УДК 51-77: 519.224: 372.851: 343.9.018

Хаврук В. О.

АНАЛІЗ ВАРІАЦІЙНОГО РЯДУ ТА ЙМОВІРНІСНИЙ ЗАКОН РОЗПОДІЛУ ЗЛОЧИНІВ ЗА ВІКОМ ПРАВОПОРУШНИКІВ

В статті в якості прикладної спрямованості теорії ймовірностей та математичної статистики обрано таке соціальне явище як злочинність, зокрема розглядається вікова структура правопорушників в Україні. Проведений аналіз варіаційного ряду розподілу злочинів за віком правопорушників на основі усереднених показників за 2017 р. в Україні. Розраховані середня арифметична, мода та коефіцієнт асиметрії ряду. Здійснено вирівнювання ряду злочинності за нормальним законом – отримана функція щільності розподілу ймовірності та побудовані відповідні гістограми та графічні залежності. Перевірено відповідність реального варіаційного ряду нормальному закону за критеріями χ^2 Пірсона, Романовського, Колмогорова. Запропонована методологія дослідження на відповідність статистичного розподілу випадкових величин теоретичному закону (розподіл Ґауса) вимагає таких основних знань, як: інтервальна оцінка параметрів випадкової величини, характеристики і властивості нормального закону розподілу, етапи перевірки гіпотези про відповідність статистики прийнятому ймовірнісному закону, основні критерії відповідності, програмні середовища Excel та Mathcad.

Ключові слова: *варіаційний ряд, випадкова величина, гіпотеза, ексцес, інтервал, ймовірність, коефіцієнт асиметрії, мода, нормальний закон розподілу (розподіл Ґауса), середньоквадратичне відхилення, середньоарифметичне відхилення.*

В закладах вищої освіти при вивчені циклу таких дисциплін, як: економіка, соціологія, кримінологія, вища математика значна увага приділяється питанню побудови математичних моделей, які з високою точністю описують явища соціальної дійсності. Сучасна наука вимагає від дослідника досконалого володіння математико-статистичним апаратом обробки та аналізу даних. На сьогодні важко собі уявити фундаментальне дослідження явищ соціальної дійсності без побудови описових та пояснювальних моделей, що теж передбачає глибоку обізнаність фахівців щодо засобів реалізації такого моделювання. Такі засоби вимагають від фахівців знань не тільки елементарних основ математики, але і поглибленого розуміння теорії ймовірностей та статистичного аналізу даних, без чого неможлива розробка нових та вдосконалення існуючих методів обробки й аналізу інформації.

У зв'язку з цим, актуальним напрямом при вивченні курсу теорії ймовірностей і математичної статистики в закладах вищої освіти є її практична спрямованість, а саме розгляд основних положень теорії з використанням статистичних даних по найважливіших соціальних явищах соціальної дійсності (наприклад злочинності). А для підвищення ефективності засвоєння курсу теорії ймовірностей і математичної статистики в частині застосування ймовірнісних законів розподілу випадкових величин (наприклад, нормального закону) необхідно навести відповідну методологію, яка б містила чітко обґрунтовані послідовні етапи дослідження явища соціальної дійсності щодо застосування теоретичного закону розподілу до варіаційного ряду.

Варіаційний ряд є групуванням за однією ознакою і з єдиним показником – змінним числом одиниць сукупності, виражених в абсолютних або відносних величинах. В юридичній сфері злочинність (кількість злочинів) групуються за регіонами, віком правопорушників, часовими проміжками вчинення злочинних діянь і т.д. Широкі можливості для аналізу варіаційних рядів злочинності дають методи теорії ймовірностей та математичної статистки.

Дослідження варіаційних рядів злочинності має важливе як теоретичне так і практичне значення. Основною проблемою обмеженого використання ймовірнісних законів розподілу для вирівнювання варіаційних рядів злочинності є складність математичних розрахунків та необхідність застосування спеціалізованих комп'ютерних програм. Зважаючи на це, правознавці, аналізуючи варіаційний ряд, обмежуються лише знаходженням таких характеристик ряду, як: мода, медіана, середнє значення. Такий підхід є обмеженим, неповним і не дозволяє проводити порівняльну характеристику різних варіаційних рядів злочинності – як наслідок, виявляти основні напрями діяльності правоохоронних органів і впливу на особистість правопорушників.

Аналіз останніх досліджень і публікацій свідчить, що правознавці визнають той факт, що в дослідженнях злочинності досить часто використовуються закони розподілу ймовірнісних величин. Так, наприклад, О. О. Гаврилов, використовуючи дані В. Г. Крупіна, наводить нормальний закон розподілу кількості вбивств та асиметричний розподіл (не даючи йому конкретну назву) кількості крадіжок на 100 тис. жителів за регіонами, а також вказує на застосуванні закону Пуасона для аналізу розподілу автотранспортних аварій, вбивств [1, с. 80-81]. Лунєєв В. В., хоча і акцентує увагу на необхідності моделювання рядів розподілу, але вказує лише на можливість вирівнювання фактичного розподілу за нормальним законом і при цьому, не обґрунтовує саму методологію отримання ймовірнісних законів [2, с. 267-268]. Більш послідовним є дослідження ймовірності вчинення повторно злочинів (рецидив), проведене М. О. Ларченко, з використанням програми STATISTICA за методом множильних оцінок Каплана-Мейєра, в результаті були отриманні закони розподілу Вейбулла [3].

Необхідно зауважити, що наукові публікації щодо застосування ймовірнісних законів розподілу для вирівнювання варіаційних рядів злочинності є поодинокими, а це, в свою чергу, вказує на необхідність проводити дослідження в даному напрямі.

Метою статті є аналіз варіаційного ряду – розподілу злочинності за віком правопорушників, знаходження ймовірнісного закону розподілу та перевірка його відповідності статистичним даним.

В якості об'єкта дослідження обраний загальний рівень злочинності та вік правопорушників в Україні на основі статистичних даних за 2017 р. Використовуючи, статистичні дані щодо віку правопорушників в Україні [4], юридичну практику, досвід та рекомендації [2, с. 262] можна сформувати усереднені показники у % розподілу злочинів за віком правопорушників (табл. 1).

Таблиця 1

Злочини, вчинені особами за віком в Україні (2017 р.)

вік, років	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
Злочини, %	1	7	12	19	26	21	7	4	2	1

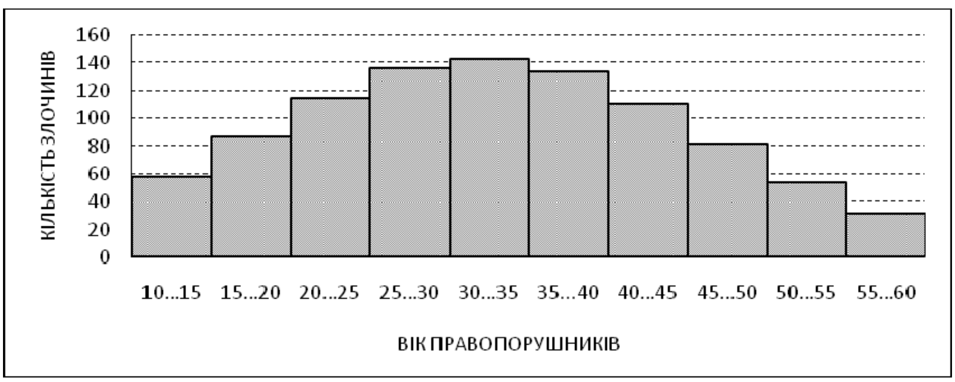

Рис. 1. Кількість злочинів за 2017 р. в Україні, вчинених особами за віком

Звернемося до загальновідомого варіаційного ряду – розподілу злочинів за віком їх суб'єктів за усередненими показниками (табл. 1; рис. 1). Представлений в табл. 1 інтервальний варіаційний ряд відбиває цілком певний зв'язок між варіюючим віком і зміною частот (відсотками осіб, що скоїли злочини). За даними світової, української і регіональної статистики спостерігається практично одна і та ж тенденція розподілу правопорушників за віком: з початку віку кримінальної відповідальності йде зростання злочинної активності, в 25-40 років (з деякими коливаннями) її рівень досягає апогею, а потім настає зниження. У цьому проявляється певна закономірність зміни частот у варіаційних рядах, що називається закономірністю розподілу, яка виявляється у великих сукупностях, де випадкові відхилення взаємно усуваються.

