STAT 352: Data Modeling and Inference

Homework 7. Due Friday, March 11th. I encourage you to type all of your solutions, though this is not necessary. However, you must scan (or photograph) any handwritten portions and upload the files to Canvas. For questions that require `R` code, you must turn in your `R` code on Canvas. Your code must in a .Rmd file.

Question 1 (Computation): *Preamble:* You are going to repeat the analysis of the of the `UCBAdmissions` data, but disaggregated for each department.

Useful Functions: Refer to `UCBAdmissions.Rmd` and Lab 7.
Do the following: Load the data using the command `data(UCBAdmissions)`.

a) Perform a likelihood ratio test for independence on the data aggregated over departments at the $\alpha = 0.01$ level. Make sure to output and report the $p-$value.

b) Construct a 0.99 level confidence interval for the log-odds for the data aggregated over departments.

c) Perform a likelihood ratio test for independence at the $\alpha = 0.01$ level for each individual department. Make sure to output and report the $p-$values.

d) Construct a 0.99 level confidence interval for the log-odds for each individual department.

e) Comment on the results of your tests and intervals. Does anything seem wrong between the aggregated and disaggregated data?

Question 2 (Theory): *Preamble:* We want to think about multiple testing. Mostly because we just did multiple tests above. You are going to derive results about three multiplicity corrections. The Sidak correction, the Bonferroni correction, and the Holm correction. You can do the computational bit of these questions by hand once you have the $p-$values from Question 1c.

Family-Wise Error Rate (FWER): Suppose we have k tests with null hypotheses $H_{0,i}$ and alternative hypotheses $H_{A,i}$. Each test has a significance level α_i, which is the probability of rejecting $H_{0,i}$ even though it is true. The tests might not be independent. We are interested in the family-wise error rate, which is given by

$$FWER = P(\textit{Rejecting any } H_{0,i}; \textit{All the } H_{0,i} \textit{ are true}).$$

In some contexts, this is more easily thought of as

$$1 - FWER = P(\textit{Failing to reject all the } H_{0,i}; \textit{All the } H_{0,i} \textit{ are true}).$$

Probability Theory Refresher: Suppose that A and B are events with positive probability.

$$\begin{aligned}
P(A|B) &= \frac{P(A \text{ and } B)}{P(B)} \\
P(A \text{ and } B) &= P(A|B)P(B) \\
&= P(B|A)P(A) \\
P(A \text{ or } B) &= P(A) + P(B) - P(A \text{ and } B)
\end{aligned}$$

If A and B are independent, then

$$\begin{aligned}
P(A|B) &= P(A) \\
P(B|A) &= P(B) \\
P(A \text{ and } B) &= P(A)P(B) \\
P(A \text{ or } B) &= P(A) + P(B) - P(A)P(B)
\end{aligned}$$

If $\neg$ represents logical negation and ^{c} represents the set complement, it is natural and common to equate $\neg B = B^c$ (this really depends on the flavor of probability theory you grew up with; some work with Boolean statements while others work with sets). Assume that $P(B) > 0$ and that $P(B) < 1$. We have further rules (the last one of which needs $0 < P(B) < 1$),

$$\begin{aligned}
P(A^c) &= 1 - P(A) \\
P(A) &= P(A \text{ and } B) + P(A \text{ and } B^c) \\
P(A) &= P(A|B)P(B) + P(A|B^c)P(B^c) \\
P(B|A) &= \frac{P(A|B)P(B)}{P(A)} \\
&= \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}
\end{aligned}$$

Do the following: Suppose that we have k tests $\delta_1, \ldots, \delta_k$ with Type I error rates $\alpha_1, \ldots, \alpha_k$.

a) The probability function is sub-additive, $P(A \text{ or } B) \le P(A) + P(B)$. Use this to show that

$$FWER \le \alpha_1 + \cdots + \alpha_k.$$

Use this to show that if all the α_i are equal to α_0 then $FWER \le k\alpha_0$.
You need to assume that the probability of rejecting any one hypothesis when all of the hypotheses are true is the same as the probability of rejecting that one hypothesis when that one hypothesis is true. Basically, the truth value of other hypotheses does not effect the probability of rejecting a particular null hypothesis. This does not say that the tests are independent, the statistics used for the tests could have some stochastic dependence on each other. This is a statement that if I am thinking about a single test, then the other tests I could consider also doing do not effect it. This is a common indifference to irrelevant information assumption. We could try to state this in terms of conditional probability, but we really cannot condition on whether a hypothesis is true because it is not random.
Hint: The word "or" relates to the word "any" and the sum rule.

