

УДК 519.213

О. В. Бурдукова, аспірант,
Ю. Г. Лега, д.т.н., професор,
О. С. Гавриш, к.ф.-м.н., доцент,
Т. В. Воробкало, к.т.н., доцент,
А. О. Іващенко, аспірант

Черкаський державний технологічний університет
б-р Шевченка, 460, Черкаси, 18006, Україна

АПРОКСИМАЦІЯ РОЗПОДІЛУ ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ МОДЕЛЯМИ БЛИЗЬКИХ ДО ГАУССІВСЬКИХ ВИПАДКОВИХ ВЕЛИЧИН

В роботі отримано аналітичні вирази апроксимуючих функцій, отриманих на основі моделей з перфорованим кумулянтним описом. Побудовано щільності розподілу різних моделей близьких до гауссівських випадкових величин, проведено порівняння з емпіричною щільністю розподілу і знайдено значення похибок апроксимації для кожного класу моделі. Показано, що моделі на основі перфорованого кумулянтного опису є ефективним інструментом для апроксимації реальних статистичних даних різної природи.

Ключові слова: апроксимація, ряд Еджворта, кумулянтні коефіцієнти, близькі до гауссівських випадкові величини, перфорація кумулянтного опису.

Вступ. При обробці статистичних даних часто зручно використовувати гауссівську модель випадкового процесу [1, 2]. Проте таке припущення інколи може приводити до помітного погіршення результатів опрацювання випадкового процесу, оскільки підібрана модель не є оптимальною. Використання стандартних негауссівських законів розподілу також не завжди повно і адекватно описують реальний процес або ж натикається на математичні труднощі подальшого опрацювання. Як компромісний варіант в роботі [3] вводяться так звані близькі до гауссівських випадкові величини (БГВВ), які описуються певним набором кумулянтних коефіцієнтів до 6-го порядку. Очевидно, що обмеження числа кумулянтів накладає обмеження на закон розподілу, тому для розв'язку практичних задач необхідно підібрати певний клас моделі БГВВ [4]. Існують різні підходи для знаходження аналітичного представлення щільності розподілу, серед яких розрізняють топографічні та апроксимативні [5]. Топографічні методи засновані на використанні розподілів Пірсона або Джонсона і враховують лише перші 4 кумулянти, що накладає обмеження на представлення деяких класів БГВВ. Для апроксимації щільності розподілу використовуються комбінації простих і достатньо добре досліджених функцій. Поширення набули полігауссівські суміші розподілів [6], які дозволяють

ефективно описувати довільні розподіли, проте можуть виникати складнощі, пов'язані з отриманням оцінок параметрів суміші і аналізом їх властивостей [5]. Щільність апроксимуючого розподілу може бути задана у вигляді усіченого ряду Еджворта [4, 7], вагові коефіцієнти якого залежать від заданого набору кумулянтних коефіцієнтів у відповідності з вибраним класом БГВВ.

Постановка задачі. Для дослідження можливості використання моделей з перфорованим кумулянтним описом для апроксимації закону розподілу використовуються статистичні дані різної природи, закон розподілу яких відрізняється від нормального. Для знаходження апроксимуючої функції і визначення оптимального класу БГВВ розглянуто два масиви вибірових значень:

- середня тривалість життя (років) в країнах світу, згідно даних ООН з 1980 року по наш час [8];
- показники контролю якості на кондитерському виробництві (індекс жирності) [9].

Для знаходження похибки апроксимації щільності розподілу випадкової послідовності моделями БГВВ необхідно отримати дві апроксимації функції щільності ймовірності:

- графічну на основі гістограми, з врахуванням реальних вибірових даних;
- у вигляді аналітичної функції (усіченого ряду Еджворта) з певним (обмеженим) на-

бором кумулянтних коефіцієнтів відповідно до заданого класу БГВВ.

$$p_{\theta}(x) = p_n(x) + \sum_{k=3}^{\infty} (-1)^k \frac{\beta_k}{k!} p_n^{(k)}(x) \quad (1)$$

де $p_n(x)$ – щільність ймовірності нормального розподілу; $p_n^{(k)}(x)$ – похідні k -го порядку нормального розподілу; β_k – квазімоменти розподілу $p_{\theta}(x)$, які пов'язані з кумулянтами κ_k співвідношеннями [7]

$$\begin{aligned} \beta_3 &= \kappa_3, & \beta_4 &= \kappa_4, & \beta_5 &= \kappa_5, \\ \beta_6 &= \kappa_6 + 10\kappa_3^2, & \beta_7 &= \kappa_7 + 35\kappa_3\kappa_4, & (2) \\ \beta_8 &= \kappa_8 + 56\kappa_3\kappa_5 + 35\kappa_4^2, \end{aligned}$$

У виразі (2) відмінними від нуля будуть лише інформативні кумулянти, які відповідають заданому класу БГВВ.

