Статистическое моделирование

Д.Е. Бурланков

Нижегородский университет им. Н.И. Лобачевского

Нижний Новгород, 2006

Фредерик У. Ланчестер

Один из основоположников автомобилестроения в Англии и теории полета самолета Фредерик У. Ланчестер (1868 – 1946), анализируя битвы Первой мировой войны, построил предельно простую систему дифференциальных уравнений, описывающих динамику уменьшения численности x и y войск двух воюющих сторон:

$$\dot{x} = -k_y\, y; \quad \dot{y} = -k_x\, x. \tag{1}$$

Это линейная система уравнений, имеющая точное решение

$$x = x_0 \operatorname{ch}\left(\frac{t}{\tau}\right) - y_0 \sqrt{\frac{k_y}{k_x}} \operatorname{sh}\left(\frac{t}{\tau}\right);$$

$$y = y_0 \operatorname{ch}\left(\frac{t}{\tau}\right) - x_0 \sqrt{\frac{k_x}{k_y}} \operatorname{sh}\left(\frac{t}{\tau}\right),$$

где $\tau = 1/\sqrt{k_x\, k_y}$. Если в какой-то момент $t = t_1$ войска y будут полностью уничтожены $y = 0$:

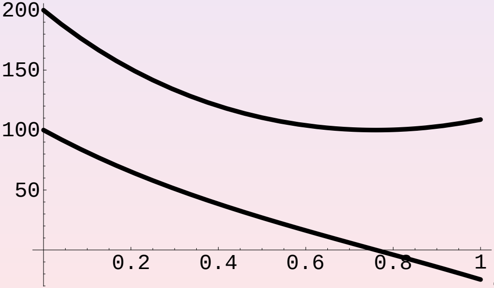
$$\operatorname{th}\left(\frac{t_1}{\tau}\right) = \frac{y_0\, \sqrt{k_y}}{x_0\, \sqrt{k_x}}.$$

Для достижения победы войсками x необходимо

$$k_x\, x_0^2 > k_y\, y_0^2. \tag{2}$$

Пример

$$x_0 = 200;\ k_x = 1; \quad y_0 = 100;\ k_y = 3; \quad \frac{k_x\, x_0^2}{k_y\, y_0^2} = \frac{4}{3}.$$



;

Мировая динамика Дж. Форрестера – 60-е годы

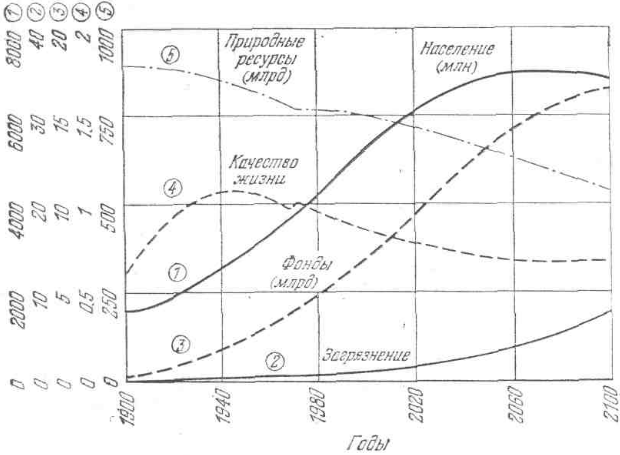
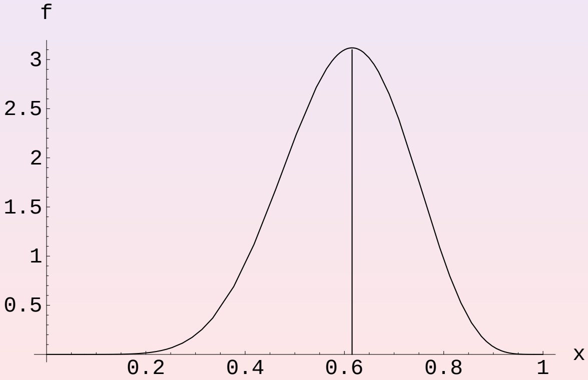


Рис. 6.1. Темп использования природных ресурсов и производство загрязнения уменьшены в 1970 г.

