

УДК 62-50

БІДЮК П.І., БОРИСЕВИЧ А.С.

## **ОЦІНЮВАННЯ ПАРАМЕТРІВ МОДЕЛЕЙ ІЗ ЗАСТОСУВАННЯМ МЕТОДУ МОНТЕ- КАРЛО ДЛЯ МАРКОВСЬКИХ ЛАНЦЮГІВ**

Розкрито існуючі основні тенденції у напрямі байєсівського оцінювання математичних і статистичних моделей процесів довільної природи із використанням чисельних методів Монте-Карло. Наявність випадкової складової у досліджуваних процесах вимагає введення в модель випадкових змінних з відповідними розподілами та застосування методів аналізу і описання випадкових процесів. Наведено приклад застосування методу Монте-Карло для марковських ланцюгів.

The current tendencies of Bayesian estimation are disclosed regarding parameter estimation of mathematical and statistical models by making use of Monte-Carlo Markov chain technique. As far as most of the processes investigated contain random component, it is necessary to introduce stochastic processes into models with respective distributions. An example of application of the MCMC technique is given.

### **Вступ**

Актуальною задачею оцінювання математичних і статистичних моделей є на сьогодні застосування байєсівської методології до процесів різної природи. Байєсівський підхід, як альтернативний класичному статистичному підходу, дозволяє точніше та повніше оцінювати моделі; він дає можливість отримувати хороші результати у тих випадках, коли використання класичних статистичних методів дуже обмежене (наприклад, випадки з короткою вибіркою статистичних даних). Байєсівській підхід відкриває нові, досить широкі можливості застосування методів математичного моделювання, а розроблені обчислювальні алгоритми оцінювання на основі генерування випадкових чисел дають змогу розв'язати поставлені задачі за допомогою сучасних обчислювальних процедур.

Класичний підхід приділяє головну увагу одержанню ефективних алгоритмів оцінювання та вивчення їх асимптотичних властивостей, які виступають підґрунтям для формування статистичного висновку на основі даних відносно великого об'єму. У випадку коротких вибірок використання результатів асимптотичної теорії представляється недостатньо обґрунтованим. Байєсівський підхід до формування статистичного висновку ґрунтується на інших теоретичних передумовах. Байєсівські методи вирізняються від класичних іншим підходом до інтерпретації дійсних параметрів моделі. Класичний виходить із того, що дійсні параметри – це не випадкові величини, а апроксимуючі їх оцінки – випадкові, оскільки вони є функціями спостережень, що містять випадкові процеси [1]. Байєсівський підхід відноситься до числа тих, що надають більш широке трактування дійсним параметрам моделі. Він

виходить із того, що параметри випадкові, тобто випадковість розглядається як іманентна властивість реального фізичного світу, вважаючи, що сам фізичний об'єкт зазнає неперервних випадкових змін. Тому шукають не випадкові оцінки, які досить близько апроксимують яку-небудь статистику випадкового параметра, наприклад, його середнє значення або моду. При практичному застосуванні вже оціненої моделі різниця практично несуттєва – дослідник працює з моделлю, яка має детерміновані коефіцієнти. Ймовірнісні властивості моделі використовуються для визначення похибки прогнозу та аналізу чутливості моделі, обчислення функції втрат і т. ін. Очевидно, що подібні обчислення можна виконувати в обох підходах.

Байєсівська методологія досліджувалась у багатьох роботах та застосовалась у різних галузях науки і техніки. Зокрема, Зельнер А. досліджував використання таких методів у економетрії [1]; Савчук В.П. аналізував надійність технічних об'єктів [2]; відомо також багато інших напрямів застосування цих методів. Використання байєсівської методології у таких випадках зводилось до великих аналітичних досліджень, які іноді вимагали ґрунтовних знань математики та статистики [3]. Швидкий розвиток комп'ютерної техніки дав можливість розпочати дослідження альтернативних обчислювальних алгоритмів, які ґрунтуються на відомих принципах генерування випадкових чисел [4 – 14] та їх використання для формування оцінок параметрів математичних і статистичних моделей випадкових процесів.

### **1. Постановка задачі**

Мета цієї статті – проаналізувати та розкрити існуючі основні тенденції та дослідження у напрямі байєсовського оцінювання математичних і статистичних моделей процесів довільної природи із використанням чисельних методів Монте-Карло. Ми припускаємо, що випадкові процеси містять детерміновану складову, яка може бути описана деякою вибраною детермінованою функцією, наприклад, авторегресією, регресією, поліномом і т. ін. Наявність випадкової складової у досліджуваному процесі вимагає введення в модель випадкових змінних з відповідними розподілами та застосування методів аналізу і описання випадкових процесів.

### **2. Суть байєсівського підходу**

Байєсівські методи розроблено в результаті систематичних спроб сформулювати та розв'язати проблеми статистичного аналізу поведінки процесів та систем різної природи на основі теореми Байєса. Передумовою до використання цієї теореми є деякі співвідношення між ймовірностями подій різного характеру та специфікації кожної з них на необхідному рівні [3].

Багато статистичних задач, незалежно від методів їх розв'язання, мають деякі загальні властивості. До отримання конкретної вибірки даних потенційно прийнятними для деякої досліджуваної ситуації розглядаються декілька ймовірнісних моделей. Після отримання даних виникає виражене у деякому числовому вигляді знання щодо відносної прийнятності цих моделей.

Відмінність байєсівської парадигми від інших статистичних підходів полягає в тому, що ще до отримання даних дослідник розглядає ступінь своєї довіри до можливих моделей та представляє її у вигляді ймовірностей. Як тільки дані отримано, теорема Байєса дозволяє досліднику розрахувати ще одну множину ймовірностей, які представляють собою переглянуті ступені довіри до можливих моделей-кандидатів із врахуванням нової інформації, що надійшла з даними.

Однією із ключових переваг байєсівського підходу є використання будь-якої початкової (ап'юріорної) інформації щодо параметрів моделі. Така інформація виражається у вигляді ап'юріорної ймовірності або функції щільності, тобто вона

приймається перед початком подальшого аналізу задачі. Надалі початкові ймовірності “переглядають”, використовуючи вибіркві дані, що знаходить своє відображення у вигляді апостеріорного розподілу оцінок параметрів чи змінних моделей.

