

С. В. Гадецька¹, В. Ю. Дубницький², Ю. І. Кушнерук³, О. І. Ходирев²¹ Харківський національний автомобільно-дорожній університет, Харків, Україна² ННІ “Каразінський банківський інститут” ХНУ імені В. Н. Каразіна, Харків, Україна³ Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна

EXCEL-ОРІЄНТОВАНІ ПРОЦЕДУРИ ВИЗНАЧЕННЯ ЕНТРОПІЇ ФУНКЦІЇ РОЗПОДІЛУ ТА ЇЇ ВІДНОСНОЇ ПАРАМЕТРИЧНОЇ ЧУТЛИВОСТІ (ЕЛАСТИЧНОСТІ) В УМОВАХ ДВОСТОРОННІХ ОБМЕЖЕНЬ НА ОБЛАСТЬ ЗНАЧЕНЬ НЕПЕРЕРВНОЇ ВИПАДКОВОЇ ВЕЛИЧИНИ

Анотація. Мета роботи. Розробка EXCEL-орієнтованого калькулятора для обчислення ентропії та її еластичності для функцій розподілу за умови обмеженої області визначення неперервної випадкової величини. **Предмет дослідження.** Функції щільності ймовірності та їх ентропії за умови двосторонніх обмежень на область визначення можливих значень випадкових величин. **Методи дослідження.** Алгоритмічний та чисельний аналіз процедур отримання чисельних значень ентропії функцій щільності неперервних випадкових величин за умови двобічних обмежень на область її визначення. **Отримані результати.** В роботі запропоновано EXCEL-орієнтований калькулятор для обчислення ентропії та її еластичності для функцій розподілу за умови обмеженої області визначення неперервної випадкової величини. Всі використані в роботі функції розподілу розподілені на три категорії залежно від того, в якій формі подана ентропія та її еластичність. До першої категорії включено функції розподілу, для яких ентропія та її еластичність можуть бути визначені в аналітичній формі. До другої категорії включено функції розподілу, для яких ентропія може бути визначена в аналітичній формі, а її еластичність – в табличній. До третьої категорії включено функції розподілу, для яких ентропія та її еластичність можуть бути визначені в табличній формі.

Ключові слова: ентропія; функції щільності ймовірності; двосторонні обмеження на область значень випадкових величин; відносна параметрична чутливість (еластичність); програмовані калькулятори; ентропійний аналіз.

Вступ

Неперервна випадкова величина X , яка має функцію розподілу $F(x)$ і щільність розподілу $f(x)$ може бути визначена на одній з трьох можливих областей визначення: $R1$ – область можливих значень задана на всій числовій осі, тобто $X \in (-\infty, +\infty)$, наприклад, нормальний розподіл; $R2$ – область можливих значень задана на додатній півосі, тобто $X \in [0, +\infty)$, наприклад, експоненціальний розподіл; $R3$ – область можливих значень задана на обмеженому інтервалі числової осі, тобто $X \in [a, b]$, наприклад, рівномірний розподіл.

У тому випадку, коли $R3 \subset R1$ або $R3 \subset R2$ розподіли можливих значень випадкової величини X називають урізаними.

Аналіз літератури. Функціонал, що має вигляд:

$$h(-\infty, \infty) = - \int_{-\infty}^{\infty} f(x) \ln f(x) dx \quad (\text{нім}) \quad (1)$$

$$H[a, b] = \begin{cases} 0, & \text{якщо } x \leq a, \\ - \int_a^b \frac{f(x)}{F(b) - F(a)} \cdot \ln \frac{f(x)}{F(b) - F(a)} dx, & \text{якщо } a < x \leq b, \\ 0, & \text{якщо } x > b. \end{cases} \quad (3)$$

В [3] отриманий загальний розв'язок задачі визначення ентропії неперервної випадкової величини при довільних варіантах її урізання. Необхідні співвідношення наведено в табл. 1. У цій же роботі наведено розв'язок цієї задачі у випадку урізаного показникового розподілу. Цей розв'язок показано у табл. 2.

Визначення ентропії для інших видів розподілу ймовірності з двосторонніми обмеженнями на область значень отримано в [4].

називають приведеною або диференціальною ентропією Шеннона. Його властивості детально розглянуті в роботі [1]. Щільність урізаної на інтервалі (a, b) випадкової величини X з функцією розподілу $F(x)$ в [2] визначена так:

$$f_{(a,b)}(x) = \begin{cases} 0, & \text{якщо } x \leq a, \\ \frac{f(x)}{F(b) - F(a)}, & \text{якщо } a < x \leq b, \\ 0, & \text{якщо } x > b. \end{cases} \quad (2)$$

Межі інтервалів урізання обумовлені фізичним змістом задачі, наприклад, величиною пропускної спроможності каналу зв'язку. В даній роботі буде розглянуто особливості побудови EXCEL-орієнтованих процедур визначення ентропії функцій, можливі значення яких задані областю $R3$, тобто їх розподіли можна вважати урізаними.

Використовуючи співвідношення (1, 2) визначимо ентропію випадкової величини, яка задана урізаним розподілом, у такому вигляді:

$$f(x) = \Gamma(\eta + \lambda) / (\Gamma(\eta)\Gamma(\lambda)) \cdot x^{\lambda-1} (1-x)^{\eta-1}, \quad (4)$$

$$\eta > 0, \quad \lambda > 0, \quad 0 \leq x \leq 1,$$

диференціальну ентропію (ентропію) визначають як

$$h = \ln \left[\frac{B(\eta, \lambda) \exp(\eta + \lambda - 2) \Psi(\eta + \lambda)}{\exp[(\lambda - 1)\Psi(\lambda)] \cdot \exp[(\eta - 1)\Psi(\eta)]} \right]. \quad (5)$$

Таблиця 1 – Диференціальна ентропія Шеннона для неперервної випадкової величини з довільним урізанням розподілом

Область визначення можливих значень неперервної випадкової величини X	Диференціальна ентропія Шеннона неперервної випадкової величини, яка має урізаний розподіл
$-\infty < a < b < +\infty$	$H[a, b] = -(F(b) - F(a))^{-1} \int_a^b f(x) \ln f(x) dx + \ln(F(b) - F(a))$
$0 < x < a$	$H[0, a] = -[F(a)]^{-1} \int_0^a f(x) \ln f(x) dx + \ln F(a)$
$a < x < +\infty$	$H[a, +\infty) = -(1 - F(a))^{-1} \int_a^{+\infty} f(x) \ln f(x) dx + \ln(1 - F(a))$

Таблиця 2 – Диференціальна ентропія Шеннона для випадкової величини з урізанням показниковим розподілом

Область визначення можливих значень величини X	Диференціальна ентропія Шеннона для випадкової величини з урізанням показниковим розподілом
$-\infty < a < b < +\infty$	$H[a, b] = \frac{e^{-\lambda a}(\lambda a + 1) - e^{-\lambda b}(\lambda b + 1)}{e^{-\lambda a} - e^{-\lambda b}} + \ln(e^{-\lambda a} - e^{-\lambda b}) - \ln \lambda$
$0 < x < a$	$H[0, a] = \frac{1 - e^{-a\lambda}(a\lambda + 1)}{1 - e^{-a\lambda}} + \ln(1 - e^{-a\lambda}) - \ln \lambda$
$a < x < +\infty$	$\ln \frac{e}{\lambda}$

Для узагальненого бета-розподілу із щільністю:

$$f(x) = \frac{1}{b-a} \cdot \frac{\Gamma(\eta + \lambda)}{\Gamma(\eta)\Gamma(\lambda)} \cdot \left(\frac{x-a}{b-a}\right)^{\lambda-1} \times \left(\frac{b-x}{b-a}\right)^{\eta-1}, \quad \eta > 0, \lambda > 0, a \leq x \leq b, \quad (6)$$

ентропію визначають за співвідношенням:

$$h = \ln \left[\frac{(b-a)B(\eta, \lambda) \exp(\eta + \lambda - 2) \Psi(\eta + \lambda)}{\exp[(\lambda - 1)\Psi(\lambda)] \cdot \exp[(\eta - 1)\Psi(\eta)]} \right]. \quad (7)$$

Для розподілу Сімпсона із щільністю:

$$f(x) = \frac{a - |x|}{a^2}, \quad -a \leq x \leq a, \quad (8)$$

ентропію визначають за співвідношенням:

$$h = \ln(a\sqrt{e}). \quad (9)$$

Для урізаного нормального розподілу із з функцією розподілу $\Phi(x; m, \sigma)$ та щільністю $\varphi(x; m, \sigma)$ за умови, що $x \in [a, b]$, ентропію визначають за співвідношенням:

$$h = \ln(Z\sigma\sqrt{2\pi}) + \frac{1}{Z} \left[\frac{a-m}{\sigma} \varphi\left(\frac{a-m}{\sigma}\right) - \frac{b-m}{\sigma} \varphi\left(\frac{b-m}{\sigma}\right) \right] + \frac{1}{2}, \quad (10)$$

де $Z = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right).$ (11)

Особливості отримання чисельних значень спеціальних функцій, а саме: гамма-функції $\Gamma(x)$,

бета-функції $B(x, y)$, дигамма-функції $\Psi(x)$, функції нормального розподілу $\Phi(x; m, \sigma)$, щільності нормального розподілу $\varphi(x; m, \sigma)$, функції ділогарифму $Li_2(x)$ детально буде розглянуто далі.

