

6 Частоты

97	Пошаговые алгоритмы вычислений
102	Представление результатов
105	Завершение анализа и выход из программы

В этой главе рассматриваются частоты, их графическое представление (столбиковые и круговые диаграммы), гистограммы и проценти́ли. Частоты — одна из команд, позволяющих непосредственно работать с диаграммами, не прибегая к командам меню Графика. Некоторые детали, касающиеся редактирования диаграмм, были рассмотрены в главе 5. *Столбиковые* и *круговые диаграммы* отображают количество объектов (*частоту*) в различных категориях дискретной переменной (имеющей небольшое число возможных значений). Однако команда Частоты позволяет работать не только с дискретными, но и с непрерывными переменными, которые могут принимать множество значений.

Далее приведены краткие определения понятий, о которых пойдет речь в этой главе:

- ▶ *Частоты*. Команда Частоты является одной из самых простых и часто используемых команд SPSS. Ее действие сводится к подсчету количества объектов в каждой категории переменной. Это называется распределением частот по категориям переменной. Если мы анализируем переменную пол, программа подсчитает распределение численности девушек и юношей среди учащихся; если используется переменная класс, получим распределение численности учащихся по классам. Выводимый результат для каждой категории включает в себя метку значения переменной, само значение переменной, частоту, процент и накопленный процент от общей частоты.
- ▶ *Столбиковые диаграммы*. Обязательное условие для применимости столбиковой диаграммы — дискретность отображаемых данных. Такими свойствами обладают переменные пол, класс, хобби и т. п. Каждая из этих переменных делит все объекты файла данных на категории. Например, переменная вуз разбивает всех учащихся на 4 группы по выбору профиля дальнейшего образования. Столбиковыми диаграммами можно отображать также непрерывные переменные. При этом к категории относятся данные, попадающие в определенный диапазон (интервал). Примерами непрерывных переменных являются средний балл отметки каждого учащегося, время финиша в гоночных состязаниях, вес пациента при медосмотре и т. п.
- ▶ *Гистограммы* используются для отображения распределения частот непрерывных переменных. Особенностью гистограмм, отличающей их от столбиковых диаграмм, является то, что каждый столбец представляет собой не отдельное значение, а диапазон значений переменной. Гистограммы могут использоваться и для дискретных данных в случаях, если число значений переменной слишком велико, чтобы отображать каждое из них отдельным столбцом диаграммы. Весь диапазон изменчивости переменной разделяют на интервалы, а затем по ним строят гистограмму. Именно таким методом следовало бы отображать распределение частот для переменных `отметка1` или `отметка2`.

- *Процентиль* показывает, какой процент распределения лежит ниже заданной величины. Например, если говорят, что процентиль значения 111 равен 75, это означает, что 75 % всех значений переменной меньше, чем 111, а 25 % — больше, чем 111.

Пошаговые алгоритмы вычислений

Для иллюстрации вычислений при помощи команды Частоты будем использовать уже знакомый вам файл `ex01.sav` (мы не приводим примеры круговых диаграмм ввиду их исключительной простоты). Количество объектов в файле равно 100. Однако сначала следует выполнить три подготовительных шага. Эти шаги позволят подготовить рабочий файл данных, запустить программу IBM SPSS Statistics 20 и открыть файл (в данном случае `ex01.sav`). Пошаговые инструкции для этого процесса приведены в главе 4 (с. 55), а подробные разъяснения — в главе 2.

После завершения шага 3 на экране должно присутствовать окно редактора данных со строкой меню и загруженным файлом `ex01.sav`.

ШАГ 4

В меню Анализ выберите команду Описательные статистики ► Частоты. На экране появится диалоговое окно, показанное на рис. 6.1.

