

УДК 004.04

Світлана Миколаївна ВИГАНЯЙЛО,

кандидат економічних наук,

доцент кафедри соціально-економічних дисциплін

Сумської філії Харківського національного університету внутрішніх справ;

ORCID: 0000-0001-5350-0728

Лариса Петрівна ПЕРХУН,

кандидат педагогічних наук, доцент

доцент кафедри економіко-математичних дисциплін та інформаційних технологій

Національної академії статистики, обліку та аудиту;

ORCID: 0000-0002-8667-2312

Аналіз програмного забезпечення для аналізу у правовій статистиці

Юристи у своїй практичній діяльності стикаються не лише з конкретними правовими фактами і подіями, а й з масовими правовими явищами і процесами, статистичний аналіз яких є необхідною умовою їхньої професійної діяльності. Студенти-правознавці, опанувавши статистичну методологію як сукупність прийомів і методів статистичного дослідження, зможуть вільно орієнтуватись у складних соціально-правових явищах і процесах, краще засвоювати інші юридичні дисципліни.

Студент-юрист має знати сучасну організацію статистичної роботи у правоохоронних органах; методи збирання статистичних даних про правопорушення і діяльність правоохоронних органів; принципи і методи обробки матеріалів статистичного спостереження; сутність узагальнювальних показників та основи їх статистичного аналізу.

Студент-юрист має вміти організовувати і проводити статистичне спостереження правових явищ; розробляти макети та будувати таблиці і графіки; обчислювати узагальнювальні статистичні показники; аналізувати статистичні матеріали про правопорушення і заходи боротьби з ними з боку правоохоронних органів; формулювати висновки і пропозиції для прийняття управлінських рішень.

Дана публікація має на меті проаналізувати існуючі сучасні програмні додатки, що використовуються для обробки та аналізу результатів, та дозволяють автоматизувати рутинну роботу.

Обробка великих масивів статистичної інформації, необхідної для аналізу правової, правоохоронної діяльності, може бути виконана лише з використанням сучасних засобів інформаційних технологій. Використання достовірної і науково-обґрунтованої інформації приводить до зменшення витрат, підвищення якості та ефективності обробки даних. Із зростанням потреби статистичного аналізу даних ринок ПЗ для статистичної обробки даних нестримно розвивається. Розглянемо найбільш поширені статистичні програми для обробки великих обсягів даних.

Метою нашого дослідження є аналіз та класифікація (щодо функціональних можливостей) статистичного програмного інструментарію, для аналізу великих сукупностей даних у правовій та правоохоронній діяльності для отримання бази для прийняття рішень, а також деякі рекомендації щодо їх вибору.

Методи статистичної обробки даних настільки різноманітні, що досліднику дуже складно розібратися у них і вибрати адекватні поставленим завданням статистичні критерії. Як правило, оптимальним є варіант, що комбінує відповідність комп'ютерного устаткування користувача системним вимогам ПЗ; відповідність можливостей ПЗ до параметрів поставлених завдань; об'єм даних для статистичного аналізу; кваліфікація користувача в області статистики. Загалом, основні вимоги до статистичного пакету - модульність; варіативність вибору способів обробки даних; простота формулювання завдання користувача; можливість автоматизованої обробки даних та зв'язків з модулями пакета; складання зрозумілого та наглядного звіту за результатами виконаного аналізу; наявність діалогового режиму роботи користувача з пакетом; можливість імпорту-експорту з іншими програмами.

Дослідження характеристик програмних засобів для математичної і статистичної обробки показав, що за функціональністю програми для статистичного аналізу можна розділити на 3 основних групи: універсальні пакети, або пакети загального призначення (STADIA, STATGRAPHICS, SPSS, STATISTICA); професійні пакети (SAS, BMDP); спеціалізовані, професійні пакети.

На ринку ПЗ існує велика кількість додатків, які пропонують допомогу у розв'язанні задач аналізу даних як у пакетному режимі, так і у вигляді бібліотек функцій, які можна використовувати в інших програмних продуктах.

SPSS для Windows (<http://www.spss.com/>) — це модульний, повністю інтегрований, у якого усіма необхідними можливостями програмний продукт, готовий до всіх етапів аналітичного процесу:

планування, збирання цих, доступу до даних та управління даними, аналізу, створення звітів створення і поширення результатів. Програмне забезпечення SPSS дозволяє проводити частотний аналіз, описову статистику, кореляційний аналіз, дисперсійний аналіз, кластерний аналіз, факторний аналіз, і навіть регресійний аналіз. З допомогою аналітичних можливостей програми SPSS можна отримати такі дані: найвигідніші сегменти ринку; стратегії позиціонування послуг щодо аналогічних послуг конкурентів; оцінка якості послуги клієнтами; перспективи розвитку;

Statistica (<https://statsoftstatistica.ru/>) — це універсальна інтегрована система, призначена для статистичного аналізу та візуалізації даних, управління базами даних, і розробки користувальницьких додатків, яка містить широкий набір процедур аналізу до застосування у наукових дослідженнях, техніці, бізнесі. Це сучасний пакет статистичного аналізу, у якому буде реалізовано всі новітні комп'ютерні і математичні методи аналізу даних. Можливість доступу до нових, нетрадиційних методів аналізу допомагає знаходити нові шляхи перевірки робочих гіпотез і дослідження даних.

