УДК 543. 42; 517. 518.45

ГАУССОВА АСИМПТОТИКА ГАММА-РАСПРЕДЕЛЕНИЯ И ПОПРАВКИ К ФОРМУЛЕ СТИРЛИНГА

ЕХИЛЕВСКИЙ СТЕПАН ГРИГОРЬЕВИЧ,
д.т.н., профессор,
ГОЛУБЕВА ОКСАНА ВАЛЕРЬЕВНА,
к.ф.-м.н., доцент
ЗАБЕЛЕНДИК ОЛЬГА НИКОЛАЕВНА,
СТРУК ТАТЬЯНА СЕРГЕЕВНА
старшие преподаватели
УО «Полоцкий государственный университет»

Аннотация. В работе развит основанный на теории информации и методе статистических моментов подход к получению формулы Стирлинга для факториалов с большими аргументами. Обоснована процедура поиска поправок к асимптотическому выражению Стирлинга, обусловленных эксцессами гамма распределения случайной величины. С ее помощью получена приближенная формула для вычисления $n!$, относительная погрешность которой убывает с ростом аргумента, как $1/n^3$.
Ключевые слова: формула Стирлинга, энтропия, гамма распределение, эксцессы высших порядков.

GAUSSIAN ASYMPTOTICS OF THE GAMMA DISTRIBUTION AND AMENDMENTS TO THE STIRLING FORMULA

Ekhilevskiy Stepan Grigoryevich,
Golubeva Oksana Valeryevna,
Zabelendik Olga Nikolaevna,
Struk Tatyana Sergeevna

Annotation. The paper develops an approach based on information theory and the method of statistical moments to obtain the Stirling formula for factorials with large arguments. The procedure for searching for corrections to the Stirling asymptotic expression due to excesses in the gamma distribution of a random variable is justified. It is used to obtain an approximate formula for calculating $n!$, the relative error of which decreases with the growth of n, as $1/n^3$.
Key words: Stirling formula, entropy, gamma distribution, higher-orders excesses.

Введение. При компьютерном моделировании природных и технологических процессов (например, динамики сорбции [1]) возникает необходимость вычисления факториалов с большими аргументами. При этом возможности прикладных пакетов имеют ограничения по максимальным значениям фигурирующих в числовых расчетах величин [2]. Преодолеть эту трудность можно разбив $n!$ на группу множителей, каждый из которых участвует в расчетах автономно. В частности, сделать это можно с помощью асимптотической формулы Стирлинга, поправки к которой необходимо учитывать для обеспечения требуемой точности вычислений. Обоснованию процедуры их получения, основанной на методе статистических моментов, посвящена данная публикация.

Получение асимптотического выражения. С помощью n-кратного интегрирования по частям и правила Лопиталя можно убедится в справедливости цепочки равенств

$$\int_0^\infty x^n e^{-x} dx = n\int_0^\infty x^{n-1} e^{-x} dx = \ldots = n! \qquad (n = 0, 1, 2, \ldots) \tag{1}$$

Разделив левую и правую части (1) на $n!$ получим равенство

$$\int_0^\infty \frac{1}{n!} x^n e^{-x} dx = 1, \tag{2}$$

которое можно интерпретировать как условие нормировки для плотности вероятности

$$f(x) = \frac{1}{n!} x^n e^{-x} \tag{3}$$

некоторой неотрицательной случайной величины X [3]. С помощью (1), (3) найдем ее математическое ожидание

$$M(X) = \int_0^\infty x f(x) dx = \frac{1}{n!}\int_0^\infty x^{n+1} e^{-x} dx = \frac{(n+1)!}{n!} = n+1, \tag{4}$$

дисперсию и среднеквадратическое отклонение:

$$M(X^2) = \int_0^\infty x^2 f(x) dx = \frac{1}{n!}\int_0^\infty x^{n+2} e^{-x} dx = \frac{(n+2)!}{n!} = (n+1)(n+2),$$

$$D(X) = M(X^2) - M(X)^2 = (n+1)(n+2) - (n+1)^2 = n+1, \tag{5}$$

$$\sigma(X) = \sqrt{D(X)} = \sqrt{n+1}\,. \tag{6}$$

То, что $M(X)$ и $D(X)$ аддитивно зависят от n, с учетом свойств математического ожидания и дисперсии, означает, что X представляет собой сумму $n+1$ независимых случайных слагаемых X_i с единичными математическими ожиданиями и среднеквадратическими отклонениями

$$M(X_i) = \sigma(X_i) = 1. \tag{7}$$

Свойством (7) обладают случайные величины, распределенные по экспоненциальному закону

$$f_i(x) = e^{-x}, \qquad (i = 1, 2, \ldots, n+1)$$

что согласуется с формулой (3) при $n = 0$. Т.е., полагая в (3) – (6)

$$m = n+1, \tag{8}$$

получим плотность вероятности суммы m независимых экспоненциально распределенных случайных слагаемых

$$f(m, x) = \frac{1}{(m-1)!} x^{m-1} e^{-x}, \qquad (m = 1, 2, \ldots) \tag{9}$$

известную как гамма распределение с m степенями свободы [3].