Розглянемо двохпараметричну щільність розподілу випадкової величини X вигляду $f(x; \lambda, \eta)$.

Нехай для параметрів щільності розподілу λ і η відомі співвідношення вигляду:

$$\begin{aligned} \lambda &= u_1(m, \sigma); \\ \eta &= u_2(m, \sigma), \end{aligned} \quad (12)$$

де m – математичне сподівання, σ – стандартне відхилення випадкової величини X , що має щільність розподілу $f(x; \lambda, \eta)$.

Для найбільш поширених видів розподілу ці співвідношення приведені в [2, 4]. У тому випадку, коли замість параметрів розподілу λ і η використовують їх оцінки $\hat{\lambda}$ і $\hat{\eta}$, які отримані за наслідками експерименту, замість співвідношення (12) слід використовувати співвідношення вигляду:

$$\begin{aligned} \hat{\lambda} &= w_1(\bar{x}, s); \\ \hat{\eta} &= w_2(\bar{x}, s), \end{aligned} \quad (13)$$

де \bar{x} – середнє значення, s – середньоквадратичне відхилення випадкової величини X . Для отримання оцінок $\hat{\lambda}$ і $\hat{\eta}$ використовують, як правило, метод моментів або метод максимуму правдоподібності. Для найбільш поширених видів розподілу ці співвідношення приведено в [5].

Підставляючи вирази (12) до функціоналу (1) отримаємо співвідношення для визначення ентропії у вигляді функції від параметрів закону розподілу випадкової величини X :

$$h_p = - \int_{-\infty}^{\infty} \left(f(x; u_1(m, \sigma), u_2(m, \sigma)) \times \ln f(x; u_1(m, \sigma), u_2(m, \sigma)) \right) dx = g(m, \sigma) \quad (14)$$

Підставляючи (13) в (1) отримаємо співвідношення для визначення ентропії у вигляді функції, аргументами якої будуть основні статистичні характеристики отриманої вибірки: її середнє значення і середньоквадратичне відхилення. Як відомо, ці характеристики існують для будь-якої вибірки незалежно від закону розподілу генеральної сукупності даних. Звідси,

$$h_s = - \int_{-\infty}^{\infty} \left(f(x; w_1(\bar{x}, s), w_2(\bar{x}, s)) \times \ln f(x; w_1(\bar{x}, s), w_2(\bar{x}, s)) \right) dx = q(\bar{x}, s) \quad (15)$$

Співвідношення (13), (14) дозволяють досліджувати залежність величини ентропії від параметрів вибраних розподілів або функцій від цих параметрів. Зокрема, в [4] визначені залежності величини ентропії від дисперсії для деяких типів розподілів. Використовуючи [5] та співвідношення (14) і (15), для визначення відносною параметричної чутливості (еластичності) ентропії по кожному з аргументів отримаємо співвідношення:

$$E_m(h_p) = \frac{\partial}{\partial m} g(m, \sigma) \frac{m}{g(m, \sigma)}; \quad (16)$$

$$E_\sigma(h_p) = \frac{\partial}{\partial \sigma} g(m, \sigma) \frac{\sigma}{g(m, \sigma)}$$

або
$$E_{\bar{x}}(h_s) = \frac{\partial}{\partial \bar{x}} q(\bar{x}, s) \frac{\bar{x}}{q(\bar{x}, s)}; \quad (17)$$

$$E_s(h_s) = \frac{\partial}{\partial s} q(\bar{x}, s) \frac{s}{q(\bar{x}, s)}.$$

Співвідношення (14), (16) слід використовувати на етапі проектування технічної системи, співвідношення (15), (17) – на етапах її випробувань та експлуатації. У [8] наведено співвідношення для отримання чисельних аналогів (16), (17) для функції відносною чутливості $E_y(x)$.

Ці співвідношення показано в табл. 3.

Таблиця 3 – Співвідношення для визначення відносною чутливості (еластичності)

Умовне позначення формули	Еластичність	Розрахункове співвідношення
Ф(3.1)	Відсоткова	$\left(\frac{y_2 - y_1}{y_1} \right) / \left(\frac{x_2 - x_1}{x_1} \right)$
Ф(3.2)	Середня дугова	$\left(\frac{2(y_2 - y_1)}{y_1 + y_2} \right) / \left(\frac{2(x_2 - x_1)}{x_1 + x_2} \right)$
Ф(3.3)	Логарифмічна	$\ln \left(\frac{y_2}{y_1} \right) / \ln \left(\frac{x_2}{x_1} \right)$

У [8] відмічено, що при малих відносних змінах ці способи обчислення величини $E_y(x)$ мало відрізняються між собою. Більш докладне дослідження області застосування кожної з формул виходить за рамки даного повідомлення. В калькуляторі передбачено використання кожної з формул з табл. 3. В даній роботі застосована формула Ф(3.3).

Таким чином, у сучасній науковій літературі, на думку авторів даного повідомлення, методи обчислення ентропії та її еластичності за умови обмеженої області визначення неперервної випадкової величин висвітлені недостатньо і потребують подальших досліджень.

Мета роботи. Розробка EXCEL-орієнтованого калькулятора для обчислення ентропії та її еластичності для функцій розподілу за умови обмеженої області визначення неперервної випадкової величини.

Предмет дослідження. Функції щільності ймовірності та їх ентропії за умови двосторонніх обмежень на область визначення можливих значень випадкових величин.

Методи дослідження. Алгоритмічний та чисельний аналіз процедур отримання чисельних значень ентропії функцій щільності неперервних випадкових величин за умови двобічних обмежень на область її визначення.

Отримані результати

Обчислення засобами системи EXCEL значень спеціальних функцій, використаних в роботі. Для розв'язання поставлених задач виникла необхідність визначення значень деяких спеціальних функцій, а саме: гамма-функції $\Gamma(x)$, бета-функції $B(x, y)$, дігамма-функції $\Psi(x)$, функції нормального розподілу $\Phi(x; m, \sigma)$, щільності нормального розподілу $\phi(x; m, \dots)$. Особливості отримання їх засобами системи EXCEL показано в табл. 4.

Значення функції $\Gamma(x)$, залежно від версії системи EXCEL, можуть бути визначені одним з трьох співвідношень, які приведено в табл. 4. У версії калькулятора, яка реалізована в даній роботі, використана формула Ф(4.1).

