

Образованието по статистика/стохастика задължителна част от културата на Европа

Марусия Божкова, Боян Пенков, Пламен Матеев *

Резюме

Този доклад е посветен на успешния завършек на усилията, положени от колектив преподаватели от Факултета по математика и информатика при Софийския Университет и сътрудници на Института по математика и информатика на БАН, за откриване на нова бакалавърска специалност “Статистика” във ФМИ. Тези усилия са продиктувани от убедеността ни, че статистическото мислене е важно за съвременния гражданин и потребител, за да му осигури възможност да упражни правото си на информиран избор. Едно от средствата за това е именно статистиката, разбирана като наука за извличане на знания от данни и се основава на непрекъснато разширяващото се ядро от математически знания, в най-тясна връзка с тези от областта на вероятностите, а напоследък все повече и с тези от компютърните науки.

2000 MSC: 01A60, 01A73, 60-03, 62-03

Ключови думи: математическа статистика, образование, Софийски университет

1 Увод

Съвременният все повече глобализиращ се свят налага все по-настоятелно нуждата от актуална информация в различните и аспекти: политически, социален, икономически, финансов и т. н. От друга страна, за да

*Работата е частично спонсорирана по договор № ВУ-МИ-105/2005 с МОН, ФНИ

бъдат правилно използвани данните, с цел извличане на информация, те трябва да се анализират и на базата на ясна числена (количествена) аргументация човек би могъл да обясни конкретни явления, да вземе решение за конкретни действия и/или мерки и да прави прогнози за бъдещо развитие. Благодарение на Закона за големите числа и на Теорията на вероятностите статистиката може да анализира много сложни факти и данни и да прави предвиждания. Със следващите примери ще илюстрираме разнообразието на задачите и универсалността на статистиката като метод за анализ на данни.

Какво е статистика? Как работи тя като метод? Как помага за решаването на практически проблеми? Нека посочим няколко примера, които могат да бъдат обект на вниманието на един статистик. От тези примери можем да изберем съществени елементи на една статистическа задача.

При прогнозиране на изхода от парламентарни избори, анкетаторите интервюират предварително определени хора (извадка) от цялата страна и записват техните мнения. Прогнозата се прави на базата на тази информация. Подобна задача възниква при пазарното проучване (Каква част от жените предпочитат козметиката на “Арома”?); в социологическите изследвания (Каква част от населението ползва домашен компютър?); в промишлеността (Каква част от производството на даден артикул е некачествена?).

Резултата от производствения процес в химически завод зависи от много фактори. Чрез наблюдение на тези фактори през определен период от време можем да съставим прогнозно уравнение за производството относно наблюдаваните фактори. Как да намерим добро прогнозно уравнение? Ако уравнението се използва за прогнозиране, то едва ли ще се получи истинската стойност на произведеното количество, т. е. при прогнозната стойност винаги правим грешка. Можем ли да ограничим тази грешка? Кои са най-важните фактори, които определят прогнозната стойност? Подобни проблеми могат да се посочат в областта на образованието, социологията, психологията и физиката.

В допълнение на прогнозирането, статистика трябва да аргументира вземането на решения, базирани на направените наблюдения. Да разгледаме проблемите при определяне ефективността на нова противогрипна ваксина. За простота нека допуснем, че 10 лица са ваксинирани с новата ваксина и се наблюдават през зимния период. От тези 10, 8 не са боледували. Ефективна ли е новата ваксина?

Примерите, на които се спряхме са различни по своята природа и

сложност, но всеки един от тях предполага прогноза и вземане на решение. Нещо повече, във всеки от тези примери съществена част са данните, получени от случайни или предварително планирани наблюдения.

Класически пример от политическия живот, който показва, че нарушаването на принципа за случаен избор води до сериозни грешки е запитването, направено от американското списание "Литературен преглед" относно изхода на предстоящите президентски избори през 1936 г. [Янев, Димитров]. Кандидати били Ф. Д. Рузвелт и А. М. Ландън. Като генерална съвкупност редакцията използвала телефонните указатели. На избраните 4 милиона адреса по цялата страна били изпратени картички с въпроси за отношението към кандидатите за президенти. Списанието изразходвало големи суми за разпращането и обработката на картичките и обявило, че за президент ще бъде избран А. М. Ландън, и то с голяма преднина в гласовете. Резултатът от изборите се оказал противоположен на тази прогноза.