Похибка апроксимації щільності розподілу моделями БГВВ може бути представлена як квадратичне відхилення центрованих нормованих щільності ймовірностей [4, 5] виду:

$$\varepsilon_{\theta} = \int_{-\infty}^{\infty} [p(x) - p_{\theta}(x)]^2 dx. \quad (3)$$

Якість підбору відповідного класу моделі БГВВ визначається найменшою розбіжністю між значеннями точності (3) апроксимації щільності реальної випадкової послідовності $p(x)$ різними моделями $p_{\theta}(x)$, в тому числі і гауссівською. Для уніфікації і зручності порівняння точності апроксимації різними моделями БГВВ їх порівнюють не між собою, а з точністю апроксимації вихідного розподілу гауссівською моделлю з щільністю розподілу $p_n(x)$

$$\varepsilon_n = \int_{-\infty}^{\infty} [p(x) - p_n(x)]^2 dx. \quad (4)$$

В цьому випадку, величина:

$$\varepsilon_a = \left| \frac{\varepsilon_{\theta}}{\varepsilon_n} \right| \quad (5)$$

характеризує наскільки ефективно (при $\varepsilon_a \rightarrow 0$) перфорована модель описує вихідний розподіл в порівнянні з гауссівською моделлю.

Метою роботи є підбір оптимальних моделей БГВВ для апроксимації експериментальних даних різної природи і знаходження кількісних показників якості апроксимації.

Результати. Розглянемо результати апроксимації вибіркового розподілу тривалості життя у світі. Використовуючи статистичні дані

[8] побудуємо гістограму розподілу вибіркового значень, а наступним кроком отримаємо емпіричну щільність розподілу шляхом апроксимації гістограми (рис. 1).

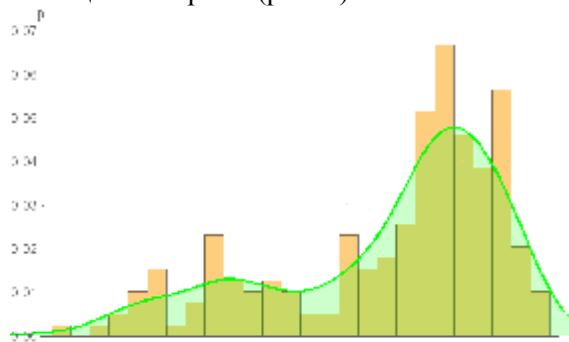


Рис. 1. Графічне представлення емпіричної щільності розподілу тривалості життя

Опишемо вихідну щільність гауссівською моделлю. Для цього необхідно отримати значення математичного сподівання та середньоквадратичне відхилення вибіркового значень

$$\begin{aligned} \mu &= \frac{1}{n} \sum_{i=1}^n x_i = 67,25, \\ \sigma &= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2} = 11,86. \end{aligned}$$

Використовуючи розраховані дані отримаємо нормальну щільність розподілу для вихідної щільності (рис. 2).

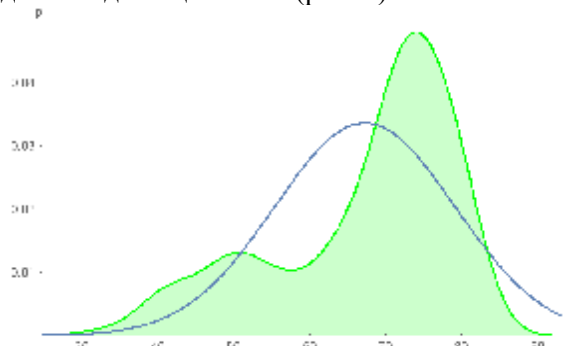


Рис. 2. Гауссівська щільність розподілу на фоні вихідної

З рис. 2 очевидно, що гауссівська модель вкрай незадовільно апроксимує досліджувані статистичні дані.

Перейдемо безпосередньо до апроксимації щільності розподілу моделями БГВВ. Для початку розглянемо представлення розподілу статистичних даних моделлю асиметричної випадкової величини при $s=2$. Застосувавши усічений ряд Еджворта, модель перфо-

рованої випадкової величини матиме щільність розподілу:

$$p_{\gamma_3(2)}(x) = p_n(x) - \frac{\kappa_3}{3!} p_n^{(3)}(x). \quad (6)$$

В позначені щільності розподілу використовується подвійний індекс, перший вказує на клас перфорованої випадкової величини, а другий вказує на степінь поліному.

