W271-Lab1 Spring 2016

Megan Jasek, Charles Kekeh and Rohan Thakur

Thursday, January 14, 2016

Question 1

ML = 36
Stat = 28
Awesome = 18
ML ∩ Stat = 22
ML ∩ Awesome = 12
Stat ∩ Awesome = 9
ML ∪ Stat ∪ Awesome = 48

1)

$$ML \cup Stat = ML + Stat - ML \cap Stat = 42$$
$$Awesome.Other = (ML \cup Stat \cup Awesome) - (ML \cup Stat) = 6$$
$$Awesome = Awesome.Other \cup (Stat \cap Awesome) \cup (ML \cap Awesome) - (ML \cap Stat \cap Awesome)$$
$$18 = 6 + 9 + 12 - (ML \cap Stat \cap Awesome)$$
$$ML \cap Stat \cap Awesome = 9$$

$$\mathbf{Pr(ML \cap Stat \cap Awesome) = \frac{ML \cap Stat \cap Awesome}{ML \cup Stat \cup Awesome} = \frac{9}{48}}$$

2)

$$Pr(Awesome|ML) = \frac{Pr(Awesome \cap ML)}{Pr(ML)}$$
$$Pr(Awesome|ML) = \frac{12}{36} = \frac{1}{3}$$
$$\mathbf{Pr(!Awesome|ML) = 1 - Pr(Awesome|ML) = 1 - \frac{12}{36} = 1 - \frac{1}{3} = \frac{2}{3}}$$

3)
Please note that this answer assumes that the question requests an answer for the data scientist being an expert in either machine learning OR statistics (Exclusive OR) and excludes experts in both fields.

$$= Pr((ML \cap Stat') \cup (ML' \cap Stat)|Awesome)$$
$$= \frac{Pr((ML \cap Awesome - ML \cap Awesome \cap Stat) \cup (Stat \cap Awesome - ML \cap Awesome \cap Stat)}{Pr(Awesome)}$$
$$= \frac{(12 - 9) + (9 - 9)}{18}$$
$$= \frac{3}{18}$$
$$= \frac{1}{6}$$

Final answer: $\mathbf{\frac{1}{6}}$

However, if a non-exclusive OR is desired, then we need to add the number of data scientists at the intersection of Awesome, Machine Learning and Statistics to our numerator, giving a result of $\frac{12}{18}$, or $\mathbf{\frac{2}{3}}$

Question 2

Pr(A)=p $\leq \frac{1}{2}$, Pr(B)=q where $\frac{1}{4} < q < \frac{1}{2}$

1)

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

Pr(A $\cup$ B) is maximized when Pr(A $\cap$ B) is minimized and P(A) and P(B) are maximized. In this case that would be:

$$min(Pr(A \cap B)) = 0, (A \ and \ B \ are \ independent)$$
$$max(Pr(A)) = 1/2$$
$$max(Pr(B)) = 1/2 - \epsilon, where \ \epsilon \ approaches \ 0$$

$$\mathbf{max(Pr(A \cup B)) = 1 - \epsilon, where\ \epsilon\ approaches\ 0}$$

Alternately, Pr(A $\cup$ B) is minimized when A and B are completely overlapping, A = B. In this case

$$min(Pr(A \cup B) = max(min(Pr(A), Pr(B))) = 1/4 + \epsilon, where \ \epsilon \ approaches \ 0$$

$$\mathbf{min(Pr(A \cup B)) = 1/4 + \epsilon, where\ \epsilon\ approaches\ 0}$$

2)

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

Pr(A|B) is maximized when A is completely contained in B. In this case, Pr(A|B) = 1.

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

Pr(A|B) is minimized when Pr(A $\cap$ B) = 0 (A and B are independent). When Pr(A $\cap$ B) = 0, then Pr(A $\cup$ B) = 0.

$$\mathbf{min(Pr(A|B)) = 0}$$

Question 3

1)
Given that the server's lifespan is a randon uniform distribution over the range [0,k], the probability of every additional year of operation is independent of the time elapsed and is equal to

$$\mathbf{Pr(1\ year\ of\ operation) = \frac{1}{k}}$$

2)
We know that:

$$E(g(x)) = \int_{x=0}^{\infty} g(x) f_x(x) dx$$

Considering g to be our refund function over time t:

$$E(g(t)) = \int_{t=0}^{1} \frac{\theta}{k} dt + \int_{t=1}^{k/2} \frac{2(k-t)^{\frac{1}{2}}}{k} dt + \int_{t=k/2}^{\frac{3k}{4}} \frac{\theta}{10k} dt + \int_{t=\frac{3k}{4}}^{\infty} 0 dt$$

$$E(g(t)) = \frac{\theta}{k}[t]_0^1 + \frac{-4}{3k}(k-t)^{\frac{3}{2}}[t]_1^{k/2} + \frac{\theta}{10k}[t]_{\frac{k}{2}}^{\frac{3k}{4}} + 0$$

$$\mathbf{E(g(t)) = \frac{\theta}{k} + \frac{4}{3k}(k-1)^{\frac{3}{2}} - \frac{4}{3k}(\frac{k}{2})^{\frac{3}{2}} + \frac{\theta}{40}}$$