Обчислення значень дігамма-функції в нашому випадку можливо тільки використовуючи її апроксимацію, наведену в [7]. В табл. 5 приведено значення функцій, що вказані в [7], і їх значення, які обчислені за апроксимуючими співвідношеннями. Припустимо, що X – нормально розподілена випадкова величина, яка має математичне сподівання m і середньоквадратичне відхилення σ . Функція її розподілу має вигляд:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp \left[-\frac{(x-m)^2}{2\sigma^2} \right] dx. \quad (18)$$

Виконавши заміну змінною згідно з умовою (18), отримаємо нормовану за середньоквадратичним відхиленням випадкову величину:

$$z = \frac{x-a}{\sigma} \Rightarrow x = \sigma \cdot z + a \Rightarrow \sigma \cdot dz = dx. \quad (19)$$

Таблиця 4 – Обчислення значень спеціальних функцій засобами системи EXCEL

Назва функції та її умовне позначення	Визначення функції	Обчислення значення функції із використанням вбудованих функцій системи EXCEL
Ф(4.1)		1. ГАММА(x)
Ф(4.2)	$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$	2. $\Gamma(x) \approx e^{-x} x^x \sqrt{2\pi/x} \times \left(1 + \frac{1}{12x} + \frac{1}{288x^2} - \frac{139}{51840x^3} - \frac{571}{2488320x^4}\right)$
Ф(4.3)		3. EXP(ГАММАНЛОГ(x))
Бета-функція Ф(4.4)	$B(\eta, \lambda) = \int_0^1 t^{\eta-1} (1-t)^{\lambda-1} dt$	$B(\eta, \lambda) = \frac{\Gamma(\eta)\Gamma(\lambda)}{\Gamma(\eta+\lambda)}$
Дігамма-функція Ф(4.5)	$\Psi(x) = \frac{d}{dx} \ln \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$	$\Psi(x) \approx \ln x - \frac{1}{2x} - \frac{1}{12x^2} + \frac{1}{120x^4} - \frac{1}{252x^6}$
Функція нормального розподілу Ф(4.6)	$\Phi(x; m, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] dx$	НОРМРАСП(x; m; σ; 1)
Функція нормального розподілу нормованої випадкової величини Ф(4.7)	$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left[-\frac{z^2}{2}\right] dz$	НОРМРАСП(z; 0; 1; 1)
Щільність нормального розподілу Ф(4.8)	$f(x; m, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right]$	НОРМРАСП(x; m; σ; 0)
Щільність нормального розподілу нормованої випадкової величини Ф(4.9)	$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right]$	НОРМРАСП(z; 0; 1; 0)
Стала Ейлера Ф(4.10)	$\gamma = \lim_{m \rightarrow \infty} \left[\sum_{i=1}^m \frac{1}{i} - \ln(m) \right]$	$\gamma = 0,5772215$
Ділогарифм Ф(4.11)	$Li_2(x) = -\int_1^x \frac{\ln t}{t-1} dt = -\int_0^x \frac{\ln(1-t)}{t} dt$	$\sum_{n=1}^{50} (-1)^n \frac{(x-1)^n}{n^2} = \sum_{n=1}^{50} \frac{(1-x)^n}{n^2}$

Таблиця 5 – Табличні значення функцій і їх значення, які обчислені за апроксимуючими співвідношеннями

Значення змінної x	Функція				
	Γ(x)	Апроксимація Γ(x)	exp[ln(Γ(x))]	Ψ(x)	Апроксимація Ψ(x)
4	6	5,99999	6	1,70611	1,70611
1,250	0,90640	0,90623	0,90640	-0,22745	-0,22782

У результаті цього перетворення отримаємо функцію нормального розподілу нормованої величини z:

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left[-\frac{z^2}{2}\right] dz. \quad (20)$$

Значення функції (20) приведено в усіх статистичних таблицях. Це означає, що обчислення значення функцій F(x) і F(z) можна виконувати за одними й тими ж самими таблицями, або по однако-вих, вбудованих в різні програмні системи коман-дах. Обчислення щільності нормального розподілу має істотні відмінності, які залежать від способу визначення випадкової величини.

Припустимо, що:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \quad (21)$$

є щільність початкової нормальної випадкової величини, і

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] \quad (22)$$

є щільність нормальної нормованої випадкової ве-личини. Тоді порівнюючи (21) і (22) отримаємо, що:

$$\sigma \cdot f(x) = f(z). \quad (23)$$

Тобто, чисельне значення щільності норма-вної випадкової величини в σ раз більше значення щільності початкової випадкової величини. Розгля-немо чисельний приклад. Припустимо, що норма-льно розподілена випадкова величина має характерис-тики: x=30, m=28, σ=2,8. Використовуємо (19):

$$z = (30 - 28) / 2,8 = 0,714.$$

Для подальших обчислень використовуємо вбудовану в EXCEL функцію, що має синтаксис: НОРМРАСП(X; середнє; стандартне_відхилення; 0). Результати обчислення показано нижче.

$$f(z = 0,714) = \text{НОРМРАСП}(0,714;0;1;0)=0,30917;$$

$$f(x = 30) = \text{НОРМРАСП}(30; 28; 2,8; 0)=0,11039;$$

$$\sigma \cdot f(x) = f(z) \Rightarrow 2,80 \cdot 0,11039 = 0,309115.$$

Таким чином, застосування статистичних таблиць або EXCEL для обчислення значень щільності нормального розподілу без урахування описаних особливостей призводить до істотних систематичних помилок при виконанні подальших розрахунків.

Обчислення значень функції ділогарифма $Li_2(x)$. У загальному випадку функцію полілогарифма $Li_u(x)$ визначають як суму нескінченного ряду:

$$\tilde{L}i_u(x) = \sum_{k=1}^{\infty} \frac{x^k}{k^u}. \quad (24)$$

Ділогарифм отримують із співвідношення (24), прийнявши величину $u=2$. В [9] наведено два способи визначення ділогарифма в інтегральній формі:

$$Li_2(x) = -\int_1^x \frac{\ln t}{t-1} dt = -\int_0^x \frac{\ln(1-t)}{t} dt. \quad (25)$$

В [9] показано, що:

$$Li_2(x) = \tilde{L}i_2(1-x). \quad (26)$$

Якщо представити кожне із співвідношень, приведених в (25), у вигляді суми нескінченних рядів, то, як показано в [9], отримаємо:

$$\sum_{n=1}^{\infty} (-1)^n \frac{(x-1)^n}{n^2} = \sum_{n=1}^{\infty} \frac{(1-x)^n}{n^2}. \quad (27)$$

Детально властивості функції $Li_2(x)$ розглянуто в [9], таблиці значень функції ділогарифма наведено в [7]. За наслідками чисельного експерименту встановлено, що в співвідношенні (27) кількість членів ряду достатньо прийняти такою, що дорівнює п'ятдесяти.

Функції щільності ймовірності і їх ентропії, які використано в роботі. Найбільш поширені функції розподілу ймовірності, що мають обмеження на

область їх визначення, подано в [2, 4, 7, 9, 11, 12]. Оскільки в деяких випадках україномовні й англійські найменування функцій щільності ймовірності не співпадають, тому в табл. 6 приведено їх відповідність. В табл. 7 наведено відомості про функції щільності розподілів, які використано в роботі. В табл. 8 приведені відомості про ентропію функцій щільності розподілів, які використані в роботі, і еластичності ентропій за параметрами цих розподілів.

Таблиця 6 – Відповідність між україномовними та англійськими назвами функцій щільності ймовірностей, які використані в роботі

№	Україномовне найменування	Англійське найменування
1	Розподіл Сімпсона	Simpson Distribution
2	Трикутний розподіл з від'ємною асиметрією	Right-angled and negatively skew version distribution
3	Трикутний розподіл з додатною асиметрією	Right-angled and positively skew version distribution
4	Синус-розподіл	Sine Wave Distribution
5	Урізаний нормальний розподіл	Truncated Normal Distribution
6	Розподіл Кумарасвами	Kumaraswamy Distribution
7	Бета-розподіл, тип I	Beta Distribution
8	Бета-розподіл, тип II	Generalized Beta Distribution
9	Розподіл арксинусу, тип I	Ark-Sine Distribution Type I
10	Розподіл арксинусу, тип II	Ark-Sine Distribution Type II
11	Розподіл арксинусу, тип III	Ark-Sine Distribution Type III
12	Косинус-розподіл	Ordinary Cosine Distribution
13	Зсунений косинус-розподіл	Raised Cosine Distribution
14	Параболічний розподіл, тип I	U-shaped parabolic distribution Type I
15	Параболічний розподіл, тип II	U-shaped parabolic distribution Type II
16	Відображений параболічний розподіл	Inverted U-shaped parabolic distribution
17	Закон рівномірно зростаючої щільності	Uniformly increasing distribution density
18	Напівеліптичний розподіл	Semi-elliptical distribution
19	Розподіл Сімпсона, тип II	Ark-Sine Distribution Type II

Таблиця 7 – Перелік функцій щільності ймовірностей, які використано в роботі

№	Назва функції	Визначення функції щільності розподілу
1	Розподіл Сімпсона [4]	$f(x) = \begin{cases} (a- x)/a^2, & \text{якщо } x \in [-a, a] \\ 0, & \text{якщо } x \notin [-a, a] \end{cases}$
2	Трикутний розподіл з від'ємною асиметрією [11]	$f(x) = \begin{cases} \frac{2}{b} \left(\frac{x-a}{b} \right), & \text{якщо } x \in [(a-b), a] \\ 0, & \text{якщо } x \notin [(a-b), a] \end{cases}$
3	Трикутний розподіл з додатною асиметрією [11]	$f(x) = \begin{cases} \frac{2}{b} \left(1 - \frac{x-a}{b} \right), & \text{якщо } x \in [a, (a+b)] \\ 0, & \text{якщо } x \notin [a, (a+b)] \end{cases}$
4	Синус-розподіл [4]	$f(x) = \begin{cases} 1/\pi \sqrt{(a^2-x^2)}, & \text{якщо } -a < x < a \\ 0, & \text{якщо } x < -a, x > a \end{cases}$

