

ОБРАЗОВАНИЕТО ПО СТАТИСТИКА/СТОХАСТИКА ЗАДЪЛЖИТЕЛНА ЧАСТ ОТ КУЛТУРАТА НА ЕВРОПА *

Марусия Славчова-Божкова[†]

Този доклад е посветен на успешния завършек на усилията на колектив преподаватели от Факултета по математика и информатика (ФМИ) при Софийския Университет и сътрудници на Института по математика и информатика (ИМИ) на БАН, за откриване на нова бакалавърска специалност “Статистика” във ФМИ. Тези усилия са продиктувани от увереността ни, че статистическото мислене може да осигури възможност на съвременния гражданин да упражни правото си на информиран избор.

1. Увод. Съвременният все повече глобализиращ се свят налага все по-настойчиво нуждата от актуална информация в различни аспекти: политически, социален, икономически, финансов и т. н. От друга страна, за да бъдат правилно използвани данните и от тях да бъде извлечена информация, те трябва да бъдат анализирани въз основа на ясна числена (количествена) аргументация. Благодарение на Теорията на вероятностите и на Закона за големите числа, статистиката може да анализира сложни ситуации и да предвижда. Следващите примери илюстрират разнообразието на задачите и универсалността на статистиката като метод за анализ на данни.

Какво е статистика? Как работи тя? 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 съществените елементи на една статистическа задача.

При прогнозиране на изхода от парламентарни избори, анкетъорите интервюират предварително определени хора (извадка) от цялата страна и отбелязват отговорите на зададени въпроси. Прогнозата се прави на базата на тази информация. Подобна задача възниква при пазарното проучване (Каква част от жените предпочитат козметиката на “Арома?”); в социологическо изследване (Каква част от населението чете вестник X ?); в промишлеността (Каква част от производството на даден артикул е некачествена?).

Резултатът от производствения процес зависи от много фактори. Чрез наблюдение на тези фактори през определен период от време можем да съставим прогноза за производството в зависимост от наблюдаваните фактори. Колкото и добро

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

[†]Авторът благодари на Б. Пенков и П. Матеев за оказаната помощ при подготовката на този доклад.

2000 Mathematics Subject Classification: 01A60, 01A73, 60-03, 62-03.

Ключови думи: математическа статистика, образование, Софийски университет.

прогнозно уравнение да сме съставили, то едва ли ще даде истинската стойност на произведеното количество, т. е. винаги ще правим грешка. Можем ли да ограничим тази грешка? Кой са най-важните фактори, които определят прогнозната стойност? Подобни проблеми могат да се посочат в областта на физиката, социологията, психологията и др.

В допълнение на прогнозирането, статистикът трябва да аргументира вземането на решения, базирани на направените наблюдения. Да вземем проблемите при определяне ефективността на нова противогрипна ваксина. За простота нека допуснем, че 10 лица са ваксинирани с новата ваксина и са наблюдавани през зимния период. От тези 10, 8 не са заболяли. Ефективна ли е новата ваксина и какво искаме да кажем с думата ефективна?

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 "Literary Digest" относно изхода на предстоящите през 1936 президентски избори [6]. Кандидати били Ф. Д. Рузвелт и А. М. Ландън. Извадката, която редакцията използвала била направена от множеството на телефонните абонати. 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 републиканецът 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 абонати не дават репрезентативна извадка от населението на страната (генералната съвкупност), 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.

В същото време демоскопите социолози Дж. Галъп и Е. Рупърт правилно прогнозирали победата на Ф. Рузвелт, основавайки я само на 4 хиляди анкети. Причината за този успех не е само правилното съставяне на извадката. Те 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 аритметична 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. 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 заплатата 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 служители заплатите се разпределят по начина, представен в таблицата:

Служители	Заплата	Общо
1	1 500	1 500
1	1 250	1 250
1	1 000	1 000
1	650	650
4	450	1 800
4	400	1 600
4	350	1 400
4	300	1 200
$N = 20$	$\bar{x} = 520$	$\Sigma = 10\ 400$

Средната заплата е 520 лева, но само 4 служители печелят повече от 520 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 служителите се разделят на две равни по брой групи - тези, които печелят повече и тези, които печелят по-малко. Тогава ще констатираме, че в това предприятие медианата на заплатата е 400 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 думите “статистика/стохастика” използвани в заглавието. “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. Якоб Бернули употребява “стохастика” като синоним на “*Ars coniectandi*” (изкуството да предвиждаме, отгатваме) в едноименното си съчинение [2]. 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 [7], е това от първата половина на XX век – събиране и предоставяне на данни при това, предимно икономически. Сега “Иконометрия” е международно

признат 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, започваща 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 през 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 по математика и механика, както и настъпилите промени след 1989 г. у нас, когато много колеги напуснаха България.

