

МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2000
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2000
*Proceedings of Twenty Ninth Spring Conference of
the Union of Bulgarian Mathematicians
Lovetch, April 3–6, 2000*

**ПРИЕМЕН ИЗПИТ ПО МАТЕМАТИКА ЗА
СПЕЦИАЛИЗИРАНИТЕ СРЕДНИ УЧИЛИЩА –
СРАВНИТЕЛЕН СТАТИСТИЧЕСКИ АНАЛИЗ НА
РАБОТАТА НА ОЦЕНЯВАЩИТЕ КОМИСИИ**

Николай Петров Такучев

Професионалният път на младите хора, кандидатстващи в специализираните средни училища зависи и от оценителските качества на проверяващите конкурсните работи.

В настоящата работа са представени резултатите от статистическо изследване на оценителската работа на четири комисии по математика за периода 1994 – 1997 г.

Предлагат се статистически параметри за индивидуалната и груповата (на комисията като цяло) грешки. На тази основа е предложен метод за оптимизация на оценителската работа на комисия от изследвания вид.

Въведение. Обект на изследване в настоящата работа са комисии, проверяващи писмените работи по математика от приемните конкурси за специализираните средни училища – езикови, математически и икономически. Въз основа на статистическия анализ на работата на индивидуалните членове и на комисии като цяло, в настоящата работа се предлагат статистически параметри, характеризиращи количествено грешките при оценяване. На базата на въведените параметри е предложен метод за селекция на комисията с цел оптимизиране на състава ѝ по отношение точността на оценяване.

С любезното съдействие на Регионалния Инспекторат на МОН – Стара Загора, за целите на настоящото изследване бяха използвани документи, свързани с работата на различни оценителски комисии, в частност тези от изпитите по математика. По-долу са изложени резултатите от изследване на 4 комисии, оценявали изпитите по математика през 1994 (две), 1995 и 1997 г. В таблица 1 са приведени статистически данни и стойности на предлаганите статистически параметри за изследваните комисии, а в таблица 2 са дадени съответните им сравнителни стойности с тези за комисии по биология и български език.

Регламент на конкурсното оценяване. В практиката на конкурсите всяка конкурсна работа се оценява от двама оценители. Всеки оценител работи в двойки с различни свои колеги. Двойката оценители оценява заедно определен брой конкурсни работи. Двамата оценители поставят по една първоначална оценка. Окончателната

Таблица 1. Статистически данни (редове 1 – 6) и стойности на предложените в работата статистически оценки (редове 7 и 8) за грешката на комисиите, оценявали изпитите по математика след 7 клас за годините 1994 (две), 1995 и 1997 г.

Статистика за комисиите, оценявали изпита по математика след 7 клас					
№		1994 (1)	1994 (2)	1995	1997
1	брой конкурсни работи	952	548	1196	1185
2	брой оценители в комисията	46	27	17	15
3	среден брой работи на оценител	20,70	20,30	70,30	79,00
4	среден брой оценителски двойки, в които участва даден оценител	1,78	1,70	5,94	8,40
5	средна оценка на всички работи	3,15	3,82	3,88	4,07
6	максимална разлика в оценките на двойка оценители, шестобална оценка	1,25	1,18	1,56	1,13
7	средна грешка за интервала от 2 до 6	0,091	0,088	0,106	0,078
8	относителна грешка на обобщения оценител (за случайно оценяване максималната относителна грешка е 10)	2,825	2,329	2,952	2,394

Таблица 2. Сравнителна статистика за работата на 14 комисии, оценявали писмени конкурсни работи по математика, биология и български език за прием в специализирани средни училища в периода 1994 – 1998 г.

Номерата на редовете в таблицата съответстват на тези в таблица 1				
	български език и литература – изпити след седми клас	биология – изпити след седми клас	биология – изпити след осми клас	математика – изпити след седми клас
брой комисии	2	4	4	4
1 (осреднени данни)	597	317	341	970
2 (осреднени данни)	29	18	18	26
3 (осреднени данни)	23	22	19	48
4 (осреднени данни)	3,2	2,6	1,8	4,5
5 (осреднени данни)	4,12	4,14	3,47	3,73
6 (максимална стойност)	2,75	2,50	3,75	1,56
7 (осреднени данни)	0,260	0,125	0,128	0,091
8 (осреднени данни)	6,10	4,03	4,29	2,63

оценка е средната стойност от двете първоначални оценки, ако те се различават с по-малко от 0,5 (по шестобалната система). В противен случай работата се оценява от арбитър. По регламент оценителите работят независимо.

