## SPSS: 2x2 Pearson's Chi-Square Test of Independence

Application: To test if two or more populations or subpopulations have different patterns of response to a qualitative/categorical DV. This can also be characterized as a test of a pattern of relationship between two qualitative/categorical variables.

Research Hypothesis: The researcher hypothesized that those stores that did not have a separate reptile department would tend to display only freshwater fish, whereas those stores that did have a separate reptile department would tend to display both freshwater and saltwater fish.

H0: There is no pattern of relationship between whether or not pet stores have separate reptile departments and whether they display only freshwater fish or both saltwater and freshwater fish in the population represented by these pet stores.

## Analyze $\rightarrow$ Descriptive Statistics $\rightarrow$ Crosstabs

- highlight the variable you want to define the rows (be sure it is qualitative/categorical) and click arrow
- highlight the variable you want to define the columns (be sure it is qualitative/categorical) and click arrow
- "Statistics" - check that you want a "Chi-square analysis"



## SPSS Syntax

CROSSTABS

| /TABLES=reptdept BY fishdept | $\leftarrow$ "row variable" BY "column variable" |
| :--- | :--- |
| ISTATISTICS=CHISQ | $\leftarrow$ get Chi-square significance test |
| /CELLS=COUNT ROW COLUMN TOTAL. | $\leftarrow$ get various row, column and/or total cell percentages (optional) |

Chi-Square Tests

|  |  |  |  |
| :--- | ---: | ---: | ---: |
|  | Value | df | Asymp. Sig. <br> (2-sided) |
| Pearson Chi-Square | $5.333^{\text {a }}$ | 1 | .021 |
| Continuity Correction | 3.000 | 1 | .083 |
| N of Valid Cases | 12 |  |  |

a. 4 cells (100.0\%) have expected count less than 5. The
minimum expected count is 3.00 .
b. Computed only for a $2 \times 2$ table

The p-value of .021 means that there is about a $2.1 \%$ chance that this result is a Type I error.

Remember, even if the printout shows it, never report $p=.000$, because that would suggest there is no possibility of a Type 1 error. Instead, report "p < .001"

Chi-square results are "suspicious" if more than $15 \%$ of the cells have expected frequencies less than 5.

The "correction for continuity" is computed for $2 \times 2$ designs, to give a better estimate of Type I error for this small design, especially when N is small.

## type or reptile department * type of fish available Crosstabulation

Count

|  | type of fish available |  | Total |
| :---: | :---: | :---: | :---: |
|  | freshwater only | fresh and saltwater |  |
| type or reptile department not separate | 5 | 1 | 6 |
| separate dept | 1 | 5 | 6 |
| Total | 6 | 6 | 12 |


| Reporting the Results: |  |  |  |
| :---: | :---: | :---: | :---: |
| Table 1 <br> Relationship between Type of Reptile Department and Type of Fish Available |  |  |  |
|  | Type of Reptile Department |  |  |
| Type of Fish Available | Notseparate Department | Separate Department | total |
| Freshwater Fish Only | 5 | 1 | 6 |
| Fresh- and Saltwater Fish | 1 | 5 | 6 |
| total | 6 | 6 | 12 |

Table 1 shows the contingency table for these variables. The sample of stores was evenly divided between the two types of reptile departments and also evenly divided between the two types of fish departments. Inspection of the table suggests that, as hypothesized, those stores with separate reptile departments tended to have both fresh- and saltwater fish, whereas, those stores without separate reptile departments tended to have only freshwater fish. However, contrary to the hypothesis, there was not a statistically significant relationship between the variables, $\mathrm{X}^{2}(1)=$ $3.00, p=.083$, when the correction for continuity was applied to this $2 \times 2$ design.