Согласно (4), (6) при больших m слева от матожидания в области возможных значений X помещается сколь угодно много среднеквадратических отклонений

$$\left(M(X)-0\right)/\sigma(X)=\sqrt{m}\xrightarrow[m\to\infty]{}\infty$$

то есть, в соответствии с неравенством Чебышева [3], область возможных значений X из полубесконечной $\left(x\in[0,\infty)\right)$ при $m\to\infty$ превращается как бы в бесконечную. На ней максимум энтропии обеспечивается нормальным распределением [3, 4]

$$f(m,x)\xrightarrow[m\to\infty]{}\frac{1}{\sqrt{2\pi m}}e^{-\frac{(x-m))^2}{2m}}. \qquad (10)$$

В справедливости последних рассуждений можно убедиться и непосредственно, вычислив с помощью (1), (9) асимметрии и эксцессы функции распределения. Для этого потребуются начальные

$$\nu_k(m)=\int_0^\infty x^k f(m,x)dx=$$

$$=\frac{1}{(m-1)!}\int_0^\infty x^{m-1+k}e^{-x}dx=\frac{(m-1+k)!}{(m-1)!}=m(m+1)...(m-1+k) \qquad (11)$$

и центральные моменты высших порядков

$$\mu_k(m)=\int_0^\infty \left(x-\nu_1(m)\right)^k f(m,x)dx=$$

$$=\sum_{i=0}^{k-2}(-1)^i C_k^i\cdot\nu_{k-i}(m)\cdot\nu_1^i(m)+(-1)^k(1-k)\nu_1^k(m), \qquad (12)$$

где C_k^i – числа сочетаний из k объектов по i.

В частности из (6), (10), (11) и (12) следует, что

$$\mu_0(m)=1,\qquad \mu_1(m)=0,\qquad \mu_2(m)=1, \qquad (13)$$

как и при любой плотности вероятности, а все асимметрии и эксцессы

$$A_{2k+1}(m)=\frac{\mu_{2k+1}(m)}{m^{(2k+1)/2}},\qquad E_{2k+2}(m)=\frac{\mu_{2k+2}(m)}{m^{k+1}}-1\cdot3\cdot...\cdot(2k+1)\qquad (k=1,2,...)$$

являются бесконечно малыми при $m\to\infty$:

$$A_3(m)=\frac{2}{m^{1/2}},\qquad A_5(m)=\frac{20}{m^{1/2}}+\frac{24}{m^{3/2}},\qquad A_7(m)=\frac{210}{m^{1/2}}+\frac{924}{m^{3/2}}+\frac{720}{m^{5/2}},...;$$

$$E_4(m)=\frac{6}{m},\qquad E_6(m)=\frac{130}{m}+\frac{120}{m^2},\qquad E_8(m)=\frac{2380}{m}+\frac{7308}{m^2}+\frac{5040}{m^3}, \qquad (14)$$

что соответствует нормальной асимптотике (10).

Подставив в (10) любое x из интервала, обеспечивающего для нормального закона практически весь вклад в энтропию (например $x=m$) получим с учетом (9)

$$\frac{1}{(m-1)!}m^{m-1}e^{-m}\xrightarrow[m\to\infty]{}\frac{1}{\sqrt{2\pi m}}. \qquad (15)$$

Откуда, пренебрегая единицами по сравнению с m, получим асимптотическое выражение для факториалов с большими аргументами

$$m!\xrightarrow[n\to\infty]{}\sqrt{2\pi m}\,m^m e^{-m}, \qquad (16)$$

известное, как формула Стирлинга [5].

Получение поправок. Уточнить правую часть (16), можно получив поправки к асимптотике (15), обусловленные эксцессами гамма распределения (асимметрии вклада не дадут, ибо их учет добавляет в (10) слагаемые, нечетные относительно математического ожидания, т.е. обращающиеся в ноль при $x = m$)

$$\frac{\mu_{2k}(\tau)}{m^k} \equiv \int_{-\sqrt{m}}^{\infty} \xi^{2k} \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} \left(1 + \sum_{i=1}^{\infty} \frac{\phi_i(\xi)}{m^i}\right) d\xi \,, \tag{17}$$

где

$$\xi(m, x) = (x - m)\big/\sqrt{m} \,, \tag{18}$$

а $\phi_i(\xi)$ – некоторые функции, раскладываемые по четным степеням ξ

$$\phi_i(\xi) = \sum_{l=0}^{3i} c_{il} \xi^{2l} \tag{19}$$

и подлежащие определению с помощью (14), при этом коэффициенты с более высокими степенями ξ оказываются равными нулю.

