

С. М. Ермаков, Т. И. Мисов

МОДЕЛИРОВАНИЕ Δ^2 -РАСПРЕДЕЛЕНИЯ*

1. Введение

Распределение Δ^2 играет важную роль в задачах уменьшения дисперсии при вычислении интегралов методом Монте-Карло [1], а также в задачах анализа систематической погрешности в регрессионном анализе. Свойства этого распределения подробно описаны [2].

Использование Δ^2 -распределения требует его моделирования. Самым простым представляется метод отбора. Он требует при каждом испытании вычисления определителя возможно большого порядка, а эффективность в общем случае определяется оценкой его максимума. При этом, неравенство Адамара дает грубую оценку.

Изучен одномерный случай полиномиальной регрессии на отрезке [3]. В этом случае оценкой определителя, с точностью до константы, служит величина определителя Вандермонда, но при $n > 2$ она довольно груба. В этом простейшем случае вычисления по методу отбора до $n = 12$ были достаточно эффективными.

Для многомерного случая до недавнего времени считалось, что моделирование Δ^2 -распределения — задача безнадежно трудная.

В данной работе для случая s -мерной области и лебеговской меры μ , авторы используют наряду с методом отбора другие методы, основанные на вычислении условных плотностей и методе композиции. Удастся значительно уменьшить трудоемкость алгоритма. Сравнение «чистого» метода отбора и предлагаемого комбинированного асимптотически показывает преимущества последнего. Авторы также предполагают, что полученные результаты удастся распространить на случай более сложных конструкций меры μ .

2. Использование условных плотностей

Пусть на множестве X элементов x определена σ -алгебра A и σ -конечная мера μ . Обозначим $\mu^n = \otimes_{k=1}^n \mu$ и предположим, что существует n линейно независимых на носителе меры μ функций (μ -линейно независимых). Если эти функции μ -интегрируемы с квадратом, что также предполагается, то они могут быть ортонормированы. Полученную ортонормированную систему обозначим $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$. Плотность

$$\Delta^2(Q) = \frac{1}{n!} \left(\det \|\varphi_i(x_j)\|_{i,j=1}^n \right)^2, \quad Q = (x_1, \dots, x_n)$$

называют плотностью по отношению к мере μ^n Δ^2 -распределения. Отметим, что требование ортогональности $\varphi_i, i = \overline{1, n}$, вызвано исключительно желанием упростить выкладки. Процедура ортогонализации Грамма—Шмидта не изменяет величины определителя.

Предлагаемая в этой статье идея моделирования Δ^2 -распределения связана с использованием условных плотностей.

Укажем сначала вид плотностей $p_k(x_1, \dots, x_k), k = \overline{1, n}$. Для этого воспользуемся хорошо известным лапласовским разложением определителя.

* Работа выполнена при финансовой поддержке РФФИ (грант № 05-01-00865).

© С. М. Ермаков, Т. И. Мисов, 2005

Лемма 2.1 (Разложение Лапласа) Пусть D — определитель n -ого порядка. Фиксируем любые m его строк $i_1, \dots, i_m$, $m < n$. Тогда

$$D = \sum_{(j_1, \dots, j_m) \in (1, \dots, n)} M_{i_1, \dots, i_m; j_1, \dots, j_m} M'_{i_1, \dots, i_m; j_1, \dots, j_m},$$

где $M_{i_1, \dots, i_m; j_1, \dots, j_m}$ — минор D , получающийся при пересечении строк с номерами $i_1, \dots, i_m$ и столбцов с номерами $j_1, \dots, j_m$, а $M'_{i_1, \dots, i_m; j_1, \dots, j_m}$ — алгебраическое дополнение $M_{i_1, \dots, i_m; j_1, \dots, j_m}$.

Утверждение леммы верно, если вместо m строк фиксировать m столбцов.

