W271-2 – Spring 2016 – **Lab 1**

Juanjo Carin, Kevin Davis, Ashley Levato, Minghu Song

January 14, 2016

Contents

Part I: Marginal, Joint, and Conditional Probabilities

Question 1

In a team of data scientists, 36 are expert in machine learning, 28 are expert in statistics, and 18 are awesome. 22 are expert in both machine learning and statistics, 12 are expert in machine learning and are awesome, 9 are expert in statistics and are awesome, and 48 are expert in either machine learning or statistics or are awesome. Suppose you are in a cocktail party with this group of data scientists and you have an equal probability of meeting any one of them.

1. **What is the probability of meeting a data scientist who is an expert in both machine learning and statistics and is awesome?**

Let N be the size of the team of data scientists (i.e., the sample size), and M, S, A the event that a data scientist is either a machine learning expert, a statistics expert, or awesome.

$$\begin{aligned}
\Pr(\mathbf{M} \cap \mathbf{S} \cap \mathbf{A}) &= \Pr(M) + \Pr(S \cap A) - \Pr(M \cup (S \cap A)) \\
&= \Pr(M) + \Pr(S \cap A) - \Pr((M \cup S) \cap (M \cup A)) \\
&= \Pr(M) + \Pr(S \cap A) - ((\Pr(M \cup S) + \Pr(M \cup A) - \Pr((M \cup S) \cup (M \cup A)))) \\
&= \Pr(M) + \Pr(S \cap A) - (\Pr(M) - \Pr(S) + \Pr(M \cap S)) - (\Pr(M) - \Pr(A) + \Pr(M \cap A)) + \Pr(M \cup S \cup A) \\
&= \Pr(M \cap S) + \Pr(M \cap A) + \Pr(M \cap A) - \Pr(M) - \Pr(S) - \Pr(A) + \Pr(M \cup S \cup A) \\
&= \frac{22}{N} + \frac{12}{N} + \frac{9}{N} - \frac{36}{N} - \frac{28}{N} - \frac{18}{N} + \frac{48}{N} = \frac{(22+12+9) - (36-28-18) + 48}{N} = \frac{43 - 82 + 48}{N} \\
&= \mathbf{\frac{9}{N}}
\end{aligned}$$

2. **Suppose you meet a data scientist who is an expert in machine learning. Given this information, what is the probability that s/he is not awesome?**

$$\Pr(\mathbf{A^c} | \mathbf{M}) = 1 - \Pr(A|M) = 1 - \frac{\Pr(A \cap M)}{\Pr(M)} = 1 - \frac{\frac{12}{N}}{\frac{36}{N}} = 1 - \frac{12}{36} = 1 - \frac{1}{3} = \mathbf{\frac{2}{3}} = \mathbf{0.6667}$$

3. **Suppose the you meet a data scientist who is awesome. Given this information, what is the probability that s/he is an expert in either machine learning or statistics?**

$$\begin{aligned}
\Pr(\mathbf{M} \cup \mathbf{S} | \mathbf{A}) &= \frac{\Pr((M \cup S) \cap A)}{\Pr(A)} = \frac{\Pr((M \cap A) \cup (S \cap A))}{\Pr(A)} \\
&= \frac{\Pr(M \cap A) + \Pr(S \cap A) - \Pr((M \cap A) \cap (S \cap A))}{\Pr(S)} = \frac{\Pr(M \cap A) + \Pr(S \cap A) - \Pr(M \cap S \cap A)}{\Pr(S)} \\
&= \frac{\frac{12}{N} + \frac{9}{N} - \frac{9}{N}}{\frac{18}{N}} = \frac{12 + 9 - 9}{18} = \frac{12}{18} \\
&= \mathbf{\frac{2}{3}} = \mathbf{0.6667}
\end{aligned}$$

Question 2

Suppose for events A and B, $\Pr(A) = p \leq \frac{1}{2}, \Pr(B) = q$, where $\frac{1}{4} < q < \frac{1}{2}$. These are the only information we have about the events.

