Math 442 – Homework Set 1

Spring 2009

p.157[4,8], p.175[3,18], p.177[8,20], p.221[8,15], p.244[14], p.251[9], p.267[9], p.280[15], p.302[6,8], p.319[15]

1. (p.157 #4) Suppose that a point (X_1, X_2, X_3) is chosen at random, that is, in accordance with a uniform probability density function, from the following set S:

$$S = \{(x_1, x_2, x_3) : 0 \leq x_i \leq 1 \text{ for } i = 1, 2, 3\}.$$

Determine:

(a) $P\left[\left(X_1 - \frac{1}{2}\right)^2 + \left(X_2 - \frac{1}{2}\right)^2 + \left(X_3 - \frac{1}{2}\right)^2 \leq \frac{1}{4}\right]$

(b) $P({X_1}^2 + {X_2}^2 + {X_3}^2 \leq 1)$.

Solution. (a) Let $A = \{(x_1, x_2, x_3) : (x_1 - \frac{1}{2})^2 + (x_2 - \frac{1}{2})^2 + (x_3 - \frac{1}{2})^2 \leq \frac{1}{4}\}$. The set S forms a unit cube in $\mathbb{R}^3$, the set A forms a sphere centered at $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ with radius $r = 1/2$. Since the points are chosen from S at random and in accordance with a uniform probability density function, the probability that a point in A is selected is equal to the volume of the sphere formed by A contained in S. In this case the entire sphere is contained by S. Therefore,

$$P\left[\left(X_1 - \tfrac{1}{2}\right)^2 + \left(X_2 - \tfrac{1}{2}\right)^2 + \left(X_3 - \tfrac{1}{2}\right)^2 \leq \tfrac{1}{4}\right] = \frac{4}{3}\pi r^3 = \frac{4\pi}{3}\left(\frac{1}{2}\right)^3 = \frac{\pi}{6} \approx 0.5236.$$

(b) Let $A = \{(x_1, x_2, x_3) : {x_1}^2 + {x_2}^2 + {x_3}^2 \leq 1\}$. Once again, the set S forms a unit cube in $\mathbb{R}^3$, the set A forms a sphere centered at $(0, 0, 0)$ with radius 1. Since the points are chosen from S at random and in accordance with a uniform probability density function, the probability that a point in A is selected is equal to the volume of the sphere formed by A contained in S. In this case, only $1/8$ of the sphere is contained by S (in other words, the portion of the sphere contained in the first quadrant of $\mathbb{R}^3$). Therefore,

$$P({X_1}^2 + {X_2}^2 + {X_3}^2 \leq 1) = \frac{1}{8}\left(\frac{4\pi}{3}\right)1^3 = \frac{\pi}{6} \approx 0.5236.$$

□

2. (p.157 #8) Suppose that the probability density function of a random variable X is as follows:

$$f_X(x) = \begin{cases} \frac{1}{n!}x^n e^{-x} & \text{for } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose also that for any given value $X = x$ $(x > 0)$, the n random variables $Y_1, \ldots, Y_n$ are independent and identically distributed and the conditional probability density function of each of them is as follows:

$$f_{Y|X}(y|x) = \begin{cases} \frac{1}{x} & \text{for } 0 < y < x, \\ 0 & \text{otherwise.} \end{cases}$$

Determine:

(a) the marginal joint probability density function of $Y_1, \ldots, Y_n$ and

(b) the conditional probability density function of X for any given values of $Y_1, \ldots, Y_n$.

Proof. (a)

$$
\begin{aligned}
f_{\mathbf{Y}}(\mathbf{y}) &= \int_{-\infty}^{\infty} f_{\mathbf{Y}|X}(\mathbf{y}|x) f_X(x) dx = \int_{-\infty}^{\infty} \prod_{i=1}^{n} \left(f_{Y_i|X}(y_i|x) \right) f_X(x)\ dx \\
&= \int_{y_{\max}}^{\infty} \left(\frac{1}{x}\right)^n \frac{1}{n!} x^n e^{-x}\ dx \qquad \text{where } y_{\max} = \max\{y_1, y_2, \ldots, y_n\} \\
&= \int_{y_{\max}}^{\infty} \frac{1}{n!} e^{-x}\ dx \quad = -\frac{1}{n!} e^{-x} \Big|_{y_{\max}}^{\infty} \quad = \frac{e^{-y_{\max}}}{n!}.
\end{aligned}
$$

That is,

$$
f_{\mathbf{Y}}(\mathbf{y}) = \begin{cases} \dfrac{e^{-y_{\max}}}{n!} & \text{if } y_i > 0 \text{ for } i = 1, 2, \ldots, n, \\ 0 & \text{otherwise.} \end{cases}
$$

(b) For $y_1, y_2, \ldots, y_n > 0$,

$$
f_{X|\mathbf{Y}}(x|\mathbf{y}) = \frac{f_{\mathbf{Y}|X}(\mathbf{y}|x) f_X(x)}{f_{\mathbf{Y}(\mathbf{y})}} \quad = \frac{\left[\prod_{i=1}^{n} f_{Y_i|X}(y_i|x)\right] f_X(x)}{f_{\mathbf{Y}}(\mathbf{y})} \quad = \frac{\left(\frac{1}{x}\right)^n \frac{1}{n!} x^n e^{-x}}{\frac{e^{-y_{\max}}}{n!}} \quad = e^{-x+y_{\max}}.
$$

That is,

$$
f_{X|\mathbf{Y}}(x|\mathbf{y}) = \begin{cases} e^{-x+y_{\max}} & \text{if } x > y_{\max} \text{ and } y_1 > 0, y_2 > 0, \ldots, y_n > 0, \\ 0 & \text{otherwise.} \end{cases}
$$

□

3. (p.175 #3) Suppose that three random variables X_1, X_2, and X_3 have a continuous joint distribution for which the joint probability density function is as follows:

$$
f_{X_1,X_2,X_3}(x_1, x_2, x_3) = \begin{cases} 8x_1x_2x_3 & \text{for } 0 < x_i < 1 \ (i = 1, 2, 3), \\ 0 & \text{otherwise.} \end{cases}
$$

Suppose also that $Y_1 = X_1$, $Y_2 = X_1X_2$, and $Y_3 = X_1X_2X_3$. Find the joint probability density function of Y_1, Y_2 and Y_3.

