 130 с. – Текст: непосредственный.
12. Справочник по теории вероятностей и математической статистике / В.С. Королюк, Н.И. Портенко, А.В. Скороход, А.Ф. Турбин. – 2-е изд., перераб. и доп. – Москва: Наука, 1985. – 640 с. – Текст: непосредственный.
13. Административно-управленческий портал: сайт. – URL: http://www.aup.ru/books/m155/4_16.htm (дата обращения: 14.12.2021). – Текст: электронный.
14. Электронный учебник по статистике: сайт. – URL: <http://statsoft.ru/home/textbook/modules/sttable.html> (дата обращения: 14.12.2021). – Текст: электронный.