

## ТЕМА 4

# АНАЛИЗ ДВУМЕРНОЙ СВЯЗИ. ТАБЛИЦЫ СОПРЯЖЕННОСТИ. КОЭФФИЦИЕНТЫ ПАРНОЙ СВЯЗИ

Лабораторные работы 7 и 8

# Лабораторная работа №7

## Таблицы сопряженности

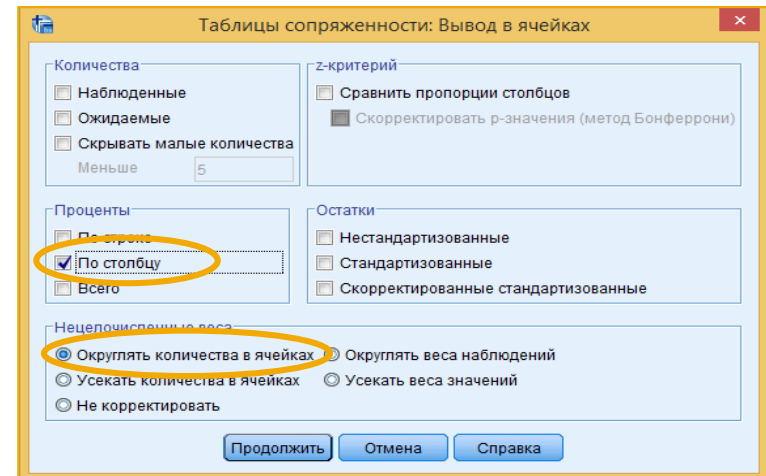
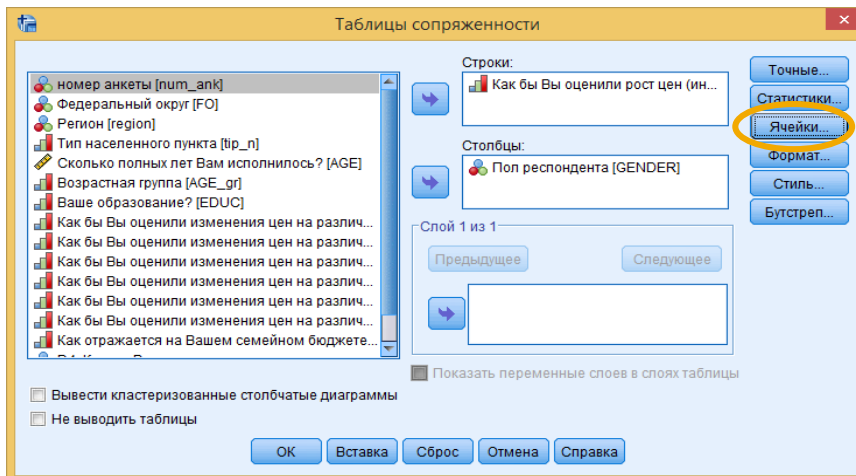
### Задание:

1. Узнать, как оценивают мужчины и женщины рост цен (инфляцию) в течение последнего месяца-двух. Определить, в каких возрастных группах мужчины и женщины оценивают рост цен (инфляцию) в течение последнего месяца-двух выше всего.
2. Выявить, носят ли различия в ответах статистически значимый характер.

**Рабочий файл:** inflation.sav

## Лабораторная работа №7

1. Запустите SPSS и откройте файл inflation.sav. Чтобы узнать, как оценивают мужчины и женщины рост цен (инфляцию) в течение последнего месяца-двух построим кросс-таблицу двух переменных – пол (**GENDER**) и рост цен (инфляцию) в течение последнего месяца-двух (**inflation**).
2. «Анализ» → «Описательные статистики» → «Таблицы сопряженности». Так как пол не может изменяться от оценки инфляции, он будет независимой переменной, которую переносим в столбцы. Переменную **inflation** переносим в строки.
3. Переходим в «Ячейки» и убираем вывод наблюдаемых частот в количестве респондентов («**observed**») и выбираем проценты по столбцам («**column**»), чтобы распределение оценки было посчитано от 100% мужчин в одном столбце и от 100% женщин в другом.



## Лабораторная работа №7

4. Получаем в окне вывода кросс-таблицу, в которой видно, что 39,8% мужчин считают инфляцию очень высокой, среди женщин таких - 43,3%. Практически половина опрошенных и мужчин, и женщин считают что инфляция умеренная. Но значимы ли эти различия или это результат случайного распределения?

|   |                         | Пол респондента |         | Всего  |
|---|-------------------------|-----------------|---------|--------|
|   |                         | Мужской         | Женский |        |
| Как бы Вы оценили рост цен (инфляцию) в течение последнего месяца-двух? | Инфляция очень высокая  | 39,8%           | 43,3%   | 41,8%  |
|   | Инфляция умеренная      | 43,8%           | 46,1%   | 45,1%  |
|   | Инфляция незначительная | 16,3%           | 10,6%   | 13,1%  |
| Всего   |                         | 100,0%          | 100,0%  | 100,0% |

Рассчитаем Хи-квадрат, нажав на кнопку «Статистика» в диалоговом меню кросс-таблицы («Анализ» → «Описательные статистики» → «Таблицы сопряженности»).



