

Санкт-Петербургский государственный университет  
Прикладная математика и информатика

Отчет по научно-исследовательской работе

Замена непрерывных распределений на дискретные для  
применения на практике

(семестр 7)

Выполнила:  
Нагуманова Карина Ильнуровна,  
группа 19.Б04-мм

Научный руководитель:  
к.ф.-м.н., доцент  
Голяндина Нина Эдуардовна.  
Кафедра статистического моделирования

Санкт-Петербург

2023

# Содержание

<b>1</b>	<b>Введение</b>	<b>3</b>
<b>2</b>	<b>Условия аппроксимации в общем случае</b>	<b>4</b>
<b>3</b>	<b>Аппроксимация нормального распределения</b>	<b>5</b>
<b>4</b>	<b>Аппроксимация логнормального распределения</b>	<b>6</b>
4.1	Способ нахождения вероятностей через математическое ожидание и дисперсию нормального распределения . . . . .	6
4.2	Непосредственная аппроксимация логнормального распределения . . . . .	7
4.3	Условие на параметры для нахождения весов при аппроксимации логнормального распределения . . . . .	10
4.4	Варианты постановки задачи . . . . .	12
4.5	Точность аппроксимации . . . . .	13
4.5.1	Неправильное использование правила 30-40-30 . . . . .	13
<b>5</b>	<b>Произведение двух логнормальных распределений</b>	<b>16</b>
5.1	Квантили вида $\pi$ , 0.5, $1 - \pi$ произведения логнормальных случайных величин . . . . .	18
<b>6</b>	<b>Сумма двух логнормальных распределений</b>	<b>21</b>
<b>7</b>	<b>Заключение</b>	<b>22</b>
	<b>Список литературы</b>	<b>22</b>
<b>8</b>	<b>Приложение</b>	<b>23</b>

# 1 Введение

В практических задачах часто требуется заменить непрерывное распределение на дискретное с сохранением математического ожидания и дисперсии. Одним из методов нахождения такого распределения для аппроксимации нормального распределения является метод Свонсона. Однако в ряде областей, например, в нефтяной промышленности распределением, описывающим запасы нефти, общепринятым является логнормальное распределение. Соответственно, реальной задачей является аппроксимация логнормального распределения.

С аппроксимируемыми случайными величинами производят сложение и умножение. Например, используем площадь дренирования пласта, среднюю чистую толщину и коэффициент извлечения углеводородов. При перемножении этих параметров получаем количество резервов нефти. Или зная запасы, параметры нефти и породы для всех залежей можно оценить профиль добычи нефти с каждой залежи и суммарный профиль, оценить экономическую эффективность проекта, которая учитывает выручку, налоги, капитальные затраты, операционные затраты, оптимальные решения по проекту. Соответственно, возникает задача находить аппроксимацию суммы и произведения по аппроксимациям исходных случайных величин.

Часто бывает на практике, что вместо настоящего распределения известны три его квантили, стандартно это 10-, 50- и 90-процентили. Задачей является нахождение по ним математического ожидания и дисперсии. Обычно задача решается построением весов для квантилей так, чтобы у полученного дискретного распределения были такие же математическое ожидание и дисперсия, как у исходного. Вообще говоря, иногда нужно, чтобы и более старшие моменты также аппроксимировались моментами построенного дискретного распределения с целью, чтобы для функций от распределений равенство математических ожиданий и дисперсий оставалось хотя бы приближенными.

Структура работы следующая:

В разделе 2 рассмотрен общий подход к трехточечной аппроксимации.

В разделе 3 аппроксимация нормального распределения, вывод правила 30-40-30.

В разделе 4 рассматривается аппроксимация логнормального распределения, условие аппроксимации и что делать, если это условие не выполняется. А также точность аппроксимации при применении правила 30-40-30 к логнормальному распределению.

В разделе 5 алгоритм аппроксимации произведения двух логнормальных распределений.

В разделе 6 алгоритм аппроксимации суммы двух логнормальных распределений.

Работа этого семестра заключена в разделах 4.3, 4.4, 4.5, 6. В моей работе использовались статьи «Swanson's Swansong» [1] и «Uncertainties impacting reserves, revenue, and costs» [2].

Кроме этого были прочитаны следующие статьи:

«Discretization, Simulation, and Swanson's (Inaccurate) Mean» [3]. В ней одна из частей исследования – сравнение различных методов дискретизации непрерывных распределений, например таких, как Extended Person-Tukey (EPT), McNamee-Celona Shortcut (MCS), Extended Swanson-Megill (ESM).

Статья «Discretization, Simulation, and the Value of Information» [4]. Из нее понятно, что данный метод дискретизации значительно недооценивает среднее значение, дисперсию и асимметрию большинства распределений, особенно логнормального, где он широко используется. И что наилучшая дискретизация зависит от контекста решения, который мы не знаем заранее.

А в статье «Performance Evaluation of Swanson’s Rule for the Case of Log-Normal Populations» [5] проводится исследование оценки эффективности метода Свонсона и сравнение с использованием равных весов. Рассмотрены различные преимущества двух методов.

## 2 Условия аппроксимации в общем случае

Пусть дана случайная величина  $\xi$  с математическим ожиданием  $m$ , дисперсией  $s^2$  и функцией распределения  $F(x)$ . Для неё заданы квантили  $x_{\pi_1}$ ,  $x_{\pi_2}$ ,  $x_{\pi_3}$ . Также есть случайная дискретная величина  $\xi_n$  с математическим ожиданием  $m_n$  и дисперсией  $s_n^2$ .

$$\xi_n : \begin{pmatrix} x_{\pi_1} & x_{\pi_2} & x_{\pi_3} \\ p_1 & p_2 & p_3 \end{pmatrix}$$

Мы хотим аппроксимировать распределение случайной величины  $\xi$  дискретным распределением  $\xi_n$ .

Нужно найти вероятности  $p_1$ ,  $p_2$ ,  $p_3$  так, чтобы следующие равенства были верными.

$$p_1 + p_2 + p_3 = 1, \quad (1)$$

$$p_1 x_{\pi_1} + p_2 x_{\pi_2} + p_3 x_{\pi_3} = m, \quad (2)$$

$$p_1 x_{\pi_1}^2 + p_2 x_{\pi_2}^2 + p_3 x_{\pi_3}^2 - m^2 = s^2. \quad (3)$$

Запишем уравнения (1)–(3) в матричной форме следующим образом

$$\begin{pmatrix} 1 & 1 & 1 \\ x_{\pi_1} & x_{\pi_2} & x_{\pi_3} \\ x_{\pi_1}^2 & x_{\pi_2}^2 & x_{\pi_3}^2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 1 \\ m \\ m^2 + s^2 \end{pmatrix}.$$

Теперь введём более изящную форму, которая подчёркивает связь вероятностей с формой распределения путём нормализации математического ожидания и дисперсии.

**Предложение 1.** Пусть верно

$$\begin{pmatrix} 1 & 1 & 1 \\ \tilde{x}_{\pi_1} & \tilde{x}_{\pi_2} & \tilde{x}_{\pi_3} \\ \tilde{x}_{\pi_1}^2 & \tilde{x}_{\pi_2}^2 & \tilde{x}_{\pi_3}^2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad (4)$$

где  $\tilde{x}_{\pi_i} = \tilde{F}^{-1}(\pi_i)$ ,  $\tilde{F}(y)$  — функция распределения  $\eta = \frac{\xi - m}{s}$ . Тогда  $m = m_n$  и  $s^2 = s_n^2$ .

**Доказательство.**

$$\begin{aligned} P(\xi \leq x_{\pi_i}) &= \pi_i, \\ P\left(\frac{\xi - m}{s} \leq \frac{x_{\pi_i} - m}{s}\right) &= \tilde{F}\left(\frac{x_{\pi_i} - m}{s}\right) = \pi_i, \end{aligned}$$

$\xi$  нормализуется так, чтобы иметь нулевое математическое ожидание и единичную дисперсию. Имеем  $x_{\pi_i} = m + s\tilde{F}^{-1}(\pi_i)$ , обозначим

$$\tilde{x}_{\pi_i} = \frac{x_{\pi_i} - m}{s} = \tilde{F}^{-1}(\pi_i). \quad (5)$$

Предположим, что  $m = m_n$  и  $s^2 = s_n^2$ , и получим систему (4). Для этого подставим (5) в уравнение (2), получаем

$$m(p_1 + p_2 + p_3) + s(p_1\tilde{x}_{\pi_1} + p_2\tilde{x}_{\pi_2} + p_3\tilde{x}_{\pi_3}) = m.$$