Тук са допуснати едновременно две грешки. Първо, телефонните указатели не дават репрезентативна извадка от населението на страната, защото абонатите са главно от по-заможни семейства. Второ, отговори са получени не от всички, а от хора, достатъчно уверени в своето мнение и свикнали да отговарят на писма - от представителите на деловия свят, които поддържали Ландън. Ако редакцията беше пристъпила по-критично към своята работа, тя би разбрала, че е открила и без това известното на всички.

В същото време социолозите Дж. Галъп и Е. Рупърт правилно прогнозирали победата на Ф. Рузвелт, обосновавайки я само на 4 хиляди анкети. Причината за този успех не е само правилното съставяне на извадката. Социолозите отчели, че обществото се разпада на социални групи, които са сравнително еднородни в отношението си към кандидатите за президенти. Затова извадката от слоя може да бъде малобройна, но със същия резултат на точност. С помощта на резултатите от изследването по слоеве може да се характеризира обществото като цяло.

Средната стойност има широка популярност като характеристика на данните, но може да ни заблуди, ако не отчитаме разпределението им. Следващите два примера ни убеждават в това.

Първият пример е по-близък до аудиторията. Нека два класа правят контролна работа и получават една и съща средна стойност от успеха: Добър (4). Но дали са еднакво добри? На пръв поглед това е така. Какво ще кажем обаче, ако в единия клас половината ученици имат Отличен

(6), а останалата половина - Слаб (2), докато в другия клас всички ученици имат Добър (4)? Този пример показва, че средната стойност не е достатъчна характеристика.

Депутати и министри определят възнаграждението си по средната работна заплата за страната. Известно е, че средната заплата може да бъде силно подвеждаща. Следващият пример илюстрира това твърдение. Нека в едно предприятие с 20 служещи заплатите се разпределят по следния начин, представен в таблицата:

Персонал	Възнаграждение	Общо
1	1 500	1 500
1	1 250	1 250
1	1 000	1 000
1	600	600
4	450	1 800
4	400	1 600
4	350	1 400
4	300	1 200
$N = 20$	$\bar{x} = 517.50$	$\Sigma = 10\,350$

Средната заплата е 517.50 лева. Но внимание, само 4 служещи печелят повече от 517.50 лева, докато 16 получават по-малко. Значи тази средна заплата не отразява правилно разпределението на заплатите. Една друга величина може да бъде от полза - това е медианата на заплатата. Тук служещите се разделят на две равни групи - тези, които печелят повече и тези, които печелят по-малко. Тогава ще констатираме, че в това предприятие медианата на заплатата е между 350 и 400 лева, и можем да приемем, че е 375 лева. Очевидно, медианата на разпределението по-добре характеризира средната заплата.

С тези примери, искаме да провокираме интереса на преподавателите по математика от началното образование до висшето, както и любопитството на младите хора към статистиката. Те са подбрани с цел да покажат, че в ежедневието си хората постоянно се сблъскват с необходимостта да извлекат нужната им информация като анализират грамотно заливащия ги поток от данни. В реалния живот постоянно се срещаме с прости на пръв поглед проблеми, за чието решение използваме елементарно статистическо знание, необходима част от математическата култура.

Нека да изясним термина “статистика/стохастика” използван в заг-

лавието. “Стохастика” произлиза от гръцкото “*στοχος*” предполагаемо, зависещо от случая и ние имаме предвид именно него, а “статистика” е употребен като синоним придобил по-широка популярност. Той е употребен като синоним на “*Ars conjectandi*” [Bernoulli, 1713] от автора на едноименното съчинение. В него той доказва, че с експерименти и анализ на получените данни могат да бъдат установени знания за “достоверността, вероятността, необходимостта и случайността на нещата”. Доказаното твърдение е популярно сега като “Закон за големите числа”.

