

Оценка корреляций выполнения заданий ЕГЭ по информатике в зависимости от формата их представления

**Левинская
Мария Александровна**

кандидат физико-математических наук;
ведущий научный сотрудник ФГБНУ «ФИПИ»,
доцент факультета прикладной математики
и физики, кафедры вычислительной математики
и программирования ФГБОУ «Московский
авиационный институт» (национальный
исследовательский университет)
levinskaja@fipi.ru

Ключевые слова: корреляция, компьютерный формат представления КИМ, практико-ориентированные задания на программирование.

В 2021 году в КИМ ЕГЭ по информатике и ИКТ были включены задания на практическое программирование (составление, отладка программы в выбранной участником среде программирования), работу с электронными таблицами и информационный поиск. Для этих заданий необходимо было использовать компьютер. Остальные задания (две трети КИМ) сохранили преемственность с материалами прошлых лет (экзамена в бланковой форме). Выполнение заданий по программированию допускалось на языках (семействах языков) C++, Java, C#, Pascal, Python, Школьный алгоритмический язык¹.

Рассмотрим изменения в статистических характеристиках КИМ ЕГЭ 2021 г. (по информатике и ИКТ)² в сравнении с предыдущим годом и особенности выполнения заданий по различным группам участников экзамена в зависимости от формата представления и практико-ориентированности задания. В статье исследованы корреляции между выполнением частей КИМ в различных регионах, группами участников с различным уровнем подготовки, группами с различной гендерной принадлежностью (юноши и девушки).

Сравнивая КИМ ЕГЭ по информатике и ИКТ по годам, следует отметить, что выборка 2020 г. меньше примерно на 10%, чем выборка 2021 г. (соответственно, 84,5 и 93,5 тысячи участников экзамена), однако по проценту выполнения КИМ (полностью и по преемственным заданиям в отдельности) учащиеся показывают приблизительно одинаковые результаты. Это позволяет сравнивать результаты ЕГЭ по информатике и ИКТ 2020 и 2021 гг. в предположении выборок одинаковой силы.

На рисунках представлены диаграммы баллов 2020 г. (рис. 1) и 2021 г. ЕГЭ в компьютерной форме (рис. 2).

На диаграмме баллов 2020 г. можно наблюдать сложение двух распределений (части 1 и части 2 с развёрнутым ответом, в которой необходимо было описать стратегию, написать/исправить программу или фрагменты кода).

¹ Методические рекомендации для учителей, подготовленные на основе анализа типичных ошибок участников ЕГЭ 2021 года по информатике и ИКТ / ФГБНУ «ФИПИ». — URL: <https://fipi.ru/ege/analiticheskie-i-metodicheskie-materialy#1/tab/173737686-5> (дата обращения 24.04.2022).

² Демоверсии, спецификации, кодификаторы ЕГЭ 2021 год / ФГБНУ «ФИПИ». — URL: <https://fipi.ru/ege/demoversii-specifikacii-kodifikatory#1/tab/151883967-5> (дата обращения 18.05.2022).

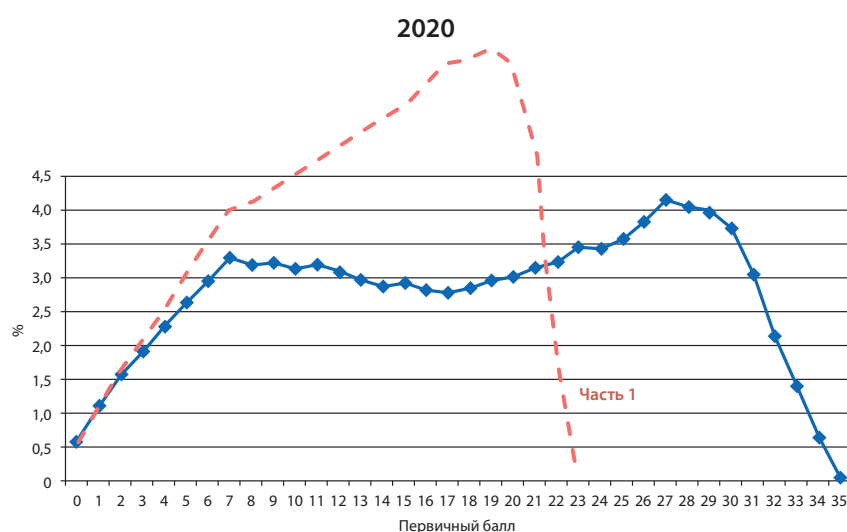


Рис. 1. Диаграмма баллов ЕГЭ по информатике и ИКТ 2020 г.