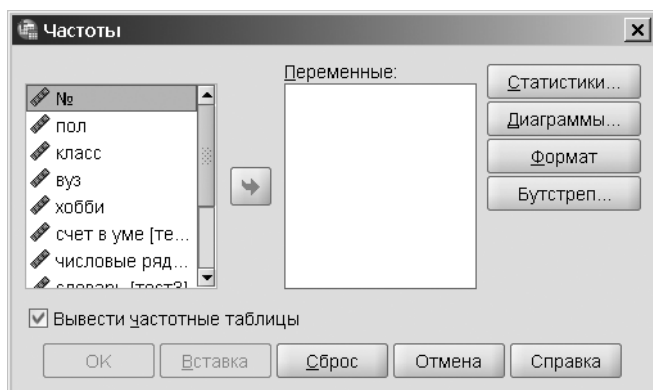


Рис. 6.1. Диалоговое окно Частоты

Частоты

Диалоговое окно Частоты типично для большинства статистических операций SPSS. В левой части окна расположен список всех доступных переменных. В нем необходимо выбрать те переменные, для которых надо вычислить распределение частот.

Для этого щелчком выделите нужную переменную, а затем щелчком на кнопке с направленной вправо стрелкой переместите ее в поле Переменные — целевой список. Если нужная переменная не видна в исходном списке, воспользуйтесь полосой прокрутки.

Для того чтобы удалить переменную из целевого списка, достаточно выделить ее, а затем воспользоваться кнопкой с направленной влево стрелкой, которая появляется на месте предыдущей кнопки при выделении любого пункта в списке Переменные. В этом случае переменная вновь переместится в исходный список. Чтобы полностью очистить список Переменные, щелкните на кнопке Сброс.

ШАГ 5

На этом шаге займемся вычислением частот для переменных пол, класс и вуз. При открытом диалоговом окне Частоты выполните следующие действия:

1. Щелкните сначала на переменной пол, чтобы выделить ее, а затем — на кнопке со стрелкой, чтобы переместить переменную в поле Переменные.
2. Повторите предыдущее действие для переменных класс и вуз.
3. Щелкните на кнопке ОК, чтобы открыть окно вывода, показанное на рис. 6.2.

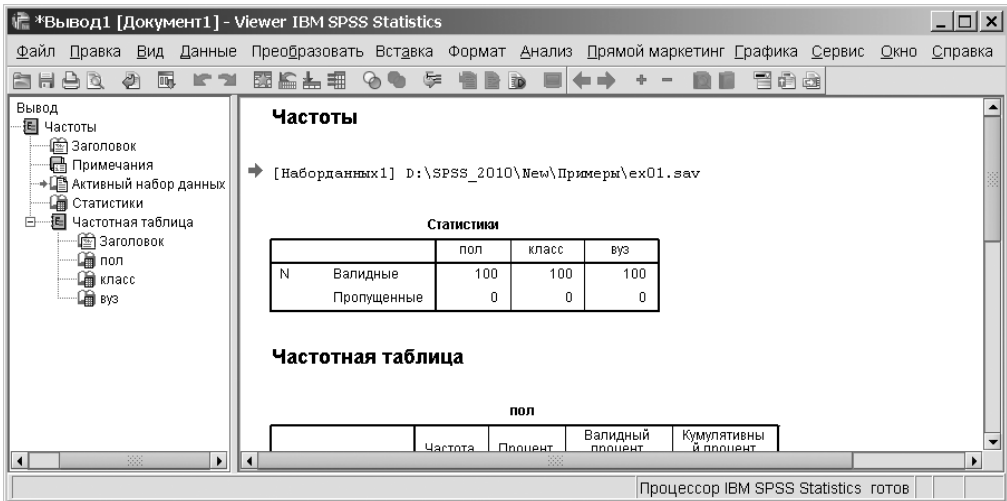


Рис. 6.2. Окно вывода программы SPSS

Переместив переменные в целевой список, вы завершаете выполнение операции щелчком на кнопке ОК. После этого программа SPSS формирует окно вывода с результатами выполнения команды. Выполнение практически всех команд статистических операций завершается именно так.