Програмне забезпечення Statistica дає змогу провадити такі процедури обробки статистичних даних: описова статистика; аналіз багатомірних таблиць; багатомірна регресія; дискримінантний аналіз; аналіз відповідностей; кластерний аналіз; факторний аналіз; дисперсійний аналіз стану і багато іншого.

StatGraphics (www.statgraphics.com). Повний універсальний статистичний пакет для діалогового аналізу (interactive analysis) статистичних даних з широкими можливостями візуалізації. Функції ПЗ: проста регресія, аналіз часових рядів, багатомірна лінійна та нелінійна регресія, авторегресійні моделі. Переваги STATGRAPHICS: поєднання наукових методів обробки різномісних даних з можливістю створення сучасної високоякісної інтерактивної графіки; широкі можливості взаємодії з іншими програмними продуктами (електронними таблицями, базами даних); високоякісна двовимірна і тривимірна графіка, інтегрована графіка, в якій усі елементи графічних представлень результатів аналізу можуть бути перетворені. Після завершення процедури статистичного аналізу даних можна обрати графічні відображення результатів, релевантні використаній процедурі аналізу.

SYSTAT (<https://systatsoftware.com/products/systat/>) - програма для статистичної обробки. Функції: ділова графіка; параметричні тести (критерії Ст'юдента, Фішера, гістограма); непараметричні тести; категоріальний аналіз; дисперсійний аналіз (одно-, дво-, багатомісний); регресійний аналіз; аналіз часових рядів; багатомісні методи, кластерний аналіз, факторний аналіз; побудова теоретичних моделей для емпіричних отриманих даних, відображення результатів у вигляді тривимісних поверхонь чи двовимісних кривих; глибокий аналіз і обробка цифрових зображень.

Minitab (www.malavida.com). Функції пакету: управління даними і файлами; електронні таблиці для поліпшення аналізу даних; регресійний аналіз - проста регресія, регресія часових рядів; потужність і розмір вибірки; таблиці і діаграми; багатомісний аналіз, а саме - факторний аналіз, кластерний аналіз, аналіз відповідностей тощо; непараметричний аналіз - різноманітні тести, включаючи вокальний тест, критерій серій, тощо; часовий ряд і прогнозування - інструмент, який допомагає виявити тенденції в даних, а також прогнозування майбутніх значень, діаграми часових рядів, експоненціальне згладжування, аналіз тенденцій; статистичне управління процесами; системний аналіз вимірів; дисперсійний аналіз - визначення відмінностей між точками даних.

NCSS (<https://www.ncss.com/>) - статистичний аналіз та побудова графіків за будь-якими даними. Розрахований на непрофесіоналів в області статистичної обробки. Інтерфейс системи багатомісний. Усі дії користувача супроводжуються підказками. Дозволяє здійснювати статистичний аналіз введених даних, давати графічну інтерпретацію отриманих результатів в зображеннях різних вимірів, імпортувати та експортувати дані в найбільш поширених форматах.

OriginPro (<https://www.originlab.com/>) - потужний пакет для аналізу результатів статистичних і наукових досліджень та вимірювань, надає засоби програмування та побудови графіків, діаграм, таблиць. Пакет програм для чисельного аналізу даних і наукової графіки, Origin створена для побудови двовимісної і тривимісної наукової графіки, яка створюється за допомогою готових шаблонів, доступних для редагування користувачем. Також можливо створювати нові власні шаблони. Після створення зображення воно може бути відредаговане за допомогою меню і діалогів. Можна експортувати отримані графіки і таблиці в ряд форматів, таких як PDF, TIFF, JPEG, GIF. В Origin можна проводити чисельний аналіз даних, включаючи різні статистичні операції, обробку сигналів, тощо.

Stata (<http://www.stata.com/>) - потужний пакет для обробки статистичних і графічних даних, позиціонується інструментом аналізу для фахівців, які займаються науковими дослідженнями. Завдяки гнучкій модульній структурі пакет можливо застосовувати для аналізу даних із різних областей знань: суспільні науки, медицина, тощо. Stata надає користувачам наступні можливості статистичного аналізу даних: динамічні факторні моделі, лінійні, узагальнені лінійні моделі і нелінійні моделі, багаторівневі

змішані моделі, моделі вибуття; узагальнені оцінки рівнянь; непараметричні методи, методи повторних вибірок і статистичного експерименту; перевірка гіпотез і робота з оціненими моделями; максимізація функцій правдоподібності, заданих користувачем; матричні команди.

