Solution to HW 12

(1) (10 pts) Sec 12.3 Problem 16

A screening test for a disease shows a positive result in 92% of all cases when the disease is actually present and in 7% of all cases when it is not. Assume that the prevalence of the disease is 1 in 600. If the test is administered to a randomly chosen individual, what is the probability that the result is positive?

Solution: Let A be the event that the test is positive, B_1 be the event that a person is infected and B_2 be the event that a person is not infected. From the condition that the prevalence of the disease is 1 in 600, we have $P(B_1) = \frac{1}{600}$ and $P(B_2) = 1 - P(B_1) = 1 - \frac{1}{600} = \frac{599}{600}$. From the condition that a positive result in 92% of all cases when the disease is actually present , we have $P(A|B_1) = \frac{92}{100}$. From the condition that a positive result in 7% of all cases when the disease is not present , we have $P(A|B_2) = \frac{7}{100}$. Since B_1 and B_2 is a partition, we have $P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) = \frac{92}{100} \cdot \frac{1}{600} + \frac{7}{100} \cdot \frac{599}{600} = \frac{92+7\cdot 599}{60000} = \frac{4285}{60000} \cong 0.07141666666$. So the probability that the result is positive is about 0.07141666666.

(2) (10 pts) Sec 12.3 Problem 18

A screening test for a disease shows a positive test result in 95% of all cases when the disease is actually present and in 20% of all cases when it is not. When the test was administered to a large number of people, 21.5% of the results were positive. What is the prevalence of the disease?

Solution: Let A be the event that the test is positive, B_1 be the event that a person is infected and B_2 be the event that a person is not infected. Assume $P(B_1) = P/100$ and $P(B_2) = (100 - P)/100$. From the condition that a screening test for a disease shows a positive test result in 95% of all cases when the disease is actually present, we have $P(A|B_1) = \frac{95}{100}$. From the condition that a screening test for a disease shows a positive test result in 20% of all cases when the disease is not present, we have $P(A|B_2) = \frac{20}{100}$. From the condition that 21.5% of the results were positive, we have $P(A) = \frac{21.5}{100}$. Since B_1 and B_2 form a partition, we have

$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2)$. Using $P(A) = \frac{21.5}{100}$, $P(A|B_1) = \frac{95}{100}$, $P(B_1) = P/100$, $P(A|B_2) = \frac{20}{100}$ and $P(B_2) = (100 - P)/100$, we have

$\frac{21.5}{100} = \frac{95}{100} \cdot \frac{P}{100} + \frac{20}{100} \cdot \frac{100-P}{100}$,

$\frac{21.5}{100} = \frac{95P+2000-20P}{10000}$,

$2150 = 75P + 2000$ and $150 = 75P$. Thus $P = 2$. So the prevalence of the disease is 1 in 50 (which is 2/100=1/50).

(3) (20 pts) A screening test for a disease shows a positive result in 95% of all cases when the disease is actually present and in 10% of all cases when it is not. Assume that the prevalence of the disease in the population is $1/50$.

(a) Find the probability that person has the disease when the test is positive.

Solution: Let A be the event that the test is positive, B_1 be the event that a person is infected and B_2 be the event that a person is not infected. From the conditions given by the problems, we have $P(A|B_1) = \frac{95}{100}$, $P(A|B_2) = \frac{10}{100}$ and $P(B_1) = \frac{1}{50}$. Since B_1 and B_2 form a partition, we have $P(B_1) + P(B_2) = 1$, Using $P(B_2) = 1 - P(B_1) = 1 - \frac{1}{50} = \frac{49}{50}$. From the Bayes formula, we have $P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1)+P(A|B_2)P(B_2)} = \frac{\frac{95}{100}\cdot\frac{1}{50}}{\frac{95}{100}\cdot\frac{1}{50}+\frac{10}{100}\cdot\frac{49}{50}} = \frac{95}{95+10\cdot 49} = \frac{95}{95+490} = \frac{95}{495} \cong 0.19192$ So the probability that person has the disease when the test is positive is about 0.19192.

(b) Find the probability that person has the disease when the test is negative.

Solution: In this problem, we want to find $P(B_1|A^c)$ which can be found by the Bayes formula, we have $P(B_1|A^c) = \frac{P(A^c|B_1)P(B_1)}{P(A^c|B_1)P(B_1)+P(A^c|B_2)P(B_2)} = \frac{\frac{5}{100}\cdot\frac{1}{50}}{\frac{5}{100}\cdot\frac{1}{50}+\frac{90}{100}\cdot\frac{49}{50}} = \frac{5}{5+90\cdot 49} = \frac{5}{5+4410} = \frac{5}{4415} \cong 0.0011325$ So the probability that person has the disease when the test is negative is about 0.0011325.

(c) Find the probability that person doesn't have the disease when the test is negative.

Solution: In this problem, we want to find $P(B_2|A^c)$ which can be found by the Bayes formula, we have $P(B_2|A^c) = \frac{P(A^c|B_2)P(B_2)}{P(A^c|B_1)P(B_1)+P(A^c|B_2)P(B_2)} = \frac{\frac{90}{100}\cdot\frac{49}{50}}{\frac{5}{100}\cdot\frac{1}{50}+\frac{90}{100}\cdot\frac{49}{50}} = \frac{5}{5+90\cdot 49} = \frac{4410}{5+4410} = \frac{4410}{4415} \cong 0.99887$ So the probability that person has the disease when the test is negative is about 0.99887.

(4) (10 pts) Sec 12.4 Problem 2

Toss a fair coin four times. Let X be the random variable that counts the number of heads. Find the probability mass function describing the distribution of X.

