

Олевский В.И.^а, Олевская Ю.Б.^б

СТОХАСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ЭЛЕКТРОФОРЕЗА ВЫСОКОМОЛЕКУЛЯРНЫХ ВЕЩЕСТВ С ИСПОЛЬЗОВАНИЕМ ПРОЦЕССА ОРНШТЕЙНА-УЛЕНБЕКА

^а ГВУЗ «Украинский государственный химико-технологический университет», г. Днепр, Украина^б НТУ «Днепропетровская политехника», г. Днепр, Украина

Целью работы является разработка статистической модели в виде трендового процесса Орнштейна-Уленбека для описания электрофореза высокомолекулярных веществ. Способность электрофореза разделять заряженные биологические макромолекулы широко применяется в научных исследованиях, химическом производстве и медицине. Несмотря на применение для оценки коэффициентов электрофоретической подвижности молекулярно-кинетической теории в виде стохастического уравнения Ланжевена и выходящих из него соотношений Эйнштейна-Смолуховского, учитывающей броуновское движение частиц среды, эти модели остаются детерминистическими. Такие модели не позволяют достичь необходимой точности при описании электрофореза высокомолекулярных веществ, особенно при анализе их различий. Процесс Орнштейна-Уленбека является единственным нетривиальным стационарным гауссовым марковским процессом и имеет свойство возвращения к среднему. Это способствовало его использованию в финансовой инженерии. В настоящее время процесс Орнштейна-Уленбека изучен достаточно глубоко и представляют интерес его различные модификации и обобщения, в частности трендовый процесс Орнштейна-Уленбека. Трендовый процесс Орнштейна-Уленбека описывает стохастический процесс, который отклоняется от детерминированного тренда и возвращается к нему со скоростью, пропорциональной отклонению. Детерминированный тренд содержит, как правило, несколько параметров, которые необходимо оценивать по данным наблюдений. В известных работах по моделированию движения диэлектриков такую оценку проводят известными статистическими методами, например, моментов или максимума правдоподобия. В случае электрофореза предлагается в качестве детерминированного тренда принять линейный тренд с коэффициентом, пропорциональным электролитической подвижности вещества в заданной среде. Статистическая обработка экспериментальных данных дает возможность получить оценки параметров, что позволяет стохастически моделировать процесс электрофореза. Определена форма модели электрофореза, разработана методика расчета параметров полученной модели по данным эксперимента по разделению белков.

Ключевые слова: процесс Орнштейна-Уленбека, электрофорез, высокомолекулярные вещества, белки.

DOI: 10.32434/2521-6406-2018-4-2-60-64

Постановка проблемы и её связь с важными научными и практическими задачами

Электрофорез — это движение коллоидных частиц в электрическом поле. Это явление впервые описано Рейссом в 1809 году [1]. В 1930-х годах электрофорез был впервые применён в

биохимии Тизелиусом для разделения белков сыворотки крови. С тех пор электрофорез занимает важное место при исследовании высокомолекулярных веществ, важнейшими из которых являются белки и нуклеиновые кислоты. Метод позволяет разделять макромолекулы по

размерам и пространственной конфигурации, что делает его удобным для количественного анализа таких веществ. Способность электрофореза разделять заряженные биологические макромолекулы широко применяется не только в исследованиях, но и в биотехнологии, химическом производстве и в медицине.

Существующие методы математического описания электрофореза опираются на модели движения круглых тел в электростатическом поле с учетом сопротивления среды и электролитической проницаемости частицы [1]. Несмотря на применение для оценки коэффициентов электрофоретической подвижности молекулярно-кинетической теории, учитывающей броуновское движение частиц среды в виде стохастического уравнения Ланжевена и следующих из него соотношений Эйнштейна-Смолуховского, эти описания остаются детерминистическими. Такие описания не позволяют достичь необходимой точности при описании электрофореза некоторых высокомолекулярных веществ, особенно при распознавании их отличий друг от друга [2].

