January 25 MIE237 Tutorial Questions

This week we'll cover the test of independence. Questions 10.86 to 10.89 are suitable practice. The TA will select from these questions. See this week's lab for advice on how to solve them using R.

In addition, here are some questions from past tests and exams.

1. The 2015 exam started with a description of a situation where a gas company collected data on a sample of gas meters. Several variables were collected, including `rust` (low, medium, high) and `brand` (A or B).

This question considered the variables `rust` and `brand` just by themselves. Here is the contingency table for these two variables with observed cell counts given along with all but one expected cell counts.

In previous years different software was used, but it should be clear what it what below. If you like you can do the analysis using R with the file `meters.csv` that is part of this tutorial's repository.

```
Rows: rust   Columns: brand
               A       B     All
low            33      77    110
               33.55   *****

medium         59     138    197
               60.09  136.91

high           30      63     93
               28.36   64.64

All            122     278    400
```

(a) Provide a 95% confidence interval for the proportion of meters with "low" level of rust, being sure to also comment (very briefly!) on any required assumptions for the calculation to be accurate.

(b) Perform the hypothesis test that answers the question "are the rows and columns independent?" being sure to also comment (very briefly!) on any required assumptions for the calculation to be accurate.