W271-Lab1 Spring 2016

Charles Kekeh

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)

$$Pr(ML \cup Stat|Awesome) = \frac{Pr((ML \cup Stat) \cap Awesome)}{Pr(Awesome)}$$

$$Pr(ML \cup Stat|Awesome) = \frac{Pr(ML \cup Awesome) + Pr(Stat \cup Awesome) - Pr(Stat \cap ML \cap Awesome)}{Pr(Awesome)}$$

$$\mathbf{Pr(ML \cup Stat|Awesome) = \frac{12 + 9 - 9}{48} = \frac{12}{48} = \frac{1}{4}}$$

Question 2

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

1)
We know that:

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

Pr(A ∪ B) is maximized when Pr(A ∩ B) is minimized (A and B are independent and Pr(A ∩ B) = 0). The maximum value is in the range:

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

From there, we also know that Pr(A ∪ B) is minimized when Pr(A ∩ B) is maximized and A and B are completely overlapping random events and Pr(A ∩ B) = Min (Pr(A), Pr(B)). In these conditions

$$\mathbf{\frac{1}{4} < Min(Pr(A \cup B) < \frac{1}{2}}$$

2)
We know that:

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

Thus Pr(A|B) is minimized when Pr(A ∩ B) = 0. And

$$\mathbf{Min(Pr(A \cup B)) = 0}$$

Similarly, Pr(A|B) is maximized when the random event B is completely overlapped by the random event A. In such cases

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

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

$$\mathbf{Pr(1yearofoperation) = \frac{1}{k}}$$

2)
We know that:

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

Considering g to be our refund function over time t:

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

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

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

3)
We know that Var(X) = E[X^2] -μ^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]

$$E[(g(x)^2] = \int_0^1 \frac{\theta^2}{k} dt + \int_1^{k/2} \frac{A^2(k-t)}{k} dt + \int_{\frac{k}{2}}^{\frac{3k}{2}} \frac{Q^2}{100k} dt$$

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

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

Question 4

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

1)

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

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

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

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

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

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

2)

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

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

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

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

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

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

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

3)

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

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)

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

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

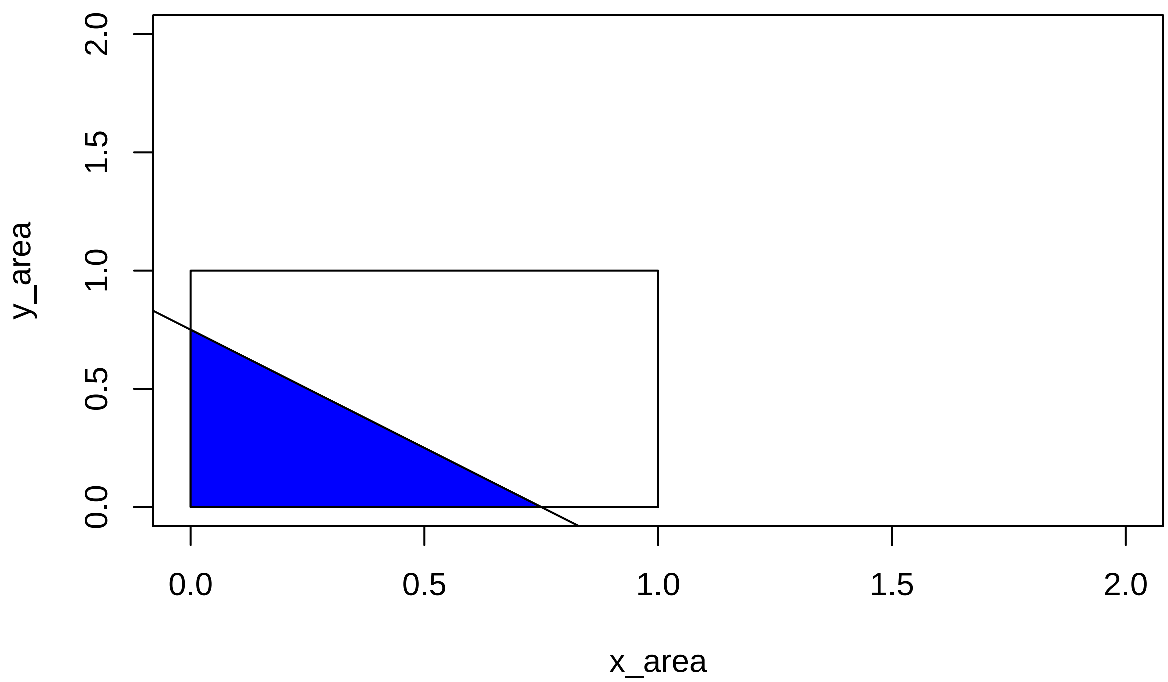
1)

```
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")
```