У виявленні реальних закономірностей розподілу полягає основна суть аналізу варіаційних рядів. Усі варіації, підпорядковуючись в своїй основі вказаній закономірності, мають багато типів особливостей (відхилень), кожна з яких пов'язана з тими або іншими причинами, встановлення яких відіграє важливу роль в статистичному аналізі.

Обставини, що визначають тип закономірностей розподілу, вивчаються на основі якісного (кримінологічного, кримінально-правового, кримінально-процесуального, адміністративно правового, цивільно-правового і т.д.) аналізу суті того або іншого явища, а саме – тих його властивостей і умов, які визначають мінливість варіюючої ознаки. Але до такого вивчення призводить лише виявлений тип закономірностей рядів розподілу.

Звернемося до даних табл. 1. Питома частка злочинців із збільшенням їх віку зростає (пряма залежність), але, досягнувши якогось рівня, незважаючи на подальше збільшення віку, знижується до мінімуму (зворотна залежність). Проте максимум питомої ваги (мода) знаходиться не посередині ряду (35 років), а зміщується до віку 32-33 років. Близько до моди розташовуються частки 25-30 років, 35-40 років і тільки потім слідує 40-45 років.

Таке зміщення до більш молодшого віку пояснюється тим, що особи даної вікової групи, не маючи стійких позитивних орієнтацій, потрапивши в складні життєві ситуації і прагнучи поліпшити своє матеріальне становище, вступають в конфлікт із законом частіше, ніж старші люди або люди більш зрілого віку. Отже, пояснення даного зрушення лежить не у фізіологічних, а в соціальних особливостях вікового характеру [2, с. 263]. Наведені пояснення лежать за межами юридичної статистики, але до них важко прийти на основі тільки логічних висновків, навіть в цьому нескладному питанні. Для цього потрібно виявити особливості реального статистичного розподілу значень ознаки. Щоб зафіксувати характер наявних відхилень, потрібно зіставити реальний розподіл з якимсь його еталоном. Такий еталон – теоретична крива розподілу, яка виражає загальну закономірність розподілу, що виключає вплив випадкових чинників. Ця крива розподілу називається кривою Лапласа-Ґауса, або нормальним розподілом. В якості еталону використовується також розподіл Пуассона і деякі інші, але вони практично не застосовуються в юридичній статистиці.

Нормальний розподіл виражається формулою [5, с. 25; 6, с. 125]:

$$P_x = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}, \qquad (1)$$

де P_x – крива нормального розподілу; x – варіанти; $\bar{x}$ – середня арифметична варіант; σ – середньоквадратичне відхилення; e і π – математичні постійні: $e \approx 2{,}7182$ і $\pi \approx 3{,}1415$.

Зрештою крива нормального розподілу залежить тільки від двох параметрів: середньої арифметичної (x) і середньоквадратичного відхилення (σ). Від їх значень залежить розташування центру розподілу кривої на вісі x і відмінності варіантів біля цього центру (рис. 2*а* і 2*б*), а також певні асиметрії лівої і правої гілок відносно центру (рис. 2*в* і 2*г*).

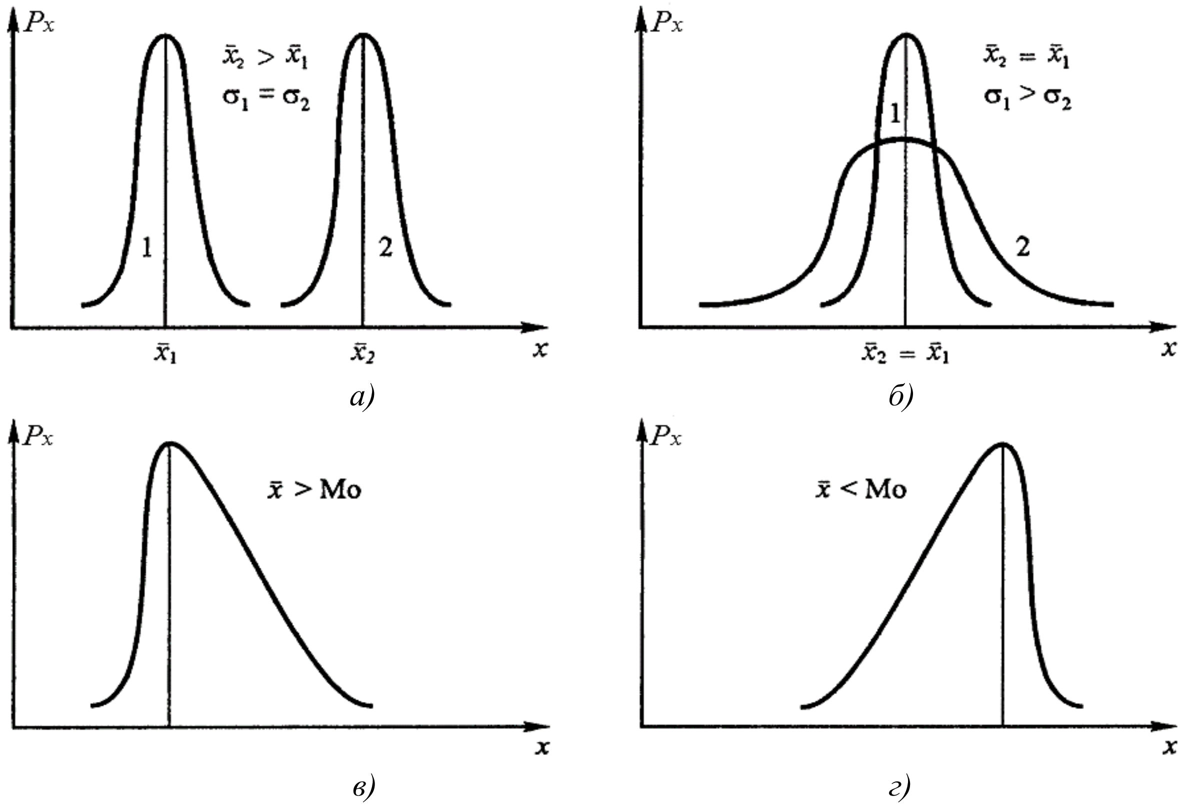

Рис. 2. Криві розподілу

У нормальному розподілі ліва і права гілки кривої симетричні, а середня арифметична, мода і медіана рівні. Проте при дотриманні цієї рівності криві можуть істотно розрізнятися між собою.

Якщо середня арифметична величина ($\overline{x}$) невелика, то крива розташовується ближче до осі ординат (P_x), якщо – велика, то крива зміщена вправо від осі Р*x* (рис. 2а, криві 1 і 2).

Якщо середньоквадратичне відхилення (*σ*) велике, то крива розподілу є високовершинною (рис. 2*б*, крива 1), що свідчить про скупчення частот в середині, і про типовість і надійність середньої. Таке положення в статистиці називають позитивним ексцесом.

Якщо середньоквадратичне відхилення невелике, то крива розподілу є низьковершинною (рис. 2*б*, крива 2), що свідчить про значну розкиданість частот ряду і недостатню надійність середньої. В статистиці вказані особливості називають негативним ексцесом.