Для розрахунку кумулянтних коефіцієнтів застосуємо наступну формулу:

$$\kappa_3 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^3 = -1644,56,$$

$$\kappa_4 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^4 = 57654,3.$$

Не дивлячись на те, що розраховане значення κ_4 істотно відрізняється від нуля, в аналітичному представленні щільності розподілу цим значенням нехтують у відповідності з вимогами опису вибраної моделі. Після підстановки значень отримаємо щільність розподілу, зображену на рис. 3.

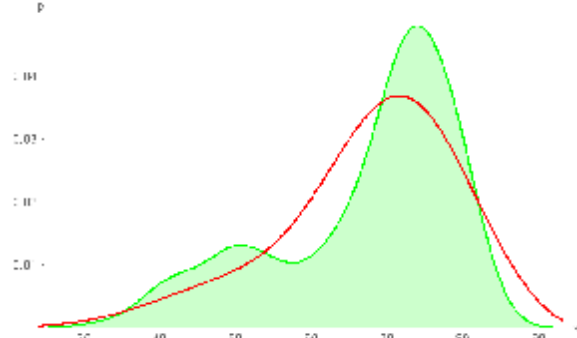


Рис. 3. Щільність розподілу асиметричної моделі при $s=2$ на фоні вихідної

Далі, для побудови апроксимаційної моделі щільності розподілу емпіричних даних застосуємо клас ексцесної випадкової величини при $s=2$, щільність розподілу якої можна представити у вигляді

$$p_{\gamma_4(2)}(x) = p_n(x) + \frac{\kappa_4}{4!} p_n^{(4)}(x). \quad (7)$$

Щільність розподілу ексцесної випадкової величини при $s=2$ виду (7), якою апроксимуються емпіричні дані, представлена на рис. 4. Розглянемо випадок, коли вхідна послідовність апроксимується асиметрично-ексцесною моделлю при $s=2$ (рис. 5) з щільністю розподілу

$$p_{\gamma_{3\gamma_4}(2)}(x) = p_n(x) - \frac{\kappa_3}{3!} p_n^{(3)}(x) + \frac{\kappa_4}{4!} p_n^{(4)}(x). \quad (8)$$

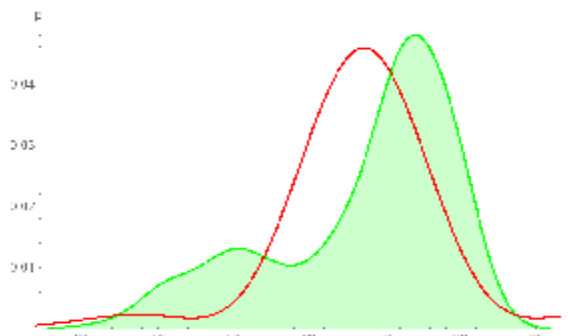


Рис. 4. Щільність розподілу ексцесної моделі при $s=2$ на фоні вихідної

Візуально жоден з розглянутих класів перфорованих величин, не підходить для якісної апроксимації вибірових даних.

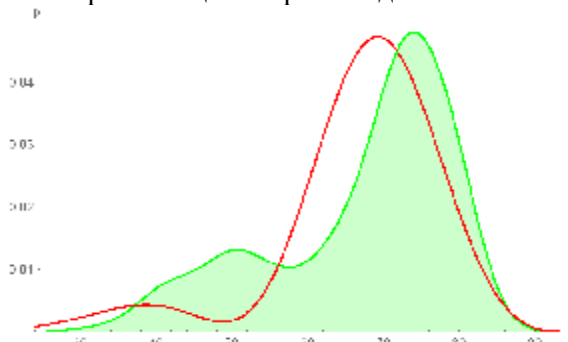


Рис. 5. Щільність розподілу асиметрично-ексцесної моделі при $s=2$ на фоні вихідної

Наступним застосуємо клас асиметричних випадкових величин при $s=3$ для побудови апроксимаційної моделі щільності розподілу:

$$p_{\gamma_3(3)}(x) = p_n(x) - \frac{\kappa_3}{3!} p_n^{(3)}(x) + \frac{10\kappa_3^2}{6!} p_n^{(6)}(x) \quad (9)$$

Порівняно з виразом виду (6) в даному виразі з'явився додатковий доданок, який уточнює модель. Графічне представлення щільності розподілу експериментальних даних і асиметричної випадкової величини при $s=3$ наведено на рис. 6.