Статистическое моделирование

При составлении сетевых графиков в строительстве для учета неопределенности начала и окончания работ используют β-распределение.



;

Теория очередей

Классическая теория очередей исходит из возможности очереди иметь любую длину от 0 до ∞.
Рассмотрение марковского процесса переходов между уровнями приводит к распределению длин очереди по Паскалю с параметром, определяемым отношению скорости прихода клиентов q к скорости обслуживания r: $\gamma = q/r$:

$$p_n = (1-\gamma)\,\gamma^n$$

с математическим ожиданием и дисперсией

$$< n >= \frac{\gamma}{1-\gamma}; \quad D_n = \frac{\gamma}{(1-\gamma)^2} =< n > (< n > +1).$$

Вероятное число мы будем представлять тройкой чисел: $[a,\ b,\ r]$, где первые два числа определяют величины значений $x_1 < x_2$:

$$a = \frac{x_1 + x_2}{2}; \quad b = \frac{|x_2 - x_1|}{2} > 0, \tag{3}$$

а третье – вероятностный параметр:

$$p(x_1) = \frac{1-r}{2}; \quad p(x_2) = \frac{1+r}{2}; \quad -1 < r < 1. \tag{4}$$

Первые семиинварианты этого распределения

$$\chi_1^B \equiv < x >^B = \frac{1-r}{2}\,(a-b) + \frac{1+r}{2}\,(a+b) = a + b\,r;$$

$$\chi_2^B = D_x^B = b^2\,(1-r^2); \tag{5}$$

$$\chi_3^B \equiv \alpha^B = -2\,b^3\,r\,(1-r^2).$$

Аппроксимация вероятными числами

Добиваясь аппроксимации вероятным числом некоторого распределения с сохранением значений первых трех семиинвариантов, нужно решить систему уравнений:

$$a + b\,r = < x >; \quad b^2\,(1 - r^2) = D_x; \quad -2\,b^3\,r\,(1 - r^2) = \alpha_x,$$

Решая эту систему, находим значения параметров бинарного распределения $[a,\ b,\ r]$, аппроксимирующего исходное:

$$k = \frac{\alpha}{2\,D}; \quad a = \chi_1 + k; \quad b = \sqrt{k^2 + D}; \quad r = -\frac{k}{b}. \qquad (6)$$

Сложение вероятных чисел

Пусть к величине x представляемой вероятным числом $[A,\, B,\, R]$ добавляется случайная величина y, также представляемой вероятным числом $[a,\, b,\, p]$. Результирующая величина $z = x + y$ принимает четыре значения

$$z_1 = A - B + a - b; \quad z_2 = A - B + a + b;$$

$$z_3 = A + B + a - b; \quad z_4 = A + B + a + b$$

с вероятностями, соответственно,

$$\frac{(1-R)(1-r)}{4};\ \frac{(1-R)(1+r)}{4};\ \frac{(1+R)(1-r)}{4};\ \frac{(1+R)(1+r)}{4}.$$

Преобразование системы из четырех случайных чисел в вероятное число (две случайных величины) проводится на основе сохранения первых трех семиинвариантов.

Нужно вычислить три момента:

$$m_1 = x_1\, p_1 + x_2\, p_2 + x_3\, p_3 + x_4\, p_4;$$

$$m_2 = x_1^2\, p_1 + x_2^2\, p_2 + x_3^2\, p_3 + x_4^2\, p_4;$$

$$m_3 = x_1^3\, p_1 + x_2^3\, p_2 + x_3^3\, p_3 + x_4^3\, p_4,$$

через которые вычисляются семиинварианты:

$$\chi_1 = m_1; \quad \chi_2 = m_2 - m_1^2;$$

$$\chi_3 = m_3 - 3\, m_1\, m_2 + 2\, m_1^3.$$

Через семиинварианты возвращется вероятное число.

