STAT2001/6039 1st Semester 2017 Mid-Semester Exam Solutions

Solution to Problem 1

(a) Let: C_i = "The committee has exactly i men"

S_i = "The sub-committee has exactly i men".

Then the required probability is

$$P(\overline{C_0 \cup C_1} \cap \overline{S_3}) = 1 - P(\overline{\overline{C_0 \cup C_1} \cap \overline{S_3}})$$

$$= 1 - P((\overline{\overline{C_0 \cup C_1}}) \cup \overline{\overline{S_3}}) \quad \text{by De Morgan's Laws}$$

$$= 1 - P(C_0 \cup C_1 \cup S_3)$$

$$= 1 - \{P(C_0) + P(C_1) + P(S_3)\} + \{P(C_0C_1) + P(C_0S_3) + P(C_1S_3)\} - P(C_0C_1S_3)$$

by the generalised additive law of probability for three events

$$= 1 - \{P(C_0) + P(C_1) + P(S_3)\} + \{0 + 0 + 0\} - 0$$

$$= 1 - \frac{\binom{10}{0}\binom{12}{7}}{\binom{22}{7}} - \frac{\binom{10}{1}\binom{12}{6}}{\binom{22}{7}} - \frac{\binom{10}{3}\binom{12}{0}}{\binom{22}{3}}$$

= 1 - 0.004644 - 0.054180 - 0.077922 - 0.863254

= $\boxed{\textbf{0.8633}}$.

(b) The required probability is

$$P(C_6 \mid S_2) = \frac{P(C_6)P(S_2 \mid C_6)}{P(S_2)} = \left\{ \frac{\binom{10}{6}\binom{12}{1}}{\binom{22}{7}} \times \frac{\binom{6}{2}\binom{1}{1}}{\binom{7}{3}} \right\} \Bigg/ \frac{\binom{10}{2}\binom{12}{1}}{\binom{22}{3}}$$

$$= \frac{0.014776245 \times 0.428571429}{0.350649351}$$

= $\boxed{\textbf{0.01806}}$.

(c) Let Y be the number of women on the sub-committee. Then

$$EY = P(C_0)E(Y \mid C_0) + P(\overline{C}_0)E(Y \mid \overline{C}_0),$$

where: $Y \sim Hyp(22,12,3) \;\Rightarrow\; EY = 3\times 12/22 = 1.636364$

$$E(Y \mid \overline{C}_0) = 3, \qquad P(C_0) = \binom{10}{0}\binom{12}{7}\Big/\binom{22}{7} = 0.004643963.$$

So $E(Y \mid \overline{C}_0) = \dfrac{EY - P(C_0)E(Y \mid C_0)}{1 - P(C_0)} = \dfrac{1.636364 - 0.004643963\times 3}{1 - 0.004643963} = \boxed{\mathbf{1.630}}$.

Note: As one might expect, $E(Y \mid \overline{C}_0)$ is slightly less than EY = 1.636, since reducing the number of women on the committee (by at least 1) tends to lower the number of women on the sub-committee (which is selected from the people on the committee).

R Code for Problem 1 (Not required)

```
# (a)
PC0=choose(12,7)/choose(22,7)
PC1=10*choose(12,6)/choose(22,7)
PS3=choose(10,3)/choose(22,3)
c(PC0,PC1,PS3,1-PC0-PC1-PS3)
# 0.004643963 0.054179567 0.077922078 0.863254393

# (b)
PC6=choose(10,6)*12/choose(22,7)
PS2gC6=choose(6,2)/choose(7,3)
PS2=choose(10,2)*12/choose(22,3)
c(PC6,PS2gC6,PS2,PC6*PS2gC6/PS2)
# 0.01477625 0.42857143 0.35064935 0.01805986

# (c)
(3*12/22 - PC0*3) / (1-PC0) # 1.630001
```

Solution to Problem 2

(a) Let Y be the number of rolls. Then the possible outcomes for which $Y = 6$ are:

123456, 132456, …, 654321.

There are 6! such outcomes, out of a total of 6^6 possibilities for the first 6 rolls.

It follows that $P(Y=6) = \frac{6!}{6^6} = \frac{5\times4\times3\times2}{6^5} = \frac{20}{6^4} = \frac{5}{324} = \boxed{\mathbf{0.01543}}$.

