

Глава VII
**ИСПОЛЬЗОВАНИЕ ПРОГРАММИРУЕМЫХ МИКРОКАЛЬКУЛЯТОРОВ ДЛЯ
АНАЛИЗА СОЦИОЛОГИЧЕСКОЙ ИНФОРМАЦИИ**

**1. Организация обработки социологической информации. Классы задач,
решаемых на ЭВМ и на программируемых микрокалькуляторах**

Бурное развитие микроэлектроники в последние годы привело к возникновению нового класса клавишных вычислительных машин – программируемых микрокалькуляторов, которые практически воплотили «...еще недавно казавшуюся фантастической мечту об ЭВМ в кармане»¹. Наряду с карманными появляются также новые типы программируемых настольных вычислительных машин, бурно развивается производство микро- и мини-ЭВМ, отличающихся от больших ЭВМ меньшими объемом памяти и скоростью выполнения операций, но значительно более дешевыми и надежными.

Анализ тенденций развития программируемых клавишных вычислительных машин показывает, что они являются перспективной группой средств вычислительной техники. Увеличение их вычислительной мощности привело к стиранию граней между ними и мини-ЭВМ², в свою очередь приближающихся по возможностям к большим ЭВМ³. Таким образом, различные типы вычислительной техники начинают равномерно заполнять разрыв между обычными калькуляторами и большими ЭВМ.

Все это приводит к изменению стратегии использования вычислительной техники. Если раньше основным был принцип централизованной обработки информации, то теперь возникает новый принцип распределенной (рассредоточенной)

[220]

обработки. «Согласно этому принципу процессорная мощность цифровой системы не концентрируется в одном месте, а рассредоточивается так, чтобы была вблизи ее *потребителей* (курсив наш. - Авт.). Данное превращение структуры системы напоминает изменение характера электропривода, произошедшее в 30-х годах, когда один большой электромотор с ременными передачами к станкам был заменен индивидуальными для каждого станка электродвигателями»⁴.

Социологическая информация не является исключением и перспективы совершенствования методов ее анализа связаны не столько с разработкой пакетов прикладных программ для больших ЭВМ с возможностью диалоговой работы, как полагают некоторые авторы⁵, сколько с совершенствованием организационных и технических средств сочетания обработки на больших ЭВМ и на микро-, мини-ЭВМ и программируемых клавишных ЭВМ (ПЭКВМ). Наш опыт организации обработки социологической информации показал, что для решения широкого класса задач даже относительно ограниченные по возможностям программируемые микрокалькуляторы (типа «Электроника БЗ-21» и «Электроника БЗ-34») значительно эффективнее, чем ЭВМ. Их использование позволяет внести существенные изменения в организацию обработки социологической информации.

С определенной долей условности обработку информации можно разделить на первичную и вторичную. Под первичной понимается обработка, исходной информацией для

¹ Трохименко Я. К., Любич Ф. Д. Инженерные расчеты на микрокалькуляторах. Киев, 1980, с. 6.

² Ландеховская Н. Г., Меньшикова Л. А. Современное состояние и тенденции развития программируемых ЭВМ. Информационный бюллетень «Приборы, средства автоматизации и системы управления». Серия ТС-2. Средства вычислительной техники и оргтехники. М., 1979.

³ Брусенков Н. П. Миникомпьютеры. М., 1979, с. 20.

⁴ Брусенцов Н. П., Миникомпьютеры. М., 1979, с. 13.

⁵ SPSS (Statistical package for the social sciences). McGraw - Hill, 1975, p. XXII.

которой служат ответы респондентов (заполненные анкеты), первичная обработка представляет собой различного рода преобразования социологической информации: расчет одномерных и многомерных распределений признаков, таксономия, классификация и т.п. Результатом вторичной обработки являются показатели, рассчитанные на основе данных первичной или выполненной ранее вторичной обработки⁶, т.е. показатели, рассчитываемые по частотам, сгруппированным данным и т.п. (средние, меры рассеивания, связи, показатели значимости).

[221]

Естественно, что вся первичная обработка производится на ЭВМ. Исключение могут составлять пилотажные исследования и экспертные опросы в тех случаях, когда число опрашиваемых не превышает 20 - 30 человек - в этом случае первичная обработка может производиться вручную. Что же касается вторичной обработки, то в настоящее время она тоже производится на ЭВМ, а класс задач, решаемых на калькуляторах, чрезвычайно узок (это преимущественно суммирование и расчет процентов). Между тем, как будет показано ниже, решение значительного класса задач вторичной обработки информации на калькуляторах намного более эффективно, чем на ЭВМ. Рассмотрим этот вопрос подробнее.

Вторичная обработка социологической информации чаще всего включает в себя расчет мер центральной тенденции, вариации и связи, расчет уровней значимости и некоторых специальных показателей⁷: различные индексы, например привлекательности профессий, удовлетворенности работой; расстояния между двумя рядами распределений и т.п. К вторичной обработке можно отнести также некоторые из методов, рассмотренных в предыдущей главе: факторный анализ, исходной информацией для которого выступает матрица корреляций, причинный анализ, некоторые из социометрических методов. Вопрос о том, какие из видов вторичной обработки целесообразно проводить на ЭВМ, а какие - на микрокалькуляторах, зависит, на наш взгляд, от объема исходной информации, по которой рассчитывается показатель (например, от размерности матрицы), от числа показателей, которые требуется рассчитать и от организации работы на ЭВМ в данном социологическом подразделении (свой ВЦ или арендуемый, есть ли возможность работы в диалоговом режиме и т.д.). Кроме того, если необходимо вычислить нестандартные, редко используемые показатели, для расчета которых исследователь не располагает программами для ЭВМ, - более целесообразно рассчитать их на микрокалькуляторе.

Дело в том, что вторичная обработка социологической информации - это итеративный процесс, тесно сливающийся с анализом информации. Ее можно описать следующей цепочкой: интерпретация данных первичной обработки - расчет показателей для проверки гипотез, возникших при этом, - интерпретация полученных показателей и выдви-

[222]

жение новых гипотез - расчет новых показателей и т.п. Для расчета показателей на ЭВМ могут понадобиться следующие виды работ: перенос необходимых данных на специальные бланки, перфорация и контроль перфорации, организация доступа к ЭВМ (от вызова необходимых программ и информации из банка данных при работе с диалоговым монитором до заказывания машинного времени), счет. В общей сложности от возникшей необходимости рассчитать некоторый показатель до его расчета может пройти от нескольких часов до нескольких дней. Исключение составляет, пожалуй, лишь работа с диалоговым монитором при наличии необходимых программ и свободного доступа к ЭВМ.

⁶ Обработка результатов вторичной обработки тоже может считаться вторичной (нет смысла вводить понятия третичной, четвертичной и прочих видов обработки).

⁷ Имеются в виду показатели специфичные именно для рассматриваемых социологами проблем и не общепринятые в статистике.

На микрокалькуляторе требуется лишь ввести программу (несколько минут) и вводить данные (с визуальным контролем по индикатору) непосредственно с клавиатуры в регистры памяти или в операционные регистры. Расчет одного показателя занимает от нескольких секунд до нескольких минут, поэтому в случае необходимости рассчитать небольшое число показателей (несколько десятков) ПКЭВМ значительно эффективней⁸.

Разумеется, при необходимости рассчитать большое число показателей (например, матриц коэффициентов корреляции, содержащих сотни коэффициентов) следует использовать ЭВМ. Другой случай предпочтительного использования ПКЭВМ – расчет редко используемых показателей. Дело в том, что процесс программирования и отладки про, грамм на ПКЭВМ значительно проще, чем на обычных ЭВМ, поэтому в данном случае может иметь смысл обработка на ПКЭВМ и достаточно больших массивов информации. Другими словами, при оценке целесообразности выбора того суммарное время, затрачиваемое на создание программы, подготовку информации и другие этапы обработки.

