Воспроизведение дискретных распределений с заданной ковариационной структурой

Новоселов А.А.

Сибирский федеральный университет

Красноярский государственный торгово-экономический институт

Воспроизведение нормального распределения

- Задана ковариационная матрица C
- Строится разложение (Холецкий, вращение)

$$C = R'R$$

- Генерируется стандартный нормальный вектор Z
- Воспроизводится требуемый вектор

$$X = R'Z$$

Двумерный пример

- Корреляция 0,5

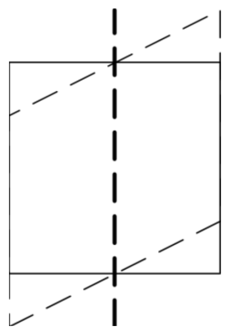
$$C = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}$$

- Разложение Холецкого

$$R' = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ \frac{1}{\sqrt{3}} & 1 \end{pmatrix} = R_2' R_1'$$

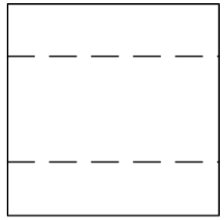
Шаги преобразования Холецкого

- Преобразование в целом $X = R_2' R_1' Z$
- Скос



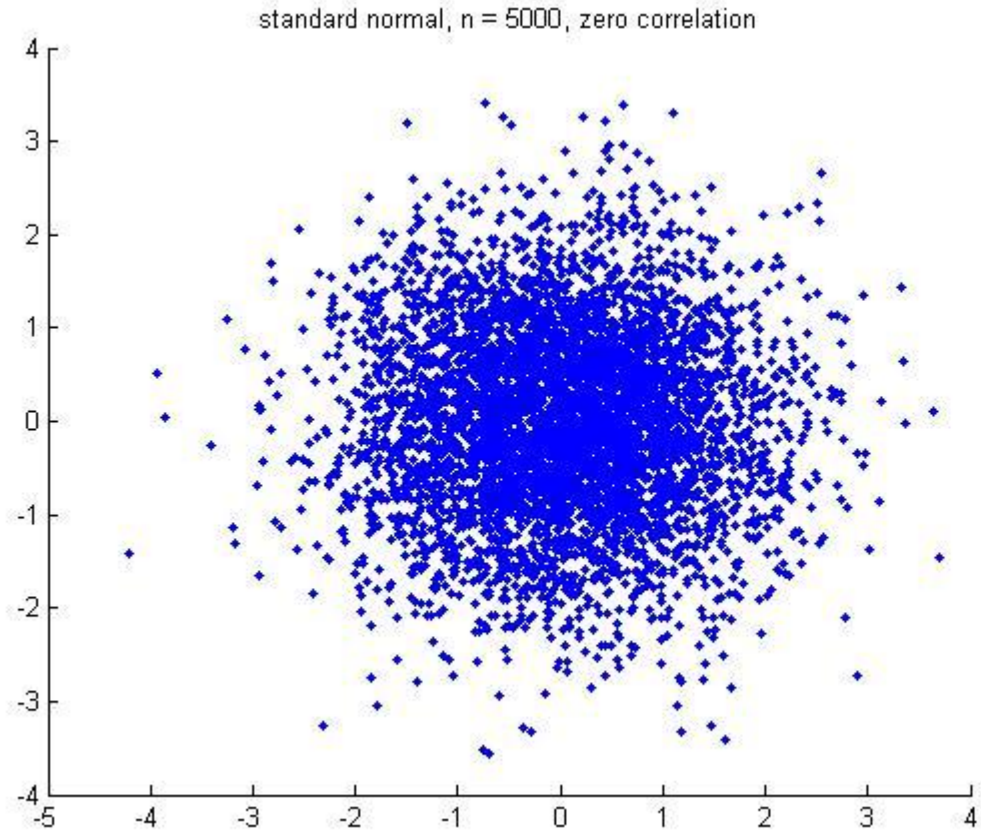
$$Y = R_1' Z$$

- Масштабирование

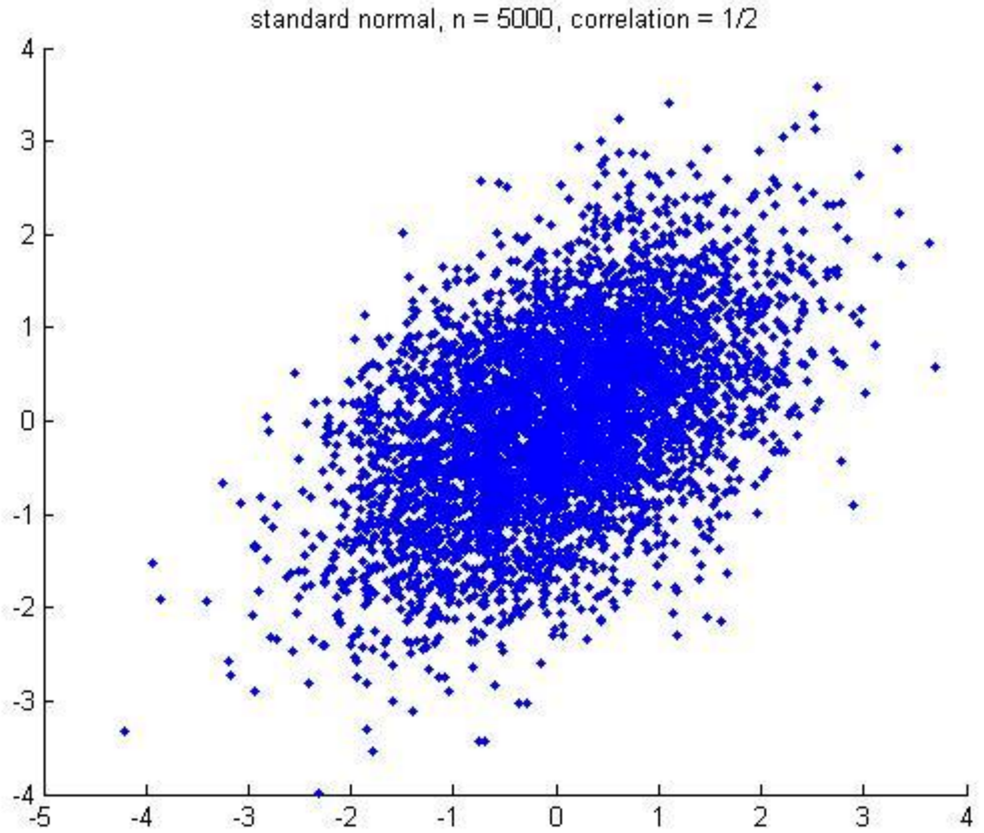


$$X = R_2' Y$$

Некоррелированный вектор Z



Коррелированный вектор X



Возможен ли такой же подход для дискретного распределения?

Нет

Задача

- Заданы дискретные маргинальные распределения компонент случайного вектора $X = (X_1, X_2)'$
- Задан коэффициент корреляции между ними ρ
- Построить совместное распределение с заданными характеристиками

Два вопроса

- Заданы маргинальные распределения. Возможен ли выбор произвольного коэффициента корреляции $\rho \in [-1, 1]$?
- Заданы маргинальные распределения и коэффициент корреляции. Единственным ли образом определяется по этой информации совместное распределение?
