Цель этой заметки — доказать теорему Иссерлиса о подсчёте ожиданий для многомерного нормального распределения. По дороге выведем функцию производящую моменты для одномерного нормального $\mathcal{N}(\mu,\sigma^2)$ и многомерного нормального $\mathcal{N}(\mu, C)$.

Ожидание экспоненты

Первая задача. Найдите $\mathbb{E}(\exp(Y))$ для нормальной $Y \sim \mathcal{N}(\mu;\sigma^2)$.

Сначала стандартизируем Y, $Y = \mu + \sigma X$, где $X \sim \mathcal{N}(0;1)$:

$$\mathbb{E}(\exp(Y)) = \mathbb{E}(\exp(\mu)\exp(\sigma X)) = \exp(\mu)\,\mathbb{E}(\exp(\sigma X)).$$

Перейдём к интегралам!

$$\mathbb{E}(\exp(\sigma X)) = \int_{\mathbb{R}} \exp(\sigma x) f(x) dx = \int_{\mathbb{R}} \exp(\sigma x)\frac{1}{\sqrt{2\pi}}\exp(-x^2/2)dx.$$

Для взятия интеграла выделим полный квадрат внутри экспоненты:

$$\sigma x - x^2/2 = -\frac{1}{2}(x^2 - 2\sigma x + \sigma^2 - \sigma^2) = -\frac{1}{2}(x-\sigma)^2 + \frac{1}{2}\sigma^2.$$

Возвращаемся к интегралу:

$$\mathbb{E}(\exp(\sigma X)) = \cdots = \int_{\mathbb{R}} \exp(\sigma^2/2)\frac{1}{\sqrt{2\pi}}\exp(-(x-\sigma)^2/2)dx = \exp(\sigma^2/2)\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\exp(-(x-\sigma)^2/2)dx.$$

Замечаем, что последний интеграл — это площадь под нормальной функцией плотности, смещённой на σ вправо. И эта площадь равна единице.

Задача раз решена,

$$\mathbb{E}(\exp(Y)) = \exp(\mu)\exp(\sigma^2/2) = \exp(\mu + \sigma^2/2).$$

Одномерная функция производящая моменты

Вторая задача. Найдите функцию производящую моменты для $Y \sim \mathcal{N}(\mu, \sigma^2)$.

По-определению, $\mathrm{mgf}_Y(u) = \mathbb{E}(\exp(uY))$.

Заметим, что $uY \sim \mathcal{N}(\mu \cdot u, \sigma^2 \cdot u^2)$.

Поэтому сразу получаем, что $\mathrm{mgf}_Y(u) = \exp(\mu u + \sigma^2 u^2/2)$.

Многомерная функция производящая моменты

Третья задача. Найдите функцию производящую моменты для случайного вектора $Y \sim \mathcal{N}(\mu, C)$.

По-определению, $\mathrm{mgf}_Y(u) = \mathbb{E}(\exp(u^T Y))$.

Заметим, что $u^T Y$ это скалярная случайная величина с нормальным распределением $\mathcal{N}(u^T\mu, u^T C u)$.

Снова быстро получаем функцию производящую моменты,

$$\mathrm{mgf}_Y(u) = \exp(u^T\mu + u^T C u/2).$$

Производящая функция выглядит как экспоненты от квадратичной функции,

$$\mathrm{mgf}_Y(u) = \exp(q(u)), \quad q(u) = u^T\mu + u^T C u/2.$$

Теорема Иссерлиса

Четвёртая задача. Для случайного вектора $Y \sim \mathcal{N}(\mu, C)$ последовательно найдите $\mathbb{E}(Y_1)$, $\mathbb{E}(Y_1Y_2)$, $\mathbb{E}(Y_1Y_2Y_3)$ и $\mathbb{E}(Y_1Y_2Y_3Y_4)$.

Вспомним, что $\mathbb{E}(Y_1) = \mathrm{mgf}'_1(0)$, $\mathbb{E}(Y_1Y_2) = \mathrm{mgf}''_{12}(0)$ и так далее.