Все эти соображения приводят к следующему разделению функций. На ЭВМ целесообразно рассчитывать статистику, сопровождающую таблицы одномерных, двухмерных и многомерных распределений признаков (меры центральной тенденции и меры вариации, рассчитываемые для каждой строки или каждого столбца таблиц, а также всевозможные

[223]

коэффициенты корреляции для таблиц). Кроме того, на ЭВМ целесообразно рассчитывать матрицы коэффициентов связей и уровней значимости признаков, например, матрицы коэффициентов связи между строками таблиц двухмерных распределений признаков. На микрокалькуляторах предпочтительно рассчитывать все показатели, определяемые не по таблицам сопряженности, а по отдельно взятым признакам или парам признаков, в частности коэффициенты корреляции для двух или нескольких признаков, меры значимости корреляций и различий между показателями. Кроме того, на калькуляторах целесообразно рассчитывать меры центральной тенденции, вариации и различные индексы, исходной информацией для которых служат средние коэффициенты корреляции и другие вторичные показатели, сопровождающие таблицы распределений признаков.

Это разделение (как и приводимые в следующем параграфе программы) относятся к использованию самого распространенного и доступного типа ПКЭВМ - «Электроники БЗ-21», поступающего в свободную продажу в магазины канцтоваров. Использование ПКЭВМ, обладающих более широкими возможностями, разумеется, расширяет класс задач, решаемых на программируемых клавишных ЭВМ. Так, например, использование настольной ПКЭВМ типа «Искра-125» (ввод программы с длиной до 100 шагов с магнитных карт, а исходных данных – с накопителей на магнитных лентах) или карманного микрокалькулятора фирмы Hewlett Packard типа HP-41C (программы длиной до 2000 шагов вводятся с магнитных карт, память для данных -319 ячеек, наличие печатающего устройства⁹) позволяет обрабатывать по относительно простым программам большие массивы информации (например, рассчитывать матрицы коэффициентов корреляции, проводить факторный анализ). Более совершенные ПКЭВМ (например, настольная клавишная ЭВМ типа HP-9830B с оперативной памятью, приблизительно равной памяти «Минск-22», устройствами записи и считывания с магнитных лент, перфокарт, с печатающим устройством, дисплеем и графопостроителем¹⁰) позволяют выполнять все виды вторичной и многие из видов первичной обработки информации.

⁸ Под эффективностью мы имеем в виду, прежде всего, экономию временных затрат; сказанное тем более касается финансовых затрат, так как стоимость, например, микрокалькулятора «Электроника БЗ-21» (80 р.) приблизительно равна стоимости аренды одного часа работы ЭВМ ЕС-1022.

⁹ Hewlett Packard. Electronic instruments and systems, 1980, # 4, 1980. (Каталог продукции).

¹⁰ Там же.

Можно предположить, что более перспективной является такая организация обработки информации, при которой в

[224]

социологических подразделениях располагается так называемая станция клавишного ввода¹¹, представляющая собой ПКЭВМ или микро-ЭВМ, снабженную дисплеем и устройством записи информации на магнитную ленту. Информацию прямо с анкет (минуя кодирование и перфорацию) вводят в ПКЭВМ или микро-ЭВМ, проверяют, редактируют и записывают на магнитную ленту (или магнитный диск). Эта лента переносится затем на большую ЭВМ, на которой производится первичная и часть видов вторичной обработки информации, результаты которой частью печатаются, а частью записываются на магнитную ленту. Затем лента опять переносится на ПЭКВМ, на которой производится детальный и углубленный анализ полученной информации, проверка гипотез, возникших при первичном анализе информации, расчет новых показателей и т.п.

2. Программы расчета статистических мер и уровней значимости

Изложенные ниже программы предназначены для работы на программируемом микрокалькуляторе «Электроника-БЗ-21» (и могут с незначительными изменениями использоваться для работы на ПКЭВМ «Электроника БЗ-34»). Микрокалькулятор «Электроника БЗ-21» функционирует в двух режимах. Нажатием клавиш Р и РП он переводится в режим программирования (рис. 30), во время которого в калькулятор вводится программа (максимальная длина программы – 60 шагов, имеются команды условного и безусловного перехода и возможность использовать подпрограммы, а также 7 ячеек обычной и 6 – так называемой «стековой» памяти). Затем нажатием клавиш Р и РР калькулятор переводится в режим работы, во время которого он автоматически производит расчеты по введенной программе или используется в качестве обычного микрокалькулятора.

Программы 2, 5, 6, 7 написаны Г. П. Талантом, программы 8, 9 и частично 1 и 11, а также использованные нами обозначения, заимствованы у Л. И. Францевича¹². Отметим,

[225]

что для работы по приведенным ниже программам желательно знакомство с инструкцией к микрокалькулятору, в частности, отметим, что появляющиеся на индикаторе микрокалькулятора числа иногда представлены в виде мантиссы и порядка числа. Например, запись [1.234567 03] на индикаторе означает $1,234567 * 10^3$, или 1234,567, а запись [3,361255—02] означает $3,361255 * 10^{-2}$, или 0,0361255. После ввода программы необходимо сначала провести расчеты для приведенного к каждой программе контрольного примера. Если полученный результат не совпадает с указанным в контрольном примере, то это означает, что при вводе

¹¹ Брусенцов Н. П. Миникомпьютеры, с. 28.

¹² Францевич Л. П. Обработка результатов биологического эксперимента на микро-ЭВМ «Электроника БЗ-21». Киев, 1979. Эту же книгу можно рекомендовать желающим освоить программирование на этом микрокалькуляторе. Наш опыт ведения семинаров для программирования на «Электронике БЗ-21» показывает, что для обучения программированию достаточно 4-х занятий.

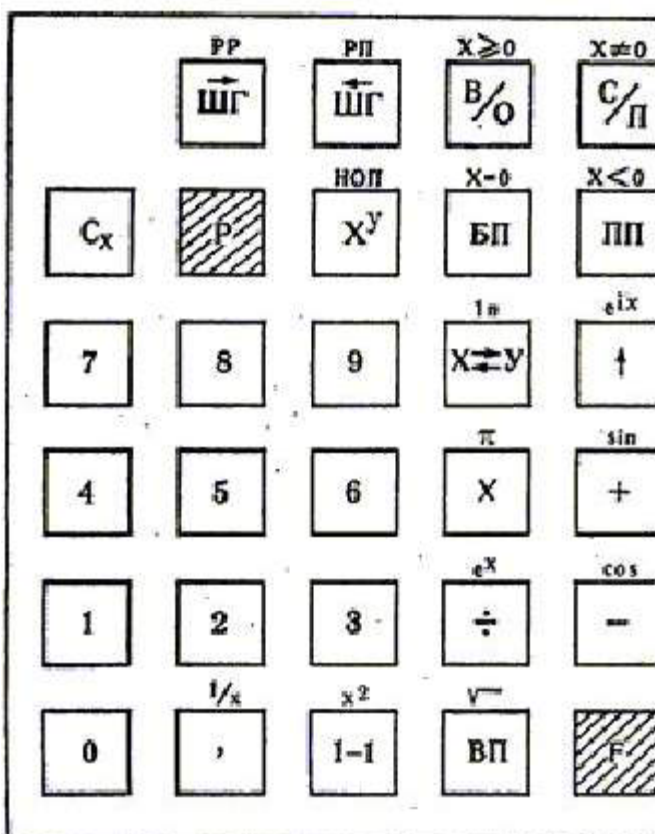


Рис. 30. Клавиатура микрокалькулятора «Электроника БЗ-21»

[226]

программы допущены ошибки, и программу следует ввести заново.