Лемма 2.2. Пусть функции $\varphi_1, \varphi_2, \dots, \varphi_n$ образуют ортонормированную систему в пространстве $L^2(X, \mu)$, где X — измеримое множество, а μ — мера Лебега. Обозначим

$$p_n(x_1, \dots, x_n) = \frac{1}{n!} \left(\det \|\varphi_i(x_j)\|_{i,j=1}^n \right)^2,$$

а

$$p_{n-k}(x_1, \dots, x_{n-k}) = \int_X p_n(x_1, \dots, x_n) dx_{n-k+1} \dots dx_n.$$

Тогда

$$p_{n-k}(x_1, \dots, x_{n-k}) = \frac{k!}{n!} \sum_{1 \leq i_1 < \dots < i_{n-k} \leq n} \left(\det \|\varphi_{i_s}(x_j)\|_{s,j=1}^{n-k} \right)^2$$

Доказательство. Воспользуемся разложением Лапласа для определителя $p_n(x_1 \dots x_n)$, фиксируя первых $n-k$ его столбцов:

$$\begin{aligned} p_n(x_1 \dots x_n) &= \frac{1}{n!} \left(\det \|\varphi_i(x_j)\|_{i,j=1}^n \right)^2 = \\ &= \frac{1}{n!} \left(\sum \left(\det \|\varphi_{i_s}(x_l)\|_{s,l=1}^{n-k} \right) \left(\det \|\varphi_{j_m}(x_r)\|_{m=1, r=n-k+1}^k \right) \right)^2, \end{aligned}$$

где суммирование ведется по $1 \leq i_1 < \dots < i_{n-k} \leq n$, а $j_1, \dots, j_k$ — дополнение набора индексов $\{i_1 \dots i_{n-k}\}$ до $\{1, 2, \dots, n\}$.

Обозначим $\left(\det \|\varphi_{i_s}(x_l)\|_{s,l=1}^{n-k} \right) =: \Delta_{i_1 \dots i_{n-k}}$. Имеем

$$\begin{aligned} p_{n-k}(x_1, \dots, x_{n-k}) &= \int_X p_n(x_1, \dots, x_n) dx_{n-k+1} \dots dx_n = \\ &= \frac{1}{n!} \int_X \left(\sum_{1 \leq i_1 < \dots < i_{n-k} \leq n} \Delta_{i_1 \dots i_{n-k}} \left(\det \|\varphi_{j_m}(x_r)\|_{m=1, r=n-k+1}^k \right) \right)^2 dx_{n-k+1} \dots dx_n = \\ &= \frac{1}{n!} \int_X \left[\sum_{1 \leq i_1 < \dots < i_{n-k} \leq n} \Delta_{i_1 \dots i_{n-k}} \left(\sum_{l=1}^k (-1)^{n+l} \varphi_{j_l}(x_n) \Delta'_{j_1 \dots j_{l-1} j_{l+1} \dots j_k} \right) \right]^2 dx_{n-k+1} \dots dx_n, \end{aligned}$$

где

$$\Delta'_{j_1 \dots j_{l-1} j_{l+1} \dots j_k} = \left(\det \|\varphi_{j_1}(x_s), \dots, \varphi_{j_{l-1}}(x_s), \varphi_{j_{l+1}}(x_s), \dots, \varphi_{j_k}(x_s)\|_{s=n-k+1}^n \right)^2.$$

Проинтегрируем сначала по x_n . Возведем в квадрат подынтегральное выражение и воспользуемся тем, что функции $\varphi_1, \varphi_2, \dots, \varphi_n$ образуют ортонормированную систему

в пространстве $L^2(X)$. Благодаря этому, попарные произведения в подынтегральном выражении, содержащие множители $\varphi_i(x_n)\varphi_j(x_n)$, $i \neq j$, обратятся в ноль. Более того, определитель $\Delta_{i_1 \dots i_{n-k}}$ под знаком внешней суммы не зависит от x_n . Поэтому его можно вынести за знак интеграла.

