

УДК 378:51-519.2

ББК 22.1

Б 79

Болтовский Д.В.

Кандидат физико-математических наук, доцент кафедры высшей математики и математической физики Национального исследовательского Томского политехнического университета, Томск, e-mail: bdv1@tpu.ru

Дисперсионный анализ показателей образовательного мониторинга в вузе (Рецензирована)

Аннотация. Проведен дисперсионный анализ динамики результатов первого семестра по высшей математике студентов очной формы обучения энергетического института Томского политехнического университета. Рассмотрение проведено в системе 3-х показателей: ВК – результаты тестового входного контроля по математике; АТТ – результаты текущей аттестации по высшей математике; ЭКЗ – результат классического экзамена. В рамках параметрического дисперсионного анализа повторных измерений оценена динамика успеваемости на институтском уровне и сделан вывод о высоко значимом различии (уровень значимости $p < 0,0005$) между ВК и ЭКЗ, а также между ВК и АТТ. При этом различие между ЭКЗ и АТТ – сильно значимое ($0,0005 < p < 0,005$). Применение непараметрического (рангового) дисперсионного анализа Фридмана с повторными измерениями подтвердило результаты параметрического дисперсионного анализа повторных измерений. На основании дисперсионного анализа для независимых выборок сделан вывод о сильно значимой неоднородности результатов АТТ и слабо статистически значимой неоднородности ($p \approx 0,05$) результатов ВК, ЭКЗ по совокупности групп. Оформлены однородные (различающиеся незначимо) кластеры групповых средних для каждого показателя. Непараметрические альтернативы однофакторного дисперсионного анализа для независимых выборок (критерий Краскела-Уоллиса и медианный тест) подтверждают сделанный на основе F-критерия вывод об уровне значимости неоднородности АТТ и смягчают значимость неоднородности ВК и ЭКЗ до слабой ($0,05 < p < 0,10$). По результатам образовательного мониторинга сформулированы предложения по повышению качества успеваемости студентов. Результаты проведенного статистического анализа могут быть учтены в рамках проходящей реформы высшего образования.

Ключевые слова: дисперсионный анализ, формы контроля успеваемости студентов, образовательный мониторинг.

Boltovskiy D.V.

Candidate of Physics and Mathematics, Associate Professor of Department of Higher Mathematics and Mathematical Physics, Tomsk National Research Polytechnic University, Tomsk, e-mail: bdv1@tpu.ru

Analysis of variance of indicators of the educational monitoring at higher school

Abstract. An analysis is made of variance of dynamics of outcomes of the first semester on higher mathematics of students of a full-time course of study at Power Institute of Tomsk Polytechnic University. Reviewing is made in system of 3 indicators: EC – outcomes of test entering control on the mathematics; CERT – outcomes of current certification for higher mathematics; EX – outcome of classical examination. Within the limits of a parametrical analysis of variance of repeated measurements, the dynamics of progress at institute level is estimated and the conclusion is drawn on highly significant distinction (a significance level $p < 0,0005$) between EC and EX, and also between EC and CERT. Thus distinction between EX and CERT is strongly significant ($0,0005 < p < 0,005$). Application of a non-parametric (rank) analysis of variance of Friedman with repeated measurements has confirmed outcomes of a parametrical analysis of variance of repeated measurements. On the basis of an analysis of variance for independent samples, the conclusion is drawn on strongly significant heterogeneity of CERT and on weak-statistically significant heterogeneity ($p \approx 0,05$) of EC and EX on a population of groups. Homogeneous (insignificantly differing) clusters of group averages for each indicator are defined. Non-parametric alternatives of one-way analysis of variance for independent statistical samples (the Kruskal-Wallis criterion and the median test) confirm the conclusion drawn on the basis of F-criterion about a significance level of CERT heterogeneity and soften the importance of EC and EX heterogeneity to weak ($0,05 < p < 0,10$). On the basis of results of educational monitoring, proposals on improvement of quality of students' progress are formulated. Outcomes of the carried out statistical analysis can be considered within the limits of passing reform of higher education.