Вместе с тем, для описания систем, имеющих некоторый временной тренд и малую случайную составляющую, в последнее время успешно используются стохастические дифференциальные уравнения. Для описания таких процессов активно используется процесс Орнштейна-Уленбека (ОУ) [3–8]. Процесс ОУ обладает значительным потенциалом моделирования временных рядов наблюдений для широкого спектра областей таких как: техника, экономика и финансы. Многие исследователи на финансовых рынках широко использовали процесс ОУ для описания динамики переменных состояния, которые стабилизируются на своих равновесных уровнях по мере приближения времени к бесконечности. Так, например, О. Васичек [8] использовал процесс ОУ для описания стохастического движения краткосрочной процентной ставки. Использование трендового процесса ОУ дает возможность учесть стохастический характер явления и определить параметры тренда не теоретически, а статистическими методами на основании экспериментальных данных, что позволяет существенно улучшить оценки параметров модельного прогноза.

Анализ положивших начало решению проблемы публикаций, нерешённых частей проблемы

В 1903 году Смолуховский привел теоретическую формулу электрофореза для тонкого двойного электрического слоя, чем были заложены

теоретические основы расчета линейного коллоидного электрофореза. В 1905 году Эйнштейн и в 1906 году Смолуховский сформулировали законы, описывающие броуновское движение коллоидных частиц, и связали константу диффузии с электрофоретической подвижностью. В 1923 году Хюккель и Дебай создали теорию экранирования электростатических взаимодействий в растворах электролитов и ввели параметр длины экранирования (или длину Дебая). В терминах Дебая и Хюккеля теория Смолуховского описывает электрофорез в концентрированных растворах электролита, где параметр экранирования велик. В 1924 году Хюккель привел теоретическое решение для задачи об электрофорезе в разбавленном растворе электролита для малых чисел Дебая. С развитием численных методов в конце прошлого века задача была решена для произвольных наборов параметров. Аналитические методы не дают достаточной точности при численной обработке результатов экспериментов, но позволяет выбрать форму тренда стохастической зависимости.

Процесс Орнштейна-Уленбека впервые был представлен в качестве модели скоростей частиц в процессе соударения с окружающими их частицами. Процесс ОУ является единственным нетривиальным стационарным гауссовским марковским процессом [4,6,7] и обладает свойством возвращения к среднему. Эти свойства способствовали его использованию в финансовой инженерии. Так, например, в работе О. Васичека [8] представлена модель для оценки мгновенной процентной ставки, а в начале нынешнего века были описаны способы использования процесса ОУ в задачах ценообразования опционов, оптимизации портфеля и теории рисков [9,10]. В настоящее время процесс ОУ изучен достаточно глубоко и представляют интерес его различные модификации и обобщения, в частности трендовый процесс ОУ.

Формулирование целей статьи

Целями этой работы является:

- 1) определение формы модели электрофореза, основанной на использовании трендового процесса Орнштейна-Уленбека;
- 2) разработка методики расчета параметров полученной модели по данным эксперимента по разделению белков в агарозном геле.

Таким образом, задачей настоящей работы является применение статистической модели в виде трендового процесса Орнштейна-Уленбека для описания электрофореза высокомолекулярных веществ.

Изложение основного материала исследования

Процесс Орнштейна-Уленбека (ОУ) является непрерывным стохастическим процессом v_t , описываемым стохастическим дифференциальным уравнением (SDE) вида

$$dv_t = k(\theta - v_t)dt + \sigma dW_t, \tag{1}$$

где θ , $k > 0$ и $\sigma > 0$ – параметры, интерпретирующие равновесный уровень, скорость достижения уровня равновесия и дисперсию переменной состояния v_t соответственно. Стандартное броуновское движение W_t рассматривается в вероятностном пространстве $(\Omega; \mathcal{F}; P)$ с фильтрацией $(\mathcal{F}_t)_{t \geq 0}$.

В процессе ОУ (1) v_t стремится к постоянному долгосрочному среднему уровню θ при стремлении времени к бесконечности. Если же процесс зависит от времени, это уравнение можно обобщить на детерминированный тренд $\mu(t)$, введя непрерывный стохастический процесс p_t , описываемый SDE вида

$$dp_t - \mu(t) = k(\mu(t) - p_t)dt + \sigma dW_t. \tag{2}$$

Уравнение (2) описывает процесс p_t , отклоняющийся от детерминированного тренда $\mu(t)$ и возвращающегося к нему со скоростью, пропорциональной отклонению. Процесс (2) называется трендовым процессом ОУ с детерминированным трендом $\mu(t)$.