Нормальний розподіл симетричний по відношенню до середньої арифметичної величини ($\overline{x}$). Проте симетричних реальних розподілів набагато менше, ніж асиметричних. В асиметричному розподілі середня арифметична, мода і медіана не співпадають, і їх відхилення однієї від іншої вимірюються за допомогою коефіцієнта асиметрії (*КА*), який розраховується за формулою [2, с. 265]:

$$KA = \frac{\overline{x} - Mo}{\sigma}, \qquad (2)$$

де *КА* – коефіцієнт асиметрії; $\overline{x}$ – середня арифметична; *Mo* – мода; *σ* – середньоквадратичне відхилення.

З формули слідує:

– якщо середня арифметична більша моди ($\overline{x}$ > *Mo*), то коефіцієнт асиметрії додатний, і це означає правосторонню асиметрію, тобто права частина кривої виявляється довша за ліву (рис. 2*в*);

– якщо середня арифметична менша моди ($\overline{x} < Mo$), то коефіцієнт асиметрії буде зі знаком мінус (від'ємний), що означає лівосторонню асиметрію, тобто ліва частина кривої довша за праву (рис. 2*г*).

В нашому випадку (табл. 1), найбільша частота скоюваних злочинів припадає на інтервал 30-35 років (рис. 1). З цього можна припустити, що ми маємо справу з від'ємним коефіцієнтом асиметрії.

Модальний інтервал дорівнює 30-35 рокам, якому відповідає 26% частоти вчинення злочинів. Мода (*Mo*) в модальному інтервалі розраховується за формулою [2, с. 266]:

$$Mo = Xo + h\frac{f_{Mo} - f_1}{(f_{Mo} - f_1) + (f_{Mo} - f_2)}, \qquad (3)$$

де *Xo* = 30 років – нижня межа модального інтервалу; h –довжина інтервалу (*h* = 5 років); f_{Mo} = 26 – частота модального інтервалу; f_1 = 19 – частота інтервалу, що передує модальному; f_2 = 21 – частота інтервалу, що слідує за модальним.

На основі наведених даних отримуємо:

$$Mo = 30 + 5 \cdot \frac{26 - 19}{(26 - 19) + (26 - 21)} \approx 32,91667$$ роки.

Величина $\overline{x} = 32,1$ року – середня арифметична інтервального ряду. Для її знаходження спочатку визначається середина кожного інтервалу шляхом складання двох його меж і ділення отриманої суми на два (наприклад, (10 +14) : 2 = 12); потім середину кожного інтервалу множимо на його частоту (12 · 0,2 = 2,4); після цього отримані добутки складаємо (загальна сума добутків середини інтервалів на частоту дорівнює 3210); розділивши цю суму (3210) на загальну суму частот (100), ми отримаємо середню арифметичну, рівну 32,1 року (табл. 2).

Таблиця 2

Розрахунок середнього арифметичного відхилення

Вік (*x*), років	**Частка злочинів (*f*)***	**Середина інтервалу ($x_{сер.}$)**	**Добуток ($f \cdot x_{сер.}$)**	**Відхилення** $(x_{сер.} - \overline{x})$	**Дисперсія** $(x_{сер.} - \overline{x})^2$
10-15	1	12,5	12,5	-19,6	384,16
15-20	7	17,5	122,5	-14,6	213,16
20-25	12	22,5	270	-9,6	92,16
25-30	19	27,5	522,5	-4,6	21,16
30-35	26	32,5	845	0,4	0,16
35-40	21	37,5	787,5	5,4	29,16
40-45	7	42,5	297,5	10,4	108,16
45-50	4	47,5	190	15,4	237,16
50-55	2	52,5	105	20,4	416,16
55-60	1	57,5	57,5	25,4	645,16
	Σ*f* = 100		Σ = 3210	x	Σ=2146,6

Це означає, що середня арифметична менша моди ($\overline{x} < Mo$ або 32,1 < 32,917), тобто ми маємо справу з незначною лівосторонньою асиметрією і від'ємним коефіцієнтом асиметрії. Для розрахунку *КА* необхідно знати середньоквадратичне відхилення. Знайдемо його з табл. 2, за формулою:

$$\sigma=\sqrt{\frac{\sum_{i=1}^{n}(x_{сер.}-\bar{x})^2}{n}}=\sqrt{\frac{2146,6}{10}}=14,651, \qquad (4)$$

де *n* – кількість інтервалів. Отже:

$$KA=\frac{\bar{x}-Mo}{\sigma}=\frac{32,1-32,91667}{14,651}=\frac{-0,81667}{14,651}=-0,05574.$$

Отже, при наявних даних $\bar{x}=32,1$ і *Mo* = 32,91667, звідки $\bar{x}$ < *Mo* та $\bar{x}$ – *Mo* = -0,81667, ми матимемо графік з незначною лівосторонньою асиметрією і від'ємним коефіцієнтом *KA* = -0,05574.

Але мода та коефіцієнт симетрії лише характеризують статистичні дані. А для з'ясування характеру розподілу випадкової величини необхідно встановити закон розподілу. Розподіл злочинів за віком правопорушників (рис. 1) дає змогу висунути припущення (гіпотезу), про нормальний закон розподілу (розподіл Ґауса).

Знаходження нормального закону розподілу ймовірнісної величини будемо здійснювати, виконуючи наступні етапи:

а) побудова полігону та гістограми відносних частот;

б) знаходження емпіричної функції розподілу й побудова її графіку;

в) обчислення числових характеристик вибірки: вибіркову середню, вибіркову дисперсію, вибіркове середнє квадратичне відхилення;

г) припускаючи, що досліджувана випадкова величина розподілена за нормальним законом, запис щільності ймовірності випадкової величини x і побудова її графіку на одному рисунку з гістограмою відносних частот (графік вирівнюючої кривої);

д) знаходження теоретичних частот нормального закону розподілу та при рівні значимості α = 0,05 перевірка узгодження гіпотези про відповідність нормальному розподілу (розподілу Ґауса) за відомими критеріями.

е) знаходження інтервальних оцінок параметра a нормального закону розподілу при довірчій ймовірності 0,95.

Порядок виконання етапів.

а) Побудова гістограми й полігону відносних частот.

Дані щодо розподілу злочинів за віком правопорушників представлені у вигляді інтервального статистичного ряду з усередненими показниками (табл. 1). Використовуючи статистичні дані табл. 1, розраховуються відносні частоти – $w=\frac{n}{100}$ (табл. 3), для наочності можна представити відносні частоти у вигляді графіків (рис. 3; рис. 4).

Таблиця 3

Інтервали і частоти вибірки

Інтервали *a*	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
Частоти *n*	1	7	12	19	26	21	7	4	2	1
Відносні частоти *w*	0,01	0,07	0,12	0,19	0,26	0,21	0,07	0,04	0,02	0,01