Необходимость создания программ специально для обработки данных социологических исследований связана с некоторыми особенностями вторичной обработки: большое количество порядковых и номинальных шкал, требующих использования ранговых и непараметрических критериев, наличие количественных признаков с заранее задаваемыми интервалами и т.д.

Другой особенностью уже технического, а не содержательного характера является своеобразие форм исходной информации – в таблицах одномерных, двухмерных и трехмерных распределений, используемых в социологических исследованиях, есть ряд уже вычисленных данных (проценты, суммы, коэффициенты), что дает возможность создавать более компактные программы и совмещать в одной программе расчет нескольких показателей.

При описании программ приняты следующие обозначения:

1) программы набираются по столбцам; после набора первого столбца на индикаторе справа появится число 10, после набора 2-го - 20, 3-го - 30 и т.д., что позволяет контролировать число введенных команд;

2) в прямоугольную рамку заключены операции, которые повторяются при вводе каждого числа или пары чисел из ряда исходных данных, пока ряд не исчерпается;

3) в круглых скобках помещены пояснения, в частности, описание выводимых на индикатор результатов и указание примерного времени автоматического счета;

4) обмен содержимым регистров x и y обозначен xu , а стековые операции обозначены с использованием соответствующих подписей на клавишах, т.е. P; P /—/

5) запятая в текстах программ и инструкций для вычисления - это не разделитель двух команд, а название соответствующей клавиши (будьте внимательны, не пропускайте ее при наборе программ!).

Среднее арифметическое, дисперсия, среднее квадратическое отклонение, коэффициент вариации, оценка дисперсии

А. Несгруппированные данные.

Представление данных

Результаты измерения: $x_1, x_2, \dots, x_i, \dots, x_n$

[227]

Результаты

Количество наблюдений n , среднее арифметическое \bar{x} , дисперсия σ^2 , среднее квадратическое отклонение σ , коэффициент вариации C_U оценка дисперсии s^2 .

Ввод программы

| | | | | | | | | | | |
|-----|-------|---------|----|------------------|-----------------|----|-----|-----|-----|----|
| В/0 | C_x | P / - / | + | БП | P4 | F2 | + | F√ | F5 | P |
| P | P2 | + | P2 | F↑ | C/Π | + | ÷ | C/Π | C/Π | PP |
| РП | P3 | P, | F3 | xy | Fx ² | ↑ | ↑ | ↑ | | |
| | P, | × | 1 | P / - / | × | P, | F5 | F4 | | |
| | C/Π | ↑ | + | ÷ | / - / | ÷ | × | ÷ | | |
| | ↑ | F2 | P3 | F ^{1/x} | ↑ | P5 | C/Π | C/Π | | |

Вычисление

Нажимаем клавиши В/О C/Π [x_i C/Π] (значение n , время счета 2 - 3 сек). Напоминаем, что запись [x_i C/Π] означает x_1 C/Π x_2 C/Π... x_n C/Π. Значение i , появляющееся на индикаторе после ввода каждого числа, помогает оператору контролировать, сколько чисел он уже ввел. После ввода всех чисел вычисляются показатели: БП P ÷ C/Π (значение \bar{x}), C/Π (значение σ^2), C/Π (значение σ), C/Π (значение C_U), C/Π (значение s^2).

Контрольный пример

Исходные данные: $x_i = 11\ 12\ 13\ 14\ 15$

Результаты: $n = 5$, $\bar{x} = 13$, $\sigma^2 = 2$, $\sigma = 1,414213$, $C_U = 0,1087856$, $s^2 = 2,5$

Формулы расчета

См.: (формулу с. 39), (I,4,1), (с. 40), (I,4,3).

Б. Сгруппированные данные.

Представление данных

Границы класса $\dots x' - x'' \dots$, (или среднее значение $\dots x_i \dots$), количество наблюдений $\dots N_i \dots$

Результаты

Количество наблюдений n , среднее арифметическое \bar{x} , дисперсия σ^2 , среднее квадратическое отклонение σ , коэффициент вариации C_U , оценка дисперсии s^2 .

Ввод программы

| | | | | | | | | | | | |
|-----|-------|----|---------|----|----|------------------|-------|----|-----|-----|----|
| В/0 | C_x | F8 | ↑ | ↑ | F4 | P / - / | × | P, | F5 | F4 | P |
| P | P2 | ÷ | P / - / | F2 | + | ÷ | / - / | + | × | ÷ | PP |
| РП | P3 | ↑ | + | + | P3 | F ^{1/x} | ↑ | P5 | C/Π | C/Π | |
| | P, | P7 | P, | P2 | БП | P4 | F2 | + | F√ | F5 | |
| | C/Π | F4 | F7 | F3 | F↑ | C/Π | + | ÷ | C/Π | C/Π | |
| | P4 | × | × | ↑ | - | Fx ² | ↑ | ↑ | ↑ | | |

[228]

Ввод константы

Если задано среднее значение класса, то 1P8;
если заданы границы, то 2P8.

Вычисление

В/О С/П (не обращайтесь внимания на числа, которые могут появиться на индикаторе).

Если задано среднее значение класса: $x'_i \uparrow N_i$ С/П. Если заданы границы класса: $x'_i \uparrow x''_i + N_i$ С/П (получим значение n).

1 БП, (запятая означает, что надо нажать клавишу, на которой изображена запятая) С/П (значение \bar{x}) С/П (значение σ^2) С/П (значение σ) С/П (значение C_U) С/П (значение s^2)

Контрольный пример

Исходные данные

| | | | | | |
|-------|----|----|----|----|----|
| x_i | 11 | 12 | 13 | 14 | 15 |
| N_i | 1 | 3 | 10 | 3 | 1 |

Ввод констант: 1 P8 $\sigma^2 = 0.7777777$ $\sigma = 0.8819171$

Результаты: $n = 18; \bar{x} = 13; C_U = 0.06783977; s^2 = 0.8235294$

Формулы расчета: См.: (с. 39), (I,4,3)

Квантили, медиана для сгруппированных данных (при равных интервалах)

Представление данных

Границы класса ($x'_i - x''_i$), расположенные в порядке убывания, N_i - количество наблюдений (частоты), N - число опрошенных (сумма частот).

Результаты

Квантиль P_p (p - доля, для которой нужно определить квантиль), медиана (рассчитывается как $p = 0,50$).

Ввод программы

| | | | | | | | | |
|-----|------------|------------|------------|--------|------------|------------|---------|-----|
| В/О | \uparrow | С/П | $Px < 0$ | \div | $-$ | \times | БП | P |
| P | F7 | P5 | F \div | $P,$ | xy | \uparrow | F / - / | PP |
| РП | \times | \div | xy | БП | F5 | P / - / | | |
| | P6 | \uparrow | P / - / | P1 | \div | $-$ | | |
| | C_x | F6 | \uparrow | xy | \uparrow | С/П | | |
| | \uparrow | $-$ | F8 | F5 | F8 | xy | | |

Вычисление

1. Занести в регистр «P8» ширину интервала: $x'_1 \uparrow x''_1 - P8$ (в случае, если таблица составлена так, что $x''_1 \neq x_{i+1}$, например, 20 — 23, 24 — 27, 28 — 31 и т.д., прибавляется единица, т.е. $x''_1 \uparrow x'_1 - 1 + P8$).

[229]

2. $x'_1 P$ (ввод нижней границы в стековый регистр¹³) NP7

3. Набираем величину квантиля: p В/О С/П (на индикаторе 0; 2 сек)

4. [N_i С/П] (на индикаторе будут появляться накопленные частоты; 2 сек.). В процессе вычисления калькулятор выделяет из введенных частот ту, которая соответствует интервалу, содержащему квантиль, и автоматически переходит к его вычислению. При этом (чтобы отличить его от накопленных частот) квантиль появляется со знаком «—». Таким образом,

¹³ В случае, если $x''_i \neq x'_{i-1}$, x'_i заменяем на $x'_i - \frac{1}{2}(x'_2 - x'_1)$, т.е. вводим C_U .

