# Final Exam

**Due: December 19, 2017**

Name: ________________________________

<table>
<thead>
<tr>
<th>1 (20 pts)</th>
<th>2 (20 pts)</th>
<th>3 (20 pts)</th>
<th>4 (20 pts)</th>
<th>5 (20 pts)</th>
<th><strong>Total (100 pts)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The questions refer to the data set Final_data_2017.txt available at:
http://wolfweb.unr.edu/homepage/zal/STAT453_Fall17/Final_data_2017.txt

This file contains a synthetic data set on a sample of air passengers. Each record corresponds to a single air trip and contains the following fields:

- **Airline** – airline selected for travel
- **Gender** – gender of the passenger
- **Destination** – destination type (international vs domestic)
- **Membership** – membership in the airline frequent flyer program
Problem 1. Present the data as a four-way contingency table.
Problem 2.

(a) Estimate the proportion of domestic trips with respect to all trips made by Alaska passengers.

\[ \hat{p}_{\text{Alaska domestic}} = \]

(b) Construct a 98% confidence interval for the true proportion of domestic trips with respect to all trips made by Alaska passengers

\[ < p_{\text{Alaska domestic}} < \]

(c) Test the hypothesis that the true proportion of domestic trip made by Alaska passengers is 0.7.

- Test:
- P-value:
- Decision:
Problem 3.

(a) Construct a three-way contingency table that cross-classifies the subjects according to Airline, Gender, and Destination.

(b) Use the table constructed in (a) to estimate the marginal and conditional odds ratios for the association between Gender and Destination. For each odds ratio, report the P-value of the exact Fisher test for independence.

\[ \theta_{\text{marginal}} = \quad , P = \]
\[ \theta_{\text{conditional, Alaska}} = \quad , P = \]
\[ \theta_{\text{conditional, United}} = \quad , P = \]

(c) Discuss the findings.
Problem 4.

(a) Construct a log-linear model to approximate the counts in the four-way table that describes the data set. (To receive a full credit, the model must beat the saturated model.)

Final model:

(b) Draw the independence graph for the model found in (a).
Problem 5. Use the model found in Problem 4 to construct a logistic regression model that predicts *Destination* using other necessary parameters as predictors.