

Дискриминантен анализ на резултатите от обучението по Висша математика в Русенски Университет

Стефка Караколева

Abstract: Discriminant Analysis of the Results of Education in Mathematics at the University of Ruse. This paper presents the application of the discriminant analysis of the results in education in Mathematics at the University of Ruse. The obtained discriminant model is applied to predict the results of teaching students in Mathematics, which are trained with traditional methods and by using computer algebra system MATLAB.

Key words: computer-based education, mathematics, statistics, discriminant analysis, SPSS, MATLAB

ВЪВЕДЕНИЕ

Предмет на изследването са резултатите от обучението на студенти по Висша математика – 3 част с използване на система за математически изчисления и визуализация MATLAB (нова методика). *Генералната съвкупност* [2,5,7] обхваща студенти от Русенски университет, които изучават Висша математика. За целта на изследването е направена серийна извадка от общо 3001 лица – студенти от 1. и 2. курс в Русенски университет „Ангел Кънчев“, обучавани през периода 2010-2014 години. Елементите за изследване във всяка серия се разделят в две хомогенни групи – **контролна група** (КГ) и **експериментална група** (ЕГ). Контролната група включва студенти, които са обучавани по традиционната методика, без използване на система за математически изчисления, а експерименталната група - студенти, обучавани по новата методика.

Формирането на извадката отговаря на всички изисквания [1,5] за формиране на извадки в педагогическите изследвания: (1) Извадката е представителна. Единиците на извадката са избрани чрез случаен подбор измежду студенти от 4 факултета и 18 специалности на Русенски университет, обучавани по Висша математика през периода 2010-2014 години, при спазване на принципа на случайност и при съблюдаване на изискването контролната и експерименталната групи да имат относително равен брой отличници, средни и слаби по успех студенти; (2) Обемът на извадката е достатъчно голям: изследваните 3001 студенти представляват около 30% от всички обучавани студенти в РУ и 45% от студентите 1 и 2 курс, изучаващи Висша математика; (3) Изборът на единиците от различни специалности е направен пропорционално на броя на студентите в съответните специалности. В изследването участват студенти, обучавани се в редовна и задочна форма, чийто брой е в съотношение, също отговарящо на съотношението в генералната съвкупност.

В настоящата статия са изложени резултати от изследванията, свързани с **дискриминантен анализ** на различията между изследваните студенти по признака „Повишаване на успеха по математика“. Получен е дискриминантен модел, чрез който за всеки нов случай (студент) може да се определи с различна степен на увереност дали той ще си повиши успеха или не, ако бъде обучаван по новата или по традиционната методика.

ИЗЛОЖЕНИЕ

1. ПОСТАНОВКА НА ЗАДАЧАТА

Целта на дискриминантния анализ е, като се използват данните за изследваните лица, да се получи дискриминантна функция (модел) от вида:

$$D = a + b_1X_1 + b_2X_2 + \dots + b_nX_n,$$

където:

- $X_i, i = 1, \dots, n$ са дискриминационни променливи;

- a – свободен член;
- $b_i, i = 1, \dots, n$ – коефициенти на дискриминантната функция.

2. ИЗСЛЕДОВАТЕЛСКИ ЕТАПИ

Построяването на модела включва следните изследователски етапи [6]:

- Избор на групираща (зависима) дискриминантна променлива;
- Избор и оценка на избора на дискриминационните променливи;
- Построяване на дискриминантния модел и определяне на неговата точност и статистическа значимост;
- Определяне на дискриминантен критерий;
- Тълкуване на получените резултати.

За решаване на изследователските задачи се използва процедурата Discriminant от менюто Classify на SPSS.

2.1. Избор на групираща (зависима) дискриминантна променлива

За дискриминантна променлива е избрана „Повишаване на успеха по математика“ с две дихотомни стойности 0 – не и 1 – да;

2.2. Оценка на избора на дискриминационните променливи

За независими (дискриминационни) променливи се задават променливите: „Методика на обучение“ (нова/стара), „Средна оценка по Висша математика 1 и 2“, „Факултет“, „Специалност“, „Пол“, „Народност“ и „Форма на обучение“, част от които отпадат в процеса на изследването.