Keywords: analysis of variance, forms of control of students' progress, educational monitoring.

Главной проблемой современного инженерного образования в условиях модернизации высшей школы является проблема повышения качества образования и оценивания качества обучения в вузах [1–3].

В связи с введением федеральных государственных образовательных стандартов (ФГОС) высшего профессионального образования (ВПО) третьего поколения весьма актуальным является вопрос о методах оценки результатов обучения студентов, о педагогиче-

ском анализе результатов обучения студентов на различных этапах обучения в рамках компетентностного подхода, об управлении качеством образования в вузе. Анализ содержания ФГОС ВПО позволяет сделать вывод, что приоритетной задачей управления качеством образования на современном этапе выступает оценка качества освоения основных образовательных программ, включающая предварительный контроль (учет итогового рейтинга обучающихся на предшествующем этапе обучения), текущий контроль успеваемости в форме контрольных работ и промежуточных аттестаций обучающихся и итоговый контроль в форме рейтинговой системы оценок с дальнейшим переводом их в экзаменационную (зачетную) оценку. Контроль качества образования приобретает характер мониторинга [1, 3, 4] или непрерывного контроля, постоянного отслеживания результатов образования, хода образовательного процесса, необходимого для систематической корректировки мероприятий по их реализации. Таким образом, мониторинг рассматривается не просто как инструмент оценки качества ВПО, а средство управления этим качеством, делающее управление образовательным процессом эффективным. Процесс мониторинга предполагает проведение измерения показателей образовательной деятельности студентов в соответствии с ФГОС ВПО. В нашем случае это измерение (оценивание) результатов образовательной деятельности студентов:

ВК – результаты входного тестирования по математике в начале текущего семестра;

АТТ1 – результаты 1-й аттестации по математике в середине текущего семестра;

АТТ2 – результаты 2-й аттестации по математике в конце текущего семестра;

ЭКЗ – результат классического экзамена в конце текущего семестра;

ЭКЗ* – скорректированные результаты ЭКЗ в середине следующего семестра.

Все числовые результаты приведены к единой 5-балльной шкале (делением результата на соответствующий максимальный результат и умножением на пять). Созданная таким образом в MS Excel база данных использовалась далее в пакете Statistica [5, 6] для статистического анализа данных.

В работах [4, 7–9] проведен компьютерный статистический сравнительный анализ результатов различных форм контроля обучения студентов вузов. В связи с этим представляет интерес дисперсионный анализ динамики результатов изучения высшей математики студентами очной формы обучения на примере 7-ми учебных групп (выборка объема $n=157$) энергетического института (ЭНИН) Томского политехнического университета, имевших более 25% должников по результатам ЭКЗ по высшей математике в объеме 1-го семестра (линейная алгебра и аналитическая геометрия + дифференциальное исчисление).

Для корректного применения F -критерия параметрического дисперсионного анализа необходимо предварительно оценить сходства наблюдаемых распределений (гистограмм), например, ВК, ЭКЗ и ЭКЗ–ВК (рис. 1) с теоретическим распределением по нормальному закону (соответствующие кривые).