Критерии хи-квадрат

|                                  | Значение            | ст.св. | Асимптотическая значимость (2-сторонняя) |
|----------------------------------|---------------------|--------|--|
| Хи-квадрат Пирсона               | 10,120 <sup>a</sup> | 2      | ,006                                     |
| Отношения правдоподобия          | 10,056              | 2      | ,007                                     |
| Линейно-линейная связь           | 6,287               | 1      | ,012                                     |
| Количество допустимых наблюдений | 1401                |        |  |



Таблицы сопряженности: Статистики

Хи-квадрат

Корреляции

Номинальная

- Коэфф. сопряженности
- Фи и V Крамера
- Лямбда
- Коэфф. неопределенности

Порядковая

- Гамма
- d Сомерса
- Тау-б Кендалла
- Тау-с Кендалла

Номинал./интерв.

- Эта

Каппа

Риск

Макнемара

Статистики Кокрена и Мантеля-Хенцеля

Проверяемое общее отношение шансов равно: 1

Продолжить Отмена Справка

Хи-квадрат = 10,120 при уровне Sig. < 0,05. Следовательно, отличия в оценках уровня инфляции мужчин и женщин носят статистически значимый характер.

## Лабораторная работа №7

- Чтобы определить, в каком возрасте мужчины и женщины оценивают рост цен (инфляцию) в течение последнего месяца-двух очень высоким, находим переменную AGE\_gr с 5 возрастными группами (18-24; 25-34; 35-44; 45-54 и 55+).
- После этого через меню **«Анализ»** → **«Таблицы»** → **«Настраиваемые таблицы»** запускаем конструктор кросс-таблиц. К левому краю **«Строки»** перетягиваем переменную inflation, к верхней границе – сначала AGE\_gr, а ниже – GENDER (как показано на рисунке). Установите расчет **«Столбцы»**, дважды щелкнув по одной из переменных таблицы.

Настраиваемые таблицы

Таблица | Заголовки | Статистические критерии | Параметры

Переменные: Нормальный Компактный Спой

Задать статистику для итогов и подытогов

|                                |             | Пол респондента   |         |         |         |         |                   |         |         |         |         |
|--------------------------------|-------------|-------------------|---------|---------|---------|---------|-------------------|---------|---------|---------|---------|
|                                |             | Мужской           |         |         |         |         | Женский           |         |         |         |         |
|                                |             | Возрастная группа |         |         |         |         | Возрастная группа |         |         |         |         |
|                                |             | 18-24             | 25-34   | 35-44   | 45-54   | 55+     | 18-24             | 25-34   | 35-44   | 45-54   | 55+     |
| Как бы Вы оценили рост цен ... | Инфляция .. | nnnn,n%           | nnnn,n% | nnnn,n% | nnnn,n% | nnnn,n% | nnnn,n%           | nnnn,n% | nnnn,n% | nnnn,n% | nnnn,n% |
|                                | Инфляция .. | nnnn,n%           | nnnn,n% | nnnn,n% | nnnn,n% | nnnn,n% | nnnn,n%           | nnnn,n% | nnnn,n% | nnnn,n% | nnnn,n% |

Категории:

Задать

Итожащие статистики:  Итожащие статистики...  Скрыть

Расположение: Столбцы

Расположение категорий: По умолчанию

Источник: Переменные столбцов

OK Вставка Сброс Отмена Справка

После построения кросс-таблицы будет видно, что самые высокие оценки характерны для мужчин 55+ лет (49,7%) и женщин 55+ лет (50,4%).

# Лабораторная работа №8

## **Проверка наличия взаимосвязи между переменными**

## Лабораторная работа №8

---

### **Задание:**

1. Проверить наличие статистической связи между полом респондента и влиянием на семейный бюджет повышения цен на продукты питания.
2. При наличии связи определить ее тесноту.

**Рабочий файл:** inflation.sav



## Лабораторная работа №8

1. Проверим гипотезу об отсутствии связи, используя критерий хи-квадрат. Если нулевая гипотеза не отвергается, то связи между переменными нет.