Окно с результатами работы программы имеет заголовок Вывод1 [Документ1]. Для просмотра результатов при необходимости можно воспользоваться полосой прокрутки в правой части окна. Структура выводимых результатов программы и основные приемы их редактирования описаны в разделе «Окно вывода и его редактирование» главы 2. Поскольку в заголовке окна видна строка меню, можно снова применить команду Частоты к выбранному подмножеству данных, для чего достаточно выбрать команду Анализ ► Описательные статистики ► Частоты и в открывшемся диалоговом окне сразу щелкнуть на кнопке ОК (ранее выбранные переменные окажутся в целевом списке автоматически), или предварительно изменить список анализируемых переменных.

Столбиковые диаграммы

Для того чтобы создать столбиковую диаграмму для дискретных данных, необходимо сначала выполнить все действия шага 5, кроме последнего, а затем в диалоговом окне Частоты щелкнуть на кнопке Диаграммы. На экране появится диалоговое окно Частоты: Диаграммы, показанное на рис. 6.3.

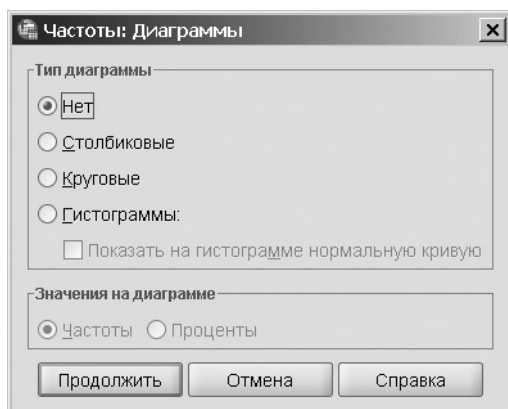


Рис. 6.3. Диалоговое окно Частоты: Диаграммы

Как видите, диалоговое окно Частоты: Диаграммы позволяет выбрать тип диаграммы с помощью переключателя. В данном примере вам понадобится установить переключатель Тип диаграммы в положение Столбиковые. В зависимости от величины, которую вы хотите использовать для отображения частот, установите переключатель Значения на диаграмме в положение Частоты или Проценты. Щелкните на кнопке Продолжить. После этого вы вернетесь в диалоговое окно Частоты и сможете завершить операцию щелчком на кнопке ОК.

ШАГ 5а

На этом шаге мы выполним построение столбиковых диаграмм для частот тех же трех переменных, пол, класс и вуз, которые использовались ранее. После выполнения шага 4 у вас должно быть открыто диалоговое окно Частоты, показанное на рис. 6.1. Если вы уже поработали с этим окном, очистите его нажатием кнопки Сброс. Для разнообразия будем заполнять целевой список двойными щелчками, а не щелчками на кнопке со стрелкой, как в предыдущем примере.

1. Дважды щелкните на переменной пол. Ее имя появится в поле Переменные. Сделайте то же самое для переменных класс и вуз и щелкните на кнопке Диаграммы, чтобы открыть диалоговое окно Частоты: Диаграммы.
2. Установите переключатель Тип диаграммы в положение Столбиковые. Если вы предпочитаете вместо частот выводить на диаграмме их доли от общей частоты в процентах, установите переключатель Значения на диаграмме в положение Проценты и щелкните на кнопке Продолжить, чтобы вернуться в диалоговое окно Частоты.
3. Щелкните на кнопке ОК, чтобы открыть окно вывода.

После выполнения этого шага программа сгенерирует три диаграммы, соответствующие выбранным переменным. При помощи полосы прокрутки можно просмотреть созданные диаграммы в окне вывода. Если возникнет необходимость отредактировать какую-либо из диаграмм, дважды щелкните на ней и обратитесь к материалу главы 5 для дальнейших действий. На рис. 6.4 показано окно вывода с диаграммой частот для переменной вуз.