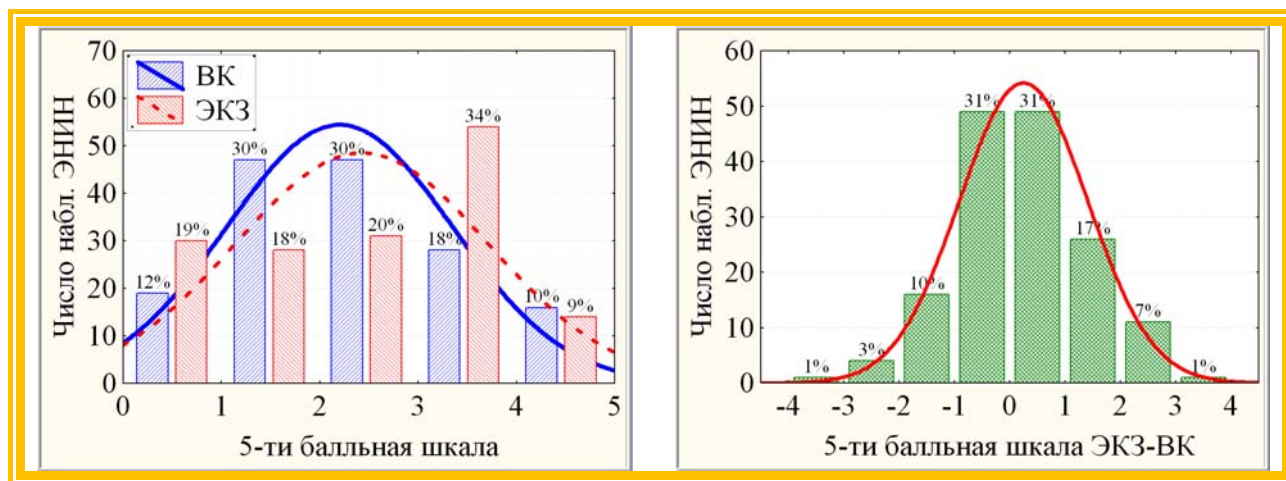


Рис. 1. Составная гистограмма ВК и ЭКЗ (слева) и ЭКЗ–ВК (справа) ЭНИН с соответствующими кривыми нормальных распределений

Проверка нормальности распределений выборок с помощью χ^2 -критерия Пирсона дает незначимое (уровень значимости $p > 0,10$) для ВК и ЭКЗ–ВК, но высоко значимое ($p < 0,0005$) для ЭКЗ отличие наблюдаемых распределений от нормального закона. В связи с нарушением в разной степени условия нормальности распределения выборок далее применялся также и непараметрический дисперсионный анализ.

Заметим, что, согласно теории измерительных шкал, балльная шкала относится к типу порядковых шкал, позволяющих ранжировать (упорядочить) результаты оценивания качества усвоения знаний студентами. Поэтому в балльной шкале оперирование средним баллом является некорректным. В связи с этим для сравнения рассматриваемых выборок предлагается использовать наряду с параметрическими также и ранговые (непараметрические) критерии, основанные на рангах, а не на средних значениях.

Совместное распределение ВК и ЭКЗ (рис. 1) может быть использовано также для визуальной оценки их различия по 5-балльной равномерной шкале: положительная динамика ($\text{ЭКЗ} > \text{ВК}$) в диапазоне (3, 4] и отрицательная – в (1, 3], а также общая положительная динамика (результаты $\text{ЭКЗ} - \text{ВК} > 0$ составляют 56%). При этом на уровне средних выборок $m_{\text{ЭКЗ}} - m_{\text{ВК}} \approx 2,416 - 2,191 \approx 0,225 = m_{\text{ЭКЗ} - \text{ВК}}$ по 5-балльной шкале. Отметим существенный разброс значений разности результатов ЭКЗ–ВК, что свидетельствует об отсутствии значимой корреляции между ВК и ЭКЗ.

Динамика успеваемости ЭНИН в переменных {ВК, АТТ1, АТТ2, ЭКЗ, ЭКЗ*} приведена на рисунке 2.