(b) Consider the following possible outcomes for which $Y = 7$:

1234556, 1235456, …, 5543216.

These outcomes all end with a 6 and have two 5s prior to that. The number of such outcomes is 6!/2. This number is obtained by first considering the number of possible arrangements of 1, 2, 3, 4, 5 and 0 in a row, which is 6!, and then noting (for example) that the *two* outcomes 123450 and 123405 are equivalent if each 0 is changed to 5.

Next, consider the following *other* possible outcomes for which $Y = 7$:

1234665, 1236465, …, 6643215.

These all end with a 5 and have two 6s prior to that. There are 6!/2 of these outcomes.

Now consider that for $Y = 7$ there are 6 possibilities for *the last number*, and then, for each of these, 5 possibilities for *the number appearing twice*, before the last number. We see that the total number of possibilities for which $Y = 7$ is $6\times5\times6!/2$.

Also, the total number of possibilities for the first 7 rolls of the die is 6^7.

It follows that $P(Y=7) = \frac{6\times5\times6!/2}{6^7} = \frac{5}{2}\times\frac{6!}{6^6} = \frac{5}{2}p_6 = \frac{25}{648} = \boxed{\mathbf{0.03858}}$.

(c) After one roll, the number of rolls until a different number comes up is geometric with parameter 5/6 and mean 6/5. Then, the number of rolls until the next new number is geometric with parameter 4/6 and mean 6/4, etc. We see that Y has expected value

$$1+\frac{6}{5}+\frac{6}{4}+\ldots+\frac{6}{1} = \frac{6}{60}(10+12+15+20+30+60) = \frac{147}{10} = \boxed{\mathbf{14.7}}.$$

Note: Another way to do parts (a) and (b) is as follows. The cdf of Y is

$$F(y) = P(Y \le y) = P(A_1 \dots A_6) \quad \text{where } A_i = \text{"At least one } i \text{ in first } y \text{ rolls"}$$

$$= 1 - P(\overline{A_1 \dots A_6})$$

$$= 1 - P(B_1 \cup \dots \cup B_6) \quad \text{where } B_i = \overline{A_i} = \text{"Absence of } i \text{ in first } y \text{ rolls"}$$

$$= 1 - \left\{ \sum_i P(B_i) - \sum_{i<j} P(B_i B_j) + \dots - \sum_{i<j<k<l<m<n} P(B_i \dots B_n) \right\}$$

$$= 1 - \left\{ \binom{6}{1} P(B_1) - \binom{6}{2} P(B_1 B_2) + \dots - \binom{6}{6} P(B_1 \dots B_6) \right\}$$

$$= 1 - 6\left(\frac{5}{6}\right)^y + 15\left(\frac{4}{6}\right)^y - 20\left(\frac{3}{6}\right)^y + 15\left(\frac{2}{6}\right)^y - 6\left(\frac{1}{6}\right)^y + 1 \times 0 .$$

Evaluation this function, we find that $F(6)$ = 0.01543 and $F(7)$ = 0.05401.

Consequently, $f(6) = F(6)$ = 0.01543 and $f(7) = F(7) - F(6)$ = 0.03858.

Solution to Problem 3

Let J = "Jim will win". Then, applying a first step analysis, we have that

$$P(J) = P(H)P(J \mid H) + P(T)P(J \mid T)$$

$$= (1/2)P(J \mid H) + (1/2)[1 - P(J)] .$$

By a similar logic, considering the 2nd and 3rd tosses in turn, we also have that:

$$P(J \mid H) = P(HH \mid H)P(J \mid HH) + P(HT \mid H)P(J \mid HT)$$

$$= (1/2)P(J \mid HH) + (1/2)P(J \mid HT)$$

$$P(J \mid HH) = P(HHH \mid HH)P(J \mid HHH) + P(HHT \mid HH)P(J \mid HHT)$$

$$= (1/2) \times 1 + (1/2)[1 - P(J \mid HT)]$$

$$P(J \mid HT) = P(HTH \mid HT)P(J \mid HTH) + P(HTT \mid HT)P(J \mid HTT)$$

$$= (1/2)1 + (1/2)[1 - P(J)] .$$

Writing $P(J)$ as p, $P(J \mid H)$ as q, $P(J \mid HH)$ as r, and $P(J \mid HT)$ as s, we obtain:

$$p = (1/2)q + (1/2)[1 - p]$$

$$q = (1/2)r + (1/2)s$$

$$r = (1/2) + (1/2)(1 - s)$$

$$s = (1/2) + (1/2)(1 - p) .$$

Solving these equations yields the required probability, $p = P(J) = \boxed{\textbf{14/25}}$.