5	Урізаний нормальний розподіл [4]	$f_i(x) = \begin{cases} \frac{1}{W\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right], & \text{якщо } a \leq x \leq b \\ 0, & \text{якщо } x < a, x > b \end{cases}, \quad W = F\left(\frac{(b-m)^2}{2\sigma^2}\right) - F\left(\frac{(a-m)^2}{2\sigma^2}\right)$
6	Розподіл Кумарасвами [4]	$f(x) = \begin{cases} abx^{a-1}(1-x^a)^{b-1}, & \text{якщо } x \in [0,1] \\ 0, & \text{якщо } x \notin [0,1] \end{cases}$
7	Бета-розподіл, тип I [4]	$f(x) = \begin{cases} \frac{1}{B(\eta,\lambda)} x^{\lambda-1}(1-x)^{\eta-1}, & \text{якщо } x \in [0,1] \\ 0, & \text{якщо } x \notin [0,1] \end{cases}$
8	Бета-розподіл, тип II [4]	$f(x) = \begin{cases} \frac{1}{b-a} \cdot \frac{1}{B(\eta,\lambda)} \left(\frac{x-a}{b-a}\right)^{\lambda-1} \left(\frac{b-x}{b-a}\right)^{\eta-1}, & \text{якщо } x \in [a,b] \\ 0, & \text{якщо } x \notin [a,b] \end{cases}$
9	Розподіл арксинусу, тип I [11]	$f(x) = \begin{cases} \frac{1}{\pi\sqrt{x(1-x)}}, & \text{якщо } x \in [0,1] \\ 0, & \text{якщо } x \notin [0,1] \end{cases}$
10	Розподіл арксинусу, тип II [11]	$f(x) = \begin{cases} \frac{1}{b\pi\sqrt{1-\left(\frac{x-a}{b}\right)^2}}, & \text{якщо } a-b < x < a+b \\ 0, & \text{якщо } x < a, x > b \end{cases}$
11	Розподіл арксинусу, тип III [12]	$f(x) = \begin{cases} \frac{1}{\pi\sqrt{\eta^2 - (x-\mu)^2}}, & \text{якщо } x \in [(\mu-\eta), (\mu+\eta)] \\ 0, & \text{якщо } x \notin [(\mu-\eta), (\mu+\eta)] \end{cases}$
12	Косинус-розподіл [11]	$f(x) = \begin{cases} \frac{1}{2b} \cos\left(\frac{x-a}{b}\right), & \text{якщо } x \in \left[a-b\frac{\pi}{2}, a+b\frac{\pi}{2}\right] \\ 0, & \text{якщо } x \notin \left[a-b\frac{\pi}{2}, a+b\frac{\pi}{2}\right] \end{cases}$
13	Зсунений косинус-розподіл [11]	$f(x) = \begin{cases} \frac{1}{2b} \left[1 + \cos\left(\pi - \frac{x+a}{b}\right)\right], & \text{якщо } x \in \left[a-b\frac{\pi}{2}, a+b\frac{\pi}{2}\right] \\ 0, & \text{якщо } x \notin \left[a-b\frac{\pi}{2}, a+b\frac{\pi}{2}\right] \end{cases}$
14	Параболічний розподіл, тип I [11]	$f(x) = \begin{cases} \frac{3}{2b^3}(x-a)^2, & \text{якщо } x \in [(a-b), (a+b)] \\ 0, & \text{якщо } x \notin [(a-b), (a+b)] \end{cases}$
15	Параболічний розподіл, тип II [11]	$f(x) = \begin{cases} \frac{6(x-a) \cdot (b-x)}{(b-a)^3}, & \text{якщо } x \in [a,b] \\ 0, & \text{якщо } x \notin [a,b] \end{cases}$
16	Відображений параболічний розподіл [11]	$f(x) = \begin{cases} \frac{3}{4b} \left[1 - \left(\frac{x-a}{b}\right)^2\right], & \text{якщо } x \in [a,b] \\ 0, & \text{якщо } x \notin [a,b] \end{cases}$
17	Закон рівномірно зростаючої щільності [12]	$f(x) = \begin{cases} \frac{2(x-a)}{(b-a)^2}, & \text{якщо } x \in [a,b] \\ 0, & \text{якщо } x \notin [a,b] \end{cases}$
18	Напівеліптичний розподіл [11]	$f(x) = \begin{cases} \frac{2}{b\pi} \sqrt{1 - \left(\frac{x-a}{b}\right)^2}, & \text{якщо } x \in [(a-b), (a+b)] \\ 0, & \text{якщо } x \notin [(a-b), (a+b)] \end{cases}$
19	Розподіл Сімпсона тип II [12]	$f(x) = \begin{cases} \frac{4(x-a)}{(b-a)^2}, & \text{якщо } a \leq x \leq \frac{a+b}{2} \\ \frac{4(b-x)}{(b-a)^2}, & \text{якщо } \frac{a+b}{2} \leq x \leq b \end{cases}$

Таблиця 8 – Ентропія функції щільності розподілу та її еластичність за параметрами розподілу

№	Назва функції	Ентропії розподілу	Еластичність ентропії за параметрами розподілу
1	Розподіл Сімпсона	$h = \ln(a\sqrt{e})$	$E_a(h) = 2/(2\ln(a) + 1)$
2	Трикутний розподіл з від'ємною асиметрією	$h = 1,4427 \ln b - 0,2786$	$E_b(h) = 1/(\ln b - 0,19311)$
3	Трикутний розподіл з додатною асиметрією	$h = 1,4427 \ln b + 0,7214$	$E_b(h) = 1/(\ln b + 0,5003)$
4	Синус-розподіл	$h = \ln(a\pi/2)$	$E_a(h) = [\ln(a\pi/2)]^{-1}$
5	Урізаний нормальний розподіл	$h = \ln(W\sigma\sqrt{2\pi}) + \frac{1}{2W} \left[\frac{a-m}{\sigma} \varphi(a,m,\sigma) - \frac{b-m}{\sigma} \varphi(b,m,\sigma) \right] + \frac{1}{2}$	Може бути визначена чисельними методами
6	Розподіл Кумарасвами	$h = -\ln \left(\frac{1}{ab} \exp \left(\frac{b-1}{b} + \frac{a-1}{a} \left(\gamma + \Psi(b) + \frac{1}{b} \right) \right) \right)$	Може бути визначена чисельними методами
7	Бета-розподіл, тип I	$T_1 = (\eta + \lambda - 2)\Psi(\eta + \lambda)$; $T_2 = [(\lambda - 1)\Psi(\lambda)]$ $T_3 = [(\eta - 1)\Psi(\eta)]$; $h = \ln \left[\frac{B(\eta, \lambda) \exp[T_1]}{\exp[T_2] \exp[T_3]} \right]$	Може бути визначена чисельними методами
8	Бета-розподіл, тип II	$h = \ln \left[\frac{(b-a) \cdot B(\eta, \lambda) \exp[T_1]}{\exp[T_2] \exp[T_3]} \right]$	Може бути визначена чисельними методами
9	Розподіл арксинусу, тип I	$h = \pi/4$	$E_a(h) = 0$
10	Розподіл арксинусу, тип II	$h = \ln(b\pi/2)$	$E_b(h) = [\ln(b\pi/2)]^{-1}$
11	Розподіл арксинусу, тип III	$A_1 = \eta^2 - \eta + \mu^2 - \mu$; $B_1 = \eta^2 + \eta + \mu^2 - \mu$ $h = -\frac{1}{\pi} \left[\ln A_1 \sqrt{A_1} - \ln B_1 \sqrt{B_1} + 2(\sqrt{A_1} - \sqrt{B_1}) \ln \pi - 2(\sqrt{A_1} - \sqrt{B_1}) \right]$	Може бути визначена чисельними методами
12	Косинус-розподіл	$\frac{1 + \ln b}{\ln 2}$	$E_b(h) = \frac{1}{1 + \ln b}$
13	Зсунений косинус-розподіл	$h = 0,5573 + 1,4427 \ln b$	$E_b(h) = \frac{1}{0,3863 + \ln b}$
14	Параболічний розподіл, тип I	$h = \frac{2}{3} - \ln \left(\frac{3}{2b} \right)$	$E_b(h) = \frac{3}{2 - 3\ln(3/(2b))}$
15	Параболічний розподіл, тип II	Може бути визначена чисельними методами	Може бути визначена чисельними методами
16	Відображений параболічний розподіл	Може бути визначена чисельними методами	Може бути визначена чисельними методами
17	Закон рівномірно зростаючої щільності	$D_1 = b^2 - 2b^2 \ln \left(\frac{2}{a-b} \right) - 3a^2 - 2a \ln(a-b)^2$ $D_2 = 2a^2 \ln(b-a) + 2ab + 2a^2 \ln 2$ $D_3 = \frac{2a \left[\ln \left(\frac{-2}{a-b} \right) - 1 \right]}{a-b}$; $h = \frac{D_1 + D_2}{2(a-b)^2} - D_3$	Може бути визначена чисельними методами
18	Напівеліптичний розподіл	$h = b\pi - \frac{1}{2\ln 2}$	$E_b(h) = \frac{2\pi \cdot b \ln 2}{2\pi \cdot b \ln 2 - 1}$
19	Розподіл Сімпсона тип II	$h = 1 - 2\ln(2/(b-a))/2$	$E_b(h) = 2/(1 - 2\ln(2/(b-a)))$