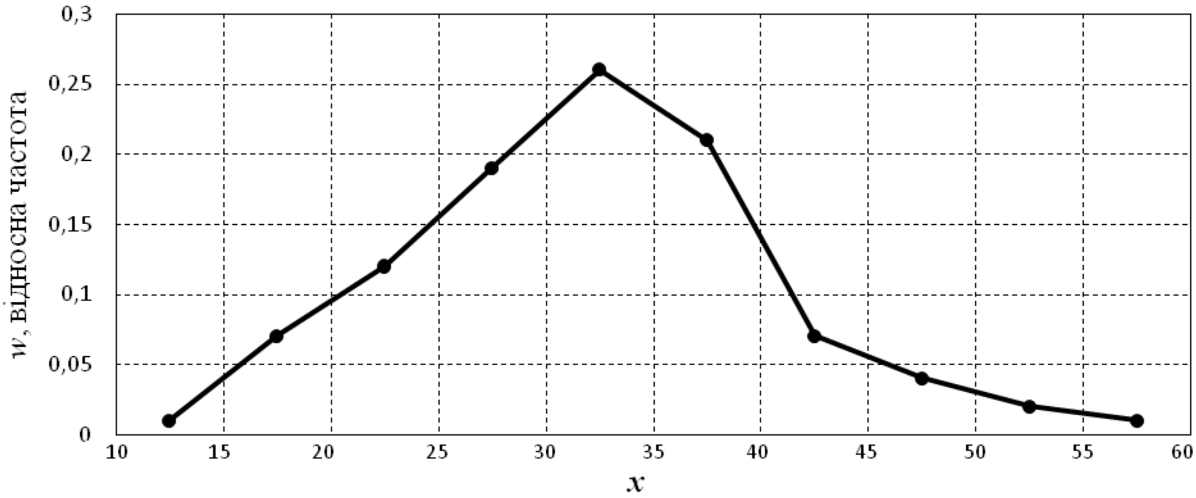

Рис. 3. Полігон відносних частот

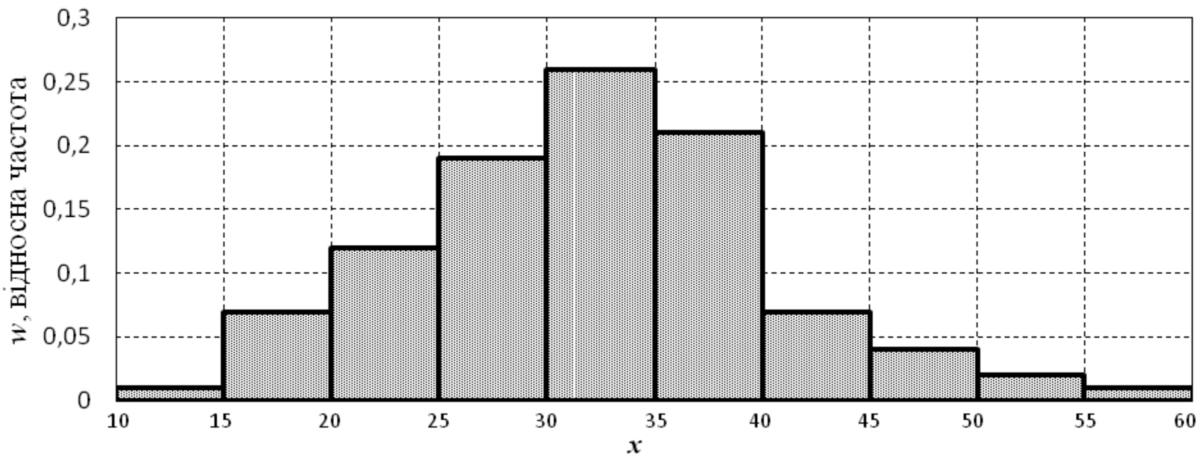

Рис. 4. Гістограма відносних частот

б) Побудуємо емпіричну (за статистичними даними) функцію розподілу випадкової величини x:

при $x \leq 10$, $F(x) = 0$; при $10 < x \leq 15$, $F(x) = 0 + 0{,}01 = 0{,}01$;

при $15 < x \leq 20$, $F(x) = 0{,}01 + 0{,}07 = 0{,}08$; при $20 < x \leq 25$, $F(x) = 0{,}08 + 0{,}12 = 0{,}20$;

при $25 < x \leq 30$, $F(x) = 0{,}20 + 0{,}19 = 0{,}39$; при $30 < x \leq 35$, $F(x) = 0{,}39 + 0{,}26 = 0{,}65$;

при $35 < x \leq 40$, $F(x) = 0{,}65 + 0{,}21 = 0{,}86$; при $40 < x \leq 45$, $F(x) = 0{,}86 + 0{,}07 = 0{,}93$;

при $45 < x \leq 50$, $F(x) = 0{,}93 + 0{,}04 = 0{,}97$; при $50 < x \leq 55$, $F(x) = 0{,}97 + 0{,}02 = 0{,}99$;

при $x > 55$, $F(x) = 0{,}99 + 0{,}01 = 1{,}00$.

Отже, маємо масив:

$$F(x)=\begin{cases} 0 & \text{при} \quad x \le 10; \\ 0,01 & \text{при} \quad 10 < x \le 15; \\ 0,08 & \text{при} \quad 15 < x \le 20; \\ 0,20 & \text{при} \quad 20 < x \le 25; \\ 0,39 & \text{при} \quad 25 < x \le 30; \\ 0,65 & \text{при} \quad 30 < x \le 35; \\ 0,86 & \text{при} \quad 35 < x \le 40; \\ 0,93 & \text{при} \quad 40 < x \le 45; \\ 0,97 & \text{при} \quad 45 < x \le 50; \\ 0,99 & \text{при} \quad 50 < x \le 55; \\ 1,00 & \text{при} \quad x > 55 \end{cases}$$

Зображаємо графік функції розподілу (рис. 5):

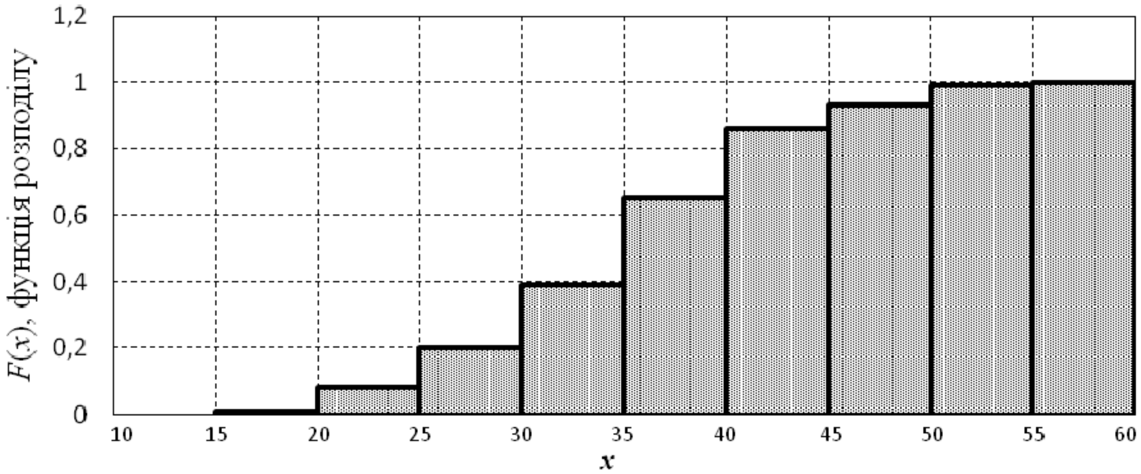

Рис. 5. Розподіл правопорушників за віком

в) Обчислимо числові характеристики вибірки: вибіркову середню, вибіркову дисперсію, вибіркове середнє квадратичне відхилення (табл. 4).

Таблиця 4

Числові характеристики вибірки

Середини інтервалів, *x*	12,5	17,5	22,5	27,5	32,5	37,5	42,5	47,5	52,5	57,5
Відносні частоти, *w*	0,01	0,07	0,12	0,19	0,26	0,21	0,07	0,04	0,02	0,01
Відносна частота на одиницю варіанти, *b* = *w*/5	0,002	0,014	0,024	0,038	0,052	0,042	0,014	0,008	0,004	0,002

Середня вибірки:

$$\overline{x}=\sum_{i=1}^{n} x_i w_i \tag{5}$$

$$\overline{x}=(12,5\cdot 0,01+17,5\cdot 0,07+22,5\cdot 0,12+27,5\cdot 0,19+32,5\cdot 0,26+37,5\cdot 0,21+42,5\cdot 0,07+$$
$$+47,5\cdot 0,04+52,5\cdot 0,02+57,5\cdot 0,01)=32,1$$

Дисперсія вибірки:

$$D\textit{в} = \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2 \cdot w_i \tag{6}$$

Dв = $(12{,}5 - 32{,}1)^2 \cdot 0{,}01 + (17{,}5 - 32{,}1)^2 \cdot 0{,}07 + (22{,}5 - 32{,}1)^2 \cdot 0{,}12 + (27{,}5 - 32{,}1)^2 \cdot 0{,}19 +$
$+ (32{,}5 - 32{,}1)^2 \cdot 0{,}26 + (37{,}5 - 32{,}1)^2 \cdot 0{,}21 + (42{,}5 - 32{,}1)^2 \cdot 0{,}07 + (47{,}5 - 32{,}1)^2 \cdot 0{,}04 +$
$+ (52{,}5 - 32{,}1)^2\ 0{,}02 + (57{,}5 - 32{,}1)^2\ 0{,}01 = 71{,}84$.