Наборов $(j_1, \dots, j_k) C_n^k$ штук. Каждому из них в подынтегральном выражении соответствуют k слагаемых, состоящих из членов с элементами $(s_1, \dots, s_{k-1}) \subset (j_1, \dots, j_k)$. Следовательно, общее количество членов становится равным kC_n^k . Из них ровно C_n^k различных. Таким образом, пользуясь ортогональностью функций $\varphi_1, \varphi_2, \dots, \varphi_n$ в $L^2(X)$, получим

$$p_{n-k}(x_1, \dots, x_{n-k}) = \frac{k}{n!} \int_X \left(\sum_{1 \leq i_1 < \dots < i_{n-k} \leq n} \Delta_{i_1 \dots i_{n-k}} \times \right. \\ \left. \times \begin{vmatrix} \varphi_{s_1}(x_{n-k+1}) & \dots & \varphi_{s_1}(x_{n-1}) \\ \vdots & \ddots & \vdots \\ \varphi_{s_{k-1}}(x_{n-k+1}) & \dots & \varphi_{s_{k-1}}(x_{n-1}) \end{vmatrix} \right)^2 dx_{n-k+1} \dots dx_{n-1}$$

Интегрируя по остальным переменным и аналогично рассуждая, приходим к утверждению леммы. Лемма доказана.

Укажем теперь вид условных плотностей $p_k(x_k | x_1 \dots x_{k-1})$, $k = 1, \dots, n$.

Следствие 2.1. Для условных плотностей $p_k(x_k | x_1 \dots x_{k-1})$, $k = 1, \dots, n$ справедливы формулы

$$p_k(x_k | x_1 \dots x_{k-1}) = \sum_{1 \leq i_1 < \dots < i_k \leq n} \frac{\Delta_{i_1 \dots i_k}^2}{(n-k+1) \sum_{1 \leq j_1 < \dots < j_{k-1} \leq n} \Delta_{j_1 \dots j_{k-1}}^2},$$

где

$$\Delta_{i_1 \dots i_s}^2 = \left(\det \|\varphi_{l_m}(x_r)\|_{m,r=1}^s \right)^2$$

ДОКАЗАТЕЛЬСТВО. Действительно, для любого $k = 1 \dots n$ имеем

$$p_k(x_k | x_1 \dots x_{k-1}) = \frac{p_k(x_1 \dots x_k)}{p_{k-1}(x_1 \dots x_{k-1})}.$$

Осталось лишь воспользоваться леммой 2.2. Следствие доказано.

3. Представление условных плотностей в виде композиции

Согласно следствию 2.1 условные плотности $p_k(x_k | x_1 \dots x_{k-1})$, $k = 1, \dots, n$ представимы в виде линейной комбинации определителей k -го порядка. Вводя соответствующую нормировку, мы получим представление условных плотностей в виде смеси:

$$\sum_{1 \leq i_1 < \dots < i_k \leq n} \frac{\sum_{(m_1 < \dots < m_{k-1}) \subset (i_1 \dots i_k)} \Delta_{m_1 \dots m_{k-1}}^2}{(n-k+1) \sum_{1 \leq j_1 < \dots < j_{k-1} \leq n} \Delta_{j_1 \dots j_{k-1}}^2} \times \\ \times \frac{\Delta_{i_1 \dots i_k}^2}{\sum_{(m_1 < \dots < m_{k-1}) \subset (i_1 \dots i_k)} \Delta_{m_1 \dots m_{k-1}}^2} = p_k(x_k | x_1 \dots x_{k-1}).$$

4. Сравнение трудоемкостей предложенного метода моделирования и метода отбора

Под трудоемкостью (сложностью) метода моделирования мы будем понимать среднее число арифметических операций, необходимое для вычисления реализации случайного вектора $Q = (x_1, \dots, x_n)$.

Если плотность распределения Δ^2 ограничена константой M , то трудоемкость простого метода отбора складывается из трудоемкостей следующих операций:

- 1) получение необходимого числа $ns + 1$ реализаций случайной величины, равномерно распределенной на $[0, 1]$;
- 2) вычисление $\varphi_i(x_j)$, $i = 1, 2, \dots, n$ в равномерно распределенных в гиперкубе случайных точках;
- 3) вычисление квадрата определителя порядка n ;
- 4) пункты 1–3 повторяются в среднем $[M]$ раз.

Лемма 4.1. *Трудоемкость моделирования Δ^2 -распределения методом отбора не превосходит величину*

$$\left(\frac{2n^3 + n + 3}{3} + cn^2 \right) \frac{n^n (\max_{x \in X, i=1, \dots, n} |\varphi_i(x)|)^{2n}}{n!}, \quad (1)$$

где c — средняя сложность однократного вычисления функции φ_i , $i = \overline{1, n}$ в любой точке $x \in X$. При условии, что все $\varphi(x)$, $x \in X$ отличны от нуля, с вероятностью 1 в (1) имеет место знак равенства.