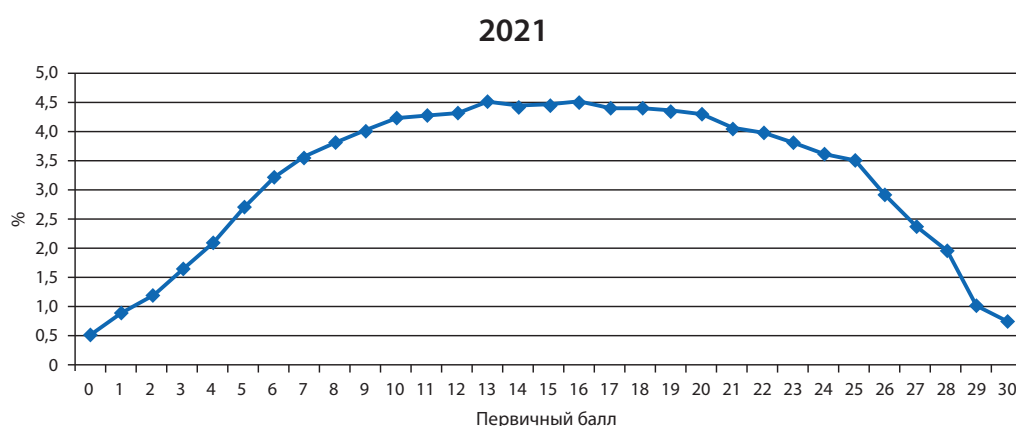


Рис. 2. Диаграмма баллов ЕГЭ по информатике и ИКТ 2021 г.

В итоговом распределении они дали два локальных максимума. Распределение баллов 2021 г. приближено к нормальному, и это говорит о том, что компьютерный формат экзамена обеспечил более равномерное распределение заданий по сложности.

В таблице 1 даны сравнительные статистические характеристики КИМ в 2020 и 2021 гг.³

В характеристиках КИМ 2020 г. можно отметить значимую отрицательную асимметрию (смещение в области высоких баллов). В распределении баллов 2021 г. асимметрия близка к нулю, что говорит об отсутствии смещения в сторону низких или высоких баллов. Коэффициенты надёжности теста при изменении формы представления заданий остались равными и на высоком уровне ($>0,9$).

Таблица 1

Статистические характеристики КИМ 2020 и 2021 гг.

	Кол-во учащихся	Сравнительный % выполнения	Асимметрия	Коэффициент надёжности (Кронбаха)
2020	84 477	51	-0.12	0,91
2021	93 541	52	-0.03	0,91

³ Пугачев В. С. Теория вероятностей и математическая статистика: учеб. пособие. — 2-е изд., исправл. и доп. — М.: Физматлит, 2002. — 496 с.