Виправлена дисперсія:

$$S^2 = \frac{n}{n-1} \cdot D\textit{в} \tag{7}$$

$S^2 = \frac{100}{100-1} \cdot 71{,}84 = 72{,}5656565656566$, де $n = 100$.

Середньоквадратичне відхилення вибірки:

$$\sigma_{\textit{в}} = \sqrt{D\textit{в}} \tag{8}$$

$\sigma_{\textit{в}} = \sqrt{71{,}84} = 8{,}47584804016684$.

г) Припускаючи, що досліджувана випадкова величина розподілена за нормальним законом, необхідно записати щільність ймовірності випадкової величини x і побудувати її графік на одному рисунку з гістограмою відносних частот (графік вирівнюючої кривої). Для нормального закону розподілу щільність ймовірності:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}, \tag{9}$$

де $a = \bar{x} = 32{,}1$; $\sigma_{\textit{в}} = 8{,}47584804016684$, тобто маємо:

$$f(x) = \frac{1}{8{,}47584804016684 \cdot \sqrt{2\pi}} e^{-\frac{(x-32{,}1)^2}{2 \cdot 8{,}47584804016684^2}}.$$

На графіку (рис. 6) для порівняння зобразимо щільність ймовірності випадкової величини x і гістограму відносних частот на одиницю варіанти (другий рядок табл. 5). Для побудови згладжуючої кривої, а також для проведення подальшої статистичної обробки складемо відповідну таблицю.

Таблиця 5

Числові характеристики вибірки

x_i	m_i	$x_i - \bar{x}$	$z_i = \frac{x_i - \bar{x}}{s}$	$\varphi(z_i)$	$y_i = f(x_i) = \frac{\varphi(z_i)}{s}$	$p_i = h \cdot y_i$	$n \cdot p_i$	$m_i - n \cdot p_i$	$\frac{(m_i - np_i)^2}{np_i}$
12,5	1	-19,6	-2,312453	0,02753	0,00325	0,0162374	1,6237357	-0,6237357	0,239599
17,5	7	-14,6	-1,72254	0,09049	0,01068	0,0533812	5,3381238	1,6618762	0,517379
22,5	12	-9,6	-1,13263	0,21006	0,02478	0,1239166	12,3916574	-0,3916575	0,012379
27,5	19	-4,6	-0,54272	0,34431	0,04062	0,2031130	20,3112951	-1,3112951	0,084657
32,5	26	0,4	0,04719	0,39850	0,04702	0,2350787	23,5078702	2,4921298	0,264197
37,5	21	5,4	0,63710	0,32566	0,03842	0,1921128	19,2112847	1,7887153	0,166543
42,5	7	10,4	1,22702	0,18792	0,02217	0,1108580	11,0857987	-4,0857987	1,505868
47,5	4	15,4	1,81693	0,07657	0,00903	0,0451695	4,5169473	-0,5169473	0,059163
52,5	2	20,4	2,40684	0,02203	0,00260	0,0129954	1,2995426	0,7004574	0,377549
57,5	1	25,4	2,99675	0,00448	0,00053	0,0026400	0,2639997	0,7360003	2,787883
Σ									6,015217

В табл. 5 $\varphi(z_i)$ обчислюється за формулою щільності стандартного нормального розподілу [5, с. 28]: $\varphi(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}$.

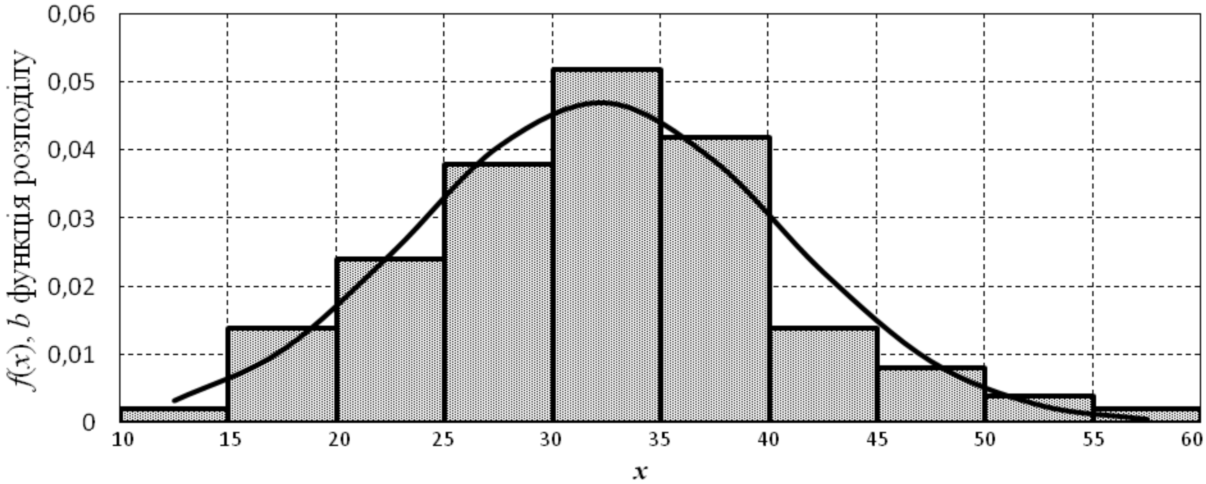

Рис. 6. Щільність ймовірності випадкової величини x та гістограма відносних частот на одиницю варіанти

д) Знаходимо теоретичні частоти нормального закону розподілу. Для цього визначаємо величину (табл. 5):

$$p_i = h \cdot y_i, \qquad (10)$$

де h – ширина інтервалу (5 років).

Теоретичні частоти розподілу злочинів щодо віку осіб за нормальним законом будуть становити (табл. 5):

$$T = n \cdot p_i, \qquad (11)$$

Перевірка правдоподібності гіпотези про приналежність статистичних даних до заданого виду ймовірнісного закону може проводитися за допомогою критеріїв: χ^2 Пірсона, Колмогорова, Романовського, Андерсона-Дарлінга, Крамера-Мізеса-Смірнова, Нікуліна, Шовене та ін.

1. Критерій χ^2 Пірсона. Статистика критерію χ^2 Пірсона має вигляд [7, с. 264]:

$$\chi_n^2 = n \sum_{i=1}^{k} \frac{(n_i / n - P_i(\theta))^2}{P_i(\theta)}, \qquad (12)$$

де $P_i(\theta)$ – ймовірність попадання випадкової величини в задані інтервали. На основі залежності (12), критерій χ^2 Пірсона записується у вигляді наступної альтернативної умови:

$$P_{розр.}(\chi^2, r) = \begin{cases} \geq \alpha - \text{гіпотеза про приналежність статистичних} \\ \text{даних до розглянутого ймовірнісного закону не відхиляється;} \\ < \alpha - \text{гіпотеза про приналежність статистичних} \\ \text{даних до розглянутого ймовірнісного закону відхиляється;} \\ \alpha = 0,05 \end{cases} \qquad (13)$$

де r – число ступенів свободи, $r = k - s$; k – число інтервалів гістограми; s – число накладених зв'язків.