Розглянемо випадкову змінну  $X$ , яка має розподіл ймовірностей, визначений у термінах невідомого параметра  $\theta$ , що належить визначеній множині можливих значень параметра  $\Theta$ . Для заданого значення  $X = x$  функція правдоподібності кожного окремого значення  $\theta$  задається як  $P(x/\theta)$ . У неперервному випадку апіорні ймовірності для множини можливих моделей відповідають, у загальному розумінні, розподілу ймовірностей на множені  $\Theta$  можливих значень параметра. Таким чином, апіорні характеристики (судження) визначають у вигляді апіорної щільності ймовірності:

$$P(\theta), \quad \theta \in \Theta \text{ такої, що } \int_{\Theta} P(\theta) d\theta = 1. \quad (1)$$

Апіорний розподіл переглядається на основі вибіркових даних  $X = x$  з метою отримання апостеріорної щільності ймовірності  $P(\theta/x)$ ,  $\theta \in \Theta$ . Відповідно, за теоремою Байеса, яку називають ще принципом зворотної ймовірності, встановлюється взаємозв'язок між  $P(x|\theta)$ ,  $P(\theta|x)$  та  $P(\theta)$ :

$$P(\theta|x) = \frac{P(x|\theta)P(\theta)}{P(x)}, \quad \theta \in \Theta, \quad (2)$$

де

$$P(x) = \int_{\Theta} P(x|\theta)P(\theta)d\theta. \quad (3)$$

Враховуючи, що знаменник у формулі (2) не залежить від  $\theta$ , досить часто залежність (2) представляють у вигляді:

$$P(\theta|x) \propto P(x|\theta)P(\theta), \quad (4)$$

де  $\propto$  – означає пропорційність.

### 3. Вибір та аналіз апіорного розподілу

Метод визначення апіорного розподілу залежить від конкретної поставленої задачі. Відрізняють випадки, коли розподіл апіорної інформації відомий, та коли розподіл параметра є “неінформативним”. У першому випадку, якщо відомо вигляд функції апіорного розподілу, то, використовуючи байесівський підхід (рівності (2) або (4)), можна знайти вигляд апостеріорного розподілу.

Існують випадки, коли апіорний та апостеріорний розподіли відносяться до одного і того ж класу розподілів. Такі розподіли називають *спряженими*. Використання таких розподілів для чисельних розрахункових методів у байесівському підході означає, що існує замкнена форма розв'язку для умовних апостеріорних розподілів [14]. Як приклад, можна розглянути випадок, коли  $X_1, \dots, X_n$  розглядається як випадкова вибірка із нормального розподілу із невідомим середнім значенням  $\mu$  та відомою дисперсією  $\sigma^2$ . Припускається, що апіорний розподіл  $\mu$  є також нормальним розподілом із середнім значенням  $\mu_0$  та дисперсією  $\sigma_0^2$ . Тоді апостеріорний розподіл  $\mu$  при заданих вибіркових значеннях  $X_1, \dots, X_n$  та заданному апіорному розподілу також буде

нормально розподіленим із середнім значенням  $\mu_*$  та дисперсією  $\sigma_*^2$ , які визначають таким чином:

$$\mu_* = \frac{\sigma^2 \mu_0 + n \sigma_0^2 \bar{X}}{\sigma^2 + n \sigma_0^2} \text{ та } \sigma_*^2 = \frac{\sigma^2 \sigma_0^2}{\sigma^2 + n \sigma_0^2}, \quad (5)$$

де  $\bar{X} = \sum_{i=1}^n X_i / n$  – вибіркове середнє.

У байєсівському аналізі досить часто зручно використовувати параметр точності  $\eta = 1/\sigma^2$  (тобто обернений до дисперсії  $\sigma^2$ ). Якщо для апіорного розподілу  $\eta_0 = 1/\sigma_0^2$ , то для апостеріорного розподілу  $\eta_* = 1/\sigma_*^2$ ; тепер рівність (5) можна представити так:

$$\eta_* = \eta_0 + n\eta \text{ та } \mu_* = \frac{\eta_0}{\eta_*} \mu_0 + \frac{n\eta}{\eta_*} \bar{X}. \quad (6)$$

Таким чином, для отриманої нормальної випадкової вибірки, інформація про  $\mu$  міститься у вибірквому середньому  $\bar{X}$ , котре є достатньою статистикою для  $\mu$ . Точність описання розподілу визначається відношенням  $X : \eta/\sigma^2 = n\eta$ . Тобто точність описання апостеріорного розподілу визначається сумою двох компонент: точністю представлення апіорного розподілу і вибірквих даних, а апостеріорне середнє є зваженим середнім апіорного середнього та вибірквого середнього з ваговим коефіцієнтом, пропорційним точності. Наведений результат свідчить про те, що внесок апіорного розподілу зменшується із зростанням розміру вибірки  $n$ . Докладний опис отримання цього результату можна знайти в [1].

Досить часто розглядають так звану байєсівську асимптотику, тобто у випадку, коли  $n \rightarrow \infty$  для апостеріорного розподілу має значення тільки правдоподібність. Це дуже легко оцінити з рівності (4), якщо її представити так [2]:

$$P(\theta | x) \propto P(x | \theta)P(\theta) = P(\theta)e^{\ln P(x|\theta)}. \quad (7)$$

Якщо припустити, що апіорний розподіл  $P(\theta)$  та функція правдоподібності  $P(x|\theta)$  є невідродженими та мають неперервні похідні і  $P(x|\theta)$  має єдиний максимум  $\theta_{ml}$ , який є оцінкою максимальної правдоподібності, то  $\ln[P(x|\theta)]$  має порядок  $n$ , а  $P(\theta)$  не залежить від об'єму вибірки. Таким чином, інтуїтивно зрозуміло, що множник правдоподібності при великих значеннях об'єму вибірки буде домінувати в апіорному розподілі.

У багатьох випадках початкова інформація про значення оцінюваних параметрів зовсім невідома, тобто невідомий вигляд апіорного розподілу  $P(\theta)$ . Іншими словами, параметр  $\theta$  є “неінформативним”. Для такого випадку запропоновано два правила вибору апіорного розподілу, які охоплюють найбільш поширені випадки [12]. Він вважає, що у випадку існування параметра на скінченному інтервалі або на інтервалі від  $-\infty$  до  $+\infty$ , його апіорна ймовірність повинна вважатись рівномірно розподіленою. Якщо ж можна обґрунтувати, що параметр приймає значення на інтервалі від 0 до  $\infty$ , то ймовірність його логарифму слід вважати рівномірно розподіленою. Перше правило Джеффріса для представлення невизначеності значення формулюється таким чином:

$$P(\theta)d\theta \sim d\theta, \quad -\infty < \theta < +\infty, \quad (8)$$

тобто  $P(\theta) \sim const$ . Цей прямокутний розподіл (або ж функція щільності імовірності) є невластим, оскільки  $\int_{-\infty}^{\infty} P(\theta)d\theta = \infty$ . Відомо, що у випадку, коли  $-\infty < \theta < +\infty$  є достовірним твердженням, то для представлення ймовірності достовірної події замість 1 використовується  $\infty$ .

Друге правило Джеффріса відноситься до параметрів, природа яких дозволяє зробити припущення, що вони приймають значення від 0 до  $\infty$ . Він пропонує прийняти, по аналогії, рівномірний розподіл логарифма параметра. Тобто якщо  $\theta = \log \sigma$ , то апіорна функція щільності імовірності для  $\theta$  буде обрана відповідно у вигляді (8). Оскільки  $d\theta = d\sigma/\sigma$ , то (8) припускає використання рівності:

$$P(\sigma)d\sigma \sim \frac{d\sigma}{\sigma}, \quad 0 < \sigma < +\infty \quad (9)$$

в якості невластної функції щільності ймовірності, яка буде представляти невизначеність значення параметра  $\sigma$ . Таким чином, для неінформативного параметра, наприклад для дисперсії, апіорний розподіл досить часто задають у вигляді  $P(\sigma) \propto 1/\sigma$ .