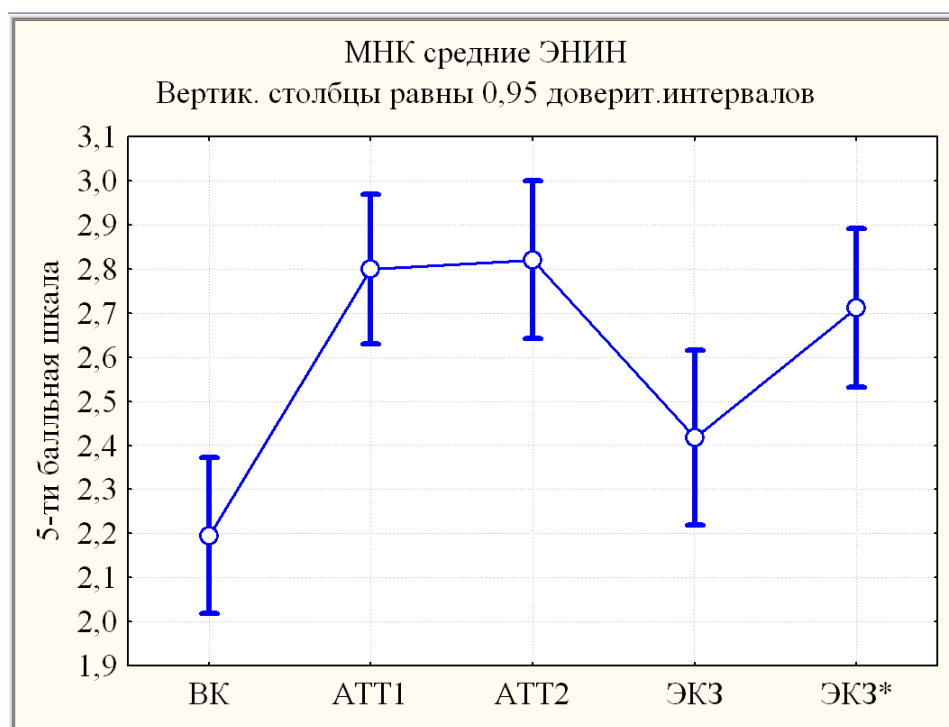


Рис. 2. Динамика успеваемости ЭНИН на уровне средних

Для оценки значимости динамики успеваемости ЭНИН применен прежде всего (с учетом проверки выборки поэлементной разности ЭКЗ–ВК на нормальность распределения) параметрический дисперсионный анализ повторных измерений. Параметрический F -критерий для зависимых выборок приводит к выводу о высоко значимой ($p < 0,0005$) динамике успеваемости ЭНИН по совокупности показателей {ВК, АТТ1, АТТ2, ЭКЗ, ЭКЗ*}. Применение непараметрического (рангового) дисперсионного анализа Фридмана с повторными измерениями подтвердило результаты параметрического дисперсионного анализа повторных измерений. В рамках параметрического дисперсионного анализа на основании критериев множественного сравнения сделан вывод о незначимом различии (уровень значимости $p > 0,10$) между АТТ1 ($m_{\text{АТТ1}} \approx 2,781$), АТТ2 ($m_{\text{АТТ2}} \approx 2,814$) и ЭКЗ* ($m_{\text{ЭКЗ*}} \approx 2,710$). При этом различие ме-

жду ВК ($m_{ВК} \approx 2,191$) и ЭКЗ ($m_{ЭКЗ} \approx 2,416$) оценивается как высоко значимое ($p < 0,0005$), а между ЭКЗ и ЭКЗ* – сильно значимое ($0,0005 < p \approx 0,001 < 0,005$). Полученные результаты подтверждаются ранговым критерием Вилкоксона для зависимых выборок.

Более детальная гистограммная динамика ЭНИН в переменных {ВК, АТТ1, АТТ2, ЭКЗ, ЭКЗ*} приведена на рисунке 3.