3)
We know that Var(X) = E[X^2] -$E[X]^2$
Thus Var(g(x)) = E[(g(x)^2] - [$E[g(X)]]^2$
We previously computed E(g(x)). We now compute E[g(x)]^2

$$E[(g(x)^2] = \int_{0}^{1} \frac{\theta^2}{k} dt + \int_{1}^{k/2} \frac{A^2(k-t)}{k} dt + \int_{\frac{k}{2}}^{\frac{3k}{4}} \frac{\theta^2}{100k} dt$$

$$\mathbf{E[(g(x)^2] = \frac{\theta^2}{k} + \frac{\theta^2}{400} - \frac{3k^2 - 8k + 4}{2k}}$$

We substract $E(g(x))^2$ as previously computed to obtain the variance, and

$$\mathbf{Var(g(x)) = \frac{\theta^2}{k} + \frac{\theta^2}{400} - \frac{3k^2 - 8k + 4}{2k} - [\frac{\theta}{k} + \frac{4}{3k}(k-1)^{\frac{3}{2}} - \frac{4}{3k}(\frac{k}{2})^{\frac{3}{2}} + \frac{\theta}{40}]^2}$$

Question 4

f(x,y) = 2e^{-x}e^{-2y} for 0 < x < ∞, 0 < y < ∞, 0 otherwise

1)

$$Pr(x > a, y < b) = \int_{y=0}^{b} \int_{x=a}^{\infty} 2e^{-x} e^{-2y} dx dy$$

$$Pr(x > a, y < b) = 2 \int_{y=0}^{b} e^{-2y} dy \int_{x=a}^{\infty} e^{-x} dx$$

$$Pr(x > a, y < b) = 2\int_{y=0}^{b} e^{-2y}(1 - [-e^{-x}]_0^a)dy$$

$$Pr(x > a, y < b) = 2e^{-a}\int_{y=0}^{b} e^{-2y}dy$$

$$Pr(x > a, y < b) = 2e^{-a}[-\frac{1}{2}e^{-2y}]_0^b$$

$$\mathbf{Pr(x > a, y < b) = e^{-a}(1 - e^{-2b})}$$

2)

$$Pr(x < y) = \int_{y=0}^{\infty}\int_{x=0}^{y} 2e^{-x}e^{-2y}dxdy$$

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

$$Pr(x < y) = 2\int_{y=0}^{\infty} e^{-2y}dy[-e^{-x}]_0^y$$

$$Pr(x < y) = 2\int_{y=0}^{\infty} e^{-2y}(1 - e^{-y})dy$$

$$Pr(x < y) = 2\int_{y=0}^{\infty} e^{-2y} - e^{-3y}dy$$

$$Pr(x < y) = 2[\frac{1}{6}e^{-3y}(2 - 3e^{y})]_0^{\infty}$$

$$\mathbf{Pr(x < y) = \frac{1}{3}}$$

3)

$$Pr(X < a) = \int_{x=0}^{a}\int_{y=0}^{\infty} 2e^{-x}e^{-2y}dxdy$$

$$Pr(X < a) = 2\int_{x=0}^{a} e^{-x}dx\int_{y=0}^{\infty} e^{-2y}dy$$

$$Pr(X < a) = 2\int_{x=0}^{a} e^{-x}dx[-\frac{1}{2}e^{-2y}]_0^{\infty}$$

$$Pr(X < a) = \int_{x=0}^{a} e^{-x}dx$$

$$\mathbf{Pr(X < a) = 1 - e^{-a}}$$

Question 5

X random variable, x a real number.
Y = a + b (X - x^2)

1)

$$E(Y) = a + bE[(x - x^2)]$$

$$E(Y) = a + bE[X^2 - 2Xx - x^2]$$

$$E(Y) = a + b[E[X^2 - 2xE[X] + x^2]]$$

E(Y) is minimized when $\frac{d}{dx}$E(Y) = 0

$$\frac{d}{dx}E(Y) = -2bE(X) + 2bx$$

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

2)

$$Whenx = E(X) : E(Y) = a + b[E[X^2] - 2(E[X])^2 + (E[X])^2]$$
$$E(Y) = a + b[E[X^2] - (E[X])^2]$$
$$\mathbf{E}(\mathbf{Y}) = \mathbf{a} + \mathbf{bVar}[\mathbf{X}]$$