У нас, все още, преобладаващото възприемане на “Статистика”, наложило се по стечение на историческите обстоятелства и утвърдено нормативно в Закон за статистиката [ДВ, 1999], е това от първата половина на XX век - събиране и предоставяне на данни при това, предимно икономически. Сега “Иконометрия” е международно признатият термин за дисциплината, обединяваща статистическите задачи в икономиката.

В следващите два параграфа ще представим ретроспекция на образованието по стохастика у нас и новата бакалавърската програма по Статистика във ФМИ, стартираща от есента на 2007 година.

2 Образованието по стохастика у нас

Без да се връщаме много назад, през 1947 г. е сформирана нова катедра в тогавашния Природо-математически факултет на СУ “Математическа статистика и застрахователна математика”, която 1963 г. се преименува в “Теория на вероятностите и математическа статистика”. През учебната 1961/62 г. е приемът на първия випуск със специализация Теория на вероятностите и математическа статистика. Известна криза във възпроизводството на специалността се появи след закриването на Единния център, както и настъпилите промени след 1989 г. у нас, които принудиха много колеги да напуснат страната.

В последните няколко години, благодарение на усилията на доцент Димитър Върндев, във ФМИ бе възстановено преподаването на дисциплини по стохастика, повече или по-малко, във всички функциониращи бакалавърски програми, а в бакалавърската програма “Приложна математика” в рамките на избираемите курсове в последната година, на практика, студентите могат да получат известна по-задълбочена подготовка в областта на стохастиката. За кратък период, поради въвеждането на магистърска и бакалавърска степени в образованието, специализация “При-

ложна статистика” можеше да се придобие в рамките на магистърската програма “Математическо моделиране в икономиката”. От 2004 година функционира самостоятелна магистърска програма “Вероятности и статистика”.

Извън ФМИ акредитирани програми “Статистика и иконометрия”, обявени в интернет-страницата на МОН, има в трите икономически висши училища - Варна, Свищов и София с бакалавърски и магистърски степени, а в Стопанския факултет на СУ само магистърска. Всички те предлагат образование тясно специализирано за целите на икономическата статистика.

3 Нова специалност

Свикнали сме да употребяваме “стохастика” като удобно съкращение за кръга задачи, включващ областите теория на вероятностите, случайни процеси, математическа статистика, анализ на данни, но за наименованието на специалността бе избрано името “Статистика”.

Новата специалност Статистика подготвя аналитични специалисти с умения за прилагане на методите на математическата статистика, съчетани със задълбочена математическа подготовка и ползване потенциала на най-новите информационни технологии. Те ще могат да решават разнообразни задачи, които се свеждат до анализ на различна по тип информация или по-точно до анализ на данни. Учебният план е разработен в съответствие с Европейските стандарти - подобласт 462 “Статистика” от Класификатора на специалностите в образованието на EUROSTAT, на традициите на ФМИ и изискванията на авторитетните статистически асоциации ASA и RSS, Консултативната група на актьорите в Европа и БАД, на Кодекса за застраховането на РБългария.

За основа на учебния план служи този на специалността “Приложна математика”. Измененията са насочени към това на студентите да бъде осигурено постепенно навлизане в областта и да получават непрекъснат поток от знания в стохастиката. В първата година има само едно изменение - нов курс “Увод в статистиката” като начално запознаване със статистиката и теорията на вероятностите. Основния курс по теория на вероятности и математическа статистика е изтеглен във втората година и е разделен на отделни модули в двата семестъра. Също така през втората година се предлага и курс по случайни процеси като задължителна

дисциплина, съобразен със съвременните приложения на стохастичното моделиране в областта на биологията, екологията, демографията, финансите, застраховането и други, свързани с управление на процеси. Приложната статистика също е разделена на два модула: Анализ на данни и регресия и Многомерни статистически модели, предвидени за пети и шести семестър съответно. В шести семестър студентите ще се запознаят в "Стохастични модели и симулации" със симулационно моделиране на конкретни реални ситуации, водещи до модели на масово обслужване, надежност, комуникационни системи и др.