ДОКАЗАТЕЛЬСТВО. Трудоемкость метода отбора $M \cdot (C_{\det} + C_{\text{func}})$, где M — максимальное значение плотности Δ^2 , $C_{\det}$ — трудоемкость вычисления определителя порядка n , а C_{func} — трудоемкость вычисления функций $\{\varphi_i(x_j)\}_{i,j=1}^n$.

Оценка максимального значения M дается неравенством Адамара:

$$\Delta^2(Q) = \frac{(\det \|\varphi_i(x_j)\|_{i,j=1}^n)^2}{n!} \leq \frac{n^n (\max_{x \in X, i=\overline{1, n}} |\varphi_i(x)|)^{2n}}{n!}.$$

Если считать, что средняя трудоемкость вычисления значения функции $\varphi_i(x)$ в точке x_j ($i, j = \overline{1, n}$), включая получение sn реализаций равномерно распределенных случайных величин, есть c , то $C_{\text{func}} = cn^2$ (мы имеем n функций, которые надо вычислить в n точках).

Осталось понять чему равняется $C_{\det}$. Возьмем матрицу $A = \|a_{ij}\|_{i,j=1}^n$. Рассмотрим сначала случай, когда все элементы матрицы A ненулевые. Приведем матрицу A к диагональному виду. Сначала умножим первую строку на $\frac{-a_{21}}{a_{11}}$ и прибавим ее ко второй. Таким образом получим ноль на первом месте во второй строчке. При этом мы совершаем $2n$ операций (n умножений и n сложений). Продолав ровно такую же процедуру с третьей по n -ую строчку, мы получим в первом столбце (за исключением первого элемента) нули. Количество совершенных операций будет равняться $2n(n-1)$, т. е. $(n(n-1)$ умножений и $n(n-1)$ сложений).

Поступаем аналогичным образом и с каждым из остальных столбцов. В k -ом ($k = \overline{2, n-1}$) столбце получаем нули, начиная с $(k+1)$ -го места, совершая при этом $2(n-k)(n-k+1)$ операций, т. е. $(n-k)(n-k+1)$ умножений и $(n-k)(n-k+1)$ сложений.

Таким образом, при приведении матрицы A к диагональному виду мы совершаем

следующее число операций:

$$\sum_{k=1}^{n-1} 2(n-k)(n-k+1) = \sum_{l=2}^n 2l(l-1) = \frac{2n(n^2-1)}{3}$$

Осталось добавить еще n умножений диагональных элементов и одно возведение полу-ченного определителя в квадрат. Если некоторые $\varphi_i(x_j)$ оказываются нулями, то число операций будет разве лишь меньше подсчитанного выше. Лемма доказана.

Сосчитаем теперь трудоемкость моделирования Δ^2 -распределения с помощью пред-ложенного нами метода.

В этой работе предлагается моделирование, включающее в себя следующие подходы:

а) метод обращения — представляем

$$\Delta^2(Q) = p_1(x_1)p_2(x_2|x_1)p_3(x_3|x_1x_2) \dots p_n(x_n|x_1x_2 \dots x_{n-1});$$

б) метод композиций — каждая из плотностей

$$p_1(x_1), p_2(x_2|x_1), p_3(x_3|x_1x_2), \dots, p_n(x_n|x_1x_2 \dots x_{n-1})$$

моделируется как смесь распределений с плотностями

$$\frac{\Delta_{i_1 i_2 \dots i_k}^2}{\sum_{(m_1 < \dots < m_{k-1}) \subset (i_1 \dots i_k)} \Delta_{m_1 \dots m_{k-1}}^2}, \quad k = \overline{1, n},$$

которые имеют вид

$$\frac{\left(\sum_{i \in \{i_1, i_2, \dots, i_k\}, 1 \leq i_1 < \dots < i_k \leq n} c_{\{i_1, i_2, \dots, i_k\}/\{i\}}^{(k)} \varphi_i(x_k) \right)^2}{\sum_{(m_1 < \dots < m_{k-1}) \subset (i_1 \dots i_k)} \Delta_{m_1 \dots m_{k-1}}^2}; \quad (2)$$