b) The Bonferroni correction sets the FWER at α and uses $\alpha_0 = \alpha/k$. Redo the tests for the disaggregated data using the $p-$values from Question 1c with the significance value $\alpha_0 = 0.01/6$ so that the FWER is controlled at level $\alpha = 0.01$. Do your substantive conclusions change?

c) Suppose that the tests $\delta_1, \ldots, \delta_k$ are independent. Show that

$$1 - FWER = (1 - \alpha_1) \times \cdots (1 - \alpha_k).$$

Use this to show that if all the α_i are equal to α_0 then $FWER = 1 - (1 - \alpha_0)^k$.
Hint: The word "and" relates to the word "all" and the product rule.

d) The Šidák correction sets the FWER at α and uses $\alpha_0 = 1 - (1 - \alpha)^{1/k}$. Redo the tests for the disaggregated data using the $p-$values from Question 1c with the significance value $\alpha_0 = 1 - 0.99^{1/6}$ so that the FWER is controlled at level $\alpha = 0.01$. Do your substantive conclusions change? Note that we are assuming that the tests are independent, though that might not actually be the case because all of the departments are from UC Berkeley.

e) One final multiple testing procedure for controlling FWER is the Holm procedure. This is sort of like an adaptive Bonferroni correction. The basic idea is to only use pieces of the significance level when thinking about null hypotheses that we have not already rejected. To do this, we need to sort the tests by increasing p-value. Let $\delta_{(1)}, \ldots, \delta_{(k)}$ be the sorted tests with increasing p-values $p_{(1)}, \ldots, p_{(k)}$. To do the Holm procedure, we set the FWER at α and perform the following steps:

1) If $p_{(1)} > \alpha/k$, fail to reject all null hypotheses.
2) Otherwise, let $1 < m \le k$ be the smallest index with $p_{(m)} > \alpha/(k - m + 1)$. Reject the null hypotheses for $i < m$ and fail to reject the null hypotheses for $i \ge m$. If no such m exists ($p_{(m)} \le \alpha/(k - m + 1)$ for all $m = 1, \ldots, k$), then reject all the null hypotheses.

Perform the Holm procedure on the six tests for independence from the departments using the $p-$values from Question 1c with FWER set at $\alpha = 0.01$. Do your substantive conclusions change?

Hint: To be more concrete about the Holm procedure. Order the p-values $p_{(1)} < p_{(2)} < p_{(3)} < \ldots < p_{(k)}$. You first compare $p_{(1)}$ to $\frac{\alpha}{k}$. If $p_{(1)} \geq \frac{\alpha}{k}$, then you fail to reject $H_{0,(1)}$ and stop and fail to reject for all of the tests. If $p_{(1)} < \frac{\alpha}{k}$, then you reject $H_{0,(1)}$ and continue making comparisons. The next comparison is $p_{(2)}$ to $\frac{\alpha}{k-1}$. If $p_{(2)} \geq \frac{\alpha}{k-1}$, then you fail to reject $H_{0,(2)}$ and stop and fail to reject for all of the tests whose p-values are bigger than $p_{(2)}$. If $p_{(2)} < \frac{\alpha}{k-1}$, then you reject $H_{0,(2)}$ and continue making comparisons. The next comparison is $p_{(3)}$ to $\frac{\alpha}{k-2}$. If $p_{(3)} \geq \frac{\alpha}{k-2}$, then you fail to reject $H_{0,(3)}$ and stop and fail to reject for all of the tests whose p-values are bigger than $p_{(3)}$. If $p_{(3)} < \frac{\alpha}{k-2}$, then you reject $H_{0,(3)}$ and continue making comparisons. And so on until you have stopped or are comparing $p_{(k)}$ to α.