3)
Y = ax + b(X - x^2)

$$E(Y) = ax + bE[(x - x^2)]$$
$$E(Y) = ax + bE[X^2 - 2Xx - x^2]$$
$$E(Y) = ax + b[E[X^2 - 2xE[X] + x^2]]$$

E(Y) is minimized when $\frac{d}{dx}$E(Y) = 0

$$\frac{d}{dx}E(Y) = a - 2bE(X) + 2bx$$

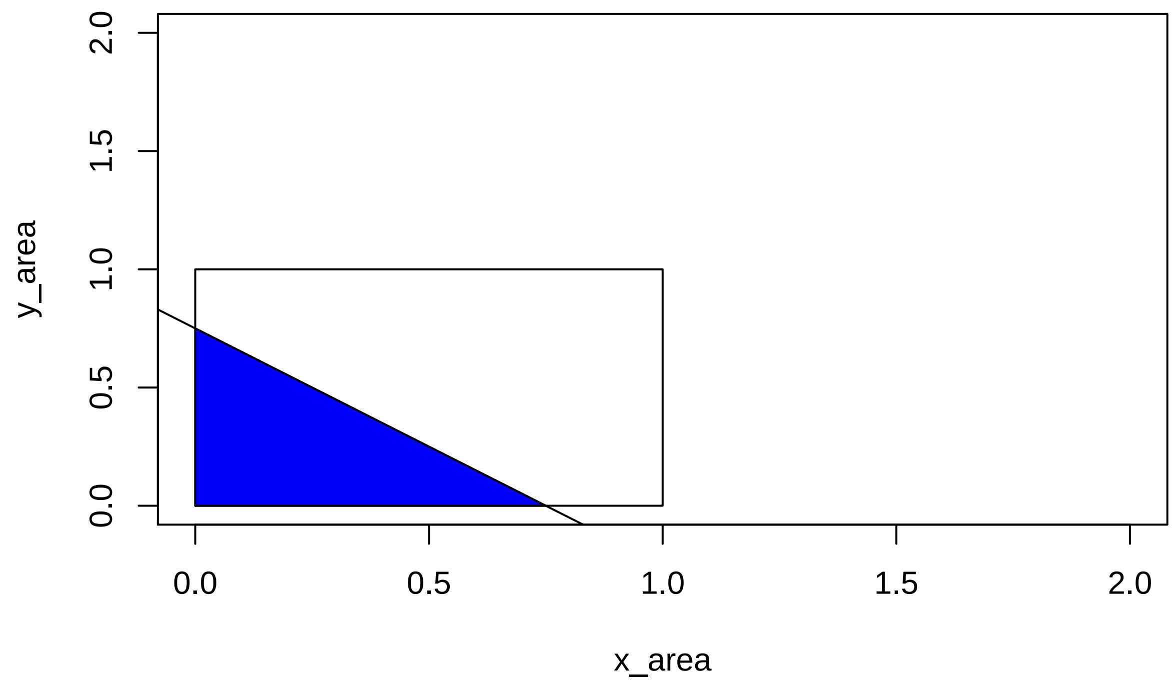
$$\frac{d}{dx}E(Y) = 0 \Rightarrow \mathbf{x} = \mathbf{E}(\mathbf{X}) - \frac{\mathbf{a}}{\mathbf{2b}}$$