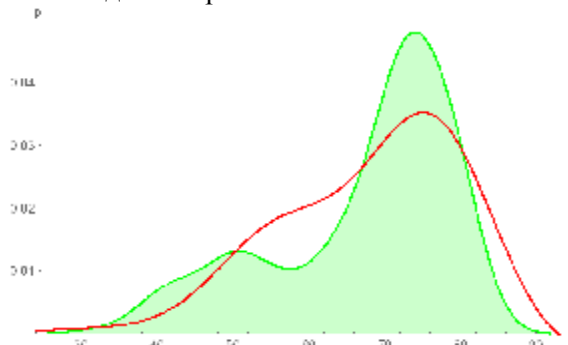


Рис. 6. Щільність розподілу асиметричної моделі при $s=3$ на фоні вихідної

Застосування класу ексцесних випадкових величин при $s=3$ для апроксимації щільності розподілу нічим не відрізняється від випадку $s=2$ і описується функцією (7), яка представлена на рис.4.

Перейдемо до побудови моделі в класі асиметрично-ексцесної випадкової величини при $s=3$

$$p_{\gamma_{3\gamma_4}(3)}(x) = p_n(x) - \frac{\kappa_3}{3!} p_n^{(3)}(x) + \frac{\kappa_4}{4!} p_n^{(4)}(x) + \frac{10\kappa_3^2}{6!} p_n^{(6)}(x). \quad (10)$$

Графік апроксимації щільності розподілу експериментальних даних асиметрично-ексцесною випадковою величиною має вигляд, представлений на рис. 7.

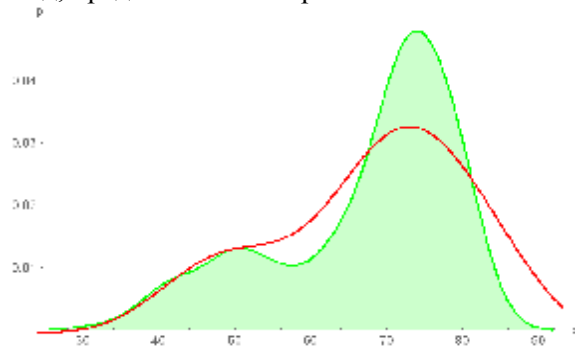


Рис. 7. Щільність розподілу асиметрично-ексцесної моделі при $s=3$ на фоні вихідної

Наступні апроксимації щільності розподілу будуть додатково містити кумулянт п'ятого порядку. Отже, модель асиметричної випадкової величини 2-го типу при $s=3$ описується щільністю розподілу

$$p_{\gamma_{3\gamma_5}(3)}(x) = p_n(x) - \frac{\kappa_3}{3!} p_n^{(3)}(x) - \frac{\kappa_5}{5!} p_n^{(5)}(x) + \frac{10\kappa_3^2}{6!} p_n^{(6)}(x). \quad (11)$$

Графік щільності розподілу (11) представлений на рис. 8.

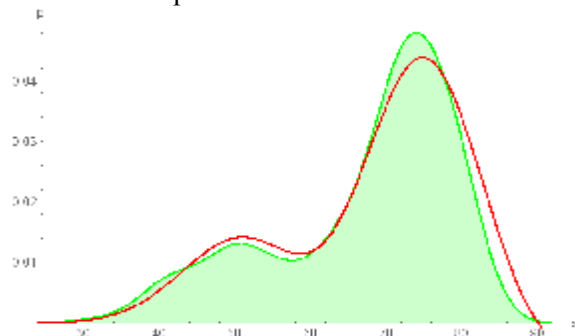


Рис. 8. Щільність розподілу асиметричної моделі 2-го типу при $s=3$ на фоні вихідної

Щільність розподілу асиметрично-ексцесної випадкової величини 3-го типу при $s=3$ представлена на рис. 9 і описується функцією

$$p_{\gamma_{3\gamma_4\gamma_5}(3)}(x) = p_n(x) - \frac{\kappa_3}{3!} p_n^{(3)}(x) - \frac{\kappa_5}{5!} p_n^{(5)}(x) + \frac{\kappa_4}{4!} p_n^{(4)}(x) + \frac{10\kappa_3^2}{6!} p_n^{(6)}(x) \quad (12)$$