Биномиальное распределение

$$< x >= n\,(a + b\,r) = A + B\,R; \quad D_x = n\,b^2\,(1 - r^2) = B^2\,(1 - R^2);$$

$$\alpha_x = -2\,n\,b^3\,r\,(1 - r^2) = -2\,B^3\,R\,(1 - R^2).$$

Из этой системы находим

$$B\,R = b\,r; \quad A = n\,a + (n-1)\,b\,r; \quad \frac{1 - R^2}{R^2} = n\,\frac{1 - r^2}{r^2},$$

откуда

$$R = \frac{r}{\sqrt{n - (n-1)\,r^2}}; \quad B = b\,\sqrt{n - (n-1)\,r^2}.$$

В симметричном случае ($r = 0$)

$$R = 0; \quad A = n\,a; \quad B = b\sqrt{n}.$$

Вероятный уровень – это бинарная система, в которой одно из чисел равно нулю, а другое x реализуется с вероятностью p. Представляется парой чисел $[a, s]$, где a – математическое ожидание, а $(s\,a) = \sqrt{D}$ – норма:

$$p = 1/(1 + s^2); \quad a = p\,x; \quad x = \frac{a}{p}; \quad D = x\,p(1 - p) = (a\,s)^2.$$

При аппроксимиации некоторого распределения вероятным уровнем – через математическое ожидание $< x >$ и дисперсию D_x:

$$a =< x >, \quad s = \frac{\sqrt{D_x}}{a}.$$

Сложение, умножение и вычитание

Сложение уровней $[a,\ s] = [a_1, s_1], [a_2, s_2],\ a_1 > 0\ a_2 > 0$:

$$a = a_1 + a_2; \quad s = \frac{1}{a}\sqrt{a_1^2\, s_1^2 + a_2^2\, s_2^2}.$$

Суммирование n одинаковых

$$z_n = [n\, a,\ s\sqrt{n}].$$

Умножение

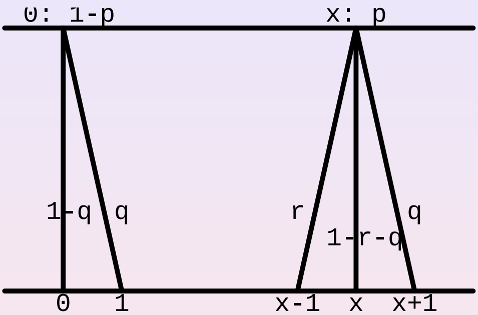
$$[a_1,\ s_1] \times [a_2,\ s_2] = [a_1\, a_2,\ s_1\, s_2].$$

Вычитание $(a > b\, p)$

$$[a,\ s] - [b,\ r] = \left[a - \frac{b}{1+s^2}, \frac{\sqrt{a^2\, s^2 + b^2\, r^2}}{a - \frac{b}{1+s^2}}\right].$$

Уровневая аппроксимация очереди

Очередь представляется уровнем x с вероятностью p, а с вероятностью $1 - p$ она имеет нулевую длину, так что математическое ожидание длины $a = x\,p$. За некоторый интервал времени с вероятностью q ее длина может увеличиться (новый клиент) от значения x до величины $x + 1$, имеющей вероятность $p\,q$ и от нулевого уровня, приняв значение 1 с вероятностью $(1 - p)\,q$. За этот же интервал она может с вероятностью r укоротиться (клиент обслужен), но только от верхнего уровня, приняв значение $x - 1$ с вероятностью $p\,(1 - q)$. И, наконец, с вероятностью 1-q-r ее длина не изменится, сохранив значение x.