Используя уравнение (1), получаем

$$s(p_1\tilde{x}_{\pi_1} + p_2\tilde{x}_{\pi_2} + p_3\tilde{x}_{\pi_3}) = 0.$$

Так как  $s \neq 0$ , то можно разделить на  $s$ , тогда получаем

$$p_1\tilde{x}_{\pi_1} + p_2\tilde{x}_{\pi_2} + p_3\tilde{x}_{\pi_3} = 0.$$

Теперь подставим (5) в уравнение (3), получаем

$$p_1(m + s\tilde{x}_{\pi_1})^2 + p_2(m + s\tilde{x}_{\pi_2})^2 + p_3(m + s\tilde{x}_{\pi_3})^2 - m^2 = s^2,$$

$$p_1\tilde{x}_{\pi_1}^2 + p_2\tilde{x}_{\pi_2}^2 + p_3\tilde{x}_{\pi_3}^2 = 1.$$

Получившиеся уравнения в матричной форме

$$\begin{pmatrix} 1 & 1 & 1 \\ \tilde{x}_{\pi_1} & \tilde{x}_{\pi_2} & \tilde{x}_{\pi_3} \\ \tilde{x}_{\pi_1}^2 & \tilde{x}_{\pi_2}^2 & \tilde{x}_{\pi_3}^2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}. \quad (6)$$

■

### 3 Аппроксимация нормального распределения

В общем случае вероятности  $p_1, p_2, p_3$  будут зависеть от математического ожидания и дисперсии, но если  $\xi \sim N(\mu, \sigma)$  имеет нормальное распределение, то  $\eta = \frac{\xi - m}{s}$  имеет нормальное стандартное распределение, которое не зависит ни от  $\mu$ , ни от  $\sigma$ .<sup>s</sup>

**Предложение 2.**  $\xi \sim N(\mu, \sigma)$ , пусть верно

$$\begin{cases} p_\pi = \frac{\delta}{2}, \\ p_{0.5} = 1 - \delta, \\ p_{1-\pi} = \frac{\delta}{2}. \end{cases} \quad (7)$$

где  $\delta = \frac{1}{\Phi^{-1}(\pi)^2}$ . Тогда  $m = m_n$  и  $s^2 = s_n^2$ .

*Доказательство.* Предположим, что  $m = m_n$  и  $s^2 = s_n^2$ , и получим систему (7).

$\Phi(y) = P(\eta = \frac{\xi - m}{s} \leq y)$  — функция распределения стандартного нормального распределения, тогда система (6) записывается как

$$\begin{pmatrix} 1 & 1 & 1 \\ \Phi^{-1}(\pi_1) & \Phi^{-1}(\pi_2) & \Phi^{-1}(\pi_3) \\ \Phi^{-1}(\pi_1)^2 & \Phi^{-1}(\pi_2)^2 & \Phi^{-1}(\pi_3)^2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}. \quad (8)$$

В частном случае симметричных квантилей вида  $\pi$ ,  $0.5$ ,  $1 - \pi$  получаем  $\Phi^{-1}(\pi) = -\Phi^{-1}(1 - \pi)$ ,  $\Phi^{-1}(0.5) = 0$ , тогда система (8) упрощается до

$$\begin{pmatrix} 1 & 1 & 1 \\ \Phi^{-1}(\pi) & 0 & -\Phi^{-1}(\pi) \\ \Phi^{-1}(\pi)^2 & 0 & \Phi^{-1}(\pi)^2 \end{pmatrix} \begin{pmatrix} p_\pi \\ p_{0.5} \\ p_{1-\pi} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

$$\begin{cases} p_\pi + p_{0.5} + p_{1-\pi} = 1, \\ (p_\pi - p_{1-\pi})\Phi^{-1}(\pi) = 0, \\ (p_\pi + p_{1-\pi})\Phi^{-1}(\pi)^2 = 1. \end{cases} \quad (9)$$

Обозначим  $\delta = \frac{1}{\Phi^{-1}(\pi)^2}$ , тогда из системы (9) получим систему (7). □

Рассмотрим случай  $\pi = 0.1$ , имеем  $\Phi^{-1}(0.1) = -\Phi^{-1}(0.9) \approx -1.28$ ,  $\Phi^{-1}(0.5) = 0$ , из уравнений системы (9) находим значения  $p_1$ ,  $p_2$ ,  $p_3$ .

$$\begin{cases} p_1 \approx 0.305, \\ p_2 \approx 0.390, \\ p_3 \approx 0.305. \end{cases}$$

Эти вероятности примерно равны  $0.3$ ,  $0.4$ ,  $0.3$ , поэтому это правило называют правилом 30-40-30 или **правилом Свонсона**.

## 4 Аппроксимация логнормального распределения

Пусть случайная величина  $\eta$  имеет логнормальное распределение, тогда случайная величина  $\xi = \ln(\eta)$  имеет нормальное распределение,  $\xi \sim N(\mu, \sigma)$ . И поэтому для нее можно использовать формулы, полученные в предыдущих разделах.

### 4.1 Способ нахождения вероятностей через математическое ожидание и дисперсию нормального распределения

Заметим, что если  $x_{\pi_1}$ ,  $x_{\pi_2}$ ,  $x_{\pi_3}$  — квантили логнормального распределения, то  $\ln(x_{\pi_1})$ ,  $\ln(x_{\pi_2})$ ,  $\ln(x_{\pi_3})$  — квантили нормального распределения. Можно взять эти квантили и использовать в способе нахождения вероятностей для нормального распределения.

Имеем следующий алгоритм.

**Алгоритм 1. Дано:** квантили  $x_{\pi_1}, x_{\pi_2}, x_{\pi_3}$  логнормальной случайной величины  $\eta$ ,  $\ln(\eta) \sim N(\mu, \sigma)$ .

**Шаги:**

1. Вычисляем значения мат. ожидания  $m$  и дисперсии  $d$  случайной величины  $\eta$ , используя известные  $x_{\pi_1}, x_{\pi_2}, x_{\pi_3}$ .
2. Выражаем параметры  $\mu$  и  $\sigma$  мат. ожидание и дисперсию соответствующего нормального распределения через параметры  $m$  и  $d$  логнормального распределения, используя следующие формулы

$$m = \exp\left(\mu + \frac{\sigma^2}{2}\right), \quad (10)$$

$$s^2 = m^2[\exp(\sigma^2) - 1]. \quad (11)$$

Заметим, что математическое ожидание логнормально распределенной случайной величины всегда положительное.

3. С помощью системы (7) находим значения вероятностей  $p_1, p_2, p_3$ .

**Результат:** вероятности  $p_1, p_2, p_3$  для  $x_{\pi_1}, x_{\pi_2}, x_{\pi_3}$  случайной величины  $\xi_n$ .

**Пример.** Пусть у нас есть логнормальная случайная величина с  $m = 2$ ,  $s^2 = 0.78125$ . Значения квантилей  $x_{10} = 1$ ,  $x_{50} = 2$ ,  $x_{90} = 3$ .

По данным формулам можно найти параметры соответствующего нормального распределения.

$$\mu = 0.69314,$$

$$\sigma = 0.42863.$$

Теперь можно найти значения  $p_1, p_2, p_3$ .

$$p_{10} = 0.371243,$$

$$p_{50} = 0.282992,$$

$$p_{90} = 0.345764.$$

## 4.2 Непосредственная аппроксимация логнормального распределения

Есть другой способ нахождения результата, полученного в разделе 4.1. Можно не переходить к нормальному распределению, а сразу вычислять вероятности для квантилей логнормального распределения.

Сначала найдём  $\tilde{F}(y)$  в терминах параметров распределения, затем найдём  $\tilde{F}^{-1}(p)$ , чтобы использовать формулу (4).

**Предложение 3.** В терминах Предложения 1 функция  $\tilde{F}^{-1}(\pi)$  выражается через  $\sigma$  как

$$\tilde{F}^{-1}(\pi) = y = \frac{\exp(\sigma\Phi^{-1}(\pi) - \frac{\sigma^2}{2}) - 1}{\sqrt{\exp(\sigma^2) - 1}}. \quad (12)$$

**Доказательство.**

$$\begin{aligned} \tilde{F}(y) &= P(\eta \leq y) = P\left(\frac{\xi - m}{s} \leq y\right) = \\ &= P(\log(\xi) \leq \log(m + sy)) = \\ &= P\left(\frac{\log(\xi) - \mu}{\sigma} \leq \frac{\log(m + sy) - \mu}{\sigma}\right) = \\ &= \Phi\left(\frac{\log(m + sy) - \mu}{\sigma}\right). \end{aligned}$$

Найдём  $\log(m + sy)$ , используя  $m = e^{\mu + \frac{\sigma^2}{2}}$  и  $s = m\sqrt{e^{\sigma^2} - 1}$ .

$$m + sy = e^{\mu + \frac{\sigma^2}{2}} + ye^{\mu + \frac{\sigma^2}{2}}\sqrt{e^{\sigma^2} - 1} = e^{\mu + \frac{\sigma^2}{2}}(1 + y\sqrt{\exp(\sigma^2) - 1}),$$

возьмем натуральный логарифм от обеих частей, получаем

$$\begin{aligned} \log(m + sy) &= \log(e^{\mu + \frac{\sigma^2}{2}}(1 + y\sqrt{\exp(\sigma^2) - 1})) = \\ &= \mu + \frac{\sigma^2}{2} + \log(1 + y\sqrt{\exp(\sigma^2) - 1}), \end{aligned}$$

