Stat 516 (Spring 2012)

Hw 1 (due Feb. 2, Thursday)

**Question 1** Suppose $P(A) = 0.45$, $P(B) = 0.32$ and $P(\bar{A} \cap B) = 0.20$.

1) Find $P(A \cap B)$

[**Sol**] Since $P(B) = P(A \cap B) + P(\bar{A} \cap B)$, $P(A \cap B) = P(B) - P(\bar{A} \cap B) = 0.32 - 0.20 = 0.12$.

2) Are the events $A$ and $B$ independent? Say why or why not

[**Sol**] No : $P(A \cap B) = 0.12$ is not equal to $P(A) \times P(B) = 0.45 \times 0.32 = 0.144$.

3) Find $P(B \mid \bar{A})$ and Find $P(A \mid A \cup B)$

[**Sol**] $P(B \mid \bar{A}) = \frac{P(\bar{A} \cap B)}{P(\bar{A})} = \frac{0.20}{1 - 0.45} = \frac{4}{11}$.

$P(A \mid A \cup B) = \frac{P(A \cap (A \cup B))}{P(A \cup B)} = \frac{P(A)}{P(A)} = 1$.

**Question 2** A fair six-sided die is rolled 6 independent times. Define an event $A_i$ as follows: $A_i$ = side $i$ is observed on the $i$-th roll, $i = 1, \ldots, 6$.

1) What is the probability that none of $A_i$ occurs?

[**Sol**] $P$(none of $A_i$ occurs) = $P(\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3 \cap \bar{A}_4 \cap \bar{A}_5 \cap \bar{A}_6) = P(\bar{A}_1) \times P(\bar{A}_2) \times P(\bar{A}_3) \times P(\bar{A}_4) \times P(\bar{A}_5) \times P(\bar{A}_6) = (\frac{5}{6})^6$.

2) What is the probability that at least one $A_i$ occurs?

[**Sol**] $P$(at least one $A_i$ occurs) = $1 - P$(none of $A_i$ occurs) = $1 - (\frac{5}{6})^6$.

**Question 3** Suppose you have been tested positive for a disease. We are interested in the probability that you actually have the disease. This probability depends on the accuracy and sensitivity of the test, and on the (prior) information on the disease. Let the false negative rate and the false positive rate of this test be $6\%$ and $2\%$, respectively. Suppose the disease is not rare in that the probability that one has the disease is $10\%$.

1) If you actually have the disease, find the probability that you have been tested positive.

[**Sol**] Let $D$ be an event that one actually has the disease, $TP$ be an event that one has been tested positive and $TN$ be an event that one has been tested negative.

Then $P(TP \mid D) = 0.94$, as the false negative rate is $6\%$.

2) Find the probability that you have been tested positive.

[**Sol**] $P(TP) = P(TP \cap (D \cup \bar{D})) = P((TP \cap D) \cup (TP \cap \bar{D})) = P(TP \cap D) + P(TP \cap \bar{D})$. 

1
\[ D = P(TP \mid D)P(D) + P(TP \mid \bar{D})P(\bar{D}) = 0.94 \times 0.2 + 0.02 \times 0.8 = 0.204. \]

3) If you have been tested positive for a disease, find the probability that you actually have the disease.
   \[ \text{[Sol]} \quad P(D \mid TP) = \frac{P(TP \cap D)}{P(TP)} = \frac{P(TP \cap D)P(D)}{P(TP)} = \frac{0.188}{0.204} = 0.9216. \]

**Question 4** Candidate A in a city election believes that 40% of the city’s voters favor him. Suppose the \( n = 25 \) voters from the city show up to vote. Note that we think of \( n = 25 \) voters as a random sample from this city.

1) Suppose \( Y \) is the number of voters who vote him. What is the probability distribution for \( Y \)? (Hint 1. \( Y \) is a discrete random variable with a well-known probability distribution)
   \[ \text{[Sol]} \quad \text{Since } Y \text{ has a binomial distribution with } n = 25 \text{ and } p = P(\text{the city’s votes favor him}) = .4 (i.e., } Y \sim b(25,0.4)) \text{, the probability distribution for } Y \text{ is } p(y) = \binom{25}{y} \cdot 0.4^y \cdot 0.6^{25-y} \text{ where } y = 0,1,\ldots,25. \]

2) What is the exact probability that candidate A will receive 32% of their votes? (Hint 2. The following information will be helpful for the probability calculation: \( P(Y \leq 9) = 0.425, P(Y \leq 8) = 0.274, P(Y \leq 7) = 0.154, P(Y \leq 6) = 0.074 \).
   \[ \text{[Sol]} \quad P(Y = 8) = P(Y \leq 8) - P(Y \leq 7) = 0.274 - 0.154 = 0.12 \]

**Question 5** A group of four components is known to contain one defective. An inspector tests the components one at a time until one defective is located. Once she locates one defective, she stops testing. Let \( Y \) denote the number of the test on which one defective is found.