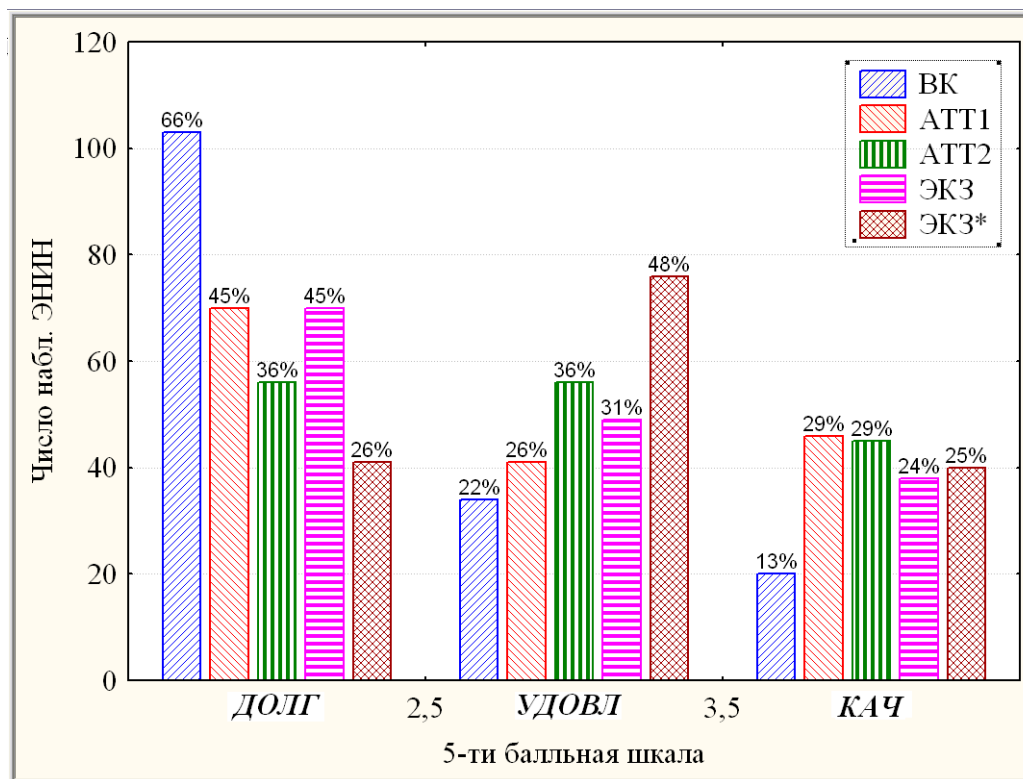


Рис. 3. Составная гистограммная динамика успеваемости ЭНИН

Согласно рисунку 3 наблюдается положительная текущая динамика успеваемости {ВК, АТТ1, АТТ2} – убывание ДОЛГ и рост как УДОВ, так и КАЧ, а также положительная динамика периода после экзаменационной ликвидации задолженностей {ЭКЗ, ЭКЗ*} – убывание ДОЛГ и рост УДОВ. Отрицательная динамика успеваемости экзаменационного периода {АТТ2, ЭКЗ}, проявившаяся в росте ДОЛГ и убывании УДОВ и КАЧ, объясняется естественной сложностью итоговой (ЭКЗ) контрольной для слабых студентов.

Выборки {ВК, АТТ1, АТТ2, ЭКЗ, ЭКЗ*} разных форм контроля успеваемости ЭНИН являются составными, состоящими из 7-ми учебных групп.

С одной стороны, аналогично динамике успеваемости ЭНИН на институтском уровне (рис. 2) можно оценить групповую динамику успеваемости ЭНИН (рис. 4). Групповая динамика успеваемости ЭНИН является очень пестрой:

- Б2 – сильно значимая положительная начальная динамика {ВК, АТТ} с последующим стабильным уровнем;
- Б1 – слабо значимая положительная начальная динамика {АТТ1, АТТ2} с последующим стабильным уровнем;
- Г3 – сильно значимая положительная начальная динамика {ВК, АТТ1} с последующей отрицательной {АТТ1, АТТ2, ЭКЗ};
- Г5 – незначимая начальная динамика {ВК, АТТ1} с последующей отрицательной {АТТ, ЭКЗ};
- А3, Г4 – значимая переменная динамика;
- А6 – незначимая динамика.

С другой стороны, применение F -критерия дисперсионного анализа для независимых выборок приводит к выводу о высоко значимой неоднородности (на уровне значимости

$p < 0,0005$) результатов АТТ1, сильно значимой неоднородности (на уровне значимости $0,0005 < p < 0,005$) результатов АТТ2 и слабо статистически значимой неоднородности (на уровне значимости $p \approx 0,05$) результатов ВК, ЭКЗ, ЭКЗ* по совокупности групп. Заметим, что непараметрические альтернативы однофакторного дисперсионного анализа для независимых выборок (критерий Краскела-Уоллиса и медианный тест) подтверждают сделанный на основе F -критерия вывод об уровне значимости неоднородности АТТ1, АТТ2 и смягчают значимость неоднородности ВК, ЭКЗ, ЭКЗ* до слабой ($0,05 < p < 0,10$) по совокупности групп. Уровень значимости различий между групповыми средними можно оценить с помощью апостериорного критерия наименьших значений разности, согласно которому можно оформить однородные (различающиеся незначимо) кластеры групповых средних в порядке их убывания для каждого показателя (рис. 4):

