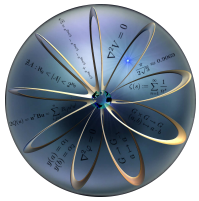

planetmath.org

Math for the people, by the people.

contingency table

Canonical name	ContingencyTable
Date of creation	2013-03-22 14:32:53
Last modified on	2013-03-22 14:32:53
Owner	CWoo (3771)
Last modified by	CWoo (3771)
Numerical id	11
Author	CWoo (3771)
Entry type	Definition
Classification	msc 62H17

Given a random sample of N observations $\mathbf{Z}_i = (Y_i, X_{i1}, \ldots, X_{ik})$ where

1. the response variables Y_i are identically distributed as Y
2. Y is categorical in nature (coming from a multinomial distribution)
3. each of the explanatory variables X_{ij} is categorical in nature

Then we can analyze the data by forming a *contingency table*. The table is customarily formed by labeling the categories for the response across the top, and then the combinations of the levels for each explanatory variable down the left-most columns. Then the cells are filled with counts or frequencies of occurrences corresponding to the specific explanatory variable level combination to the left and the response to the top.

The simplest example of a contingency table is where the response variable Y comes from a binomial distribution (with two possible responses r_1 and r_2) and there is only one explanatory variable X, which has only two levels, A_1 and A_2. This is an instance of a *2 way contingency table*:

	r_1	r_2
A_1	n_{11}	n_{12}
A_2	n_{21}	n_{22}

where n_{ij} corresponds to the count or frequency of level A_i and response r_j.

Example A penny P and a quarter Q are each tossed separately 100 times. The outcome for each toss is recorded, H for head and T for tail. The numbers of heads and tails obtained from the tosses are recorded in the following 2 way table:

Coin Type	H	T
P	45	55
Q	56	44

A *3 way contingency table* consists of one response variable and two explanatory variables.

Example Four dice are used in an experiment to test whether they are more or less the "same" (having the same probability distribution). Each die comes from a combination of one of two casinos, C_1 and C_2, made by

one of two manufacturers, M_1 and M_2. Furthermore, no two dice have the same combination. Now, each dice is tossed 120 times and the outcomes (1 through 6) are recorded. We have the following contingency table:

Casino	Manufacturer	1	2	3	4	5	6
C_1	M_1	17	21	19	20	22	21
	M_2	22	20	20	19	21	18
C_2	M_1	21	18	18	23	21	19
	M_2	20	19	18	21	21	21

The explanatory variables are the casinos (C's) and the manufacturers (M's), and the response variable is the number appearing on the top face of a thrown die (1 through 6).