В конце периода имеется пять значений с вероятностями

Значение	0	1	x-1	x	x+1
Вероятность	(1-p)(1-q)	(1-p)q	p r	p (1-q-r)	p q

x и p определяются из условия неизменности, для чего должны остаться неизменными математическое ожидание и дисперсия.

$$< x >= q + p\,(x - r) = p\,x;$$
$$m_2 = q + 2\,p\,q\,x + p\,(r - 2\,r\,x + x^2) = p\,x^2.$$

Отсюда $p = q/r$, откуда следует, что стационарное решение имеется лишь при $q < r$. Далее находим

$$x = \frac{r}{r - q}; \quad a = \frac{q}{r - q} = \frac{\gamma}{1 - \gamma}; \quad \gamma = \frac{q}{r}.$$

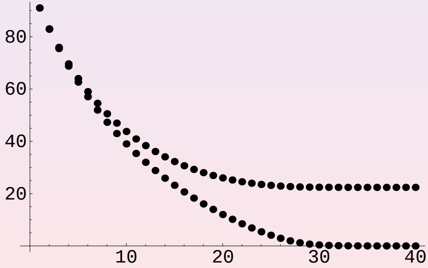
Математическое ожидание совпадает с точным значением. Дисперсия $D = a$ отличается от точного значения $D = a(a + 1)$

Статистическая Ланчестерская динамика

Численность войск определяем уровнями $[a_1, s_1]$ и $[a_2, s_2]$. Коэффициенты эффективности также представляются вероятными уровнями $[k_1, r_1]$ и $[k_2, r_2]$. Пример:

$z1 = [100, 0.33]; m1 = [0.1, 0.65];$

$z2 = [100, 0.33]; m2 = [0.1, 1.53];$

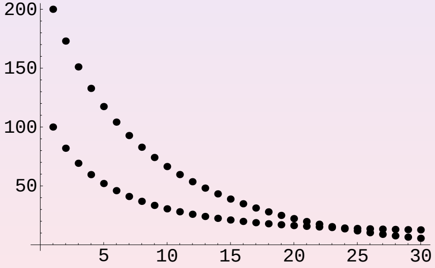


Первая армия побеждает за счет меньшего разброса в эффективности (верхняя кривая).

Вот результат победы армии с меньшей численностью, но лучшей эффективностью:

$z1 = [200, 0.33]; m1 = [0.1, 2];$

$z2 = [100, 0.33]; m2 = [0.3, 0.5];$



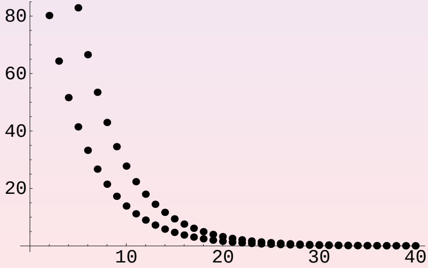
При *критическом Ланчестерском соотношении*

$$k_1 \, n_1^2 = k_2 \, n_2^2$$

при одинаковых разбросах наблюдается динамика Ланчестера, затянутая во времени:

$z1 = [200, 0.1]; m1 = [0.1, 0.1];$

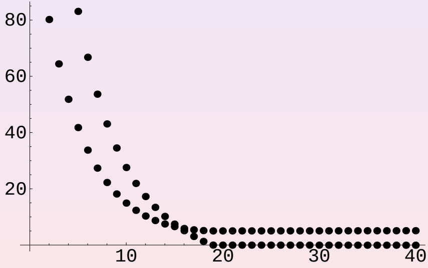
$z2 = [100, 0.1]; m2 = [0.4, 0.1];$



При Ланчестерском соотношении победа определяется минимумом разброса эффективности:

$z1 = [200, 0.1]; m1 = [0.1, 0.3];$

$z2 = [100, 0.1]; m2 = [0.4, 0.2];$



Выводы

- Вероятные числа моделируют любое распределение, сохраняя первых три семиинварианта.
- Вероятные уровни моделируют рвспределения положительных случайных чисел с сохранением математического ожидания.
- Операции сложения, вычитания, суммирования, умножения, деления, и возведения в степень вероятных чисел имеют результатом вероятные числа. То же и для вероятных уровней.
- Это свойство позволяет легко модифицировать программы динамического моделирования в программы вероятностного моделирования.
- Работа с вероятными числами и уровнями упрощает работу экспертов.