Основни предпоставки и използвани понятия. Принципно процесът на оценяване може да се сравни с процесите на измерване във физиката. При процеса на оценяване човекът-оценител играе ролята на своеобразен измерителен уред за

знания, който съпоставя числена оценка на определено количество знания. Всяка оценка включва в себе си грешка, генерирана в процеса на оценяване. Хората са различни и по отношение на способностите им на оценители – едни грешат по-често и/или в по-голяма степен от други.

Под *грешка* по-долу ще се разбира разликата между две оценки – поставената от оценителя и неизвестната точна оценка, съответстваща на знанията, показани в дадена конкурсна работа. Грешките са случайни величини. Осреднената грешка от множеството оценявания, в които оценителят е участвал е неговата *индивидуална грешка*.

Съвкупността от оценките може да се разглежда като получена от единен измерителен уред – съвкупността от оценители в комисията, наричан по-долу *обобщен оценител*. Обобщеният оценител има своя грешка, която е осреднена стойност от индивидуалните грешки на оценителите в комисията. Оценителските характеристики на обобщения оценител дават представа за оценителските качества на средностатистическия член на комисията.

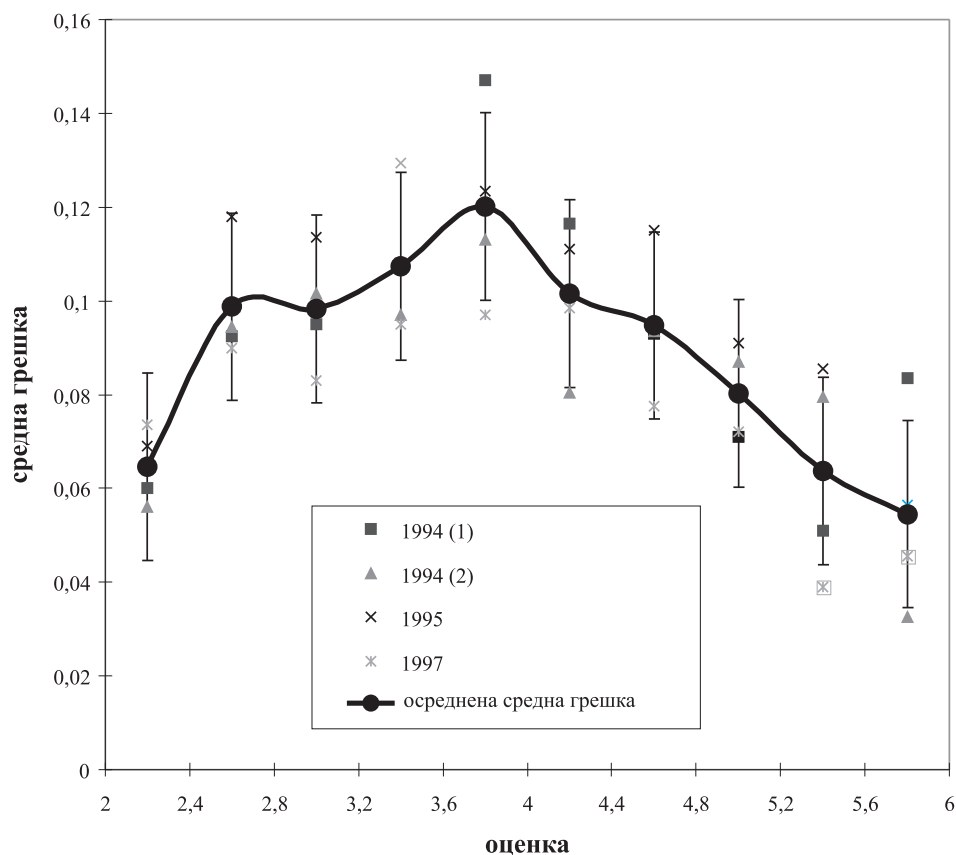
Индивидуална грешка. Разликата в двете независимо поставени оценки за дадена работа е толкова по-голяма, колкото по-големи са индивидуалните грешки на оценителите от двойката. Стандартното отклонение на средната оценка в случая се изчислява по формулата:

$$\text{Стандартно отклонение} = (2 \cdot (\text{средна оценка} - \text{една от двете оценки})^2 / 2)^{1/2},$$