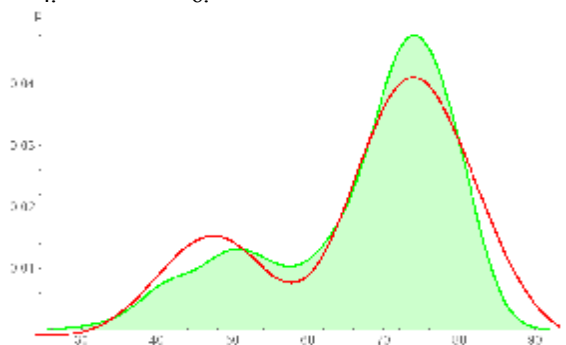


Рис. 9. Щільність розподілу асиметрично-ексцесної 3-го типу моделі при $s=3$ на фоні вихідної

При апроксимації експериментальних даних моделями БГВВ при $s=4$ функції розподілу, отримані на основі усечених рядів Еджворта, мають ділянки, на яких набувають від'ємних значень, тому не можуть розглядатися як функції щільності ймовірності.

Отже для досліджуваної задачі, з аналізу графіків можна зробити висновок, що для опису щільності розподілу вибіркового даного тривалості життя у світі моделями БГВВ з різною глибиною перфорації, найкращі результати апроксимації досягаються при використанні асиметричної моделі 2-го типу при $s=3$, представленій функцією (11) на рис. 8, втім для неупередженого прийняття моделі необхідно розрахувати числові характеристики відхилення модельованої щільності від реальної за виразом (3). Маємо, при $s=2$

$$\varepsilon_{\gamma_3(2)} = 9,837 \cdot 10^{-3}, \quad \varepsilon_{\gamma_4(2,3)} = 12,243 \cdot 10^{-3},$$

$$\varepsilon_{\gamma_{3\gamma_4}(2)} = 11,632 \cdot 10^{-3},$$

при $s=3$

$$\varepsilon_{\gamma_3(3)} = 8,826 \cdot 10^{-3}, \quad \varepsilon_{\gamma_{3\gamma_4}(3)} = 7,947 \cdot 10^{-3},$$

$$\varepsilon_{\gamma_{3\gamma_5}(3)} = 5,165 \cdot 10^{-3}, \quad \varepsilon_{\gamma_{3\gamma_4\gamma_5}(3)} = 6,563 \cdot 10^{-3}.$$

Числові дані про похибки апроксимації різними моделями БГВВ підтверджують правильність візуального сприйняття про найкращий варіант підбору перфорованої моделі, а саме моделі, заданої в класі асиметричної

моделі 2-го типу з набором кумулянтних коефіцієнтів γ_3, γ_5 .

Розрахуємо, як перфоровані моделі описують вихідний розподіл в порівнянні з гауссівською моделлю у відсотках:

$$\varepsilon_{a(\gamma)} = \frac{|\varepsilon_n - \varepsilon_\gamma|}{|\varepsilon_n|} \cdot 100\%. \quad (13)$$

Маємо, при $s=2$

$$\varepsilon_{a(\gamma_3(2))} \approx 33\%, \quad \varepsilon_{a(\gamma_4(2,3))} \approx 17\%,$$

$$\varepsilon_{a(\gamma_3\gamma_4(2))} \approx 21\%,$$

при $s=3$

$$\varepsilon_{a(\gamma_3(3))} \approx 40\%, \quad \varepsilon_{a(\gamma_3\gamma_4(3))} \approx 46\%,$$

$$\varepsilon_{a(\gamma_3\gamma_5(3))} \approx 65\%, \quad \varepsilon_{a(\gamma_3\gamma_4\gamma_5(3))} \approx 56\%.$$

Загалом при $s=3$ точність апроксимації щільності розподілу різними моделями з перфорованим кумулянтним описом зростає порівняно з випадком $s=2$ за рахунок додатково врахування впливу відповідних інформативних кумулянтів. Слід також враховувати, що при збільшенні степені s перфорована модель може призводити до значних спотворень щільності розподілу і як наслідок не може бути використана.

Очевидно що перфорована модель не завжди гарно описує вихідний розподіл, проте можна підібрати оптимальний клас моделей

БГВВ, які забезпечують найменшу похибку апроксимації щільності розподілу.

Для зручності сприйняття оцінки значень похибки і порівняння їх між собою побудуємо табл. 1, дані якої характеризують точність апроксимуючих моделей.

Розглянемо результати апроксимації вибірових даних якості на кондитерському виробництві [9]. По аналогії з розв'язком попередньої задачі побудуємо гістограму розподілу вибірових значень і апроксимуємо її засобами Wolfram Mathematica для отримання емпіричної щільності розподілу (рис. 10).

