Advanced Chi-Square Workshop

Meredith Dodd <meredith.dodd@woz-u.com>

Thu 7/30/2020 11:37 AM

To: Meredith Dodd <meredith.dodd@woz-u.com>

```
# Is there any difference in the type of cereal and the manufacturer? Independent chi-square

library("gmodels")
library("dplyr")

CrossTable(cereal$mfr, cereal$type, fisher=TRUE, chisq=TRUE, expected=TRUE, sresid = TRUE, format="SPSS")

# A is much more likely to produce a hot cereal than any other manufacturer.

# From my experience, there are not many makers of hot cereal (oatmeal) or many different types, so I expect that there is a 70/30 ratio of cold to hot cereal. Let's test that with an independent chi-square

## Data Wrangling

cereal %>% group_by(type) %>% summarise(count=n())

## Run the analysis

observed = c(74, 3)
expected = c(0.7, 0.3)
chisq.test(x=observed, p = expected)

## Looks like they are significantly different from that

# McNemar Chi-Square - look at whether the number of homes that have upholstery changed over time from 1700 to 1770

CrossTable(upholstery$TimePoint, upholstery$Upholstery, fisher=TRUE, chisq=TRUE, mcnemar=TRUE, expected=TRUE, sresid=TRUE, format="SPSS")

## Although it is significant, when we actually look at the standardized residauls, there's nothing over the absolute value of 2, so not really
```

Meredith Dodd, Ph.D. | Data Science Program Chair and Instructor
meredith.dodd@woz-u.com
o: 480-291-8068
powered by
WOZ U
https://woz-u.com