Немного заранее подготовимся! Во-первых, в нуле $q(0) = 0$, $\exp(q(0)) = 1$.

Найдём первую производную $q'_1(u) = \mu_1 + c_1^T u$, где c_1 — первый столбец матрицы C. Для наглядности перепишем её в скалярном виде

$$q'_1(u_1, u_2, \ldots, u_n) = \mu_1 + c_{11}u_1 + c_{12}u_2 + \cdots + c_{1n}u_n.$$

В нуле первая производная равна $q'_1(0) = \mu_1$ соответствующему ожиданию.

Вторая производная $q''_{12}(u) = c_{12}$ тождественно равна соответствующей ковариации.

Третья производная $q'''_{123} = 0$ тождественно равна нулю, ведь $q(u)$ — квадратичная функция.

А теперь считаем ожидания по очереди,

$$\mathbb{E}(Y_1) = \mathrm{mgf}'_1 = \exp(q)q'_1 = 1 \cdot \mu_1 = \mu_1.$$

Пока что ничего неожиданного, мы же сами обозначили $\mathbb{E}(Y_1)$ как μ_1.

Пойдём дальше!

$$\mathbb{E}(Y_1Y_2) = \mathrm{mgf}''_{12} = \exp(q)(q'_1q'_2 + q''_{12}) = \mu_1\mu_2 + c_{12}.$$

Это тождество, верное для любый случайных величин, не только для нормальных, $\mathbb{E}(Y_1Y_2) = \mathbb{E}(Y_1)\,\mathbb{E}(Y_2) + \mathbb{C}\mathrm{ov}(Y_1, Y_2)$.

Продолжаем,

$$\mathbb{E}(Y_1Y_2Y_3) = \mathrm{mgf}'''_{123} = \exp(q)(q'_1q'_2q'_3 + q'_3c_{12} + q'_1c_{23} + q'_2c_{13}) = \mu_1\mu_2\mu_3 + \mu_1c_{23} + \mu_2c_{13} + \mu_3c_{12}.$$

Каждое слагаемое идёт с единичным весом. Каждое слагаемое содержит все индексы от единицы до максимального ровно по одному разу. Все варианты перемножения μ_i и c_{jk} присутствуют. Множитель μ_i «съедает» один индекс, а c_{jk} «съедает» два индекса.

Если какой-то индекс повторяется, то его надо повторить :)

$$\mathbb{E}(Y_1Y_2Y_2) = \mathrm{mgf}'''_{122} = \exp(q)(q'_1q'_2q'_2 + q'_2c_{12} + q'_1c_{22} + q'_2c_{12}) = \mu_1\mu_2\mu_2 + \mu_1c_{22} + \mu_2c_{12} + \mu_2c_{12}.$$

Фанаты могут убедиться, что

$$\mathbb{E}(Y_1Y_2Y_3Y_4) = \mu_1\mu_2\mu_3\mu_4 + \mu_1\mu_4c_{23} + \mu_2\mu_4c_{13} + \mu_3\mu_4c_{12} + \mu_1\mu_3c_{24} + \mu_1\mu_2c_{34} + \mu_2\mu_3c_{14} + c_{12}c_{34} + c_{13}c_{24} + c_{14}c_{23}.$$

Пример.

Вектор Y имеет совместное нормальное распределение,

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 5 & -1 & -2 \\ -1 & 6 & -3 \\ -2 & -3 & 7 \end{pmatrix} \right).$$

Найдите $\mathbb{E}(Y_1Y_2Y_3)$.

Решение:

$$\mathbb{E}(Y_1Y_2Y_3) = \mu_1\mu_2\mu_3 + \mu_1c_{23} + \mu_2c_{13} + \mu_3c_{12} = 1 \cdot 2 \cdot 3 + 1 \cdot (-3) + 2 \cdot (-2) + 3 \cdot (-1) = -2.$$