Оценката на избора на дискриминационните признаци се извършва чрез тест за равенство на средните стойности в групите - критерий Wilk's Lambda [6]. Установено е за кои от признаците средните стойности на дискриминантната функция се различават значимо в изследваните групи.

Построеният дискриминантен модел трябва да отразява точното разделяне на изследваните групи. За да се осигури точността на модела, като изходна хипотеза се издига H_0 : Средните стойности на дискриминантната функция в изследваните групи са равни. Верността на изходната хипотеза се определя от стойността на показателя Significance. При това, нулевата хипотеза се приема, ако тестовата величина Sig има стойност, по-голяма от избрания риск за грешка (в случая 0.05) и съответната променлива следва да отпадне от анализа. Стойност на $Sig < 0.05$ доказва погрешност на нулевата хипотеза и статистическа значимост на различията на средните стойности на дискриминантната функция в групите.

Таблица 1.
Оценка на дискриминационните променливи

	Wilk's Lambda	F	df1	df2	Sig.
Нова/стара методика	,975	49,150	1	1933	,000
Средна оценка VM1+2	,897	221,808	1	1933	,000
Факултет	,995	10,054	1	1933	,002
Специалност	,983	33,167	1	1933	,000
Пол	,995	9,234	1	1933	,002
Народност	,999	2,052	1	1933	,152
Форма на обучение	1,000	,056	1	1933	,814

За изследваната извадка се установява, че признаците „форма на обучение“ и „народност“ не притежават необходимите дискриминиращи свойства, затова те

отпадат от анализа и като дискриминационни променливи се разглеждат останалите пет фактора, таблица 1.

2.3. Построяване на дискриминантния модел

Построяването на дискриминантния модел се състои в изчисляване и анализ на коефициентите на дискриминантната функция [6,7]. Качеството на модела се измерва чрез стойността на каноничния коефициент на корелация (Canonical correlation) между изчислените стойности на дискриминантната функция и реалната принадлежност на случаите към отделните групи. Получената с SPSS стойност на този коефициент 0.371 показва, че е налице умерена корелация. За проверка на теста за значимост на различията на средните стойности на дискриминантната функция в изследваните групи се използва показателя Wilk's Lambda, за който е получена със SPSS значима стойност 0.862, което показва висока значимост на различията на средните стойности на дискриминантната функция в групите.

Изчислените коефициенти на дискриминантната функция, фигура 1-(а), се използват за построяване на дискриминантен модел:

$$D = -0.79588 X_1 + 0.95650 X_2 - 0.08836 X_3 + 0.08159 X_4 + 0.23670 X_5 - 2.76278, \quad (1)$$

където:

- D е дискриминантна функция, показваща дали успехът се повишава или не;
- X_1 е „нова/стара методика“;
- X_2 е „средна оценка по Висша математика 1 и 2“;
- X_3 е „факултет“;
- X_4 е „специалност“;
- X_5 е „пол“.

Canonical Discriminant Function Coefficients

	Function
	1
Нова/стара методика	-.79588
Средна оценка ВМ1+2	.95650
Факултет	-.08836
Специалност	.08159
Пол	.23670
(Constant)	-2.76278

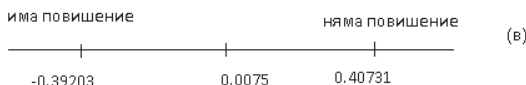
(а)

Functions at Group Centroids

	Function
	1
Увеличаване на успеха	.40731
няма повишение на успеха	-.39203
има повишение на успеха	-.39203
няма повишение на успеха	.40731

Unstandardized canonical discriminant functions evaluated at group means

(б)