первое отрицательное число, появившееся на индикаторе, представляет собой искомое значение квантиля Pp . Никакого другого смысла минус не имеет и его следует отбросить. В случае, если оператор не обратит внимания на то, что на индикаторе появилось отрицательное число и будет продолжать вводить N_i С/П на индикаторе все равно будет восстанавливаться значение полученного квантиля.

5. Если по той же таблице необходимо рассчитать несколько квантилей, то можно поступить двояким способом. Во-первых, можно повторять пп. 3 и 4 заново. Во-вторых, можно начать вычисления с наибольшего квантиля, записывая накопленные частоты. В этом случае вычисления последующих квантилей производится следующим образом: $F6$ (на индикаторе появляется значение произведения pN). По этой величине определяем из таблицы соответствующие N_i и нижнюю границу N_i . Затем $N_i P5 x'_i P$ (не обращать внимания на появляющиеся при этом числа). Определяем по таблице предыдущую накопленную частоту $sumF_{i-1}$. Далее $sumF_{i-1} \uparrow F6$ БП Р4 С/П (получим значение квантиля; 2 сек).

Контрольный пример

| | | | | | | | | |
|--------------------|----------------|-------|-------|-------|-------|-------|-----|-----|
| Границы интервалов | $x'_i - x''_i$ | 20-24 | 24-28 | 28-32 | 32-36 | 36-40 | ... | N |
| Частоты | σ | 135 | 295 | 291 | 307 | 204 | ... | 204 |

Медиана $P_{0.50} = 35.51791$ («—» отбрасывается).

[230]

Формулы расчета¹⁴

$$Pp = x_i + \frac{pN - sumF_{i-1}}{N_i} (x''_i - x'_i),$$

где x'_i - нижняя граница интервала, содержащего частоту pN

N_i - частота этого интервала

$sumF_{i-1}$ - накопленная к i -му интервалу.

Энтропийная мера вариации

Представление данных

N_1, N_2, \dots, N_k - частоты одномерного распределения признака

Результаты

Энтропийная мера вариации c_v , сумма частот N

Ввод программы

| | | | | | | | | |
|-----|-----|-----|----|-----|----|-----|-----|----|
| В/О | Сх | ↑ | P3 | F5 | F3 | P1n | /—/ | P |
| P | P3 | P1n | F2 | 1 | ↑ | — | C/П | PP |
| РП | P4 | X | ↑ | + | F4 | ↑ | | |
| | P5 | ↑ | F4 | P5 | + | F5 | | |
| | C/П | F3 | + | БП | ↑ | P1n | | |
| | P2 | + | P4 | F ↑ | F4 | + | | |

¹⁴ В случае, если $x''_i \neq x'_{i+1}$, формула несколько видоизменяется: x заменяется на $x''_i \neq x'_{i+1}$, а $(x''_i - x'_i)$ на $(x''_i - x'_i + 1)$: программа остается без изменений, требуется лишь выполнить указания в примечаниях к пп. 1 и 2.

Вычисление

В/О С/П 2 сек. [N_i С/П (5 сек., на индикаторе значение i)] БП Р4 С/П (получим N_i ; 5 сек.) F4 (получим N).

Контрольный пример

Варианты

Результат y_2

Частоты

5 7

$N=39$

Формулы расчета

$$N = \sum_{i=1}^k N_i$$

$$\varepsilon = -\frac{1}{\ln k} \sum_{i=1}^k \frac{N_i}{N} \ln \frac{N_i}{N} = -\frac{1}{\ln k} \left(\frac{1}{N} \sum_{i=1}^k N_i \ln N_i - \ln N \right)$$

[231]

Мода для сгруппированных данных

Представление данных

Дано одномерное распределение признака X , N_i - частоты, x'_i и x''_i - границы класса, I_i - ширина интервала (см. сноску к предыдущей программе), l - интервал, содержащий моду.

Результаты: Мода (M_0)

Ввод программы

| | | | | | |
|-----|----|----|------|-----|----|
| В/О | ↑ | — | — | × | Р |
| Р | F4 | + | ↑ | ↑ | РР |
| РП | + | Р, | Р/—/ | F2 | |
| | ↑ | F5 | ÷ | + | |
| | F5 | ↑ | ↑ | С/П | |
| | xy | F4 | F3 | | |

Вычисление

$x_1 P_2 I_1 P_3 N_{l-1} P_4 N_l P_5 N_{l+1} B / O C / П (M_0; 4 \text{сек.})$

Контрольный пример

| | | | | |
|---------|-------|--------|---------|---------|
| Признак | 60-80 | 80-100 | 100-120 | 120-140 |
| Частоты | | 110 | 204 | 50 |

$M_0=107,5806$

Формулы расчета – см.: (1,3,3).

Коэффициент корреляции r , уровень значимости t

Представление данных

Число объектов n , x_1, x_2, \dots, x_n - значение переменной X , y_1, y_2, \dots, y_n значения переменной Y для каждого объекта, N_x и N_y - сумма значений переменной X и Y соответственно.

$$N_x = \sum_{i=1}^n x_i \quad N_y = \sum_{i=1}^n y_i$$

Результаты

Коэффициент корреляции между признаками X и Y – r , уровень значимости t .

[232]

Ввод программы

| | | | | | | | | | | | |
|-----|----|-------------------|-----------------|-----|-----------------|-----------------|----|----|-----|-------------------|----|
| В/О | P3 | F ¹ /x | × | С/П | F5 | F6 | F7 | ↑ | xy | F4 | P |
| P | xy | xy | P6 | P3 | + | + | + | F7 | ÷ | 2 | PP |
| РП | P2 | × | F3 | xy | P5 | P6 | P7 | × | С/П | + | |
| | × | P5 | Fx ² | P2 | F2 | F3 | БП | FV | 1 | F ¹ /x | |
| | ↑ | F2 | × | × | Fx ² | Fx ² | P3 | ↑ | — | БП | |
| | F4 | Fx ² | P7 | ↑ | ↑ | ↑ | F6 | F5 | P7 | P7 | |

Вычисление

n | — | P4 N_x ↑ N_y В/О С/П (4 сек.)

[x_i ↑ y_i C/П] (4 сек). Не обращайтесь внимания на появляющиеся при этом на

индикаторе числа. После ввода всех пар значений переменной X и Y определяем r и t. БП ВП С/П (получаем r; 3 сек). P5 Fx² С/П (получаем уровень значимости t; 4 сек).

Контрольный пример

Исходные данные:

| | | | | | | | |
|-------|----------------|----|----|----|----|----|----------------------|
| n = 5 | x _i | 11 | 12 | 13 | 14 | 15 | N _x = 65 |
| | y _i | 22 | 22 | 23 | 25 | 25 | N _y = 117 |

Результаты

r = 0.9383148 = 9.383148 · 10⁻¹,

t = 4.700095

Так как при 3-х степенях свободы критическое значение t на уровне 5% равно 3,2, а на уровне 1% - 5,8, данное значение r значимо на уровне 5%.

Формулы расчета См. (II,5,3), (V,8,1)

Критические значения

См. табл. И Приложение 3.

***Коэффициент корреляции Спирмена (ρ) при отсутствии связанных рангов¹⁵,
уровень значимости t***

Представление данных

n - число объектов,

R₁^(x), R₂^(x), ..., R_n^(x) - ранги 1-го, 2-го, ..., n-го объекта соответственно по переменной X

R₁^(y), R₂^(y), ..., R_n^(y) - ранги 1-го, 2-го, ..., n-го объекта соответственно по переменной Y

[233]

Результаты

Количество введенных пар рангов n (для контроля правильности ввода), коэффициент корреляции ρ, уровень значимости t

Ввод программы.

