

Д. Айстраханов

м.Київ

**ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ МОДЕЛЮВАННЯ
ПРОФЕСІЙНОЇ КОМПЕТЕНТНОСТІ ВИПУСКНИКІВ
ПРОФЕСІЙНО-ТЕХНІЧНИХ ЗАКЛАДІВ**

Анотація. У статті розглядаються питання можливостей та правомірності застосування статистичних пакетів STATISTICA та Eviews, що використовуються у статистичних, економічних, фінансових та інших дослідженнях, до процесів автоматизації моделювання професійної компетентності випускників професійно-технічних навчальних закладів.

Ключові слова: програмний пакет, моделювання.

Айстраханов Д.

**ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ МОДЕЛИРОВАНИЯ
ПРОФЕССИОНАЛЬНОЙ КОМПЕТЕНТНОСТИ ВЫПУСКНИКОВ
ПРОФЕССИОНАЛЬНО-ТЕХНИЧЕСКИХ ЗАВЕДЕНИЙ**

Аннотация. В статье рассматриваются вопросы возможностей и правомерности применения статистических пакетов STATISTICA и Eviews, используемых в статистических, экономических, финансовых и других исследованиях, к процессам автоматизации моделирования профессиональной компетентности выпускников профессионально-технических учебных заведений.

Ключевые слова: программный пакет, моделирование.

D. Aistrahamov

**PROFESSIONAL SOFTWARE FOR SIMULATION OF
COMPETENCE GRADUATES OF VOCATIONAL-TECHNICAL
SCHOOLS**

Abstract. The article examines the opportunities and legality of the use of statistical packages STATISTICA and Eviews, used for statistical, economic,

financial and other studies, the processes of modeling professional competence of graduates of vocational-technical schools.

Keywords: software package, modeling professional competence.

Постановка проблеми. В останні десятиліття в педагогічних роботах усе більшого поширення набувають дослідження, пов'язані з моделюванням понятійних, процесуальних, структурних і концептуальних характеристик та окремих «сторін» навчання й виховання в межах національного соціокультурного простору на загальноосвітньому й професійно-орієнтованому рівнях. Стрімкий розвиток інформаційних відносин в сучасному суспільстві зумовлює виняткову актуальність моделювання в педагогіці – галузі знань, безпосередньо пов'язаній з інформаційними процесами. Так, з 2001 року по теперішній час в Україні захищено 228 дисертацій, в яких використовується термін «моделювання». Але В.П. Беспалько справедливо вважає, що методи дослідження педагогічної науки “ще сильно контрастують на фоні загальних успіхів природничих і точних наук” [1, с. 4]. Зважаючи на це, завдання вчених-педагогів вбачається у тому, “...щоб скоріше вивести її зі стану традиційних описових рекомендацій і словесних побудов на шлях моделювання, сучасного експерименту і практичного застосування результатів” [1, с. 4]. Математичне моделювання професійної компетентності випускників ПТНЗ є однією з необхідних умов удосконалення професійно-технічної освіти.

Метою даної статті є обґрунтування використання програмних пакетів, аналіз їх можливостей для вирішення завдань моделювання професійної компетентності випускників професійно-технічних навчальних закладів.

Виклад основного матеріалу. У сучасній світовій практиці фахівці різних галузей знань широко застосовують статистичний апарат для ретельного аналізу результатів роботи. Персональні комп'ютери відкривають статистичні методи широкому колові користувачів, яке на 90% відсотків

складається здебільшого з непрофесійних статистиків. Ці методи стали доступні не теоретично, а в високоінтегрованих середовищах, де весь етап аналізу від введення даних, імпорту їх з інших баз даних до побудови Автозвіту і збереження їх у форматі HTML повністю автоматизований.

Ринок пропонує користувачам більше тисячі статистичних програм [2]. Короткий огляд програмних продуктів наведено у табл. 1.

Таблиця 1.

Класифікація статистичних пакетів

Типи	Російські	Зарубіжні
Професійні	Немає	SAS, BMDP
Універсальні	STADIA, Olymp	STATGRAPHICS, SPSS STATISTICA, S-PLUS
Спеціалізовані	Mesosaur, DataScore, Класс- Мастер, Эвреста, САНИ	Велика різноманітність

Вибір пакета для оброблення даних залежить від характеру поставлених завдань, обсягу оброблюваного матеріалу, кваліфікації користувачів, обладнання і т. д. При виборі статистичних пакетів загального призначення можна керуватися наступними вимогами [3]. Пакет повинен:

- містити достатньо повний набір стандартних статистичних методів;
- бути простим для швидкого оволодіння та використання;
- відповідати вимогам можливості перетворення та організації зберігання даних і можливості обміну з іншими базами даних;
- мати широкий набір засобів графічного представлення даних;
- надавати зручні можливості для включення у звіти вихідних даних, графіків, проміжних та кінцевих результатів розрахунків;
- бути недорогим.