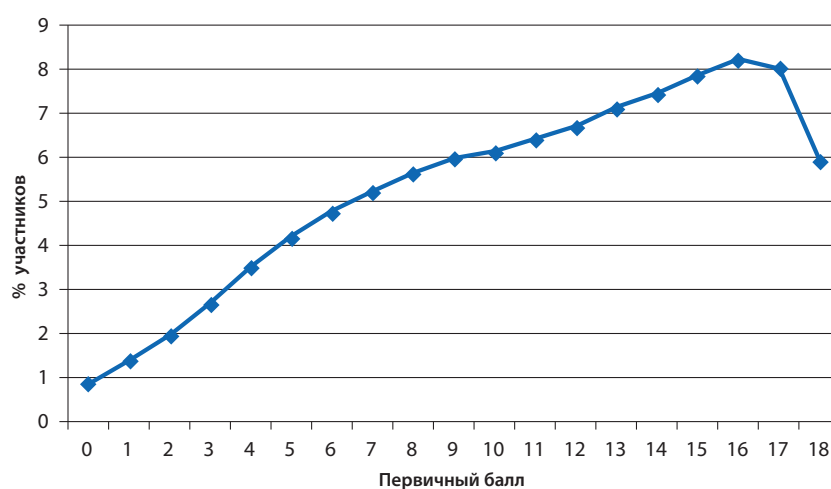


Рис. 3. Диаграмма баллов за часть КИМ без специализированного ПО

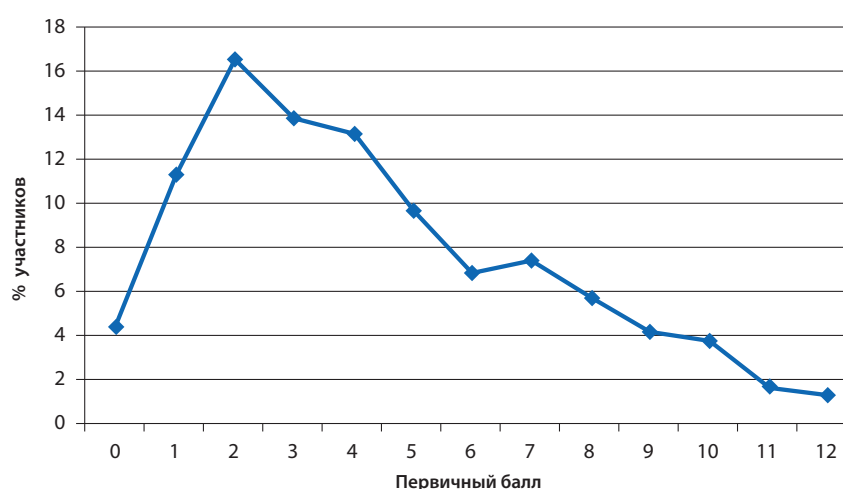


Рис. 4. Диаграмма баллов за часть КИМ, требующую специализированного ПО

Рассмотрим выполнение КИМ 2021 г. по частям (часть 1 — практико-ориентированные задания, требующие специализированного программного обеспечения (ПО), и часть 2 — остальные задания). На рисунке 3 представлена диаграмма баллов за часть КИМ, включающую в себя задачи, не требующие специализированного ПО. На рисунке 4 представлена диаграмма баллов за часть КИМ, включающую в себя практико-ориентированные задания (требующие специализированного ПО).

В таблице 2 приведено сравнение характеристик этих частей экзаменационной работы.

В приведённых статистических характеристиках и на графиках можно отметить отрицательную асимметрию (смещение в области высоких баллов) части КИМ с заданиями, в которой не требуется специализированное ПО (средний процент выполнения ~62%) и значимую положительную асимметрию (смещение в области

Таблица 2

Статистические характеристики частей КИМ

	Средний % выполнения	Асимметрия	Корреляция балла между частями КИМ
Не требуется ПО	62	–0,38	0,8
Требуется ПО	36	0,64	

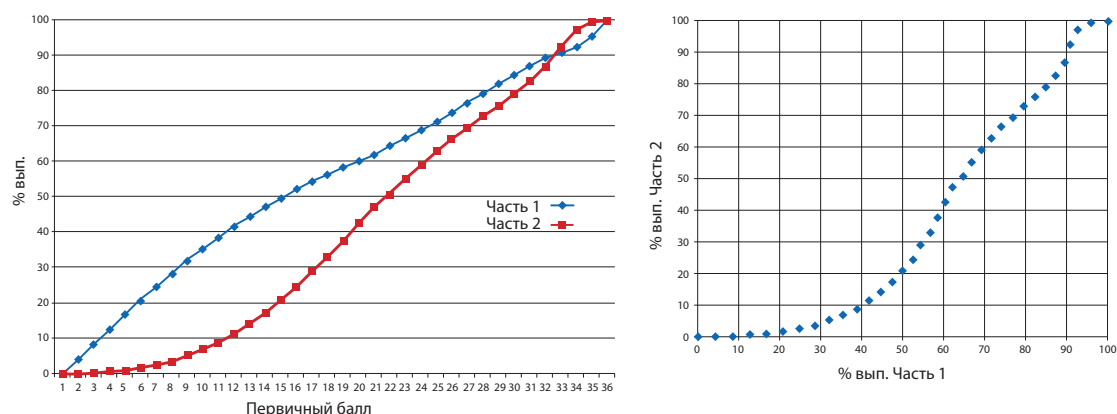


Рис. 5. Соотношение процента выполнения КИМ 2020 г. по частям

низких баллов) части КИМ с практико-ориентированными заданиями (средний процент выполнения $\sim 36\%$). Корреляция баллов между этими частями КИМ по всей выборке показывает достаточно сильную положительную связь: $\sim 0,8$.

Графики зависимости среднего процента выполнения части 1 и части 2 от итогового балла в КИМ ЕГЭ 2020 г. и соотношение процента выполнения по частям представлены на рисунке 5.

Аналогичные графики зависимости среднего процента выполнения части КИМ ЕГЭ 2021 г. с использованием специализированного ПО и части КИМ с заданиями без использования ПО представлены на рисунке 6.