- $H_0$ : Нет связи между признаками
- $H_1$ : Связь статистически значима

Комбинационная таблица Как отражается на Вашем семейном бюджете повышение цен на продукты питания? \* Пол респондента

% в Пол респондента

|   |  | Пол респондента |         | Всего  |
|---|--|-----------------|---------|--------|
|   |  | Мужской         | Женский |        |
| Как отражается на Вашем семейном бюджете повышение цен на продукты питания? | Для нашей семьи это очень существенно, нам придется ограничивать себя в тратах на продукты | 23,5%           | 31,4%   | 27,9%  |
|   | Это наносит урон нашему бюджету, но на нашем питании это не отразится                      | 36,3%           | 39,8%   | 38,2%  |
|   | Это не существенно влияет на наш бюджет  | 28,9%           | 23,3%   | 25,8%  |
|   | Такое повышение цен для нашего бюджета значения не имеет                                   | 11,3%           | 5,5%    | 8,1%   |
| Всего   |  | 100,0%          | 100,0%  | 100,0% |

Критерии хи-квадрат

|                                  | Значение            | ст.св. | Асимптотическая значимость (2-сторонняя) |
|----------------------------------|---------------------|--------|--|
| Хи-квадрат Пирсона               | 29,061 <sup>a</sup> | 3      | ,000                                     |
| Отношения правдоподобия          | 29,068              | 3      | ,000                                     |
| Линейно-линейная связь           | 27,461              | 1      | ,000                                     |
| Количество допустимых наблюдений | 1496                |        |  |

а. Для числа ячеек 0 (0,0%) предполагается значение, меньше 5. Минимальное предполагаемое число равно 53,71.

**Связь является статистически значимой**

## Лабораторная работа №8

2. Если связь есть, то можно определить тесноту связи, используя подходящие для номинальных переменных статистики («Анализ» → «Описательные статистики» → «Таблицы сопряженности» → «Статистика»).

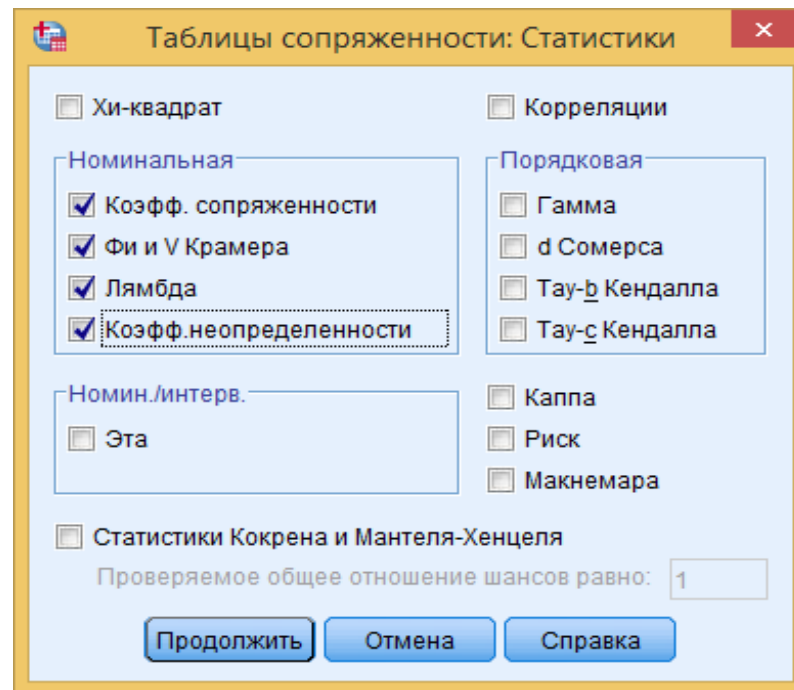
### Коэффициент сопряженности Пирсона

(contingency coefficient) - мера тесноты связи в таблицах сопряженности любого размера. Равен нулю для независимых переменных, максимум = 1.

**Фи-коэффициент (Phi coefficient,  $\phi$ )** используют для измерения тесноты связи переменных в таблицах 2x2. Равен нулю, если связь отсутствует. Равен плюс или минус единице для полностью зависимых переменных.

**V-коэффициент Крамера (Cramer's-V)** – мера тесноты связи, используемая в таблицах сопряженности любого размера (для таблиц 2x2 совпадает с фи). Равен 0 для независимых переменных, равен 1 для полностью зависимых переменных.

**Lambda** использует прогноз, основанный на выборе модального значения. Равен 1 для полностью зависимых переменных, однако может быть равен 0, даже когда переменные зависимы.



## Лабораторная работа №8

---

3. Определить тесноту позволяют полученные значения коэффициентов, равные около 0,1. Их величина позволяет сказать, что переменные обладают связью, но ее нельзя назвать тесной.

### Симметричные меры

|                                  | Значение | Примерная<br>Знач. |
|----------------------------------|----------|--------------------|
| Номинал/номинал    Фи            | ,139     | ,000               |
| V Крамера                        | ,139     | ,000               |
| Количество допустимых наблюдений | 1496     |                    |