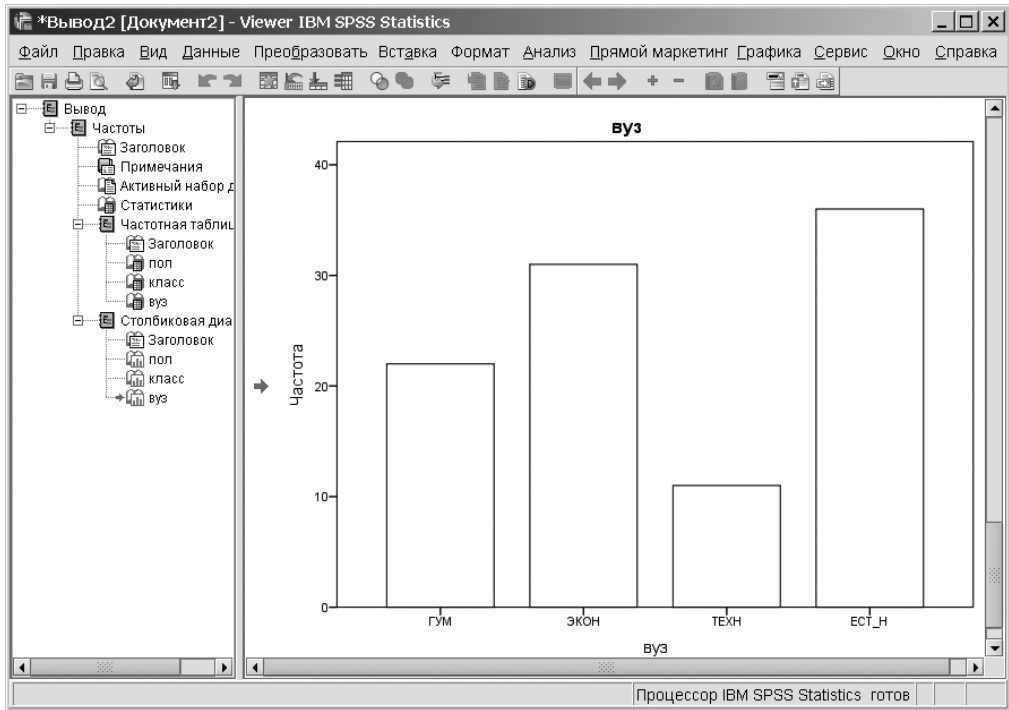


Рис. 6.4. Пример столбиковой диаграммы

Гистограммы

Процесс построения гистограмм аналогичен процессу построения столбиковых диаграмм. Единственным различием, которое следует всегда иметь в виду, является тот факт, что гистограммы предназначены для отображения распределения непрерывных переменных. Таким образом, для переменных пол, класс и вуз гистограммы непригодны. В то же время для переменной отметка2, отражающей среднюю отметку для каждого учащегося, гистограмма является весьма удобным средством описания распределения частот. Далее приведен пример построения гистограммы для переменной отметка2.

ШАГ 56

После выполнения шага 4 у вас должно быть открыто диалоговое окно Частоты, показанное на рис. 6.1. При необходимости повторите шаги 1–4. Если в целевом списке диалогового окна Частоты после работы над материалом предыдущего раздела остались переменные, щелкните на кнопке Сброс.

1. Щелкните сначала на переменной отметка2, чтобы выделить ее, а затем — на кнопке со стрелкой, чтобы переместить переменную в поле Переменные, и, наконец, на кнопке Диаграммы, чтобы открыть диалоговое окно Частоты: Диаграммы.
2. Установите переключатель Тип диаграммы в положение Гистограммы, если необходимо, установите флажок Показывать на гистограмме нормальную кривую и щелкните на кнопке Продолжить, чтобы вернуться в диалоговое окно Частоты.
3. В диалоговом окне Частоты снимите флажок Вывести частотные таблицы и щелкните на кнопке ОК.

