

СРАВНИТЕЛЕН АНАЛИЗ МЕЖДУ РЕАЛНИ ИЗПИТНИ ТЕСТОВЕ

Драго Михалев, Райна Алашка

michalev@abv.bg, alraina@abv.bg

ВТУ „Тодор Каблешков”, ул. „Гео Милев”158, София 1574
БЪЛГАРИЯ

Резюме: Направен е сравнителен анализ между четири реални изпитни теста. Пресметнати са корелационните зависимости между тях по задачи и теми. Направен е дисперсионен анализ за значимостта на получения бал от тестовете.

Ключови думи: тест, изпит, корелационен анализ, дисперсионен анализ, образование, статистика.

УВОД

Направен е сравнителен анализ между четири реални теста за прием в ВТУ „Тодор Каблешков” проведени на 21 юни 2008 г., 29 юли 2008 г., 21 юни 2009 г. и 28 юли 2009 г. (виж [5]). За краткост ще ги означаваме съответно с 1 тест (изпит), 2 тест (изпит), 3 тест (изпит), 4 тест (изпит). Всеки от тези тестове съдържа 20 задачи с избираем отговор и 10 задачи със свободен отговор. Към всяка задача с избираем отговор са дадени 4 възможности за отговор, от които точно една е правилният отговор. За правилен отговор на задачите с избираем отговор се получават – 3 точки, за грешен – 0 точки, за неотбелязан отговор – 1 точка. За задачите със свободен отговор се дават 6 точки при правилен отговор и 0 точки при грешен или неотбелязан отговор. Пресметнати са основните числови характеристики за всички тестове както това е направено в [4], [5] за 4 тест. Задачите от тестовете са разпределени на 6 основни теми (виж таблица 1). Пресметнати са корелационните зависимости между тях по задачи и теми. Направен е дисперсионен анализ за значимостта на получения бал от тестовете.

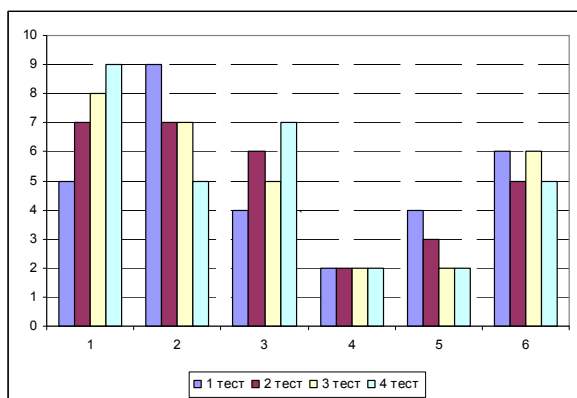
2. ОСНОВНИ ЕМПИРИЧНИ СТАТИСТИКИ НА ТЕСТОВЕТЕ

1. Структура на тестовете

Таблица 1

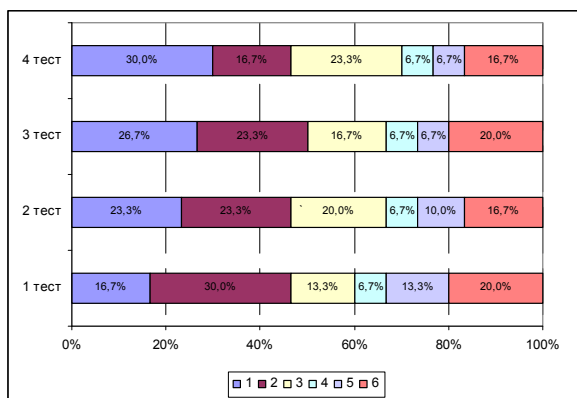
№	ТЕМА	Брой задачи			
		1 тест	2 тест	3 тест	4 тест
1	Алгебрични действия. Степенна, логаритмична и показателна функция. Прогресии.	5	7	8	9
2	Уравнения, неравенства, системи.	9	7	7	5
3	Планиметрия.	4	6	5	7
4	Стереометрия.	2	2	2	2
5	Тригонометрия.	4	3	2	2
6	Комбинаторика. Класическа вероятност. Граница на функция. Производна на функция.	6	5	6	5
Общо:		30	30	30	30

На фигура 1 е показан броят на задачите по теми за четирите теста.



Фиг.1

На фигура 2 е показана структурата на тестовите по броя на задачите от различните темите за четирите теста.