Question 6

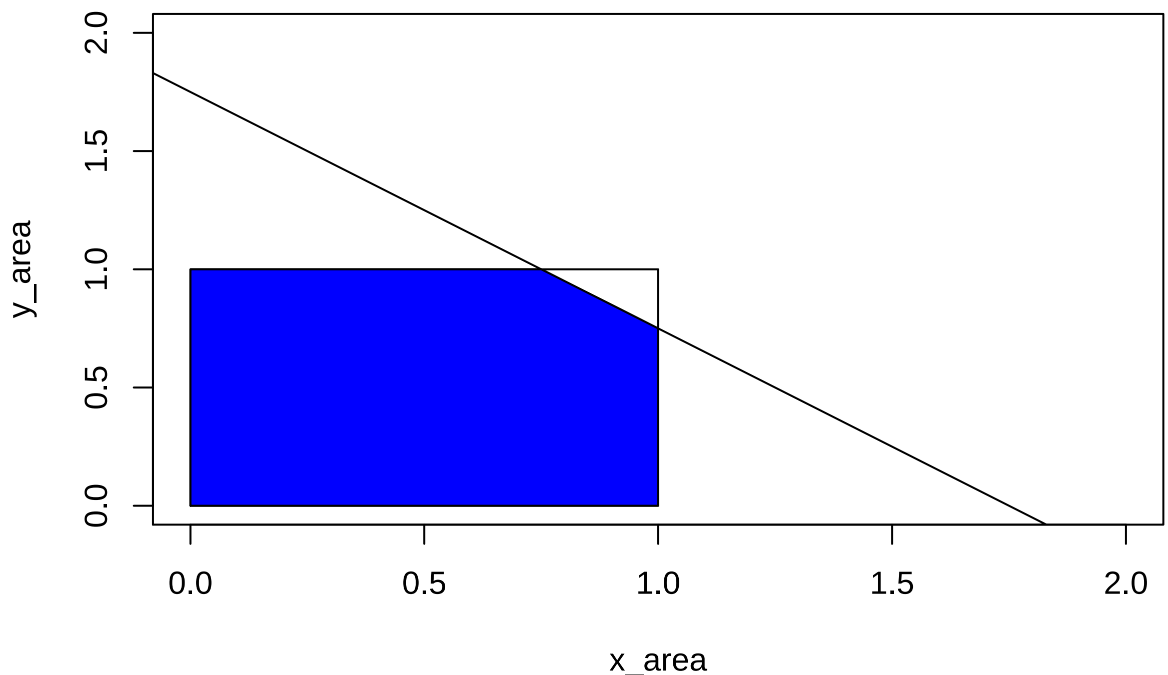
X, Y independent continuous variables, uniform over [0..1]
Z = X + Y

1)

```r
x_area = c(0:2)
y_area = c(0:2)

plot(x_area, y_area, type = "n")

xx = c(0, 1, 1, 0)
yy = c(0, 0, 1, 1)
polygon(xx, yy, density = 0, border = "black")

abline(.75, -1)
xz = c (0, .75, 0)
yz = c(0, 0, .75)
polygon(xz, yz, col = "blue", border = "black")
```




```r
plot(x_area, y_area, type = "n")

xx = c(0, 1, 1, 0)
yy = c(0, 0, 1, 1)
polygon(xx, yy, density = 0, border = "black")

abline(1.75, -1)
xz = c (0, 1, 1, .75, 0, 0)
yz = c(0, 0, .75, 1, 1, 0)
polygon(xz, yz, col = "blue", border = "black")
```



2)
From the areas above, we derive that:
For 0 <= z <= 1:

$$\mathbf{Pr(Z < z) = \frac{z^2}{2}}$$

For 1 < z <= 2:

$$\mathbf{Pr(Z < z) = 1 - \frac{(2-z)^2}{2}}$$

Hence:
For 0 <= z <= 1:

$$\mathbf{f(z) = \frac{d}{dz}\frac{z^2}{2} = z}$$

For 1 < z <= 2:

$$\mathbf{f(z) = \frac{d}{dz}1 - \frac{(2-z)^2}{2} = 2 - z}$$

Question 7

1)
In order to compute the expected number of dice rolls for both scenarios (player wins or house wins), we must first compute the probability of the player rolling an x in subsequent rounds if the game does not end with the first roll of the dice.

In the scenario, that the game goes beyond the first round, let us assume that the player needs p successes out of 36 in order to successfully roll x.

Therefore, if x has value 4 or 10, p = 3 successes. If x has value 5 or 9, p = 4 successes. if x has value 6 or 8, p = 5 successes.

Calculating the expected value of p:

$$E(p) = 3 * Pr(Player\ rolls\ 4\ or\ 10) + 4 * Pr(Player\ rolls\ 5\ or\ 9) + 5 * Pr(Player\ rolls\ 6\ or\ 8)$$
$$E(p) = 3 * \frac{6}{36} + 4 * \frac{8}{36} + 5 * \frac{10}{36}$$

```
library(MASS)
value = 3*6/36 + 4*8/36 + 5*10/36
print(fractions(value))
```

```
## [1] 25/9
```

Therefore, we expect to get 25/9 (2.78) successes out of 36 for the player after the first roll. Therefore the expected probability of success is:

$$Pr(p\ successes) = \frac{25}{9} * \frac{1}{36}$$

```
library(MASS)
value = 25/(9*36)
print(fractions(value))
```

```
## [1] 25/324
```

Now, in order to compute the expected number of dice rolls given the player wins, we need the following values:

$$Pr(House\ wins\ in\ 1\ roll) = Pr(2,\ 3\ or\ 12) = \frac{4}{36}$$
$$Pr(Player\ wins\ in\ 1\ roll) = Pr(7\ or\ 11) = \frac{8}{36}$$

Therefore,

$$Pr(Game\ proceeds\ to\ second\ roll) = 1 - (\frac{4}{36} + \frac{8}{36}) = \frac{24}{36}$$
$$Pr(House\ wins\ on\ second\ roll) = Pr(7) = \frac{6}{36}$$
$$Pr(Player\ wins\ on\ second\ roll) = Pr(x) = \frac{25}{324}$$
$$Pr(Game\ proceeds\ to\ third\ roll) = 1 - (\frac{6}{36} + \frac{25}{324}) = \frac{245}{324}$$

In every subsequent round, the house wins with probability 6/36, the player wins with probability 25/324 and the game continues with probability 245/324.