тогда

$$\frac{\log(m + sy) - \mu}{\sigma} = \frac{\sigma}{2} + \frac{\log(1 + y\sqrt{\exp(\sigma^2) - 1})}{\sigma}.$$

То есть можно выразить

$$\tilde{F}(y) = \Phi\left(\frac{\log(m + sy) - \mu}{\sigma}\right) = \Phi\left(\frac{\sigma}{2} + \frac{\log(1 + y\sqrt{\exp(\sigma^2) - 1})}{\sigma}\right).$$

Далее можно найти  $\Phi^{-1}(\pi)$ .

$$\Phi\left(\frac{\sigma}{2} + \frac{\log(1 + y\sqrt{\exp(\sigma^2) - 1})}{\sigma}\right) = \pi,$$

$$\Phi^{-1}(\pi) = \frac{\sigma}{2} + \frac{\log(1 + y\sqrt{\exp(\sigma^2) - 1})}{\sigma}.$$

Теперь можно найти  $\log(1 + y\sqrt{\exp(\sigma^2) - 1})$ .

$$\log(1 + y\sqrt{\exp(\sigma^2) - 1}) = \sigma\Phi^{-1}(\pi) - \frac{\sigma^2}{2},$$



$$1 + y\sqrt{\exp(\sigma^2) - 1} = \exp(\sigma\Phi^{-1}(\pi) - \frac{\sigma^2}{2}).$$

В итоге получаем

$$\tilde{F}^{-1}(\pi) = y = \frac{\exp(\sigma\Phi^{-1}(\pi) - \frac{\sigma^2}{2}) - 1}{\sqrt{\exp(\sigma^2) - 1}}.$$

■

**Предложение 4.** Параметр  $\sigma$  для логнормального распределения выражается через значения квантилей, как

$$\sigma = \frac{\log\left(\frac{x_{\pi_2}}{x_{\pi_1}}\right)}{\Phi^{-1}(\pi_2) - \Phi^{-1}(\pi_1)}. \quad (13)$$

**Доказательство.** Покажем, что дисперсию логнормального распределения можно вычислить из отношения двух квантилей.

$$\begin{aligned} P(\xi \leq x_\pi) &= \pi, \\ P\left(\frac{\log(\xi) - \mu}{\sigma} \leq \frac{\log(x_\pi) - \mu}{\sigma}\right) &= \pi. \end{aligned}$$

Следовательно,

$$\Phi\left(\frac{\log(x_\pi) - \mu}{\sigma}\right) = \pi,$$

и тогда

$$\log(x_\pi) = \mu + \sigma\Phi^{-1}(\pi). \quad (14)$$

С помощью двух квантилей мы можем исключить  $\mu$  из соответствующих уравнений. Пусть есть  $\pi_1$ -ый и  $\pi_3$ -ый квантили со значениями  $x_{\pi_1}$  и  $x_{\pi_3}$ .

$$\begin{aligned} \log(x_{\pi_1}) &= \mu + \sigma\Phi^{-1}(\pi_1), \\ \log(x_{\pi_3}) &= \mu + \sigma\Phi^{-1}(\pi_3). \end{aligned}$$

Вычтем из второго уравнения первое, получаем

$$\log\left(\frac{x_{\pi_3}}{x_{\pi_1}}\right) = \sigma(\Phi^{-1}(\pi_3) - \Phi^{-1}(\pi_1)).$$

И в итоге получаем

$$\sigma = \frac{\log\left(\frac{x_{\pi_2}}{x_{\pi_1}}\right)}{\Phi^{-1}(\pi_2) - \Phi^{-1}(\pi_1)}.$$

■

**Алгоритм 2. Дано:** квантили  $x_{\pi_1}, x_{\pi_2}, x_{\pi_3}$  логнормальной случайной величины  $\eta$ ,  $\ln(\eta) \sim N(\mu, \sigma)$ .

**Шаги:**

1. Выражаем параметр  $\sigma$  из отношения  $x_{\pi_3}$  к  $x_{\pi_1}$ , используя формулу (13).
2. Вычисляем значения  $\tilde{F}^{-1}(\pi)$  для случайной величины  $\eta$  по формуле (12).
3. С помощью системы (4) находим значения вероятностей  $p_1, p_2, p_3$ .

**Результат:** вероятности  $p_1, p_2, p_3$  для  $x_{\pi_1}, x_{\pi_2}, x_{\pi_3}$  случайной величины  $\xi_n$ .

**Замечание 1.** Результаты Алгоритмов 1 и 2 совпадают.

**Пример.**

Посчитаем пример для  $\frac{x_{90}}{x_{10}} = 3$ . По формулам из этого раздела получаем

$$\sigma = \frac{\log(\frac{x_{90}}{x_{10}})}{\phi^{-1}(0,9) - \phi^{-1}(0,1)} \approx 0.428626,$$

$$\tilde{F}^{-1}(p) = \frac{\exp(\sigma\phi^{-1}(p) - \frac{\sigma^2}{2}) - 1}{\sqrt{\exp(\sigma^2) - 1}}$$

$$\tilde{F}^{-1}(0.9) \approx 1.2915826424,$$

$$\tilde{F}^{-1}(0.1) \approx -1.0539640761,$$

$$\tilde{F}^{-1}(0.5) \approx -0.1954343914.$$

Из системы (4) находим вероятности  $p_{10}, p_{50}, p_{90}$ .

$$p_{10} = 0.371243,$$

$$p_{50} = 0.282992,$$

$$p_{90} = 0.345764.$$

### 4.3 Условие на параметры для нахождения весов при аппроксимации логнормального распределения

Мы рассмотрели способы вычисления вероятностей для квантилей при аппроксимации логнормального распределения. Но эти вероятности находятся не при любом  $\sigma$ . Выясним, какое должно быть ограничение на этот параметр.

**Предложение 5.** Положительные вероятности  $p_1, p_2, p_3$  для аппроксимации логнормальной случайной величины  $\eta$  с квантилями вида  $x_\pi, x_{0.5}, x_{1-\pi}$  существуют только при условии

$$\exp(\sigma^2) + \exp(-\sigma^2) - \exp\left(\Phi^{-1}(\pi)\sigma - \frac{\sigma^2}{2}\right) - \exp\left(\Phi^{-1}(1-\pi)\sigma - \frac{\sigma^2}{2}\right) \leq 0,$$

**Доказательство.**

$$\ln(\eta) \sim N(\mu, \sigma^2), \quad \tilde{F}(y) = P(\eta \leq y).$$

С помощью формулы (12) найдем  $\tilde{F}^{-1}(\pi_i)$  и сделаем следующие обозначения

$$\tilde{F}^{-1}(\pi) = t_1, \quad \tilde{F}^{-1}(0.5) = t_2, \quad \tilde{F}^{-1}(1 - \pi) = t_3.$$

Теперь рассмотрим систему (6), запишем ее через  $t_1, t_2, t_3$  и выразим вероятности  $p_1, p_2, p_3$ .

$$\begin{aligned} p_2(t_2 - t_3) &= p_1(t_3 - t_1) - t_3, \\ p_1(t_1^2 - t_3^2) + p_2(t_2^2 - t_3^2) &= 1 - t_3^2. \end{aligned}$$

Тогда получаем

$$\begin{aligned} p_1(t_1^2 - t_3^2) + (t_2 + t_3)(p_1(t_3 - t_1) - t_3) &= 1 - t_3^2, \\ p_1(t_1 - t_3)(t_1 + t_3) + (t_2 + t_3)(p_1(t_3 - t_1) - t_3) &= 1 - t_3^2. \end{aligned}$$

$$p_1 = \frac{1 + t_2 t_3}{(t_1 - t_3)(t_1 + t_3)}, \quad (15)$$

$$p_2 = \frac{p_1(t_3 - t_1) - t_3}{t_2 - t_3} = \frac{1 + t_1 t_3}{(t_2 - t_1)(t_2 + t_3)}, \quad (16)$$

$$p_3 = 1 - p_1 - p_2. \quad (17)$$

Все вероятности должны быть положительными, подставим в формулы для вероятностей значения переменных  $t_1, t_2, t_3$ , которые ищутся по формуле (11).

