Homework Assignment Number Four Solutions

Problem 1.
You have a computer code that generates random integers in the range 20 to 40, inclusive.
(a) If the random variable, x, is the value of the random integer, what PDF describes the distribution of x?
(b) What is the probability that 25<x<35?
(c) What is the mean value of the random integer?
(d) What is the variance of the integer?

Solution:
(a) This problem requires the discrete uniform PDF.
The situation describes means the numbers one through ten have equal probability to be drawn. Therefore, we use the discrete uniform PDF,

$$f(x;k) = \frac{1}{k}$$

where k = 21 and $x \in \{20,21,22,...39,40\}$.

(b) The probability that 25<x<35 is given by

$$P(25 < \mathrm{x} < 35) = \sum_{\mathrm{x}=25}^{35} f(x;21) = \frac{1}{21}\sum_{\mathrm{x}=26}^{34} 1 = \frac{9}{21} = 0.4286$$

(c) What is the mean value of the random integer?

$$\mu_x = \sum_{i=1}^{k} x_i f(x;k) = \frac{1}{k}\sum_{i=1}^{k} x_i = \frac{1}{21}\sum_{x=20}^{40} x = \frac{1}{21}\left(20+21+22+...+39+40\right) = 30.0$$

(d) What is the variance of the integer?

$$\sigma^2 = \sum_{i=1}^{k}\left(x_i-\mu\right)^2 f(x;k) = \frac{1}{k}\sum_{i=1}^{k}\left(x_i-\mu\right)^2 = \frac{1}{21}\sum_{i=1}^{21}\left(x_i-30.0\right)^2 = 36.6\overline{6}$$

Problem 2.
According to Chemical Engineering Progress (Nov. 1990) approximately 30% of all pipework failures in chemical plants are caused by operator error.
(a) If x is a random variable that describes the number of pipework failures caused by operator error, what PDF will describe x?
(b) What is the probability that out of the next 20 pipework failures, at least 10 are due to operator error?
(c) What is the probability that no more than 4 out of 20 such failures are due to operator error?
(d) What is the probability that out of the next 20 pipework failures, exactly 5 are due to operator error?

Solution:
(a) This problem requires the binomial PDF because it satisfies the three criteria for the binomial distribution, namely (1) each experiment is the same and independent, (2) there are 2 outcomes, and (3) the probability of success, p, is the same for each experiment. p = 0.30 = probability that operator error caused pipework failure. n=20. x = number of successes with a range from 0 to 20.

(b) Find probability that at least 10 of next 20 are successes (operator error caused pipework failure), $P(X \geq 10)$.

In general, we use the formula:

$$P(X \geq 10) = \sum_{x=10}^{20} f(x) = \sum_{x=10}^{20} b(x;n,p) = \sum_{x=10}^{20} b(x;20,0.30)$$

If we have a computer program that calculate the binomial PDF, then this is no problem and we have:

$$P(X \geq 10) = \sum_{x=10}^{20} b(x;20,0.30) = 0.047962$$

If we want to use the tables in the back of Walpole, Myers, and Myers, then we have a tougher job. Since we have a greater than or less than sign in our probability, we know we must use the CUMULATIVE binomial PDF. Because CUMULATIVE PDFs give $P(X \leq x)$, we must change the greater than sign to a less than sign, because that's all the table gives.

$$P(X \geq 10) = P(X \leq 20) - P(X \leq 9)$$

We know that:

$$P(X \leq r) \equiv B(r;n,p) = \sum_{x=0}^{r} b(x;n,p)$$

so

$$P(X \geq 10) = B(20,20,0.3) - B(9,20,0.3)$$

and we know $B(r;n,p)$ are given in Table A.1 of the appendix of WMM. So

$$P(X \geq 10) = 1.0000 - 0.9520 = 0.048$$

(c) Find $P(X \leq 4)$. From Table A.1 of the appendix of WMM

$$P(X \leq 4) = B(4,20,0.3) = 0.2375$$

Or, using the computer program:

$$P(X \leq 4) = \sum_{x=0}^{4} b(x;20,0.30) = 0.237508$$

(d) What is the probability that out of the next N pipework failures, exactly 5 are due to operator error?

$$P(X = x) = b(x;n,p) = \binom{n}{x} p^x q^{n-x}$$

$$P(X = 5) = b(5,20,0.3) = \binom{20}{5} 0.3^5 0.7^{15} = 0.1789$$

Problem 3.
According to genetics theory, a certain cross of guinea pigs will result in red, black, and white offspring in the ratio 8:4:8.