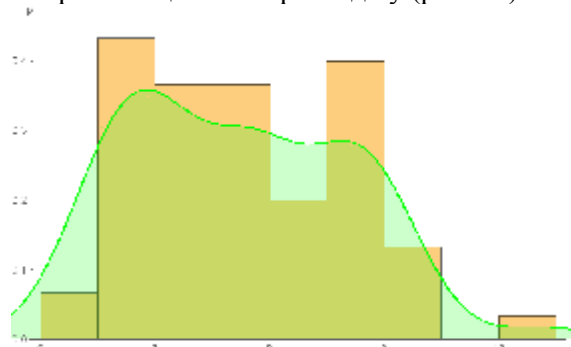


Рис. 10. Графічне представлення емпіричної щільності розподілу якості індексу жирності

Апроксимуємо вихідні дані гауссівською моделлю.

Таблиця 1

Похибки апроксимації статистичних даних моделями БГВВ

Статистичні дані про середню тривалість життя (років) в країнах світу				
	точність апроксимуючих моделей			
Гауссівська апроксимація	$14,743 \cdot 10^{-3}$			
Степінь поліному	s=2		s=3	
Моделі перфорованих випадкових величин	ε_0	$\varepsilon_{a(\gamma)}, \%$	ε_0	$\varepsilon_{a(\gamma)}, \%$
Асиметрична 1 типу $\{\gamma_3\}$	$9,837 \cdot 10^{-3}$	33	$8,826 \cdot 10^{-3}$	40
Ексцесна 1 типу $\{\gamma_4\}$	$12,243 \cdot 10^{-3}$	17	$12,243 \cdot 10^{-3}$	17
Асиметрично-ексцесна 2 типу $\{\gamma_3, \gamma_4\}$	$11,632 \cdot 10^{-3}$	21	$7,947 \cdot 10^{-3}$	46
Асиметрична 2 типу $\{\gamma_3, \gamma_5\}$			$5,165 \cdot 10^{-3}$	65
Асиметрично-ексцесна 3 типу $\{\gamma_3, \gamma_4, \gamma_5\}$			$6,563 \cdot 10^{-3}$	56
Статистичні дані про показники контролю якості на кондитерському виробництві				
	точність апроксимуючих моделей			
Гауссівська апроксимація	$3,224 \cdot 10^{-3}$			
Степінь поліному	s=2			
Моделі перфорованих випадкових величин	ε_0	$\varepsilon_{a(\gamma)}, \%$		
Асиметрична 1 типу $\{\gamma_3\}$	$2,963 \cdot 10^{-3}$	8		
Ексцесна 1 типу $\{\gamma_4\}$	$0,356 \cdot 10^{-3}$	89		
Асиметрично-ексцесна 2 типу $\{\gamma_3, \gamma_4\}$	$0,237 \cdot 10^{-3}$	93		

Для цього розраховуємо значення математичного сподівання та середньоквадратичне відхилення вибірових даних:

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i = 7,738,$$

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2} = 0,96.$$

Використовуючи розраховані значення, будемо апроксимувати гауссівську щільність розподілу (рис. 11).

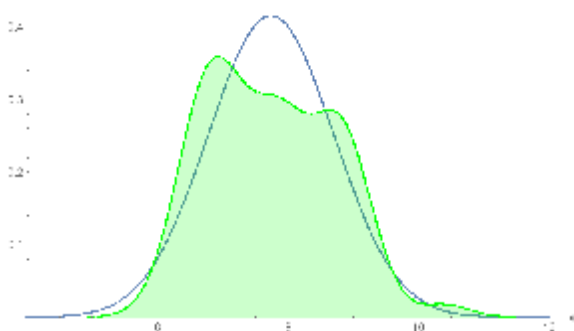


Рис. 11. Гауссівська щільність розподілу на фоні вихідної

Перейдемо до апроксимації статистичних даних про показник контролю якості моделями перфорованих випадкових величин. Використаємо модель асиметричної випадкової величини при $s=2$. Спочатку розраховуємо кумулянтні коефіцієнти:

$$\kappa_3 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^3 = 0,302,$$

$$\kappa_4 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^4 = 2,017.$$

Використовуючи ряд Еджворта для переходу від кумулянтного опису до опису у вигляді щільності розподілу, отримаємо функцію виду (6), графік якої наведено на рис. 12.