1. **What are the maximum and minimum possible values for $\Pr(A \cup B)$?**

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

The maximum value of $\Pr(A \cup B)$ occurs when $A \cap B$ is the smallest set possible. In this case, since $\Pr(A) \leq \frac{1}{2}$ and $\Pr(B) < \frac{1}{2}$, it would be $A \cap B = \varnothing$ (if A and B were disjoint sets, which might be the case), so:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) = \Pr(A) + \Pr(B) - \Pr(\varnothing) = p + q - 0 = p + q$$

which could have a maximum value close to 1 (i.e., $A \cup B \approx \Omega$), in case $\Pr(A) = \frac{1}{2}$ and $\Pr(B) \approx \frac{1}{2}$.

The minimum value of $\Pr(A \cup B)$ occurs when $A \cup B$ is the largest set possible, A or B. In this case, since $\Pr(A)$ does not have a lower bound, that would happen when $A \subseteq B$ and (consequently) $A \cap B = A$, which would lead to:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A) = \Pr(B) = q$$

whose minimum value is greater than $\frac{1}{4}$.

In summary,

$$\mathbf{\frac{1}{4}} < \Pr(\mathbf{A} \cup \mathbf{B}) < \mathbf{1}$$

2. **What are the maximum and minimum possible values for $\Pr(A|B)$?**

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

If $B \subseteq A$ (which would imply that the lower bound for p is also $\frac{1}{4}$), then:

$$\Pr(A|B) = \frac{\Pr(B)}{\Pr(B)} = 1$$

As seen in the previous part, since $\Pr(A) \leq \frac{1}{2}$ and $Pr(B) < \frac{1}{2}$, it might occur that $A \cap B = \varnothing$, and hence:

$$\Pr(A|B) = \frac{\Pr(\varnothing)}{\Pr(B)} = \frac{0}{q} = 0$$

irrespective of the value of q.

In summary,

$$\mathbf{0} \leq \Pr(\mathbf{A}|\mathbf{B}) \leq \mathbf{1}$$

Part II: Random Variables, Expectation, Conditional Exp.

Question 3

Suppose the life span of a particular server is a continuous random variable, t, with a uniform probability distribution between 0 and k year, where $k \leq 10$ is a positive integer.

The server comes with a contract that guarantees a full or partial refund, depending on how long it lasts. Specifically, if the server fails in the first year, it gives a full refund denoted by θ. If it lasts more than 1 year but fails before $\frac{k}{2}$ years, the manufacturer will pay $x = \$A(k-t)^{1/2}$, where A is some positive constant equal to 2 if $t \leq \frac{k}{2}$. If it lasts between $\frac{k}{2}$ and $\frac{3k}{4}$ years, it pays $\frac{\theta}{10}$.

1. **Given that the server lasts for $\frac{k}{4}$ years without failing, what is the probability that it will last another year?**
2. **Compute the expected payout from the contract, $E(x)$.**
3. **Compute the variance of the payout from the contract.**

Question 4

Continuous random variables X and Y have a joint distribution with probability density function $f(x, y) = 2e^{-x}e^{-2y}$ for $0 < x < \infty, 0 < y < \infty$ and 0 otherwise.

1. **Compute $\Pr(X > a, Y < b)$, where a, b are positive constants and $a < b$.**

$$\Pr(X > a, Y < b) = \int_{x=a}^{\infty} \int_{y=0}^{b} f(x, y) dx dy$$

$$= \int_{x=a}^{\infty} \int_{y=0}^{b} 2e^{-x}e^{-2y} dx dy = 2 \left(\int_{x=a}^{\infty} e^{-x} dx \right) \left(\int_{y=0}^{b} e^{-2y} dy \right)$$

$$= 2 \left[-e^{-x} \right]_{x=a}^{\infty} \left[-\frac{e^{-2y}}{2} \right]_{y=0}^{b} = \left[e^{-x} \right]_{x=a}^{\infty} \left[e^{-2y} \right]_{y=0}^{b} = (0 - e^{-a}) \left(e^{-2b} - 1 \right)$$