Детерминированный тренд $\mu(t)$ содержит, как правило, несколько параметров, которые необходимо оценивать по данным наблюдений [3–7]. В известных работах по моделированию движения диэлектриков такую оценку проводят известными статистическими методами, например, моментов или максимума правдоподобия [6,7].

Процесс электрофореза при детерминистическом описании характеризуется постоянной скоростью движения частиц – так называемой электролитической подвижностью [1,2]. В случае электрофореза предлагается в качестве детерминированного тренда принять линейный тренд с коэффициентом, пропорциональным электролитической подвижности вещества в заданной среде α [1] вида $\mu(t) = k_1 \alpha t$. Тогда SDE (2) согласно стохастическим уравнениям Ито имеет решение вида

$$p_t = k_1 \alpha t + p_0 e^{-kt} + \sigma \int_0^t e^{k(t-u)} dW_u. \tag{3}$$

При произвольном p_0 асимптотическое распределение p_t является гауссовым со средним значением $k_1 \alpha t$ и дисперсией $\sigma_1^2 = \sigma^2 / 2k$.

Распределение плотности вероятности перехода для этого процесса в произвольный момент времени согласно уравнению Фоккера-Планка имеет вид

$$p(x, t | y, s) = \sqrt{\frac{\varphi}{\pi}} e^{-\varphi \psi(x, t | y, s)}, \tag{4}$$

где $\varphi = \frac{e}{4\sigma_1^2}$,

$$\psi(x, t | y, s) = \left((k_1 \alpha t - x) + (k_1 \alpha s - y) e^{k(s-t)} \right)^2.$$

Статистическая обработка экспериментальных данных дает возможность получить оценки параметров, что позволяет стохастически моделировать процесс электрофореза. Данные электрофореза (рис. 1) согласно разработанной нами методики [2] могут быть представлены как выборка относительных частот контрастности для ячеек разбиения изображения полосы и приближают экспериментальную функцию распределения вероятности (рис. 2).

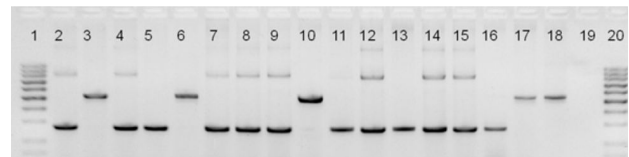


Рис. 1. Электрофореграмма ДНК кукурузы в полиакриламидном геле

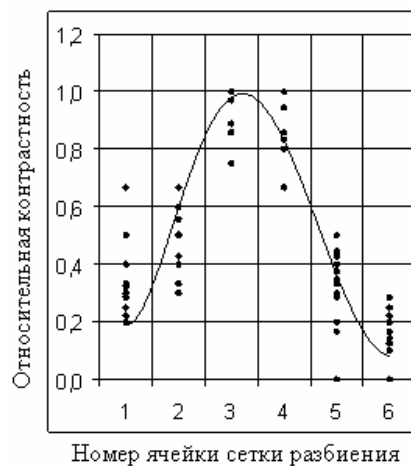


Рис. 2. Примерное распределение контрастности по продольным срезам полосы (сплошная линия – приближение кривой нормального распределения)

Для фиксированного времени t значения $k_1\alpha$ и σ_1^2 определяются как среднее и оценочная дисперсия выборки, что достаточно для анализа результатов электрофореза.

В случае необходимости рассмотрения траекторий движения частиц вещества может быть применена методика [7].

Выводы

Предложен способ построения статистической модели в виде трендового процесса Орнштейна-Уленбека для описания электрофореза высокомолекулярных веществ. Определена форма модели электрофореза. Разработана методика расчета параметров полученной модели по данным эксперимента по разделению белков.