ВК: {Г5, Б1, А6, Б2, Г3}, {А6, Б2, Г3, А3, Г4} так, что Г5 от А3 или Б1 от Г4 отличаются статистически значимо (на уровне значимости $0,005 < p < 0,05$);

АТТ1: {Г3}, {Б2, Г5, Б1, А3}, {Б1, А3, Г4, А6} так, что Г3 от Б2 или Б2 от Г4 отличаются статистически значимо (на уровне значимости $0,005 < p < 0,05$);

АТТ2: {Б2, Б1, Г3}, {Б1, Г3, Г5}, {Г5, А6, А3, Г4} так, что Б2 от Г5 или Г3 от Г4 отличаются статистически значимо (на уровне значимости $0,005 < p < 0,05$);

ЭКЗ: {Б2, Б1, Г3, Г5}, {Б1, Г3, Г5, А6}, {Г3, Г5, А6, Г4, А3}, так, что Б2 от А6 или А3 от Б1 отличаются статистически значимо (на уровне значимости $0,005 < p < 0,05$);

ЭКЗ*: {Б2, Б1, Г3, Г5, Г4, А6}, {Г3, Г5, Г4, А6, А3}, так, что Б1 от А3 отличаются статистически значимо (на уровне значимости $0,005 < p < 0,05$).