Використання EXCEL-орієнтованого калькулятора для визначення ентропії функцій розподілу в умовах обмежень на область значень неперервної випадкової величини. Функції ентропії й еластичності, які показано в табл. 8, можна згрупувати за двома ознаками: за способом їх завдання і за кількістю параметрів, необхідних для визначення їх чисельних

значень. Групування функцій за способом їх завдання приведено в табл. 9.

Групування типів розподілів за кількістю параметрів, які необхідні для обчислення ентропії, показано в табл. 10.

Перелік параметрів, необхідних для обчислення ентропії, визначено в табл. 11.

Таблиця – 9. Групування функцій ентропії та еластичності за способом завдання*)

Спосіб завдання функції	Функція ентропії	Функція еластичності
Аналітичний	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 17, 18, 19	1, 2, 3, 4, 10, 12, 13, 14, 18, 19
Табличний	15, 16	5, 6, 7, 8, 11, 15, 16, 17

*) Номери типів розподілів співпадають з номерами, наведеними в табл. 6.

Таблиця – 10. Групування типів розподілів за кількістю параметрів, які необхідні для обчислення ентропії*)

Кількість параметрів, необхідних для обчислення ентропії	Типи розподілів
Однопараметричні розподіли	1, 2, 3, 4, 10, 12, 13, 14, 18
Двохпараметричні розподіли	6, 7, 11, 15, 16, 17, 19
Чотирьохпараметричні розподіли	5, 8

*) Номери типів розподілів співпадають з номерами, наведеними в табл. 6.

Таблиця 11 – Параметри, які необхідні для обчислення ентропії

№	Тип розподілу	Параметри для обчислення ентропії
1	Розподіл Сімпсона	a
2	Трикутний розподіл з від'ємною асиметрією	b
3	Трикутний розподіл з додатною асиметрією	b
4	Синус-розподіл	a
5	Урізаний нормальний розподіл	a, b, m, σ
6	Розподіл Кумарасвами	a, b
7	Бета-розподіл, тип I	η, λ
8	Бета-розподіл, тип II	η, λ, a, b
10	Розподіл арксинусу, тип II	b
11	Розподіл арксинусу, тип III	η, λ
12	Косинус-розподіл	b
13	Зсунений косинус-розподіл	b
14	Параболічний розподіл, тип I	b
15	Параболічний розподіл, тип II	a, b
16	Відображений параболічний розподіл	a, b
17	Закон рівномірно зростаючої щільності	a, b
18	Напівеліптичний розподіл	b
19	Розподіл Сімпсона, тип II	a, b

З даних, які наведено в табл. 9–11, витікає, що розробка спеціалізованого калькулятора істотно розширює можливості системи EXCEL в процесі ентропійного аналізу даних. Необхідність такого розширення обґрунтовано в [15].

Калькулятор, який запропоновано у даній роботі, побудовано засобами табличного процесора MS Excel. Діалог користувача з калькулятором реалізовано на основі стандартного для MS Excel інструмента "Користувацька Форма". На першому робочому аркуші книги Excel розміщено кнопку виклику переліку розподілів, при натисненні якої на екран виводиться перша форма (рис. 1) – перелік типів розподілів, усього 19 найменувань. Їх перелік наведено на рис. 1.

Завдяки використанню елементів керування формою типу "прапорець" (CheckBox) є можливість вибрати з переліку декілька розподілів, розставивши "прапорці" перед найменуваннями потрібних, і після натиснення кнопки "ОБЧИСЛИТИ" перейти до послідовного налаштування параметрів кожного з них.

При цьому для кожного типу розподілу на екран виводиться форма з власним набором параметрів. Для цілого ряду розподілів передбачається внесення даних від одного до кількох спостережень.

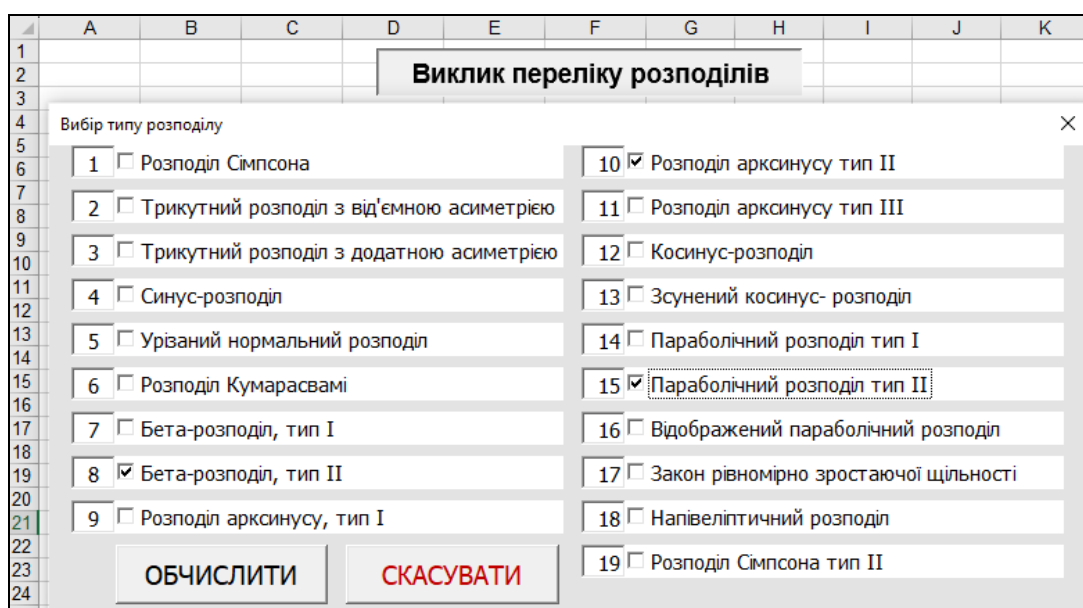


Рис. 1. Скрін-копія вікна виклику форми з переліком розподілів за допомогою відповідних кнопок (Fig. 1. Screen – a copy of the screen window of the form call screen with a list of distributions using the appropriate buttons)

Тому заголовок рядка параметрів першого спостереження у вигляді "d1" виділено чорним шрифтом, а другого і третього ("d2" і "d3") – сірим. Після початку введення даних другого спостереження його заголовок також стає чорним. Якщо перейти до введення даних третього спостереження, виконується скролінг, тобто, рядок першого спостереження зникає з вікна форми, а його місце займає рядок другого спостереження. При цьому на місці третього з'являється рядок четвертого спостереження і так далі. Кількість спостережень не обмежується.

Кнопка "ПУСК" завершує введення параметрів розподілу і забезпечує перехід до аркуша з обчисленими значеннями ентропії за всіма спостереженнями. Для кожного розподілу створюється окремий аркуш результатів. Це обумовлено особливостями різних розподілів і різним набором результатів, які можуть містити таблицю з обчисленою ентропією, еластичністю, а також графіки ентропії та еластичності (набір залежить від типу розподілу). При введенні параметрів таких розподілів користувач за допомогою "прапорців" може вибирати, у якому вигляді йому потрібні результати.

Визначення ентропії і її еластичності для функції, заданої в аналітичному вигляді, розглянемо на прикладі функції щільності розподілу арксинуса тип II (№10).