Відзначимо таку важливу властивість, як чутливість апіорного розподілу. При виборі різних апіорних розподілів отримують різні апостеріорні розподіли. У такому випадку корисно оцінити ступінь впливу отриманих відмінностей. На практиці досить часто отриманий апостеріорний розподіл використовують надалі в якості апіорного, при цьому беручи його в деякий степінь  $\alpha$ , де  $0 < \alpha < 1$ . Для реальних практичних задач використовують навіть експертну думку щодо типу апіорного розподілу або деякі прийняті в літературі погодження та попередні публікації.

### 5. Обчислювальні методи байєсівського аналізу

Байєсівський підхід досить широко використовує інформацію про ймовірнісний розподіл параметрів. До кінця 80-х років минулого сторіччя в якості розрахункових методів для виведення байєсовських оцінок параметрів та їх апостеріорних розподілів використовували аналітичні методи: спряжені апіорні розподіли та апроксимацію [14]. З початку 90-х років завдяки стрімкому розвитку комп'ютерних технологій почали поширюватися зовсім нові методи обчислень, які ґрунтуються на безпосередньому генеруванні (моделюванні вибірки) необхідних вимірів за апостеріорним розподілом.

Генерування випадкових величин із заданим розподілом – сучасний підхід, який дає змогу працювати в умовах, коли існують асимптотичні аналітичні результати для властивостей оцінок та їх статистичних розподілів, але невідомо, які властивості будуть мати оцінки при малих вибірках даних. Виділяють два загальних підходи до моделювання: історичне моделювання та моделювання за принципом Монте-Карло [14]. Термін “метод Монте-Карло” (запропонований Джоном фон Нейманом та Станіславом Уламом у 1940-х рр.) відноситься до моделювання процесів із використанням генератора випадкових чисел. Сьогодні методи Монте-Карло утворюють захоплюючий розділ сучасної математики та статистичних чисельних експериментів. Назва методу з'явилась від назви міста, широко відомого своїми казино. Метод моделювання Монте-Карло виник від того факту, що поняття “кількість шансів” було використано з метою знаходження інтегралів від складних рівнянь, що використовувались при розробці перших ядерних бомб (інтеграли квантової механіки).

Інтеграли складних розподілів, які формувались із великих вибірок випадковими числами із декількох розподілів, апроксимувались згенерованими даними.

Методи моделювання Монте Карло поділяються на ітеративні та неітеративні. До неітеративного моделювання відносять метод генерування вибірки за важливістю (importance sampling) та метод відбраковки або прийняття вибірки (rejection or acceptance sampling) [5].

**5.1. Неітеративні методи Монте-Карло.** Розглянемо деякий ймовірнісний розподіл або ймовірнісну щільність розподілу  $p(x)$  для дискретного або неперервного процесу. Розподіл  $p(x)$  називають ще цільовим розподілом або цільовою щільністю розподілу. Необхідно згенерувати вибіркові значення  $\{X_i\}_{i=1}^N$  за розподілом  $p(x)$  та знайти їх статистичну оцінку, тобто оцінити математичне сподівання функції  $\phi(x)$  за розподілом, наприклад:

$$E_p[\phi] = \int \phi(x)p(x)dx. \quad (10)$$

Припустимо, що  $x$  – вектор із  $\mathbb{R}^n$  з компонентами  $X_i$ , а  $\phi(x)$  – деяка функція. Для вибіркових значень  $\{X_i\}_{i=1}^N$  можна розрахувати:

$$\Phi = \frac{1}{N} \sum_i \phi(X_i). \quad (11)$$

Оскільки вектор  $\{X_i\}_{i=1}^N$  згенеровано за  $p(x)$ , то математичне сподівання  $\Phi$  дорівнює  $E_p[\phi]$ ; при збільшенні кількості вибіркових значень  $N$  дисперсія  $\Phi$  пропорційно зменшується.

Моделювання вибірки за важливістю (importance sampling) – це не метод генерування вибіркових значень  $\{X_i\}_{i=1}^N$  за розподілом  $p(x)$ , а тільки метод оцінювання сподівання  $\phi(x)$ , представленого виразом (10). Основна ідея методу полягає у тому, що моделювання виконується за іншим, простішим для виконання розрахунків апроксимуючим розподілом, наприклад  $q(x)$ , який досить близький до цільового розподілу  $p(x)$ . Алгоритм моделювання складається із таких основних кроків:

- вибирають вибіркові значення  $X_i$ , згенеровані за  $q(x)$ ;
- з того, що моделювання відбувалось за “помилковим” розподілом формують ваговий коефіцієнт:

$$\omega_i = \frac{p(X_i)}{q(X_i)}; \quad (12)$$

- оцінку шуканої величини  $E_p[\phi]$  знаходять за виразами:

$$\Phi = \frac{1}{N} \sum_i \omega_i \phi(X_i) \quad \text{або} \quad \Phi = \frac{\sum_i \omega_i \phi(X_i)}{\sum_i \omega_i}. \quad (13)$$

Основним недоліком такого методу є те, що відносно зростання дисперсії зумовлене непостійними ваговими коефіцієнтами, наявна велика залежність від вибору апроксимуючого розподілу  $q(x)$ , а також те, що всі вибіркові значення будуть нівелюватись невеликою кількістю значень із великими ваговими коефіцієнтами.

Метод відбраковки або прийняття вибірки (rejection or acceptance sampling) полягає знову ж таки в моделюванні за іншим апроксимуючим розподілом, але перерахунок вагового коефіцієнта виконується в процесі генерування, утримуючи тільки частину змодельованих точок вимірів. Основний алгоритм генерування вибірки можна визначити такими кроками:

1. Відшукання масштабуючої константи  $M$ , такої що

$$\forall x, \quad p(x) \leq M q(x). \quad (14)$$

2. Незалежно вибирають вибіркові значення  $X_i$ , згенеровані по  $q(x)$  та  $U_i$  рівномірно розподілені на інтервалі  $[0,1]$ .

3. Якщо  $M U_i q(X_i) < p(X_i)$ , то приймають  $X_i$ , в іншому випадку значення  $X_i$  відкидається (відбраковується, звідси і назва методу) та повертаються до кроку 2.

4. Кроки 2-3 повторюють до тих пір, поки не буде знайдено  $N$  значень  $X_i$ .

Недоліком цього алгоритму, як і в попередньому випадку, є проблеми при великих розмірах вибірки; алгоритм виробляє “прийнятні” випадкові числа тільки за  $1/M$  часу, якщо розподіли не співпадають. Якщо  $M$  занадто велике, то багато змодельованих значень будуть відкидатись.

Таким чином, описані вище методи працюють задовільно тільки у тих випадках, коли обраний для генерування розподіл  $q(x)$  подібний до цільового розподілу  $p(x)$ . Однак, у більшості складних задач дуже важко створити окремий розподіл, який задовольняв би усім властивостям.