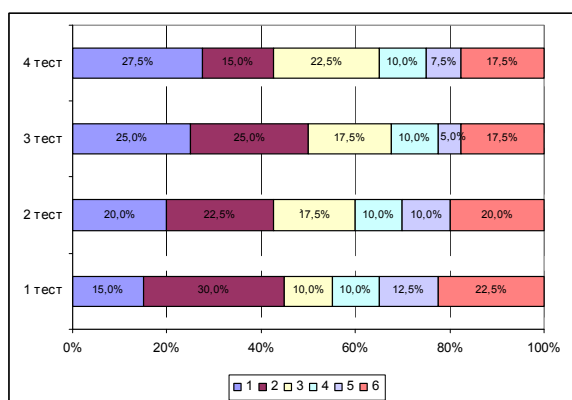
Фиг.2

В таблица 2 е даден максималният брой точки, които може да получи кандидатът по всяка тема за различните тестове.

Таблица 2

№ на темата	Максимален бал за темата			
	1 тест	2 тест	3 тест	4 тест
1	18	24	30	33
2	36	27	30	18
3	12	21	21	27
4	12	12	12	12
5	15	12	6	9
6	27	24	21	21
Общо:	120	120	120	120

На фигура 3 е показана структурата на теста по брой точки от различните теми за четирите теста.

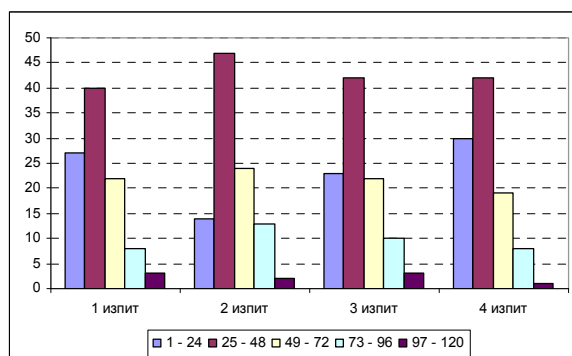


Фиг.3

2. Анализ на получените оценки

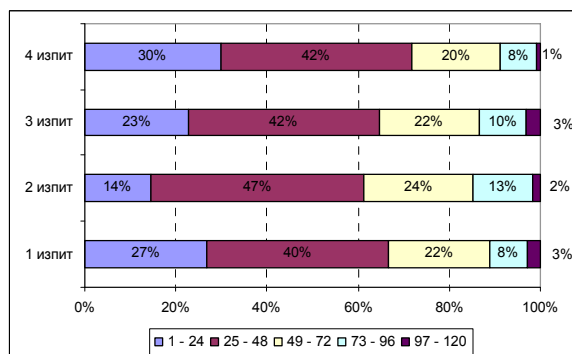
Групираме получените тестови балове, за всеки от тестовите в пет групи с еднаква ширина на интервала.

На фигура 4 са построени сравнителни хистограми на абсолютните честоти, приведени към 100 кандидата по групи за всеки от тестовите.



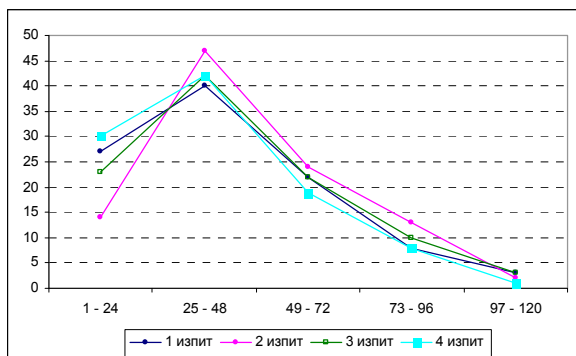
Фиг.4

На фигура 5 са дадени линейни диаграми за процентното съотношение на кандидатите по групи за всеки от тестовите.

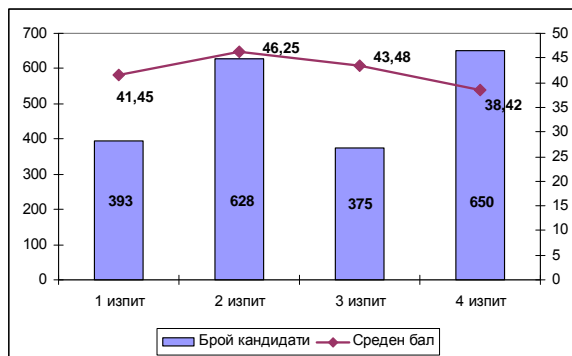


Фиг.5

На фигура 6 са съпоставени полигоните за различните тестове по групи.