На практиці, як правило, використовують декілька систем, призначених для статистичного дослідження даних: SAS, SPSS, S-плюс,

STATISTICA та ін. Виявилося, що статистичний аналіз дуже зручно проводити в середовищі Windows. Це середовище дуже природне для нього. Перевага STATISTICA в тому, що вона із самого початку заснована на Windows–технологіях, а не переносилася з великих комп'ютерів. Можна сказати, що STATISTICA є Windows–середовищем, призначеним для всебічного проведення статистичного аналізу даних.

Є наступні 7 причин [4], по яких краще вибирати STATISTICA серед інших систем: це передусім повна узгодженість зі стандартами Windows, легка навченість (багато в чому пов'язана з реалізацією Windows–технологій, що відображають інтуїтивні уявлення статистиків про середовище аналізу даних), мінімальні вимоги до комп'ютера, унікальна наукова і презентаційна графіка, вичерпний набір класичних та сучасних методів статистики, представлених у системі, повний переклад системи на російську мову і, що важливо, ціна, яка, завдяки політиці компанії розробника, доступна широкому колові користувачів.

Розробником STATISTICA є американська фірма StatSoft Inc. (США). У ній і реалізований так званий графічно–орієнтований підхід до аналізу даних, сенс якого в тому, щоб отримувати всебічне візуальне представлення даних на всіх етапах статистичного оброблення та на основі цього вибирати наступний крок аналізу.

Система побудована за модульним принципом і містить усі відомі методи статистичного аналізу: від описової статистики та побудови таблиць спряженості до просунутих методів класифікації та аналізу цензурованих даних (аналізу виживаності). Дані можуть бути легко експортовані в популярні бази даних й імпортовані з них за допомогою кількох клацань миші.

У STATISTICA є чудовий засіб – імовірнісний калькулятор, користуватися яким так само просто, як звичайним калькулятором, і який дає змогу швидко обчислювати відсоткові точки різних розподілів, а також розраховувати критичні значення різних статистик. У STATISTICA є сотні

типів графіків, призначених для візуалізації вихідних даних, розвідувального аналізу, графічного виводу результатів і вибору наступних напрямів аналізу. Крім стандартних типів графіків, у STATISTICA є велика кількість спеціалізованих статистичних графіків. Графіки можуть автоматично змінюватися при зміні пов'язаного з ними файлу даних. Особливо важливим є засіб ПЕНЗЕЛЬ для візуального аналізу даних. Вбудована в систему мова програмування STATISTICA BASIC дає змогу користувачеві розширювати можливості системи, програмувати власні оригінальні методи.

Також на практиці широко використовується економетричний пакет EViews. Наведемо порівняльний аналіз цих двох відомих статистичних пакетів: STATISTICA і Eviews [5,6]. Порівняння проведемо залежно від розв'язуваних завдань, простоти й зручності їх виконання, таким чином визначаючи переваги і недоліки цих пакетів. Відразу зазначимо, що обидві програми адаптовані до операційної системи Windows, підтримують технології OLE, drop-down тощо, є компактними і займають мало дискового простору. Обидві програми мають зручні графічні інтерфейси.

Програма STATISTICA має модульну структуру, тобто складається з модулів, кожен з яких використовується для вирішення свого конкретного класу завдань, а саме: аналіз часових рядів і прогнозування, множинна регресія, нелінійне оцінювання, факторний аналіз, моделювання структурними рівняннями, непараметрична статистика, дисперсійний аналіз (ANOVA / MANOVA), дискримінантний функціональний аналіз. Кілька модулів об'єднані в групу «промислова статистика: контроль якості, аналіз процесів, планування експерименту».

При завантаженні пакета програм STATISTICA і створенні нового файлу з'являється електронна таблиця, у якій стовпці є змінними, а рядки – спостереженнями. Зручність введення даних у програмі STATISTICA зумовлена тим, що файл таблиці подібний до аналогічного з програми Excel. Дана програма дає змогу імпортувати дані з інших Windows додатків і

програм DOS, таких як: MS Excel, MS Access, FoxPro, Paradox, dBASE, CSV, SPSS, а також з файлів *. txt.

На відміну від STATISTICA пакет Eviews не має модульної системи, але містить так зване вікно робочого файла, де можна зберігати низку об'єктів.