- Ответ на оба вопроса отрицателен

Дискретная сетка значений

$$\begin{array}{cccc} \bullet & \bullet & \cdots & \bullet \\ \cdots & \cdots & \cdots & \cdots \\ \bullet & \bullet & \cdots & \bullet \\ \bullet & \bullet & \cdots & \bullet \end{array}$$

Вращение сетки с целью достижения необходимой корреляции недопустимо; при этом разрушается сама сетка

Решение проблемы достигается «переливанием» вероятностной меры, что формально означает использование смеси некоторых базовых распределений

$$\begin{array}{cc|ccc} x_{21} & q_1 & r_{11} & \cdots & r_{m1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{2n} & q_n & r_{1n} & \cdots & r_{mn} \\ \hline & & p_1 & \cdots & p_m \\ & & x_{11} & \cdots & x_{1m} \end{array}$$

Пример 1: узкий диапазон корреляций

Сетка совместного распределения

0	1/2	•	•	•
1	1/2	•	•	•
		1/3	1/3	1/3
		0	1	2

Моменты маргинальных распределений

$$\mathbf{E}X_1 = 1, \quad \mathbf{E}X_2 = 1/2$$

$$\sigma_1 = \sqrt{2/3}, \quad \sigma_2 = 1/2$$

Диапазон значений коэффициента корреляции

$$\boxed{-\sqrt{\frac{2}{3}} \le \rho \le \sqrt{\frac{2}{3}}}$$

Пример 1, типовые распределения

Независимое распределение

$$\begin{array}{cc|ccc} 0 & 1/2 & 1/6 & 1/6 & 1/6 \\ 1 & 1/2 & 1/6 & 1/6 & 1/6 \\ \hline & & 1/3 & 1/3 & 1/3 \\ & & 0 & 1 & 2 \end{array}$$

Комонотонное распределение

$$\begin{array}{cc|ccc} 0 & 1/2 & 1/3 & 1/6 & 0 \\ 1 & 1/2 & 0 & 1/6 & 1/3 \\ \hline & & 1/3 & 1/3 & 1/3 \\ & & 0 & 1 & 2 \end{array}$$