$$\begin{aligned} p_1 &= \frac{1 + \frac{\left(\exp\left(-\frac{\sigma^2}{2}\right) - 1\right) \left(\exp(\Phi^{-1}(1 - \pi)\sigma - \frac{\sigma^2}{2}) - 1\right)}{\exp(\sigma^2) - 1}}{\frac{\exp\left(\Phi^{-1}(\pi)\sigma - \frac{\sigma^2}{2}\right) - \exp\left(\Phi^{-1}(1 - \pi)\sigma - \frac{\sigma^2}{2}\right)}{\sqrt{\exp(\sigma^2) - 1}} \cdot \frac{\exp\left(\Phi^{-1}(\pi)\sigma - \frac{\sigma^2}{2}\right) - \exp\left(-\frac{\sigma^2}{2}\right)}{\sqrt{\exp(\sigma^2) - 1}}} = \\ &= \frac{\exp(\sigma^2) + \exp(\Phi^{-1}(1 - \pi)\sigma - \sigma^2) - \exp(-\frac{\sigma^2}{2}) - \exp(\Phi^{-1}(1 - \pi)\sigma - \frac{\sigma^2}{2})}{\exp(\Phi^{-1}(0.1) * 2\sigma - \sigma^2) - \exp(\Phi^{-1}(\pi)\sigma - \sigma^2) - \exp(-\sigma^2) + \exp(\Phi^{-1}(1 - \pi)\sigma - \sigma^2)}. \\ p_2 &= \frac{1 + \frac{(\exp(\Phi^{-1}(\pi)\sigma - \frac{\sigma^2}{2}) - 1)(\exp(\Phi^{-1}(1 - \pi)\sigma - \frac{\sigma^2}{2}) - 1)}{\exp(\sigma^2) - 1}}{\frac{\exp(-\frac{\sigma^2}{2}) - \exp(\Phi^{-1}(\pi)\sigma - \frac{\sigma^2}{2})}{\sqrt{\exp(\sigma^2) - 1}} \cdot \frac{\exp(-\frac{\sigma^2}{2}) - \exp(\Phi^{-1}(1 - \pi)\sigma - \frac{\sigma^2}{2})}{\sqrt{\exp(\sigma^2) - 1}}} = \\ &= \frac{\exp(\sigma^2) + \exp(-\sigma^2) - \exp(\Phi^{-1}(\pi)\sigma - \frac{\sigma^2}{2}) - \exp(\Phi^{-1}(1 - \pi)\sigma - \frac{\sigma^2}{2})}{\exp(-\sigma^2) - \exp(\Phi^{-1}(1 - \pi)\sigma - \sigma^2) - \exp(\Phi^{-1}(\pi)\sigma - \sigma^2) + \exp(-\sigma^2)} = \end{aligned}$$

$$= \frac{\exp(\sigma^2) + \exp(-\sigma^2) - \exp\left(\frac{\sigma^2}{2}\right) (\exp(\Phi^{-1}(\pi)\sigma) + \exp(-\Phi^{-1}(\pi)\sigma))}{2 \exp(-\sigma^2) - \exp(-\sigma^2) (\exp(-\Phi^{-1}(\pi)\sigma) + \exp(\Phi^{-1}(\pi)\sigma))}.$$

Вероятности  $p_1$  и  $p_3$  положительные при любом параметре  $\sigma$ . Рассмотрим знаменатель  $p_2$ .

$$\begin{aligned} & 2 \exp(-\sigma^2) - \exp(-\sigma^2) (\exp(-\Phi^{-1}(\pi)\sigma) + \exp(\Phi^{-1}(\pi)\sigma)) = \\ & = \exp(-\sigma^2) (2 - \exp(-\Phi^{-1}(\pi)\sigma) - \exp(\Phi^{-1}(\pi)\sigma)) = \\ & = - \frac{\exp(-\sigma^2) (\exp(\Phi^{-1}(\pi)\sigma) - 1)^2}{\exp(\Phi^{-1}(\pi)\sigma)}. \end{aligned}$$

Числитель и знаменатель дроби положительные при любом значении параметра  $\sigma$ . Значит, весь знаменатель  $p_2$  отрицательный. Из условия отрицательности числителя получаем следующее ограничение на  $\sigma$ .

$$\exp(\sigma^2) + \exp(-\sigma^2) - \exp\left(\frac{\sigma^2}{2}\right) (\exp(\Phi^{-1}(\pi)\sigma) + \exp(-\Phi^{-1}(\pi)\sigma)) \leq 0.$$

■

Например, для  $\pi = 0.1$  получаем ограничение  $\sigma \leq 0.6913$ . Посмотрим, какому коэффициенту асимметрии соответствует это значение  $\sigma$ .

$$\begin{aligned} \gamma_3 &= \sqrt{\exp(\sigma^2) - 1} (\exp(\sigma^2) + 2), \\ \gamma_3 &= 2.82778. \end{aligned}$$

Ограничение на  $\sigma$  становится слабее при уменьшении значения  $\pi$ .

Рассмотрим  $\pi = 0.05$ , получаем ограничение  $\sigma \leq 1.04585$ .

$$\gamma_3 = 7.02529.$$

## 4.4 Варианты постановки задачи

**Задача:** имеются квантили  $x_\pi$ ,  $x_{0.5}$ ,  $x_{1-\pi}$  логнормальной случайной величины  $\eta$ . Нужно уметь считать её математическое ожидание и дисперсию.

Варианты решения задачи:

1. Не переходить к аппроксимации дискретной случайной величиной, а сразу же из двух уравнений вида (14), записанных для двух квантилей, найти значения параметров  $\mu$  и  $\sigma$  нормальной случайной величины  $\ln(\eta) \sim N(\mu, \sigma)$ . Далее по формулам (10) и (11) вычислить значения мат. ожидания  $m$  и дисперсии  $s^2$  случайной величины  $\eta$ .

2. Перейти к трехточечной аппроксимации дискретной случайной величиной  $\xi_n$ , у которой мат. ожидание и дисперсия равны мат. ожиданию  $m$  и дисперсии  $s^2$  случайной величины  $\xi$ . И считать значения  $m$  и  $s$  через квантили  $x_\pi$ ,  $x_{0.5}$ ,  $x_{1-\pi}$  и вероятности  $p_1$ ,  $p_2$ ,  $p_3$ .
3. Если условие для положительных вероятностей не выполняется, можно воспринимать задачу не как поиск вероятностей для  $\xi_n$ , а как поиск коэффициентов для  $x_\pi$ ,  $x_{0.5}$ ,  $x_{1-\pi}$  таких, чтобы параметры, полученные по формулам (2) и (3), были равны мат. ожиданию и дисперсии  $\eta$ .

## 4.5 Точность аппроксимации

Предлагаемые методы аппроксимации логнормального распределения не работают при  $\sigma \leq 0.6913$ . На практике часто используют правило Свонсона 30-40-30 для аппроксимации логнормального распределения, поэтому посмотрим на точность 30-40-30. Особенно это важно при  $\sigma \geq 0.6913$ .

### 4.5.1 Неправильное использование правила 30-40-30

**Предложение 6.** Ошибка аппроксимации мат.ожидания логнормального распределения с помощью правила 30-40-30 равна

$$\frac{|m_1 - m_2|}{m_1} = \frac{\left| \exp\left(\frac{\sigma^2}{2}\right) - \frac{1}{2(\Phi^{-1}(0.1))^2} (\exp(\sigma\Phi^{-1}(0.1)) - 1 + \exp(\sigma\Phi^{-1}(0.9))) + 1 \right|}{\exp\left(\frac{\sigma^2}{2}\right)}$$

и не зависит от параметра  $\mu$ .

**Доказательство.** Выразим ошибку аппроксимации мат.ожидания логнормального распределения через параметры  $\mu$  и  $\sigma$ .

$$m_1 = \exp\left(\mu + \frac{\sigma^2}{2}\right).$$

Имеем следующие квантили

$$x_\pi = \exp(\mu + \sigma\Phi^{-1}(0.1)),$$

Точные значения вероятностей

$$p_1 = p_3 = \frac{1}{2(\Phi^{-1}(0.1))^2},$$

$$p_2 = 1 - \frac{1}{(\Phi^{-1}(0.1))^2}.$$

Тогда мат.ожидание аппроксимации равно

$$\begin{aligned}
m_2 &= \frac{1}{2(\Phi^{-1}(0.1))^2} \exp(\mu + \sigma\Phi^{-1}(0.1)) + \\
&+ \left(1 - \frac{1}{(\Phi^{-1}(0.1))^2}\right) \exp(\mu + \sigma\Phi^{-1}(0.5)) + \frac{1}{2(\Phi^{-1}(0.1))^2} \exp(\mu + \sigma\Phi^{-1}(0.9)) = \\
&= \frac{1}{2(\Phi^{-1}(0.1))^2} \exp(\mu)(\exp(\sigma\Phi^{-1}(0.1)) - 1 + \exp(\sigma\Phi^{-1}(0.9))) + \exp(\mu).
\end{aligned}$$

Получили ошибку

$$\begin{aligned}
&\frac{|m_1 - m_2|}{m_1} = \\
&= \frac{\left| \exp\left(\mu + \frac{\sigma^2}{2}\right) - \frac{1}{2(\Phi^{-1}(0.1))^2} \exp(\mu)(\exp(\sigma\Phi^{-1}(0.1)) - 1 + \exp(\sigma\Phi^{-1}(0.9))) + \exp(\mu) \right|}{\exp\left(\mu + \frac{\sigma^2}{2}\right)} = \\
&= \frac{\left| \exp\left(\frac{\sigma^2}{2}\right) - \frac{1}{2(\Phi^{-1}(0.1))^2} (\exp(\sigma\Phi^{-1}(0.1)) - 1 + \exp(\sigma\Phi^{-1}(0.9))) + 1 \right|}{\exp\left(\frac{\sigma^2}{2}\right)} =
\end{aligned}$$

■

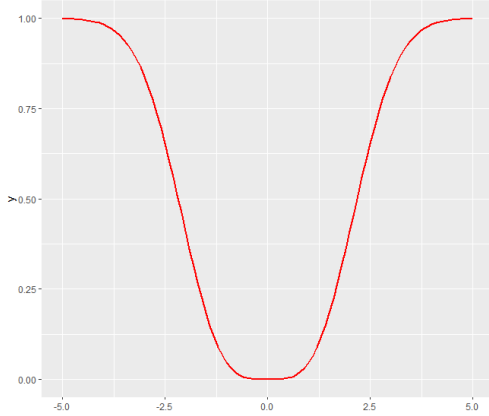


Рис. 1: Ошибка аппроксимации мат.ожидания.