1) Find the probability distribution for \( Y \) (represent it by a table).
   \[ \text{[Sol]} \quad \text{The possible values of } Y = 1,2,3,4. \text{ Then } P(Y = 1) = \frac{1}{4}, P(Y = 2) = \frac{3}{4} \times \frac{1}{3} = \frac{1}{4}, \]
   \[ P(Y = 3) = \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{4}, P(Y = 4) = \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} \times \frac{1}{1} = \frac{1}{4}. \]

<table>
<thead>
<tr>
<th>( Y )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(y) )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{4} )</td>
</tr>
</tbody>
</table>

2) Find the probability that the number of the test on which one defective is found is more than 2.
   \[ \text{[Sol]} \quad P(Y > 2) = P(Y = 3) + P(Y = 4) = \frac{1}{2}. \]

3) Find \( E(Y) \) and \( V(Y) \).
Question 6 Customers arrive at a checkout counter in a department store according to a Poisson distribution at an average of seven per hour. Let \( Y \) be the number of customers arriving per hour.

1) What is the expected value of \( Y \) and the variance of \( Y \)?

[Sol] Since \( Y \) has a Poisson distribution at an average of seven per hour, \( E(Y) = \lambda = 7 \). Then \( Y \sim \text{Poi}(7) \).

\[
E(Y) = 7 \text{ and } V(Y) = 7.
\]

2) Find the probability that at least two customers arrive.

[Sol] \[
P(Y \geq 2) = 1 - P(Y < 2) = 1 - P(Y \leq 1) = 1 - e^{-7}7^0/0! - e^{-7}7^1/1!.
\]

3) Find the probability that exactly two customers arrive in the 2-hour period of time: between 1:30 p.m. and 3:30 p.m.

[Sol] Let \( Y \) be the number of customers that arrive in a given two-hour period of time. Then \( Y \) has a Poisson distribution with \( \lambda = 2 \times 7 \) and \( P(Y = 2) = \frac{14^2}{2!}e^{-14} \).

Question 7 Suppose we have the discrete random variable \( Y \) with the following probability distribution:

\[
P(Y = y) = p(y) = \begin{cases} 
1/8 & y = 0, \\
3/8 & y = 1, \\
c & y = 2, \\
1/4 & y = 3, \\
0 & \text{otherwise}.
\end{cases}
\]

1) Find the constant \( c \).

[Sol] Since \( Y \) is a discrete random variable, \( \sum_{y} p(y) = 1 \). Then \( 1 = 1/8 + 3/8 + c + 1/4 \) and \( c = 1/4 \).

2) Find the probability that \( Y \) is less than 3.
\[ \text{[Sol]} \ P(Y < 3) = P(Y = 0) + P(Y = 1) + P(Y = 2) = 1/8 + 3/8 + 1/4 = 3/4. \]

3) Find \( E(Y) \) and \( V(Y) \).
\[ \text{[Sol]} \ E(Y) = \sum_{y=0}^{3} y p(y) = 0 \cdot (1/8) + 1 \cdot (3/8) + 2 \cdot (1/4) + 3 \cdot (1/4) = 13/8, \]
\[ V(Y) = E(Y^2) - E(Y)^2 = \sum_{y=0}^{3} y^2 p(y) - (13/8)^2 = 29/8 - 169/64 = 63/64. \]

4) Find \( E(2Y + 6) \) and \( V(-2Y + 6) \).
\[ \text{[Sol]} \ E(2Y + 6) = 2E(Y) + 6 = 2(13/8) + 6 = 37/4, \]
\[ V(-2Y + 6) = (-2)^2 V(Y) = 63/16. \]

Question 8 Let the cumulative distribution function (C.D.F.) of a continuous random variable \( Y \) be
\[ F(y) = \begin{cases} 0 & y < 0, \\ \frac{y}{8} & 0 \leq y < 2, \\ \frac{y^2}{16} & 2 \leq y < 4, \\ 1 & y \geq 4. \end{cases} \]

1) Find the probability density function (p.d.f) of \( Y \), \( f(y) \).
\[ \text{[Sol]} \ f(y) \equiv \frac{\partial F(y)}{\partial y} = \begin{cases} \frac{1}{8} & 0 \leq y < 2, \\ \frac{y}{8} & 2 \leq y < 4, \\ 0 & \text{elsewhere}. \end{cases} \]

2) Find the expected value of \( Y \), \( E(Y) \).
\[ \text{[Sol]} \ E(Y) = \int_{-\infty}^{\infty} y f(y) dy = \int_{0}^{2} y(1/8)dy + \int_{2}^{4} y(y/8)dy = 31/12. \]