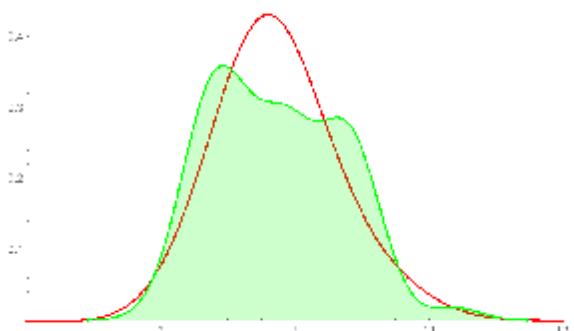


Рис. 12. Щільність розподілу асиметричної моделі при $s=2$ на фоні вихідної

Наступним застосуємо клас ексцесних випадкових величин для апроксимації щільності розподілу, який описується щільністю розподілу виду (7) з кумулянтним 4-го порядку, розрахованим для досліджуваної вибірки (рис. 13).

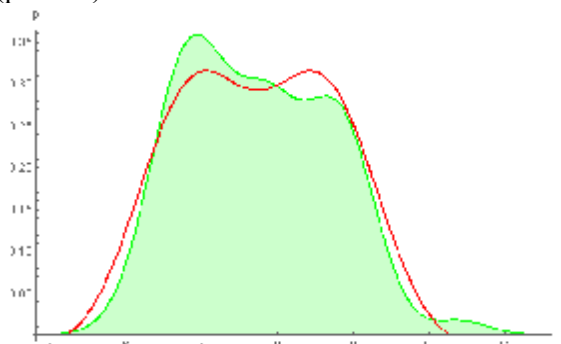


Рис. 13. Щільність розподілу ексцесної моделі при $s=2$ на фоні вихідної

Далі розглянемо апроксимацію асиметрично-ексцесною моделлю при $s=2$, що описується виразом (8), а графік щільності розподілу представлений на рис. 14.

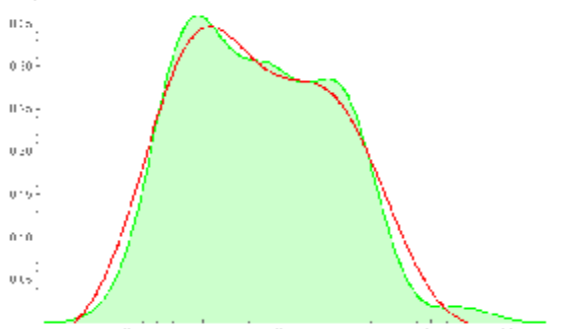


Рис. 14. Щільність розподілу асиметрично-ексцесної моделі при $s=2$ на фоні вихідної

Для кількісної оцінки неузгодженості реального і апроксимуючого розподілів знайдемо числові характеристики відхилення за виразом (3) і зведемо в табл. 1. Також для полегшення інтерпретації результатів в таблиці наведені дані порівняння якості апроксимації моделями БГВВ порівняно з гауссівською моделлю, розраховані по формулі (13).

З графіків, зображених на рис. 12-14, і з даних, наведених у табл. 1, видно, що для статистичних даних показника контролю якості на кондитерському виробництві оптимальною є апроксимація моделлю асиметрично-ексцесної величини при $s=2$. Також непогані результати отримані для ексцесної моделі при $s=2$. Оскільки ексцесна модель має меншу кількість параметрів ніж асиметрично-

ексцесна, а результати апроксимації дуже близькі між собою, то для спрощення подальшої обробки статистичних даних можна використовувати більш просту модель.

Висновки. При використанні різних моделей перфорованих випадкових величин для апроксимації статистичних даних необхідно враховувати емпіричні значення кумулянтів. Врахування великої кількості кумулянтів, як правило, уточнює модель і покращує точність апроксимації, проте одночасно з цим приводить до її ускладнення. При збільшенні глибини перфорації моделі (штучному зануленні кумулянтів вищих порядків до $2s$ -го порядку) апроксимаційна функція може приймати від'ємні значення і не задовольняє вимогам до щільності ймовірності і як наслідок не може застосовуватися.

При виборі апроксимаційної моделі доцільно керуватися трьома факторами:

- близькістю графічного представлення апроксимаційної функції до емпіричної щільності розподілу;
- найменшою похибкою апроксимації ε_0 або найбільшим виграшем порівняно з гауссівською апроксимацією $\varepsilon_{a(\gamma)}$;
- при майже однакових результатах точності рекомендується використовувати більш просту модель з меншою кількістю параметрів.