Фиг.6



Фиг.7

3. Основни статистики на тестовете

В таблица 3 и таблица 4 са дадени средните статистически величини и отклонение за бала по тестове (виж [1], [2], [3], [4]) съответно за негрупиран и групиран данни.

◆ *Негрупиран данни*

Таблица 3

Изпит	Брой изпитвани	Средно	Стандартно отклонение	Мода	Медиана
1 изпит	393	42,10	23,59	15	37
2 изпит	628	46,34	21,93	30	41
3 изпит	375	43,41	23,47	27	38
4 изпит	650	39,24	21,46	18	35

◆ *Групиран данни*

Таблица 4

Изпит	Брой изпитвани	Средно	Стандартно отклонение	Мода	Медиана
1 изпит	393	41,45	24,40	35,23	38,99
2 изпит	628	46,25	22,65	39,04	43,20
3 изпит	375	43,48	24,64	36,76	40,57
4 изпит	650	38,42	22,47	33,39	36,56

От фигура 6 и таблица 4 се вижда, че стойностите на средното, медианата и модата за четирите теста са близо и между тях съществуват неравенствата $M_o < M_e < \bar{X}$, което означава, че разпределенията им са ляво асиметрични.

На фигура 7 е представена хистограма според броя кандидати положили изпитите и полигонът за средния получен бал от тях.

3. АНАЛИЗ НА КАЧЕСТВАТА НА ЗАДАЧИТЕ ОТ ТЕСТА

1. Трудност

В таблица 5 и таблица 6 са пресметнати коефициентите на трудност (виж [2], [4],[5]).

Таблица 5

№ на темата	Коефициент на трудност на задачите спрямо всички изпитвани			
	1 тест	2 тест	3 тест	4 тест
1	48,2%	47,6%	35,9%	38,8%
2	45,7%	48,5%	48,7%	48,5%
3	41,7%	48,9%	38,2%	36,8%
4	4,2%	9,2%	10,0%	6,4%
5	24,9%	19,3%	21,6%	8,4%
6	26,3%	24,2%	31,4%	22,4%
Общо:	36,2%	38,8%	35,7%	33,0%

Таблица 6

№ на темата	Коефициент на трудност на задачите спрямо отговорилите			
	1 тест	2 тест	3 тест	4 тест
1	57,2%	60,8%	44,3%	49,0%
2	49,4%	55,8%	51,6%	54,1%
3	63,6%	63,0%	45,9%	47,0%
4	4,2%	9,2%	10,0%	6,4%
5	33,8%	27,8%	42,0%	11,9%
6	30,0%	27,3%	41,5%	27,0%
Общо:	42,8%	47,0%	43,1%	40,2%

Можем да отбележим, че задачите от четвърта тема „Стереометрия” са много трудни за кандидатите във всички години (тестове). Това говори, че в учебните програми на средния курс трябва да се отдели повече часове за по-доброто обучение по тази тема или да се дават задачи с по-малка

трудност в тестовете по тази тема, за повишаване разграничителната способност на теста.

2. Успеваемост

В таблица 7 е даден средният бал получен от кандидатите за една задача от тема получен в различните тестове.

Таблица 7

№ на темата	Среден бал за задачата			
	1 тест	2 тест	3 тест	4 тест
1	1,70	1,73	1,39	1,51
2	1,65	1,89	1,79	1,65
3	1,60	1,85	1,67	1,55
4	0,25	0,55	0,60	0,38
5	1,06	0,97	1,13	0,57
6	1,27	1,19	1,33	0,92
Общо:	1,40	1,54	1,45	1,31

В таблица 8 е даден средният бал (средният брой получени точки от кандидатите) по темите получени за различните тестове.

Таблица 8

№ на темата	Среден бал за темата			
	1 тест	2 тест	3 тест	4 тест
1	8,49	12,08	11,09	13,60
2	14,89	13,23	12,53	8,24
3	6,38	11,08	8,33	10,88
4	0,50	1,10	1,20	0,77
5	4,22	2,90	2,27	1,14
6	7,61	5,95	8,00	4,62
Общо:	42,10	46,34	43,41	39,24

В таблица 9 е даден процентът на средния получен бал спрямо максимално възможния бал по темите получен за различните тестове.

Таблица 9

№ на темата	Процент на средния бал за темата спрямо максималният			
	1 тест	2 тест	3 тест	4 тест
1	47,17%	50,33%	36,95%	41,20%
2	41,37%	49,00%	41,77%	45,76%
3	53,18%	52,74%	39,64%	40,30%
4	4,20%	9,16%	10,00%	6,38%
5	28,14%	24,20%	37,78%	12,68%
6	28,20%	24,81%	38,08%	22,01%
Общо:	35,09%	38,62%	36,17%	32,70%

3. Корелация на даден тест с всички останали тестове

♦ Корелация по задачи

В таблица 10 е дадена общата матрица на корелация между тестовете, като корелацията е смятана спрямо всички задачи от тях. (виж [1], [4],[5]).