Now, computing expected number of dice rolls given the player wins:

$$\begin{aligned} E(Y|Player\ Wins) = 1 * Pr(Player\ wins\ in\ 1\ roll) \\ +2 * Pr(Game\ proceeds\ to\ second\ roll) * Pr(Player\ wins\ on\ second\ roll) + \\ 3 * Pr(Game\ proceeds\ to\ third\ roll) * Pr(Player\ wins\ on\ third\ roll) + \cdots \end{aligned}$$

Computing for first 10 rolls

```r
value = 1 * 8/36 + 2 * 24/36 * 25/324 + 3 * 24/36 * 245/324 * 25/324 + 4 * 24/36 *
    ((245/325)^2) * 25/324 + 5 * 24/36 * ((245/325)^3) * 25/324 + 6 * 24/36 *
    ((245/325)^4) * 25/324 + 7 * 24/36 * ((245/325)^5) * 25/324 + 8 * 24/36 *
    ((245/325)^6) * 25/324 + 9 * 24/36 * ((245/325)^7) * 25/324 + 10 * 24/36 *
    ((245/325)^8) * 25/324
print(round(value, 4))
```

```
## [1] 1.0495
```

Now, computing expected number of dice rolls given the house wins:

$$\begin{aligned}E(Y|House\ Wins) = 1 * Pr(House\ wins\ in\ 1\ roll) +\\ 2 * Pr(Game\ proceeds\ to\ second\ roll) * Pr(House\ wins\ on\ second\ roll) +\\ 3 * Pr(Game\ proceeds\ to\ third\ roll) * Pr(House\ wins\ on\ third\ roll) + \cdots\end{aligned}$$

Computing for first 10 rolls

```r
value = 1 * 4/36 + 2 * 24/36 * 6/36 + 3 * 24/36 * 245/324 * 6/36 + 4 * 24/36 *
    ((245/325)^2) * 6/36 + 5 * 24/36 * ((245/325)^3) * 6/36 + 6 * 24/36 * ((245/325)^4) *
    6/36 + 7 * 24/36 * ((245/325)^5) * 6/36 + 8 * 24/36 * ((245/325)^6) * 6/36 +
    9 * 24/36 * ((245/325)^7) * 6/36 + 10 * 24/36 * ((245/325)^8) * 6/36
print(round(value, 4))
```

```
## [1] 1.898
```

Final answer: Since the probability of winning for the house is greater than that of the player for every round after round 10, and we know that the expected number of rolls given the player wins is less than the expected number of rolls given the house wins for 10 dice rolls, we know that expected number of rolls given the player wins is less than the expected number of rolls given the house wins for infinite rolls

2)

$$\begin{aligned}E(payoff) = 100 * Pr(Player\ wins\ in\ 1\ roll) +\\ 80 * Pr(Game\ proceeds\ to\ second\ roll) * Pr(Player\ wins\ on\ second\ roll) +\\ 60 * Pr(Game\ proceeds\ to\ third\ roll) * Pr(Player\ wins\ on\ third\ roll) +\\ 40 * Pr(Game\ proceeds\ to\ fourth\ roll) * Pr(Player\ wins\ on\ fourth\ roll) +\\ 0 * Pr(Game\ proceeds\ to\ fifth\ roll) * Pr(Player\ wins\ on\ fifth\ roll)\end{aligned}$$

Computing result:

```r
value = 100 * 8/36 + 80 * 24/36 * 25/324 + 60 * 24/36 * 245/324 * 25/324 + 40 *
    24/36 * ((245/325)^2) * 25/324 + 0 * 24/36 * ((245/325)^3) * 25/324
print(round(value, 2))
```

```
## [1] 29.84
```

Now, computing expected net profit:

$$E(Net\ profit) = E(Payoff\ -\ Cost) = E(payoff) - E(Cost) = 29.84 - 20 = 9.84$$

Final answer: The expected net profit of the game is $9.84

Question 8

E(Y_1) = E(Y_2) = ... = E(Y_n) = μ
Var(Y_1) = Var(Y_2) = ... = Var(Y_n) = σ^2

1)