Оцінимо узгодженість обраного теоретичного розподілу із статистичними даними відповідно до критерію Пірсона (χ^2-квадрат). Критерій узгодження Пірсона – найбільш часто використовуваний критерій для перевірки простої гіпотези про закон розподілу (гіпотезу, що однозначно фіксує розподіл спостережень, називають простий (у ній мова йде про одне значення параметра), а якщо ні, то – складної). Підсумовуючись величини стовпчика 10 в табл. 5, одержуємо величину: $\chi^2_{розр.} = \sum_{i=1}^{10} \frac{(m_i - np_i)^2}{np_i} = 6{,}015217$, що характеризує міру розходження теоретичного і статистичного розподілів [5, с. 28; 6, с. 129; 8, с. 273]. Чим більша величина χ^2-квадрат, тим більше це розходження. Для ступеня вільності $K = k - 3 = 10 - 3 = 7$, де $k = 10$ – число інтервалів, знаходимо [4, с. 352] $\chi^2_{табл.} = 14{,}1$. Оскільки, $\chi^2_{розр.} < \chi^2_{табл.}$, тоді гіпотеза про належність статистичних даних до нормального закону розподілу приймається.

Задана величина критичного значення рівня значущості $\alpha_{кр.} = P\left(\chi^2 \geq \chi^2_{розр.}\right) = 0{,}05$, який повинен бути перевершений для прийняття гіпотези про закон розподілу. Необхідно вказати, що що прийняття гіпотези завжди відбувається на деякому суб'єктивно прийнятому рівні значущості і ґрунтується на значеннях кінцевої вибірки (в даному випадку $n = 100$). Як правило орієнтуються на критичне значення рівня значущості $\alpha = 0{,}01 \ldots 0{,}05$ ($\alpha = 0{,}1$ відповідає м'якшому підходу до оцінки гіпотези, $\alpha = 0{,}05$ відповідає більш строгому підходу до оцінки гіпотези).

За таблицею [9, с. 352] критичних точок χ^2 – розподілу, виходячи з числа ступенів свободи $r = k - 3 = 7$ ($k = 3$ – число розрядів) і реалізованого значення $\chi^2_{розр.} = 6{,}015217$, знаходимо величину: $\alpha = P\left(\chi^2 \geq \chi^2_{розр.}\right) = P\left(\chi^2 \geq 6{,}015217\right) \approx 0{,}54$, тобто ймовірність того, що величина розподілена по закону χ^2 перевершить $\chi^2_{розр.}$. Таблиці дають не дуже точні значення α. Перевіримо на точність знайдене в таблиці значення $\alpha = 0{,}54$ в Mathcad 15:

$$\varphi(x,r) := \frac{1 \cdot x^{\frac{r-2}{2}} \cdot e^{\frac{-x}{2}}}{2^{\frac{r}{2}} \cdot \Gamma\left(\frac{r}{2}\right)}; \qquad \alpha := \int_{6.015217}^{\infty} \varphi(x,7)\,dx = 0.538.$$

Оскільки ймовірність, $\alpha > \alpha_{кр.}$ (0,538 > 0,05), тоді можна вважати, що емпірично прийнятий теоретичний нормальний розподіл не протирічить статистичним даним і гіпотеза про вид розподілу і його параметри може бути прийнята. Іншими словами, прийнята гіпотеза не протирічить наявним вибірковим даним на рівні значущості $\alpha = 0{,}05$ або рівні надійності (рівні довіри) 1 – 0,05 = 0,95.

2. Критерій Романовського [10, с. 96].

$$K_P = \frac{\chi^2 - K}{\sqrt{2K}} = \begin{cases} \leq 3 \text{ – гіпотеза про приналежність статистичних} \\ \text{даних до розглянутого ймовірнісного закону} \\ \text{не відхиляється;} \\ 3 > \text{ – гіпотеза про приналежність статистичних} \\ \text{даних до розглянутого ймовірнісного закону відхиляється} \end{cases} \qquad (14)$$

де χ^2 – критерій Пірсона; K – число інтервалів.

За критерієм Романовського:

$$K_P = \left|\frac{\chi^2 - K}{\sqrt{2K}}\right| = \left|\frac{6{,}015217 - 7}{\sqrt{2 \cdot 7}}\right| = 0{,}263194327594243 < 3$$

, отже гіпотеза про

приналежність статистичних даних щодо розподілу злочинів за віком правопорушників до нормального закону розподілу за критерієм Романовського приймається.

3. Критерій Колмогорова. В критерії Колмогорова відстань між емпіричним $F_n(x)$ і теоретичним $F_n(x, \theta)$ розподілом має вигляд [7, с. 183; 11; 12]:

$$D_n = \sup_{|x|<\infty} | F_n(x) - F_n(x,\theta) | , \qquad (15^1)$$

де n – обсяг вибірки. З рівності (15^1) критерій Колмогорова без поправки Большева записується у вигляді альтернативної умови [5, с. 28]:

$$P(\lambda) = P\left\{ \max | F^*_{розр.}(x) - F_{теор.}(x) | / \sqrt{n} \right\} = \begin{cases} \geq P_Д \text{ – гіпотеза про приналежність статистичних даних до розглянутого ймовірнісного закону не відхиляється;} \\ < P_Д \text{ – гіпотеза про приналежність статистичних даних до розглянутого ймовірнісного закону відхиляється} \end{cases} \qquad (15^2)$$

де n – обсяг вибірки; $F^*_{розр.}(x)$ – розрахункове значення функції розподілу; $F_{теор.}(x)$ – теоретичне значення функції розподілу; $P_Д$ – довірча ймовірність.

За критерієм Колмогорова:

$$P(\lambda) = P\left\{ \max | F^*_{розр.}(x) - F_{теор.}(x) | / \sqrt{n} \right\} = \left| \frac{-4,0857987}{\sqrt{100}} \right| = 0,40857987 ,$$

де n = 100 – обсяг вибірки. Для значення λ = 0,40857987 ймовірність α = 0,997032 ($\lambda<\alpha$) критерій статистично не значимий. Це означає, що з ймовірністю 0,0,997032 можна стверджувати, що відхилення фактичних (емпіричних) частот від теоретичних є випадковими. Отже, є підстави затверджувати, що емпіричний розподіл підкоряється нормальному розподілу.

е) Знаходимо інтервальні оцінки параметра а нормального закону розподілу. Довірчу ймовірність приймаємо 0,95. Ймовірність того, що абсолютна величина відхилення менша додатного числа δ:

$$P\left(| x - a | < \delta \right) = 2 \cdot \Phi\left(\frac{\delta}{\sigma} \right) , \qquad (16)$$

де x – неперервна випадкова величина; a – математичне очікування; δ – границя відхилення абсолютної випадкової величини; $\Phi(t)$ – функція Лапласа (стандартизовані таблиці); σ – середньоквадратичне відхилення.

В нас $\sigma(x)$ = 10, a = $M(x)$ = 32,1. Величина ймовірності задана: $2 \cdot \Phi\left(\frac{\delta}{10} \right) = 0,95$,

$2\cdot\Phi\left(\frac{\delta}{10}\right)=0,475$. З стандартизованої таблиці знаходимо: $\frac{\delta}{10}=1,96$, δ = 19,6. Отже, з ймовірністю 0,95 параметр а нормального розподілу (математичне очікування генеральної сукупності) знаходиться в інтервалі:

($a - \delta$; $a + \delta$) = (32,1 – 19,6; 32,1 + 19,6) = (12,5; 51,7).