3) Find \( P(1 \leq Y \leq 3) \).
\[ \text{[Sol]} \ P(1 \leq Y \leq 3) = \int_{1}^{2} (1/8)dy + \int_{2}^{3} (y/8)dy = 7/16 \text{ or } P(1 \leq Y \leq 3) = F(3) - F(1) = \frac{3^2}{16} - \frac{1}{8} = 7/16. \]

4) Find \( P(Y \geq 1 \mid Y \leq 3) \).
\[ \text{[Sol]} \ P(Y \geq 1 \mid Y \leq 3) = \frac{P(1 \leq Y \leq 3)}{P(Y \leq 3)} = \frac{F(3) - F(1)}{F(3)} = \frac{7/16}{9/16} = 7/9. \]

Question 9 The grade point averages of a large population of college students are normally distributed with mean 2.4 and standard deviation 0.8.
1) What fraction of the students will possess a grade point average in excess of 3.0?
   [Sol] Since \( Y \sim N(2.4, 0.8^2) \), \( P(Y > 3.0) = P(Z > \frac{3.0 - 2.4}{0.8}) = P(Z > 0.75) = 0.2266 \).

2) Suppose that three students are randomly selected from the student body. What is the probability that all three will possess a grade point average in excess of 3.0?
   [Sol] \( P( \text{all three will possess a grade point average in excess of 3.0} ) = [P(Y > 3.0)]^3 = 0.2266^3 = 0.6798 \).

**Question 10** Let \( Y_1 \) and \( Y_2 \) be continuous random variables with the following probability density function:

\[
f(y_1, y_2) = \begin{cases} 
6(1 - y_2) & 0 \leq y_1 \leq y_2 \leq 1, \\
0 & \text{elsewhere.}
\end{cases}
\]

1) Find the marginal density function of \( Y_1 \) and \( Y_2 \) (Note: Specify the range of \( Y_1 \) and \( Y_2 \) completely).
   [Sol] \( f_1(y_1) = \int_{-\infty}^{y_2} f(y_1, y_2) dy_2 = \int_{y_1}^{1} 6(1 - y_2) dy_2 = 3(1 - y_1)^2 \) where \( 0 \leq y_1 \leq 1 \),

\[
f_2(y_2) = \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 = \int_{y_1}^{y_2} 6(1 - y_2) dy_1 = 6y_2(1 - y_2) \) where \( 0 \leq y_2 \leq 1 \).

2) Are \( Y_1 \) and \( Y_2 \) independent? Why?
   [Sol] No. i) the range of \( Y_1 \) and \( Y_2 \) are related to each other, and ii) \( f(y_1, y_2) \neq f_1(y_1) \times f_2(y_2) \).

3) Find the conditional density function for \( Y_2 \) given \( Y_1 = y_1 \) (Note: Specify the range of \( Y_2 \) completely).
   [Sol] \( f(y_2 \mid y_1) = \frac{f(y_1, y_2)}{f_1(y_1)} = \frac{2(1-y_2)}{(1-y_1)^2} \) where \( y_1 \leq Y_2 \leq 1 \).

**Question 11** The number of defects per yard in a certain fabric, \( Y \), was known to have a Poisson distribution with parameter \( \lambda \). The parameter \( \lambda \) was assumed to be a random variable with a density function given by

\[
f(\lambda) = \begin{cases} 
k \, e^{-\lambda/2} & \lambda \geq 0, \\
0 & \text{elsewhere.}
\end{cases}
\]

1) Find the value of \( k \) that makes \( f(\lambda) \) a probability density function (Hint. compare \( f(\lambda) \) with well-known probability distributions).
   [Sol] If you compare \( f(\lambda) \) with an exponential distribution, we can see that \( k = 1/2 \).
In other words, \( \lambda \sim \text{exponential}(2) \). Another way is to use the definition of a probability
density function, which is \(1 = \int_{-\infty}^{\infty} f(\lambda)d\lambda = \int_{0}^{\infty} ke^{-\lambda/2} = k \int_{0}^{\infty} e^{-\lambda/2} = 2k\).

2) Find \(E(Y \mid \lambda)\) and \(V(Y \mid \lambda)\).
[\text{Sol}] Given \(\lambda, Y \sim \text{poisson}(\lambda)\). Then \(E(Y \mid \lambda) = \lambda\) and \(V(Y \mid \lambda) = \lambda\).

3) Find \(E(Y)\) and \(V(Y)\).
[\text{Sol}] \(E(Y) = E(E(Y \mid \lambda)) = E(\lambda) = 2\) (because \(\lambda \sim \text{exponential}(2)\)), \(V(Y) = V(E(Y \mid \lambda)) + E(V(Y \mid \lambda)) = V(\lambda) + E(\lambda) = 4 + 2 = 6\) (because \(\lambda \sim \text{exponential}(2)\))