(a) If {x} is a set of random variables that describes the number of red, black, and white offspring, what PDF describes {x}?
(b) Find the probability that among 8 offspring 5 will be red, 2 black, and 1 white.

Solution:
(a) This problem requires the multinomial distribution because it satisfies the three multinomial criteria: (1) each experiment is the same and independent, (2) there are 3 distinct outcomes, and (3) the probability of success, p, is the same for each experiment.
The first outcome, x_1, is red, with $p_1 = 0.4$. The second outcome, x_2, is black, with $p_2 = 0.2$. The third outcome, x_3, is white, with $p_3 = 0.4$.
(b) Find the probability that for n=8, x_1=5, x_2=2, x_3=1.
For the multinomial distribution:

$$P(\{X = x\}) = m(\{x\}; n, \{p\}, k) = \binom{n}{x_1, x_2 ... x_k} \prod_{i=1}^{k} p_i^{x_i}$$

$$P(\{X = x\}) = \binom{8}{5,2,1} 0.4^5 0.2^2 0.4^1 = 0.0275$$

Problem 4.
An urn contains 3 green balls, 2 blue balls, and 4 red balls. A random sample of 5 balls is selected.
(a) If {x} is a set of random variables that describes the number of balls of each color drawn, what PDF describes {x}?
(b) Find the probability that in the sample of 5 balls, both blue balls and at least 1 red ball are selected.

Solution:
This requires the multivariate hypergeometric distribution because the experiment is done without replacement. We have 3 green 2 blue, 4 red balls. We select 5 without replacement. Find the probability that in that 5 you have 2 blue and at least 1 red.
The multivariate hypergeometric distribution is:

$$h(\{x\}; N, n, \{a\}) = \frac{\binom{a_1}{x_1}\binom{a_2}{x_2}\binom{a_3}{x_3}\cdots\binom{a_k}{x_k}}{\binom{N}{n}}$$

What possible combinations of 5 will satisfy this problem?
(2 blue, 1 red, 2 green) and (2 blue, 2 red, 1 green) and (2 blue, 3 red, 0 green)

$$\frac{\binom{2}{2}\binom{4}{1}\binom{3}{2}}{\binom{9}{5}} + \frac{\binom{2}{2}\binom{4}{2}\binom{3}{1}}{\binom{9}{5}} + \frac{\binom{2}{2}\binom{4}{3}\binom{3}{0}}{\binom{9}{5}} = \frac{1 \cdot 4 \cdot 3 + 1 \cdot 6 \cdot 3 + 1 \cdot 4 \cdot 1}{126} = \frac{17}{63} = 0.2698$$

Problem 5.
Population studies of biology and the environment often tag and release subjects in order to estimate size and degree of certain features in the population. 10 animals of a certain population thought to be near extinction are caught,

tagged and released in a certain region. After a period of time, a random sample of 15 of the types of animals are caught. There are only 25 of the animals in the region.
(a) If x is a random variable that describes the number of animals caught both times, what PDF describes x?
(b) What is the probability that five of the animals caught in the second batch had been caught in the first batch?

Solution:
(a) This problem requires the hypergeometric distribution because the sampling is done without replacement. N=25 animals. k=10 tagged animals. n=15. Find P(x=5). The single variable hypergeometric distribution is

$$h(x;N,n,k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \quad \text{for x = 0,1,2...n}$$

(b)

$$h(5;25,15,10) = \frac{\binom{10}{5}\binom{25-10}{15-5}}{\binom{25}{15}} = 0.2315$$

Problem 6.
Among 150 IRS employees in a large city, only 30 are women. (a) If 10 of the employees are chosen at random to provide free tax assistance to residents of the city, use the binomial approximation to the hypergeometric PDF to find the probability that at least three women were selected.

