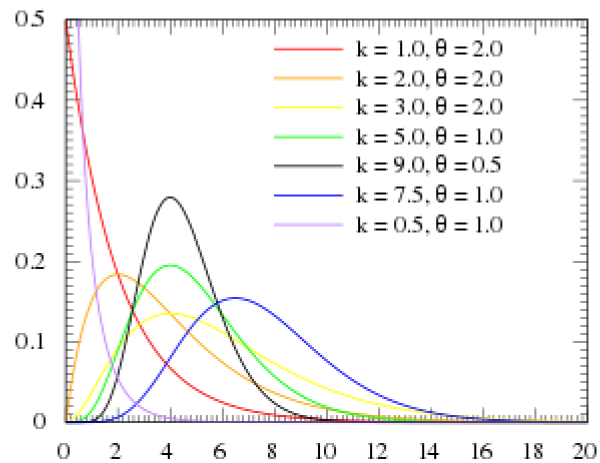
Гамма-распределение

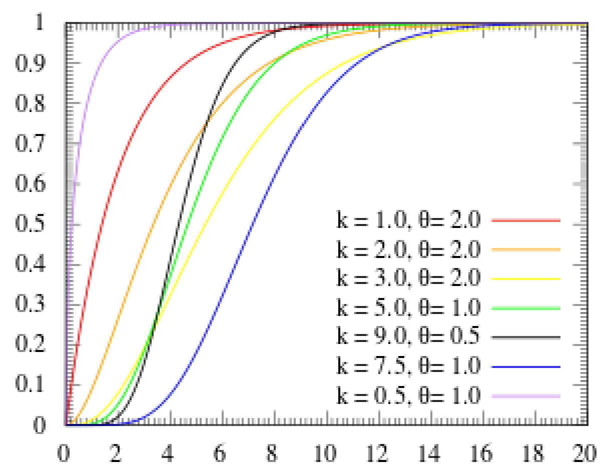
Га́мма-распределе́ние в теории вероятностей — это двухпараметрическое семейство абсолютно непрерывных распределений. Если параметр k принимает целое значение, то такое гамма-распределение также называется **распределе́нием Эрла́нга**.

Гамма распределение

Плотность вероятности



Функция распределения



Обозначение	$\Gamma(k,\theta)$ или $\Gamma(\alpha,\beta)$[1]	
Параметры	$k>0,\ \theta>0$	$\alpha>0,\ \beta>0$
Носитель	$x\in(0,\infty)$	$x\in(0,\infty)$
Плотность вероятности	$\frac{1}{\Gamma(k)\theta^k}x^{k-1}e^{-\frac{x}{\theta}}$	$\frac{\alpha^\beta x^{\beta-1}}{\Gamma(\beta)}e^{-\alpha x}$
Функция распределения	$\frac{1}{\Gamma(k)}\gamma\left(k,\frac{x}{\theta}\right)$	$\frac{1}{\Gamma(\beta)}\gamma(\beta,\alpha x)$
Математическое ожидание	$k\theta$	$\frac{\beta}{\alpha}$
Медиана	Отсутствует явное выражение в закрытой форме	Отсутствует явное выражение в закрытой форме

Содержание

Определение

Пусть распределение случайной величины X задаётся плотностью вероятности, имеющей вид

$$f_X(x)=\begin{cases} x^{k-1}\frac{e^{-x/\theta}}{\theta^k\,\Gamma(k)}, & x\geq 0\\ 0, & x<0\end{cases},$$

где $\Gamma(k)$ — гамма-функция Эйлера.

Тогда говорят, что случайная величина X имеет гамма-распределение с положительными параметрами θ и k. Пишут $X\sim\Gamma(k,\theta)$.

Замечание. Иногда используют другую параметризацию семейства гамма-распределений. Или вводят третий параметр — сдвиг.

Мода	$(k-1)\theta$ при $k \geq 1$	$\frac{\beta-1}{\alpha}$ при $\beta \geq 1$
Дисперсия	$k\theta^2$	$\frac{\beta}{\alpha^2}$
Коэффициент асимметрии	$\frac{2}{\sqrt{k}}$	$\frac{2}{\sqrt{\beta}}$
Коэффициент эксцесса	$\frac{6}{k}$	$\frac{6}{\beta}$
Информационная энтропия	$k + \ln\theta + \ln\Gamma(k) + (1-k)\psi(k)$	$\beta - \ln\alpha + \ln\Gamma(\beta) + (1-\beta)\psi(\beta)$
Производящая функция моментов	$(1-\theta t)^{-k}$ при $t < \frac{1}{\theta}$	$\left(1 - \frac{t}{\alpha}\right)^{-\beta}$ при $t < \alpha$
Характеристическая функция	$(1-\theta it)^{-k}$	$\left(1 - \frac{it}{\alpha}\right)^{-\beta}$

Моменты

Математическое ожидание и дисперсия случайной величины X, имеющей гамма-распределение, имеют вид

$$\mathbb{E}[X] = k\theta,$$
$$\mathbb{D}[X] = k\theta^2.$$

Свойства гамма-распределения

- Если $X_1, \ldots, X_n$ — независимые случайные величины, такие что $X_i \sim \Gamma(k_i, \theta),\ i = 1, \ldots, n$, то

$$Y = \sum_{i=1}^{n} X_i \sim \Gamma\left(\sum_{i=1}^{n} k_i, \theta\right).$$

- Если $X \sim \Gamma(k, \theta)$, и $a > 0$ — произвольная константа, то

$$aX \sim \Gamma(k, a\theta).$$

- Гамма-распределение бесконечно делимо.