Результати визначення їх чисельного значення показано на рис. 2 та 3.

Рис. 2. Форма налаштування параметрів розподілу арксинуса тип II і бажаного переліку результатів обчислення (Fig. 2. Screen – a copy of the screen window of the form of setting the parameters of arcsine distribution type II and the desired list of calculation results)

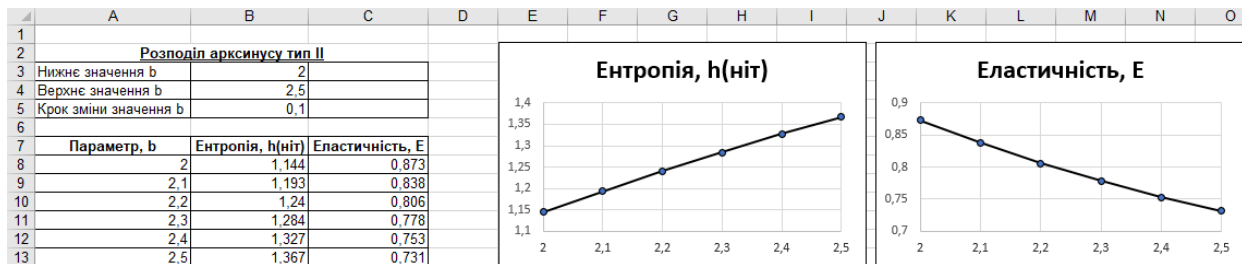


Рис. 3. Скрін-копія аркуша з результатами обчислень ентропії та її еластичності для розподілу арксинуса тип II (Fig. 3. Screen – a copy of the sheet with the results of calculations of entropy and its elasticity for the arcsine distribution type II)

Визначення ентропії для функцій, заданих в аналітичній формі, та їх еластичностей, заданих в табличній формі, розглянемо на прикладі функції щільності Бета-розподілу тип II (№8).

Результати обчислень їх чисельного значення показано на рис. 4 та 5.

Визначення ентропії і еластичності, заданих в табличному вигляді, розглянемо на прикладі функції параболічного розподілу тип II (№15).

Результати обчислення наведено на рис. 6 та 7.

Чисельне значення ентропії для функцій, заданих в табличному вигляді більш детально розглянемо на прикладі функції щільності параболічного розподілу тип II.

Рис. 4. Форма налаштування параметрів Бета-розподілу тип II (Fig. 4. Screen – a copy of the screen window of the form of setting the parameters of Beta distribution type II and the desired list of calculation results)

Спостереження	η	λ	a	b	Ентропія, h(ніт)
d1	3	1,5	1	2	0,273
d2	3	1,6	1	2	0,259
Еластичність відносно параметру					
η					
λ	*				-0,815
a					
b					

Рис. 5. Скрін-копія аркуша з результатами обчислень ентропії та її еластичності для Бета-розподілу тип II (Fig. 5. Screen – a copy of the screen window with the results of calculating the entropy and elasticity for the Generalized Beta Distribution)

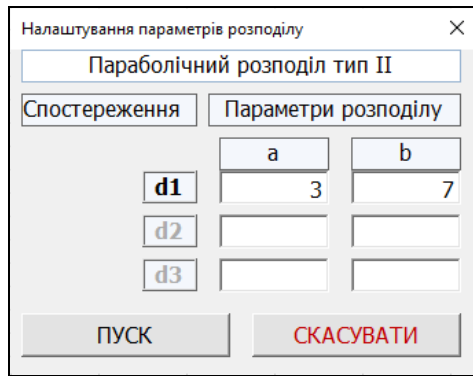


Рис. 6. Форма налаштування параметрів параболічного розподілу тип II

(Fig. 6. A screenshot of the sheet with the results of setting the parameters of the parabolic distribution type II)

	A	B	C	D
1				
2	Параболічний розподіл тип II			
3	Спостереження		Параметри розподілу	
4		a	b	Ентропія, h(ніт)
5	d1	3	7	1,261

Рис. 7. Скрін-копія аркуша з результатами обчислень ентропії та її еластичності для параболічного розподілу тип II

(Fig. 7. Screenshot – a copy of t with the results of calculating the entropy and elasticity for the U-shaped parabolic distribution Type II)

Використовуючи (1) і функцію щільності розподілу наведену в табл. 7, визначимо ентропію як співвідношення:

$$h = - \int_a^b \frac{6(x-a) \cdot (b-x)}{(b-a)^3} \ln \frac{6(x-a) \cdot (b-x)}{(b-a)^3} dx. \quad (28)$$

Для символічного інтегрування (28) використовували систему аналітичного програмування DERIVE 6. Результати приведені в наступних співвідношеннях:

$$h = \ln[(a-b)^2] - \frac{\Omega_1 + \Omega_2 + \Omega_3}{3(a-b)^3} \pm \frac{\pi \cdot ib^2(3a-b)}{(a-b)^3}; \quad (29)$$

$$\Omega_1 = (3a^2(a-3b)\ln(b-a)); \quad (30)$$

$$\Omega_2 = 3b^2(3a-b)\ln(a-b); \quad (31)$$

$$\Omega_3 = (a-b)^3(3\ln 6-5) = 0,3753(a-b). \quad (32)$$

Ентропія, що визначена у співвідношенні (29), представлена в комплексній формі, хоча за сучасними уявленнями ентропія – дійсне додатне число.

Слід зазначити, що інтегрування співвідношення (28) системою Mathcad 15 також дало результат у вигляді комплексного числа. Можна припустити, що подібне явище пов'язане з виникненням особливостей на нижній і верхній границях інтегрування співвідношення (28). Детальніше дослідження цього факту виходить за рамки даного повідомлен-

ня. Чисельне інтегрування співвідношення (28) виконано вказаними системами при $a=3$ і $b=7$ та дало однакові результати:

$$h = - \int_3^7 \left(\frac{6(x-3) \cdot (7-x)}{(7-4)^3} \times \ln \frac{6(x-3) \cdot (7-x)}{(7-4)^3} \right) dx = 1,261 \text{ (ніт)}. \quad (33)$$

Оскільки основне завдання роботи – отримання програмного продукту, що використовує тільки можливості системи EXCEL, то для обчислення значень ентропії параболічного розподілу тип II (№15) та відображеного параболічного розподілу (№16) використовували чисельне інтегрування методом трапецій у варіанті, який викладено в [14].

Для чисельного інтегрування співвідношення (28) з метою виключення можливості появи особливостей при обчисленні значень підінтегральної функції на її границях, представимо його у вигляді:

$$h = - \int_{a-\varepsilon}^{b+\varepsilon} \left(\frac{6[x-(a-\varepsilon)] \cdot [(b+\varepsilon)-x]}{(b-a)^3} \times \ln \frac{6[x-(a-\varepsilon)] \cdot [(b+\varepsilon)-x]}{(b-a)^3} \right) dx \quad (34)$$

за умови, що $\varepsilon = 1 \cdot 10^{-k}$; $k = 1, \dots, 4$.

Залежність величини ентропії від зміни меж інтегрування на величину ε наведена в табл. 12.

Таблиця 12 – Залежність величини ентропії від величин границь інтегрування

Величина k	Параметри розподілу		Величина ентропії h
	a	b	
1	2,9	7,1	1,297
2	2,99	7,01	1,266
3	2,999	7,001	1,262
4	2,9999	7,0001	1,261

Таким чином, при зміні інтервалу інтегрування на величину $\varepsilon = \pm 0,0001$ отриманий результат практично співпадає з результатами застосування систем DERIVE 6 і Mathcad 15.

Розглянемо застосування калькулятора для визначення довірчого інтервалу оцінки ентропії параболічного розподілу типу II. У [12] приведені оцінки параметрів цього розподілу, які отримані методом максимуму правдоподібності:

$$\hat{a} = \bar{x} - s\sqrt{5}; \quad \hat{b} = \bar{x} + s\sqrt{5}. \quad (35)$$

Прийнявши, як і в попередньому прикладі $a=3$ і $b=7$, отримаємо, що середньоквадратичне відхилення $S = 0,894$.

В [13] запропоновано визначати нижню межу S_n і верхню межу S_e довірчого інтервалу точкової оцінки S за співвідношеннями:

$$\begin{aligned} s_n &= k_n s; \\ s_e &= k_e s. \end{aligned} \quad (36)$$

Спосіб визначення цих коефіцієнтів і їх чисельні значення, в залежності від кількості спостережень, приведено в табл. 13.