т.е. стандартното отклонение е равно на половината от разликата между оценките на двамата оценители. В настоящата работа за характеристика на индивидуалната грешка на даден оценител се използва средната стойност от стандартните отклонения (включително нулевите) взети по абсолютна стойност, за оценките с негово участие.

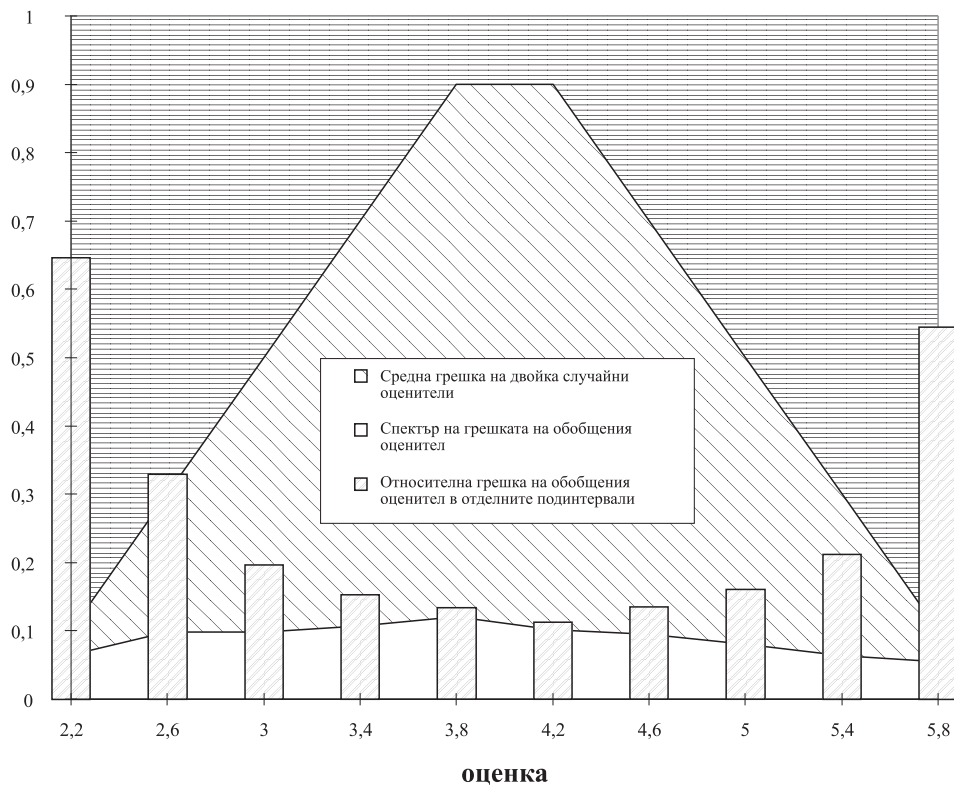
Тази характеристика е функция както на индивидуалната грешка на конкретен оценител, така и на индивидуалните грешки на останалите оценители, участвали в двойка с него в процеса на оценяване. Колкото по-голям е броят на конкурсните работи, в оценяването на които участва оценителят и с колкото по-голям брой други оценители е в двойка (вж. редове 3 и 4 в табл. 1 и 2), толкова в по-голяма степен зависимостта на предложената характеристика от индивидуалните грешки на останалите оценители се превръща в зависимост от средната грешка за комисията (грешката на обобщения оценител). Т.е. с точност до еднаквата за всички оценители средна грешка на обобщения оценител, предложената характеристика ще зависи единствено от индивидуалната грешка. За краткост по-долу тя е наричана просто индивидуална грешка.