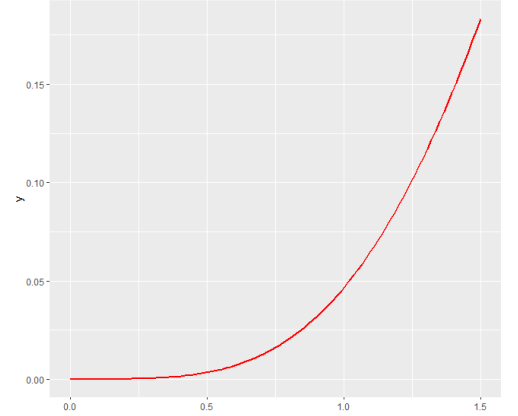


Рис. 2: Ошибка аппроксимации мат.ожидания,  $\sigma \leq 1.5$ .

**Предложение 7.** Ошибка аппроксимации дисперсии логнормального распределения с помощью правила 30-40-30 равна

$$\begin{aligned}
\frac{|d_1 - d_2|}{d_1} &= \left| \exp(\sigma^2)(\exp(\sigma^2 - 1)) - \frac{1}{2(\Phi^{-1}(0.1))^2} \exp(2\sigma\Phi^{-1}(0.1)) - \right. \\
&\quad \left. - \left(1 - \frac{1}{(\Phi^{-1}(0.1))^2}\right) \exp(2\sigma\Phi^{-1}(0.5)) - \right. \\
&\quad \left. - \frac{1}{2(\Phi^{-1}(0.1))^2} \exp(2\sigma\Phi^{-1}(0.9)) + m_2^2/2\mu \right| / \exp(\sigma^2)(\exp(\sigma^2 - 1))
\end{aligned}$$

и не зависит от параметра  $\mu$ .

**Доказательство.** Выразим аппроксимации дисперсии через параметры распределения.

$$d_1 = \exp(2\mu + \sigma^2)(\exp(\sigma^2 - 1)).$$

$$d_2 = \frac{1}{2(\Phi^{-1}(0.1))^2} \exp(2\mu + 2\sigma\Phi^{-1}(0.1)) +$$

$$+ \left(1 - \frac{1}{(\Phi^{-1}(0.1))^2}\right) \exp(2\mu + 2\sigma\Phi^{-1}(0.5)) + \frac{1}{2(\Phi^{-1}(0.1))^2} \exp(2\mu + 2\sigma\Phi^{-1}(0.9)) - m_2^2.$$

Получили ошибку

$$\begin{aligned} \left| \frac{d_1 - d_2}{d_1} \right| &= \left| \exp(2\mu + \sigma^2)(\exp(\sigma^2 - 1)) - \frac{1}{2(\Phi^{-1}(0.1))^2} \exp(2\mu + 2\sigma\Phi^{-1}(0.1)) - \right. \\ &\quad \left. - \left(1 - \frac{1}{(\Phi^{-1}(0.1))^2}\right) \exp(2\mu + 2\sigma\Phi^{-1}(0.5)) - \right. \\ &\quad \left. - \frac{1}{2(\Phi^{-1}(0.1))^2} \exp(2\mu + 2\sigma\Phi^{-1}(0.9)) + m_2^2 \right| / \exp(2\mu + \sigma^2)(\exp(\sigma^2 - 1)) = \\ &= \left| \exp(\sigma^2)(\exp(\sigma^2 - 1)) - \frac{1}{2(\Phi^{-1}(0.1))^2} \exp(2\sigma\Phi^{-1}(0.1)) - \right. \\ &\quad \left. - \left(1 - \frac{1}{(\Phi^{-1}(0.1))^2}\right) \exp(2\sigma\Phi^{-1}(0.5)) - \right. \\ &\quad \left. - \frac{1}{2(\Phi^{-1}(0.1))^2} \exp(2\sigma\Phi^{-1}(0.9)) + m_2^2/2\mu \right| / \exp(\sigma^2)(\exp(\sigma^2 - 1)). \end{aligned}$$

■

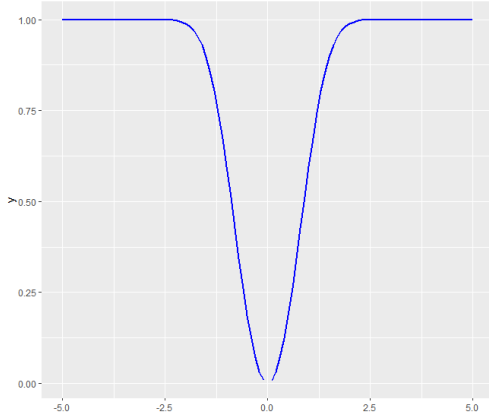


Рис. 3: Ошибка аппроксимации дисперсии.

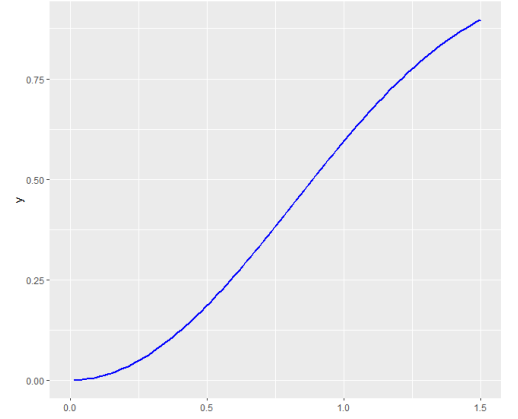


Рис. 4: Ошибка аппроксимации дисперсии,  $\sigma \leq 1.5$ .

## 5 Произведение двух логнормальных распределений

Нам доступен метод объединения любых логнормально распределенных случайных величин. Эта процедура применяется в нефтяной промышленности, она выполняется быстро и может быть выполнена вручную. Например, используем площадь дренирования пласта, среднюю чистую толщину и коэффициент извлечения углеводородов. При перемножении этих параметров получаем количество резервов нефти.

Рассмотрим произведение любых двух логнормально распределенных случайных величин.

$$\begin{aligned}\ln(\xi_1) &\sim N(\mu_1, \sigma_1^2), \\ \ln(\xi_2) &\sim N(\mu_2, \sigma_2^2).\end{aligned}$$

Введем следующие обозначения:

$x_\pi, x_{0.5}, x_{1-\pi}$  — квантили случайной величины  $\xi_1$ ,

$y_\pi, y_{0.5}, y_{1-\pi}$  — квантили случайной величины  $\xi_2$ .

**Предложение 8.** При перемножении квантилей  $x_\pi$  и  $y_\pi$  двух логнормальных случайных величин  $\xi_1$  и  $\xi_2$  получается квантиль случайной величины  $\xi_1\xi_2$  вида  $z_q$ , где

$$q = P(\xi_1\xi_2 < x_\pi y_\pi) = \Phi\left(\frac{\Phi^{-1}(\pi)(\ln(x_{0.5}) + \ln(y_{0.5}) - \ln(x_\pi) - \ln(y_\pi))}{\sqrt{(\ln(x_{0.5}) - \ln(x_\pi))^2 + (\ln(y_{0.5}) - \ln(y_\pi))^2}}\right). \quad (18)$$

**Доказательство.** Выразим параметры распределений  $\mu_1, \mu_2, \sigma_1, \sigma_2$  через квантили. По определению квантиля  $P(\xi_1 < x_\pi) = \pi$ .