Данные графики показывают прямую зависимость соответствующих частей КИМ: корреляция в ЕГЭ 2020 г. составляет 0,96, в ЕГЭ 2021 г. $-0,92$. При этом рассматриваемые части КИМ-2021, в отличие от КИМ-2020, однозначно на всём диапазоне первичного балла различаются

в сложности. На рисунке видно превышение сложности в части КИМ с заданиями, требующими специализированного ПО.

Рассмотрим выполнение КИМ информатики и ИКТ ЕГЭ разных лет (и, соответственно, разного формата представления) участниками различных регионов. Процент выполнения всей работы по информатике и ИКТ ЕГЭ 2020 г. (в бланковом формате) и процент выполнения соответствующей работы в 2021 г. (в компьютерном формате) по регионам показан на рисунке 7.

Можно отметить сильную корреляцию процента выполнения работы, представленной в различной форме, по всем регионам. Значение коэффициента корреляции (Пирсона) — более 0,92. При исключении нескольких регионов с малыми выборками и регионов с относительно сильным изменившимся количеством участников значение коэффициента корреляции (Пирсона) в 76 регионах возрастает и равно 0,96.

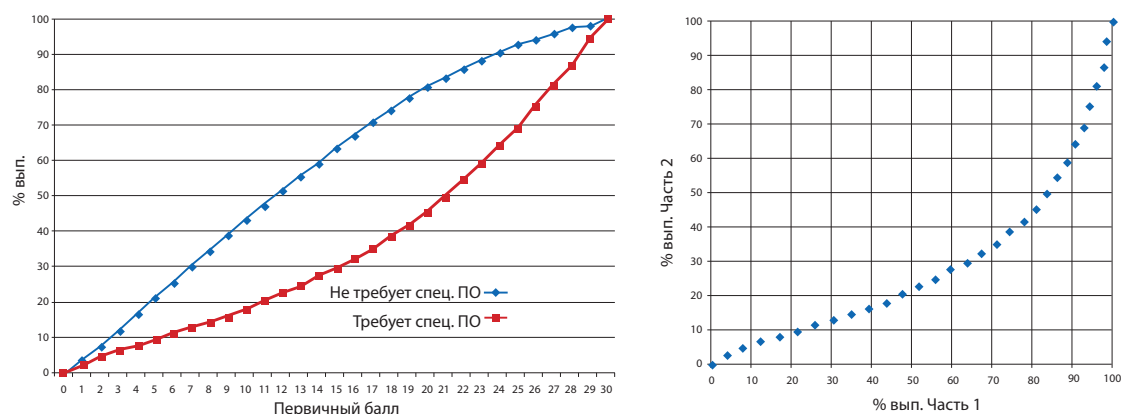


Рис. 6. Соотношение процента выполнения КИМ 2021 г. по частям

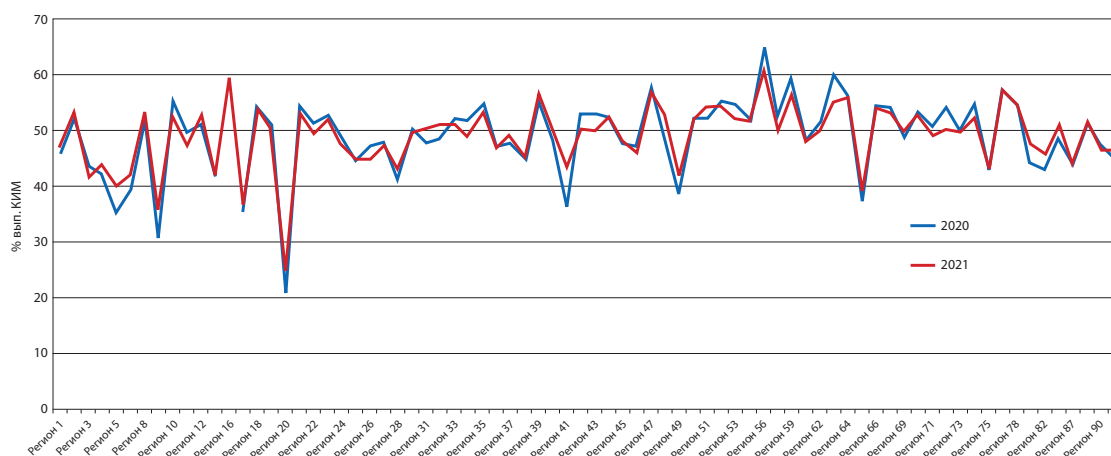


Рис. 7. Выполнение КИМ ЕГЭ в 2020–2021 гг. по регионам

Исследуем корреляцию итогового балла с баллами за **различные части КИМ**. Можно отметить одинаково сильную корреляцию итогового процента выполнения работы 2020 г. с процентом выполнения части КИМ 2021 г., требующей использования специального ПО (включающей задания на практическое программирование), и остальной части работы. Однако корреляции процента выполнения отдельных взятых частей работы с итоговым процентом выполнения КИМ 2020 года ниже, чем корреляции за всю работу ($\sim 0,91$), что говорит о составляющих, оценивающих независимые навыки и знания в каждой части.