¹⁵ В случае связанных рангов следует воспользоваться программой для вычисления r - это всегда позволит определить ρ независимо от того, связаны ранги или нет (гл. 11, §4).

| | | | | | | | | | | | |
|-----|----------------|----|----------------|----|----|----|-----------------|-----|----|-----|----|
| В/О | C _x | F8 | P7 | 1 | 4 | F6 | × | /—/ | F7 | xy | P |
| P | P8 | + | БП | + | × | — | С/П | P7 | xy | ÷ | PP |
| РП | P7 | P8 | 0 | P6 | ↑ | ↑ | P5 | F6 | + | С/П | |
| | С/П | F7 | xy | × | F5 | F7 | Fx ² | 3 | FV | | |
| | × | 1 | P7 | P5 | + | + | 1 | — | ↑ | | |
| | ↑ | + | C _x | F8 | ↑ | 3 | — | ↑ | F5 | | |

Чистка

В/О С/П

Вычисление

[$R_i^{(x)} \uparrow R_i^{(y)}$; С/П] (после ввода каждой i -й пары на индикаторе появляется текущее значение i , что облегчает оператору правильность ввода; 3 сек). После ввода всех пар чисел на индикаторе должно появиться заданное число n .

БП 2 С/П (получим ρ ; 5 сек),

С/П (получим t ; 5 сек).

Контрольный пример

Исходные данные:

$n=12$

| | | | | | | | | | | | | |
|-------------|---|---|----|---|---|---|---|----|---|----|----|---|
| $R_i^{(x)}$ | 2 | 8 | 12 | 3 | 1 | 6 | 7 | 10 | 4 | 9 | 11 | 5 |
| $R_i^{(y)}$ | 6 | 5 | 10 | 7 | 3 | 4 | 9 | 8 | 1 | 11 | 12 | 2 |

Результаты

$$\rho = 0.7062935$$

$$t = 3.155016$$

Формулы расчета

См. (V,8,4)

Критические значения. Табл. И, Приложение 3.

Коэффициент корреляции τ , уровень значимости z

Представление данных

n - число объектов

x_1, x_2, \dots, x_n - значения переменной X (или ранги по X) для каждого объекта

y_1, y_2, \dots, y_n - значения переменной Y (или же ранги по Y) для каждого объекта.

[234]

Результаты

Коэффициент корреляции τ , для несвязных и для связных рангов, уровень значимости z (для случая несвязных рангов).

Ввод программы

| | | | | | | | | | | | |
|-----|------|----|-----|------|------|-----|-----|----|-------------|-----|----|
| В/О | С/П | F3 | F8 | — | БП | + | ПП | 9 | /—/ | + | P |
| P | P, | — | БП | × | 2 | ↑ | P8 | + | $Px \neq 0$ | ↑ | PP |
| РП | F2 | ↑ | PO | ↑ | P/—/ | F7 | С/П | ↑ | + | F7 | |
| | — | P, | С/П | P/—/ | × | xy | 1 | F8 | ↑ | + | |
| | P/—/ | × | ↑ | + | FV | ÷ | 0 | БП | × | P7 | |
| | ↑ | ПП | 1 | P, | 2 | С/П | + | 4 | FV | В/О | |

Вычисление

1. $C_x P7$ (чистка) БП FO $x_1 P2 y_1 P3 [x_i \uparrow y_i C/\Pi]$ (4 сек.)

Действия рамке выполняются для $i = 2, 3, \dots, n$. Затем $x_2 P2 y_2 P3 [x_i \uparrow y_i C/\Pi]$ ($i = 3, 4, \dots, n$) $x_3 P2 y_3 P3 [x_i \uparrow y_i C/\Pi]$ ($i = 4, 5, \dots, n$) и так далее.

Последний раз $x_{n-1} P2 y_{n-1} P3 x_n \uparrow y_n C/\Pi$

2. $nP3$ БП F \times , C/П (2 сек.) $xу P8 /—/ P$,

3. Пункт 3 выполняется только при наличии связанных рангов по X – в противном случае переходим сразу к п.6

4. F8 /—/ P

5. Пункт 5 выполняется лишь в том случае если есть связанные ранги по Y, в противном случае переходим сразу к п. 6

$[g_i C/\Pi]$ (2 сек.), g_i - число связанных рангов в i -группе связанных рангов по признаку Y

6. P /—/ \uparrow БП F4C/П (значение τ ; 3 сек.)

7. C/П (2 сек.) F3 4 ν C/П (значение z; 4 сек.)

Контрольный пример

Исходные данные

| | | | | | |
|-------|---|---|-----|-----|---|
| x_i | 5 | 1 | 3 | 4 | 2 |
| y_i | 3 | 1 | 4,5 | 4,5 | 2 |

$n=5$ $g=2$ для переменной Y

Результаты: $\tau = 0.5270463$, $z = 0.9797959$

Формулы для расчета¹⁶

$$\tau = \frac{P-Q}{\sqrt{\frac{n(n-1)}{2} - T_x \sqrt{\frac{n(n-1)}{2} - T_y}} = \frac{S}{\frac{1}{2} \sqrt{\left[n(n-1) - \sum_i f_i(f_i-1) \right] \left[n(n-1) - \sum_i g_i(g_i-1) \right]}}$$

$$Z = \frac{S^*}{\frac{1}{18} n(n-1)(2n+5)} = \frac{S^*}{\frac{1}{2} \sqrt{\frac{4n+10}{9} n(n-1)}}$$

$$\text{где } S^* \begin{cases} S+1 & \text{при } S > 0 \\ S & \text{при } S = 0 \\ S-1 & \text{при } S < 0 \end{cases}$$

$$S = P - Q$$

Критические значения

См. табл. А Приложение 3.

Примечание Пункт 1 программы вычисления τ довольно трудоемкий, иногда имеет смысл определить значение S, вычисляемое в пункте 1, вручную (в этом случае объекты ранжируются по значениям одной из переменных). Далее S заносится в «P7» и вычисление повторяется, начиная с п. 2.

Линейная корреляция и регрессия

Представление данных

x_1, x_2, \dots, x_n - значения признака X

y_1, y_2, \dots, y_n - значения признака Y

¹⁶ Формулы для z в случае связанных рангов см. в кн.: Кендэл М. Ранговые корреляции. М., 1975, с. 66.

Средние x и y , коэффициент корреляции r , количество пар объектов n и параметры a и b линейного уравнения $y = ax + b$

Ввод программы

| | | | | | | | | | | | |
|-----|-----|----|----|----|-----------------|----------------|----------------|-----|----|------|----|
| В/О | С/П | P, | F2 | P, | Fx ² | P2 | P/—/ | P7 | ↑ | xy | P |
| P | P2 | ПП | ПП | 1 | + | P/—/ | ПП | F3 | F2 | P/—/ | PP |
| PP | xy | F4 | F4 | + | P, | P3 | C _x | ПП | ÷ | — | |
| | P3 | F2 | ПП | P, | ↑ | ПП | F6 | P8 | ↑ | /—/ | |
| | × | ПП | P4 | БП | F3 | C _x | × | С/П | F4 | ↑ | |
| | ↑ | P4 | + | PO | В/О | P6 | F√ | P4 | × | В/О | |

Чистка

↑ /—/ (шесть раз)

Вычисление

В/О С/П [x_i ↑ y_i С/П] (6 сек)

P /—/ (значение n) БП P5 С/П (10 сек)

[236]

F7 ÷ (получим r) F6 ÷ (значение a) ↑ P/—/ P/—/ P/—/ (значение \bar{x}) × ↑ P/—/ P/—/ (значение y) xy (значение b)

Контрольный пример

Исходные данные

| | | | | | |
|-------|----|----|----|----|----|
| x_i | 11 | 12 | 13 | 14 | 5 |
| y_i | 22 | 22 | 23 | 25 | 25 |

Результаты:

$n = 5, r = 0.9383148, a = 0.9, \bar{x} = 13, y = 23.4, b = 11.7$

Формулы расчета (см. II,5,3), (III,1,11), (III,1,12).