Поряд з комерційними статистичними пакетами існує велика кількість безкоштовних статистичних програм і додатків. При цьому ряд програм не тільки не поступається, а і перевершує по функціональності комерційні додатки. Розглянемо основні безкоштовні програми статистичної обробки даних.

Макрос-додаток **XLSTAT-Pro** (<http://www.xlstat.com>) для MS Excel, який включає в себе більше 50 статистичних функцій, включаючи аналіз виживаності.

OpenEpi ([https://openepi.software.informer.com/2.3/#googtrans\(en|ru\)](https://openepi.software.informer.com/2.3/#googtrans(en|ru))) - набір статистичних функцій, що дозволяє швидко застосувати відносно прості і часто використовувані статистичні тести, можна використовувати в он-лайн режимі на сайті розробника, або інсталиувати на комп'ютер. Плюсом пакета є набір функцій для розрахунку статистичної потужності, чисельності груп, генерації випадкових чисел, а також можливість розрахунку статистичної значущості на підставі групових статистик.

PSPP (<http://lostapp.ru/soft/pspp>) - за зовнішнім виглядом і функціональністю дуже нагадує SPSS (власне, і назва пакета є дзеркальним відображенням), при цьому повністю безкоштовний.

SOFA (<http://lostapp.ru/soft/sofa-statistics>) - дозволяє виконувати основні статистичні тести, але не дає можливості проводити регресійний аналіз. Однією з відмінних особливостей пакету є швидке створення різних типових графіків і не потребують форматування підсумкових таблиць, а також можливість виконувати користувальницькі скрипти на Python.

AtteStat (<http://offext.ru/library/science/development/87.aspx>) - програма призначена для професійної статистичної обробки даних в різних областях діяльності. Вона виконана у вигляді надбудови до популярних електронних таблиць Microsoft Excel. Конструктивно програма складається з функціонально незалежних модулів, об'єднаних загальним інтегратором, коректно інтегрується в інтерфейс електронних таблиць, не уповільнює швидкодію комп'ютерної системи в процесі роботи, не вимагає ніяких налаштувань і не конфліктує з іншими програмами. Повнота методів забезпечує якісний аналіз даних різної природи. Найбільш затребувані модулі: Аналіз часових рядів і прогнозування. Параметрична статистика. Непараметрична статистика. Перевірка нормальності. Обробка пропущених даних. Багатовимірне шкалювання. Експертні оцінки. Інформаційний аналіз. Факторний аналіз. Точні критерії. Описова статистика. Кореляційний аналіз. Кластерний аналіз. Дисперсійний аналіз.

Висновки Розглянуто та запропоновано до використання ряд програмних продуктів та дана коротка змістовна характеристика. Інформаційні технології та їх складові частини розвиваються невпинно та стрімко, відповідно ринок статистичних програмних продуктів проаналізований на поточний момент. У разі потреби потужного, загальноновизнаного пакету з простим і зрозумілим інтерфейсом, краще скористатися SPSS. Для початківців і професіоналів, яким потрібна підказка і розвинена документація російською мовою, можна рекомендувати STATISTICA. Для невибагливих користувачів, які обмежуються у своїх дослідженнях стандартними статистичними методами можна рекомендувати Excel. Для ознайомлення із статистичною обробкою даних у навчальному процесі можна рекомендувати ознайомлення із більшістю розглянутих програмних продуктів. Досліджено найбільш популярні та функціональні програмні продукти з засобами статистичного аналізу даних та розглянуто інструменти, які дозволять в умовах великих обсягів інформації здійснити обробку та аналіз даних з використанням сучасних інформаційних технологій. Вибір оптимального програмного забезпечення залежить від особливостей і параметрів конкретного дослідження правової галузі.

Список бібліографічних посилань

1. Виганяйло С.М. Правова статистика. Навчальний посібник для студентів юридичних вузів для юридичних спеціальностей студентів 1-2 курсу спеціальності 081 «Право» та 262 «Правоохоронна діяльність» денної та заочної форми навчання. Суми, Видавничий дім «Ельдорадо»: 2019. – 145 с.

2. Гойко О. В. Практичне використання пакета STATISTICA для аналізу медико-біологічних даних : навчальний посібник для студентів вищих навчальних закладів / О. В. Гойко. - Київ, 2004. - 76с.

3. Правова статистика: Підручник / Моїсеєв Є.М., Джужа О.М., Василевич В.В. та інш.; За загальною редакцією професора О.М.Джужі. – Київ:Атіка, 2008. - 326с.

Одержано 8.02.2021