Список літератури

1. Ван Трис Г. Теория обнаружения, оценок и модуляции. Т. 3. Обработка сигналов в радио- и гидролокации и прием случайных гауссовских сигналов на фоне помех: пер. с англ., под ред. В. Т. Горяинова. М.: Сов. радио, 1977. 664 с.
2. Левин Б. Р. Теоретические основы статистической радиотехники. Кн. 1. М.: Сов. радио, 1969. 752 с.
3. Кунченко Ю. П. Полиномиальные оценки параметров близких к гауссовским случайных величин. Ч. 1. Стохастические полиномы, их свойства и применение для нахождения оценок параметров. Черкассы: ЧИТИ, 2001. 133 с.
4. Гавриш О. С., Заболотний С. В., Бурдукова Е. В., Иващенко А. А. Критерий аппроксимации статистических данных моделями на основе перфорированного кумулянтного описания. *Обробка сигналів і не-*

гауссівських процесів: тези доп. V Міжнар. наук.-практ. конф. Черкаси: ЧДТУ, 2015. С. 14–17.

5. Берегун В. С., Берегун В. С., Красильников О. І. Апроксимативні методи знаходження щільності імовірностей. *Електроніка і зв'язь*. 2010. № 4 (57). С. 51–55.
6. Заболотний С. В., Чепинога А. В. Апроксимація емпіричних розподілів поліноміальних статистик полігауссовими моделями. *Новітні технології в телекомунікаціях*: тези V Міжнар. наук.-техн. симпозиуму. К.: ДУІКТ, 2012. С. 86–88.
7. Малахов А. Н. Кумулянтный анализ негауссовских случайных процессов и их преобразований. М.: Сов. радио, 1978. 376 с.
8. https://en.wikipedia.org/wiki/List_of_countries_by_life_expectancy – List of countries by life expectancy
9. <http://www.statistica.ru/local-portals/industry-analytics/example/1558/> – Контроль качества на кондитерском производстве.

References

1. Van Tris, G. (1977) The theory of detection, estimates and modulation, in Goryainov, V. T. (ed.). Vol. 3. Signal processing in radar and sonar and the reception of random Gaussian signals against interference. Moscow: Sov. radio, 664 p. [in Russian].
2. Levin, B. R. (1969) Theoretical foundations of statistical radio engineering. Book 1. Moscow: Sov. radio, 752 p. [in Russian].
3. Kunchenko, Yu. P. (2001) Polynomial estimates of parameters of close to Gaussian random variables. Part 1. Stochastic polynomials, their properties and applications to find parameters estimates. Cherkassy: ChITI, 133 p. [in Russian].
4. Havrysh, O. S., Zabolotnii, S. V., Burduкова, O. V., Ivashchenko, A. A. (2015) The criterion of approximation of statistical data by models based on perforated cumulant description. *Obrobka syhnaliv i nehaussivskiyh procesiv*: theses of the V international scientific-practical conference. Cherkassy: ChDTU, pp. 14–17 [in Russian].
5. Berehun, V. S., Berehun, V. S., Krasylnikov, O. I. (2010) Approximation methods of finding probability density. *E'lektronika i svyaz'*, 4 (57), pp. 51–55 [in Ukrainian].

6. Zabolotnii, S. V., Chepynoha, A. V. (2012) Approximation of empirical distributions of polynomial statistics polygaussian models. *Novitni tehnologiyi v telekomunikatsiyah: theses of the V international scientific and technical symposium*. Kiev, pp. 86–88 [in Ukrainian].
7. Malahov, A. N. (1978) Non-Gaussian cumulant analysis of random processes and their transformations. Moscow: Sov. radio, 376 p. [in Russian].
8. https://en.wikipedia.org/wiki/List_of_countries_by_life_expectancy – The list of countries by life expectancy.
9. <http://www.statistica.ru/local-portals/industry-analytics/example/1558/> – Quality control at confectionery.

O. V. Burdukova, Ph.D.,
Yu. G. Lega, Dr.Tech.Sc., professor,
O. S. Havrysh, Ph.D., associate professor,
T. V. Vorobkalo, Ph.D., associate professor,
A. O. Ivashchenko, Ph.D.
Cherkasy State Technological University
Shevchenko blvd, 460, Cherkasy, 18006, Ukraine

APPROXIMATION OF EXPERIMENTAL DATA DISTRIBUTION BY CLOSE TO GAUSSIAN RANDOM VARIABLE MODELS

In the paper analytic expressions of approximating functions on the basis of models with perforated cumulant description are obtained.

Densities of the distribution of various models, close to the Gaussian random variables, are built, the comparison with empirical distribution density is made and the value of approximation errors for each model class is found. It is shown that the models based on perforated cumulant description are an effective tool for approximating real statistical data of different nature.

Keywords: *approximation, number of Edgeworth, cumulant coefficients, close to the Gaussian random variables, perforation of cumulant description.*

Статтю представляє Ю. Г. Лега, д.т.н., професор