Линейная корреляция и регрессия при группировке в классы

Представление данных

| Частоты | | | | $N_1 \dots$ | $N_i \dots$ |
|--------------------|------------|-----|---|-------------|-------------|
| Середины признаков | интервалов | для | X | $x_1 \dots$ | $x_i \dots$ |
| | | | Y | $y_1 \dots$ | $y_i \dots$ |

Результаты

Общее число респондентов N , средние значение \bar{x} и \bar{y} , коэффициент корреляции r и параметр линейного уравнения $y = ax + b$

Ввод программы

| | | | | | | | | | | | |
|-----|-----|-----|----|-----|------|----------------|----------------|-----|----|------|----|
| В/О | P2 | ÷ | ÷ | ↑ | ↑ | P/—/ | P/—/ | P7 | ↑ | xy | P |
| P | ПП | ПП | F4 | ПП | P/—/ | F2 | ПП | F3 | F2 | P/—/ | PP |
| PP | P4 | F÷ | ПП | F÷ | + | В/О | C _x | ПП | ÷ | — | |
| | С/П | С/П | ÷ | С/П | ↑ | ПП | F6 | P8 | ↑ | /—/ | |
| | P3 | P4 | F4 | F3 | P, | C _x | × | С/П | F4 | ↑ | |
| | ПП | ПП | ÷ | × | xy | P6 | F√ | P4 | × | В/О | |

Чистка

C_x P /—/ (6 раз)

Вычисление [N_i В/О С/П (2-3 сек.) x_i С/П (4 сек.) y_i С/П (6 сек.)]

P /—/ (получим N) P2 P/—/ P3 БП 5С/П (10сек.)

F7 ÷ (получим r) F6 ÷ (получим a) ↑ P/—/ P/—/ P/—/ (значение \bar{x}) × ↑ P/—/ P/—/ (значение y) ху (значение b).

[237]

Контрольный пример

| | | | | |
|-------|----|----|----|----|
| N_i | 10 | 3 | 3 | 10 |
| x_i | 10 | 10 | 20 | 20 |
| y_i | 10 | 20 | 10 | 20 |

Результаты:

$$N = 26; r = 0.5384615; a = 0.5384615; \bar{x} = 15; \bar{y} = 15; b = 6.923077$$

Формулы расчета

$$N = \sum_{i=1}^k N_i; \sum_{i=1}^k x_i = \sum_{i=1}^k N_i x_i \quad (k - \text{число интервалов})$$

$$\sum_{i=1}^k x_i^2 = \sum_{i=1}^k N_i x_i^2; \sum_{i=1}^k y_i = \sum_{i=1}^k N_i y_i;$$

$$\sum_{i=1}^k y_i^2 = \sum_{i=1}^k N_i y_i^2; \sum_{i=1}^k x_i y_i = \sum_{i=1}^k N_i x_i y_i;$$

Остальные формулы те же, что и в предыдущем случае (т.е. без группировки в классы).

Программа для вычисления по таблице $k \times l$ значения χ^2 , коэффициента Чупрова T, коэффициента Крамера T.

Исходные данные: Таблица $k \times l$ (см. § 2 гл. 11)

Результаты. Значение χ^2 , T и T_c

Ввод программы

| | | | | | | | | | |
|-----|-----------------|----|----|-----|----|----|-----|-----|----|
| В/0 | С/П | F2 | БП | ↑ | ↑ | × | P6 | F√ | P |
| P | ху | + | P0 | F3 | F5 | ↑ | С/П | F√ | PP |
| РП | Fх ² | ↑ | F8 | × | × | F6 | F4 | ↑ | |
| | ху | F8 | ↑ | P6 | F√ | ху | ↑ | F6 | |
| | ÷ | + | 1 | С/П | ↑ | ÷ | F5 | × | |
| | ↑ | P8 | — | F4 | F3 | F√ | ÷ | С/П | |

Вычисление

Чистка C_x P8. NP3. Больше из чисел $(k-1)$ и $(l-1)$ запоминаем в 4-й (P4) меньше - в 5-й ячейке (P5).

Считаем по столбцам N (y_1) P2 N₁₁ ↑ N (x_1) В/О С/П С/П и далее в цикле $[N_{i1} \uparrow N(x_i) \text{С/П}] i = 2, 3, \dots, k$ (2 сек). Затем для второго столбца N (y_2) P2 $[N_{i2} \uparrow N(x_i) \text{С/П}] i = 1, 2, \dots, k$ и т.д. После ввода всех столбцов БП P × С/П (получим χ^2)

С/П (получим T) С/П (получим T^2).

[238]

Контрольный пример

| | y_1 | y_2 | y_3 | $N(x_i)$ |
|-------|-------|-------|-------|----------|
| x_1 | 29 | 36 | 15 | 80 |
| x_2 | 14 | 24 | 2 | 40 |

| | | | | |
|----------|----|----|----|-----|
| $N(x_i)$ | 43 | 60 | 17 | 120 |
|----------|----|----|----|-----|

C_x P8 120 P3 2P4 1P5 Далее 43P2 29 ↑ 80 В/О С/П С/П 14 ↑ 40 С/П 60P2 36 ↑ 80 С/П 24 ↑ 40 С/П 17P2 15 ↑ 80 С/П 2 ↑ 40 С/П Все столбцы введены ВП P × С/П (получим $\chi^2 = 4,770432$) С/П ($T = 0,1676604$) С/П ($T_c = 0,1993830$)

Формулы расчета
См. (II,2,1), (II,2,4), (II,2,5)
Уровни значимости
Табл. Б Приложение 3.

Коэффициенты связи для таблицы 2×2 : Q, Φ и χ^2

Представление данных:

| | Y_1 | Y_2 |
|-------|----------|----------|
| x_1 | N_{11} | N_{12} |
| x_2 | N_{21} | N_{22} |

Результаты

Коэффициент Юла Q, коэффициент контингенции Φ, показатель χ^2 .

Ввод программы

В/О xy ↑ F3 F6 + F3 ↑ — xy + P P
P + F6 + × ↑ ↑ F5 P8 ÷ C/П PP
PП P6 × ↑ P6 F6 F4 × — C/П ↑
P7 P6 F7 F2 × × P5 + F8 F7
F3 F2 + ↑ F√ P4 ↑
+ ↑ P7 F4 P6 F2 F4 F8 F6 C/П

Вычисление $N_{11}P2 N_{12}P3N_{21}P4N_{22}P5$ В/О С/П (получили Q; 8 сек.) С/П (получили Φ; 2 сек.) $F\chi^2$ С/П (получили χ^2 ; 1 сек.)

[239]

Контрольный пример:

Исходные данные: таблица

| | |
|---|---|
| 1 | 2 |
| 3 | 4 |

Результаты:

$Q = -0,20$ $\Phi = -0,08908708$ $\chi^2 = 0,07936508$

Формулы расчета

См (с. 86), (с. 88), (II,2,1).

Коэффициенты частной корреляции (для r и τ).

Представление данных:

Коэффициенты парной корреляции r_{01} , r_{02} и r_{12} (или τ_{01} , τ_{02} и τ_{12}) для коэффициента частной корреляции $r_{01.2}$ или $\tau_{01.2}$; коэффициенты $r_{01.2}$, $r_{03.2}$ и $r_{13.2}$ для коэффициента $r_{01.23}$ и т.д.