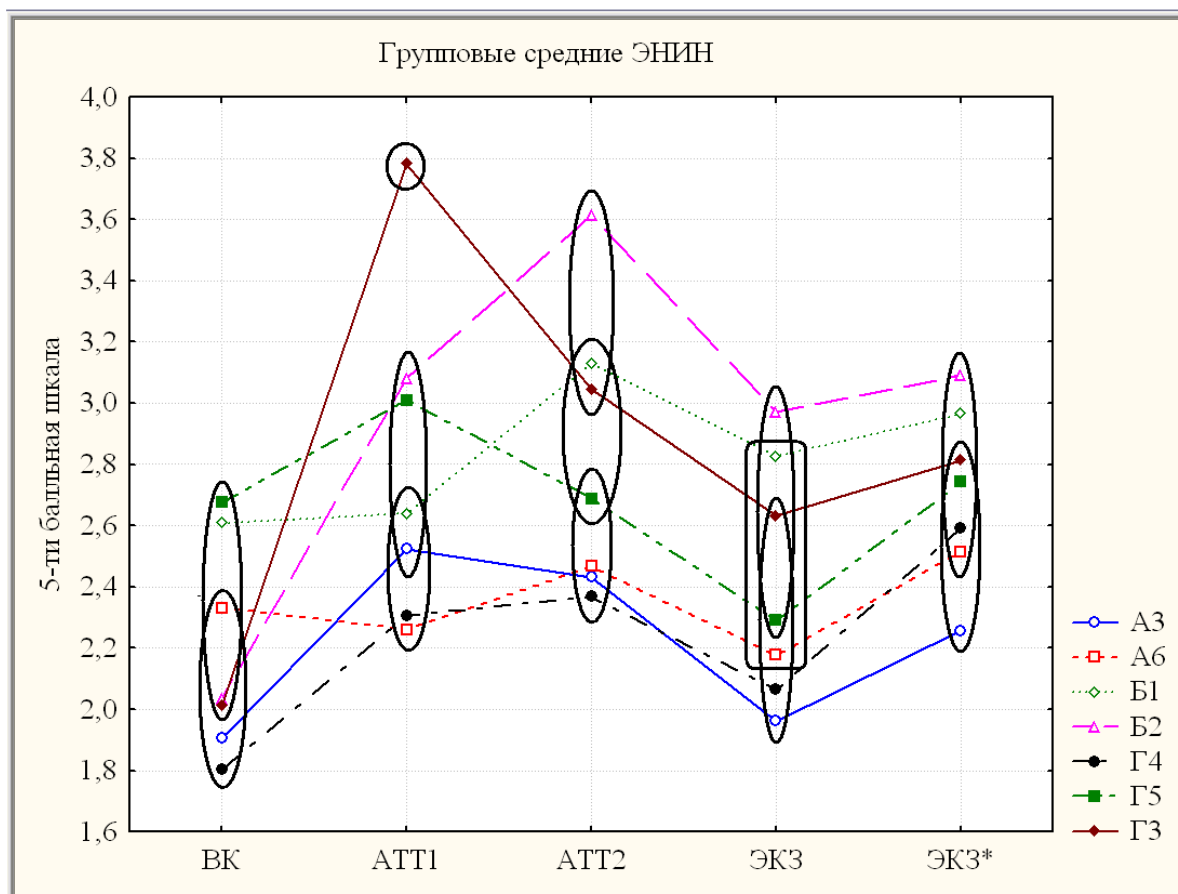


Рис. 4. Групповая динамика успеваемости ЭНИН

По результатам образовательного мониторинга ЭНИН можно сформулировать следующие предложения по повышению качества успеваемости студентов:

- усиление работы приемной комиссии по повышению качества набора в силу низких результатов ВК: 66% набора имеют «неуд» (рис. 3);

- увеличение аудиторных часов на изучение сложной программы высшей математики в силу слабо «удовл» результатов текущего контроля (рис. 2–4);
- введение составной итоговой отчетности по высшей математике в объеме 1-го семестра в связи с составным характером семестровой программ: линейная алгебра и аналитическая геометрия (зачет) + дифференциальное исчисление (экз.) для ликвидации разрыва АТТ2–ЭКЗ (рис. 2–4);
- продление практики периода после экзаменационной ликвидации задолженностей {ЭКЗ, ЭКЗ*}, дающей возможность доосвоить сложную программу высшей математики и имеющей положительную динамику (рис. 2–4).

Выводы

1. В рамках параметрического дисперсионного анализа повторных измерений оценена динамика успеваемости ЭНИН на институтском уровне: сделан вывод о незначимом различии (уровень значимости $p > 0,10$) между АТТ1 ($m_{\text{АТТ1}} \approx 2,781$), АТТ2 ($m_{\text{АТТ2}} \approx 2,814$) и ЭКЗ* ($m_{\text{ЭКЗ*}} \approx 2,710$). При этом различие между ВК ($m_{\text{ВК}} \approx 2,191$) и ЭКЗ ($m_{\text{ЭКЗ}} \approx 2,416$) оценивается как высоко значимое ($p < 0,0005$), а между ЭКЗ и ЭКЗ* – сильно значимое ($0,0005 < p \approx 0,001 < 0,005$).

2. Применение непараметрического (рангового) дисперсионного анализа Фридмана с повторными измерениями подтвердило результаты параметрического дисперсионного анализа повторных измерений.

3. На основании дисперсионного анализа для независимых выборок сделан вывод о высоко значимой неоднородности (на уровне значимости $p < 0,0005$) результатов АТТ1, сильно значимой неоднородности (на уровне значимости $0,0005 < p < 0,005$) результатов АТТ2 и слабо статистически значимой неоднородности (на уровне значимости $p \approx 0,05$) результатов ВК, ЭКЗ, ЭКЗ* по совокупности групп. Оформлены однородные (различающиеся незначимо) кластеры групповых средних в порядке их убывания для каждого показателя.

4. Непараметрические альтернативы однофакторного дисперсионного анализа для независимых выборок (критерий Краскела-Уоллиса и медианный тест) подтверждают сделанный на основе F -критерия вывод об уровне значимости неоднородности АТТ1, АТТ2 и смягчают значимость неоднородности ВК, ЭКЗ, ЭКЗ* до слабой ($0,05 < p < 0,10$) по совокупности групп.

5. По результатам образовательного мониторинга ЭНИН сформулированы предложения по повышению качества успеваемости студентов.