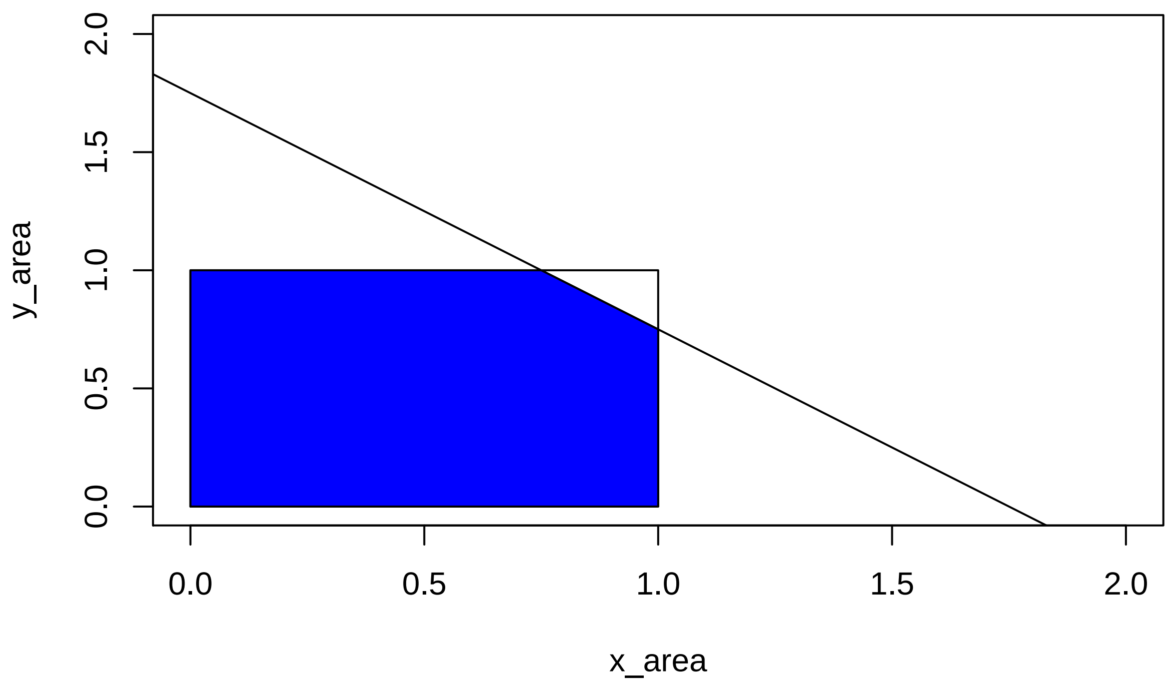



```
plot(x_area, y_area, type = "n")

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)

Event Class	Sum of Dices	Events in Class	Pr(Sum of Dices)
House wins	2	(1,1)	$\frac{1}{36}$
	3	(1,2)(2,1)	$\frac{1}{18}$
	12	(6,6)	$\frac{1}{36}$
You win	7	(3,4)(4,3)(5,2)(2,5)(1,6)(6,1)	$\frac{1}{6}$
	11	(5,6)(6,5)	$\frac{1}{18}$
X	4	(2,2)(3,1)(1,3)	$\frac{1}{12}$
	5	(2,3)(3,2)(1,4)(4,1)	$\frac{1}{9}$
	6	(3,3)(4,2)(2,4)(5,1)(1,5)	$\frac{5}{36}$
	8	(4,4)(6,2)(2,6)(3,5)(5,3)	$\frac{5}{36}$
	9	(3,6)(6,3)(5,4)(4,5)	$\frac{1}{9}$
	10	(5,5)(6,4)(4,6)	$\frac{1}{12}$

We can now define

$$\begin{gathered}
E(Y_{Playerwins}) = Pr(Playerwinsinone) * 1 + \\
\sum_{n=0}^{\infty} (Pr(4))^2 (1 - (Pr(4) + Pr(7)))^n (n+2) + \\
\sum_{n=0}^{\infty} (Pr(5))^2 (1 - (Pr(5) + Pr(7)))^n (n+2) + \\
\sum_{n=0}^{\infty} (Pr(6))^2 (1 - (Pr(6) + Pr(7)))^n (n+2) + \\
\sum_{n=0}^{\infty} (Pr(8))^2 (1 - (Pr(8) + Pr(7)))^n (n+2) + \\
\sum_{n=0}^{\infty} (Pr(9))^2 (1 - (Pr(9) + Pr(7)))^n (n+2) + \\
\sum_{n=0}^{\infty} (Pr(10))^2 (1 - (Pr(10) + Pr(7)))^n (n+2)
