

Таблица сопряженности (Contingency table)

Синонимы: Таблица контингентности, Факторная таблица

В статистике — метод представления многомерного частотного распределения двух признаков, предназначенный для исследования статистической связи между ними. Широко используется в областях, где требуется статистическая обработка результатов исследований и построение аналитических моделей, особенно в бизнес-аналитике. Впервые термин был введен Карлом Пирсоном в 1904 году.

Например, в представленной ниже таблице сопряженности показано распределение заемщиков, допустивших и не допустивших просрочку, по переменной «Пол».

Пол\Просрочка	Допущена	Не допущена	Всего
Мужской	12	40	52
Женский	4	44	48
Всего	16	84	100

Таблицы сопряженности являются удобным средством изучения статистических зависимостей. Они часто используются для проверки гипотез о наличии связи между двумя признаками с использованием теста Фишера или критерия согласия Пирсона хи-квадрат.

По строкам таблицы сопряженности размещаются значения одного признака, а по столбцам — другого (если признаки числовые, то они должны быть предварительно преобразованы к интервальному представлению). На пересечении строки и столбца указывается частота совместного появления f_{ij} соответствующих значений признаков x_i и y_j .

Сумма частот по строке f_i называется **маргинальной частотой строки**, а по столбцу f_j — **маргинальной частотой столбца**. Сумма маргинальных частот равна размеру выборки n , а их распределение представляет собой одномерное распределение переменной, образующей строки или столбцы таблицы.

В таблицах сопряженности могут быть представлены как абсолютные, так и относительные частоты (в долях или процентах). Относительные частоты могут рассчитываться по отношению:

- к маргинальной частоте по строке;
- к маргинальной частоте по столбцу;
- к объему выборки.