### 5.2. Ітераційні методи Монте-Карло для марковських ланцюгів

Ітераційні методи моделювання Монте-Карло ґрунтуються на ідеї побудови Марковського ланцюга та на відміну від попередніх методів використовують варіанти (кандидати) розподілу  $q(x)$  (щільності розподілу), які залежать тільки від поточного стану  $X_i$  [5, 13]. До цих методів можна віднести генерування вибірки за Гіббсом, алгоритм Метрополіса, Метрополіса-Хастінгса на інші їх модифікації.

Розглянемо стохастичний процес  $\{X_t\}$ , де  $X_t \in \Theta$ . Він марковський, якщо має ту властивість, що при заданому значенні  $X_t$ , значення  $X_h$  при  $h > t$  не залежать від значень  $X_s$  при  $s < t$ . Інакше кажучи,  $\{X_t\}$  – марковський процес, якщо його умовна функція розподілу задовольняє такій умові:

$$P(X_h / X_s, s \leq t) = P(X_h / X_t), \quad h > t. \quad (15)$$

Якщо  $\{X_t\}$  – дискретний стохастичний процес, то його основна характеристика має вигляд:

$$P(X_h/X_t, X_{t-1}, \dots) = P(X_h/X_t), \quad h > t. \quad (16)$$

Якщо  $A$  є підмножиною  $\Theta$ , то функція

$$P_t(\theta, h, A) = P(X_h \in A/X_t = \theta), \quad h > t \quad (17)$$

є функцією перехідної ймовірності марковського процесу. Якщо перехідна ймовірність залежить від  $h-t$ , а не від  $t$ , то процес має стаціонарну перехідну ймовірність.

Основою моделювання за допомогою марковського ланцюга є побудова марковського процесу, для якого стаціонарний розподіл переходів визначається функцією  $P(\theta/X)$ . Процес моделювання досить тривалий, він продовжується до тих пір, поки розподіл поточних значень процесу не наблизиться до стаціонарного розподілу переходів. Виходить, що для заданого  $P(\theta/X)$  може бути сконструйована велика кількість марковських ланцюгів із заданими параметрами. Методи, які використовують моделювання випадкових величин марковським ланцюгом з метою одержання розподілу  $P(\theta/X)$ , відносять до методів Монте-Карло для марковських ланцюгів (МКМЛ).

Метод генерування вибірки Гіббса представляє собою витончений спосіб формування вибірки із спільних розподілів багатовимірних змінних шляхом застосування багаторазових вибірок із визначених одновимірних умов. Наприклад, для двовимірної спільної щільності розподілу  $f(x, y)$  при розрахунку використовують умовні щільності розподілів  $f(x/y)$  та  $f(y/x)$ . Основний алгоритм генерування можна описати таким чином:

1. Вибір початкових значень  $X_0$  та  $Y_0$  у відповідності з прийнятим розподілом.
2. Генерування  $Y_{t+1}$  за умовним розподілом  $f(y/X_t)$ .
3. Генерування  $X_{t+1}$  за умовним розподілом  $f(x/Y_{t+1})$ .
4. Багатократне повторення кроків 2-3 для того, аби ланцюг досяг збіжності до свого стаціонарного розподілу.

Таким чином, одержуємо набір значень  $(X_t, Y_t)$ , який при  $t = 1, \dots, N, N \rightarrow \infty$  буде збігатись до спільного розподілу  $f(x, y)$ . На практиці використовують достатньо велику кількість вимірів  $N$ , при цьому відкидають перші  $M$  ( $M < N$ ) випадкових значень ітерацій Гіббса, які вважають зразками для випробувань на відмову (зазвичай  $M$  вибирають як 10-20 % від  $N$ ). Випробування на відмову використовується для забезпечення близькості вибірових значень до спільного розподілу  $f(x, y)$ . У загальному випадку збіжність методу Гіббса можна записати так:

$$\frac{1}{N-M} \sum_{t=M+1}^N g(X_t, Y_t) \rightarrow \int g(x, y) f(x, y) dx dy, \quad N \rightarrow \infty. \quad (18)$$

Зауважимо також, що розглянутий метод представляє собою проходження по довгому ланцюзі із збереженням всіх випадкових вимірів після випробувань на відмову для



одержання вибірових значень Гіббса. Крім того, можна запускати декілька відносно коротких ланцюгів, використовуючи різні початкові значення та відносно малі  $N$ . Випадкове значення останньої ітерації Гіббса в кожному ланцюзі використовується для формування поточного значення вибірки Гіббса.

З попередніх викладок можна зробити висновок, що вибірка Гіббса має перевагу в декомпозиції багатовимірної задачі оцінювання у декілька задач меншої розмірності шляхом застосування повних умовних розподілів параметрів. Ця особливість робить вибірку Гіббса простою і широко застосовною. Однак, досить часто недостатньо зводити всі значення вибірки Гіббса до одновимірної задачі. Коли параметри сильно корельовані, доводиться вимірювати їх спільно. Для того, аби переконатись у збіжності методу Гіббса досить часто процедуру повторюють декілька разів із різними початковими значеннями, щоб переконатись у збіжності алгоритму.

Інший метод – алгоритм Метрополіса застосовується, наприклад, у тих випадках, коли ймовірнісний розподіл відомий за винятком нормуючої константи. Припустимо, що необхідно одержати випадкову вибірку з розподілу із цільовою щільністю  $f(x)$ , що містить складну нормуючу константу; при цьому пряме одержання вибірки або трудомістке або нездійсненне. Згідно з процедурою вибирають апроксимуючий розподіл, для якого легко генерувати випадкові значення, тобто визначають кандидата на функцію щільності,  $q(v, x)$ . Алгоритм Метрополіса генерує послідовність випадкових вимірів з апроксимуючого розподілу  $q(v, x)$ , розподіл якої збігається до  $f(x)$ . Алгоритм можна представити такими кроками:

1. Вибрати щільність апроксимуючого розподілу  $q(v, x)$ .
2. Задати поточний стан ланцюга: вибираючи вибірове значення  $X_t$ , згенерувати значення  $V$  за розподілом  $q(v, X_t)$  (який називають ще “стрибокподібним”). Такий розподіл повинен бути симетричним, тобто  $q(v, x) = q(x, v)$  для всіх  $v$  та  $x$ .
3. Розрахувати показник пропускнуї спроможності (або значення стрибка):

$$r = \frac{f(V)}{f(X_t)}. \quad (19)$$

4. Якщо  $r \geq 1$ , то встановлюють  $X_{t+1} = V$ ; інакше, якщо  $r < 1$ , то

$$X_{t+1} = \begin{cases} V & \text{з ймовірністю } r \\ X_t & \text{з ймовірністю } 1 - r \end{cases} \quad (20)$$

Повторення алгоритму багато разів при деяких регулярних нежорстких умовах забезпечує збіжність послідовності  $\{X_t\}$  до розподілу  $f(x)$ .