Последната година семинарът "Статистическа лаборатория" дава възможност на студентите да подготвят решение на конкретна реална задача в избрана от тях област. Целта е да придобият опит в дейностите съпътстващи едно статистическо изследване, като самостоятелно преминат през етапите му от откриването и формулирането на изследователската задача до изготвянето на представяне с предложение за решение.

Чрез различни практикуми се създават условия за професионално овладяване на съвременния статистически софтуер. Според своите предпочитания студентите могат да избират курсове според областта на бъдещата си реализация измежду: Статистика с R, Финансова математика, Елементи от теория на информацията, Модели на смъртност, Времени редове, Структурно моделиране, Биостатистика, Статистически софтуер, Застрахователна математика, Изчислителна статистика (практикум), Алгоритми за статистически процедури, Статистически методи в социалните науки, Планиране на експеримента, Приложение на статистически методи - БДС, Проучвателна статистика (Data mining).

Завършването на специалността става с полагане на бакалавърски държавен изпит.

Завършилите успешно образователно-квалификационна степен бакалавър по статистика могат да продължат обучението си в магистърска степен по всички специалности на факултета и специалности, към които са се насочили като област на приложение на статистическата си квалификация, а също да се реализират главно като аналитични специалисти - математик-статистик или актюер в банки, застрахователни компании, експерт в промишлеността, търговията, агробизнеса, в държавната администрация, консултанти по статистика в научни институти и преподаватели във висши училища.

4 Заключение

Надяваме се програмата да спомогне за просперитета на нацията с увеличаване на културата в областта на стохастиката. Благодарим на всички, които взеха участие в нейното създаване и бъдеща работа.

5 Литература

[Bernoulli, 1713] Bernoulli, J., Ars Conjectandi, (opus posthumum) Paris IV, Basel, 1713, бълг.превод Б.Пенков в Б.Пенков (ред. и съст.) , Бернули, Лаплас, Колмогоров Вероятности, НИ, София, 1982.

[Laplace, 1820] Laplace, P.S., Theorie analytique des probabilités. Introduction., Paris, 1820, бълг.превод В.Сотиров в Б.Пенков (ред. и съст.), Бернули, Лаплас, Колмогоров Вероятности, НИ, София, 1982.

[ISCED, 1997] UNESCO Official Text, International Standard Classification of Education - ISCED, 1997.

[ЗС, 1999] Закон за статистиката, Държавен вестник бр. 57, 25.06.1999 (последно изм. ДВ бр. 88, 04.11.2005.

[СУ, 1980] Д.Цонев и колектив, История на Софийския университет, ред.Г.Наумов, НИ, София, 1980.

[МИ, 2002] В. Митев и колектив, превод от френски на Юношеска енциклопедия Larousse, Издателство "Слънце", София.

[Б. Димитров, Н. Янев, 1997] Вероятности и статистика, Университетско издателство "Св. Кл. Охридски".

Марусия Божкова

*Факултет по математика и информатика
Софийски университет "Св. Кл. Охридски",
бул. Дж. Баучер, № 5
1164 София, България
e-mail: bojкова@fmi.uni-sofia.bg*

Пламен Матеев

*Факултет по математика и информатика
Софийски университет "Св. Кл. Охридски",
бул. Дж. Баучер, № 5
1164 София, България
e-mail: pmat@fmi.uni-sofia.bg*

Боян Пенков

*Факултет по математика и информатика
Софийски университет "Св. Кл. Охридски",
бул. Дж. Баучер, № 5
1164 София, България
e-mail: boyanipenkov@yahoo.com*

The Education in Statistics/Stochastics - a Compulsory Part of European Culture

This report is dedicated to the successful finalization of the efforts of the team, consisting of professors from the Faculty of Mathematics and Informatics (FMI) at the Sofia University and collaborators from the Institute of Mathematics and Informatics of the Bulgarian Academy of Sciences, directed towards announcement of a new Bachelor speciality "Statistics" at FMI. These efforts are the consequence of our deep belief that statistical reasoning/thinking is a compulsory part of the citizens' and consumers' culture, that can assure them the right of equal opportunities in making informed choice.