Обратите особое внимание на необходимость снять флажок Вывести частотные таблицы. Этот флажок отвечает за наличие в выводимых данных команды Частоты всей неграфической информации. Если оставить его установленным при использовании непрерывных переменных, размер таблицы частот будет определяться всем множеством различных значений переменной, что редко оказывается полезным для исследователя. Поэтому при анализе непрерывных данных флажок Вывести частотные таблицы лучше снимать. Установленный флажок Показывать на гистограмме нормальную кривую указывает на то, что поверх гистограммы будет нанесена нормальная кривая.

Описательные статистики и процентили

Если вернуться к диалоговому окну Частоты и щелкнуть в нем на кнопке Статистики, на экране появится диалоговое окно Частоты: Статистики, показанное на рис. 6.5. В этом окне представлены такие показатели, как процентили и описательные статистики (подробно описательные статистики рассматриваются в главе 7). Из трех приведенных далее примеров вы узнаете, каким образом создать гистограмму, получить описательные статистики и вычислить процентили с заданным шагом или заданной величины. Во всех трех примерах используется переменная отметка2, а начинаются они с диалогового окна, полученного после выполнения шага 4 (см. рис. 6.1).

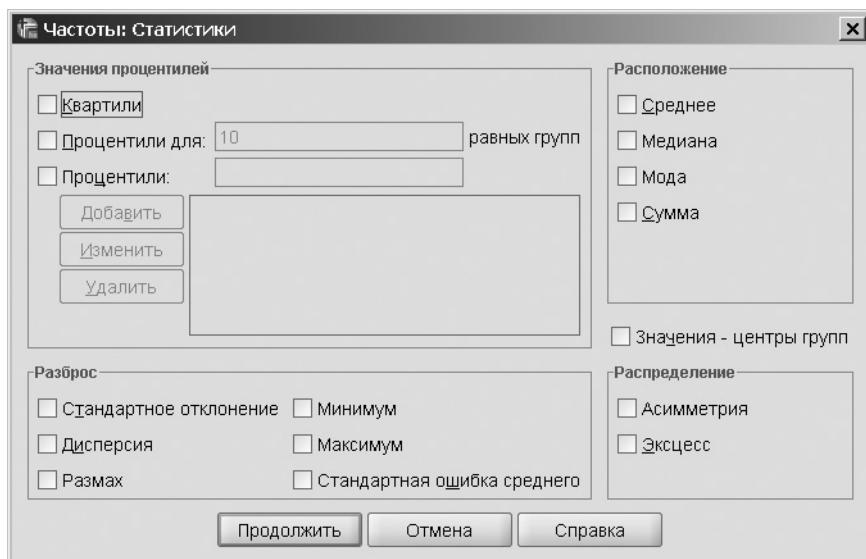


Рис. 6.5. Диалоговое окно Частоты: Статистики

ШАГ 5В

На этом шаге мы построим гистограмму распределения частот и вычислим для него среднее значение, стандартное отклонение, асимметрию и эксцесс распределения:

1. Щелкните сначала на переменной отметка2, чтобы выделить ее, затем — на кнопке со стрелкой, чтобы переместить переменную в поле Переменные, и, наконец, на кнопке Диаграммы, чтобы открыть диалоговое окно Частоты: Диаграммы.

2. Установите переключатель Тип диаграммы в положение Гистограммы и щелкните на кнопке Продолжить, чтобы вернуться в диалоговое окно Частоты.
3. Щелкните на кнопке Статистики, чтобы открыть диалоговое окно Частоты: Статистики, показанное на рис. 6.5.
4. В области Расположение установите флажок Среднее, в области Распределение — флажки Асимметрия и Экссесс, в области Разброс — флажок Стандартное отклонение и щелкните на кнопке Продолжить, чтобы вернуться в диалоговое окно Частоты.
5. В диалоговом окне Частоты снимите флажок Вывести частотные таблицы и щелкните на кнопке ОК.