Грешка на обобщения оценител. Разликата между двете оценки за дадена работа може да се разглежда и като следствие на неточността на обобщения оценител. По-долу като характеристика на грешката на обобщения оценител (груповата грешка) се прилага стандартното отклонение (половината от разликата между двете оценки). Под *средна грешка* на обобщения оценител в определен интервал от оценки се разбира средната стойност от всички грешки на обобщения оценител в интервала. Изменението на средната грешка в интервала на оценките (от 2 до 6), наричано по-долу *спектър на грешката*, е показано на фиг. 1 и на втори план на фиг. 2.



Фиг. 1. Осреднен спектър на грешката за обобщения оценител на 4 комисии по математика за годините 1994 (две), 1995 и 1997 г.

С точки са отбелязани осреднените в 10 подинтервала средни грешки на изследваните комисии. Плътната линия е степенен регресионен модел на осреднената средна грешка за всички комисии. Прави впечатление устойчивата форма на спектъра на грешката през годините – тя не зависи от състава на комисията и от годината на изпита, а само от вида на комисията и критериите на оценяване. Тази особеност на спектъра на грешката е още по-подчертана при комисии по биология и български език и би могла да се използва като обобщена характеристика на грешката за комисии от даден тип. Обща особеност на спектрите на грешката за всички видове комисии е намалението на грешката в двата края на интервала на оценките. Максимумът на грешката за комисии по математика е близо до оценка 4. Таблица 2, ред 7 показва, че средната грешка на комисии по математика е съответно 2,9 и 1,4 пъти по-малка от тази на комисии по български и биология.



Фиг. 2. Обобщен оценител за комисиие по математика. Спектър на грешката (светлата област зад колоните), средна грешка на двойка случайни оценители (по-тъмната област на трети план) и относителна грешка в 10 подинтервала на интервала на оценките (колони на преден план)

Случайни оценители и относителна грешка. Преценка за грешката на обобщения оценител може да се направи и относително – като се сравни с грешките на възможно най-грешащия оценител.

Оценката на *случайния оценител* е случайно число от интервала (2 – 6). Максималната грешка на двойката случайни оценители има просто статистическо разпределение (Симпсън) и се изчислява лесно. Средната ѝ грешка е половината от максималната. На фиг. 2 трапецовидната зона на трети план илюстрира разпределението на средната грешка на двойка случайни оценители.

Отношението на средната грешка на обобщения оценител в даден подинтервал от интервала на оценките с тази на двойката случайни оценители в подинтервала е *относителната грешка* на обобщения оценител за подинтервала. Нормално относителната грешка на обобщения оценител в подинтервала е по-малка от единица (колони на фиг. 2). Тя достига стойност 1 само когато средната грешка на обобщения оценител в подинтервала достигне тази на двойката случайни оценители. Под относителна грешка в целия интервал на оценките по-долу ще се разбира сумата от относителните грешки във всички подинтервали.

Относителната грешка е мерило за точността на оценяване. Колкото по-малка е тя, толкова по-точно е оценяването. За обобщения оценител относителната грешка за целия интервал на оценките е 0 при оценяване без грешка и 10 при оценяване със средна грешка, съпадаща с тази на двойката случайни оценители.

Относителната грешка на комисиите по математика е съответно 2,3 и 1,6 пъти по-малка от тази на комисиите по български и биология (табл. 2, ред 8).

Аномални стойности на грешката. Относителната грешка може да превиши стойност 1, ако оценителите в двойките оценяват тенденциозно – завишават или занижават оценката над стойността на индивидуалната си грешка (вж. табл. 2, ред 6). Стойности на относителната грешка, надвишаващи единица, се наричат по-долу аномални. Аномално високите грешки зачестяват в двата края на интервала на оценките, но осреднено, относителната грешка на обобщения оценител за комисиите по математика остава под единица за всеки от подинтервалите на интервала на оценките (за разлика от тази на част от комисиите по биология и български език).

Аномалното надценяване на едни кандидати е еквивалентно на недооценяването на другите. То променя резултатите от класирането и би следвало да се намали до възможния минимум.

Примерен метод за оптимизация на работата на оценителска комисия. Въведените по-горе статистически параметри – индивидуална грешка, средна грешка на обобщения оценител и относителна грешка дават възможност работата на оценяващата комисия да бъде оптимизирана с цел намаляване на грешката в три направления:

1. оптимизация на броя членове на комисията;
2. оптимизация на критериите за оценяване;
3. оптимизация на състава на комисията.