СПИСОК ЛИТЕРАТУРЫ

1. *Theory of DNA electrophoresis* / G.W. Slater, S. Guillouzi, M.G. Gauthier, J.F. Mercier, M. Kenward, L.C. McCormick, F. Tessier // *Electrophoresis*. – 2002. – Vol.23. – P.791-816.
2. *Olevskiy V.I., Smetanin V.T., Olevska Yu.B.* Fuzzy method of recognition of high molecular substances in evidence-based biology // *AIP Conference Proceedings*. – 2017. – Vol.1895. – PP.070003-1-070003-9. DOI:10.1063/1.5007392
3. *Lasky P., Ditlevsen S.* A review of the methods for signal estimation in stochastic diffusion leaky integrate-and-fire neuronal models // *Biol. Cybern.* – 2008. – Vol.99. – P.253-262.
4. *Lo A., Wang J.* Implementing option pricing models when asset returns are predictable // *J. Financ.* – 1995. – Vol.1. – P.87-129.
5. *Lucia J., Schwartz E.* Electricity prices and power derivatives: evidence from the Nordic power exchange // *Rev. Deriv. Res.* – 2002. – Vol.5. – P.5-50.
6. *Rujivan S.* A closed-form formula for the conditional moments of the extended CIR process // *J. Comput. Appl. Math.* – 2016. – Vol.297. – P.75-84, DOI: 10.1016/j.cam.2015.11.001.
7. *Deng Y., Barros A., Grall A.* Residual useful life estimation based on a time-dependent Ornstein-Uhlenbeck process // *Chem. Engineer. Trans.* – 2013. – Vol.33. – P.325-330.
8. *Vasicek O.* An equilibrium characterization of the term structure // *J. Financ. Econ.* – 1977. – Vol.5. – P.177-188.
9. *Wilhelm M., Winter C.* Finite element valuation of swing options // *J. Comput. Financ.* – 2008. – Vol.11. – P.107-132.
10. *Zhang B., Grzelak L., Oosterlee C.* Efficient pricing of commodity options with early-exercise under the Ornstein-Uhlenbeck process // *Appl. Numer. Math.* – 2012. – Vol.62. – P.91-111.

Поступила в редакцию 02.11.2018

СТОХАСТИЧНЕ МОДЕЛЮВАННЯ ЕЛЕКТРОФЕРЕЗУ ВИСОКОМОЛЕКУЛЯРНИХ РЕЧОВИН З ВИКОРИСТАННЯМ ПРОЦЕСУ ОРНШТЕЙНА-УЛЕНБЕКА

Олевський В.І., Олевська Ю.Б.

Метою роботи є розробка статистичної моделі у вигляді трендового процесу Орнштейна-Уленбека для опису електрофорезу високомолекулярних речовин. Здатність електрофорезу розділяти заряджені біологічні макромолекули широко застосовується в наукових дослідженнях, хімічному виробництві та медицині. Незважаючи на застосування для оцінювання коефіцієнтів електрофоретичної рухливості молекулярно-кінетичної теорії у вигляді стохастичного рівняння Ланжевена і виходячих з нього співвідношень Ейнштейна-Смолуховського, що враховує броунівський рух частинок середовища, ці моделі залишаються детерміністичними. Такі моделі не дозволяють досягти необхідної точності при описі електрофорезу високомолекулярних речовин, особливо при аналізі їх відмінностей. Процес Орнштейна-Уленбека є єдиним нетривіальним стаціонарним гаусовим марковським процесом і має властивість повернення до середнього. Це сприяло його використанню у фінансовій інженерії. В даний час процес Орнштейна-Уленбека вивчений досить глибоко і становлять інтерес його різні модифікації і узагальнення, зокрема трендовий процес Орнштейна-Уленбека. Трендовий процес Орнштейна-Уленбека описує стохастичний процес, який відхиляється від детермінованого тренду і повертається до нього зі швидкістю, пропорційною відхиленню. Детермінований тренд містить, як правило, кілька параметрів, які необхідно оцінювати за даними спостережень. У відомих роботах з моделювання руху діелектриків таку оцінку проводять відомими статистичними методами, наприклад, моментів або максимуму правдоподібності. У разі електрофорезу пропонується в якості детермінованого тренду прийняти лінійний тренд з коефіцієнтом, пропорційним електролітичній рухливості речовини в заданому середовищі. Статистична обробка експериментальних даних дає можливість отримати оцінки параметрів, що дозволяє стохастично моделювати процес електрофорезу. Визначено форму моделі електрофорезу, розроблена методика розрахунку параметрів отриманої моделі за даними експерименту з розділення білків.

Ключові слова: процес Орнштейна-Уленбека, електрофорез, високомолекулярні речовини, білки.

