SOLUTIONS: 4.1 Probability Distributions and 4.2 Binomial Distributions

1. The following table contains a probability distribution for a random variable X.

x	1	2	3	4	5
$P(x)$	0.16	0.22	0.28	0.20	0.14

a. Find the expected value (mean) of X.

x	$P(x)$	$xP(x)$
1	0.16	0.16
2	0.22	0.44
3	0.28	0.84
4	0.20	0.80
5	0.14	0.70
	$\Sigma P(x) = 1$	$\Sigma xP(x) = 2.94$

$$\mu = 2.94$$

b. Find the variance and standard deviation of X.

x	$P(x)$	x^2	$x^2P(x)$
1	0.16	1	0.16
2	0.22	4	0.88
3	0.28	9	2.52
4	0.20	16	3.20
5	0.14	25	3.50
	$\Sigma P(x) = 1$		$\Sigma x^2P(x) = 10.26$

$$\text{Variance: } \sigma^2 = \Sigma x^2 P(x) - \mu^2 = 10.26 - (2.94)^2 = 10.26 - 8.64 = 1.62$$

$$\text{Standard Deviation: } \sigma = \sqrt{1.62} = 1.27$$

2. The random variable X represents the number of calls per hour to a call center in the last month. The probability distribution for X is below.

x	0	1	2	3	4	5	6	7
$P(x)$	0.01	0.10	0.26	0.33	0.18	0.06	?	0.03

a. What is the missing value in the table?
Since a probability distribution is given, all of the numbers in the second row should add up to 1. The sum of the numbers that are given is

$$0.01 + 0.10 + 0.26 + 0.33 + 0.18 + 0.06 + 0.03 = 0.97.$$

Therefore, the missing value is $1 - 0.97 = 0.03$.

b. Find the mean of X.

x	$P(x)$	$xP(x)$
0	0.01	0.00
1	0.10	0.10
2	0.26	0.52
3	0.33	0.99
4	0.18	0.72
5	0.06	0.30
6	0.03	0.18
7	0.03	0.21
	$\Sigma P(x) = 1$	$\Sigma xP(x) = 3.02$

$$\mu = 3.02$$

c. Find the variance and standard deviation of X.

x	$P(x)$	x^2	$x^2P(x)$
0	0.01	0	0.00
1	0.10	1	0.10
2	0.26	4	1.04
3	0.33	9	2.97
4	0.18	16	2.88
5	0.06	25	1.50
6	0.03	36	1.08
7	0.03	49	1.47
	$\Sigma P(x) = 1$		$\Sigma x^2P(x) = 11.04$

$$\text{Variance: } \sigma^2 = \Sigma x^2P(x) - \mu^2 = 11.04 - (3.02)^2 = 11.04 - 9.12 = 1.92$$

$$\text{Standard Deviation: } \sigma = \sqrt{1.92} = 1.39$$

d. Interpret your results.

The average number of calls per hour over the last month is represented by the expected value (mean), 3.02. The standard deviation measures the 'average' spread of X, the number of calls per hour, around the mean.

3. In a quality control analysis, the random variable X represents the number of defective products per each batch of 100 products produced.

Defects (x)	0	1	2	3	4	5
Batches	95	113	87	64	13	8

a. Use the frequency distribution above to construct a probability distribution for X. The total number of batches is

$$95 + 113 + 87 + 64 + 13 + 8 = 380.$$

Thus,

x	0	1	2	3	4	5
$P(x)$	$\frac{95}{380} = 0.25$	$\frac{113}{380} \approx 0.30$	$\frac{87}{380} \approx 0.23$	$\frac{64}{380} \approx 0.17$	$\frac{13}{380} \approx 0.03$	$\frac{8}{380} \approx 0.02$

b. Find the mean of this probability distribution.

x	$P(x)$	$xP(x)$
0	0.25	0.00
1	0.30	0.30
2	0.23	0.46
3	0.17	0.51
4	0.03	0.12
5	0.02	0.10
	$\Sigma P(x) = 1$	$\Sigma xP(x) = 1.49$

$$\mu = 1.49$$

c. Find the variance and standard deviation of this probability distribution.

x	$P(x)$	x^2	$x^2P(x)$
0	0.25	0	0.00
1	0.30	1	0.30
2	0.23	4	0.92
3	0.17	9	1.53
4	0.03	16	0.48
5	0.02	25	0.50
	$\Sigma P(x) = 1$		$\Sigma x^2P(x) = 3.73$

Variance: $\sigma^2 = \Sigma x^2P(x) - \mu^2 = 3.73 - (1.49)^2 = 3.73 - 2.12 = 1.51$

Standard Deviation: $\sigma = \sqrt{1.51} = 1.23$

4. A surgery has a success rate of 75%. Suppose that the surgery is performed on three patients.

   a. What is the probability that the surgery is successful on exactly 2 patients?

      The number of successful surgeries, X can be represented by a binomial distribution with $n = 3$ trials, success probability $p = 0.75$ and failure probability $q = 1 - p = 0.25$. Therefore,

$$P(2) =_n C_2 p^2 q^{n-2} =_3 C_2 (0.75)^2 (0.25)^1 \approx 0.422.$$

   b. Let X be the number of successes. What are the possible values of X? $0, 1, 2, 3$

   c. Create a probability distribution for X.

      We must find $P(0), P(1), P(2)$, and $P(3)$. (We already have $P(2)$).