Правило прийняття та відхилення згенерованого значення для цього алгоритму може ґрунтуватися на таких умовах: (i) якщо стрибок з  $X_t$  до  $V$  збільшує щільність розподілу, то приймається  $V$  як  $X_{t+1}$ ; (ii) якщо стрибок зменшує щільність розподілу, то покладають  $X_{t+1} = V$  з ймовірністю, рівною показнику  $r$ , і встановлюють  $X_{t+1} = X_t$  в іншому випадку з ймовірністю  $(1 - r)$ . Прикладами симетричної щільності  $q(v, x)$  можуть слугувати нормальний розподіл та розподіл  $t$ -Стюдента для середніх значень параметрів моделі.

У 1970 році Хастінгс узагальнив алгоритм Метрополіса двома шляхами. По-перше, стрибкоподібний розподіл, тобто апроксимуючу щільність розподілу, не обов'язково обирати симетричною. По-друге, правило стрибка змінюється таким чином:

$$r = \frac{f(V)/q(V, X_t)}{f(X_t)/q(X_t, V)} = \frac{f(V)q(X_t, V)}{f(X_t)q(V, X_t)}. \quad (21)$$

Таку модифікацію називають алгоритмом Метрополіса-Хастінгса. Аналогічним чином за допомогою алгоритму Метрополіса-Хастінгса можна генерувати випадкові виміри за будь-якою функцією щільності, при цьому симетричність розподілу не вимагається.

Як і у випадку вибірки Гіббса необхідно використовувати достатню кількість значень  $N$  та виключати з розгляду перші  $M$  значень. Алгоритм Метрополіса-Хастінгса можна представити так:

$$\frac{1}{N-M} \sum_{t=M+1}^N g(X_t) \rightarrow \int g(x)f(x)dx, \quad N \rightarrow \infty. \quad (22)$$

Для алгоритму Метрополіса-Хастінгса вибір апроксимуючої щільності розподілу є дуже важливим. Вона не повинна мати занадто велику або малу дисперсію, повинна бути достатньо близькою до цільового розподілу. Якщо показник пропускнуої спроможності занадто великий, то ланцюг буде виробляти занадто багато малих кроків в околі локальних викидів, збільшуючи кореляцію та час збіжності. Таким чином, незалежні вибіркові значення отримують лише із великими інтервалами. Якщо ж цей показник занадто низький, то ланцюг буде "застрягати" в окремих місцях. Оптимальним значенням для показника пропускнуої спроможності вважається 20-50 %.

Зауважимо, що індексами  $i=1\dots N$  ми позначали вище незалежні вибіркові значення, що генеруються за відповідним розподілом. Позначення  $t=1\dots N$  відповідають послідовностям станів Марковського ланцюга. Алгоритм Метрополіса-Хастінгса, як узагальнений варіант алгоритму Метрополіса та вибірки Гіббса, на відміну від методу відбраковки вибірки, при моделюванні  $N$  ітерацій не виробляє  $N$  незалежних вибіркових значень за цільовим розподілом, а самі вибіркові значення корелюють. Це пояснюється тим, що методи МКМЛ включають у себе марковський процес, у якому згенеровано послідовність станів  $\{X_t\}$ , кожен елемент якого  $X_t$  має ймовірність розподілу, що залежить від попереднього значення  $X_{t-1}$ . За значний проміжок часу марковський ланцюг повинен суттєво переміститись для того, щоб ефективно згенерувати незалежні вибіркові значення цільового розподілу.

Вибір функції пропозиції щільності дозволяє уникати вищеописаної кореляційної залежності між вибірковими значеннями завдяки деяким ідеям та прийомам. Зокрема, якщо обрати в якості пропозиції щільності  $q(v, x) = q(v)$ , то отримаємо незалежну вибірку. Якщо покласти  $q(v, x) = f(x-v)$ , то отримуємо алгоритм Метрополіса-Хастінгса із випадковим блуканням. Досить корисним буває застосування композицій та комбінацій, тобто використання декількох видів пропозицій розподілу, які вибирають випадковим чином. Наприклад, формують незалежну вибірку за апіорним розподілом з метою подальшого аналізу апостеріорного розподілу на великих кроках.

Для визначення нижньої границі кількості ітерацій, необхідних для того, щоб вибіркові значення були незалежними, для методу Метрополіса-Хастінгса досить часто застосовують емпіричне правило. Якщо найбільша довжина у просторі ймовірних

станів дорівнює  $L$ , то для отримання незалежних вибірових значень за алгоритмом Метрополіса-Хастінгса, де апроксимуючий розподіл моделюється випадковим блуканням із розміром кроку  $\varepsilon$ , він повинен виконуватись щонайменше  $T \cong (L/\varepsilon)^2$  ітерацій.

До переваг описаних вище методів МКМЛ відносять:

- задача, яка реалізується за методами МКМЛ, є легкою у розв'язанні у тому смислі, що для отримання результату необхідно прогнати цей метод визначене число разів;
  - використовуючи алгоритм Метрополіса-Хастінгса, можна уникнути розрахунку граничних розподілів та нормуючих констант у щільностях розподілу;
  - можна отримати функції параметрів  $\phi(\theta)$ , виходячи із імітованих (змодельованих) розподілів  $\phi(\theta_i)$ ;
  - можливе відтворення багатьох природних форм та інших нестандартних особливостей процесів та явищ, що моделюються.
- До недоліків методів МКМЛ відносять:
- нередукованість, аперіодичність та властивості апостеріорних розподілів залишатися встановленими аналітично;
  - діагностика збіжності методів все ще залишається малозрозумілою, можлива демонстрація того, що отриманий ланцюг не буде збігатись;
  - вибірові значення, отримані за методами МКМЛ, можуть бути досить сильно автокорельованими; тому необхідно мати великий розмір вибірки, порядку  $(10^3 - 10^6)$  згенерованих значень.

Подолання описаних недоліків вважається у значній мірі творчим процесом. Діагностика збіжності може включати в себе дослідження графіків та тестування рівності розподілів між різними частинами ланцюга. Автокореляції можуть бути дещо скореговані новою параметризацією та групуванням вибірових значень. Ланцюг можна “перезапустити”, тобто повторити виконання алгоритму, використовуючи різні початкові значення та/або різні генератори початкових випадкових чисел. Для перевірки збіжності можна також скористатись показником зменшення масштабу, який розраховується за виразом:

$$R = \sqrt{\frac{\text{загальна дисперсія}}{\text{дисперсія всередині ланцюга}}} \quad (23)$$

Якщо значення  $R$  наближається до 1, то це забезпечує збіжність.

Однак, питання збіжності методів МКМЛ все ще залишаються відкритими. На сьогодні багато дослідників займаються теоретичними дослідженнями збіжності цих методів.

## 6. Перевірка гіпотез та вибір моделі

У багатьох випадках виникає задача порівняння альтернативних гіпотез та моделей. При наявності точних формулювань гіпотез або моделей, що порівнюються, практичний спосіб їх порівняння залежить від цілі аналізу, стану апіорної інформації та від наявності функції втрат. Розглянемо критерії вибору альтернативних моделей у загальному випадку.