Преобразуем эту вероятность так, чтобы ее можно было записать через функцию распределения стандартного нормального распределения, следующим образом:

$$P(\xi_1 < x_\pi) = P(\ln(\xi_1) < \ln(x_\pi)) = P\left(\frac{\ln(\xi_1) - \mu_1}{\sigma_1} < \frac{\ln(x_\pi) - \mu_1}{\sigma_1}\right).$$

Так как  $\xi_1$  распределена логнормально с параметрами  $\mu_1$  и  $\sigma_1^2$ , то

$$\frac{\ln(\xi_1) - \mu_1}{\sigma_1} \sim N(0, 1).$$

Следовательно, можно записать логарифм квантиля, как:

$$\ln(x_\pi) = \sigma_1 \Phi^{-1}(\pi) + \mu_1. \quad (19)$$

Аналогично для  $x_{50}$ , получаем, что

$$\mu_1 = \ln(x_{0.5}). \quad (20)$$

Используя формулы (19) и (20) можно выразить значение  $\sigma_1$ .

$$\sigma_1 = \frac{\ln(x_\pi) - \ln(x_{0.5})}{\Phi^{-1}(\pi)}. \quad (21)$$



Аналогичные действия проводим для  $\xi_2$  и тогда получаем

$$\frac{\ln(y_\pi) - \mu_2}{\sigma_2} = \Phi^{-1}(\pi), \quad (22)$$

$$\mu_2 = \ln(y_{0.5}). \quad (23)$$

Используя формулы (22) и (23) можно выразить значение  $\sigma_2$ ,

$$\sigma_2 = \frac{\ln(y_\pi) - \ln(y_{0.5})}{\Phi^{-1}(\pi)}. \quad (24)$$

Теперь рассмотрим случайную величину  $\eta = \xi_1 \xi_2$ . Мы хотим вычислить, каким квантилем для  $\eta$  является произведение квантилей  $x_\pi$  и  $y_\pi$ .

Для этого надо найти, чему равна вероятность  $P(\xi_1 \xi_2 < x_\pi y_\pi)$ .

$$\begin{aligned} P(\xi_1 \xi_2 < x_\pi y_\pi) &= P(\ln(\xi_1) + \ln(\xi_2) < \ln(x_\pi) + \ln(y_\pi)) = \\ &= P\left(\frac{\ln(\xi_1) + \ln(\xi_2) - (\mu_1 + \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} < \frac{\ln(x_\pi) + \ln(y_\pi) - (\mu_1 + \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right). \end{aligned}$$

Так как  $\xi_1$  распределена логнормально с параметрами  $\mu_1$  и  $\sigma_1^2$ , а  $\xi_2$  распределена логнормально с параметрами  $\mu_2$  и  $\sigma_2^2$ , то

$$\begin{aligned} \ln(\xi_1) + \ln(\xi_2) &\sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2), \\ \frac{\ln(\xi_1) + \ln(\xi_2) - (\mu_1 + \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} &\sim N(0, 1). \end{aligned}$$

Используя формулы (19) и (21), выразим  $\ln(x_\pi)$  и  $\ln(y_\pi)$ .

$$\ln(x_\pi) = \mu_1 + \Phi^{-1}(\pi)\sigma_1,$$

$$\ln(y_\pi) = \mu_2 + \Phi^{-1}(\pi)\sigma_2.$$

Тогда можно записать

$$\begin{aligned} &P(\xi_1 \xi_2 < x_\pi y_\pi) = \\ &= P\left(\frac{\ln(\xi_1) + \ln(\xi_2) - (\mu_1 + \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} < \frac{(\mu_1 + \Phi^{-1}(\pi)\sigma_1) + (\mu_2 + \Phi^{-1}(\pi)\sigma_2) - (\mu_1 + \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) = \\ &= P\left(\frac{\ln(\xi_1) + \ln(\xi_2) - (\mu_1 + \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} < \frac{\Phi^{-1}(\pi)(\sigma_1 + \sigma_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) = \\ &= \Phi\left(\frac{\Phi^{-1}(\pi)(\sigma_1 + \sigma_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right). \end{aligned}$$

Используя формулы (21) и (24), перепишем эту дробь через значения квантилей.

$$\begin{aligned} \frac{\Phi^{-1}(\pi)(\sigma_1 + \sigma_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} &= \frac{\Phi^{-1}(\pi) \left( \frac{(\ln(x_{0.5}) - \ln(x_\pi)) + (\ln(y_{0.5}) - \ln(y_\pi))}{-\Phi^{-1}(\pi)} \right)}{\sqrt{\frac{(\ln(x_{0.5}) - \ln(x_\pi))^2 + (\ln(y_{0.5}) - \ln(y_\pi))^2}{(\Phi^{-1}(\pi))^2}}} = \\ &= \frac{(\ln(x_{0.5}) - \ln(x_\pi)) + (\ln(y_{0.5}) - \ln(y_\pi))}{\sqrt{(\ln(x_{0.5}) - \ln(x_\pi))^2 + (\ln(y_{0.5}) - \ln(y_\pi))^2} \cdot \Phi^{-1}(\pi)}. \end{aligned}$$

Тогда получаем следующую формулу

$$P(\xi_1 \xi_2 < x_\pi y_\pi) = \Phi \left( \frac{\Phi^{-1}(\pi)(\ln(x_{0.5}) + \ln(y_{0.5}) - \ln(x_\pi) - \ln(y_\pi))}{\sqrt{(\ln(x_{0.5}) - \ln(x_\pi))^2 + (\ln(y_{0.5}) - \ln(y_\pi))^2}} \right).$$

■

Таким образом, с помощью формулы (18) можно посчитать, какой квантиль получается при перемножении  $\pi$ -ых квантилей.

**Следствие 1.** При перемножении квантилей  $x_{0.5}$  и  $y_{0.5}$  получается снова 0.5-ый квантиль.

**Доказательство.** Из раздела 4.2 знаем, что  $P(\xi_1 \xi_2 < x_{0.5} y_{0.5})$  можно написать следующим образом:

$$P(\xi_1 \xi_2 < x_{0.5} y_{0.5}) = \Phi \left( \frac{\ln(x_{0.5}) + \ln(y_{0.5}) - (\mu_1 + \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} \right).$$

Но по формуле (14) в числителе получается 0.  
Значит,

$$P(\xi_1 \xi_2 < x_{0.5} y_{0.5}) = \Phi(0) = 0.5.$$

■

## 5.1 Квантили вида $\pi$ , 0.5, $1 - \pi$ произведения логнормальных случайных величин

Как по каким-то произвольным получившимся квантилям, полученным при перемножении данных квантилей для двух логнормальных случайных величин, найти нужные нам, такие же, как исходные  $\pi$ , 0.5,  $1 - \pi$  квантили произведения этих двух случайных величин? Сначала нужно понять на какой прямой лежат точки вида  $(x_\pi; \Phi^{-1}(\pi))$ .

Для этого рассмотрим следующий QQ-плот:

$$\{x_i, F_\eta^{-1}(F_\xi(x_i))\}_{i=1}^n.$$

Как связаны параметры нормального распределения, квантили которого откладываются по оси  $X$ , и параметры прямой, на которой лежат точки этого QQ-плота?

Ось  $X$ :  $\xi \sim N(a, \sigma^2)$ .

Ось  $Y$ :  $\eta \sim N(0, 1)$ .

Возьмем две точки и построим по ним уравнение прямой.

$$(F_{\xi}^{-1}(0.1), F_{\eta}^{-1}(0.1)),$$

$$(F_{\xi}^{-1}(0.5), F_{\eta}^{-1}(0.5)).$$

$$\Phi\left(\frac{x_p - a}{\sigma}\right) = p \quad \Rightarrow \quad \frac{x_p - a}{\sigma} = \Phi^{-1}(p).$$

Получаем, что

$$x_p = a + \sigma\Phi^{-1}(p).$$

Для первой точки возьмем  $p = 0.1$ .

$$(a + \sigma\Phi^{-1}(0.1); \Phi^{-1}(0.1)).$$

Для второй точки возьмем  $p = 0.5$ .

$$(a + \sigma\Phi^{-1}(0.5); \Phi^{-1}(0.5)) \quad \Rightarrow \quad (a; 0).$$

Составим уравнение прямой:

$$\frac{x - a}{(a + \Phi^{-1}(0.1)\sigma) - a} = \frac{y}{\Phi^{-1}(0.1)}, \quad \frac{x - a}{\Phi^{-1}(0.1)\sigma} = \frac{y}{\Phi^{-1}(0.1)}.$$

Следовательно,

$$\sigma y = x - a,$$

Получили уравнение прямой на которой лежат точки данного QQ-плота:

$$y = \frac{x - a}{\sigma}. \quad (25)$$

**Предложение 9.** Зная квантили  $x_{\pi}$ ,  $x_{0.5}$ ,  $x_{1-\pi}$  случайной величины  $\xi_1$  и квантили  $y_{\pi}$ ,  $y_{0.5}$ ,  $y_{1-\pi}$  случайной величины  $\xi_2$  можно найти квантили  $z_{\pi}$ ,  $z_{0.5}$ ,  $z_{1-\pi}$  случайной величины  $\xi_1\xi_2$ , как

$$z_{\pi} = \exp(\sigma\Phi^{-1}(\pi) + a),$$

$$z_{0.5} = x_{0.5}y_{0.5},$$

$$z_{1-\pi} = \exp(\sigma\Phi^{-1}(1 - \pi) + a),$$

где  $a$  и  $\sigma$  – параметры прямой  $y = \frac{x - a}{\sigma}$ , на которой лежат точки  $(\ln(x_{\pi}y_{\pi}), t)$  и  $(\ln(x_{0.5}y_{0.5}), 0)$ ,

$$t = \frac{\Phi^{-1}(\pi)((\ln(x_{0.5}) + \ln(y_{0.5})) - (\ln(x_{\pi}) + \ln(y_{\pi})))}{\sqrt{(\ln(x_{0.5}) - \ln(x_{\pi}))^2 + (\ln(y_{0.5}) - \ln(y_{\pi}))^2}}.$$

**Доказательство.** С помощью формулы (18) можно посчитать, какой получается квантиль для случайной величины  $\xi_1\xi_2$ , если перемножить квантили  $x_\pi$  и  $y_\pi$  исходных случайных величин.