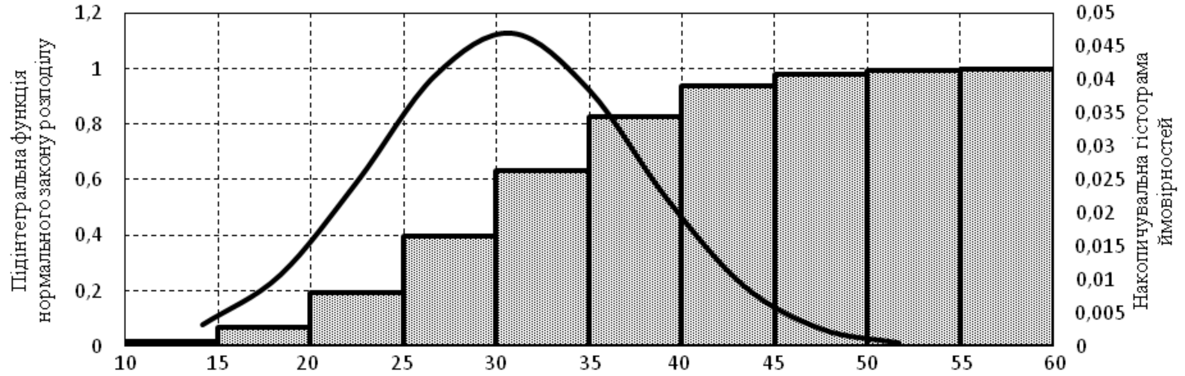

Рис. 7. Підінтегральна функція нормального закону розподілу та накопичувальна гістограма ймовірностей

Таблиця 6

Значення щільності $f(z)$ та функції $F(z)$ нормального закону розподілу правопорушників за віком

x_i	m_i	$x_i-\bar{x}$	$z_i=\frac{x_i-\bar{x}}{s}$	$\varphi(z_i)$	$F(z_i)$	$y_i=f(x_i)==\frac{\varphi(z_i)}{s}$	$p_i = h\cdot y_i$	$n\cdot p_i$
12,5	1	-19,6	-2,312453	0,02753	0,010376366	0,00325	0,0162374	1,6237357
17,5	7	-14,6	-1,72254	0,09049	0,042485572	0,01068	0,0533812	5,3381238
22,5	12	-9,6	-1,13263	0,21006	0,128684832	0,02478	0,1239166	12,3916574
27,5	19	-4,6	-0,54272	0,34431	0,293661301	0,04062	0,2031130	20,3112951
32,5	26	0,4	0,04719	0,39850	0,518819101	0,04702	0,2350787	23,5078702
37,5	21	5,4	0,63710	0,32566	0,737970146	0,03842	0,1921128	19,2112847
42,5	7	10,4	1,22702	0,18792	0,890092463	0,02217	0,1108580	11,0857987
47,5	4	15,4	1,81693	0,07657	0,965386084	0,00903	0,0451695	4,5169473
52,5	2	20,4	2,40684	0,02203	0,991954391	0,00260	0,0129954	1,2995426
57,5	1	25,4	2,99675	0,00448	0,998635628	0,00053	0,0026400	0,2639997

Отже, згідно статистичних відомостей за 2017 р. про розподіл злочинів за віком правопорушників проведений повний розрахунок варіаційного ряду, а саме: моди, коефіцієнта асиметрії, та знайдений нормальний закон розподілу віку правопорушників в Україні $f(x)=\frac{1}{8,47584804016684\cdot\sqrt{2\pi}}e^{-\frac{(x-32,1)^2}{2\cdot 8,47584804016684^2}}$ (рис. 7; табл. 6).

Висновки. Таким чином, на основі аналізу варіаційного ряду розподілу злочинності за віком правопорушників, з'ясовано, що:

– найбільша частота скоюваних злочинів припадає на інтервал 30-35 років (26%);

– середня арифметична інтервального ряду становить $\bar{x}=32,1$ року;

– модальна величина дорівнює $Mo \approx 32,91667$ роки;

– коефіцієнт асиметрії від'ємний (KA = -0,05574) – графік розподілу злочинів за віком правопорушників має незначну лівосторонню асиметрію;

– вирівнювання можна виконати за нормальним законом розподілу (розподіл Ґауса), щільність ймовірності якого: $f(x)=\frac{1}{8{,}47584804016684\cdot\sqrt{2\pi}}e^{-\frac{(x-32{,}1)^2}{2\cdot 8{,}47584804016684^2}}$;

– за критеріями χ^2 Пірсона, Романовського, Колмогорова гіпотеза про відповідність розподілу злочинів за віком правопорушників нормальному закону розподілу (розподіл Ґауса) приймається.

Перспективи подальших наукових досліджень. Отриманий нормальний закон розподілу злочинів за віком правопорушників з щільністю ймовірності: $f(x)=\frac{1}{8{,}47584804016684\cdot\sqrt{2\pi}}e^{-\frac{(x-32{,}1)^2}{2\cdot 8{,}47584804016684^2}}$ та перевірка його на відповідність за критеріями χ^2 Пірсона, Романовського, Колмогорова не дозволяє однозначно стверджувати, що це єдино можливий ймовірнісний закон. А тому, доцільно провести додаткові дослідження щодо можливості вирівнювання за іншими ймовірнісними законами розподілу, наприклад, на вирівнювання варіаційного ряду за логістичним розподілом, провести їх порівняння за коефіцієнтами кореляції та обрати найбільш "точний" з них.

Перспективним напрямками для розрахунку ймовірнісних законів розподілу є реалізація в програмному середовищі MATLAB.

Саме ці обставини вимагають в майбутньому проводити додаткові дослідження варіаційних рядів злочинності та розрахунки параметрів всіх можливих ймовірнісних законів з їх перевіркою за різними критеріями відповідності.

Використана література:

1. *Гаврилов О. А.* Курс правовой информатики : учеб. для вузов / О. А. Гаврилов. – Москва : Норма, Инфра-М, 2000. – 432 с.
2. *Лунеев В. В.* Юридическая статистика : учеб. / В. В. Лунеев. – 2-е изд., перераб. и доп. – Москва : Юристъ, 2007. – 394 с.
3. *Ларченко М. О.* Дослідження ймовірності рецидиву за допомогою математичних методів / М. О. Ларченко // Науковий вісник Ужгородського університету: Серія Право. – Випуск 23, Частина I, Том 3. – С. 46-51.
4. Генеральна прокуратура України: Довідка про стан та структуру кримінальних правопорушень на території України [Електронний ресурс]. – Режим доступу : http://www.gp.gov.ua/ua/stst2011.html?dir_id=112173&libid=100820
5. *Рублева Г. В.* Математическая статистика: статистические критерии проверки гипотез : учеб.-метод. пособ. / Г. В. Рублева. – Тюмень : Изд-во Тюменского госуд. ун-та. – 2014. – 50 с.
6. *Воскобойников Ю. Е.* Математическая статистика (с примерами в Excel) : учеб, пособ / Ю. Е. Воскобойников, Е. И. Тимошенко. – 2-е изд. перераб. и доп. – Новосибирск : НГА-СУ (Сибстрин), 2006. – 152 с.
7. *Лемешко Б. Ю.* Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход : монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. – Новосибирск : Изд-во НГТУ, 2011. – 888 с. (серия "Монографии НГТУ").
8. *Бартел Лендерт ван дер Варден.* Математическая статистика / Бартел Лендерт ван дер Варден ; пер. с нем. Л. Н. Большева. – Москва : Изд-во иностр. лит., 1960. – 436 с.
9. *Єлейко В. І.* Економетричний аналіз діяльності підприємств : навч. посіб. / В. І. Єлейко, Р. Д. Боднар, М. Я. Демчишин. – Тернопіль : Навчальна книга-Богдан, 2011. – 362 с.
10. *Романовский В. И.* Элементарный курс математической статистики / В. И. Романовский. – 2-е изд., перер. и доп. – Москва ; Ленинград : Госпланиздат, 1939. – 360 с.

11. *Kolmogoroff A. N.* Sulla determinazione empirica di una legge di distribuzione / A. N. Kolmogoroff // Giornale dell` Istituto Italiano degly Attuari. – 1933. – Vol. 4. – № 1. – P. 83-91.
12. *Lilliefors H. W.* On the Kolmogorov-Smirnov test for normality with mean and variance unknown / H. W. Lilliefors // Journal of the American Statistical Association. – 1967. – Vol. 62. – P. 399-402.