Результаты

Коэффициенты частной корреляции любого порядка (для вычисления коэффициентов n -ого порядка сначала вычисляются коэффициенты $(n-1)$ -го порядка)

Ввод программы

| | | | | | | | |
|-----|-----------------|------|----|-----------------|----|-----|----|
| V/O | P/—/ | × | 1 | ↑ | × | C/Π | P |
| P | ↑ | ↑ | ↑ | F3 | F√ | | PP |
| PΠ | Fx ² | P/—/ | F2 | Fx ² | ↑ | | |
| | P2 | xy | — | — | F4 | | |
| | P/—/ | — | P2 | ↑ | xy | | |
| | P3 | P4 | 1 | F2 | ÷ | | |

Вычисление.

Для коэффициентов первого порядка $r_{01}P, r_{02}P, r_{12}P$, В/О С/Π (получим $r_{01,2}$, 7 сек).

Для коэффициентов второго порядка $r_{01,3}P, r_{02,3}P, r_{13,2}P$, В/О С/Π (получим $r_{01,2,3}$; 7 сек) и т.д.

Аналогично вычисляются коэффициенты $\tau_{01,2}; \tau_{01,2,3}$ и т.д.

Контрольный пример

$r_{01,3} = 0.620$ $r_{03,2} = 0.240$ $r_{13,2} = -0.171$ $r_{01,2,3} = 0.6911214$

Формулы расчета

(III,2,3 - III,2,5), (III,2,6), (III,2,7)

Коэффициент множественной корреляции

Представление данных

Частные коэффициенты корреляции, необходимые для расчета коэффициента множественной корреляции $R_{0.12...n} : r_{01}; r_{02,1}; r_{03,1,2}; \dots; r_{0n,12...(n-1)}$

[240]

Результаты

Коэффициент $R_{0.12...n}$

Ввод программы.

| | | | | | |
|-----|-----------------|----|----|-----|----|
| V/O | 1 | xy | БΠ | F√ | P |
| P | P2 | — | P↑ | C/Π | PP |
| PΠ | C/Π | ↑ | 1 | | |
| | Fx ² | F2 | ↑ | | |
| | ↑ | × | F2 | | |
| | 1 | P2 | — | | |

Вычисление

В/О С/Π r_{01} С/Π (3 сек) $r_{02,1}$ С/Π (3 сек)... $r_{0n,12...(n-1)}$ С/Π (3 сек) После ввода всех частных коэффициентов корреляции БΠ P × С/Π (получим $R_{0.12...n}$; 3 сек)

Контрольный пример.

$r_{01} = 0.705$ $r_{02,1} = 0.479$ $r_{03,1,2} = 0.342$ $R_{0.1,2,3} = 0.8110240$

Формулы расчета

(III,3,7)

Значимость различия долей (процентов)

Представление данных

v_1, v_2 - сравниваемые доли признаков (проценты)

n_1, n_2 - объемы выборок, по которым рассчитывались v_1 и v_2 соответственно.

Результаты

Критерий значимости различий долей z (может использоваться, если выполняется каждое из следующих условий: $n_1 \geq 50, n_2 \geq 50, np > 5, n(1-p) > 5$, формула для вычисления приведена ниже).

Ввод программы

| | | | | | | | | | |
|-----|----|------|------|------|----|-----|----|-----------------|----|
| В/О | P4 | F1/x | F3 | ↑ | F7 | /—/ | × | Fx ² | P |
| P | xy | P7 | F1/x | P/—/ | + | ↑ | F√ | F√ | PP |
| РП | P5 | ↑ | P4 | + | ↑ | F8 | ↑ | C/П | |
| | — | F4 | ↑ | P5 | F5 | + | F6 | | |
| | F6 | × | F5 | F4 | xy | ↑ | xy | | |
| | F2 | P, | × | ↑ | ÷ | F5 | ÷ | | |

Счет

Если v_1 и v_2 выражены в долях, вводим: 1P8, если в процентах: 100P8

$n_1 P2 \ n_2 P3 \ v_1 \uparrow \ v_2 \ В/О \ С/П$ (получим z ; 9 сек)

Контрольный пример

Исходные данные:

[241]

$$n_1 = 392, \quad v_1 = 0.296, \quad v_1 = 29.6\%, \\ n_2 = 277 \quad v_2 = 0.209 \quad \text{или} \quad v_2 = 20.9\%$$

Результат: $z = 2,526962$

Формулы расчета

См (V,5,1).

Критические значения

См. табл. А Приложения 3.

Значимость различий средних арифметических

Представление данных

$\bar{x}_1, s_1, \bar{x}_2, s_2$ - средние арифметические и оценки средних квадратических отклонений, рассчитанные по выборкам объема n_1 и n_2 соответственно.

Результаты,

Критерий значимости различий t , число степеней свободы f

Ввод программы

| | | | | | | | | | |
|-----|----------------|-----------------|-----|-----------------|-----------------|-----------------|----|-----|----|
| В/О | C _x | С/П | F5 | Fx ² | С/П | Fx ² | 1 | ↑ | P |
| P | P3 | ПП | ÷ | F√ | Fx ² | P4 | + | F5 | PP |
| РП | P5 | F4 | 2 | ↑ | xy | F3 | ↑ | + | |
| | С/П | F3 | — | F3 | P2 | + | F4 | P5 | |
| | ПП | Fx ² | С/П | F√ | ÷ | P3 | xy | В/О | |
| | F4 | ↑ | — | ÷ | ↑ | F2 | ÷ | | |

Вычисление

В/О С П (1 сек) $n_1 \uparrow s_1$ С/П (≈ 4 сек)

$n_2 \uparrow s_2$ С/П (получим v ; ≈ 5 сек)

$\bar{x}_1 \uparrow \bar{x}_2$ С/П (получим критерий t ; 1 сек)

Контрольный пример

Исходные данные:

$$\bar{x}_1 = 15, s_1 = 4, \bar{x}_2 = 11, s_2 = 0.1, n_1 = 30, n_2 = 65$$

Результаты:

$$v = 29,01788$$

$$t = 5,476435$$

Формулы расчета

См. (V,6,1), (V,6,2)

Критические значения

См. табл. А Приложение И.

Значимость различий двух коэффициентов корреляции

Представление данных

Коэффициенты корреляции r_1 и r_2 рассчитанные по выборкам n_1 и n_2 соответственно.

[242]

Результаты

Уровень значимости различий коэффициентов корреляции z .

Ввод программы

| | | | | | | | |
|------|------|------|-----------------|------|-----|---|----|
| В/О | ↑ | 3 | FV | ↑ | FV | 1 | P |
| P | 3 | — | P2 | P, | ↑ | + | PP |
| PП | — | F1/x | ПП | ÷ | F2 | 2 | |
| | F1/x | ↑ | , | FV | ÷ | — | |
| | P2 | F2 | ПП | Pln | C/П | ÷ | |
| P/—/ | + | , | Fx ² | P/—/ | В/О | | |

Вычисление

Не обращая внимания на числа, которые могут появиться из стека, выполняем:

$r_1 P, r_2 P, n_1 P, n_2$ В/О С/П (получим z ; 9 сек)

Контрольный пример

Исходные данные:

$$r_1 = 0.6 \quad n_1 = 28$$

$$r_2 = 0.8 \quad n_2 = 23$$

Результаты: $z = 1,351550$

Формулы расчета

См. (V,9,1).

Критические значения

См. табл. А приложения 3

Значимость различий распределений (критерий χ^2)

Исходные данные

Таблица распределений двух совокупностей респондентов по некоторому признаку X:

| | |
|---------------------|-----|
| $k_1 k_2 \dots k_n$ | K |
| $l_1 l_2 \dots l_n$ | L |

где k_i и l_i - частоты; k_i - число респондентов из первой совокупности, выбравших i -ю градацию при ответе на вопрос X; l_i - аналогичная частота для второй группы респондентов. K и L -- суммы частот:

$$K = \sum_{i=1}^n k_i \quad L = \sum_{i=1}^n l_i$$

Результаты

[243]

Критерий χ^2 , показывающий значимость различий распределений признака X в 2 группах респондентов¹⁷.