Обозначим  $z_\pi$ ,  $z_{0.5}$ ,  $z_{1-\pi}$  — квантили случайной величины  $\eta$ . Тогда по Следствию 1 имеем  $x_{0.5}y_{0.5} = z_{0.5}$ .

Нужно вычислить значения  $z_\pi$  и  $z_{1-\pi}$ . Введем обозначение:

$$t = \frac{\Phi^{-1}(\pi)((\ln(x_{0.5}) + \ln(y_{0.5})) - (\ln(x_\pi) + \ln(y_\pi)))}{\sqrt{(\ln(x_{0.5}) - \ln(x_\pi))^2 + (\ln(y_{0.5}) - \ln(y_\pi))^2}}.$$

Тогда с помощью точек  $(\ln(x_\pi y_\pi), t)$  и  $(\ln(x_{0.5}y_{0.5}), 0)$  можно найти параметры  $a$  и  $\sigma$  прямой, на которой они лежат, по формуле (25).

$$\frac{\ln(x_{0.5}y_{0.5}) - a}{\sigma} = 0 \quad \Rightarrow \quad a = \ln(x_{0.5}y_{0.5}),$$

$$\frac{\ln(x_\pi y_\pi) - a}{\sigma} = t,$$

$$\sigma = \frac{\ln(x_\pi y_\pi) - a}{t} = \frac{\ln(x_\pi y_\pi) - \ln(x_{0.5}y_{0.5})}{t}.$$

Так как точки  $(\ln(z_\pi), \Phi^{-1}(\pi))$  и  $(\ln(z_{1-\pi}), \Phi^{-1}(1 - \pi))$  тоже лежат на этой прямой, то мы можем вычислить значения  $\ln(z_\pi)$  и  $\ln(z_{0.5})$ , зная уравнение прямой, следующим образом:

$$\frac{\ln(z_\pi) - a}{\sigma} = \Phi^{-1}(\pi),$$

$$\ln(z_\pi) = \sigma\Phi^{-1}(\pi) + a,$$

$$\frac{\ln(z_{1-\pi}) - a}{\sigma} = \Phi^{-1}(1 - \pi),$$

$$\ln(z_{1-\pi}) = \sigma\Phi^{-1}(1 - \pi) + a.$$

И, наконец, находим  $z_\pi$  и  $z_{1-\pi}$ .

$$z_\pi = \exp(\sigma\Phi^{-1}(\pi) + a),$$

$$z_{1-\pi} = \exp(\sigma\Phi^{-1}(1 - \pi) + a).$$

■

Как теперь найти математическое ожидание  $\eta = \xi_1\xi_2$ ? Случайные величины  $\xi_1$  и  $\xi_2$  распределены логнормально. Их произведение — случайная величина  $\eta$  тоже имеет логнормальное распределение, поэтому

$$\ln(\eta) = \ln(\xi_1\xi_2) = \ln(\xi_1) + \ln(\xi_2) \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

В разделе 3 было описано, как искать математическое ожидание и дисперсию. Можно использовать метод Свонсона аппроксимации нормального распределения для  $\ln(\eta)$ . Для этого надо взять не сами квантили  $z_\pi$ ,  $z_{0.5}$  и  $z_{1-\pi}$ , а их логарифмы. Соответствующие вероятности  $p_1$ ,  $p_2$ ,  $p_3$  можно найти с помощью системы (7), так как данные квантили симметричны.

## 6 Сумма двух логнормальных распределений

Рассмотрим сумму двух логнормальных случайных величин.

$$\ln(\xi_1) \sim N(\mu_1, \sigma_1^2),$$

$$\ln(\xi_2) \sim N(\mu_2, \sigma_2^2),$$

$$\eta = \xi_1 + \xi_2.$$

Дано: квантили  $x_\pi$ ,  $x_{0.5}$ ,  $x_{1-\pi}$  случайной величины  $\xi_1$  и квантили  $y_\pi$ ,  $y_{0.5}$ ,  $y_{1-\pi}$  случайной величины  $\xi_2$ .

Нужно найти квантили  $z_\pi$ ,  $z_{0.5}$ ,  $z_{1-\pi}$  случайной величины  $\eta$ , а также вычислить вероятности  $p_1$ ,  $p_2$ ,  $p_3$  такие, что мат. ожидание и дисперсия исходной логнормальной случайной величины равны мат. ожиданию и дисперсии дискретной аппроксимации.

Берем симметричные квантили, а именно  $\pi = 0.1$ . Чтобы найти  $z_{10}$ ,  $z_{50}$ ,  $z_{90}$  будем использовать аппроксимацию суммы логнормальных распределений логнормальным распределением.  $\ln(\eta) \sim N(\mu, \sigma)$ .

У нас есть следующие ограничения на параметры:  $\mu_1, \mu_2 < 12$ ,  $\sigma_1, \sigma_2 < 1.5$ . Пусть мы нашли аппроксимацию суммы двух логнормальных величин, тогда с учетом этих ограничений её значения  $\mu$  и  $\sigma$  тоже будут иметь свои ограничения. При этом, чтобы найти значения вероятностей  $p_1$ ,  $p_2$ ,  $p_3$  нужно, чтобы выполнялось то же условие, что в разделе 4.3. А именно,  $\sigma < 0.6913$ .

Альтернатива: Если это ограничение на  $\sigma$  не выполняется и мы не можем вычислить положительные вероятности, то можно не переходить к аппроксимации дискретным распределением, а вычислить значения мат. ожидания и дисперсии  $\eta$  с помощью квантилей  $z_\pi$ ,  $z_{0.5}$ ,  $z_{1-\pi}$  по формулам вида (14).

Имеем следующий алгоритм для решения задачи.

**Алгоритм 3.** Дано: Квантили  $x_\pi$ ,  $x_{0.5}$ ,  $x_{1-\pi}$  — квантили  $\xi_1$ ,  $y_\pi$ ,  $y_{0.5}$ ,  $y_{1-\pi}$  — квантили  $\xi_2$ .

1. Найти параметры  $\mu_1$ ,  $\sigma_1$ ,  $\mu_2$  и  $\sigma_2$  через значения квантилей, используя формулы раздела 4.2.
2. Вычислить значения мат. ожидания  $m$  и дисперсии  $s^2$  случайной величины  $\xi_1 + \xi_2$ , как суммы  $m = m_1 + m_2$ ,  $d = d_1 + d_2$ , где  $m_1, d_1$  — мат.ожидание и дисперсия  $\xi_1$ , а  $m_2, d_2$  — случайной величины  $\xi_2$ . Они пересчитываются аналогично  $m$  и  $d$ .

3. Выразить параметры  $\mu$  и  $\sigma$  нормального распределения через параметры  $m$  и  $d$  логнормального распределения, используя формулы (10) и (11).
4. Вычислить, какой квантиль получается при сложении  $x_\pi$  и  $y_\pi$ , используя следующую формулу

$$\begin{aligned} P(\xi_1 + \xi_2 < x_\pi + y_\pi) &= P(\ln(\xi_1 + \xi_2) < \ln(x_\pi + y_\pi)) = \\ &= P\left(\frac{\ln(\xi_1 + \xi_2) - \mu}{\sigma} < \frac{\ln(x_\pi + y_\pi) - \mu}{\sigma}\right) = \\ &= \Phi\left(\frac{\ln(x_\pi + y_\pi) - \mu}{\sigma}\right). \end{aligned}$$

5. Найти значения квантилей  $z_{10}$ ,  $z_{50}$ ,  $z_{90}$  по Алгоритму 2.

**Результат:** вероятности  $p_1$ ,  $p_2$ ,  $p_3$  для квантилей  $z_{\pi_1}$ ,  $z_{\pi_2}$ ,  $z_{\pi_3}$  случайной величины  $\xi_1 + \xi_2$ .

## 7 Заключение

Таким образом, были получены следующие результаты: методы аппроксимации нормального и логнормального распределений, условие на  $\sigma$  для аппроксимации логнормального распределения, точность аппроксимации логнормального правилом 30-40-30, методы аппроксимации суммы и произведения двух логнормальных распределений.

При этом возникали проблемы с тем, что аппроксимировать дискретным распределением получается только при ограниченных значениях параметра  $\sigma$  и тем, что для суммы логнормальных результат имеет ошибки, так как сумма логнормальных распределений не является логнормальным распределением.