Розглядається множина параметричних моделей  $M_1, M_2, \dots, M_m$ , які описують вибірові дані  $X_1, \dots, X_n$  умовними розподілами  $f(x | \theta, M_i)$ ,  $i = 1..m$ . Граничний розподіл даних припускає, що:

$$p(x | M_i) = \int_{\Theta} f(x | \theta, M_i) P(\theta | M_i) d\theta. \quad (24)$$

Головним механізмом формування статистичного висновку є байєсівський фактор (коефіцієнт), який представляє собою відношення апостеріорних ймовірностей до апіорних:

$$BF(i, j) = \frac{P(M_i | x) / P(M_j | x)}{P(M_i) / P(\theta | M_j)} = \frac{P(x | M_i)}{P(x | M_j)}, \quad i, j = 1..m. \quad (25)$$

Перевага однієї моделі над іншою визначається у відповідності із значенням байєсівського фактора. Якщо воно суттєво перевищує 1, то приймається відповідне рішення. Для вибору моделі використовують критерій найбільшої граничної щільності розподілу  $p(x | M_i)$ , що відповідає умові  $BF(i, j) > 1$ .

Досить часто у розрахунках використовують апроксимацію (наближення) байєсівського фактора, а саме інформаційний критерій Байєса-Шварца (Schwarz Bayesian information criterion), який визначається так:

$$-2 \ln BF \approx \Delta BIC = -2 \ln \frac{\sup_{M_i} f(x | \theta, M_i)}{\sup_{M_j} f(x | \theta, M_j)} - (p_j - p_i) \ln n, \quad (26)$$

де  $n$  – величина вибірки;  $\dim M_i = p_j$ ,  $i, j = 1..m$  – розмірність моделі.

Байєсівські методи порівняння та вибору альтернативних моделей і гіпотез складають уніфіковану множину принципів, які вважають функціональними та застосовними до широкого класу задач. Такі методи дозволяють враховувати при порівнянні та перевірці гіпотез та моделей апіорну інформацію. Як показано у [1] байєсівський підхід до перевірки гіпотез та моделей є єдиним підходом, який передбачає вибір дії на основі максимізації очікуваної корисності.

## 7. Прогнозний розподіл

Найважливішим аспектом байєсівського підходу, як повноцінного підходу до формування статистичного висновку, є прогнозування. Байєсівська парадигма дає достатній об'єм інформації для прогнозування. Зокрема, головним інструментом виступає ймовірнісний розподіл, який використовується в якості прогнозного.

Якщо  $x^*$  – новий вимір, то можна записати сукупний розподіл  $x^*$  та параметрів  $\theta$  за умови заданої вибірки даних  $\{X_t\}$ :

$$P(x^*, \theta | x) = P(x^* | \theta, x) P(\theta | x). \quad (27)$$

Для отримання прогнозного розподілу  $P(x^* | x)$ , як умовної функції, необхідно проінтегрувати рівність (27) по  $\theta$ , яка і буде прогнозним розподілом. Спираючись на рівності (2) і (3) можна записати:

$$P(x^* | x) = \int_{\Theta} P(x^*, \theta | x) d\theta = \int_{\Theta} P(x^* | \theta) P(\theta | x) d\theta. \quad (28)$$

Можна побачити, що прогнозний розподіл  $P(x^* | x)$  не залежить від  $\theta$ . Рівність (28) вказує на те, що прогнозний розподіл (або прогноуюча функція розподілу ймовірності) розглядається як середнє умовних прогнозних розподілів  $P(x^* | \theta)$ , причому ваговою функцією слугує апостеріорний розподіл для  $\theta$ , тобто  $P(\theta | x)$ .

### 8. Приклади застосування МСМС методів до моделей стохастичної волатильності

Модель стохастичної волатильності Тейлора С. відноситься до класу моделей, в яких враховуються зміни дисперсії та коваріації. Для дискретного часу модель волатильності в достатньо простому варіанті має такий вигляд:

$$\begin{aligned} y_t &= \exp\left(\frac{h_t}{2}\right)\varepsilon_t, \text{ де } \varepsilon_t \sim \text{Norm}(0,1), \\ h_{t+1} &= \mu + \varphi(h_t - \mu) + \eta_t, \text{ де } \eta_t \sim \text{Norm}(0, \sigma_\eta^2), \end{aligned} \quad (28)$$

де  $y_t$  – середня відкоригована доходність активу на час  $t$ ,  $h_t$  – логарифм волатильності (мінливості) на момент часу  $t$ ,  $\text{Norm}(\cdot, \cdot)$  – нормальний розподіл. Параметр  $\mu$ , як середнє значення  $h_t$ , відіграє роль масштабуючого коефіцієнта;  $\varphi$  відображає стійкість волатильності,  $\sigma_\eta$  – волатильність логарифмованої змінної. Передбачається, що  $\varepsilon_t$  та  $\eta_t$  не корелюють між собою, а процес  $h_t$  – стаціонарний, ( $|\varphi| < 1$ ), та має початкову умову  $h_0 \sim \text{Norm}[\mu, \sigma_\eta^2 / (1 - \varphi^2)]$  [12].

Перше рівняння моделі (28) у просторі стану визначає умовні розподіли спостережень із заданими невідомими станами  $h_t$ , тобто це рівняння вимірів:

$$y_t | h_t = \exp\left(\frac{h_t}{2}\right)\varepsilon_t, \quad \varepsilon_t \sim \text{Norm}(0,1). \quad (29)$$

Невідомі стани визначаються відповідно до марковських переходів за часом та задаються відповідним рівнянням стану:

$$h_t | h_{t-1}, \mu, \varphi, \sigma_\eta^2 = \mu + \varphi(h_{t-1} - \mu) + \eta_t, \quad \eta_t \sim \text{Norm}(0, \sigma_\eta^2), \quad (30)$$

із початковою умовою  $h_0$ .

Байєсівська модель складається із спільного апріорного розподілу з усіх неспостережувальних змінних, трьох параметрів  $\mu, \varphi, \sigma_\eta^2$ , невідомих станів  $h_0, h_1, \dots, h_n$  та спільного розподілу спостережень, які представлені у вигляді даних щодо середньої доходності  $y_1, \dots, y_n$ . Байєсівський висновок ґрунтується на апостеріорному розподілі неспостережуваних параметрів при заданих даних, тобто на визначенні спільної ймовірності розподілу:

$$p(\mu, \varphi, \sigma_\eta^2, h_0, h_1, \dots, h_n) = p(\mu, \varphi, \sigma_\eta^2) p(h_0 | \mu, \varphi, \sigma_\eta^2) \prod_{t=1}^n p(h_t | h_{t-1}, \mu, \varphi, \sigma_\eta^2). \quad (31)$$

При цьому припускається незалежність апіорних розподілів параметрів  $\mu, \varphi$  та  $\sigma_\eta^2$ .