$$= \mathbf{e^{-a} \left(1 - e^{-2b} \right) = e^{-a} - e^{-a-2b}}$$

2. **Compute $\Pr(X < Y)$.**

$$\Pr(X < Y) = \int_{x=0}^{y} \int_{y=0}^{\infty} f(x, y) dx dy$$

$$= \int_{x=0}^{y} \int_{y=0}^{\infty} 2e^{-x}e^{-2y} dx dy = 2 \int_{y=0}^{\infty} \left(\int_{x=0}^{y} e^{-x} dx \right) e^{-2y} dy$$

$$= 2 \int_{y=o}^{\infty} \left[-e^{-x} \right]_{x=0}^{y} e^{-2y} dy = 2 \int_{y=0}^{\infty} \left(1 - e^{-y} \right) e^{-2y} dy = 2 \int_{y=0}^{\infty} \left(e^{-2y} - e^{-3y} \right) dy$$

$$= 2 \left[-\frac{e^{-2y}}{2} + \frac{e^{-3y}}{3} \right]_{y=0}^{\infty} = 2 \left[0 - \left(-\frac{1}{2} + \frac{1}{3} \right) \right] = 2 \left(\frac{1}{2} - \frac{1}{3} \right) = 1 - \frac{2}{3}$$

$$= \mathbf{\frac{1}{3}}$$

3. **Compute** $\Pr(X < a)$**.**

$$\begin{aligned}
\Pr(X < a) &= \int_{x=0}^{a}\int_{y=0}^{\infty} f(x,y)dxdy \\
&= \int_{x=0}^{a}\int_{y=0}^{\infty} 2e^{-x}e^{-2y}dxdy = 2\left(\int_{x=0}^{a} e^{-x}dx\right)\left(\int_{y=0}^{\infty} e^{-2y}dy\right) \\
&= 2\left[-e^{-x}\right]_{x=0}^{a}\left[-\frac{e^{-2y}}{2}\right]_{y=0}^{\infty} = \left[e^{-x}\right]_{x=0}^{a}\left[e^{-2y}\right]_{y=0}^{\infty} = \left(e^{-a}-1\right)(0-1) \\
&= \mathbf{1 - e^{-a}}
\end{aligned}$$

Question 5

Let X be a random variable and x be a real number. A linear function of the squared deviation from x is another random variable, $Y = a + b(X - x)^2$, where a and b are some positive constant.

1. **Find the value of x that minimizes $E(Y)$. Show that your result is really the minimum.**

The *Law of the unconscious statistician* states that:

$$E\left[g(X)\right] = \int g(x)f_X(x)dx$$

So, if we call $\mu = E(X)$ and $\sigma^2 = Var(X)$ (and knowing that $Var(X) = E\left(X^2\right) - \left(E(X)\right)^2 = E\left(X^2\right) - \mu^2$):

$$\begin{aligned}
E(Y) &= \int \left(a + b(X-x)^2\right) f(X)dX \\
&= \int \left(a + b(X^2 + x^2 - 2xX\right) f(X)dX \\
&= \left(a + bx^2\right)\int f(X)dX - 2bx\int Xf(X)dX + b\int X^2 f(X)dX \\
&= (a + bx^2)1 - 2bx\mu + b(\sigma^2 - \mu^2) \\
&= bx^2 - 2b\mu x + \left(a + b(\sigma^2 + \mu^2)\right)
\end{aligned}$$

Hence:

$$\frac{dE(Y)}{dx} = 2bx - 2b\mu = 2b(x-\mu) = 2b\left(x - E(X)\right)$$

And consequently:

$$\frac{dE(Y)}{dx} = 0 \Rightarrow \mathbf{x = E(X)}$$

2. **Find the value of $E(Y)$ for the choice of x you found in (1)?**

We just have to substitute in the last expression of $E(Y)

$$\begin{aligned}
\mathbf{E(Y)} &= \int \left(a + b(X-\mu)^2\right) f(X)dX \\
&= b\mu^2 - 2b\mu^2 + \left(a + b(\sigma^2 + \mu^2)\right) = -b\mu^2 + a + b\sigma^2 + b\mu^2 \\
&= \mathbf{a + b\sigma^2}
\end{aligned}$$

