Stat 204 *Exam 1* *Name:*

All statements should be expressed in complete sentences, and every numerical answer should be justified by showing how it was obtained. Writing R code that will actually calculate the numerical answer is definitely encouraged, followed by displaying the numerical answer your code returns. Recall that the point of this exam is to demonstrate your mastery of complete analytic processes.

I pledge that I have neither given nor received unauthorized aid on this exam. Pledged: .....

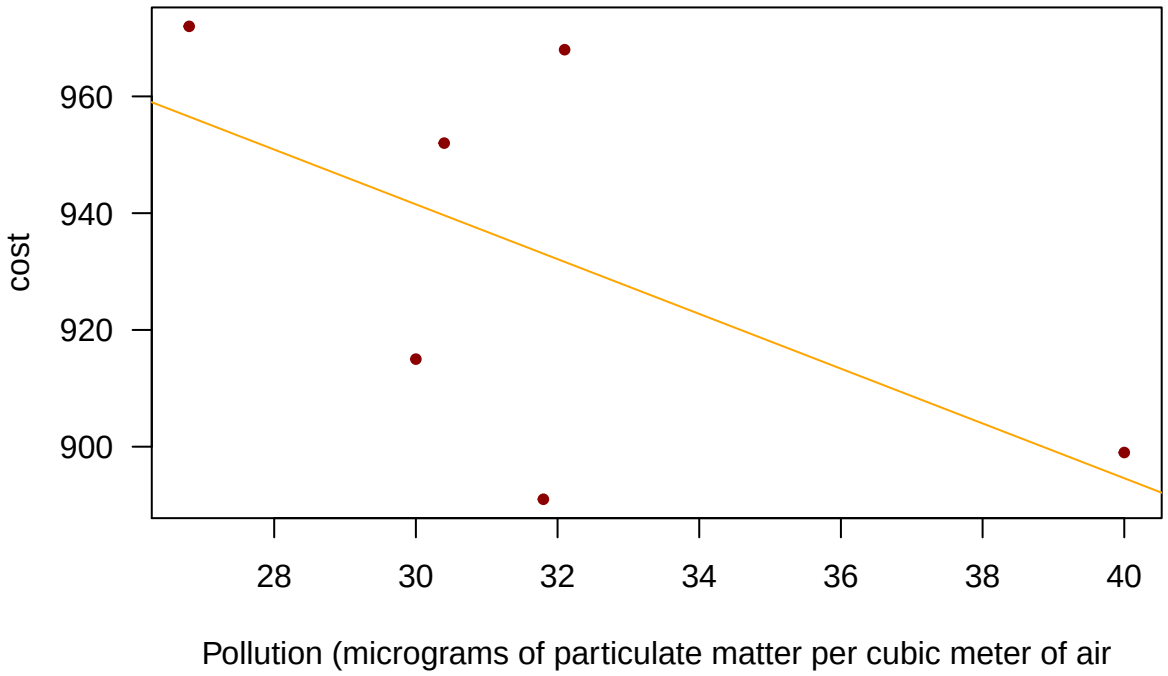
1. (Peck, 1.53, p.34; 4 points) Two hundred women were selected at random from the membership of the American Association of University Women (a large professional organization for women). Each woman selected was asked if she was married. The 160 married women were asked questions about church attendance and whether or not they would rate their marriage as happy. The proportion of those attending church regularly who rated their marriage as happy was significantly higher than the proportion of those not attending church regularly who rated their marriage as happy.

Peck, Roxy (2014-01-01). Statistics: Learning from Data (with JMP Printed Access Card) (Page 34). Cengage Learning. Kindle Edition.

   a. *Is the study described an experiment? If so, what is the explanatory variable and what is the response variable? If not, explain why it is not an experiment.*

   b. *From this study alone, is it reasonable to conclude that regular church attendance is the cause of the observed difference in rating marriages as happy.? Justify your answer.*

2. (Peck, 4.25, pp.185-186; 4 points) The article "Air Pollution and Medical Care Use by Older Americans" (Health Affairs [2002]: 207–214) gave data on a measure of pollution (micrograms of particulate matter per cubic meter of air) and the cost of medical care.