Для визначення апіорних розподілів обираємо неінформативний розподіл для  $\mu$ :  $\mu \sim Norm(0, 10)$ ; для  $\varphi$  покладемо  $\varphi = 2\varphi^* - 1$ , де  $\varphi^* \sim Beta(\alpha, \beta)$ , тобто це бета-розподіл із параметрами  $\alpha = 20$  та  $\beta = 1,5$ , які дозволяють отримувати значення  $\varphi$  у межах  $|\varphi| < 1$ . Для  $\sigma_\eta^2$  вибираємо спряжений зворотній гамма-розподіл із такими параметрами:  $\sigma_\eta^2 \sim InvGamma(5; 0,01)$ .

Для спільного розподілу (31) величину  $p(h_t | h_{t-1}, \mu, \varphi, \sigma_\eta^2)$  можна визначити за рівнянням стану (30). Правдоподібність  $p(y_1, \dots, y_n | \mu, \varphi, \sigma_\eta^2, h_0, h_1, \dots, h_n)$  визначається рівнянням вимірів (29) та припущеннями умов незалежності:

$$p(y_1, \dots, y_n | \mu, \varphi, \sigma_\eta^2, h_0, h_1, \dots, h_n) = \prod_{t=1}^n p(y_t | h_t). \quad (31)$$

Далі, за теоремою Байеса, сукупний апостеріорний розподіл неспостережувальних параметрів при заданих даних буде пропорційним до апіорного розподілу та правдоподібності, тобто:

$$p(\mu, \varphi, \sigma_\eta^2, h_0, \dots, h_n | y_1, \dots, y_n) \propto p(\mu) p(\varphi) p(\sigma_\eta^2) p(h_0 | \mu, \varphi, \sigma_\eta^2) \times \prod_{t=1}^n p(h_t | h_{t-1}, \mu, \varphi, \sigma_\eta^2) \prod_{t=1}^n p(y_t | h_t). \quad (32)$$

Розподіл для  $y_t$  можна розглядати як розподіл  $t$ -Стюдента із нульовим середнім, змінною дисперсією  $\sigma_t^2$  та невизначеними ступенями свободи  $\mathcal{G}$  для похибок спостережень:  $y_t \sim t - Student(0, \sigma_t^2, \mathcal{G})$ . Апіорний розподіл для  $\mathcal{G}$  можна обрати із розподілу Хі-квадрат за допомогою перетворення  $\mathcal{G} = \mathcal{G}^* + 2$ , де  $\mathcal{G}^* \sim \chi^2$  (8). Розподіл для  $y_t$  часто розглядають як нормальний розподіл із нульовим середнім та змінною дисперсією  $\sigma_t^2$ , тобто:  $y_t \sim Norm(0, \sigma_t^2)$ .

За допомогою даної моделі дослідимо стійкість ресурсної бази банку на прикладі аналізу поточних рахунків клієнтів. Необхідно проаналізувати волатильність клієнтських рахунків на можливість відтоку грошових коштів та подальшого визначення їх стійкості. Початкові дані, які представляють собою щоденні залишки грошових коштів клієнтів  $\{x_t\}$  за півтора роки (приблизно 370 значень), необхідно перетворити на часовий ряд, який буде вказувати середню відкориговану дохідність утримуемого активу, тобто середню мінливість залишків, за допомогою перетворення:

$$y_t = \log x_t - \log x_{t-1} - \frac{1}{n} \sum_{i=1}^n (\log x_i - \log x_{i-1}), t = 1, \dots, n. \quad (33)$$

На рис. 1 зображено динаміку часового ряду  $y_t$ .

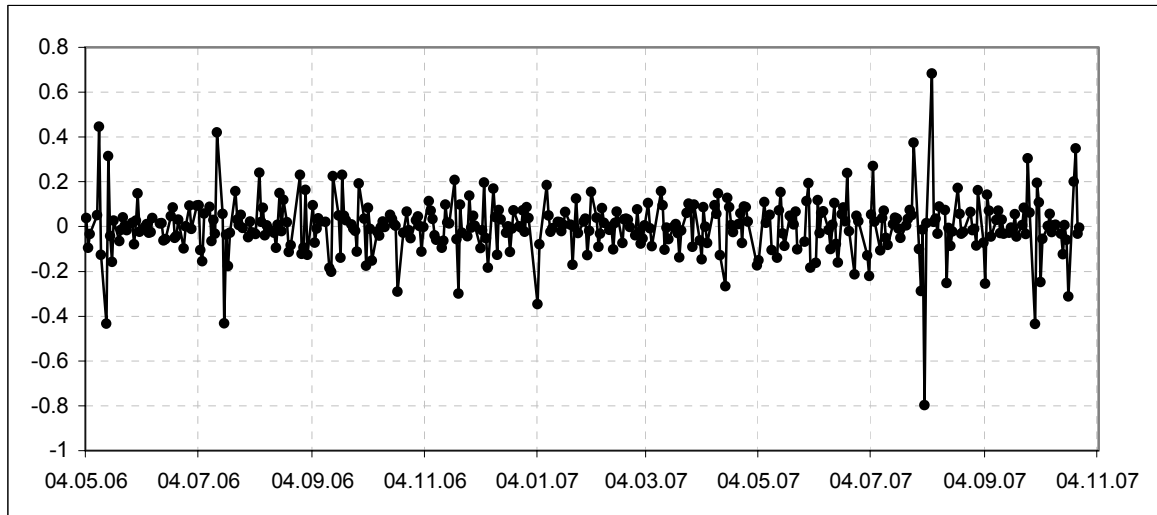


Рис. 1. Динаміка часового ряду  $y_t$

Для реалізації моделі використана програмна реалізація алгоритму Гіббса в системі WinBUGS (версія 1.4.3), яка є вільно доступним програмним продуктом. Результати роботи програми та алгоритму Вибірки Гіббса подано у таблиці 1.

Таблиця 1

**Байєсівське оцінювання параметрів моделі, обчислених за допомогою WinBUGS**

Параметр	Середнє значення	Стандартне відхилення	Похибка	2,5 %	Медіана	97,5 %	Початок уточн.	Об'єм даних
$1/\sigma_\eta^2$	622,1	230,2	21,24	232,6	595,6	1144,0	10001	10000
$\mu$	-4,486	0,1896	0,00804	-4,739	-4,511	-4,116	10001	10000
$\vartheta^*$	4,252	1,472	0,08974	2,146	3,97	7,966	10001	10000
$\varphi^*$	0,9881	0,006572	3,377E-4	0,973	0,989	0,9981	10001	10000

Значення, що відповідають 2,5 % і 97,5 % , утворюють довірчий інтервал для оцінок, а початок уточнення оцінок означає, що перші 10000 випадкових значень відхиляються; тобто фактичні оцінки формуються, починаючи з 10001-го згенерованого значення.

На рис. 2а та 2б представлені вибірки значень параметрів моделі та їх апостеріорні розподіли (вигляд функції щільності розподілу), отримані в результаті застосування методу МКМЛ.

