

1 Introduction

We want to examine nutrition by income group to determine whether we have evidence from the BRFSS dataset that those in higher income brackets tend to have better nutrition.

2 Questions

Question 1. This question will focus on a particular aspect of nutrition: whether an individual drinks regular soda or pop that contains sugar. Our goal is to understand whether the likelihood of drinking soda or pop with sugar is correlated with income category. The variables of interest are *income_cat* and *ssbsugr2_cat*:

- *income_cat* records a value of 1 if an individual has income of \$25,000 or below, a value of 2 if an individual has income between \$25,000 and \$75,000, and a value of 3 if an individual has income above \$75,000.
- *ssbsugr2_cat* records a value of 0 if an individual does not drink soda or pop, 1 if an individual drinks one soda or pop per day, and 2 if an individual drinks more than one soda or pop per day.

(Part a) Create a twoway relative frequency table for the *income_cat* variable and the *ssbsugr2_cat* variable.

(Part b) State the null and alternative hypotheses in order to test whether there is an association between these two variables.

(Part c) Create a table for the expected counts.

(Part d) Find all nine contributions to the chi-squared test statistic and record them in a table.

(Part e) What is the chi-square test statistic?

(Part f) What are the degrees of freedom for the test? What is the p-value?

(Part g) Using a 5% significance level, what is the conclusion of the test? Be specific.

(Part h) If there is an association between drinking regular soda or pop that contains sugar and income, describe how the two variables are related.

Question 2. This question will focus on a different aspect of nutrition: how often an individual drinks sugar-sweetened fruit drinks. These types of drinks include Kool-aid, lemonade, sweet tea, ad sports or energy drinks such as Gatorade and Red Bull. The variables of interest are *income_cat* and *ssbfrut3_cat*:

- *income_cat* records a value of 1 if an individual has income of \$25,000 or below, a value of 2 if an individual has income between \$25, 000 and \$75,000, and a value of 3 if an individual has income above \$75,000.
- *ssbfrut3_cat* records a value of 0 if an individual does not drink sugary fruit drinks 1 if an individual drinks one sugary fruit drink per day, and 2 if an individual drinks more than one sugary fruit drink per day.

- (a) Create a twoway relative frequency table for the *income_cat* variable and the *ssbfrut3_cat* variable.
- (b) State the null and alternative hypotheses in order to test whether there is an association between these two variables.
- (c) Create a table for the expected counts.
- (d) Find all nine contributions to the chi-squared test statistic and record them in the table below.
- (e) What is the chi-square test statistic?
- (f) What are the degrees of freedom for the test? What is the p-value?
- (g) Using a 5% significance level, what is the conclusion of the test? Be specific.
- (h) If there is an association between drinking sugary fruit drinks and income, describe how the two variables are related.