Об'єктна структура робочого вікна дає змогу працювати одночасно з різними типами інформації. Управління об'єктами здійснюється за допомогою процедур (procs), які, в свою чергу, можуть самі створювати нові об'єкти. Кожен об'єкт містить конкретний вид інформації: низку даних, коефіцієнти, графіки та діаграми, моделі, результати обчислень, таблиці (матриці) тощо.

Крім того, програма Eviews має командний рядок, куди вводяться певні команди, які дають змогу здійснювати статистичний аналіз даних. Команди можуть бути збережені в окремому файлі, що дає змогу дослідникові запам'ятати ретроспективу виконуваних дій.

На відміну від програми STATISTICA в Eviews перш, ніж ввести дані, необхідно задати їхній формат, далі створити об'єкт типу ряд, задати кількість змінних і кількість спостережень. Пакет Eviews дає змогу працювати з вісьмома типами даних (річні, піврічні, квартальні, місячні, тижневі (5 днів), тижневі (7 днів), щоденні й недатовані спостереження). Процедура введення і опису даних в додатку Eviews складніша, ніж в STATISTICA, тому бажано імпортувати дані з інших додатків. Програма дає змогу працювати з даними, запозиченими з таких програм, як: MS Access, Gauss, ODBS, SAS, SPSS, MS Excel, Stata, ACSII, HTML.

Аби легко візуалізувати як вихідні параметри моделі, так і кінцеві результати обчислень, необхідно побудувати різні графіки, діаграми, коррелограми, спектрограми і т.п. Перегляд у графічній інтерпретації вихідних даних у вигляді лінійних графіків, гістограм розподілів імовірності, діаграм і т.п. для обох програм здійснюється за допомогою команд меню. Перегляд результатів у вигляді графіків в програмі STATISTICA проводиться

безпосередньо з робочого модуля за допомогою відповідних кнопок, причому цей додаток дає змогу задати опцію побудови графіка автоматично після кожної проведеної процедури, також є можливість перегляду даних у різних масштабах. Побудова коррелограм АКФ і ЧАКФ у пакеті STATISTICA здійснюється тільки в різних вікнах, в Eviews же їх можна переглянути і в одному вікні, але в STATISTICA вказані межі білого шуму. Результати моделювання в Eviews у графічному режимі можна переглянути за допомогою команд основного меню робочого файлу.

Перегляд числових характеристик досліджуваних даних (середнє значення, стандартне відхилення, ексцес, ймовірність тощо) у додатку Eviews здійснюється за допомогою команд меню, щоб зробити це, в STATISTICA необхідно зайти в окремий модуль Основні статистики / Таблиці (Basic Statistics / Tables) (тут же є критичні значення різних розподілів). Набір описових статистик для обох пакетів однаковий.

У програмі STATISTICA оцінка коефіцієнтів одно– і багатофакторної лінійної регресії здійснюється в окремому модулі Множинна регресія (Multiple regression). Результати проглядаються в окремому вікні, де є коефіцієнти, які оцінені методом найменших квадратів, коефіцієнт детермінації, статистика Фішера оцінки значущості регресії, статистики Стьюдента оцінки значущості коефіцієнтів, коефіцієнт кореляції (матриця кореляцій), статистика Дарбіна Уотсона. Істотними недоліками програми STATISTICA є: по-перше, той факт, що оцінка коефіцієнтів простої регресії проводиться тільки методом найменших квадратів; по-друге, визначення наявності гетероскедастичності залишків доводиться проводити в окремому модулі (а саме, за допомогою тесту Спірмена в модулі непараметричні статистики).

На відміну від STATISTICA пакет Eviews дає змогу проводити оцінку регресії не тільки методом найменших квадратів, а й методами максимальної правдоподібності, зваженим і нелінійним методами найменших квадратів, досить просто набрати назву методу в командному рядку при оцінці

коефіцієнтів моделі. До того ж, Eviews дає змогу зробити поправку на гетероскедастичність із врахуванням характеру залежності помилок від незалежної змінної. За допомогою команди меню гетероскедастичність визначається тестом Уайта. Зазначимо, що за допомогою командного рядка можна задати методику виявлення гетероскедастичності методами Парку, Глейзера та ін.

Проблему мультиколінеарності факторів можна подолати в програмі STATISTICA двома способами: знайти оцінки методом головних компонент (реалізовується в модулі Факторний аналіз (Factor Analysis)) або використовувати гребневу регресію (можливо тільки для STATISTICA версії старші 6.0).