Peck, Roxy (2014-01-01). Statistics: Learning from Data (with JMP Printed Access Card) (Page 185). Cengage Learning. Kindle Edition.

a. *What is the equation of the least squares regression line?*

```
pollution.lm$coefficients
```

```
## (Intercept)   pollution
## 1082.244014   -4.691073
```

b. *What is the predicted cost of medical care for a region that has a pollution measure of 35?*

c. *What is the approximate change in the cost of medical care associated with a 1 unit increase in the measure of pollution?*

d. *Would you use the least squares regression line to predict the cost of medical care for a region that has a pollution measure of 60? Why or why not?*

3. (Peck, 5.49, p.255; 4 points) Lyme disease is the leading tick-borne disease in the United States and Europe. Diagnosis of the disease is difficult and is aided by a test that detects particular antibodies in the blood. The article "Laboratory Considerations in the Diagnosis and Management of Lyme Borreliosis" (American Journal of Clinical Pathology [1993]: 168–174) used the following notation:

   + represents a positive result on the blood test
   − represents a negative result on the blood test
   L represents the event that the patient actually has Lyme disease
   L^c represents the event that the patient actually does not have Lyme disease

Peck, Roxy (2014-01-01). Statistics: Learning from Data (with JMP Printed Access Card) (Page 255). Cengage Learning. Kindle Edition.

Create a hypothetical 1000 table based on the data in that article.

Lyme	L	L^c	Total
+	1.94	29.94	31.88
−	0.13	967.99	968.12
$Total$	2.07	997.93	1000

Using the information in this table, calculate the following probabilities.

a. $P(L)$

b. $P(+ \mid L)$

c. $P(+ \mid L^c)$

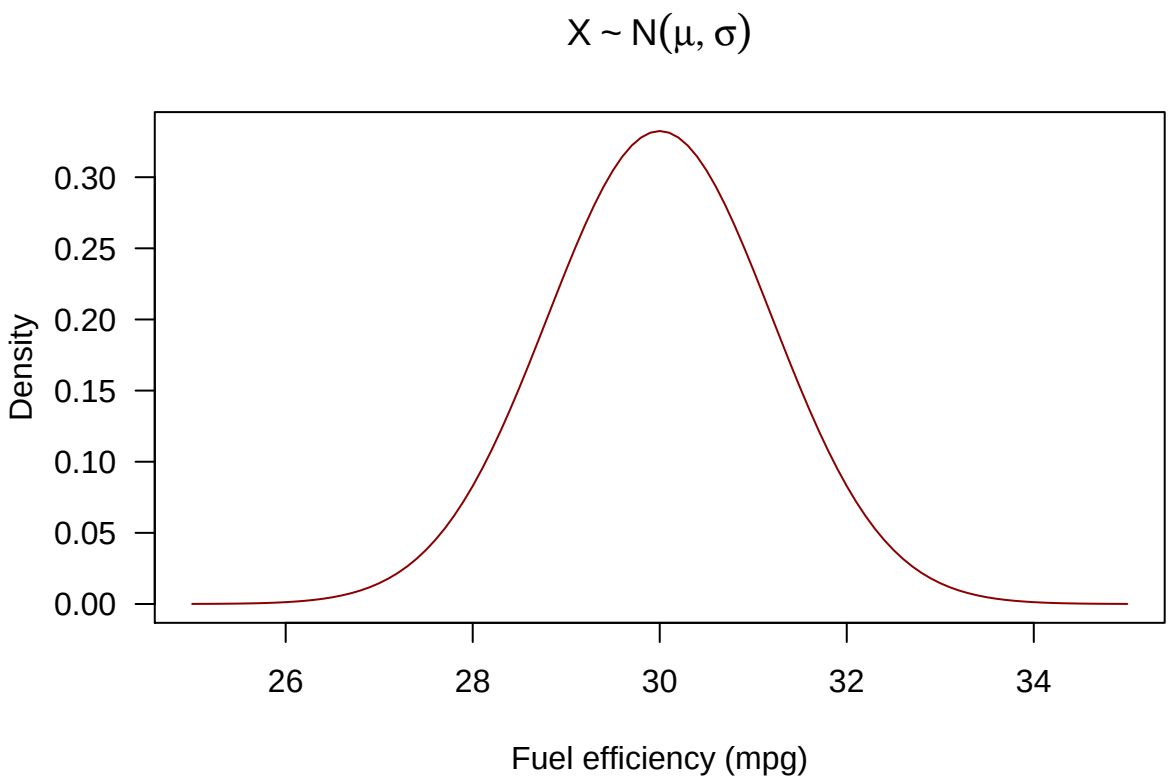
d. $P(L \mid +)$

4. (Peck, 6.58, p.314; 4 points) Suppose that fuel efficiency (miles per gallon, mpg) for a particular car model under specified conditions is normally distributed with a mean value of 30.0 mpg and a standard deviation of 1.2 mpg.