Використовуючи співвідношення (15), (33), (34), визначимо нижню h_n і верхню h_e границю довірчих інтервалів точкової оцінки ентропії h , яка обчислена у співвідношенні (33).

Результати обчислень наведено в табл. 14.

З даних, які наведено в цій таблиці можна зробити наступний висновок.

Для умов чисельного прикладу (33):

– збільшення даних з 10 до 100 скорочує розмір довірчого інтервалу оцінки ентропії в 2,7 рази,

– збільшення даних з 100 до 1000 скорочує розмір довірчого інтервалу оцінки ентропії лише в 1,6 рази.

Наведений приклад свідчить, що запропонований в даному повідомленні калькулятор може бути використаний для планування експерименту, пов'язаного з ентропійним аналізом даних.

Таблиця 13 – Визначення чисельних значень коефіцієнтів s_n та s_e

Кількість спостережень, n	Кількість степенів свободи, df	Відсоткові точки розподілу, $\chi_{df,0,025}^2$	Величина $k_n = \sqrt{df / \chi_{df,0,025}^2}$	Відсоткові точки розподілу, $\chi_{df,0,975}^2$	Величина $k_e = \sqrt{df / \chi_{df,0,975}^2}$
10	9	19,022	0,687	2,700	1,825
100	99	128,422	0,878	73,610	1,161
1000	999	1088,487	0,958	913,301	1,046

Таблиця 14 – Нижня і верхня границі довірчого інтервалу точкової оцінки ентропії h

Кількість спостережень, n	Довірча ймовірність								Розмір довірчого інтервалу $\Delta = h_e - h_n$
	Нижня границя, $\alpha = 0,025$				Верхня границя, $\alpha = 0,975$				
	s	\hat{a}	\hat{b}	h_n	s	\hat{a}	\hat{b}	h_e	
10	0,615	3,545	6,454	0,940	1,631	1,353	8,647	1,862	0,922
100	0,784	3,142	6,857164	1,186	1,161	2,404	7,596	1,522	0,336
1000	0,856	2,973	7,026	1,205	1,046	2,660	7,338	1,417	0,212

Висновки

1. В роботі запропоновано EXCEL-орієнтований калькулятор для обчислення ентропії та її еластичності для функцій розподілу за умови обмеженої області визначення неперервної випадкової величини.

2. Всі використані в роботі функції розподілу розподілені на три категорії залежно від форми подання ентропії та її еластичності.

3. До першої категорії включено функції розподілу, для яких ентропія та її еластичність може бути визначена в аналітичній формі, а саме:

- розподіл Сімпсона,
- трикутний розподіл з від'ємною асиметрією,
- трикутний розподіл з додатною асиметрією,
- синус-розподіл,
- розподіл арксинусу тип I,
- розподіл арксинусу тип II,
- косинус-розподіл,
- зсунений косинус-розподіл,
- параболічний розподіл тип I,
- напівеліптичний розподіл,
- розподіл Сімпсона тип II.

4. До другої категорії включено функції розподілу для яких ентропія може бути визначена в аналітичній формі, а її еластичність може бути визначена в табличній формі, а саме:

- урізаний нормальний розподіл,
- розподіл Кумарасвами,
- Бета-розподіл тип I,
- Бета-розподіл тип II,
- закон рівномірно зростаючої щільності, розподіл арксинусу тип III.

5. До третьої категорії включено функції розподілу, для яких ентропія та її еластичність можуть бути визначені в табличній формі, а саме:

- параболічний розподіл тип II,
- відображений параболічний розподіл.

6. Наведено основні відомості про структуру запропонованого калькулятора та приклади його застосування.

7. Викладено методику визначення довірчого інтервалу точкової оцінки ентропії.

8. Наведено приклад використання запропонованого в даному повідомленні калькулятора для планування експерименту, пов'язаного з ентропійним аналізом даних.

REFERENCES

1. Martin, N. and England, J. (1981), *Mathematical Theory of Entropy*, Addison – Wesley Publishing Company, London, 350 p., available at: <https://www.osti.gov/biblio/5157083>.
2. Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994), *Continuous Univariate Distributions*, Vol. 1, JohnWiley&sons, inc, New York, 706 p., doi: <http://dx.doi.org/10.1080/00401706.1995.10484392>.
3. Gadetska, S.V. and Dubnitskiy, V. Yu. (2018), “Vyznachenja dyferencialnoji entropiji vypadkovoji velychyny, jaka zadana urizanyim pokaznykovym rozpodilom [Deifinon of the differential entropy of a random value given by a truncated distribution]”, *Control, Navigation and Communication Systems*, Vol. 5, Issue 51, pp. 69–73, doi: <http://dx.doi.org/10.26906/SUNZ.2018.5.069>.
4. Michlowicz, J. V., Nichols, J. M. and Bucholtz, F. (2014), *Handbook of differential entropy*, A. CHAPMAN & HALL, New York, 220 p., doi: <https://doi.org/10.1201/b15991>.
5. Dubnytskyi V. Y. and Skorikova I. H. (2015), “The determination of distribution parameters of one-dimensional continuous random variable according to its initial characteristics by financial risks modelling”, *Financial and Credit Activity Problems of Theory and Practice*, No1(18), pp. 212-225, doi: <http://dx.doi.org/10.18371/FCAPTP.V1I18.46129>.
6. Gadetska, S., Dubnitskiy, V. and Kobulin, A. (2018), “Parametrychna elastychnist entropii Shenmona, Tsallisa ta Reni [Parametric elasticity of entropy of Shannon, Tsallis and Renyi]”, *Control, Navigation and Communication Systems*, Issue 3(49), pp. 61-66, doi: <http://dx.doi.org/10.26906/SUNZ.2018.3.061>.
7. (1972), *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, Edited by Milton Abramowitz and Irene A. Stegun, National Bureau of Standards, USA, Washington, 1046 p., available at: https://personal.math.ubc.ca/~cbm/aands/abramowitz_and_stegun.pdf.
8. Zamkov, O. O., Tolstopyatenko, A. V. and Cheremykh, Yu. N. (2001), “*Matematicheskiye metody v ekonomike*” [*Mathematical methods in economics*], Delo i Servis, 368 p., available at: <https://www.at.alleng.org/d/econ/econ186.htm>.
9. Lewin, L. (1981), *Polylogarithms and associated functions*] NORTH HOLLAND, New York • Oxford, 359 p., available at: <https://www.worldcat.org/title/polylogarithms-and-associated-functions/oclc/6666619>.
10. Kumaraswamy, P. (1980), “A generalized probability density function for double-bounded random processes”, *Journal of Hydrology*, Vol. 46 (1-2), pp. 79-88, doi: [http://dx.doi.org/10.1016/0022-694\(80\)90036-0](http://dx.doi.org/10.1016/0022-694(80)90036-0).
11. Horst, Rinne (2010), *Location-Scale Distributions, Linear Estimation and Probability, Plotting Using MATLAB*, Department of Economics and Management Science, Justus-Liebig-University, Giessen, 260 p., available at: http://geb.uni-giessen.de/geb/volltexte/2010/7607/pdf/RinneHorst_LocationScale_2010.pdf.
12. Vadzinsky, R. N. (2001), “*Spravochnik po veroyatnostnym raspredelenijam*” [*Handbook of Probability Distributions*], NAUKA, Moscow, 295 p.
13. Hemelrijk, J. (1977), *Orienterende Cursus Mathematische Statistiek*, Mathematisch Centrum, Amsterdam, 247 p., available at: <https://core.ac.uk/download/pdf/301666803.pdf>.
14. Kundrat, A. M. and Kundrat, M. M. (2014), “*Naukovo – tekhnichni obchyslennia zasobamy MathCAD ta MS EXCEL [Scientific-technical calculations by means of Mathsad and MC EXCEL]*”, NUVHP, Rivno, 2014. 252 p. available at: <https://ep3.nuwm.edu.ua/1760/1/734733%20zah.pdf>.
15. Gadetska, S., Dubnitskiy, V., Khodyrev, A. and Kushneruk, Yu. (2023), “Excel-orientovanyi kalkuliator dlia obchyslennia rezultativ entropiinoho analizu danykh, shcho rozpodileni po katehoriakh [Excel-oriented calculator for calculating results of entropy analysis of data distributed by categories]”, *Advanced Information Systems*, Vol. 7, No. 2, pp. 28–40, doi: <https://doi.org/10.20998/2522-9052.2023.2.05>.