През последните няколко години, благодарение на усилията на доцент Димитър Въндев, 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 можеха да получат известна по-задълбочена подготовка в областта на стохастиката. За кратък период, поради въвеждането на магистърска и бакалавърска степени в образованието, специализация “Приложна 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 висши икономически училища – във 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.

Необходимо е да отбележим какво е отношението към образованието по стохастика в международен план. Стандартите за училищна програма по математика в САЩ [3] отделят място на “Измерване” и “Вероятност и анализ на данни” наравно с “Аритметика”, “Геометрия” и “Алгебра” от предучилищна възраст до последната година на средното образование. Дословно ще цитираме аргументите за това:

“Measurement. The study of measurement is crucial in the school mathematics curriculum because of its practicality and pervasiveness in so many aspects of life. The Measurement Standard includes understanding the attributes, units, systems, and processes of measurement as well as applying the techniques, tools, and formulas to determine measurements. Measurement can serve as a way to integrate the different strands of mathematics because it offers opportunities to learn about and apply other areas of mathematics such as number, geometry, functions, and statistical ideas.”

“Data Analysis and Probability. Reasoning statistically is essential to being an informed citizen and consumer. The Data Analysis and Probability Standard calls for students to formulate questions and collect, organize, and display relevant data to answer these questions. Additionally, it emphasizes learning appropriate statistical methods to

analyze data, making inferences and predictions based on data, and understanding and using the basic concepts of probability. ”

Още през 1983 г. Paul D. Minton [1] е забелязал необходимостта от въвеждане на самостоятелна програма по Статистика в американските университети:

“Usually the name Statistics does not appear in either the list of undergraduate departments or that of undergraduate major from which entering students choose. Statistics is thus not seen as a discipline, nor as a possible career. . . Statistics has completely missed the opportunities for recruitment that occur during freshman college enrollment. The availability of undergraduate degree programs would remedy this.”

През 1999 г. в [1] ясно са формулирани целите и съдържанието на бакалавърска програма (BSc) по 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 със Световните и Европейски стандарти – подобласт 462 “Статистика” от Класификатора на специалностите в образованието на UNESCO и EUROSTAT ([3] и [5]), с традициите на ФМИ и с изискванията на авторитетните статистически асоциации ASA и RSS, с Консултативната група на актьорите в Европа и БАД, с 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 курса “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 преминават през етапите му от откриването и формули-

рането 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 в 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 банка, застрахователна компания, експерт в промишлеността, търговията, агробизнеса, в държавната администрация, консултант по статистика в научни институти и преподавател във висши училища.

4. Заключение. Надяваме се програмата да намери привърженици между кандидат-студентите и бъдещите ѝ възпитаници да намерят достойното си място в обществото. Убедени сме, че тя ще спомогне за развитие на културата в областта на стохастиката.

ЛИТЕРАТУРА

- [1] Amstat News. PUNSEW meets to Foster Growth in Undergraduate Education in Statistics, June 1999, Issue 265, 19.
- [2] J. BERNOULLI. Ars Conjectandi, Basel, 1713 (бълг. превод: Бернули, Лаплас, Колмогоров Вероятности, София, 1982).
- [3] Eurostat. Classification of fields of education and training (1999)
<http://ec.europa.eu/eurostat/ramon/nomenclatures/> (12.01.2007)
- [4] NCTM. Principles and Standards for School Mathematics,
<http://www.nctm.org/standards/>(12.01.2007).
- [5] UNESCO Official Text, International Standard Classification of Education – ISCED, 1997.
<http://www.unesco.org/>(12.01.2007).
- [6] Б. ДИМИТРОВ, Н. ЯНЕВ. Вероятности и статистика. Университетско издателство "Св. Кл. Охридски", 1997.
- [7] Закон за статистиката. Държавен вестник бр. 57, 25.06.1999 (последно изм. ДВ бр. 88, 04.11.2005).

Марусия Божкова
ФМИ – СУ “Св. Кл. Охридски”
бул. Дж. Баучер, № 5
1164 София, България
e-mail: bojкова@fmi.uni-sofia.bg

и
Институт по математика и информатика
Секция Вероятности и статистика
ул. Акад. Г. Бончев, бл. 8
1113 София
e-mail: bojкова@math.bas.bg

THE EDUCATION IN STATISTICS/STOCHASTICS – A COMPULSORY PART OF EUROPEAN CULTURE

Maroussia Slavtchova-Bojkova

This report is dedicated to the successful finalization of the efforts of a team from the Faculty of Mathematics and Informatics (FMI) of Sofia University and from the Institute of Mathematics and Informatics (IMI) of the Bulgarian Academy of Sciences, directed towards announcement of a new bachelor degree in “Statistics” at FMI. These efforts are the consequence of our deep belief that statistical reasoning is a compulsory part of the citizens’ and consumers’ culture, that can assure them equal opportunities in making informed choice.