3. **Suppose $Y = ax + b(X - x)^2$. Find the values of x that minimizes $E(Y)$. Show that your result is really the minimum.**

$$\begin{aligned} E(Y) &= \int \left(ax + b(X-x)^2\right) f(X)dX \\ &= \int \left(ax + b(X^2 + x^2 - 2xX\right) f(X)dX \\ &= (ax + bx^2) \int f(X)dX - 2bx \int Xf(X)dX + b \int X^2 f(X)dX \\ &= (ax + bx^2)1 - 2bx\mu + b(\sigma^2 - \mu^2) \\ &= bx^2 + (a - 2b\mu)x + b(\sigma^2 + \mu^2) \end{aligned}$$

Hence:

$$\frac{dE(Y)}{dx} = 2bx + (a - 2b\mu) = 2b(x - \mu) + a = 2b\left(x - E(X)\right) + a$$

And consequently:

$$\frac{dE(Y)}{dx} = 0 \Rightarrow \mathbf{x = E(X) - \frac{a}{2b}}$$

Question 6

Suppose X and Y are independent continuous random variables, where both of which are uniformly distributed between 0 and 1. Let random variable $Z = X + Y$.

1. **Choose a value of z between 0 and 2, and draw a graph depicting the region of the $X - Y$ plane for which Z is less than z.**
2. **Derive the probability density function, $f(z)$.**

Note that

$$f_X(x) = f_Y(y) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & Otherwise \end{cases}$$

Because X and Y are independent, the probability density function of the sum of X and Y is equal to the convolution of $f_X(x)$ and $f_Y(y)$ described by

$$f_Z(z) = \int_0^1 f_X(z - y) f_Y(y) dy$$

If $0 \leq z \leq 1$

$$f_Z(z) = \int_0^z dy = z$$

And if $1 < z \leq 2$

$$f_Z(z) = \int_{z-1}^1 dy = 2 - z$$

This give us

$$f_Z(z) = \begin{cases} z & 0 \le z \le 1 \\ 2 - z & 1 < z \le 2 \\ 0 & Otherwise \end{cases}$$

Question 7

In a casino, you pay the following game. A pair of fair, ordinary 6-faced dice are rolled. If the sum of the dice is 2, 3, or 12, the house wins. If it is 7 or 11, you win. If it is any other number x, the house rolls the dice again until the sum is either 7 or x. If it is 7, the house wins. If it is x, you win. A game ends if one of the two players wins. Let Y be the number of rolls needed until the game ends.

1. **Is the expected number of rolls given that you win more than, equal to, or less than the expected number of rolls given that house wins (in a game)? The steps to arrive at your answer numerically need to be clearly shown.**

On the first roll, the house wins 4 games, so their are 36-4 = 32 games.

$$E(\text{Rolls}|\text{Win 1st Roll}) = \frac{8}{32} + \frac{24}{32}(E+1) = 4$$

On subsequent rolls, we have

$$P(X) = \begin{pmatrix} 4 & 5 & 6 & 8 & 9 & 10 \\ \frac{3}{36} & \frac{4}{36} & \frac{5}{36} & \frac{5}{36} & \frac{4}{36} & \frac{3}{36} \end{pmatrix}$$

giving 3 unique scenarios. Similarly to roll one we have

$$E(\text{Rolls}|\text{Win 2nd Roll}) = \frac{3}{30} + \frac{27}{30}(E+1) = 10$$
$$E(\text{Rolls}|\text{Win 2nd Roll}) = \frac{4}{30} + \frac{26}{30}(E+1) = 7.5$$
$$E(\text{Rolls}|\text{Win 2nd Roll}) = \frac{5}{30} + \frac{25}{30}(E+1) = 6$$

And

$$E(\text{Rolls}|\text{You Win}) = \frac{4(12) + 10(6) + 7.5(8) + 6(10)}{36} = 6.\overline{3}$$