Связь с другими распределениями

- Гамма-распределение является распределением Пирсона типа III[2].
- Экспоненциальное распределение является частным случаем гамма-распределения:

$$\Gamma(1, 1/\lambda) \equiv \mathrm{Exp}(\lambda).$$

- Если $X_1, \ldots, X_k$ — независимые экспоненциальные случайные величины, такие что $X_i \sim \mathrm{Exp}(\lambda),\ i = 1, \ldots, k$, то

$$Y = \sum_{i=1}^{k} X_i \sim \Gamma(k, 1/\lambda).$$

- Распределение хи-квадрат является частным случаем гамма-распределения:

$$\Gamma\left(\frac{n}{2}, 2\right) \equiv \chi^2(n).$$

- Согласно центральной предельной теореме, при больших k гамма-распределение может быть приближено нормальным распределением:

$$\Gamma(k,\theta) \approx \mathrm{N}(k\theta, k\theta^2) \text{ при } k \to \infty.$$

- Если X_1, X_2 — независимые случайные величины, такие что $X_i \sim \Gamma(k_i, 1),\ i = 1, 2$, то

$$\frac{X_1}{X_1 + X_2} \sim \mathrm{B}(k_1, k_2).$$

- Распределение Рэлея заменой переменной сводится к гамма-распределению.
- Обычное распределение Вейбулла заменой переменной сводится к гамма-распределению.
- Распределение Накагами заменой переменной сводится к гамма-распределению.
- Естественным обобщением гамма-распределения является усеченное гамма-распределение.

Моделирование гамма-величин

Учитывая свойство масштабирования по параметру θ, указанное выше, достаточно смоделировать гамма-величину для $\theta = 1$. Переход к другим значениям параметра осуществляется простым умножением.

Используя тот факт, что распределение $\Gamma(1, 1)$ совпадает с экспоненциальным распределением, получаем, что если U — случайная величина, равномерно распределённая на интервале (0, 1], то $-\ln U \sim \Gamma(1, 1)$.

Теперь, используя свойство k-суммирования, обобщим этот результат:

$$\sum_{i=1}^{n} -\ln U_i \sim \Gamma(n, 1),$$

где U_i — независимые случайные величины, равномерно распределённые на интервале (0, 1].

Осталось смоделировать гамма-величину для $0 < k < 1$ и ещё раз применить свойство k-суммирования. Это является самой сложной частью.

Ниже приведён алгоритм без доказательства. Он является примером выборки с отклонением.

1. Положить m равным 1.
2. Сгенерировать V_{2m-1} и V_{2m} — независимые случайные величины, равномерно распределённые на интервале (0, 1].
3. Если $V_{2m-1} \le v_0$, где $v_0 = \dfrac{e}{e+\delta}$, перейти к шагу 4, иначе к шагу 5.
4. Положить $\xi_m = \left(\dfrac{V_{2m-1}}{v_0}\right)^{\frac{1}{\delta}},\ \eta_m = V_{2m}\xi_m^{\delta-1}$. Перейти к шагу 6.
5. Положить $\xi_m = 1 - \ln \dfrac{V_{2m-1} - v_0}{1 - v_0},\ \eta_m = V_{2m} e^{-\xi_m}$.
6. Если $\eta_m > \xi_m^{\delta-1} e^{-\xi_m}$, то увеличить m на единицу и вернуться к шагу 2.
7. Принять $\xi = \xi_m$ за реализацию $\Gamma(\delta, 1)$.

Подытожим:

$$\theta\left(\xi - \sum_{i=1}^{[k]} \ln U_i\right) \sim \Gamma(k, \theta),$$

где $[k]$ является целой частью k, а ξ сгенерирована по алгоритму, приведённому выше при $\delta = \{k\}$ (дробная часть k); U_i и V_l распределены как указано выше и попарно независимы.

Примечания

1. Родионов, 2015, с. 29.
2. Королюк, 1985, с. 134.

Литература

- *Лагутин М.Б.* Наглядная математическая статистика. — М.: Бином, 2009. — 472 с.
- *Жуковский М.Е., Родионов И.В.* Основы теории вероятностей. — М.: МФТИ, 2015. — 82 с.
- *Жуковский М.Е., Родионов И.В., Шабанов Д.А.* Введение в математическую статистику. — М.: МФТИ, 2017. — 109 с.
- *Королюк В.С., Портенко Н.И., Скороход А.В., Турбин А.Ф.* Справочник по теории вероятностей и математической статистике. — М.: Наука, 1985. — 640 с.