STOCHASTIC MODELING OF HIGH-MOLECULAR SUBSTANCES ELECTROPHORESIS USING THE ORNSTEIN-UHLENBECK PROCESS

Olevsky V.I.^a, Olevskaya Yu.B.^b,

^a Ukrainian State University of Chemical Technology, Dnipro, Ukraine

^b Dnipro Polytechnic, Dnipro, Ukraine

The aim of the work is to develop a statistical model in the form of the Ornstein-Uhlenbeck trend process for describing the electrophoresis of high-molecular substances. The ability of electrophoresis to divide charged biological macromolecules is widely used in scientific research, chemical production and medicine. Despite the use of the molecular-kinetic theory for estimation of the coefficients of the electrophoretic mobility in the form of the stochastic Langevin equation and the Einstein-Smoluchowski relations emerging from it, which takes into account the Brownian motion of particles of the medium, these models remain deterministic. Such models do not allow one to achieve the required accuracy in the description of electrophoresis of high-molecular substances, especially when analyzing their differences. The Ornstein-Uhlenbeck process is the only non-trivial stationary Gaussian Markov process and has the property of returning to the average. This contributed to its use in financial engineering. At present, the Ornstein-Uhlenbeck process has been studied quite deeply and various modifications and generalizations are of interest, in particular the Ornstein-Uhlenbeck trend process. The Ornstein-Uhlenbeck trend process describes a stochastic process that deviates from the deterministic trend and returns to it with a speed proportional to the deviation. Deterministic trend, as a rule, contains several parameters that must be estimated from observational data. In well-known works on modeling the motion of dielectrics, such an estimate is carried out by well-known statistical methods, for example, by method of moments or maximum likelihood estimation. In the case of electrophoresis, we propose to adopt a linear trend with a coefficient proportional to the electrolytic mobility of the substance in a given environment as a deterministic trend. Statistical processing of experimental data makes it possible to obtain parameter estimates, which allows stochastically simulating the process of electrophoresis. The shape of the electrophoresis model has been determined, and methods have been developed for calculating the parameters of the obtained model according to the data of an experiment on the separation of proteins.

Keywords: Ornstein-Uhlenbeck process, electrophoresis, high-molecular substances, proteins.

REFERENCES

1. Slater G.W., Guillouzi, S., Gauthier M.G., Mercier J.F., Kenward M., McCormick L.C., Tessier F. Theory of DNA electrophoresis. *Electrophoresis*, 2002, vol.23, pp.791-816.
2. Olevskiy V.I., Smetanin V.T., Olevska Yu.B. Fuzzy method of recognition of high molecular substances in evidence-based biology. *AIP Conference Proceedings*, 2017, vol.1895, pp.070003-1–070003-9. DOI: 10.1063/1.5007392.
3. Lasky P., Ditlevsen S. A review of the methods for signal estimation in stochastic diffusion leaky integrate-and-fire neuronal models. *Biol. Cybern.*, 2008, vol.99, pp.253-262.
4. Lo A., Wang J. Implementing option pricing models when asset returns are predictable. *J. Financ.*, 1995, vol.1, pp.87-129.
5. Lucia J., Schwartz E. Electricity prices and power derivatives: evidence from the Nordic power exchange. *Rev. Deriv. Res.*, 2002, vol.5, pp.5-50.
6. Rujivan S. A closed-form formula for the conditional moments of the extended CIR process. *J. Comput. Appl. Math.*, 2016, vol.297, pp.75-84, DOI:10.1016/j.cam.2015.11.001.
7. Deng Y., Barros A., Grall A. Residual useful life estimation based on a time-dependent Ornstein-Uhlenbeck process. *Chem. Engineer. Trans.*, 2013, vol.33, pp.325-330.
8. Vasicek O. An equilibrium characterization of the term structure. *J. Financ. Econ.*, 1977, vol.5, pp.177-188.
9. Wilhelm M., Winter C. Finite element valuation of swing options. *J. Comput. Financ.*, 2008, vol.11, pp.107-132.
10. Zhang B., Grzelak L., Oosterlee C. Efficient pricing of commodity options with early-exercise under the Ornstein-Uhlenbeck process. *Appl. Numer. Math.*, 2012, vol.62, pp.91-111.