СПИСОК ЛІТЕРАТУРИ

1. Martin N., England J. *Mathematical Theory of Entropy*. London: Addison – Wesley Publishing Company, 1981. 350 p. URL: <https://www.osti.gov/biblio/5157083>.
2. Johnson N. L., Kotz S., Balakrishnan, N. *Continuous Univariate Distributions, Volume 1*. New York: JohnWiley&sons, inc, 1994. 706 p. DOI: <http://dx.doi.org/10.1080/00401706.1995.10484392>.
3. Гадецька С. В., Дубницький В. Ю. Визначення диференціальної ентропії випадкової величини, яка задана урізаним показниковим розподілом. *Системи управління, навігації та зв'язку*. Полтава: ПНТУ, 2018, вип. 5(51). С. 69–73. DOI: <http://dx.doi.org/10.26906/SUNZ.2018.5.069>.
4. Michlowicz J. V., J. M. Nichols J. M., Bucholtz F. *Handbook of differential entropy*. New York: A. CHAPMAN & HALL, 2014. 220 p. DOI: <https://doi.org/10.1201/b15991>.
5. Dubnytskyi V. Y., Skorikova I. H. The determination of distribution parameters of one-dimensional continuous random variable according to its initial characteristics by financial risks modelling. *Financial and Credit Activity Problems of Theory and Practice*. 2015. No. 1(18). P. 212-225. DOI: <http://dx.doi.org/10.18371/FCAPTP.V1I18.46129>.
6. Гадецька С. В., Дубницький В. Ю., Кобилін А. М. Параметрична еластичність ентропії Шеннона, Тсалліса та Ренї. *Системи управління, навігації та зв'язку*. Полтава: ПНТУ, 2018. Вип. 3(49). С. 61-66. DOI: <http://dx.doi.org/10.26906/SUNZ.2018.3.061>.
7. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. Edited by Milton Abramowitz and Irene A. Stegun. Washington: National Bureau of Standards, USA, 1972. 1046 p. URL: https://personal.math.ubc.ca/~cbm/aands/abramowitz_and_stegun.pdf.
8. Замков О. О., Толстопятенко А. В., Черемых Ю. Н. *Математические методы в экономике*. М.: Дело и Сервис, 2001. 368 с. URL: <https://www.at.alleng.org/d/econ/econ186.htm>.
9. Lewin L. *Polylogarithms and associated functions*. New York, Oxford : NORTH HOLLAND, 1981. 359 p. URL: <https://www.worldcat.org/title/polylogarithms-and-associated-functions/oclc/6666619>.
10. Kumaraswamy P. A generalized probability density function for double-bounded random processes. *Journal of Hydrology*. Vol. 46 (1-2). P. 79-88. DOI: [http://dx.doi.org/10.1016/0022-694\(80\)90036-0](http://dx.doi.org/10.1016/0022-694(80)90036-0).

11. Horst Rinne. Location–Scale Distributions, Linear Estimation and Probability, PlottingUsing MATLAB. Giessen: Justus-Liebig-University, Department of Economics and Management Science, 2010. 260 p. http://geb.uni-giessen.de/geb/volltexte/2010/7607/pdf/RinneHorst_LocationScale_2010.pdf.
12. Вадзинский Р. Н. Справочник по вероятностным распределениям. М.: НАУКА, 2001. 295 с. URL: <https://read.in.ua/book141065>.
13. Hemelrijk J. Orienterende Cursus Mathematische Statistiek. Amsterdam : Mathematisch Centrum, 1977. 247 p. URL: <https://core.ac.uk/download/pdf/301666803.pdf>.
14. Кундрат А. М., Кундрат М. М. Науково-технічні обчислення засобами MathCAD та MS EXCEL. Рівне: НУВГП, 2014. 252 с. URL: <https://ep3.nuwm.edu.ua/1760/1/734733%20zah.pdf>.
15. Гадецька С. В., Дубницький В. Ю., Кушнерук Ю. І., Ходирев О. І. Excel-орієнтований калькулятор для обчислення результатів ентропійного аналізу даних, що розподілені по категоріях. *Сучасні інформаційні системи*. 2023. Т. 7, № 2. С. 28-40. DOI: <https://doi.org/10.20998/2522-9052.2023.2.05>.

Received (Надійшла) 17.05.2023

Accepted for publication (Прийнята до друку) 23.08.2023

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Гадецька Світлана Вікторівна – кандидат фізико-математичних наук, доцент, доцент кафедри вищої математики, Харківський національний автомобільно-дорожній університет, Харків, Україна;

Svitlana Gadetska – PhD in Physics and Mathematics, Associate Professor, Associate Professor of Department of Higher Mathematics of Kharkiv National Automobile and Highway University, Kharkiv, Ukraine;

e-mail: svgadetska@ukr.net; ORCID ID: <https://orcid.org/0000-0002-9125-2363>.

Дубницький Валерій Юрійович – кандидат технічних наук, старший науковий співробітник, старший науковий співробітник Харківського навчально-наукового інституту “Каразінський банківський інститут” Харківського національного університету ім. В. Н. Каразіна, Харків, Україна;

Valeriy Dubnitskiy – PhD in Engineering Senior Researcher Senior Researcher of “Karazin Banking Institute” of V.N. Karazin Kharkiv National University, Kharkiv, Ukraine;

e-mail: dubnitskiy@gmail.com; ORCID ID: <https://orcid.org/0000-0003-1924-4104>.

Кушнерук Юрій Іонович – кандидат технічних наук, доцент, доцент Інституту цивільної авіації Харківського національного університету Повітряних Сил імені Івана Кожедуба, Харків, Україна;

Yuri Kushneruk – Candidate of Technical Sciences, Associate Professor, Associate professor of Civil Aviation Institute of Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine;

e-mail: kyshneryk_ui@ukr.net; ORCID ID: <https://orcid.org/0000-0001-5844-7137>.

Ходирев Олександр Іванович – старший викладач Харківського навчально-наукового інституту “Каразінський банківський інститут” Харківського національного університету ім. В. Н. Каразіна, Харків, Україна;

Alexander Khodyrev – Senior Lecturer of “Karazin Banking Institute” of V.N. Karazin Kharkiv National University, Kharkiv, Ukraine;

e-mail: khodyrevmjk3758@gmail.com; ORCID ID: <https://orcid.org/0000-0001-9871-9440>.

**EXCEL-oriented procedures for determining the entropy of a distribution function
and its relative parametric sensitivity (elasticity) under the conditions
of two-sided restrictions on the range of values of a continuous random variable**

Svitlana Gadetska, Valeriy Dubnitskiy, Yuri Kushneruk, Alexander Khodyrev

Abstract. The goal of the work. Development of an EXCEL-oriented calculator for calculating entropy and its elasticity for distribution functions under the condition of a limited domain of definition of a continuous random variable. **Subject of study.** Probability density functions and their entropy under the condition of bilateral restrictions on the domain of possible values of random variables. **Research methods.** Algorithmic and numerical analysis of procedures for obtaining numerical values of entropy of density functions of continuous random variables under the condition of two-way restrictions on the area of its definition. **The obtained results.** The work proposes an EXCEL-oriented calculator for calculating entropy and its elasticity for distribution functions under the condition of a limited area of definition of a continuous random variable. All distribution functions used in the work are divided into three categories depending on the form in which entropy and its elasticity are presented. The first category includes distribution functions for which entropy and its elasticity can be determined analytically, namely: Simpson distribution, Right-angled and negatively skew version distribution, Right-angled and positively skew version distribution, Sine Wave Distribution, Ark – Sine Distribution Type I, Ark – Sine Distribution Type II, Ordinary Cosine Distribution, Raised Cosine Distribution, U-shaped parabolic distribution Type I, Semi-elliptical distribution, Ark – Sine Distribution Type II. Another category includes distribution functions for which entropy can be defined in the analytical group, its elasticity can be defined in the tabular part, namely: Truncated Normal Distribution, Kumaraswamy Distribution, Beta Distribution, Generalized Beta Distribution, Uniformly increasing distribution density, Ark-Sine Distribution Type III. The third category includes distribution functions for which entropy and its elasticity can be determined in tabular form. Basic information about the structure of the proposed calculator and examples of its application are provided. The method of determining the confidence interval of the point estimate of entropy is described. An example of using the calculator proposed in this message for planning an experiment related to entropy data analysis is given.

Keywords: entropy; probability density functions, bilateral restrictions on the range of values of random variables; relative parametric sensitivity (elasticity); programmable calculators; entropy analysis.