Оцінка нелінійної регресії у програмі STATISTICA проводиться в окремому модулі Нелінійне оцінювання (Non-linear estimation), тут можна як задати вид залежності самостійно, так і скористатися наявними: регресія логіт / пробитий, регресія експоненціального зростання, кус очно-лінійна регресія. Для оцінки коефіцієнтів нелінійної регресії довільного виду використовуються ітеративні методи, такі як квазі-ньютонівський, Хука-Дживса, симплексний та ін. Результатами оцінки є лише індекс детермінації і статистика Фішера. Підбір виду гладкої функції можна здійснити тільки на основі візуального аналізу графіків. На відміну від цього, у програмі Eviews підібрати нелінійну функцію, яка найкраще відображає залежність, можна на основі тесту Бокса-Кокса. Оцінка коефіцієнтів здійснюється на основі нелінійного МНК (NLS) і зваженого МНК (WLS).

Моделі бінарного вибору (логіт / пробитий моделі) легко оцінюються в пакеті STATISTICA в модулі Нелінійне оцінювання, вихідними даними слугують логарифмічна функція правдоподібності, обмежена логарифмічна функція правдоподібності, χ^2 -статистика, оцінені методом максимальної правдоподібності параметри моделі. На відміну від STATISTICA Eviews дає змогу будувати не тільки моделі бінарного вибору, а й моделі множинного вибору як з упорядкованими, так і з неупорядкованими альтернативами. Для

цього просто в полі вибору методу оцінювання треба зробити вибір на користь методу, відповідного шуканій моделі. Вихідними параметрами служать χ^2 -статистика, псевдо-коефіцієнт детермінації, логарифмічні функції правдоподібності.

Аналіз часових рядів у програмі STATISTICA здійснюється в модулі Аналіз Тимчасових рядів / Прогнозування (Time Series analysis / forecasting). Даний модуль містить такі методи дослідження часових рядів: модель авторегресії проінтегрованого змінного середнього, моделі інтервенції для АРПСС, експоненційне згладжування і прогнозування (адаптивні моделі прогнозування), аналіз розподілених лагів, сезонна декомпозиція і спектральний аналіз ряду. Крім цього, є вікно перетворення початкового ряду, яке дає змогу здійснювати різні алгебраїчні операції з рядом, брати різниці різного порядку, виділяти тренд методом ковзних середніх.

До переваг STATISTICA перед Eviews треба зарахувати побудову адаптивних моделей прогнозування. У програмі STATISTICA є зручна таблиця, в якій можна визначити тип моделі (адитивна, мультиплікативна, з лінійним трендом, що включає сезонну компоненту), а також задати параметри адаптації в трьох режимах: «вручну», автоматично (на підставі критерію збіжності за мінімальною середньою квадратичною помилкою, вибір за допомогою мережі різних мінімальних помилок. В Eviews аналогічна процедура побудови адаптивних моделей досить ускладнена: потрібно визначити тип моделі, метод оцінювання параметрів, вибрати початкові значення адаптаційних коефіцієнтів, критерій збіжності процесу розрахунку і т.д. в кількох діалогових вікнах. Також більш переважна побудова регресійних моделей з розподіленими лагами в програмі STATISTICA, ніж в Eviews, бо процедура побудови дуже проста в застосуванні: потрібно лише задати значення максимального лага і ступінь апроксимації полінома (лаги Алмон). Виділення тренду з ряду за допомогою гладких функцій та моделювання сезонної компоненти на основі гармонік в оболонці STATISTICA виконується у модулі Нелінійна оцінка. Процедуру

згладжування тренда за допомогою різних ковзних середніх легко виконати в обох програмах. Щодо визначення наявності тренда у вихідних даних, то його можна здійснити в обох програмах тільки за допомогою спектрального аналізу, або коррелограм АКФ і ЧАКФ, спеціального тесту (наприклад, тест Форстера–Стюарта) програми не проводять. Але в програмі Eviews, за допомогою менеджера тестів, таку перевірку можна здійснити, задавши відповідні команди в командному рядку.

Істотним недоліком програми STATISTICA є відсутність можливостей перевірки ряду на стаціонарність, і визначити його стаціонарність можна лише на основі візуального аналізу графіка ряду, спектрограми і коррелограм АКФ і ЧАКФ. На відміну від цього, пакет Eviews дає змогу провести як звичайний, так і розширений тести Дікі-Фуллера на наявність стаціонарності (тести одиничного кореня).