## Список литературы

- [1] Keith G. Swanson's Swansong.— Текст: электронный // stochastic: [сайт].— URL: <https://www.stochastic.dk/post/swanson-s-swansong> (дата обращения: 23.12.2021).
- [2] Uncertainties impacting reserves, revenue, and costs.— Текст: электронный // AAPG Wiki: [сайт].— URL: [https://wiki.aapg.org/Uncertainties impacting reserves, revenue, and costs](https://wiki.aapg.org/Uncertainties_impacting_reserves_revenue_and_costs) (дата обращения: 27.05.2022).
- [3] Bickel, J. Eric, Lake, Larry W., and John Lehman. "Discretization, Simulation, and Swanson's (Inaccurate) Mean."SPE Econ Mgmt 3 (2011): 128–140. doi: <https://doi.org/10.2118/148542-PA>.

- [4] Bickel, J. Eric. "Discretization, Simulation, and the Value of Information." Paper presented at the SPE Annual Technical Conference and Exhibition, Denver, Colorado, USA, October 2011. doi: <https://doi.org/10.2118/145690-MS>.
- [5] Moghadasi, Maryam and Jerry L. Jensen. "Performance Evaluation of Swanson's Rule for the Case of Log-Normal Populations." (2014). DOI:10.1007/978-3-642-32408.

## 8 Приложение

На C++ были реализованы следующие полезные на практике функции.

•

Дано: значения квантилей  $x_{\pi_i}$ , математическое ожидание  $m$ , дисперсия  $s^2$  непрерывной случайной величины.

Задача: найти вероятности  $p_i$  такие, что непрерывное распределение можно заменить дискретным с данными квантилями и полученными весами с сохранением математического ожидания и дисперсии.

Решение описано в разделе 2.

Система:

$$\begin{pmatrix} 1 & 1 & 1 \\ x_{\pi_1} & x_{\pi_2} & x_{\pi_3} \\ x_{\pi_1}^2 & x_{\pi_2}^2 & x_{\pi_3}^2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 1 \\ m \\ m^2 + s^2 \end{pmatrix}.$$

Функция:

vector<double> P (double  $m$ , double  $s$ , double  $x_{\pi_1}$ , double  $x_{\pi_2}$ , double  $x_{\pi_3}$ ).

•

Дано: вероятности  $\pi_i$ .

Задача: найти вероятности  $p_i$  для дискретного распределения, заменяющего исходное нормальное распределение, с любыми тремя квантилями  $x_{\pi_1}$ ,  $x_{\pi_2}$  и  $x_{\pi_3}$ .

Решение описано в разделе 3.

Система:

$$\begin{pmatrix} 1 & 1 & 1 \\ \Phi^{-1}(\pi_1) & \Phi^{-1}(\pi_2) & \Phi^{-1}(\pi_3) \\ \Phi^{-1}(\pi_1)^2 & \Phi^{-1}(\pi_2)^2 & \Phi^{-1}(\pi_3)^2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Функция:

vector<double> PNormal (double  $\pi_1$ , double  $\pi_2$ , double  $\pi_3$ ).

•

Дано: вероятность  $\pi$ .

Задача: найти вероятности  $p_i$  для дискретного распределения, заменяющего исходное нормальное распределение, в случае симметричных квантилей вида  $\pi$ , 0.5 и  $1 - \pi$ .

Решение описано в разделе 3 с помощью системы (2).

Формулы:

$$\begin{cases} p_\pi = \frac{1}{2\Phi^{-1}(\pi)^2}, \\ p_{0.5} = 1 - \frac{1}{\Phi^{-1}(\pi)^2}, \\ p_{1-\pi} = \frac{1}{2\Phi^{-1}(\pi)^2}. \end{cases}$$

Функция:

vector<double> PNormalSim (double  $\pi$ ).

•

Дано: параметры нормального распределения  $\mu$  и  $\sigma$ , соответствующего логнормальному распределению.

Задача: найти параметры этого логнормального распределения  $m$  и  $s$ .

Решение получено из определений логнормального распределения и соответствующего ему нормального распределения.

Формулы:

$$\begin{aligned} m &= \exp\left(\mu + \frac{\sigma^2}{2}\right), \\ s^2 &= m^2(\exp(\sigma^2) - 1). \end{aligned}$$

Функции:

double M (double  $\mu$ , double  $\sigma$ ),

double S (double  $\mu$ , double  $\sigma$ ).

•

Дано: вероятности  $\pi_1$ ,  $\pi_2$ , значения квантилей  $x_{\pi_1}$ ,  $x_{\pi_2}$ .

Задача: найти дисперсию логарифмически нормального распределения через квантили дискретного распределения, которое его заменяет.



Решение описано в разделе 4.2, получена формула (2).

Формула:

$$\sigma = \frac{\log\left(\frac{x_{\pi_2}}{x_{\pi_1}}\right)}{\Phi^{-1}(\pi_2) - \Phi^{-1}(\pi_1)}.$$

Функция:

double Sig (double  $\pi_1$ , double  $\pi_2$ , double  $x_{\pi_1}$ , double  $x_{\pi_2}$ ).

•

Дано: вероятность  $\pi$ , значения квантилей  $x_\pi$ ,  $x_{0.5}$ .

Задача: найти дисперсию логарифмически нормального распределения через квантили дискретного распределения, которое его заменяет в случае симметричных квантилей.

Решение получено как частый случай формулы (2).

Формула:

$$\sigma = \frac{\ln(x_\pi) - \ln(x_{0.5})}{\Phi^{-1}(\pi)}.$$

Функция:

double SigSim (double  $\pi$ , double  $x_\pi$ , double  $x_{0.5}$ ).

•

Дано: вероятность  $\pi$ , параметры нормального распределения  $\mu$ ,  $\sigma$ .

Задача: найти квантили логнормальной случайной величины, зная параметры соответствующего нормального распределения в случае симметричных квантилей.

Решение описано в разделе 5.

Формулы:

$$\begin{aligned}\ln(x_\pi) &= \mu + \Phi^{-1}(\pi)\sigma, \\ \ln(x_{0.5}) &= \mu, \\ \ln(x_{1-\pi}) &= \mu + \Phi^{-1}(1 - \pi)\sigma.\end{aligned}$$

Функция:

double lnX (double  $\pi$ , double  $\mu$ , double  $\sigma$ ).

•

Дано: вероятность  $\pi$ , дисперсии  $\sigma_1$  и  $\sigma_2$  нормальных случайных величин.

Задача: понять, какой квантиль получается при перемножении квантилей логнормальных случайных величин, через дисперсии соответствующих нормальных случайных величин в случае симметричных квантилей.

Решение описано в разделе 5.

Формулы:

$$P(\xi_1 \xi_2 < x_\pi y_\pi) = \Phi \left( \frac{\Phi^{-1}(\pi)(\sigma_1 + \sigma_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} \right),$$

$$q = \frac{\Phi^{-1}(\pi)(\sigma_1 + \sigma_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}}.$$

Функция:

double ProbPr (double  $\pi$ , double  $\sigma_1$ , double  $\sigma_2$ ).

•

Дано: вероятность  $\pi$ , квантили  $x_\pi$ ,  $x_{0.5}$  и  $y_\pi$ ,  $y_{0.5}$ .

Задача: понять, какой квантиль получается при перемножении  $\pi$ -ых квантилей логнормальных случайных величин, через логарифмы  $\pi$ -го и 0.5-го квантилей.

Решение описано в разделе 5, получена формула (??).

Формулы:

$$P(\xi_1 \xi_2 < x_\pi y_\pi) = \Phi \left( \frac{\Phi^{-1}(\pi)(\ln(x_{0.5}) + \ln(y_{0.5}) - \ln(x_\pi) - \ln(y_\pi))}{\sqrt{(\ln(x_{0.5}) - \ln(x_\pi))^2 + (\ln(y_{0.5}) - \ln(y_\pi))^2}} \right),$$

$$q = \frac{\Phi^{-1}(\pi)(\ln(x_{0.5}) + \ln(y_{0.5}) - \ln(x_\pi) - \ln(y_\pi))}{\sqrt{(\ln(x_{0.5}) - \ln(x_\pi))^2 + (\ln(y_{0.5}) - \ln(y_\pi))^2}}.$$

Функция:

double ProbPrX (double  $\pi$ , double  $x_\pi$ , double  $x_{0.5}$ , double  $y_\pi$ , double  $y_{0.5}$ ).

•

Дано: вероятность  $\pi$ , квантили  $x_\pi$ ,  $x_{0.5}$  и  $y_\pi$ ,  $y_{0.5}$ .

Задача: найти значения  $\pi$ -го квантиля для произведения двух логнормально распределенных случайных величин через их квантили.

Решение описано в разделе 7.

Формулы:

$$z_{\pi} = \exp \left( \frac{\ln(x_{\pi}y_{\pi}) - \ln(x_{0.5}y_{0.5})}{q} \Phi^{-1}(\pi) + \ln(x_{0.5}y_{0.5}) \right),$$

$$q = \Phi^{-1}(\mathrm{P}(\xi_1\xi_2 < x_{\pi}y_{\pi})).$$

Функция:

double Q (double  $\pi$ , double  $x_{\pi}$ , double  $x_{0.5}$ , double  $y_{\pi}$ , double  $y_{0.5}$ ).