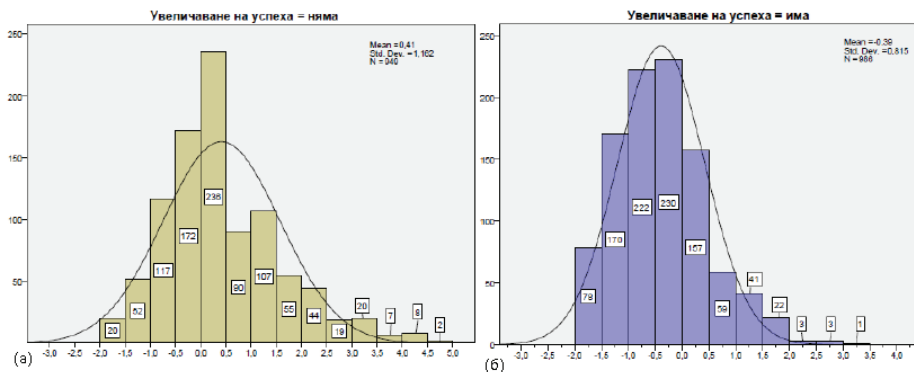
Фигура 1: Канонични коефициенти и групови центроиди на дискриминантния модел

2.4. Определяне на дискриминантен критерий

Важен етап от дискриминантния анализ е определянето на дискриминантния критерий [6] за принадлежност на отделен случай към всяка от групите, т.е. за какви стойности на функцията D може да се твърди (с определена степен на увереност), че успехът на нов студент ще се повиши и за какви стойности на D – няма да се повиши. За определяне на дискриминантния критерий, се използват груповите центроиди, т.е. средните стойности на дискриминантната функция в изследваните групи, фигура 1–(б). Въз основа на тези стойности се построяват интервалите, които показват принадлежност на нов случай към едната или другата група, фигура 1–(в). В случая се оказва, че отрицателните стойности на D показват повишаване на успеха. За точно определяне на съответствието между стойности на дискриминантната

функция, групова принадлежност и вероятност на прогнозата, се използват съхранените чрез Save променливи на дискриминантния анализ.

Точността на различията между изследваните групи зависи и от разсейването на стойностите в изследваните групи. Това разсейване се определя при анализ на хистограмите на разпределението на дискриминантната функция, фигура 2–(а),(б).



Фигура 2: Хистограми на разпределението на дискриминантната функция

2.5. Тълкуване на получените резултати

Полученият дискриминантен модел (1) се използва за прогнозиране [6] на резултатите от обучението с традиционна или нова методика за нов студент, за който са известни: средна оценка по Математика 1 и 2, факултет, специалност и пол.

Таблица 2.

Стойности на дискриминантната функция D , прогноза за повишаване на успеха по нова и стара методика и степен на увереност (%) за мъже, спец. КСТ, фак. ЕЕА

Нова методика	Средна оценка по Висша математика 1 и 2							
	2	2.50	3	3.50	4	4.50	5	5.50
D	-1.5924	-1.1142	-0.6359	-0.1577	0.3205	0.7988	1.2771	1.7553
Повишение	има	има	има	има	няма	няма	няма	няма
Увереност	78%	71%	63%	53%	56%	65%	73%	80%

(а) Нова методика, мъже, спец. КСТ

Стара методика	Средна оценка по Висша математика 1 и 2							
	2	2.50	3	3.50	4	4.50	5	5.50
D	-0.7966	-0.3183	0.1599	0.6382	1.1164	1.5947	2.0729	2.5512
Повишение	има	има	няма	няма	няма	няма	няма	няма
Увереност	66%	56%	53%	62%	71%	78%	84%	88%

(б) Стара методика, мъже, спец. КСТ

В таблица 2–(а),(б) са дадени получените стойности на дискриминантната функция, прогнози за това дали студентът ще си повиши успеха или не, както и съответните им вероятности (степен на увереност) за студенти–мъже от специалност „Компютърни системи и технологии“ (КСТ) при прилагане на нова или традиционна методика. Получени са данни за всички изследвани 18 специалности.