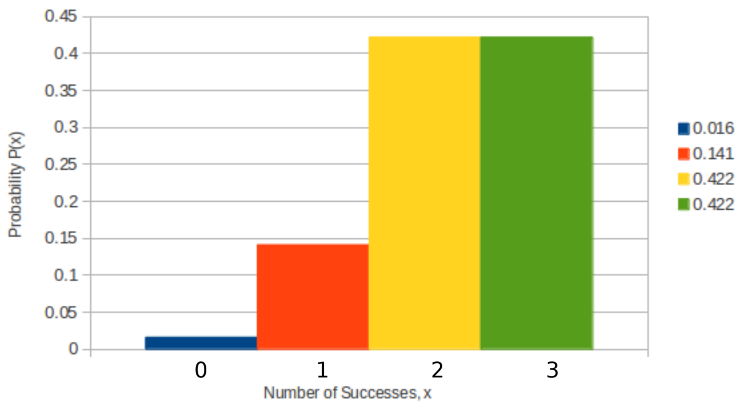
$$P(0) =_n C_0 p^0 q^{n-0} =_3 C_0 (0.75)^0 (0.25)^3 = \frac{1}{64} \approx 0.016$$
$$P(1) =_n C_1 p^1 q^{n-1} =_3 C_1 (0.75)^1 (0.25)^2 = \frac{9}{64} \approx 0.141$$
$$P(2) =_n C_2 p^2 q^{n-2} =_3 C_2 (0.75)^2 (0.25)^1 = \frac{27}{64} \approx 0.422$$
$$P(3) =_n C_3 p^3 q^{n-3} =_3 C_3 (0.75)^3 (0.25)^0 = \frac{27}{64} \approx 0.422$$

      (Note: the sum of the probabilities in decimal form is 1.001 (not exactly 1.000, as it should be). This is due to round-off error. The sum of the probabilities, expressed as fractions, is, indeed 1 because when expressed as fractions, these probabilities are exact.)

      The probability distribution is:

x	0	1	2	3
$P(x)$	$\frac{1}{64}$	$\frac{9}{64}$	$\frac{27}{64}$	$\frac{27}{64}$

d. Graph the probability distribution for X using a histogram.



e. Find the mean of X.

$$\mu = np = 3 \cdot 0.75 = 2.25$$

(This formula for the mean only works for a binomial distribution.)

f. Find the variance and standard deviation of X.

$$\text{Variance: } \sigma^2 = npq = 3 \cdot 0.75 \cdot 0.25 = 0.5625$$

$$\text{Standard Deviation: } \sigma = \sqrt{npq} = \sqrt{0.5625} = 0.75$$

(These formulas for variance and standard deviation only work for a binomial distribution.)

5. New York City typically has rain on about 16% of days in October.

a. What is the probability that it will rain on exactly 5 days in October? 15 days?

The number of rainy days, X can be represented by a binomial distribution with $n = 31$ trials (the number of days in the month of October), success probability $p = 0.16$ (representing a rainy day) and failure probability $q = 1 - p = 0.84$.

Although it is not quite true that the weather (rain or not) is independent from day to day, we shall assume it is quite close to being independent, in order to use the binomial distribution.

Therefore,

$$P(5) =_n C_5 p^5 q^{n-5} =_{31} C_5 (0.16)^5 (0.84)^{26} \approx 0.191.$$

$$P(15) =_n C_{15} p^{15} q^{n-15} =_{31} C_{15} (0.16)^{15} (0.84)^{16} \approx 0.0000213.$$

b. What is the mean number of days with rain in October?

$$\mu = np = 31 \cdot 0.16 = 4.96$$

c. What is the variance and standard deviation of the number of days with rain in October?

$$\text{Variance: } \sigma^2 = npq = 31 \cdot 0.16 \cdot 0.84 = 4.17$$

$$\text{Standard Deviation: } \sigma = \sqrt{npq} = \sqrt{4.17} = 2.04$$

6. From past records, a store finds that 26% of people who enter the store will make a purchase. 18 people enter the store during a one-hour period.

a. What is the probability that exactly 10 customers make a purchase? 18 customers? 3 customers?

The number of customers that make a purchase, X, can be represented by a binomial distribution with $n = 18$ trials (the total number of customers), success probability $p = 0.26$ (representing a customer who makes a purchase) and failure probability $q = 1 - p = 0.74$.

Again, it is not quite true that the customers' decisions to make a purchase are independent, as for example, their conversations among each other or with the sales people may influence other customers around. Nevertheless, we shall assume these decisions are quite close to being independent, in order to use the binomial distribution.

Therefore,

$$P(10) =_n C_{10} p^{10} q^{n-10} =_{18} C_{10} (0.26)^{10} (0.74)^8 \approx 0.00555.$$

$$P(18) =_n C_{18} p^{18} q^{n-18} =_{18} C_{18} (0.26)^0 (0.74)^{18} \approx 0.0000000000295.$$

$$P(3) =_n C_3 p^3 q^{n-3} =_{18} C_3 (0.26)^3 (0.74)^{15} \approx 0.157.$$

b. Find the expected number of customers who make a purchase.

$$\mu = np = 18 \cdot 0.26 = 4.68$$

c. Find the variance and standard deviation of the number of customers who make a purchase.

$$\text{Variance: } \sigma^2 = npq = 18 \cdot 0.26 \cdot 0.74 = 3.46$$

$$\text{Standard Deviation: } \sigma = \sqrt{npq} = \sqrt{3.46} = 1.86$$