Тождество (17) должно выполняться при любых, в том числе и бесконечно больших m. При этом важно, что $\int_{-\infty}^{-\sqrt{m}} \xi^{2n} \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} d\xi$ является бесконечно малой более высокого порядка, чем $1\big/m^i$, в чем легко убедиться с помощью правила Лопиталя

$$\lim_{m\to\infty} \frac{\int_{-\infty}^{-\sqrt{m}} \xi^{2n} \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} d\xi}{1\big/m^i} = \lim_{m\to\infty} \frac{m^n e^{-\frac{m}{2}}}{-i\sqrt{2\pi}\big/m^{i+1}} = \frac{1}{-i\sqrt{2\pi}} \lim_{m\to\infty} \frac{m^{n+i+1}}{e^{m/2}} = 0$$

С учетом этого обстоятельства выполним интегрирование в (17)

$$\frac{\mu_{2k}(\tau)}{m^k} \equiv I(k) + \sum_{i=1}^{\infty} \frac{1}{m^i} \sum_{l=0}^{3i} I(k+l) c_{il} \,, \tag{20}$$

где

$$I(n) = \int_{-\infty}^{\infty} x^{2n} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \begin{cases} 1 \cdot 3 \cdot \ldots \cdot (2n-1), & n = 1, 2, \ldots \\ 1, & n = 0 \end{cases}$$

В частности, полагая в (20) $k = 0, 1, 2, 3$ и приравнивая коэффициенты при $1\big/m$ в левой и правой части полученных равенств, запишем с учетом (13), (14), (19) систему линейных алгебраических уравнений относительно коэффициентов c_{1l}. Ее решение, записанное в матричной форме, имеет вид

$$\hat{c}_1 = \hat{a}_3^{-1} \begin{pmatrix} 0 \\ 0 \\ 6 \\ 130 \end{pmatrix} = \begin{pmatrix} -1/12 \\ 1 \\ -7/12 \\ 1/18 \end{pmatrix}, \tag{21}$$

где

$$\hat{c}_i = \begin{pmatrix} c_{i0} \\ c_{i1} \\ \dots \\ c_{i\,3i} \end{pmatrix}, \qquad \hat{a}_n = \begin{pmatrix} I(0) & I(1) & \dots & I(n) \\ I(1) & I(2) & \dots & I(n+1) \\ \dots & \dots & \dots & \dots \\ I(n) & I(n+1) & \dots & I(2n) \end{pmatrix}. \tag{22}$$

Если попытаться в (19) увеличить старшую степень $\phi_1(\xi)$ на две единицы, в (20) также следует положить $k=4$. В результате, с учетом разложения (14) для $E_8(m)$, вместо (21) получим

$$\begin{pmatrix} c_{10} \\ c_{11} \\ c_{12} \\ c_{13} \\ c_{14} \end{pmatrix} = \hat{a}_4^{-1} \begin{pmatrix} 0 \\ 0 \\ 6 \\ 130 \\ 2380 \end{pmatrix} = \begin{pmatrix} -1/12 \\ 1 \\ -7/12 \\ 1/18 \\ 0 \end{pmatrix},$$

что подтверждает правильность верхнего предела суммирования в (19).

С точки зрения решаемой задачи (уточнение формулы Стирлинга) в (21) существенно лишь найденное $c_{10} = -1/12$, как значение $\phi_1(\xi)$ при $x = m$ (см. (18), (19)). С учетом этого обстоятельства, а также формул (17) – (19), (21) вместо (15) получим

$$\frac{1}{(m-1)!} m^{m-1} e^{-m} \xrightarrow[m\to\infty]{} \frac{1}{\sqrt{2\pi m}} \left(1 - \frac{1}{12m}\right),$$

что не противоречит имеющимся оценкам погрешности формулы Стирлинга [5]

$$\frac{(m-1)! - \sqrt{2\pi m}\, m^{m-1} e^{-m}}{(m-1)!} \xrightarrow[m\to\infty]{} \frac{1}{12m}.$$

Согласно (19), чтобы получить поправку порядка $1/m^2$ кроме имеющихся в (14) потребуются эксцессы

$$E_{10}(m) = \frac{44100}{m} + \frac{303660}{m^2} + \frac{623376}{m^3} + \frac{362880}{m^4}, \tag{23}$$

$$E_{12}(m) = \frac{866250}{m} + \frac{11098780}{m^2} + \frac{47324376}{m^3} + \frac{76998240}{m^4} + \frac{39916800}{m^5}. \tag{24}$$

Последовательно полагая в (20) $k = 0, 1, 2, 3, 4, 5, 6$ и приравнивая в получаемых таким образом уравнениях коэффициенты при $1/m^2$ с помощью (14), (23), (24) найдем

$$c_{20} = \Delta_1^{(6)} / \Delta^{(6)} = 1/288$$

где $\Delta^{(6)}$ – определитель матрицы $\hat{a}_6$, а $\Delta_1^{(6)}$ получается из него заменой элементов первого столбца свободными членами соответствующей системы линейных алгебраических уравнений 0, 0, 0, 120, 7308, 303660, 11098780.