ШАГ 5г

Чтобы вычислить процентилю переменной отметка2 с шагом 10, необходимо разделить шкалу процентов на 10 равных промежутков с границами, равными 10, 20, 30 и т. д.:

1. Щелкните сначала на переменной отметка2, чтобы выделить ее, затем — на кнопке со стрелкой, чтобы переместить переменную в поле Переменные, и, наконец, на кнопке Статистики, чтобы открыть диалоговое окно Частоты: Статистики, показанное на рис. 6.5.
2. В области Значения процентов установите флажок Процентилю для, в поле рядом с флажком введите значение 10 и щелкните на кнопке Продолжить, чтобы вернуться в диалоговое окно Частоты.
3. В диалоговом окне Частоты снимите флажок Вывести частотные таблицы и щелкните на кнопке ОК.

ШАГ 5д

Наконец, чтобы вычислить набор произвольных процентов (в данном случае 5, 70 и 95), выполните следующие действия:

1. Щелкните сначала на переменной отметка2, чтобы выделить ее, затем — на кнопке со стрелкой, чтобы переместить переменную в поле Переменные, и, наконец, на кнопке Статистики, чтобы открыть диалоговое окно Частоты: Статистики, показанное на рис. 6.5.
2. В области Значения процентов установите флажок Процентилю, в поле рядом с флажком введите значение 5 и щелкните на кнопке Добавить.
3. Введите в то же поле числа 70 и 95, по очереди добавляя их в список процентов, и щелкните на кнопке Продолжить, чтобы вернуться в диалоговое окно Частоты.
4. В диалоговом окне Частоты снимите флажок Вывести частотные таблицы и щелкните на кнопке ОК.

Обратите внимание на то, что квартили (25-й, 50-й и 75-й процентилю) можно вычислить одним действием, установив флажок Квартили в области Значения процентов диалогового окна Частоты: Статистики (см. рис. 6.5).

Представление результатов

Несмотря на то что выводимые программой SPSS результаты занимают не очень много места, как правило, их можно представить в еще более компактной форме (см. раздел «Окно вывода и его редактирование» в главе 2).

На возможные небольшие различия между данными, которые вы видите на экране своего компьютера, и иллюстрациями этой книги можно не обращать внимания.

Частоты

Фрагмент окна вывода после выполнения шага 5 приведен на рис. 6.6.

пол					
		Частота	Процент	Валидный процент	Кумулятивный процент
Валидные	ЖЕН	61	61,0	61,0	61,0
	МУЖ	39	39,0	39,0	100,0
	Итого	100	100,0	100,0	

класс					
		Частота	Процент	Валидный процент	Кумулятивный процент
Валидные	А	33	33,0	33,0	33,0
	Б	35	35,0	35,0	68,0
	В	32	32,0	32,0	100,0
	Итого	100	100,0	100,0	

вуз					
		Частота	Процент	Валидный процент	Кумулятивный процент
Валидные	ГУМ	22	22,0	22,0	22,0
	ЭКОН	31	31,0	31,0	53,0
	ТЕХН	11	11,0	11,0	64,0
	ЕСТ_Н	36	36,0	36,0	100,0
	Итого	100	100,0	100,0	

Рис. 6.6. Фрагмент окна вывода после выполнения шага 5

Далее дана трактовка терминов, используемых программой в окне вывода:

- ▶ Частота — количество объектов, соответствующих каждой категории (градации) переменной;
- ▶ Процент — процент от общей численности (с учетом пропусков). Если бы в файле были пропущенные значения, их процент был бы указан в предпоследней строке;
- ▶ Валидный процент — процент значений для каждой категории за вычетом пропущенных значений;
- ▶ Кумулятивный процент — накопленный процент величины Валидный процент к данному значению переменной;
- ▶ Валидные — список значений переменной;
- ▶ Итого — сумма значений в столбце.