По-долу са приведени конкретни примери на базата на данните за четирите изследвани комисии.

1. Оптимизиране на броя членове на комисията. В табл. 1, ред 7 са приведени средните грешки за целия интервал на оценките на обобщените оценители за различните комисии. Средната грешка е функция в частност и на темпа на работа – претоварването на оценителите я увеличава.

Ако организационно се изключи обменът на информация между оценителите в двойката, „намаляващо“ средната грешка, то тя може да се използва като критерий за оптималност на темпа на работа. При зададен срок за оценяване на конкурсните работи и предварително опитно установен оптимален темп за оценяване, може да се направи преценка за оптималния брой на проверителите в комисията. Например при очаквани около 1000 писмени работи (2000 оценки), очакван срок за работа на комисията 7 дни и оптимален темп на работа 19 оценени работи за ден (за комисията с най-малка средна грешка, тази през 1997 г.), оптималният брой членове на комисията е $2000/(7 \times 19) = 15$ оценители, колкото са били в действителност членовете на комисията през 1997 г. (табл. 1, ред 2).

2. Оптимизиране на критериите за оценяване. Относителната грешка е индикатор за областите от спектъра на грешката, в които е необходимо детайлизи-

ране на критериите за оценяване, за да се изключат тенденциите за надценяване или подценяване. На фиг. 2 с колони е показано разпределението на относителната грешка за осреднения обобщен оценител, съответстващ на комисиите по математика (от фиг. 1), в 10 подинтервала на интервала на оценките. „Зоната на увереността“ – с ниска относителна грешка при оценяването, е в подинтервала 2,8 – 5,2 и е заградена от две зони на „нарастваща неувереност“ – в посока към двата края на интервала на оценките, т.е. комисиите по математика се нуждаят от детайлизиране на критериите за оценки от подинтервалите 2 – 3 и 5 – 6.

3. Оптимизиране на състава на оценяващата комисия. Индивидуалните грешки на всеки от оценителите в комисията могат да бъдат изчислени и съпоставени след приключване на работата на комисията. Ако двойките оценители са формирани по случаен начин, оценителите в двойката са работили независимо и всеки оценител е участвал в няколко двойки, то получените резултати позволяват сравнение на индивидуалните оценителски качества на членовете на комисията. Изводите от подобни сравнения трябва да се правят внимателно, да се основават на участие на оценителя в поне две комисии, като се търсят и допълнителни източници на информация за работата на конкретния оценител (например работните тетрадки) и за организацията на работата в комисията като цяло.

На фиг. 3 са приведени индивидуалните грешки (във вид на светли и тъмни колони) на оценителите, участвали в две от изследваните комисии по математика. Зад колоните е показана тъмна ивица, чийто хоризонтален горен край съответства на средната грешка на обобщения оценител за 1995 година. Индивидуалните грешки за предходната година са приведени към тези за 1995 г. чрез умножение с нормировъчен коефициент, равен на отношението на средните грешки за двете комисии.

От фигурата се вижда, че оценители 6 и 8 имат ниско ниво на индивидуалната си грешка. В следващи години те могат да бъдат привлечени като арбитражи.