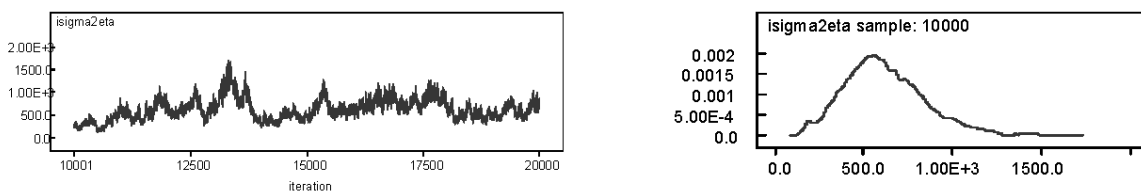


Рис. 2а. Вибірки значень параметрів моделі та їх апостеріорні розподіли

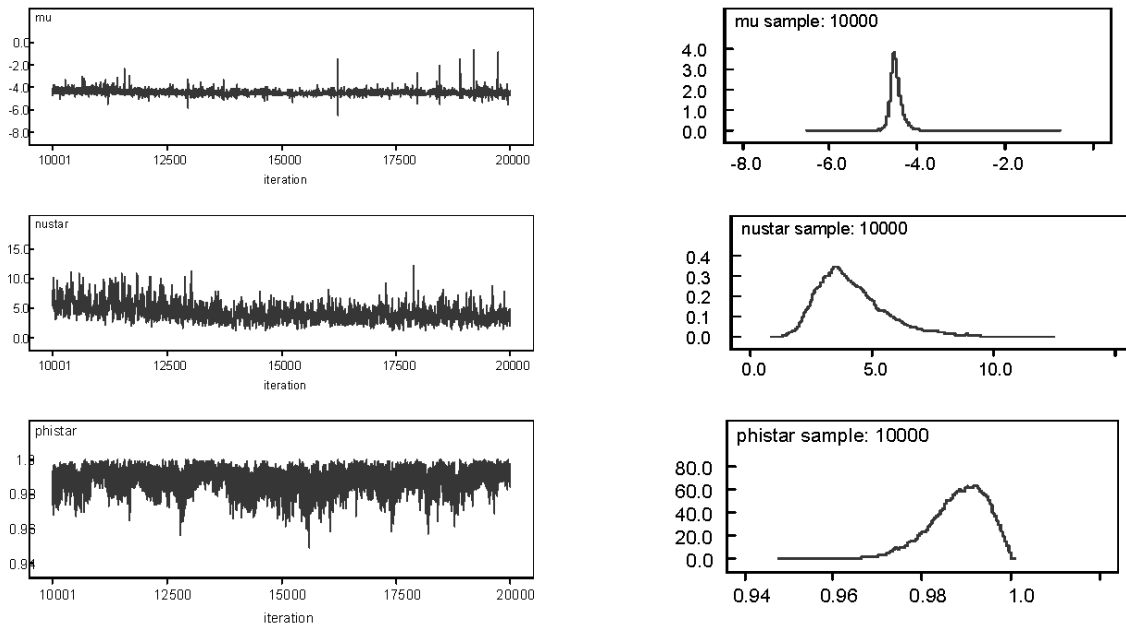


Рис. 26. Вибірки значень параметрів моделі та їх апостеріорні розподіли

## 9. Висновки

Застосування байєсівського підходу до формування статистичного висновку дозволяє зовсім по іншому сприймати та досліджувати оцінювані моделі. Він дає можливість оперувати не тільки з отриманими оцінками, а також із відповідними ймовірнісними розподілами, застосовувати наявні у різних формах апріорні знання дослідника щодо оцінок параметрів моделі. Це дає змогу отримувати більші об'єми вихідної інформації та чіткіше описувати структуру та інші характеристики досліджуваної моделі.

Методика застосування обчислювальних алгоритмів Монте-Карло, які ґрунтуються на генеруванні випадкових вимірів, тісно поєднана з байєсівською методологією. Цим розв'язується задача генерування вимірів за необхідним ймовірнісним розподілом. Описані методи значно розширюють та поліпшують можливості байєсівського аналізу та сфери його застосування. Однією з переваг такого підходу є створення та аналіз моделей за даними різної, зокрема малої, розмірності.

Розглянутий приклад показує, що для оцінювання таких складних нелінійних моделей, як модель стохастичної волатильності (подібні моделі іноді досить складно оцінювати за допомогою відомих класичних методів), дозволяє отримати прийнятні оцінки, оперуючи апріорними та апостеріорними розподілами параметрів та використовуючи алгоритми МКМЛ.

У цілому методи Монте-Карло для марковських ланцюгів трансформуються у відносно нескладні для програмування алгоритми, спрямовані на реалізацію методів байєсівського аналізу. При цьому варто зазначити, що методи даного класу ще потребують поглибленого аналізу та осмислення у майбутньому. Зокрема, необхідно дослідити точність оцінювання складних нелінійних моделей, збіжність та їх прогнозні характеристики, виконати порівняння з іншими, більш відомими методами.

## ЛІТЕРАТУРА

1. Зельнер А. Байесовские методы в эконометрии. – М.: Статистика, 1980. – 434 с.
2. Савчук В.П. Байесовские методы статистического оценивания: Надежность технических объектов. – М.: Наука, 1989. – 328 с.



3. Справочник по прикладной статистике /Под ред. Ллойда Э., Ледермана У. – М.: Финансы и статистика, 1989. – 525 с.
4. Bergman N. Recursive Bayesian Estimation: Navigation and Tracking Applications / Linkoping University (Sweden), 1999, No.579. – 219 p.
5. Besag J. Markov Chain Monte Carlo for Statistical Inference // Working Paper, Center for Statistics and the Social Sciences, 2001, No.9. – 25 p.
6. Carlin B.P., Louis T.A. Bayes and Empirical Bayesian Methods for Data Analysis. – London: Chapman and Hall, 1996. – 418 p.
7. Chib S., Nardari F., Shepard N. MCMC Methods for generalized SVM / TR OX1 1NF, Oxford, UK, 1998. – 24 p.
8. Dueker M. Kalman Filtering with Truncated Normal State Variables for Bayesian Estimation of Macroeconomic Models / Working Paper, 2005, No. 057B. – 55p.
9. Geweke, J., Tanizaki, H. Bayesian estimation of state-space models using M-H algorithm and Gibbs sampling // Computational Statistics and Data Analysis, 2001, v.37, No.2, pp. 151-170.
10. Бард Й. Нелинейное оценивание параметров. – М.: Статистика, 1979. – 349 с.
11. Kim S., Shephard N., Chib S. Stochastic volatility: Likelihood inference and comparison with ARCH models // Review of Economic Studies, 1998, 65, pp. 361- 393.
12. MacKay D. J.C. Information Theory, Inference, and Learning Algorithms – Cambridge: Cambridge University Press, 2003. – 640 p.
13. Nigel Da Costa Lewis. Market Risk Modeling. Applied Statistical Methods for Practitioners. – London: Risk Waters Group Ltd. , 2003. – 238 p.
14. Tsay R. S. Analysis of financial time series. – New York : Wiley & Sons, Inc. – 2002. – 448 p.