\end{gathered}$$

$$E(Y_{Housewins}) = Pr(Housewinsinone) * 1 + \sum_{n=0}^{\infty} Pr(X) Pr(7) (1 - Pr(7))^n (n+2)$$

We know:

$$\begin{gathered}
\frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + 4x^3 + \cdots \\
\frac{1}{(1-x)^2} - 1 = 2x + 3x^2 + 4x^3 + \cdots \\
\frac{1}{x(1-x)^2} - \frac{1}{x} = 2 + 3x + 4x^2 + 5x^3 \cdots \\
\frac{1-(1-x)^2}{x(1-x)^2} = 2 + 3x + 4x^2 + 5x^3 \cdots
\end{gathered}$$

Thus:

$$E(Y_{PlayerWins}) = Pr(Playerwin) * 1 + \Pr(4)^2 \frac{1-(Pr(4)+Pr(7))^2}{(Pr(4)+Pr(7))^2(1-(Pr(4)+Pr(7)))} +$$
$$\Pr(5)^2 \frac{1-(Pr(5)+Pr(7))^2}{(Pr(5)+Pr(7))^2(1-(Pr(5)+Pr(7)))} +$$
$$\Pr(6)^2 \frac{1-(Pr(6)+Pr(7))^2}{(Pr(6)+Pr(7))^2(1-(Pr(6)+Pr(7)))} +$$
$$\Pr(8)^2 \frac{1-(Pr(8)+Pr(7))^2}{(Pr(8)+Pr(7))^2(1-(Pr(8)+Pr(7)))} +$$
$$\Pr(9)^2 \frac{1-(Pr(9)+Pr(7))^2}{(Pr(9)+Pr(7))^2(1-(Pr(9)+Pr(7)))} +$$
$$\Pr(10)^2 \frac{1-(Pr(10)+Pr(7))^2}{(Pr(10)+Pr(7))^2(1-(Pr(10)+Pr(7)))} +$$

$$E(Y_{HouseWins}) = Pr(Housewin) * 1 + \Pr(X)Pr(7)\frac{1-(Pr(7)+Pr(X))^2}{(1-(Pr(7)+Pr(X)))(Pr(7)+Pr(X))^2}$$