3. ИЗВОДИ

На основата на получените резултати от дискриминантния анализ, могат да се направят следните изводи:

- Наблюдава се ясно очертаваща се тенденция в полза на „новата методика“ по отношение на повишаване на успеха по математика. Във всяка от групите, повишаване на успеха се наблюдава 2–3 пъти по-често при студенти, обучавани по новата методика, отколкото при тези, обучавани по традиционната методика. Например, мъжете от специалности КСТ в 4 от 8 случая повишават успеха си, ако се обучават по „нова методика“, докато ако бъдат обучавани по „традиционна методика“, едва в 2 от 8 случая се очаква да се повиши успеха и то само за оценки от 2 до 2.50;
- В случаите на прогноза за повишаване на успеха, вероятностите за това са по-високи при обучение по „нова методика“, отколкото съответните вероятности за увеличаване на успеха при обучение по „стара методика“. Това означава, че студент, обучен по „нова методика“ е по-вероятно да повиши успеха си, отколкото ако се обучава по „стара методика“;
- В случаите на прогноза за липса на повишение на успеха, вероятностите за това са по-ниски при обучени по „нова методика“, отколкото съответните вероятности за липса на повишение на успеха при обучение по „стара методика“. Това означава, че студент, обучен по „нова методика“ е по-малко вероятно да не повиши успеха си, отколкото ако се обучава по „стара методика“;
- Като се използват хистограмите на дискриминантната функция (фигура 2), може да се направи извод, че за студенти с еднакви характеристики, например при определена стойност на $D \in [-1.5; -1]$, броят на случаите, в които успехът се повишава (170 или 17% от всички, при които има повишение) е над 3 пъти по-голям от студентите със същите характеристики (52 случая или 5% от всички без повишение), които не повишават успеха си.

ЗАКЛЮЧЕНИЕ

Вследствие на изследване резултатите от обучението на студенти по Висша математика и извършения задълбочен дискриминантен анализ на КГ и ЕГ, е получен дискриминантен модел (1), който е приложен за прогнозиране на резултатите от обучението според признака „повишаване на успеха по математика“ на студент, обучен по новата или традиционната методики.

Резултатите от дискриминантния анализ показват убедително предимствата на новата методика пред традиционната.

Настоящата статия е част от задълбочено изследване ефективността от прилагането на компютърни програмни системи за изчисление и визуализация в обучението по математика [3,4].

ЛИТЕРАТУРА

[1] Бижков, Г., Краевски, В. *Методология и методи на педагогическите изследвания*, Университетско издателство „Св. Климент Охридски“, София, 2007.

[2] Гочева-Илиева, С. Г. *Вероятности и статистика*, Университетско издателство „Паисий Хилендарски“, Пловдив, 2013.

[3] Караколева, С. *Доказване ефективността на компютърно съпроводено обучение по Висша математика в Русенски Университет*, Научни трудове на РУ, том 54, серия 6.1, Русе, 2015.

[4] Караколева, С. *Изследване резултатите от обучението по Висша математика в Русенски Университет чрез класификационни дървета*, Научни трудове на РУ, том 54, серия 6.1, Русе, 2015.

[5] Клаус, Г., Ебнер, Х. *Основи на статистиката за психолози, педагози и социолози*, Наука и изкуство, София, 1971.

[6] Моосмюлер, Г., Ребик, Н. *Маркетинговите изследвания с SPSS*, ИНФРА-М, Москва, 2009, ISBN: 978-5-16-002811-8.

[7] Харалампиев, К. *Въведение в основните статистически методи за анализ*, ИК „Йозеф Кнехт“, София, 2007, ISBN: 978-954-07-2847-6.

За контакти:

Стефка Караколева, Катедра “Приложна математика и статистика”, Русенски университет “Ангел Кънчев”, тел: 082-888 606, e-mail: skarakoleva@uni-ruse.bg