Рассматривая корреляции процента выполнения КИМ по частям внутри регионов, можно отметить, что в регионах с низким средним баллом связь между решением практико-ориентированных заданий и остальной частью КИМ слабее, чем в остальных регионах. С возрастанием среднего процента выполнения работы (или силы региона) в среднем увеличивается коэффициент корреляции между итоговым баллом и процентом выполнения части КИМ с практико-ориентированными заданиями (см. диаграмму рассеяния зависимости среднего балла за работу и корреляции итогового балла и балла за практико-ориентированную часть на рисунке 8). Однако эта связь умеренная, и нельзя назвать её сильной по всем регионам РФ⁴.

Рассмотрим процент выполнения различных частей КИМ по группам участников с **различным уровнем подготовки**. Выделим четыре уровня подготовки участников:

- 1) от 0 до 5 баллов (ниже минимального);
- 2) от 6 до 14 баллов (низкий);
- 3) от 15 до 22 баллов (средний);
- 4) от 23 до 30 баллов (высокий).

На рисунке 9 для этих групп представлены проценты выполнения частей КИМ в зависимости от практико-ориентированности заданий.

В среднем процент выполнения части 1 (содержащей задания, требующие специального ПО) и части 2 (задания, не требующие ПО) возрастает с увеличением уровня подготовки. Однако внутри данных групп связь процента выполнения первой и второй частей работы практически отсутствует (корреляция $< 0,2$ для всех групп). Для группы с уровнем подготовки ниже минимального корреляция для большинства регионов либо незначима, либо отрицательна (увеличение процента выполнения практико-ориентированных заданий влечёт уменьшение процента выполнения остальных заданий, и наоборот). Для остальных групп с уровнем подготовки 2, 3, 4 внутри большинства регионов корреляции также малые и незначимые, что показывает наличие дополнительных факторов, по которым следует разделять выборку для выявления более сильных связей.

Рассмотрим процент выполнения задания КИМ по группам участников с **различной**

⁴ Тюрин Ю. Н., Макаров А. А. Анализ данных на компьютере: учебное пособие. — Новое изд. — М.: МЦНМО, 2016. — 368 с.

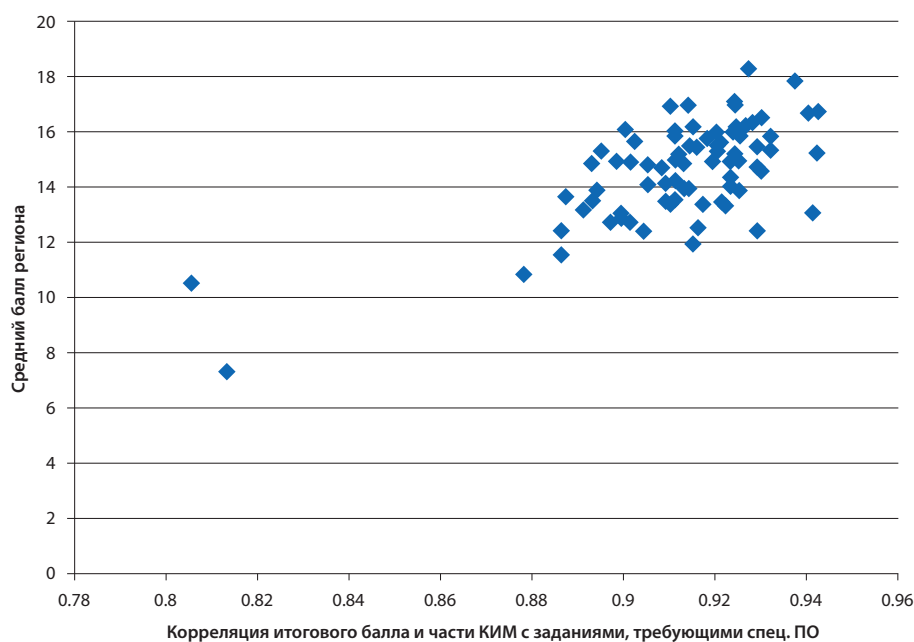


Рис. 8. Корреляции итогового балла с практико-ориентированной частью по регионам

гендерной принадлежности (юноши и девушки). Процент девушек, принимающих участие в экзамене по информатике и ИКТ, от года к году незначительно уменьшается и остаётся на уровне 21–24% (см. рис. 10).