в) метод отбора — каждая из последних плотностей моделируется методом отбо-ра, поэтому сложность их моделирования равна $\left(\max_{x_k \in X, i = \overline{1, k}} |\varphi_i(x_k)| \right)^2 C_k$, где C_k — сложность нахождения коэффициентов $c_{\{i_1, i_2, \dots, i_k\}/\{i\}}^{(k)}$. Покажем, что эти коэффициен-ты получаются итеративно и укажем процедуру их нахождения.

Замечание. В некоторых случаях более удобным может быть моделирование по-следних плотностей методом обращения. Этот случай подробно не исследуется.

Лемма 4.2. Условная плотность k -го порядка $p_k(x_k|x_1x_2 \dots x_{k-1})$ представима в виде

$$\sum_{1 \leq i_1 < \dots, i_k \leq n} \frac{\sum_{j=1}^k \left(c_{\{i_1, i_2, \dots, i_k\}/\{i_j\}}^{(k)} \right)^2}{(n-k+1) \sum_{1 \leq j_1 < \dots, j_{k-1} \leq n} \left(c_{j_1 j_2 \dots j_{k-1}}^{(k)} \right)^2} \times \frac{\left(\sum_{m=1}^k (-1)^{m+1} c_{\{i_1, i_2, \dots, i_k\}/\{i_m\}}^{(k)} \varphi_{i_m}(x_k) \right)^2}{\sum_{j=1}^k \left(c_{\{i_1, i_2, \dots, i_k\}/\{i_j\}}^{(k)} \right)^2}, \quad (3)$$

где

$$c_{\{i_1, i_2, \dots, i_k\}/\{i_j\}}^{(k)} = \sum_{s=1, s \neq j}^k (-1)^{s+1} c_{\{i_1, \dots, i_k\}/\{i_j, i_s\}}^{(k-1)} \varphi_{i_s}(x_{k-1}). \quad (4)$$

ДОКАЗАТЕЛЬСТВО. Проведем индукцию по k .

База: $k = 2$. Имеем $p_1(x_1) = \frac{1}{4} \sum_{i=1}^n \varphi_i^2(x_1)$. Следовательно, $c^{(1)} = 1$. Рассмотрим вторую условную плотность

$$p_2(x_2|x_1) = \sum_{i_1 \neq i_2} \frac{\varphi_{i_1}^2(x_1) + \varphi_{i_2}^2(x_1)}{(n-1) \sum_{l=1}^n \varphi_l^2(x_1)} \frac{\begin{vmatrix} \varphi_{i_1}(x_1) & \varphi_{i_1}(x_2) \\ \varphi_{i_2}(x_1) & \varphi_{i_2}(x_2) \end{vmatrix}}{\varphi_{i_1}^2(x_1) + \varphi_{i_2}^2(x_1)}.$$

Для нее $c_m^{(2)} = \varphi_m(x_1) = \varphi_m(x_1) \cdot c^{(1)}$, $m = \overline{1, n}$. База получена.

Сделаем индукционный переход. Пусть лемма верна для $k = s$. Докажем, что она справедлива и при $k = s + 1$.

Из (3) при $k = s$ видно, что $c_{\{i_1, i_2, \dots, i_s\}/\{i_j\}}^{(s)}$ — это минор $(s-1)$ -го порядка матрицы $\|\varphi_i(x_j)\|_{i,j=1}^s$, полученный вычеркиванием j -ой строки и последнего $(s$ -го) столбца.

Рассмотрим теперь условную плотность $(s+1)$ -ого порядка. Раскладываем определитель $\Delta_{i_1 i_2 \dots i_{s+1}}^2$ по $(s+1)$ -ому столбцу. Получим линейную комбинацию $\varphi_i(x_s)$, $i = \overline{1, s+1}$ с коэффициентами $c_{\{i_1, i_2, \dots, i_{s+1}\}/\{i_j\}}^{(s+1)}$. Ясно, что эти коэффициенты — миноры порядка s матрицы $\|\varphi_i(x_j)\|_{i,j=1}^{s+1}$, полученные вычеркиванием j -ой строки и последнего $((s+1)$ -ого) столбца. Разлагая теперь каждый из этих миноров по s -ому столбцу, получим соотношение (4). Лемма доказана.