Гистограммы

На рис. 6.7 приведены результаты выполнения шага 5б. По горизонтальной оси отложены значения переменной с шагом, равным 0,125 (ширина каждого столбца тоже равна 0,125). Каждое указанное значение соответствует середине столбца. На вертикальной оси гистограммы отложены частоты диапазонов. Справа от гистограммы помещены вычисленные параметры: Среднее значение и Стандартное отклонение распределения, а также общее число объектов (N). Кроме того, на гистограмму нанесена нормальная

кривая, поскольку при выполнении процедуры был установлен флажок Показывать на гистограмме нормальную кривую. Гистограмма без труда может быть сделана более наглядной. В частности, для данной гистограммы желательно укрупнение интервалов и, соответственно, уменьшение количества столбцов. По вопросам редактирования гистограмм обратитесь к главе 5 (см. рис. 5.12).

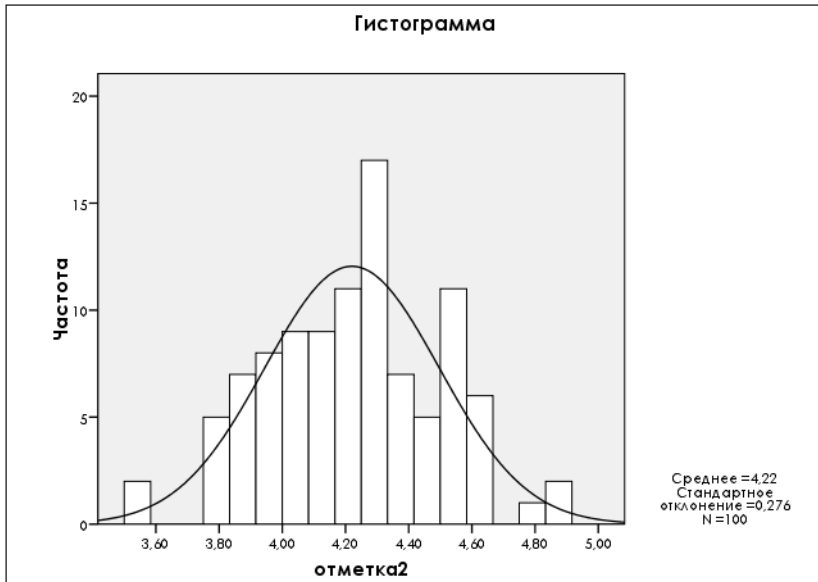


Рис. 6.7. Фрагмент окна вывода после выполнения шага 5б

Описательные статистики и процентиля

Результаты выполнения шагов 5г и 5д приведены на рис. 6.8.

Статистики		
отметка2		
N	Валидные	100
	Пропущенные	0
Среднее		4,2205
Стд. отклонение		,27589
Асимметрия		-,007
Стд. ошибка асимметрии		,241
Экссесс		-,357
Стд. ошибка эксцесса		,478

Статистики		
отметка2		
N	Валидные	100
	Пропущенные	0
Процентили	5	3,8000
	70	4,3500
	95	4,6500

Рис. 6.8. Фрагменты окна вывода после выполнения шагов 5г и 5д

Описательные статистики рассматриваются подробнее в главе 7, поэтому мы не станем обсуждать их здесь. Обратите внимание: если в вывод результатов включены Асимметрия и Экссесс, для них по умолчанию вычисляются стандартные ошибки (Стд. ошибка).

Что касается процентилей, то их можно трактовать следующим образом: 5 % значений переменной отметка не превышают 3,8 (5 % учащихся имеют отметку не выше 3,8), 70 % значений не превышают 4,35 (30 % учащихся имеют отметку выше 4,35) и т. д.

Завершение анализа и выход из программы

Отредактируйте содержимое окна вывода в соответствии со своими предпочтениями: скройте лишнюю информацию, исправьте таблицы и пр. (см. раздел «Окно вывода и его редактирование» в главе 2). При необходимости отредактируйте диаграммы (см. главу 5).

Для дальнейшего использования результата все содержимое окна вывода или его фрагменты можно сохранить в файле *.spv, экспортировать в другой формат (например, Word), перенести в документ Word или вывести на печать (подробности см. в разделе «Сохранение, экспорт, перенос и печать результатов» главы 2).

Для выхода из программы выберите в меню Файл команду Выход.