Note 1: Since Jim goes first, he has a better than even chance of winning ($p > 1/2$).

Note 2: We also find that:

$$q = P(J \mid H) = 17/25$$
$$r = P(J \mid HH) = 16/25$$
$$s = P(J \mid HT) = 18/25.$$

Thus, getting heads on the first toss improves Jim's chances ($q > p$). It also makes intuitive sense that $q = P(J \mid H)$ is the average of $r = P(J \mid HH)$ and $s = P(J \mid HT)$.

Note 3: Another working is to simultaneously consider the next two tosses of the coin given that heads come up first. This logic (plus the first equation above) yields:

$$P(J) = P(H)P(J \mid H) + P(T)P(J \mid T)$$
$$= (1/2)P(J \mid H) + (1/2)[1 - P(J)]$$
$$P(J \mid H) = P(HTT \mid H)P(J \mid HTT) + P(HTH \mid H)P(J \mid HTH)$$
$$+P(HHT \mid H)P(J \mid HHT) + P(HHH \mid H)P(J \mid HHH)$$
$$= (1/4)[1 - P(J)] + (1/4) \times 1$$
$$+(1/4)[1 - P(J \mid HT)] + (1/4) \times 1$$
$$P(J \mid HT) = P(HTH \mid HT)P(J \mid HTH) + P(HTT \mid HT)P(J \mid HTT)$$
$$= (1/2) \times 1 + (1/2)[1 - P(J)].$$

Writing $P(J)$ as p, $P(J \mid H)$ as q, and $P(J \mid HT)$ as s (as previously), we obtain:

$$p = (1/2)q + (1/2)[1 - p]$$
$$q = (1/4)(1 - p) + (1/4) + (1/4)(1 - s) + (1/4)$$
$$s = (1/2) + (1/2)(1 - p).$$

We now have three equations in three unknowns (rather than four equations in four unknowns as previously). Solving these equations yields the same solution, $p = 14/25$.

Solution to Problem 4

(a) Let "0" denote nobody in the Waskit family having criplea, let "3" denote all three persons in the Waskit family having criplea, etc. Then the required probability is

$$P(3\,|\,\bar{0}) = \frac{P(3\bar{0})}{P(\bar{0})} = \frac{P(3) - P(30)}{P(\bar{0})} = \frac{P(3) - 0}{1 - P(0)} = \frac{(1/5)^3}{1 - (4/5)^3} = \frac{1}{61} = \boxed{\mathbf{0.01639}}.$$

Note: As one might expect, this is greater than the *unconditional* probability that all three Waskit family members have criplea, namely $P(3) = (1/5)^3$ = 1/125 = 0.00800.

(b) Let R be the event that a randomly chosen member of the Waskit family has criplea. Then the required probability is

$$P(3\,|\,R) = \frac{P(3)P(R\,|\,3)}{P(R)} = \frac{(1/5)^3 \times 1}{1/5} = (1/5)^2 = 1/25 = \boxed{\mathbf{0.04}}.$$

Note 1: This may also be deduced more simply as follows. By independence, the probability that both of the other two members of the family have criplea is $(1/5)^2$.

Note 2: It may seem that **(a)** and **(b)** should have the same answer. But there is a difference between the two types of conditioning. This difference can be clarified by considering the following parallel but simpler scenario. Two fair coins are tossed and covered up. Find the probability that *both coins have come up heads* given that:

(i) *at least one of the coins has come up heads*; and

(ii) *we uncover one of the coins randomly, and see that it has come up heads*.

To answer **(i)** we consider the possibilities TT, TH, HT and TT, exclude TT, and so obtain 1/3. The answer to **(ii)** is the probability that the other coin shows heads: 1/2.