Антикомонотонное распределение

$$\begin{array}{cc|ccc} 0 & 1/2 & 0 & 1/6 & 1/3 \\ 1 & 1/2 & 1/3 & 1/6 & 0 \\ \hline & & 1/3 & 1/3 & 1/3 \\ & & 0 & 1 & 2 \end{array}$$

Пример 1: все распределения

Распределение общего вида при заданных маргинальных распределениях

$$
\begin{array}{cc|ccc}
0 & 1/2 & u & v & 1/2-u-v \\
1 & 1/2 & 1/3-u & 1/3-v & u+v-1/6 \\
\hline
 & & 1/3 & 1/3 & 1/3 \\
 & & 0 & 1 & 2
\end{array}
$$

Ограничения на параметры

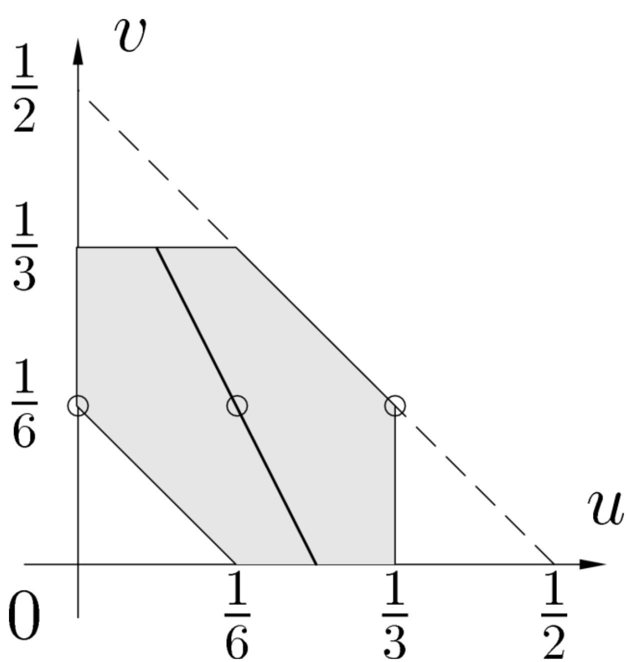
$$0 \le u \le 1/3, \;\; 0 \le v \le 1/3$$
$$1/6 \le u+v \le 1/2$$

	Антикомонотонное	**Независимое**	**Комонотонное**
u	0	1/6	1/3
v	1/6	1/6	1/6

Пример 1: все распределения

Область допустимых значений параметров в плоскости

$$(u, v)$$



Некоррелированные распределения

$$2u + v = \frac{1}{2}$$

Распределения с заданной корреляцией

$$\sqrt{6}\left(2u + v - \frac{1}{2}\right) = \rho$$

Пример 2: типовые распределения

Сетка совместного распределения

0	2/5	•	•	•
1	3/5	•	•	•
		1/3	1/2	1/6
		0	1	2

Независимое распределение

0	2/5	2/15	1/5	1/15
1	3/5	1/5	3/10	1/10
		1/3	1/2	1/6
		0	1	2

Комонотонное распределение

0	2/5	1/3	1/15	0
1	3/5	0	13/30	1/6
		1/3	1/2	1/6
		0	1	2

Антикомонотонное распределение

0	2/5	0	7/30	1/6
1	3/5	1/3	4/15	0
		1/3	1/2	1/6
		0	1	2

Пример 2: все распределения

Распределение общего вида при заданных маргинальных распределениях

$$
\begin{array}{cc|ccc}
0 & 2/5 & u & v & 2/5-u-v \\
1 & 3/5 & 1/3-u & 1/2-v & u+v-7/30 \\
\hline
 & & 1/3 & 1/2 & 1/6 \\
 & & 0 & 1 & 2
\end{array}
$$

Ограничения на параметры

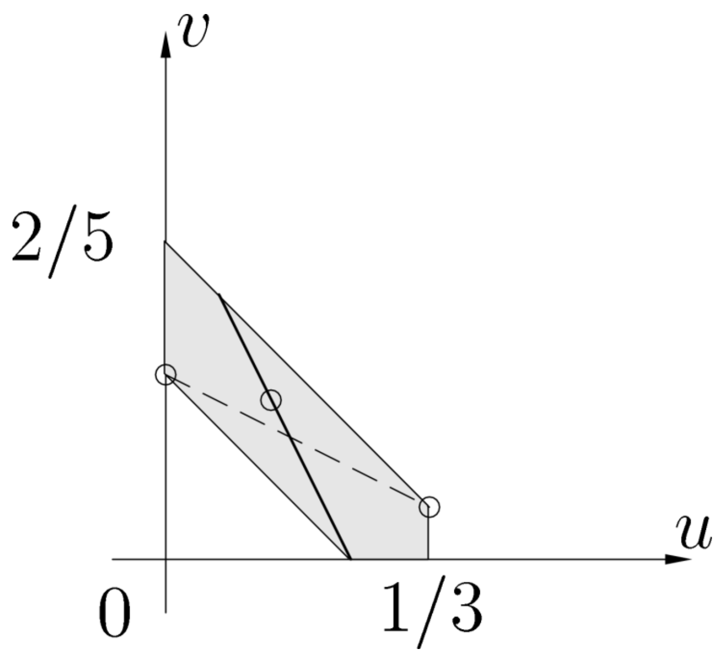
$$0 \le u \le 1/3, \;\; 0 \le v \le 1/2$$
$$7/30 \le u+v \le 2/5$$

