One-way Frequency Tables

One-way frequency refers to a tabulation of the data which only examines one categorical variable at a time. The Procedure FREQ can preform the tabulation of this simple structure, and produce tests for equal proportions across the categories. The most common statistical method to perform the specified test is the one-way frequency P^2 statistic. PROC FREQ can compute this.

As an example of this, we will look at the null hypothesis that the regions in the sample data have an equal number of respondents. A summary of the data reveals that the number of observations for each region is:

CN	EA	MW	RM	PC
11	60	47	35	39

The Pearson P^2 statistic is based on the formula $(O_i - E_i)^2/E_i$, where O_i and E_i are the observed and expected counts for the ith category, respectively. The expected count is based on the null hypothesis of equal proportions among the categories and is simply

SAS Work Shop
Categorical Data Procedures
Handout #3

Statistical Programs
College of Agriculture

HTTP://WWW.UIDAHO.EDU/AG/STATPROG

an average given by (11+60+47+35+39)/5 = 38.4. The ratios for each category are then added to form the P^2 statistic. The appropriate SAS codes to do this follows:

```
DATA REGION;
     INPUT REGION $ COUNT;
     CARDS;
CN 11
EA 60
MW 47
RM 35
PC 39
;

PROC FREQ;
     WEIGHT COUNT;
     TABLES REGION / CHISQ;
```

The results of the sample code are printed below:

REGION	Frequency	Percent	Cumulative Frequency	Cumulative Percent
CN	11	5.73	11	5.73
EA	60	31.25	71	36.98
MW	47	24.48	118	61.46
PC	39	20.31	157	81.77
RM	35	18.23	192	100.00

SAS Work Shop
Categorical Data Procedures
Handout #3

Statistical Programs
College of Agriculture

HTTP://WWW.UIDAHO.EDU/AG/STATPROG

```
        Chi-Square Test

    for Equal Proportions

    ----------------------------

Chi-Square      33.9375

DF              4

Pr > ChiSq      <.0001


    Sample Size = 192
```

Here, SAS reports the counts (frequencies) and percentages for each category along with their cumulative counterparts. Following this is the test information. The chi-squared value is given with it's associated degrees of freedom (equal to the number of categories minus one) and the p-value for the test. In this case, the p-value is quite small. Therefore, the null hypothesis is rejected and the conclusion is made that the proportions among regions are not equal. This should not be unexpected given the values for CN and EA.