В таблице 3 указан средний балл за КИМ ЕГЭ по информатике и ИКТ в 2019–2021 гг. среди участников в разбивке по гендерному признаку.

Средний балл (и, соответственно, процент выполнения работы) среди девушек

выше, чем у юношей, как во всей выборке, так и по регионам со значимой выборкой. На диаграммах баллов за 2020 г. (рис. 11) и 2021 г. (рис. 12) видно смещение распределения баллов у девушек в сторону высоких баллов.

Однако в 2021 г. при большей практико-ориентированности задач процент юношей, получивших крайние высокие баллы (27–30), выше, чем девушек. Соответственно, и в группе с высоким уровнем

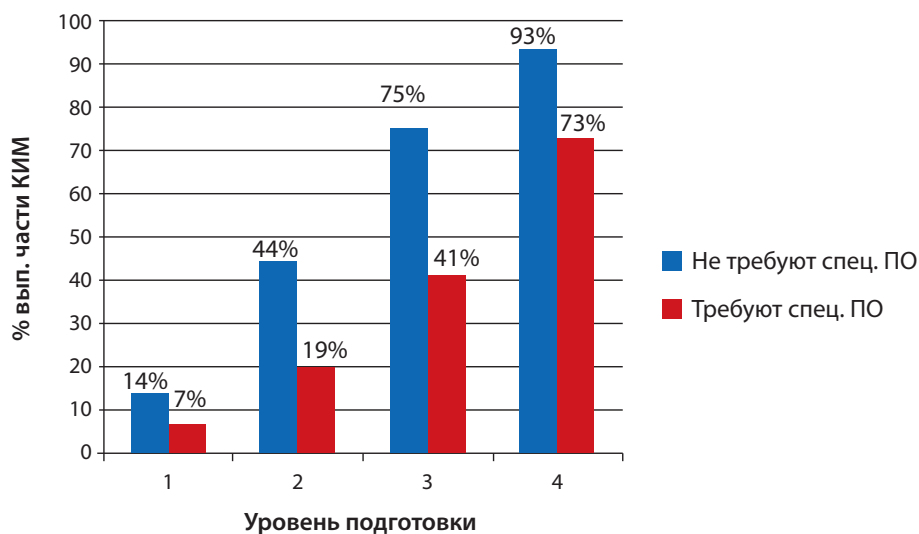


Рис. 9. Выполнение частей КИМ по уровням подготовки

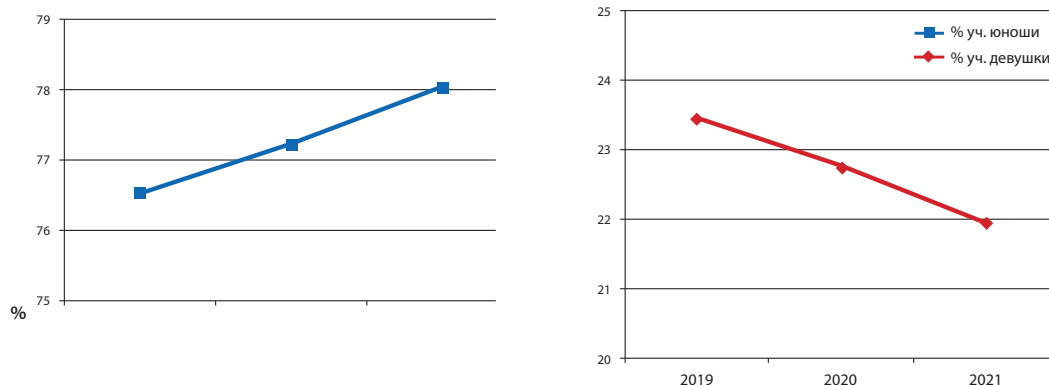


Рис. 10. Гендерный состав участников по годам

Таблица 3

Средний балл за КИМ ЕГЭ 2019–2021 гг. (юноши и девушки)

	Итоговый балл		
	2019 г.	2020 г.	2021 г.
Юноши	18	18	15
Девушки	20	19	16

подготовки процент юношей незначительно выше.

Анализируя процент выполнения отдельных заданий, можно отметить, что девушки хуже юношей справляются с заданиями на написание и отлаживание алгоритмов (например, задания 6 и 22, КТ — 1.1.4 «читать и отлаживать программы на языке программирования»,

19–21 — алгоритм логической игры, задания 24–27, в которых требуется специализированное ПО).