Побудова моделей АРСС в програмі STATISTICA виконується за допомогою модуля Аналіз часових рядів / Прогнозування. Ідентифікація моделі проводиться або на основі візуального аналізу коррелограм АКФ і ЧАКФ, враховуючи їхні специфічні особливості, або спектрів випадкових процесів. У програмі Eviews ідентифікацію моделі АРСС можна зробити ще й за допомогою Q-статистики (тест Лjung-Бокса). До того ж, є можливість порівняти дві значущі моделі АРСС, тобто зробити їх селекцію по критеріях Шварца і Акайке. Eviews виконує ці критерії на основі дисперсії помилки, яка повинна бути мінімальною. Тут зазначимо, що використання тестів Лjung-Бокса та інформаційних критеріїв у програмі STATISTICA можливе тільки для версій старших 6.0 в окремих модулях. Побудувати моделі стаціонарних рядів, що враховують зміну дисперсії (моделі ARCH і GARCH), можна тільки в програмі Eviews, STATISTICA таких можливостей не має. Крім того, знайдену модель АРСС за допомогою ARCH методу можна протестувати на гетероскедастичну складову. Принагідно зазначимо, що у версіях Eviews, старших 5.1, є можливості оцінки та специфікації моделей

TARCH і E-GARCH (моделює умовну дисперсію як асиметричну функцію, представлену в логарифмічному вигляді).

Програма Eviews дає змогу провести специфікацію векторної авторегресійної моделі та моделі виправлення помилки (дослідження коінтеграції між кількома змінними). Перевірка коінтеграції у програмі здійснюється за допомогою процедури Йохансена, яка визначає кількість векторів коінтеграції у групі часових рядів і забезпечує оцінки максимальної правдоподібності векторів коінтеграції і векторів швидкості приведення.

Зазвичай моделювання нестационарних рядів проводиться на основі моделі АРСС, де порядок інтегрування є порядком взяття різниці. У програмі STATISTICA за допомогою процедури перетворення ряду обчислюють різниці до тих пір, поки ряд не стане стаціонарним, причому стаціонарність можна перевірити лише на основі візуального аналізу, далі ідентифікують і будують модель АРСС. У додатку Eviews для побудови моделі АРСС використовується розширений тест Дікі-Фуллера, причому перевірка стаціонарності проводиться автоматично після взяття різниці першого або другого порядків (відповідний порядок треба вказати в діалоговому вікні). Звідси можна зробити висновок, що побудова моделі АРСС краще проводити в Eviews.

Висновок. Отже, останнім часом набув значного поширення в економіко-соціологічних дослідженнях аналіз панельних даних. Програма Eviews на сьогодні є одним з небагатьох статистичних пакетів, який дає змогу досліджувати панельні дані, причому, зазначимо, що така можливість є лише у програм, старших п'ятої версії. Візуальний аналіз представлених даних можна проводити як врахуванням структури панелі, так із врахуванням об'єднання даних. Eviews дає змогу легко зробити специфікацію моделі, що враховує лаг, а також оцінити модель з випадковими ефектами, використовуючи тест Хаусмана. Процедуру оцінки можна зробити за допомогою зваженого узагальненого МНК. Після оцінки побудованої моделі

по панельних даних можуть бути проаналізовані й відображені графічно фіксовані та випадкові ефекти.

Для завдань автоматизації моделювання професійної компетентності випускників професійно-технічних навчальних закладів рекомендується освоїти пакет Eviews, бо сфера застосування його охоплює всі аспекти сучасної теорії та практики економетричного моделювання. Високі функціональні можливості при обробці кількісних змінних дають підставу говорити про Eviews як про надійний інструмент для прогнозування показників. Так, для прогнозування фінансових часових рядів EViews, крім традиційних інструментів прогнозування, дає змогу використовувати аналіз відгуку на імпульси та моделювання умовної гетероскедастичності, як заходи волатильності. Додатково підкреслимо, що EViews надає можливість будувати прогноз відразу ж після побудови моделі.

Література:

1. Беспалько В.П. Основы теории педагогических систем. (Проблемы и методы психолого-педагогического обеспечения технических обучающих систем) [текст] / В.П.Беспалько. – Воронеж: Изд-во Воронежского ун-та, 1977. – 304 с.
2. <http://www.piter-press.ru/attachment.php>
3. Общая теория статистики: Статистическая методология в изучении коммерческой деятельности: Учебник / Под ред. О.Э. Башиной, А.А. Спирина. – 5-е изд., доп. и перераб. – М.:Финансы и статистика, 2000. – 440 с.
4. <http://www.rusmedserv.com/medstat/article002.htm>
5. Математические методы прогнозирования экономических показателей: учеб. пособ. / [Саяпова А. Р., Гусельникова Е. А., Лакман И. А., Шамуратов Н.М.]. – Уфа, 2000 – 128 с.
6. Шамуратов Н. М. Эконометрика: учеб. пособие / Шамуратов Н. М., Лакман И. А. – Уфа, 2007 – 118 с.