Ввод программы

| | | | | | | | | | |
|----|----------------|-----------------|-----------------|----|-----|----|----|-----|----|
| ВО | C _x | xy | F4 | F6 | F7 | F2 | ↑ | C/Π | P |
| P | P7 | Fx ² | Fx ² | + | + | ↑ | ↑ | | PP |
| РП | C/Π | ↑ | ↑ | ↑ | P7 | F3 | — | | |
| | P4 | F2 | F3 | F5 | БП | + | ↑ | | |
| | ↑ | + | + | + | P↑ | P8 | F8 | | |
| | P5 | P6 | ↑ | ↑ | C/Π | F7 | × | | |

Вычисление

Заносим K в P2; L в P3. В/О C/Π (1 сек)

[$k_i \uparrow l_i$ C/Π] ($t_1 \approx 4$ сек) БП P5 C/Π (получим χ^2)

Контрольный пример

Исходные данные:

| | | | |
|----|----|----|----|
| 29 | 36 | 15 | 80 |
| 14 | 24 | 2 | 40 |

Результат: $\chi^2 = 4,770444$

Формулы расчета

$$\chi^2 = (K + L) \left[\sum_{i=1}^n \left(\frac{k_i^2}{K} + \frac{l_i^2}{L} \right) \frac{1}{k_i + l_i} - 1 \right]$$

Критические значения

См. табл. Б Приложение 3.

Что касается расчета специальных показателей, то во многих случаях (в частности, при расчете индексов удовлетворенности, престижа и т.п.) можно воспользоваться программой вычисления средних арифметических для сгруппированных данных (вместо середин интервалов взять значения весов тех или иных вариантов ответа). При большом числе индексов при неизменных весах для ускорения расчетов целесообразно использовать специальные программы. Приведем пример такого рода программы.

Расчет индекса для пятибалльных шкал

Представление данных

Одномерное распределение по признаку, имеющему 5 вариантов частоты: N_1, N_2, N_3, N_4, N_5 . Сумма частот равна N. Каждый из вариантов имеет вес: $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$

[244]

Результат и формула расчета

Некоторый индекс

$$I = \frac{1}{N} \sum_{i=1}^5 \beta_i N_i$$

Ввод программы

¹⁷ В цит. книге Л. Францевича есть программа расчета критерия в случае, если значение K и L не заданы.



Вычисление

а) $\beta_1 P2 \beta_2 P3, \beta_3 P4 \beta_4 P5 \beta_5 P6,$

б) NP7

в) В/О С/П (1 сек.) N_1 С/П (3 сек) N_2 С/П (3 сек) N_3 С/П (3 сек) N_4 С/П (3 сек) N_5

С/П (получим индекс I).

При неизменных баллах для расчета нового индекса начинать с п. б., а при неизменном N - с пункта в.

Контрольный пример

| | | | | | | |
|-----------|----|-----|----|------|----|----------|
| β_i | 1 | 0.5 | 0 | -0.5 | -1 | |
| N_i | 20 | 20 | 30 | 10 | 10 | $N = 90$ |

Результат: $I = 0,1666666$

Ниже приводится программа расчета евклидовых расстояний между векторами, характеризующими социологические объекты (респондентов, групп респондентов). Такого рода расстояния используются для определения близости объектов в некотором пространстве признаков (типичный пример - разбиение массива опрошенных на некоторые типы в соответствии с содержательными гипотезами и проверка того, насколько удачно введены типы, путем расчета расстояний между ними). Эта программа представляет интерес тем, что результатом счета является не одно число, а матрица расстояний 4×4 (т.е. 6 чисел, так как матрица симметрична).

[245]

Программа расчета матрицы расстояний между 4-мя векторами

Представление данных

Вектора:

$$X = \{x_1, x_2, \dots, x_n\}$$

$$Y = \{y_1, y_2, \dots, y_n\}$$

$$Z = \{z_1, z_2, \dots, z_n\}$$

$U = \{u_1, u_2, \dots, u_n\}$ где u - число признаков, x_i, y_i, z_i, u_i — значение i -го признака для векторов X, Y, Z, U соответственно.

Результаты

Матрица расстояний M между векторами X, Y, Z, U:

$$M = \begin{bmatrix} d_{XX} & d_{XY} & d_{XZ} & d_{XU} \\ d_{YX} & d_{YY} & d_{YZ} & d_{YU} \\ d_{ZX} & d_{ZY} & d_{ZZ} & d_{ZU} \\ d_{UX} & d_{UY} & d_{UZ} & d_{UU} \end{bmatrix}$$

Ввод программы

| | | | | | | | | | | | |
|-----|-------|-------|-------|--------|--------|--------|--------|------|--------|------|------|
| В/О | C_x | $P,$ | $P,$ | ПП | $F5$ | $F4$ | ПП | $P,$ | $F8$ | $P,$ | P |
| P | $P6$ | C_x | $C/П$ | FC_x | $P8$ | $P8$ | FC_x | $P6$ | — | $P6$ | PP |
| ПП | $P,$ | $P,$ | $F2$ | $F4$ | ПП | ПП | $F4$ | БП | Fx^2 | В/О | |
| | C_x | C_x | $P7$ | $P8$ | FC_x | FC_x | $P7$ | $F2$ | ↑ | | |
| | $P,$ | $P,$ | $F3$ | ПП | $F3$ | $F5$ | ПП | $F7$ | $F6$ | | |
| | C_x | C_x | $P8$ | FC_x | $P7$ | $P8$ | FC_x | ↑ | + | | |

Вычисление

1. В/О С/П (4 сек).
2. $x_i P2$ $y_i P3$ $z_i P4$ $u_i P5$ С/П ($i = 1, 2, \dots, n$; 18 сек)
3. $F\sqrt{\quad}$ (получим расстояние d_{yz}), $P, F\sqrt{\quad}$ (расстояние d_{xz}), $P, F\sqrt{\quad}$ (расстояние d_{xu}), $P, F\sqrt{\quad}$ (d_{yz}), $P, F\sqrt{\quad}$ (d_{yu}) $P, F\sqrt{\quad}$ (d_{zu})

Расстояния главной диагонали равны нулю, остальные расстояния получаем из соображений симметрии ($d_{ij} = d_{ji}$)

Контрольный пример

Исходные данные:

$$X = \{1.7.3\}$$

$$Y = \{2.4.4\}$$

$$Z = \{1.1.3\}$$

$$U = \{5.5.2\}$$

[246]

Результаты после округления

| | | | | |
|----|-------|-------|-------|-------|
| M= | 0 | 3.317 | 6.000 | 4.582 |
| | 3.317 | 0 | 3.317 | 3.742 |
| | 6.000 | 3.317 | 0 | 5.742 |
| | 4.582 | 3.742 | 5.745 | 0 |

Формулы расчета

$$d_{xy} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Остальные расстояния вычисляются аналогично.

Приведенные программы дают достаточно полное представление о возможностях ПК ЭВМ. Мы полагаем, что использование ПК ЭВМ для вторичной обработки социологической информации, особенно для расчета уровня значимости различий, целесообразно при любых формах организации обработки, в том числе и при наличии собственного ВЦ с диалоговым режимом работы. Дальнейшее совершенствование ПК ЭВМ (увеличение объема памяти и быстродействия, и фиксирование программ на магнитных пластинках, избавляющие от необходимости вводить при включении программу заново) существенно расширит сферу применения микрокалькуляторов. Простота обращения, не предполагающая специальной подготовки, а также возможность использования их в качестве обычного микрокалькулятора, позволяет предположить, что программируемые клавишные ЭВМ станут для социологов таким же настольным средством анализа информации, как микрокалькуляторы.

[247]