Рассмотрим корреляции двух частей КИМ (с практико-ориентированными заданиями и заданиями, не требующими специального ПО и практических навыков программирования) на группах по уровню подготовки с разбиением

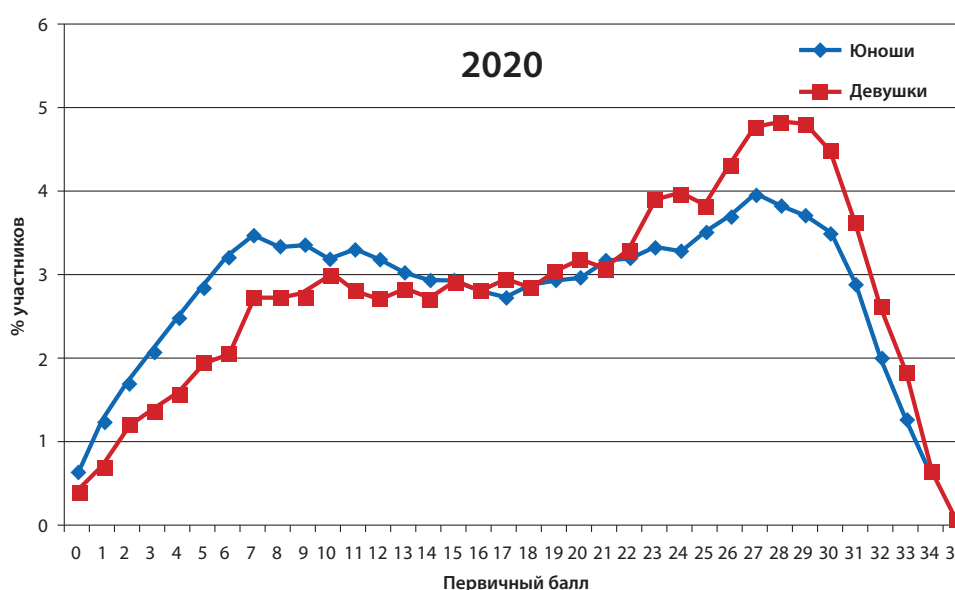


Рис. 11. Диаграмма баллов 2020 г. (юноши и девушки)

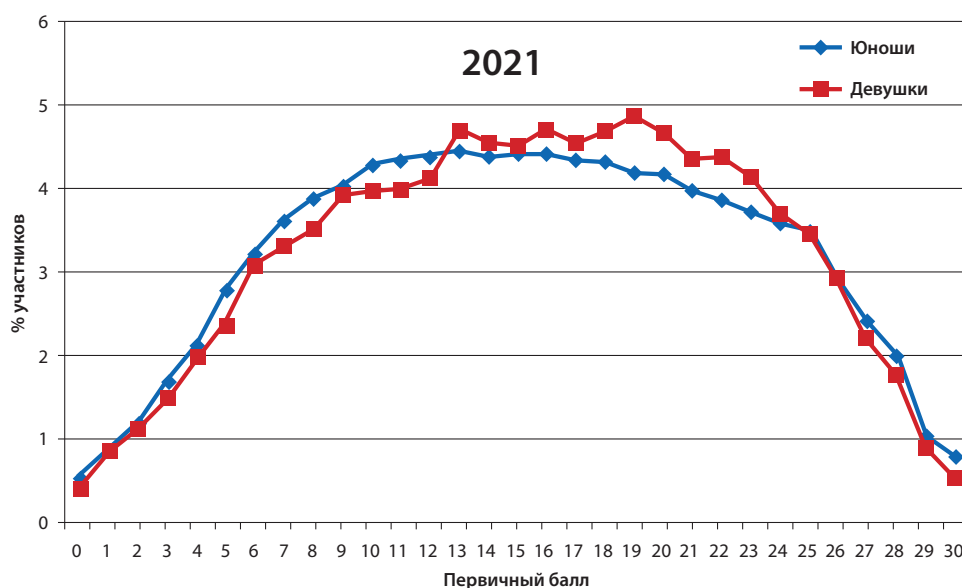


Рис. 12. Диаграмма баллов 2021 г. (юноши и девушки)

по гендерному признаку. Возьмём группы с высоким уровнем подготовки, так как в группах с низким уровнем подготовки корреляции слабые или незначимые, в том числе из-за высокого процента нулевых баллов в части КИМ, требующих специализированного ПО (~36%). Можно отметить, что в группах с высоким уровнем подготовки (по всем регионам в отдельности и по выборке в целом) между процентом выполнения разных частей КИМ корреляции у подвыборки девушек слабее, чем у юношей. То есть у участниц экзамена нельзя отметить сильной связи между высокими баллами за задания, не требующие специализированного ПО,

и за практико-ориентированные задания. На рисунке 13 показана диаграмма баллов за практико-ориентированные задания в группе с высоким уровнем подготовки (в разбивке по гендерному признаку).

