STAT 630, HW 9
Due: Tues, November 16

Reading: *OpenIntro* Sections 6.3 and 6.4

Concept Questions

Please refer to lecture 13 notes.

Exercise 1. A news article reports that "Americans have differing views on two potentially inconvenient and invasive practices that airports could implement to uncover potential terrorist attacks." This news piece was based on a survey conducted among a random sample of 1,137 adults nationwide, interviewed by telephone November 7-10, 2010, where one of the questions on the survey was "Some airports are now using 'full-body' digital x-ray machines to electronically screen passengers in airport security lines. Do you think these new x-ray machines should or should not be used at airports?" Below is a summary of responses (answers to this question) based on party affiliation.[1]

	Republican	Democrat	Independent	Total
Should	264	299	351	914
Should not	38	55	77	170
Don't know/No answer	16	15	22	53
Total	318	369	450	1137

(a) Write the null and alternative hypotheses for a chi-square test of independence between party affiliation and opinion on full-body scans.

(b) Assuming independence, calculate the expected counts for the boxed cells.

(c) The test statistic is $\chi^2 = 4.36$. What is the p-value?

(d) What is the conclusion of this hypothesis test?

(e) The conclusion of the test may be incorrect, meaning a testing error was made. If an error was made, was it a Type I or a Type II Error?

Exercise 2. Microhabitat factors associated with forage and bed sites of barking deer in Hainan Island, China were examined from 2001 to 2002. In this region woods make up 4.8% of the land, cultivated grass plot makes up 14.7%, and deciduous forests makes up 39.6%. Of the 426 sites where the deer forage, 4 were categorized as woods, 16 as cultivated grassplot, and 61 as deciduous forests. The table below summarizes these data.[2]

Woods	Cultivated grassplot	Deciduous forests	Other	Total
4	16	61	345	426

Conduct a chi-square test of goodness-of-fit to test whether barking deer prefer to forage in certain habitats over others. Write the null and alternative hypotheses, tabulate the expected counts, report the test statistic and p-value, and make a conclusion. Also discuss whether the conditions for the test are satisfied.

[1] `https://www.cbsnews.com/news/poll-4-in-5-support-full-body-airport-scanners/`

[2] Liwei Teng et al. "Forage and bed sites characteristics of Indian muntjac (Muntiacus muntjak) in Hainan Island, China". In: *Ecological Research* 19.6 (2004), pp. 675–681.

Data Analysis Questions

Please refer to lab 8.

The following exercises use the `GSS2002` data set from the `resampledata` library.

```
library(resampledata)
data("GSS2002")
```

Exercise 3. Compute a 95% confidence interval for the difference between the proportion of males and females that favor the death penalty. Interpret the interval.

Exercise 4. Consider the categorical variables `Education` (highest level of education) and `SpendMilitary` (position on government spending on military).

(a) Make a contingency table between the two variables (the table should have 5 rows and 3 columns).

(b) Use `prop.table()` to make a contingency table of the row proportions (where the rows are the different categories for `Education`). Based on the table of row proportions, does there appear to be a relationship between the two categorical variables?

(c) Conduct a chi-square test of independence between `Education` and `SpendMilitary`. Write down the null and alternative hypotheses and your conclusion based on the p-value.

(d) Extract the table of expected counts for the chi-square test you conducted in part (c). Using this table, are the conditions for the test satisfied?