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

For W to be un unbiased estimator of μ:

$$E(W) = \mu$$

$$\Rightarrow E(\sum_{i=1}^{n} a_i Y_i) = \mu$$

$$\Rightarrow \sum_{i=1}^{n} a_i E(Y_i) = \mu$$

$$\Rightarrow \sum_{i=1}^{n} a_i \mu = \mu$$

$$\Rightarrow \mu \sum_{i=1}^{n} a_i = \mu$$

$$\Rightarrow \mathbf{\sum_{i=1}^{n} a_i = 1}$$

2)

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

$$Var(W) = \sum_{i=1}^{n} a_i^2 Var(Y_i)$$

$$\mathbf{Var(W) = \sigma^2 \sum_{i=1}^{n} a_i^2}$$

3)
We know that:

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

Multiply each side of the equation by σ^2 and rewrite the inequality by swapping which terms are on each side. It does not change the value of the inequality by multiplying each side by a positive number and σ^2 is a positive number.

$$\sigma^2 \sum_{i=1}^{n} a_i^2 \geq \sigma^2 \frac{1}{n}(\sum_{i=1}^{n} a_i)^2$$

Substitute the variance found in part 2 on the left-hand side:

$$Var(W) \geq \sigma^2 \frac{1}{n}(\sum_{i=1}^{n} a_i)^2$$

When W is unbiased, we know that:

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

Substitute the above value in to the right-hand side:

$$Var(W) \geq \frac{\sigma^2}{n}$$

We know that:

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

Substitute this value in to the right-hand side of the equation and the proof is complete.

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

Question 9

$\mathrm{W}_1 = (\frac{n-1}{n})\bar{Y}$

$\mathrm{W}_2 = \mathrm{k}\bar{Y}$

1)

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

Similarly:

$$bias(W_2) = E(k\bar{Y}) - \mu$$
$$bias(W_2) = kE(\bar{Y}) - \mu$$
$$bias(W_2) = k\mu - \mu$$
$$\mathbf{bias(W_2) = \mu(k-1)}$$

Which is a consistent estimator?
W_1 is a consistent estimator of μ because as n goes to ∞, the difference between W_1 and μ goes to 0.
W2 is not a consistent estimator of μ.

2)

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

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

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

Similarly:

$$Var(W_2) = Var(k\bar{Y})$$

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

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

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

Which estimator has lower variance?

The estimator that has lower variance depends on the values of n and k as follows:
For n = 1, k > 0, Var($\mathbf{W}_1$) is lower
For n > 1 and k = $\frac{n-1}{n}$, Var($\mathbf{W}_1$) = Var($\mathbf{W}_2$)
For n > 1 and k > $\frac{n-1}{n}$, Var($\mathbf{W}_1$) is lower
For n > 1 and k < $\frac{n-1}{n}$, Var($\mathbf{W}_2$) is lower

Question 10

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

1)

$$E(\bar{Y}) = E[\frac{\sum_{i=1}^{n} Y_i}{n}]$$

We know that E(X) is a linear function, thus:

$$E(\bar{Y}) = \frac{1}{n}\sum_{i=1}^{n} E[Y_i]$$

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

$$\Rightarrow \mathbf{E(\bar{Y}) = \mu = E[Y_i], y = 1, \cdots, n}$$

2)

$$Var(\bar{Y}) = Var[\frac{\sum_{i=1}^{n} Y_i}{n}]]$$

We know that Var[$\sum_{i=1}^{n}$a$_i X_i$] = $\sum_{i=1}^{n} a_i^2$Var(X_i), thus:

$$Var(\bar{Y}) = \frac{1}{n^2}\sum_{i=1}^{n}(Var[Y_i])$$

$$\Rightarrow \mathbf{Var(\bar{Y}) = \frac{1}{n^2}n\sigma^2 = \frac{1}{n}Var[Y_i], y = 1, \cdots, n}$$

3)

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

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

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

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

$$E[\widehat{\sigma^2}] = \frac{1}{n}\sum_{i=1}^{n}[E(Y_i^2) - \frac{2}{n}E(Y_i\sum_{j=1}^{n} Y_j) + E(\bar{Y}^2)]$$

$$E[\widehat{\sigma^2}] = \frac{1}{n}\sum_{i=1}^{n}[E(Y_i^2) - \frac{2}{n}E(Y_i\sum_{j=1,j\neq i}^{n} Y_j + Y_iY_i) + E(\bar{Y}^2)]$$

$$E[\widehat{\sigma^2}] = \frac{1}{n}\sum_{i=1}^{n}[E(Y_i^2) - \frac{2}{n}(E(Y_i\sum_{j=1,j\neq i}^{n} Y_j) + E(Y_i^2)) + E(\bar{Y}^2)]$$