For the house we have

$$E(\text{Rolls}|\text{Win 1st Roll}) = \frac{4}{28} + \frac{24}{28}(E+1) = 7$$
$$E(\text{Rolls}|\text{Win 2nd Roll}) = \frac{6}{33} + \frac{27}{33}(E+1) = 5.5$$
$$E(\text{Rolls}|\text{Win 2nd Roll}) = \frac{6}{32} + \frac{26}{32}(E+1) = 5.(\bar{3})$$
$$E(\text{Rolls}|\text{Win 2nd Roll}) = \frac{6}{31} + \frac{25}{31}(E+1) = 5.1(\bar{6})$$

And

$$E(\text{Rolls}|\text{House Win}) = \frac{7(12) + 5.5(6) + 5.\bar{(3)}(8) + 5.1\bar{(6)}(10)}{36} = 5.87$$

Thus the expected rolls given you win is greater than the expected rolls given the house wins.

2. **Suppose it takes \$20 to play, and the payoff is \$100, \$80, \$60, \$40, \$0 if you win in the 1st, 2nd, 3rd, 4th, 5th round, respectively. That is, if you win in the 1st round, you are paid \$100 (so your net profit is \$80), if you win in the 2nd round, you are paid \$80, etc. Derive the expected payoff function of a game.**

Part III: Statistical Estimation and Statistical Inference

In classical statistics, parameters are unknown constants whereas estimators are functions of samples and are random variables. The questions in this section are designed to clarify the relationship between parameters and estimators, and explore the properties that different estimators may have.

Question 8

Let $Y1, \ldots, Yn$ be n random variables, such that any two of them are uncorrelated, and all share the same mean μ and variance σ^2. Let Y be the average Y_i, which is also a random variable.

Define the class of linear estimators of μ by

$$W = \sum_{i=1}^{n} a_i Y_i$$

where the a_i are constants.

1. **What restriction on the a_i is needed for W to be an unbiased estimator of μ?**

$$W = a_1 E(Y_1) + a_2 E(Y_2) + ... + a_n E(Y_n)$$

And $Y_i = \mu$ ∀ i ∈ n.

Therefore

$$W = (a_1 + a_2 + ... + a_n)\mu$$

For W to be an unbiased estimator of μ, $a_1 + a_2 + ...a_3 = 1$

2. **Find $Var(W)$.**

$$Var(W) = a_1^2 Var(Y_1) + a_2^2 Var(Y_2) + ... + a_n^2 Var(Y_n)$$

and Var(Y_i) = $\sigma^2 \forall \in$ n. Therefore

$$Var(W) = (a_1^2 + a_2^2 + ... + a_n^2)\sigma^2$$

3. **Given a set of numbers $a_1, \dots, a_n$, the following inequality holds:**

$$\frac{1}{n}\left(\sum_{i=1}^{n} a_i\right)^2 \leq \sum_{i=1}^{n} a_i^2$$

Use this inequality, along with the previous parts of this question, tho show that $Var(W) \geq Var(\bar{Y})$ whenever W is unbiased. We say that $\bar{Y}$ is the best linear unbiased estimator (BLUE).

When W is unbiased (ie $(a_1 + a_2 + ... + a_n) = 1$), the expression becomes

$$\frac{1}{n} \leq a_1^2 + a_2^2 + ... + a_n^2$$

Multiplying by σ we have

$$Var(\bar{Y}) \leq \frac{\sigma}{n} = \sigma^2(a_1^2 + a_2^2 + ... + a_n^2) = Var(W)$$

or

$$Var(W) \geq Var(\bar{(Y)})$$

Question 9

Let $\bar{Y}$ denote the average of n independent draws from a population distribution with mean μ and variance σ^2. Consider two alternative estimators of μ: $W_1 = \frac{n-1}{n}\bar{Y}$ and $W_2 = k\bar{Y}$, where $0 < k < 1$.