Заметим, что для вычисления коэффициентов $c_{\{i_1, i_2, \dots, i_k\}/\{i_j\}}^{(k)}$ нужно совершить $2k - 3 + c(k-1)$ операций ($k-1$ вычисление функции $\varphi_i(x_j)$, $k-1$ умножение и $k-2$ сложения; величины $c_{\{i_1, i_2, \dots, i_k\}/\{i_j\}}^{(k-1)}$, $j \neq l$ нам известны). С другой стороны, на вычисление коэффициентов в смеси, в виде которой представляется плотность (3) отводится:

- 1) 0 операций, если $k = 1$;
- 2) $2C_n^{k-1} + kC_n^k$ операций, если $k = 2, 3, \dots, n-1$;
- 3) $2C_n^{n-1} = 2n$ операций, если $k = n$.

Следовательно, трудоемкость C_τ моделирования случайной величины τ с плотностью распределения $p_k(x_k|x_1 x_2 \dots x_{k-1})$, представимой в виде смеси есть

$$\sum_{1 \leq i_1 < \dots < i_k \leq n} \frac{\sum_{(l_1, l_2, \dots, l_{k-1}) \subset (i_1, \dots, i_k), l_1 < \dots < l_{k-1}} \Delta_{i_1 l_2 \dots l_{k-1}}^2}{(n-k+1) \sum_{1 \leq j_1 < \dots < j_{k-1} \leq n} \Delta_{j_1 j_2 \dots j_{k-1}}^2} C_1,$$

где C_1 — сумма трудоемкостей моделирования распределения с плотностью (2) и моделирования дискретного распределения, с помощью которого строится смесь. Если моделировать это дискретное распределение дихотомией, то его трудоемкость составит $\log_2 C_n^k$ (количество вероятностей для k -ой условной плотности есть C_n^k). Пусть $C_\varphi^{(k)} = \left(\max_{x_k \in X, j=\overline{1, k}} |\varphi_{i_j}(x_k)| \right)^2$. Тогда в силу того, что было показано выше

$$C_1 = \begin{cases} 0, & k = 1, \\ (k(c+2) - (c+3) + 2C_n^{k-1} + kC_n^k) C_\varphi^{(k)}, & 1 < k < n, \\ (n(c+4) - (c+3)) C_\varphi^{(n)}, & k = n. \end{cases}$$

Обозначив $S_{i_1, \dots, i_k} = \frac{\sum_{(l_1, l_2, \dots, l_{k-1}) \subset (i_1, \dots, i_k), l_1 < \dots < l_{k-1}} \Delta_{i_1 l_2 \dots l_{k-1}}^2}{(n-k+1) \sum_{1 \leq j_1 < \dots < j_{k-1} \leq n} \Delta_{j_1 j_2 \dots j_{k-1}}^2}$ и $C_k = (k(c+2) -$

$(c + 3) + 2C_n^{k-1} + kC_n^k + \log_2 C_n^k$, получим, что для $k = \overline{2, n-1}$

$$C_\tau = C_k \sum_{1 \leq i_1 < \dots < i_k \leq n} S_{i_1, \dots, i_k} C_\varphi^{(k)} \leq C_k C_\varphi^{(k)}.$$

Аналогичные оценки справедливы и для $k = 1$ и $k = n$.

Условных плотностей $p_1(x_1), p_2(x_2|x_1), \dots, p_n(x_n|x_1x_2 \dots x_{n-1})$ n штук, причем трудоемкость моделирования k -ой из них зависит только от k и $C_\varphi^{(k)}$. Следовательно, трудоемкость моделирования случайной величины, имеющей плотность распределения $\Delta^2(x_1, x_2, \dots, x_n)$, равна $C = \left(\sum_{k=2}^{n-1} C_k \right) C_\varphi^{(n)}$. Проведя несложные выкладки, получаем, что

$$\sum_{k=2}^{n-1} C_k = \frac{c+2}{2}n^2 - \frac{c+8}{2}n - 3 + (n+4)2^{n-1} + \log_2 \prod_{k=2}^{n-1} C_n^k.$$