References:

1. *Gavrilov O. A.* Kurs pravovoy informatiki : ucheb. dlya vuzov / O. A. Gavrilov. – Moskva : Norma, Infra-M, 2000. – 432 c.
2. *Luneev V. V.* Yuridicheskaya statistika : ucheb. / V. V. Luneev. – 2-e izd., pererab. i dop. – Moskva : Yurist, 2007. – 394 s.
3. *Larchenko M. O.* Doslidzhennia ymovirnosti retsydyvu za dopomohoiu matematychnykh metodiv / M. O. Larchenko // Naukovyi visnyk Uzhhorodskoho universytetu: Seriia Pravo. – Vypusk 23, Chastyna I, Tom 3. – S. 46-51.
4. Heneralna prokuratura Ukrainy: Dovidka pro stan ta strukturu kryminalnykh pravoporushen na terytorii Ukrainy [Elektronnyi resurs]. – Rezhym dostupu : http://www.gp.gov.ua/ua/stst2011.html?dir_id=112173&libid=100820
5. *Rubleva G. V.* Matematicheskaya statistika: statisticheskie kriterii proverki gipotez : ucheb.-metod. posob. / G. V. Rubleva. – Tyumen : Izd-vo Tyumenskogo gosud. un-ta. – 2014. – 50 s.
6. *Voskoboynikov Yu. Ye.* Matematicheskaya statistika (s primerami v Excel) : ucheb, posob / Yu. Ye. Voskoboynikov, Ye. I. Timoshenko. – 2-e izd. pererab. i dop. – Novosibirsk : NGA-SU (Sibstrin), 2006. – 152 s.
7. *Lemeshko B. Yu.* Statisticheskiy analiz dannykh, modelirovanie i issledovanie veroyatnostnykh zakonomernostey. Kompyuternyy podkhod : monografiya / B. Yu. Lemeshko, S. B. Lemeshko, S. N. Postovalov, Ye. V. Chimitova. – Novosibirsk : Izd-vo NGTU, 2011. – 888 s. (seriya “Monografii NGTU”).
8. Bartel Lendert van der Varden. Matematicheskaya statistika / Bartel Lendert van der Varden ; per. s nem. L. N. Bolsheva. – Moskva : Izd-vo inostr. lit., 1960. – 436 s.
9. *Ieleiko V. I.* Ekonometrychnyi analiz diialnosti pidpryiemstv : navch. posib. / V. I. Yeleiko, R. D. Bodnar, M. Ya. Demchyshyn. – Ternopil : Navchalna knyha-Bohdan, 2011. – 362 c.
10. *Romanovskiy V. I.* Elementarnyy kurs matematicheskoy statistiki / V. I. Romanovskiy. – 2-e izd., perer. i dop. – Moskva ; Leningrad : Gosplanizdat, 1939. – 360 s.
11. *Kolmogoroff A. N.* Sulla determinazione empirica di una legge di distribuzione / A. N. Kolmogoroff // Giornale dell` Istituto Italiano degly Attuari. – 1933. – Vol. 4. – № 1. – P. 83-91.
12. *Lilliefors H. W.* On the Kolmogorov-Smirnov test for normality with mean and variance unknown / H. W. Lilliefors // Journal of the American Statistical Association. – 1967. – Vol. 62. – P. 399-402.

Хаврук В. А. Анализ вариационного ряда и вероятностный закон распределения преступлений по возрасту правонарушителей.

В статье в качестве прикладной направленности теории вероятностей и математической статистики избрано такое социальное явление как преступность, в частности рассматривается возрастная структура правонарушителей в Украине. Проведенный анализ вариационного ряда распределения преступлений по возрасту правонарушителей на основе усредненных показателей за 2017 г. в Украине. Рассчитаны средняя арифметическая, мода и коэффициент асимметрии ряда. Сравнение вариационного ряда распределения правонарушителей, совершивших преступления, с нормальным законом распределения позволяет утверждать, что преступность в Украине “молодеет”, а именно набольшее количество преступлений совершаются в возрасте 30-35 лет, при этом имеется отрицательный коэффициент асимметрии, который указывает на общую несимметричность реального распределения. Полную картину определения характера распределения преступности по возрасту правонарушителей дает лишь выравнивание вариационного ряда по известным теоретическим законам распределения случайных величин. В статье осуществлено выравнивание ряда преступности за нормальным законом. Для этого были проведены следующие действия: рассчитаны относительные частоты; построены полигон и гистограмма относительных частот; проведены все необходимые расчёты – числительные характеристики *случайной величины и плотности стандартного нормального распределения сведены в таблицу; получена функция плотности распределения вероятности; построены графики функции распределения, плотности и гистограмма вероятностей случайной величины, а также накопительная гистограмма вероятностей.*

Проверено соответствие реального вариационного ряда нормальному закону за критериями χ2 Пирсона, Романовского, Колмогорова.

Предложенная методология исследования на соответствие статистического распределения случайных величин теоретическому закону (распределение Гаусса) требует таких основных знаний, как: интервальная оценка параметров случайной величины, характеристики и свойства нормального закона распределения, этапы проверки гипотезы о соответствии статистики принятому вероятностному закону, основные критерии соответствия, программные среды Excel и MathCAD.

***Ключевые слова**: вариационный ряд, случайная величина, гипотеза, эксцесс, интервал, вероятность, коэффициент асимметрии, мода, нормальный закон распределения (распределение Гаусса), среднеквадратичное отклонение, среднеарифметическое отклонение.*

Khavruk V. O. Analysis of the variation series and the probable law of distribution of crimes on age of offenses.

In the article, as an applied orientation of probability theory and mathematical statistics, such a social phenomenon as crime is chosen, in particular the age structure of offenders in Ukraine is considered. The analysis of the variational series of distribution of crimes by age of offenders on the basis of averaged indicators for 2017 in Ukraine is carried out. The arithmetic mean is calculated, the mod and the asymmetry ratio of the row. A regularization of a number of criminality has been performed under normal law - the probability distribution density function has been obtained and appropriate histograms and graphical dependencies have been constructed. The correspondence of the real variation series to the normal law according to the criteria χ2 of Pearson's, Romanovskiy, and Kolmogoroff is checked. The proposed research methodology for the correspondence of the statistical distribution of random variables to the theoretical law (Gaussian distribution) requires such basic knowledge as: interval estimation of the random variable parameters, characteristics and properties of the normal distribution law, stages of testing the hypothesis of the compliance of statistics with the accepted probabilistic law, the main compliance criteria, Excel and MathCAD environments.

***Keywords:** variational series, random variable, hypothesis, kurtosis, interval, probability, asymmetry coefficient, mode, normal probability law (Gaussian distribution), mean deviation, arithmetic mean deflection.*

УДК 378.147.091.3:[51:005.336.2-057.875]:37.091.12.011.3-051:51/53(043.5)

Чкана Я. О.

ВИЗНАЧЕННЯ РІВНІВ СФОРМОВАНОСТІ МАТЕМАТИЧНОЇ КОМПЕТЕНТНОСТІ МАЙБУТНІХ ВЧИТЕЛІВ МАТЕМАТИКИ

Стаття присвячена визначенню основних критеріїв та рівнів сформованості математичної компетентності майбутніх вчителів математики у процесі вивчення фахових дисциплін. Критерії сформованості математичної компетентності розглядаємо як сукупність сутнісних ознак, які дають можливість зробити висновки про стан та рівень її сформованості, при цьому ступінь сформованості визначається в конкретних показниках, які характеризуються в свою чергу низкою ознак. Обґрунтовано, що критеріями сформованості математичної компетентності визначено когнітивний, процедурно-технологічний, інтелектуальний. У статті також обґрунтовано їх