1. **Compute the biases of both W_1 and W_2. Which estimator is consistent?**

$$W_1 = (\frac{n-1}{n})\bar{Y}$$

$$E(W_1) = \frac{n-1}{n}E(\bar{Y})$$

$$= \frac{n-1}{n}\mu$$

Therefore

$$Bias(W_1) = \frac{n-1}{n}\mu - \mu$$

$$nBias(W_1) = -\mu$$

$$Bias(W_1) = \frac{-\mu}{n}$$

Note that as $n \to \infty$, Bias$(W_1) \to 0$.

$$W_2 = k\bar{Y}$$
$$E(W_2) = kE(\bar{Y})$$
$$E(W_2) = k\mu$$

And

$$Bias(W_2) = k\mu - \mu$$
$$Bias(W_2) = \mu(k-1)$$

Given that 0 < k < 1, -μ < Bias(W_2) < 0.
Note that W_1 is consistent, as E(W_1) converges to μ as $n \to \infty$.

2. **Compute $Var(W_1)$ and $Var(W_2)$. Which estimator has lower variance?**

$$Var(W_1) = (\frac{n-1}{2})^2 Var(\bar(Y))$$
$$Var(W_1) = \frac{n-1}{n}\frac{n-1}{n}\frac{\sigma^2}{n}$$
$$Var(W_1) = \Big(\frac{(n-1)^2}{n^3}\Big)\sigma^2$$

$$Var(W_2) = k^2 Var(\bar(Y))$$
$$Var(W_2) = \frac{k^2}{n}\sigma^2$$

Note that by choosing k sufficiently close to 0, the variance of W_2 can be made arbitrarily small. Thus W_2 has the lower variance.

Question 10

Given a random sample $Y_1, Y_2, \ldots, Y_n$ from some distribution $F(\cdot)$ with mean μ and variance σ^2, where both μ and σ^2 are unknown parameters.

Let $\bar{Y}$ be the average of the sample. Consider the following estimator for σ^2:

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2$$

1. **Show that $E(\bar{Y}) = E(Y_i)\ \forall i \in 1, 2, \ldots, n$**

$$\begin{aligned} E(\hat{Y}) &= \frac{1}{n}\sum_{i=1}^{n} E(Y_i) \\ &= \frac{1}{n}[E(Y_1) + E(Y_2) + ... + E(Y_n)] \end{aligned}$$

Because $Y_1 \ldots\ Y_n$ are from the same distribution with mean μ,

$$E(\hat{Y}) = \frac{1}{n}[\mu + \mu + ... + \mu]$$

Thus

$$E(Y_1) = E(Y_2) = \cdots = E(Y_n) = E(\hat{Y})$$

2. **Show that** $Var(\bar{Y}) = \frac{1}{n}Var(Y_i)\ \forall i \in 1, 2, \ldots, n$

$$\begin{aligned} Var(\hat{Y}) &= \sum_{i=1}^{n} Var(\frac{Y_i}{n}) \\ &= Var\left(\frac{1}{n}Y_1 + \frac{1}{n}Y_2 + \cdots + \frac{1}{n}Y_n\right) \\ &= \frac{1}{n^2}Var(Y_1) + \frac{1}{n^2}Var(Y_2) + \cdots + \frac{1}{n^2}Var(Y_n) \end{aligned}$$

Since $Y_1 \ldots Y_n$ are from the same distribution with variance σ^2

$$= \frac{1}{n^2}[\sigma^2 + \sigma^2 + ... + \sigma^2]$$

Note that σ^2 occurs n times

$$= \frac{1}{n^2}[n\sigma^2]$$

Thus

$$Var(\hat{Y}) = \frac{1}{n}Var(Y_i)\forall i \in 1, 2, \ldots, n$$

3. **Compute the expectation of $\widehat{\sigma^2}$ in terms of n and σ^2. In your derivation, make sure make use of the *i.i.d.* property and identify where you use it.**

$$\begin{aligned} E(\widehat{\sigma^2}) &= E\left(\frac{(Y_i - \bar{Y})^2}{n}\right) \\ nE(\widehat{\sigma^2}) &= E\big(\sum_{i=1}^{n} Y_i^2 - 2Y_i\bar{Y} + \bar{Y}^2\big) \\ &= E\left[\sum_{i=1}^{n} Y_i^2\right] - E\left[2\bar{Y}\sum_{i=1}^{n} Y_i\right] + E\left[\bar{Y}^2\sum_{i=1}^{n} 1\right] \end{aligned}$$