Solution:
(a) This problem requires the binomial approximation to the hypergeometric, as stated in the problem statement. N=150 employees. k=30 women. n=10. $p = \frac{30}{150} = 0.2$. Find $P(X \geq 3)$

If we solve this using our code,

$$P(X \geq 3) = \sum_{x=3}^{10} b(x;10,0.2) = 0.322200$$

If we solve this using the Tables, then we must rearrange:

$$P(X \geq 3) = P(X \leq 10) - P(X \leq 2)$$

We know that:

$$P(X \leq r) \equiv B(r;n,p) = \sum_{x=0}^{r} b(x;n,p)$$

so

$$P(X \geq 3) = B(10,10,0.2) - B(2,10,0.2)$$

$$P(X \geq 3) = 1 - 0.6778 = 0.3222$$

Problem 7.
A scientist inoculates several mice, one at a time, with a disease germ until he finds 2 that have contracted the disease. If the probability of contracting the disease is 1/6.
(a) If x is a random variable that describes the number of mice which must be inoculated, what PDF describes x?
(b) What is the probability that 8 mice are required?
(c) What is the probability that between 8 and 10 mice, inclusive, are required?

Solution:
(a) This problem requires the negative binomial distribution with k=2, x=8, and p=1/6.

$$b^*(x;k,p) = \binom{x-1}{k-1} \mathrm{p}^{\mathrm{k}} \mathrm{q}^{\mathrm{x-k}} \quad \text{for x} = \mathrm{k}, \mathrm{k}+1, \mathrm{k}+2...$$

(b) What is the probability that 8 mice are required?

$$b^*(8;2,\frac{1}{6}) = \binom{8-1}{2-1}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{8-2} = 0.0651$$

(c) What is the probability that between 8 and 10 mice, inclusive, are required?

$$P(8 \le X \le 10) = \sum_{x=8}^{10} b^*(x;2,\frac{1}{6}) = 0.185279$$

Problem 8.
The acceptance scheme for purchasing lots containing a large number of batteries is to test no more than 75 randomly selected vatteries and to reject a lot if a sing battery fails. Suppoe the probability of a failure is 0.001.

(a) What is the probability that a lot is accepted?
(b) What is the probability that a lot is rejected on the 20^{th} attempt?
(c) What is the probability that it is rejected in 10 or less trials?

Solution:
(a)
This problem requires the binomial PDF where n = 75, $x \ge 1$, and p=0.001.

$$P(x \ge 1) = 1 - P(x < 1) = P(x = 0)$$

$$P(X = x) = b(x;n,p) = \binom{n}{x} p^x q^{n-x}$$

$$P(X = 1) = b(1;75,0.001) = \left(\frac{75!}{0!75!}\right)\left(0.001^0\right)\left(0.999^{75}\right) = (1)(1)\left(0.999^{75}\right) = 0.9277$$

(b) What is probability that lot is rejected on the 20^{th} test?

The probability that a lot is rejected on the 20^{th} test assumes that the first 19 tests passed. So, this calls for the geometric distribution.

$$g(x;p) = \mathrm{pq}^{\mathrm{x-1}} \quad \text{for x} = 1,2,3...$$

$$g(20;0.001) = (0.001)0.999^{19} = 0.000981$$

(c) What is probability that lot is rejected in 10 or less trials?
The probability that a lot fails in 10 or less trials is just the cumulative geometric distribution with $x \le 10$

$$P(X \le 10) \equiv G(10;0.001) = \sum_{i=1}^{10} g(i;0.001) = 0.001(0.999)^{i-1}$$

IN MATLAB, create the file geo.m and geoprob.m that look like this:

```
function prob = geo(x,p)
prob = p*(1-p)^(x-1);

function f = geoprob(a,c,p)
f = 0.0;
for x = a:1:c
   f = f + geo(x,p);
end
```

At the command line, type:
» geoprob(0,10,0.0001)

ans = 0.0100

Problem 9.

On average a certain intersection results in 3 traffic accidents per month. What is the probability that for any given month at this intersection
(a) exactly 5 accidents occur?
(b) less than 3 accidents occur?
(c) at least 2 accidents occur?

Solution:

This problem requires the Poisson distribution with $\lambda = 3\frac{accidents}{month}$, $t = 1\,month$ and x=5 for part (a), x=5 for part (b),and x$\ge$2 for part (c).

$$p(x;\lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!} \quad \text{for } x = 0,1,2............$$

(a)

$$p(5;3) = \frac{e^{-3}(3)^5}{5!} = 0.1008$$

(b) From table A.2, WMM:

$$P(x < 3) = P(x \le 2) = \sum_{i=0}^{2} p(i;3) = 0.4232$$

(c) From table A.2, WMM:

$$P(x \ge 2) = 1 - P(x < 2) = 1 - P(x \le 1) = 1 - \sum_{i=0}^{1} p(i;3) = 0.8009$$