Найденное c_{20} позволяет уточнить формулу для приближенного вычисления факториалов

$$(m-1)! \xrightarrow[m\to\infty]{} \frac{\sqrt{2\pi m}\, m^{m-1} e^{-m}}{1 - \dfrac{1}{12m} + \dfrac{1}{288m^2}}, \tag{25}$$

снизив оценку относительной погрешности до

$$\frac{\sqrt{2\pi m}\, m^{m-1} e^{-m}}{(m-1)!\left(1-\frac{1}{12m}+\frac{1}{288m^2}\right)} - 1 \xrightarrow[m\to\infty]{} \frac{c_{30}}{m^3},$$

где c_{30} – константа в разложении $\phi_3(\xi)$ по обратным степеням m (см. (19)). Ее значение, найденное с помощью описанной выше процедуры уже для $m=2$ $(n=1)$ позволяет обеспечить погрешность менее 0,1%. Фрагмент программы с результатами соответствующих расчетов представлен на рис.1.

$$h(m) := -1 + \frac{\sqrt{2\pi m}}{(m-1)!} \cdot m^{m-1} \cdot \frac{e^{-m}}{1 - \frac{1}{12m} + \frac{1}{288m^2}}$$

$m := 2..8$

$t(m) := (h(m)+1)\cdot(m-1)!$ $\quad t(2) = 1.00031$ $\quad t(3) = 2.00019$

$t(4) = 6.0002467877$

$$g(m) := -1 + \frac{\sqrt{2\pi m}}{(m-1)!} \cdot m^{m-1} \cdot \frac{e^{-m}}{1 - \frac{1}{12m}}$$

$t(5) = 24.000508665$

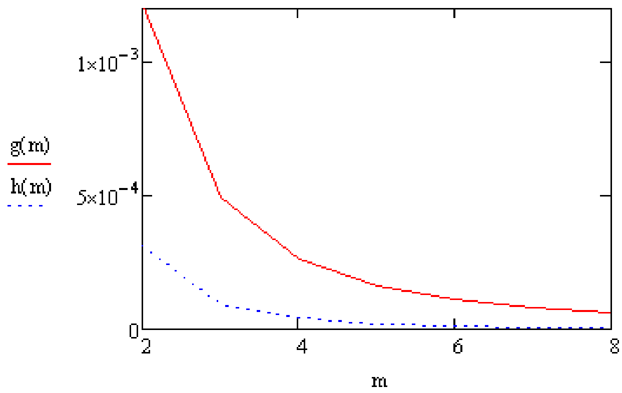

Рис. 1. Относительная погрешность формулы Стирлинга, исправленной с учетом эксцессов 6-го (сплошная кривая) и 12-го (пунктир) порядков гамма распределения

Выводы. В работе развит основанный на теории информации и методе статистических моментов подход к получению формулы Стирлинга для факториалов с большими аргументами. Обоснована процедура поиска поправок к формуле Стирлинга, обусловленных эксцессами гамма распределения случайной величины. С ее помощью получена приближенная формула для вычисления $n!$, относительная погрешность которой убывает с ростом аргумента, как $1/n^3$.

Список литературы

1. Ехилевский, С.Г. Нестационарная задача динамики сорбции углекислого газа в регенеративном патроне изолирующего респиратора / С.Г. Ехилевский // Вестн. Фонда фундам. исследований. – 2019. – № 3(89) – С. 57–65.
2. Аладьев, В.З. Вычислительные задачи на персональном компьютере / В.З. Аладьев, Н.А. Гершгорн. – Киев.: Техника,1991. – 246 с.
3. Гнеденко, Б.В. Курс теории вероятностей / Б.В. Гнеденко. – М. : Наука, 1969. – 400 с.
4. Ехилевский, С.Г. Экстремальность энтропии, формула Стирлинга и закон больших чисел / С.Г. Ехилевский, Т.С. Рудькова, О.В. Голубева // Прикладная математика и информатика: современные исследования в области естественных и технических наук : материалы IV науч.-практ. междунар. конф. (школы-семинара) молодых ученых, Тольятти, 23–25 апр. 2018 г. / Тольят. гос. ун-т. – Тольятти, 2018. – Ч. 1. – С. 461–464.
5. Корн, Г. Справочник по математике для научных работников и инженеров / Г. Корн, Т. Корн. – М.: Наука, 1968.