Peck, Roxy (2014-01-01). Statistics: Learning from Data (with JMP Printed Access Card) (Page 314). Cengage Learning. Kindle Edition.

Let X be a random variable with a normal distribution with mean 30.0 mpg and standard deviation 1.2 mpg,

$$X \sim N(\mu = 30.0, \sigma = 1.2)$$



a. *What is the probability that the fuel efficiency for a randomly selected car of this model is between 29 and 31 mpg?*

b. *Find a number x^* such that 95% of all cars of this model have efficiencies exceeding x^* (i.e., $P(x > x^*) = 0.95$).*

5. (Peck, 6.88, p.329; 4 points) You are to take a multiple-choice exam consisting of 10 questions with five possible responses to each question. Suppose that you have not studied and so you randomly select one of the five answers for each question. Let x represent the number of correct responses on the test.

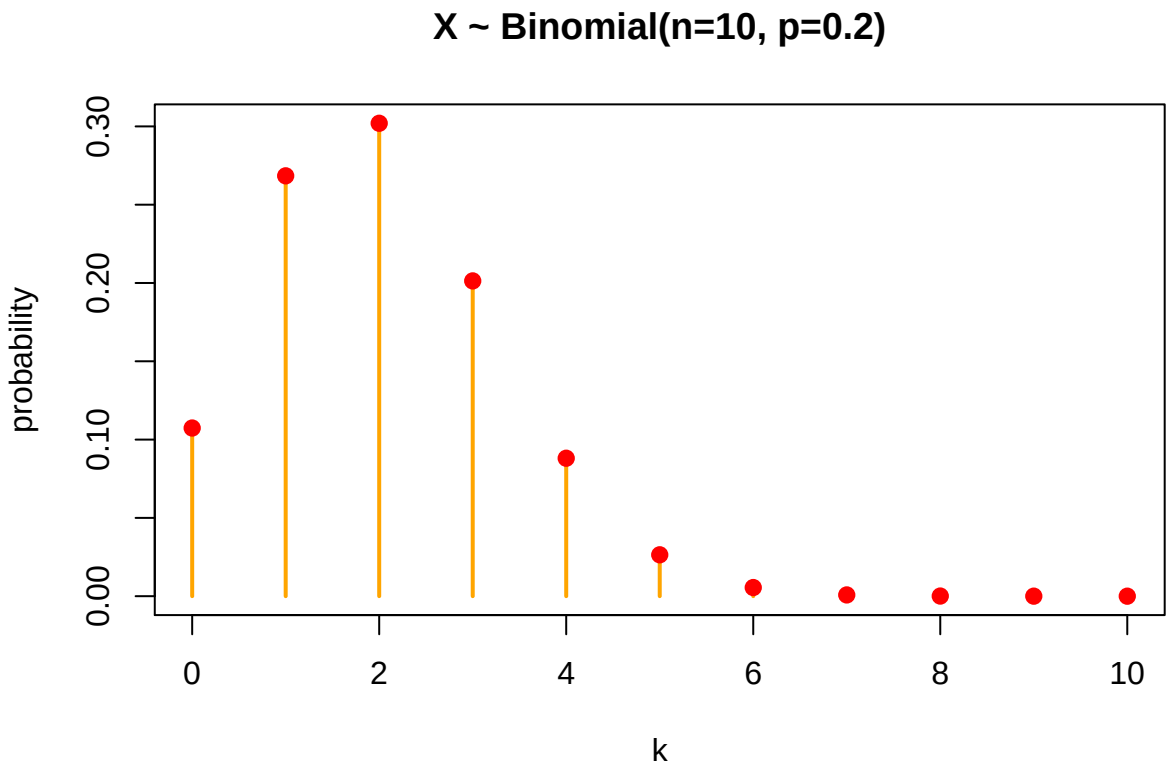
Peck, Roxy (2014-01-01). Statistics: Learning from Data (with JMP Printed Access Card) (Page 329). Cengage Learning. Kindle Edition.

```
n <- 10
p <- 0.20
```

Let X be a random variable having a binomial distribution with $n = 10$ and $p = 0.20$, so

$$X \sim Binomial(n, p).$$

Distribution.



a. *Calculate* $P(X \geq 7)$.

b. *Another student who studies diligently has an 85% chance of answering each question correctly. For this second student, calculate* $P(X \geq 7)$.