```
pr.house.wins.in.one <- 4/36
pr.player.wins.in.one <- 8/36
pr.seven <- 6/36
pr.x.events <- c(3/36, 4/36, 5/36/ 5/36, 4/36, 3/36)
pr.x.plus.seven.events <- pr.x.events + pr.seven
exp.y.player.wins <- pr.player.wins.in.one +
    sum((pr.x.events)^2 * (1 - pr.x.plus.seven.events^2)/((1- pr.x.plus.seven.events)*pr.x.plus.seven.e

exp.y.house.wins <- pr.house.wins.in.one +
    pr.seven*sum(pr.x.events)*(1 - (pr.seven + sum(pr.x.events))^2)/
    ((1 - (pr.seven + sum(pr.x.events)))*(pr.seven + sum(pr.x.events))^2)

print(sprintf("E(Y_Player_Wins) = %f", exp.y.player.wins))
```

```
## [1] "E(Y_Player_Wins) = 0.908914"
```

```
print(sprintf("E(Y_House_Wins) = %f", exp.y.house.wins))
```

```
## [1] "E(Y_House_Wins) = 0.437681"
```

2)

$$E(Payoff) = 100 * Pr(Y=1) + 80 * Pr(Y=2) + 60 * Pr(Y=3) + 40 * Pr(Y=4) + 0 * Pr(Y=5)$$
$$Pr(Y=1) = Pr(Playerwinsinone)$$

Question 8

$\mathrm{E}(\mathrm{Y}_1) = \mathrm{E}(\mathrm{Y}_2) = \ldots = \mathrm{E}(\mathrm{Y}_n) = \mu$
$\mathrm{Var}(\mathrm{Y}_1) = \mathrm{Var}(\mathrm{Y}_2) = \ldots = \mathrm{Var}(\mathrm{Y}_n) = \sigma^2$

1)
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)

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

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

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

3)
We know that:

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

Knowing also that:

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

We conclude:

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

When W is unbiased, we know that:

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

Thus:

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

and

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

Question 9

W$_1$ = $(\frac{n-1}{n})\bar{Y}$
W$_2$ = k$\bar{Y}$

1)

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

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

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

$$\mathbf{bias(W_1)} = \frac{\mu}{\mathbf{n}}$$

Similarly:

$$bias(W2) = E(k\bar{Y}) - \mu$$

$$bias(W2) = kE(\bar{Y}) - \mu$$

$$bias(W2) = k\mu - \mu$$

$$\mathbf{bias(W2)} = \mu(\mathbf{k-1})$$

2)

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

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

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

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

Similarly:

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

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

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

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

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:

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

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

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

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:

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

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

3)

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

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

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

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

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

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

$$\widehat{\sigma^2} = \frac{1}{n}(n\sigma^2 + n\frac{\sigma^2}{n})$$

$$\widehat{\sigma^2} = \mathbf{\frac{n+1}{n}} \sigma^{\mathbf{2}}$$

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

5)
We know that:

$$\widehat{\sigma^2} = \frac{n+1}{n}\sigma^2$$

is a biased estimator. Thus:

$$\frac{n}{n+1}\widehat{\sigma^2}$$

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)

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

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

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

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}):

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

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

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

$WeconcludethatW_1isunbiasedfor\theta$

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

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

We know that

$$E(Y) = E(\bar{Y}) = E[E(Y|X)] = E[\theta X] = \theta E(X)$$

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

$$\Rightarrow \mathbf{E}(\mathbf{W_2}) = \theta$$

Question 12

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

$$n = 900$$

$$\bar{Y} = -32.8$$

$$s = 466.4$$

$$t = \frac{\bar{Y}}{\frac{s}{\sqrt{n}}} = -2.150538$$

$$p(z \leq t) = 0.0158$$

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

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

The value of R also confirms the small practical effect despite the test being statistically significant because of the high sample size.

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:

$$z_Y = \frac{\frac{188}{429} - .5}{se(\bar{Y})} = -2.57861$$

$$\Rightarrow p(z) = 0.0058$$

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