Note that $\sum_{i=1}^{n} Y_i$ = n$\bar{Y}$

$$\begin{aligned} nE(\widehat{\sigma^2}) &= E\left[\sum_{i=1}^{n} Y_i^2\right] - nE\left[\bar{Y}^2\right] \\ &= nE[Y_i^2] - nE[\bar{Y}^2] \\ E(\widehat{\sigma^2}) &= E[Y_i^2] - E[\bar{Y}^2] \end{aligned}$$

Note that for independent and identically distributed random variables $Y_i \dots Y_n$ with sample mean $\bar{Y}$

$$E[Y^2] = E[\bar{Y^2}] = \frac{1}{n}\sigma^2 - \mu^2$$

Substituting into our equation gives

$$\begin{aligned} E(\widehat{\sigma^2}) &= \left[\sigma^2 + \mu^2\right] - \left[\frac{1}{n}\sigma^2 + \mu^2\right] \\ &= \frac{(n-1)}{n}\sigma^2 \end{aligned}$$

4. **Is this an unbiased estimator for σ^2?**

Since $E(\widehat{\sigma^2}) \neq \sigma^2$, this is a biased estimator.

5. **If not, what function of $\widehat{\sigma^2}$ produce an unbiased estimator?**

$$E(\widehat{\sigma^2}) = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2$$

Question 11

Wooldridge's textbook: Appendix C, Question 4*i*, *ii*, *iii*.

4i. We have random variable Z = $\frac{Y}{X}$

$$E(Z|X) = E(\frac{Y}{X}|X)$$

Using property CE.2

$$\begin{aligned} E(Z|X) &= \frac{1}{X}E(Y|X) \\ &= \theta X/X \\ &= \theta \end{aligned}$$

Question 12

Wooldridge's textbook: Appendix C, Question 6.

6. **You are hired by the governor to study whether a tax on liquor has decreased average liquor consumption in your state. You are able to obtain, for a sample of individuals selected at random, the difference in liquor consumption (in ounces) for the years before and after the tax. For person i who is sampled randomly from the population. Y_i denotes the change in liquor consumption. Treat these as a random sample from a Normal(μ, $|sigma^2$) distribution.**

i. H_0: $\mu = 0$.
ii. H_a: $\mu < 0$.

iii.

$$t = \frac{\bar(y) - \mu}{s/\sqrt{n}} = \frac{-32.8}{466.4/\sqrt{900}} \approx -2.11$$

p = Pr(z ≤ -2.11) = .017

Because 0.01 < p < 0.05, we can reject H_0 at the 5% level but not at the 1% level.

iv. Given our estimate of μ, we are estimating the tax was associated with a 32.8 oz decrease in alcohol consumption. From a health perspective, this is less than a 1 oz decrease in consumption of liquor per week, a fairly insignificant difference.

v. This analysis assumes that other factors that influence alcohol consumption over the period of interest have remained constant.

Question 13

Wooldridge's textbook: Appendix C, Question 8. In addition, answer the following questions:

8. **For a given player, the outcome of a particular shot can be modeled as a Bernoulli variable: if Y_i is the outcome of shot i, then $Y_i = 1$ if the shot is made, and $Y_i = 0$ if the shot is missed. Let θ denote the probability of making any particular three-point shot attempt. The natural estimator of u is $\bar{Y}$ = FGM/FGA.**

i.

$$\theta = 188/429 \approx 0.438$$

ii. For each trial

$$Var(Y_i) = \theta(1-\theta)$$

Thus for the sample mean we have

$$Var(\bar{Y}) = \frac{\theta(1-\theta)}{n}$$

$$sd(\bar{Y}) = \sqrt{\frac{\theta(1-\theta)}{n}}$$

iii. For $\theta = 0.5$, we have $t = (\bar{Y} - 0.5)/se(\bar{Y})$

$$se = \sqrt{0.438(1-.438)/429} \approx 0.024$$

$$t = (.438 - 0.5)/0.024 \approx -2.58$$

$$p \approx Pr(Z \leq -2.58) \approx 0.004$$

At the 1% significance level, we reject the null hypotheses that $\theta = 0.5$

1. **Define Type I error.**

Rejecting the null hypothesis when it is true.