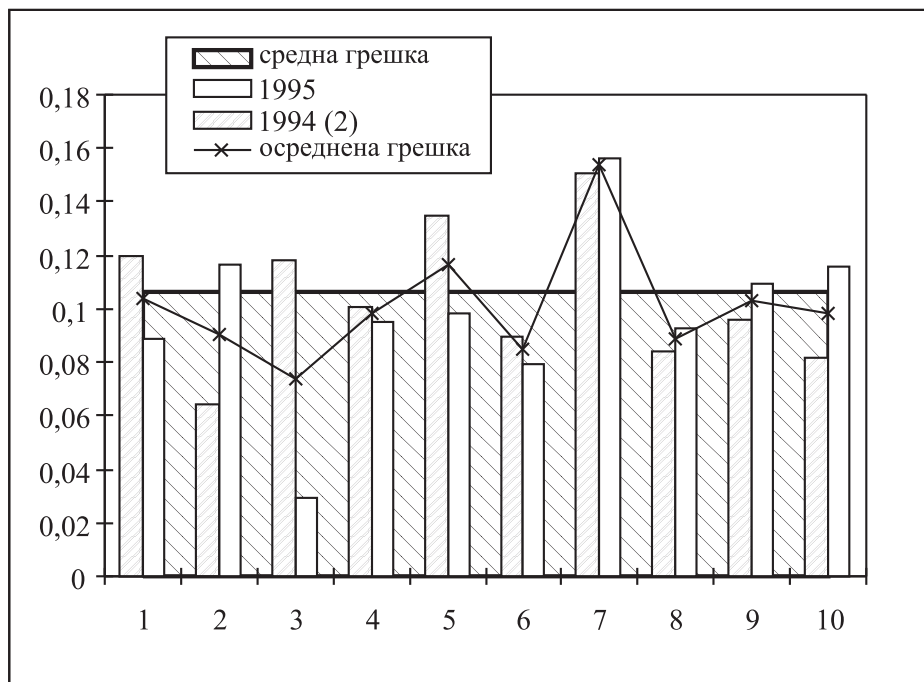
Оценител 7 има устойчиво завишена индивидуална грешка и би следвало да се обмисли замяната му при формирането на следващи комисии.

В документацията за комисията от 1995 г. се забелязват и организационни проблеми:

Оценител 3 е участвал в 3 двойки – и в трите с един и същи партньор. Това е довело до индивидуална грешка, трикратно по-ниска от средната, вероятно в следствие на съгласуване на оценките, още повече, че индивидуалната му грешка от предходната година е висока. Същият принцип на самосформиране на двойките се забелязва и при другите оценители в тази комисия. Това води до изкривяване на оценката за случайната грешка. До същия резултат води и крайно неравномерното разпределение на работата по оценителите – някои са участвали в десет оценителски двойки, други само в една.

Най-добрият вариант е комисия с относително постоянен през годините състав, оптимизиран по брой на членовете и индивидуалните им грешки. Ниската средна грешка на такава комисия би могла да се достигне и поддържа чрез ежегодна ограничена селекция на състава.

Средната грешка на обобщения оценител като критерий за арбитражиране. Ако разликата в шестобалните оценки на двойката оценители надвиши 0,5,



Фиг. 3. Грешки на оценителите, участвали в две комисии по математика. Индивидуалните грешки са дадени във вид на колони. Горната ограничителна линия на тъмната област зад колоните показва средната грешка за 1995 г.

конкурсната работа се оценява и от арбитър. Приетата неясно от какви съображения критериална стойност 0,5 (еднаква за всички видове оценителски комисии) би следвало да получи интерпретация, обвързваща я със средната грешка на обобщения оценител за даден вид комисия. Тази стойност превишава петкратно средната грешка за комисии по математика (табл. 1, ред 7), но само двукратно средната грешка на комисии по български език и литература (ред 7 на табл. 2). Т.е. изискванията към оценителите по математика са относително по-либерални. От статистическа гледна точка конкурсните работи по математика биха се оценявали от арбитър при същото ниво на грешката, както тези по български език, ако критерият за арбитриране по математика се измени на 0,2.

Николай Петров Такучев
 Тракийски университет, Студентски град
 катедра „Математика и Физика“
 Стара Загора 6000
 e-mail: npt@af.uni-sz.bg

**ADMISSION EXAMINATION IN MATHEMATICS FOR THE
SPECIALIZED HIGH SCHOOLS – COMPARATIVE STATISTICAL
ANALYSIS OF THE ASSESSMENT COMMISSIONS WORK**

Nikolay Petrov Takoutchev

The professional realization of the young people, applying in the specialized high schools depends to a certain extent on the abilities of the assessment commission members.

This paper presents the results of statistical investigation of the assessment abilities of four commissions for admission examination in mathematics for the period 1994 – 1997.

Statistical parameters for the errors are proposed – both individual for the commission members and general for the commission as a whole.

On this basis an objective method for comparing individual assessment abilities is proposed, e.g. a method for commission selection, aiming at optimization of the assessment accuracy.