Solution to Problem 5

(a) The mgf of X is $m(t) = Ee^{Xt} = \sum_{x=\pm1,\pm2,\ldots} e^{xt}\frac{1}{2^{|x|+1}} = \sum_{x=1}^{\infty} e^{xt}\frac{1}{2^{x+1}} + \sum_{x=-\infty}^{-1} e^{xt}\frac{1}{2^{-x+1}}$

$$= \frac{1}{2}\sum_{x=1}^{\infty}\left(\frac{e^t}{2}\right)^x + \frac{1}{2}\sum_{x=-\infty}^{-1}\left(2e^t\right)^x$$

$$= \frac{1}{2}\sum_{x=1}^{\infty}\left(\frac{e^t}{2}\right)^x + \frac{1}{2}\sum_{r=1}^{\infty}\left(\frac{1}{2e^t}\right)^r \qquad \text{(where } r = -t\text{)}$$

$$= \frac{1}{2}\left(\frac{1}{1-e^t/2} - 1\right) + \frac{1}{2}\left(\frac{1}{1-1/(2e^t)} - 1\right).$$

Thus, $\boxed{m(t) = (2-e^t)^{-1} + (2-e^{-t})^{-1} - 1}$, which equals $\boxed{\mathbf{1.0306}}$ at $t = 0.1$.

Note: We may also express the mgf in other ways, such as $m(t) = \left(\frac{3}{e^t + e^{-t} - 1} - 2\right)^{-1}$.

(b) First, $m'(t) = -(2-e^t)^{-2}(-e^t) - (2-e^{-t})^{-2}e^{-t} - 0 = e^t(2-e^t)^{-2} - e^{-t}(2-e^{-t})^{-2}$.

So the expected value of X is equal to $\mu = m'(0) = 1(2-1)^{-2} - 1(2-1)^{-1} = \boxed{0}$.

Also, $m''(t) = e^t(-2)(2-e^t)^{-3}(-e^t) + e^t(2-e^t)^{-2} - e^{-t}(-2)(2-e^{-t})^{-3}e^{-t} + e^{-t}(2-e^{-t})^{-1}$

So $\mu_2' = m''(0) = 1(-2)(2-1)^{-3}(-1) + 1(2-1)^{-2} - 1(-2)(2-1)^{-3}1 + 1(2-1)^{-1}$

$$= 2+1+2+1 = 6.$$

It follows that $\sigma^2 = \mu_2' - \mu^2 = \boxed{6}$.

Note: Alternatively, and more simply, we have that $EX = 0$ (by symmetry), and then

$$VX = EX^2 = \sum_{x=\pm1,\pm2,\ldots} x^2\frac{1}{2^{|x|+1}} = \frac{1}{2}\sum_{x=\pm1,\pm2,\ldots} x^2\frac{1}{2^{|x|}} = \frac{1}{2}\left(2\sum_{x=1,2\ldots} x^2\frac{1}{2^x}\right) \text{ by symmetry}$$

$$= \sum_{y=1,2\ldots} y^2\frac{1}{2^y} \quad \text{(after changing } x \text{ to } y\text{)}$$

$$= EY^2 \quad \text{where } Y \sim Geo(1/2)$$

$$= VY^2 + (EY)^2$$

$$= \frac{1/2}{1-(1/2)^2} + \left(\frac{1}{1/2}\right)^2 = 2+4 = 6.$$

R Code for Problem 5 (Not required)

```
tval=0.1; mval=1/(2-exp(tval)) + 1/(2-exp(-tval)) - 1; mval # 1.030638
1/(-2+3/(exp(tval)+exp(-tval)-1))  # 1.030638

X11(w=8,h=5)

tv=seq(-0.25,0.25,0.001); mv = 1/(2-exp(tv)) + 1/(2-exp(-tv)) - 1

plot(tv,mv,type="l", xlim=c(-0.2,0.2), lwd=3,
        xlab="t", ylab="m(t)", ylim=c(1,1.2)  )

abline(h=seq(1,1.2,0.01),lty=3);  abline(v=seq(-0.2,0.2,0.05),lty=3)

points(tval, mval,pch=16, cex=1.5)
```