Solution: The sample space is
$\Omega = \{TTTT, HTTT, THTT, TTHT, TTTH, HHTT, HTHT, HTTH, THHT, THTH, TTHH, THHH, HTHH, HHTH, HHHT, HHHH\}$. We have $|\Omega| = 16$. Since $\{x|X(x) = 0\} = \{TTTT\}$, we have $P(X = 0) = \frac{1}{16}$.

Since $\{x|X(x) = 1\} = \{HTTT, THTT, TTHT, TTTH\}$, we have $P(X = 1) = \frac{4}{16}$. Since $\{x|X(x) = 2\} = \{HHTT, HTHT, HTTH, THHT, THTH, TTHH\}$, we have $P(X=2)=\frac{6}{16}$. Since $\{x|X(x)=3\}=\{THHH,HTHH,HHTH,HHHT\}$, we have $P(X=3)=\frac{4}{16}$. Since $\{x|X(x)=4\}=\{HHHH\}$, we have $P(X=4)=\frac{1}{16}$.

(5) (10 pts) Sec 12.4 Problem 10
Suppose the probability mass function of a discrete random variable X is given by the following table:

x	$P(X=x)$
-1	0.2
-0.5	0.25
0.1	0.1
0.5	0.1
1	0.35

Find and graph the corresponding distribution function $F(x)$.
Solution: The distribution function id

$$F(x) = \begin{array}{rl} 0 & x<-1 \\ 0.2 & -1\le x<-0.5 \\ 0.45 & -0.5\le x<0.1 \\ 0.55 & 0.1\le x<0.5 \\ 0.65 & 0.5\le x<1 \\ 1 & 1\le x \end{array}$$

(6) (10 pts) Sec 12.4 Problem 12
Let X be a random variable with distribution function

$$F(x) = \begin{array}{rl} 0 & x<0 \\ 0.05 & 0\le x<1.3 \\ 0.30 & 1.3\le x<1.7 \\ 0.85 & 1.7\le x<1.9 \\ 0.90 & 1.9\le x<2 \\ 1.0 & 2\le x \end{array}$$

Determine the probability mass function of X.
Solution:

$$P(x) = \begin{cases} 0.05 & x = 0 \\ 0.25 & x = 1.3 \\ 0.55 & x = 1.7 \\ 0.35 & x = 1.9 \\ 0.1 & x = 2 \end{cases}$$

(7) (15 pts) Sec 12.4 Problem 16
The following table contains the per plant in a sample of size 30:

15	27	13	2	0	16
26	0	2	1	17	15
21	13	5	0	19	25
12	11	0	16	22	1
12	11	0	16	22	1

(a) Find the relative frequency distribution.
Solution: The relative frequency distribution is

number of aphids	15	27	13	2	0	16	26	1	17
relative frequency	2/30	1/30	2/30	2/30	6/30	2/30	1/30	3/30	2/30
number of aphids	21	5	19	25	12	11	22	28	9
relative frequency	1/30	1/30	1/30	1/30	1/30	1/30	1/30	1/30	1/30

(b) Compute the average value by (i) averaging the values in the table directly and (ii) using the relative frequency distribution obtained in (a)
Solution: (i) You should get $\frac{334}{30} \cong 11.13333$.
(ii) The average value is $(2/30)\cdot 15+(1/30)\cdot 27+(2/30)\cdot 13+(2/30)\cdot 2+(6/30)\cdot 0+(2/30)\cdot 16+(1/30)\cdot 26+(3/30)\cdot 1+(2/30)\cdot 17+(1/30)\cdot 21+(1/30)\cdot 5+(1/30)\cdot 19+(1/30)\cdot 25+(1/30)\cdot 12+(1/30)\cdot 11+(1/30)\cdot 22+(1/30)\cdot 28+(1/30)\cdot 9 = \frac{334}{30} \cong 11.13333$

(8) (15 pts) Sec 12.4 Problem 20
Suppose that the probability mass function of a discrete random variable X is given by the following table:

x	$P(X = x)$
0	0.3
1	0.3
2	0.1
3	0.1
4	0.2

(a) Find $E(X)$.
Solution: $E(X) = \sum_x xP(X = x) = 0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2 \cdot P(X = 2) + 3 \cdot P(X = 3) + 4 \cdot P(X = 4)$
$= 0 \cdot 0.3 + 1 \cdot 0.3 + 2 \cdot 0.1 + 3 \cdot 0.1 + 4 \cdot 0.2$
$= 0.3 + 0.2 + 0.3 + 0.8 = 1.6$.

(b) Find $E(X^2)$.
Solution: $E(X^2) = \sum_x x^2 P(X = x) = 0^2 \cdot P(X = 0) + 1^2 \cdot P(X = 1) + 2^2 \cdot P(X = 2) + 3^2 \cdot P(X = 3) + 4^2 \cdot P(X = 4)$
$= 0 \cdot 0.3 + 1 \cdot 0.3 + 4 \cdot 0.1 + 9 \cdot 0.1 + 16 \cdot 0.2$
$= 0.3 + 0.4 + 0.9 + 3.2 = 4.8$.

(c) Find $E(2X - 1)$.
Solution:$E(2X - 1) = \sum_x (2x - 1)P(X = x) = (2 \cdot 0 - 1) \cdot P(X = 0) + (2 \cdot 1 - 1)P(X = 1) + (2 \cdot 2 - 1)P(X = 2) + (2 \cdot 3 - 1)P(X = 3) + (2 \cdot 4 - 1)P(X = 4)$
$= (-1) \cdot 0.3 + 1 \cdot 0.3 + 3 \cdot 0.1 + 5 \cdot 0.1 + 7 \cdot 0.2$
$= -0.3 + 0.3 + 0.3 + 0.5 + 1.4 = 2.2$.