We know that:

$$Cov(X,Y) = E(XY) - \mu_X\mu_Y$$

$$\Rightarrow E(XY) = Cov(X,Y) + \mu_X\mu_Y$$

We also know that:

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

$$\Rightarrow Cov(X,X) = E[(X-\mu_X)^2] = Var(X)$$

And

$$E(XX) = E(X^2) = Var(X) + \mu_X^2$$

Going back to $E[\widehat{\sigma^2}]$, we use the iid property to reduce $E(Y_i\sum_{j=1,j\neq i}^{n} Y_j) + E(Y_i^2)$ since Y_i and Y_j are identical and independent varaibles. Thus:

$$E(Y_i\sum_{j=1,j\neq i}^{n} Y_j) = \sum_{j=1,j\neq i}^{n} E(Y_iY_j) = \sum_{j=1,j\neq i}^{n} E(Y_i)E(Y_j) = (n-1)\mu^2$$

And, using our finding above about $E(X^2)$, we have:

$$E(Y_i^2) = \sigma^2 + \mu^2$$

And

$$E(\bar{Y}^2) = \frac{\sigma^2}{n} + \mu^2 = \frac{\sigma^2 + n\mu^2}{n}$$

Thus:

$$E[\widehat{\sigma^2}] = \frac{1}{n}\sum_{i=1}^{n}[\sigma^2 + \mu^2 - \frac{2}{n}[(n-1)\mu^2 + \sigma^2 + \mu^2] + \frac{\sigma^2 + n\mu^2}{n}]$$

$$E[\widehat{\sigma^2}] = \frac{1}{n}\sum_{i=1}^{n}[\sigma^2 + \mu^2 - \frac{2}{n}[n\mu^2 + \sigma^2] + \frac{\sigma^2 + n\mu^2}{n}]$$

$$E[\widehat{\sigma^2}] = \frac{1}{n}\sum_{i=1}^{n}[\frac{(n-1)\sigma^2}{n}]$$

$$\mathbf{E}[\widehat{\sigma^2}] = \frac{\mathbf{(n-1)\sigma^2}}{\mathbf{n}}$$

4)
We have shown that $E(\widehat{\sigma^2}) \neq \sigma^2$ and we conclude that $\widehat{\sigma^2}$ is a biased estimator for σ^2.

5)
We know that: $\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ is a biased estimator of σ^2. Thus:

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

is an unbiased estimator of σ^2

Question 11

X, Y positive random variables. E(Y|X) = θ X

i)
We know that Z = $\frac{Y}{X}$
We first compute E(Z|X)

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

Using E(a(X) Y + b(X))|X) = a(X)E(Y|X) + bX, we derive:

$$E(Z|X) = \frac{1}{X}E(Y|X) = \frac{\theta X}{X} = \theta$$

Then, we knwo that E[E(Z|X)] = E(Z). Thus:

$$E(\theta) = E(Z)$$

Θ being a constant:

$$E(\theta) = \theta$$

and

$$\mathbf{E(Z)} = \theta$$

ii)
$W_1 = n^{-1}\sum_{i=1}^{n}\frac{Y_i}{X_i}(X_i, Y_i) : i = 1, 2, \cdots, n$ is the estimator.
We compute $E(W_1)$:

$$E(W_1) = n^{-1}E[\sum_{i=1}^{n}(\frac{Y_i}{X_i})]$$

$$E(W_1) = n^{-1}\sum_{i=1}^{n}E(Y_i/X_i)$$

$$\mathbf{E(W_1) = n^{-1}[nE(Z)] = E(Z)} = \theta$$

We conclude that W_1 is unbiased for θ

iii)
$W_2 = \frac{\bar{Y}}{\bar{X}}$

$$W_2 = \frac{n^{-1}\sum_{i=1}^{n}Y_i}{n^{-1}\sum_{i=1}^{n}X_i}$$

$$\Rightarrow W_2 = \frac{Y_1 + Y_2 + Y_3 + \cdots + Y_n}{X_1 + X_2 + X_3 + \cdots + X_n}$$

whereas:

$$W_1 = \frac{Y_1}{X_1} + \frac{Y_2}{X_2} + \frac{Y_3}{X_3} + \cdots \frac{Y_n}{X_n}$$

$W_2 and W_1$ are thus different estimators

Now, showing that

$$E(\bar{Y}) = E(Y)$$

$$E(\bar{Y}) = E(n^{-1} * [Y_1 + Y_2 + \cdots + Y_n])$$
$$\Rightarrow E(\bar{Y}) = n^{-1} * E(Y_1 + Y_2 + \cdots + Y_n)$$
$$\Rightarrow E(\bar{Y}) = n^{-1} * [E(Y_1) + E(Y_2) + \cdots + E(Y_n)]$$
$$\Rightarrow E(\bar{Y}) = n^{-1} * n * \theta X$$
$$\Rightarrow E(\bar{Y}) = \theta E(X)$$