На диаграмме видно смещение распределения баллов за эти задания у юношей в сторону высоких баллов. При этом диаграмма баллов за остальные задания имеет смещение в сторону высоких баллов у девушек (рис. 14).

Рассматривая связь выполнения заданий из разных частей КИМ, также можно установить, что участницы, набравшие высокие баллы за задания, не требующие специализированного ПО и практических

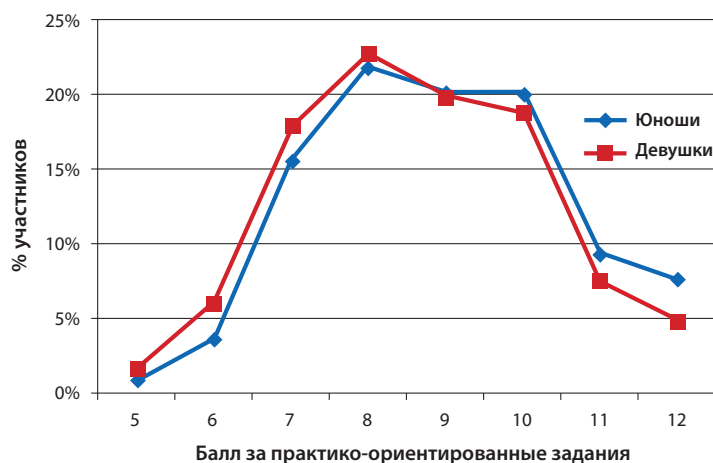


Рис. 13. Распределение баллов за практико-ориентированные задания

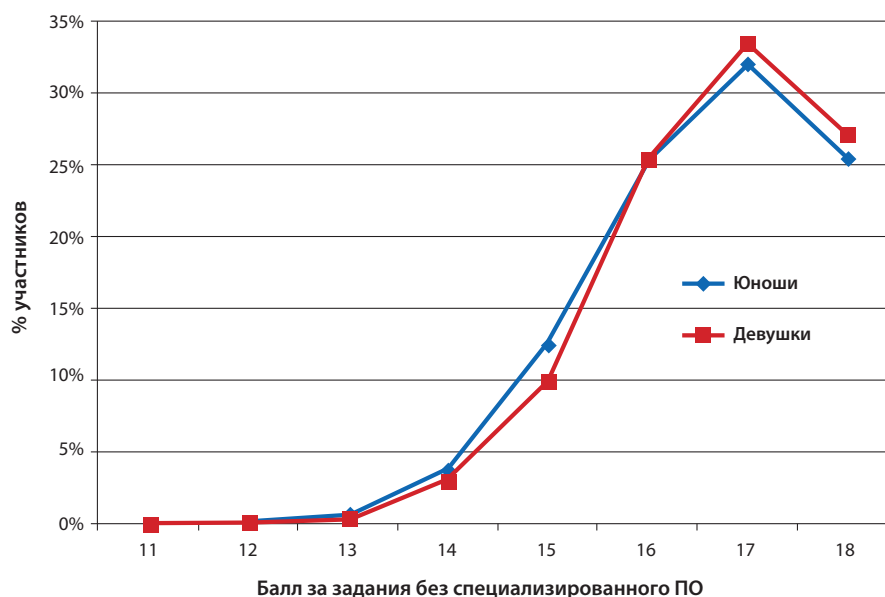


Рис. 14. Распределение баллов за задания без специализированного ПО

навыков, в среднем хуже справляются с практико-ориентированными заданиями, чем юноши (набравшие тот же балл).

Выводы

Таким образом, рассмотренные зависимости процента выполнения КИМ ЕГЭ по годам показывают, что новые практико-ориентированные задания (с использованием специального ПО) не изменили в целом характеристики работы по информатике и ИКТ. Выявлена высокая корреляция процента выполнения

КИМ по регионам за два года. Рассмотренные корреляции выполнения работы по частям КИМ и по группам участников позволили увидеть некоторую латентную переменную в итоговом проценте выполнения, относящуюся к практико-ориентированным заданиям. Статистический анализ выполнения различных частей КИМ в различных группах участников выявил высокую зависимость процента выполнения практико-ориентированных заданий от принадлежности к гендерной группе для участников с высоким уровнем подготовки.