STAT 104
Solution Set to Assignment 3

(Marking scheme: 35 points for questions with 5 bonus points for challenge: Total = 40)

4.6) (1 point)
Using the Applet

a) After the first 40 tosses, the proportion of heads is 0.38 .
The count of heads is 15 out of 40.
The difference between the count of heads and 20 is 5.

b) After the first 120 tosses, the proportion of heads is 0.43 .
The count of heads is 51 out of 120.
The difference between the count of heads and 60 is 9.

c) After the first 240 tosses, the proportion of heads is 0.45 .
The count of heads is 108 out of 240.
The difference between the count of heads and 120 is 12.

After the first 480 tosses, the proportion of heads is 0.48 .
The count of heads is 229 out of 480.
The difference between the count of heads and 240 is 11.

(1 point for correct account of results)

4.10) (4 points)

a) S is the random variable for the gender of the subject.
S = { Male , Female }
(1 point)

b) E is the r.v. for effort on the RPE scale after 10 minutes of exercise.
E = {6, 7, 8, …. , 19, 20 }
(1 point –showing of discrete events or use of … implying integers in between)

c) V is the r.v. for max volume of oxygen consumed per minute during exercise.
V = $\{ [2.5 , 6] \} \ni \mathfrak{R}$
(i.e. the set of all the real numbers between 2.5 and 6, inclusive)
(1 point – mention of continuous range, or reference to real numbers)

d) M is the r.v. for the maximum heart rate, measured in beats per minute.
M = $\{ \mathfrak{Z}^{+} \}$
(i.e. the set of all positive integers, assuming that measured heart rate is rounded up to the 'nearest beat per minute')
(1 point – for above or anything sensible / closely equivalent)

4.14) (2 points)

P (American has O type blood) = 0.45
P (Chinese has O type blood) = 0.35

i) P (Both American **and** Chinese have O type blood)
= P (American has O type blood) * P (Chinese has O type blood)
= 0.45 * 0.35
= 0.1575
(1 point)

ii) P (Both American **and** Chinese have same type blood)

= P (Both American **and** Chinese have O type blood **or**
Both American **and** Chinese have A type blood **or**
Both American **and** Chinese have B type blood **or**
Both American **and** Chinese have AB type blood)

= 0.45*0.35 + 0.4*0.27 + 0.11*0.26 + 0.04*0.12
= 0.2989
(1 point)

4.16) (2 points for full table. No need to show sum of probabiltiies)

	Probability
Michigan State	0.3000
Michigan	0.0000
Minnesota	0.0000
Northwestern	0.0000
Penn State	0.0000
Iowa	0.1556
Illinois	0.1556
Purdue	0.1556
Indiana	0.0778
Ohio State	0.0778
Wisconsin	0.0778
Sum Probabilities	1.0000

4.20) (4 points – 1 for each part)

a) The sum of the probabilities for all the categories is 1.0, which confirms legitimate assignment.

b) P(A) = 0.000 + 0.003 + 0.060 + 0.062
= 0.125

c) $\mathbf{B^c}$ means the complement of event B, which in turn means 'the set of all events except for event B'. Literally, this means everyone who is not white.

$\mathbf{P(B^c) =}$ 1- P(B) (using the complement rule)
P(B) = 0.06 + 0.691 = 0.751
Therefore $\mathbf{P(B^c) = 0.249}$

d) "The person chosen is a non-hispanic white" can be expressed as the intersection of events of $\mathbf{A^c \cap B}$.
$\mathbf{P(A^c \cap B) = 0.691}$

4.32) (2 points – 1 for each part)

a) P (None of the 5 calls reaches a live person) = $\mathbf{(1\text{-}0.2)^5 = 0.32768}$
b) P (None of the 5 calls made to NYC reaches a person) = $\mathbf{(1\text{-}0.08)^5 = 0.65908}$

4.40) (3 points – 1 for each part)

b) Discrete, as the RPE scale can only take integer values between 6 to 20 inclusive.

c) Continuous, as the oxygen consumed per minute (measured in l/min) can take any real number value between 2.5 to 6 inclusive.

d) Depends on whether fractional 'beats per minute' are allowed in measurements. If the measurements are integer numbers only then in is discrete. If fractional measurements are allowed then it is continuous.

4.50) (4 points – 1 for each part)

a) P(A and B support funding and C opposes it) = 0.6 * 0.6 * 0.4 = **0.144**

b)

Student A	Student B	Student C	Probability of combination
Supports	Supports	Supports	**0.216**
Supports	Supports	**Oppose**	**0.144**
Supports	**Oppose**	Supports	**0.144**
Supports	**Oppose**	**Oppose**	**0.096**
Oppose	Supports	Supports	**0.144**
Oppose	Supports	**Oppose**	**0.096**
Oppose	**Oppose**	Supports	**0.096**
Oppose	**Oppose**	**Oppose**	**0.064**

c)

X	Probability
3	**0.064**
2	**0.288**
1	**0.432**
0	**0.216**

d) The majority of the advisory board opposes funding occurs when X is greater than or equal to 2.

P (X >= 2) = 0.288 + 0.064 **= 0.352**

4.56) (2 points – 1 for each part)

a) **P(p hat >= 0.16)** = 1 - ϕ^{-1}[(0.16-0.15)/0.0092] = **13.85%**

b) **P(0.14 <= p hat <= 0.16) =** ϕ^{-1}[(0.16-0.15)/0.0092] - ϕ^{-1}[(0.14-0.15)/0.0092] **=72.29%**

1.82) (2 points – 1 for each part)

mean = 266
s.d. = 16

a) Between 234and 298 days (ie between the 2.5% and 97.5% tail ends, which is approximately $\mu \pm 2\sigma$)
Note more precisely, it should be $\mu \pm 1.96\sigma$ which gives an interval of 234.64 to 297.36.

b) The shortest 2.5% of pregnancies are 234 days or less.
The longest 2.5% of pregnancies last 298 days or greater.

1.86) (1 point)

Eleanor's z-score = (680-500)/100 = 1.8
Gerald's z-score = (27-18)/6 = 1.5

Eleanor's z-score is higher, and hence she has the higher score.

1.94) **(1 point)**

Percentage of scores above 800 = 1 - ϕ^{-1}[(800-533)/115] = **1.0123%**

1.100) **(2 points)**

1 - ϕ^{-1}[(475-mean)/s.d.] = 10%
(475-mean)/s.d. = 1.282

ϕ^{-1}[(25-mean)/s.d.] = 10%
(25-mean)/s.d. = -1.282

Simultaneous equations…

500 – 2*mean = 0
mean = 250
s.d. = (475-250)/1.282 = **175.6**

1.106) **(4 points – 1 for each part)**

a) Area is 25%.
z-value for 1st quartile = -0.67449
z-value for 3rd quartile = +0.67449
b)
Score for 1st quartile = 100+(-0.67449*15) = 89.88
Score for 2nd quartile = 100
Score for 3rd quartile = 100+(+0.67449*15) = 110.12

c) IQR = 2*0.67449 = 1.3490

c) Q1 – 1.5 IQR = -0.674 – 1.5 x 1.349 = -2.698
Q3 + 1.5 IQR = 0.674 + 1.5 x 1.349 = 2.698
P[Z = -2.70] = 0.0035, P[Z = 2.70] = 0.0035
Which implies the proportion of outliers is about 0.007

1.108) (1 point)

This normal quantile plot suggests that the data is not normally distributed since the points fall on a curved pattern. The plot suggests that the data set has light tails since the extreme values aren't as extreme as would be expected with a normal distribution.