2. **What is the probability of Type I error of this test?**

0.01

3. **Define Type II error.**

Accepting the null hypothesis when it is false.

4. **What is the probability of Type II error when using this decision rule, assuming the "true" population proportion is $\theta^* = 0.45$.**

At the 1% significance level, Z_{crit} = -2.33. And

$$-2.33 = (\theta^*_{crit} - 0.5)/0.24$$
$$\theta^*_{crit} = .444.$$
$$P(\theta^*_{crit} \geq .444|\mu = 0.45 \text{ and } \sigma = 0.024) = Pr(Z \geq -.25)) \approx 0.60.$$

The type II error rate in this scenario is 0.60.

5. **Define the power of the test (in general terms).**

The power of a test is the probability that we correctly reject the null hypothesis, ie. Pr(Reject | Null is False).

6. **Calculate the power of this test, again assuming the "true" population proportion is $\theta^* = 0.45$.**

The power is 1 - 0.65 = 0.40

```
plot(cars)
```

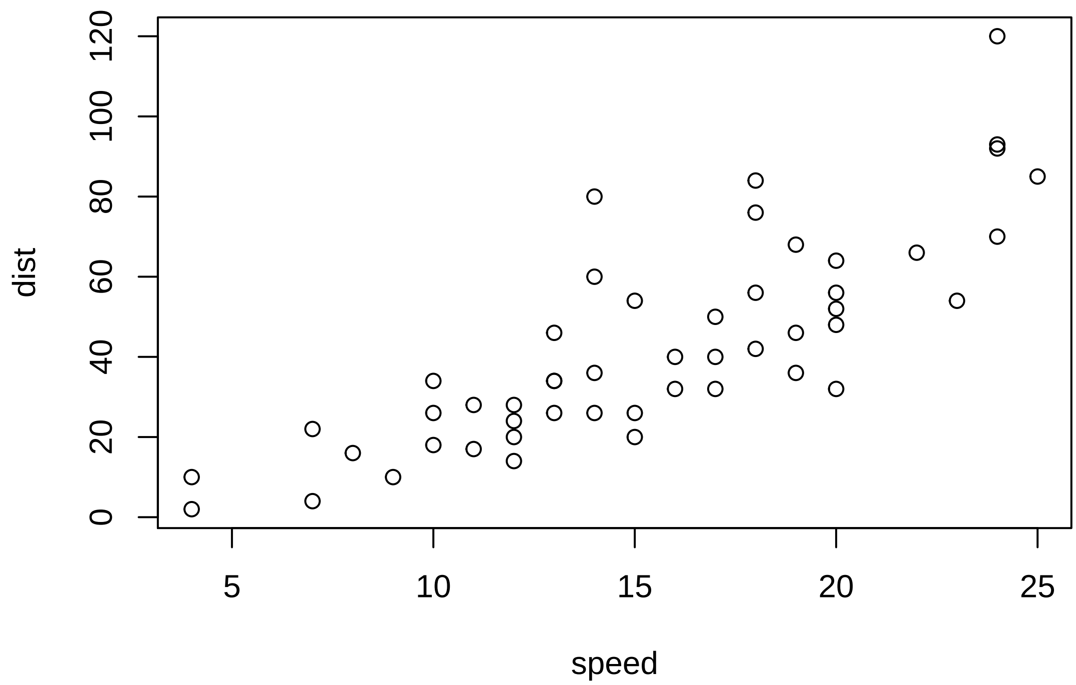

Figure 1: Some plot