Similarly,

$$E(\bar{X}) = E(X)$$

Therefore,

$$E(W_2) = E[\frac{\bar{Y}}{\bar{X}}]$$
$$\Rightarrow E(W_2) = E[\frac{\theta E(X)}{E(X)}] = E(\theta)$$
$$\Rightarrow \mathbf{E(W_2)} = \theta$$

Question 12

i)
The null hypothesis is that $\mu = 0$
ii)
The alternative hypothesis hypothesis is that $\mu < 0$
iii)

$$\mu = 0, when\ the\ null\ hypothesis\ is\ true$$
$$n = 900$$
$$\bar{Y} = -32.8$$
$$s = 466.4$$
$$t = \frac{\bar{Y} - \mu}{\frac{s}{\sqrt{n}}} = -2.109777$$
$$p(z \leq t) = 0.0174$$

Thus, we reject the null hypothesis at the 5% significance level as $p(z \leq t) \leq 0.05$ We cannot reject the null hypothesis at the 1% significance level as $p(z \leq t) \geq 0.01$

iv)
The effect size is inferior to 10% of the variance of the State Liquor Consumption variable. That's an indication of a small practical effect. We also compute the correlation coefficient as

$$r = \sqrt{\frac{t^2}{t^2 + DF}}$$
$$r = \sqrt{\frac{-2.11^2}{-2.11^2 + 899}} = 0.07$$

The value of R also confirms the small practical effect despite the test being statistically significant because of the high sample size.
v)
What has been assumed is that the other determinates of liquor consumption have had no net effect over the two-year period that was analyzed.

Question 13

$Y_i = 1$ Shot made. $Y_i = 0$ Shot missed. $\theta = Pr(Makinga3ptshot)$ Bernouilli distribution. $\bar{Y} = \frac{FGM}{FGA}$ estimator of θ

i)

$$\theta = \frac{188}{429} = .4382284$$

ii)
Because Y has a Bernouilli distribution: $E(Y) = \theta$ Let $Y_i i \in 1, \cdots, n$ be an occurance of a free throw. We know that each Y_i is a Bernouilli variable. We can define:

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

Thus:

$$Var(\bar{Y}) = (\frac{1}{n})^2 Var(\sum_{i=1}^{n} Y_i)$$
$$Var(\bar{Y}) = (\frac{1}{n})^2 \sum_{i=1}^{n} Var(Y_i)$$

Thus:

$$Var(\bar{Y}) = (\frac{1}{n})^2 n\theta(1-\theta)$$
$$\Rightarrow Var(\bar{Y}) = \frac{\theta(1-\theta)}{n}$$

And:

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

iii)
We know $se(\bar{\gamma}) = \sqrt{\frac{\bar{\gamma}(1-\bar{\gamma})}{n}}$ And $\frac{\bar{\gamma}-\theta}{se(\bar{Y})} \equiv Normal(0,1)$ We compute:

$$z_Y = \frac{\frac{188}{429} - .5}{se(\bar{Y})} = -2.588303$$
$$\Rightarrow p(z) = 0.0048$$

The p-value is significant at the 1% significance level and we reject the null hypothesis.

Additional Questions

1)
Type I error is the probability of a false positive or the probability of rejecting the null hypothesis when it is true. It's the probability of saying there is a result when there is not one.

2)
The probablity of the type I error is the significance level and is 1%.

3)
Type II error is the probability of a false negative, or the probablity of failing to reject the null hypothesis when it is false.

4)
We know that this is a one-tailed test and that the null is rejected if z is above -2.33.

Finding the critical value of Y which makes the null get rejected:

$$-2.33 = Z_{critical} = \frac{\bar{Y}_{critical} - 0.5}{se(\bar{Y})}$$

```
value = qnorm(0.01) * sqrt(188/429 * (1 - 188/429)/429) + 0.5
print(round(value, 2))
```

```
## [1] 0.44
```

Now, probability of a type II error is the probability of getting a Y value greater than 0.44

$$Pr(Z > \frac{0.44 - 0.45}{se(\bar{Y})})$$

Computing this, we get:

```
value2 = 1 - pnorm((value - 0.45)/sqrt(188/429 * (1 - 188/429)/429))
print(round(value2, 2))
```

```
## [1] 0.59
```

Therefore, the probability of Type II error of this test is 0.59

5)
The power of the test is the probabilty of rejecting the null hypothesis when it is actually false.

6)

$$Power = 1 - Pr(Type\ II\ Error)$$
$$Power = 1 - 0.59$$
$$Power = 0.41$$