Таблица 10

	1	2	3	4
1	1,000	0,416	0,417	0,533
2	0,416	1,000	0,287	0,481
3	0,417	0,287	1,000	-0,053
4	0,533	0,481	-0,053	1,000

♦ Корелация по теми

В таблица 11 е дадена общата матрица на корелация между тестовете, като корелацията е смятана по определените 6 теми. (виж [1], [4],[5]).

Таблица 11

	1	2	3	4
1	1,000	0,950	0,937	0,913
2	0,950	1,000	0,947	0,990
3	0,937	0,947	1,000	0,908
4	0,913	0,990	0,908	1,000

Получената ниска корелация в таблица 10 се обяснява с разместване на местата на задачите от различните теми в тестовете. От таблица 11 се вижда висока корелация между тестовете по задачите групирани в теми.

4. Дисперсионен анализ

Целта е да се установи дали различните тестове водят до различни резултати.

Нулевата хипотеза гласи, че всички средни са равни, т.е. средният резултат на кандидати при различните тестове би бил еднакъв.

Алтернативната хипотеза е, че резултатът от различните тестове би бил различен.

Определяме равнище на значимост $\alpha = 0,01$.

Пресмятаме:

Междугрупова дисперсия:

$$(1) \quad \sigma_M^2 = \frac{n \sum_{j=1}^k (\bar{X}_j - \bar{X}_0)^2}{k-1} \approx 1,45$$

Вътрешногрупова дисперсия:

$$(2) \quad \sigma_B^2 = \frac{\sum_{j=1}^k \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2}{n-k} \approx 24,47$$

Изчислява се отношението:

$$(3) \quad F = \frac{\sigma_M^2}{\sigma_B^2} = 0,059$$

По таблицата на Стюдънт се установява F_T . При 3 и 20 степени на свобода и при $P(t) = 0,99$, $F_T = 4,94$

Съпоставят се емпиричната с табличната стойност за F . Имаме $F < F_T$, т.е. с риск за грешка 1% може да се твърди, че различните тестове водят до еднакви резултати.

4. ИЗВОДИ

Целта на тестовете е да се определи група от кандидати, които имат добри математически познания и владеят работата с математически справочник и те да се разграничат от кандидатите с по ниски действителни способности. Друга цел на тестовете е да се ограничи възможността от много на брой отлични оценки.

Проведените корелационен и дисперсионен анализ показват, че различните тестове измерват едно и също количество знание от кандидатите, като отделят еднакви групи от кандидати за прием във ВТУ „Тодор

Каблешков”. Тестовете са подходящо избрани и осъществяват първоначалните си цели.

Всички тестове имат ляво асиметрично разпределение, което ограничава получаването на много на брой високи оценки.

Забележка: Тези изводи са направени при предположение че кандидатите се явяват по едно и също време на изпит, без да отчитаме различната степен на подготвеност на кандидатите по години, което е свързано и с различни учебни програми .

ЛИТЕРАТУРА:

- [1] Калоянов Т., Статистика, Издателство „Тракия-М”, София, 2004 г.
- [2] Стоименова Е., Измерителни качества на тестовете, София, 2000 г.
- [3] Улучев Р., Михалев Д., Приложна математика, ВТУ „Тодор Каблешков”, София, 2008 г.
- [4] Алашка Р., Михалев Д., Основни числови характеристики на реален изпитен тест, София, 2009 г.
- [5] Алашка Р., Статистически анализ за оценка качеството на изпитен тест, София, 2009 г.
- [6] Конкурсни тестове по математика – 2008 година, 2009 година, София, - http://vtu.bg/i_admission_b.html .

COMPARATIVE ANALYSIS BETWEEN ACTUAL EXAMINATION TESTS

Drago Michalev, Rayna Alashka
michalev@abv.bg, alraina@abv.bg

HST “Todor Kableshkov”, 158 Geo Milev Street, 1574 Sofia
BULGARIA

Key words: test, examination, correlation analysis, dispersion analysis, statistics, education.

Abstract: A comparative analysis between the four real trials test is made. The correlation between them and tasks and topics are calculated. Dispersion analysis of the importance of the resulting examination marks of the tests is made.