Тем самым мы доказали следующую теорему

Теорема 4.1. *Для описанного метода трудоемкость моделирования Δ^2 -распределения равна*

$$C = \left(\frac{c+2}{2}n^2 - \frac{c+8}{2}n - 3 + (n+4)2^{n-1} + \log_2 \prod_{k=2}^{n-1} C_n^k \right) C_\varphi^{(n)}.$$

Сопоставим полученные в теореме 4.1 и лемме 4.1 оценки трудоемкостей. Для этого сначала, ради уменьшения сложности вычислений, оценим $\log_2 \prod_{k=2}^{n-1} C_n^k$ сверху:

$$\log_2 \prod_{k=2}^{n-1} C_n^k = \sum_{k=2}^{n-1} \log_2 C_n^k \leq \sum_{k=2}^{n-1} C_n^k = 2^n - n - 2.$$

От этого трудоемкость моделирования методом обращения только возрастет. Тем не менее, покажем, что даже при таком грубом завышении, предложенный метод дает заметный выигрыш по сравнению с простым методом отбора. Для определенности положим трудоемкость вычисления функции $\varphi_i(x_j)$ $c = 1$. Тем самым, рассмотрим отношение

$$\frac{\frac{3n^2}{2} - \frac{11n}{2} - 5 + (n+6)2^{n-1}}{\left(\frac{2n^3}{3} + n^2 - \frac{2n}{3} \right) \frac{n^n}{n!}}.$$

При таком выборе c это выражение максимально, поскольку коэффициент перед n^2 в числителе есть $c/2 - 1$, а в знаменателе c . Поэтому с ростом c мы будем получать все меньшие числа. Это подчеркивает преимущества предложенного метода, особенно, если моделирование $\varphi_i(x)$ достаточно трудоемко (напомним, что характеристикой этой трудоемкости является именно c):

$n = 2$	0,375
$n = 3$	0,427
$n = 4$	0,129
$n = 5$	0,066
$n = 6$	0,035
$n = 7$	0,019
$n = 8$	0,011
$n = 9$	$7,1 \cdot 10^{-3}$
$n = 10$	$3,9 \cdot 10^{-4}$
$n = 20$	$5,6 \cdot 10^{-5}$
$n = 50$	$1,26 \cdot 10^{-9}$

Таким образом, при больших значениях n (если не учитывать то, что трудоемкость в нашей схеме зависит от $\max |\varphi_i(x)|$, $i = \overline{1, n}$ в квадрате, а не в степени $2n$ как в простом отборе) предложенный нами метод моделирования дает серьезный выигрыш по сравнению с простым отбором.

Итак, мы показали, что распределение Δ^2 может быть промоделировано не только методом отбора (как до сего времени считалось), но и комбинацией методов обращения, отбора и композиций, причем с огромным выигрышем. Это дает возможность пользоваться распределением Δ^2 в множестве практических задач.

Summary

S. M. Ermakov, T. I. Missov. Simulation of the Δ^2 -distribution.

This article discusses the issue of effective simulation of the Δ^2 -distribution. Till recently it has been considered that the only applicable to the Δ^2 -distribution simulation method is the rejection one. The authors propose an approach, making use of the inversion, composition, and rejection method at the same time, which significantly decreases simulation complexity. The gain is obtained thanks to the structure of the conditional densities, which was first discovered by the authors. The paper presents the simulation algorithm, proofs of the theorems about the form of the conditional densities, and a complexity comparison between the suggested and the pure rejection method.

Литература

1. *Ермаков С. М., Золотухин В. Г.* Полиномиальные приближения и метод Монте-Карло // Теория вероятностей и ее применения (1960) 5, № 4, с. 473–475.
2. *Ермаков С. М.* Метод Монте-Карло и смежные вопросы. М., 1971. 327 с.
3. *Ermakov S. M., Schwabe R.* On Randomizing Estimators in Linear Regression Models // Advances in stochastic simulation methods. 2000. S. 305–314.

Статья поступила в редакцию 19 апреля 2005 г.