	Антикомонотонное	**Независимое**	**Комонотонное**
u	0	2/15	1/3
v	7/30	1/5	1/15

Пример 2: все распределения

Область допустимых значений параметров в плоскости

$$(u, v)$$



Границы корреляции

$$-\frac{7}{\sqrt{102}} \leq \rho \leq \frac{8}{\sqrt{102}}$$

Некоррелированные распределения

$$2u + v = \frac{7}{15}$$

Распределения с заданной корреляцией

$$\frac{30}{\sqrt{102}}\left(2u + v - \frac{7}{15}\right) = \rho$$

Класс Фреше

- Маргинальные распределения p, q
- Класс Фреше $\mathcal{F}(p, q)$
- Комонотонное распределение

$$R^{+}(p, q) \in \mathcal{F}(p, q)$$

- Антикомонотонное распределение

$$R^{-}(p, q) \in \mathcal{F}(p, q)$$

- Независимое распределение

$$R^{\diamond}(p, q) \in \mathcal{F}(p, q)$$

Класс Фреше: структура

- Совокупность распределений с корреляцией ρ

$$\mathcal{F}_\rho(p,q) \subseteq \mathcal{F}(p,q)$$

- Границы коэффициента корреляции

$$\rho \in [\rho_{\min}, \rho_{\max}]$$

- Крайние распределения

$$R^-(p,q) \in \mathcal{F}_{\rho_{\min}}(p,q), \quad R^+(p,q) \in \mathcal{F}_{\rho_{\max}}(p,q)$$

- Объединение

$$\mathcal{F}(p,q) = \bigcup_{\rho \in [\rho_{\min}, \rho_{\max}]} \mathcal{F}_\rho(p,q)$$

Смеси

Задан коэффициент корреляции $\rho \in [\rho_{\min}, \rho_{\max}]$

- Произвольная корреляция

$$R_\rho^{-+}(p,q) = \lambda_\rho R^+(p,q) + (1-\lambda_\rho) R^-(p,q)$$

$$\lambda_\rho = \frac{\rho - \rho_{\min}}{\rho_{\max} - \rho_{\min}}$$

- В примере 2 для $\rho = 0$: $\lambda_0 = 7/15$

$$R_0^{-+}(p,q) : \quad \begin{array}{cc|ccc} 0 & 2/5 & 7/45 & 7/45 & 4/45 \\ 1 & 3/5 & 8/45 & 31/90 & 7/90 \\ \hline & & 1/3 & 1/2 & 1/6 \\ & & 0 & 1 & 2 \end{array}$$

Смеси +

- Положительная корреляция

$$R_\rho^+(p,q) = \lambda_\rho^+ R^+(p,q) + (1-\lambda_\rho^+)R^\diamond(p,q)$$

$$\lambda_\rho^+ = \frac{\rho}{\rho_{\max}}$$

- В примере 2 для $\rho = 4/\sqrt{102}$: $\lambda_\rho = 1/2$

$$R_\rho^+(p,q) : \quad \begin{array}{cc|ccc} 0 & 2/5 & 7/30 & 2/15 & 1/30 \\ 1 & 3/5 & 1/10 & 11/30 & 2/15 \\ \hline & & 1/3 & 1/2 & 1/6 \\ & & 0 & 1 & 2 \end{array}$$

Смеси -

- Отрицательная корреляция

$$R^-_\rho(p,q) = \lambda^-_\rho R^-(p,q) + (1-\lambda^-_\rho) R^\diamond(p,q)$$

$$\lambda^-_\rho = \frac{\rho}{\rho_{\min}}$$

- В примере 2 для $\rho = -4/\sqrt{102}$: $\qquad \lambda_\rho = 4/7$

$$R^-_\rho(p,q) : \quad \begin{array}{cc|ccc} 0 & 2/5 & 2/35 & 23/105 & 13/105 \\ 1 & 3/5 & 29/105 & 59/210 & 3/70 \\ \hline & & 1/3 & 1/2 & 1/6 \\ & & 0 & 1 & 2 \end{array}$$