Примечания:

1. Разработка и организация системы мониторинга качества образовательной деятельности студентов в высшем учебном заведении / С.Ф. Багаутдинова, Н.И. Левшина, Л.Н. Санникова В.И. Турченко // Фундаментальные исследования. 2014. № 1-0. С. 109–114.
2. Богачева А.Г. Проблемы управления качеством образования в вузе // Фундаментальные и прикладные исследования: проблемы и результаты. 2015. № 17. С. 103–107.
3. Внутривузовский мониторинг формирования компетенций у студентов / В.В. Кошкин, А.С. Масленников, Л.А. Стешина, Н.Н. Старыгина // Alma mater (Вестник высшей школы). 2015. № 2. С. 77–80.
4. Мониторинг качества обучения студентов физико-технических направлений подготовки / Л.И. Терехина, И.А. Цехановский, В.Ф. Зальмеж, В.Н. Задорожный // Современные проблемы науки и образования. 2016. № 2. URL: <http://www.science-education.ru/ru/article/view?id=24277> (дата обращения: 09.06.2016)

References:

1. Development and organization of the monitoring system of the quality of educational activity of students at higher schools / S.F. Bagautdinova, N.I. Levshina, L.N. Sannikova, V.I. Turchenko // Fundamental Research. 2014. No. 1-0. P. 109–114.
2. Bogacheva A.G. Problems of quality control of education at higher school // Fundamental and Applied Research: Problems and Results. 2015. No. 17. P. 103–107.
3. Inter-university monitoring of formation of competencies of students / V.V. Koshkin, A.S. Maslennikov, L.A. Steshina, N.N. Starygina // Alma mater (Journal of the Higher School). 2015. No. 2. P. 77–80.
4. Monitoring of the quality of teaching students of physical and technical training areas / L.I. Terekhina, I.A. Tsekhanovskiy, V.F. Zalmez, V.N. Zadorozhny // Modern Problems of Science and Education. 2016. No. 2. URL: <http://www.science-education.ru/ru/article/view?id=24277> (date of the reference: 09/06/2016)

5. Боровиков В.П. STATISTICA. Искусство анализа данных на компьютере: для профессионалов. СПб.: Питер, 2003. 688 с.
6. Халафян А.А. STATISTICA 6. Статистический анализ данных: учебник. М.: Бином-Пресс, 2008. 512 с.
7. Болтовский Д.В. Факторный подход кластеризации результатов оценивания математических знаний в системе высшего образования // Современные проблемы науки и образования. 2014. № 3. URL: www.science-education.ru/117-13284 (дата обращения: 09.06.2016)
8. Терехина Л.И. Применение дисперсионного анализа для оценки значимости динамики показателей успеваемости в ВУЗе // Современные проблемы науки и образования. 2014. № 3. URL: www.science-education.ru/117-13282 (дата обращения: 24.04.2015)
9. Терехина Л.И. Применение факторного подхода кластеризации результатов мониторинговой оценки знаний по математике в вузе // Современные проблемы науки и образования. 2015. № 1-1. URL: <http://www.science-education.ru/ru/article/view?id=18605> (дата обращения: 23.02.2016)
5. Borovikov V.P. STATISTICA. Art of the data analysis on computer: for professionals. SPb.: Piter, 2003. 688 pp.
6. Khalafyan A.A. STATISTICA 6. Statistical analysis of data: a textbook. M.: Binom-Press, 2008. 512 pp.
7. Boltovskiy D.V. Factor approach to clustering of the results of assessment of mathematical knowledge in the system of higher education // Modern Problems of Science and Education. 2014. No. 3. URL: www.science-education.ru/117-13284 (date of reference 09/06/2016)
8. Terekhina L.I. The use of dispersion analysis for the assessment of the significance of the dynamics of academic performance at higher school // Modern Problems of Science and Education. 2014. No. 3. URL: www.science-education.ru/117-13282 (date of access: 04/24/2015)
9. Terekhina L.I. The use of the factor approach to clustering of the results of assessment of mathematical knowledge at higher school // Modern Problems of Science and Education. 2015. No. 1-1. URL: <http://www.science-education.ru/ru/article/view?id=18605> (date of access: 23/02/2016)