Solution. Inverting the given equations we obtain new equations of the following form:

$$
x_1 = y_1, \qquad x_2 = \frac{y_2}{y_1}, \qquad x_3 = \frac{y_3}{y_2}.
$$

This implies $0 < y_1 < 1,\ 0 < y_2 < y_1$ and $0 < y_3 < y_2$. That is

$$
\mathcal{B} = \{(y_1, y_2, y_3) : 0 < y_3 < y_2 < y_1 < 1\}.
$$

The Jacobian of transformation is

$$
J = \begin{vmatrix} 1 & 0 & 0 \\ -y_2/{y_1}^2 & 1/y_1 & 0 \\ 0 & -y_3/{y_2}^2 & 1/y_2 \end{vmatrix} = \frac{1}{y_1y_2},
$$

and so,

$$
f_{Y_1,Y_2,Y_3}(y_1, y_2, y_3) = \begin{cases} \dfrac{8y_3}{y_1y_2} & \text{for } 0 < y_3 < y_2 < y_1 < 1, \\ 0 & \text{otherwise.} \end{cases}
$$

□

4. (p.176 #18) Let the conditional probability density function of X given Y be $f_{X|Y}(x|y) = 3x^2/y^3$ for $0 < x < y$ and 0 otherwise. Let the marginal probability density function of Y be $f_Y(y)$, where $f_Y(y) = 0$ for $y \le 0$, but is otherwise unspecified. Let $Z = X/Y$. Prove that Z and Y are independent and find the marginal probability density function of Z.

   Solution. Let us transform the conditional distribution of X given $Y = y$ into the conditional distribution of Z given $Y = y$, since conditioning on $Y = y$ allows us to treat Y as the constant y. Because $X = ZY$, the inverse transformation is $x = zy$. The Jacobian dx/dz is y and the conditional probability density function of Z given $Y = y$ is

$$f_{Z|Y}(z|y) = yf_{X|Y}(zy|y) = \begin{cases} 3z^2 & \text{for } 0 < z < 1, \\ 0 & \text{otherwise.} \end{cases}$$

   Since $f_{Z|Y}(z|y)$ does not depend on y, Z is independent of Y and the marginal probability density function of Z is $f_Z(z) = f_{Z|Y}(z|y)$. □

5. (p.177 #8) Suppose that an electronic system comprises four components, and let X_j denote the time until component j fails to operate ($j = 1, 2, 3, 4$). Suppose that X_1, X_2, X_3, and X_4 are independent and identically distributed random variables, each of which has a continuous distribution with distribution function $F_X(x)$. Suppose that the system will operate as long as both component 1 and at least one of the other three components operate. Determine the distribution function of the time until the system fails to operate.

   Solution. Let the random variable Y denote the time until the *system* fails to operate. The probability that component 1 has failed at (or before) a specified time y is given by $F_{X_1}(y) = P(X_1 \le y)$; that is, the time of failure X_1 has occurred before time y. As X_2, X_3, and X_4 are independent and identically distributed random variables the probability that all three components have failed at (or before) a specified time y is given by the product of their distribution functions. That is, $P(X_2 \le y, X_3 \le y, X_4 \le y) = F_{X_2}(y)F_{X_3}(y)F_{X_4}(y) = (F_{X_1}(y))^3$. The event of interest $\{Y \le y\}$ is equivalent to $\{X_1 \le y\} \cup \{X_2 \le y, X_3 \le y, X_4 \le y\}$. Then using the identity $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,

$$\begin{aligned} F_Y(y) &= P(Y \le y) = P(X_1 \le y) + P(X_2 \le y, X_3 \le y, X_4 \le y) - P(X_1 \le y, X_2 \le y, X_3 \le y, X_4 \le y) \\ &= F_{X_1}(y) + \prod_{i=2}^{4} F_{X_i}(y) - \prod_{i=1}^{4} F_{X_i}(y) \\ &= F_{X_1}(y) + (F_{X_1}(y))^3 - (F_{X_1}(y))^4. \end{aligned}$$

   □

6. (p.178 #20) Suppose that the random variables X, Y, and Z have the following joint probability density function:

$$f_{X,Y,Z}(x, y, z) = \begin{cases} 2 & \text{for } 0 < x < y < 1 \text{ and } 0 < z < 1, \\ 0 & \text{otherwise.} \end{cases}$$

   Evaluate $P(3X > Y|1 < 4Z < 2)$.

   Solution. We note that $(X, Y) \perp Z$ (X and Y are independent of Z). Thus $P(3X > Y|1 < 4Z < 2) = P(3X > Y)$. The region $3x > y$ contained in the support set of (X, Y) is depicted in the following graph (Figure 1) Then integrate over A. Note that the integral must be split at $x = 1/3$,

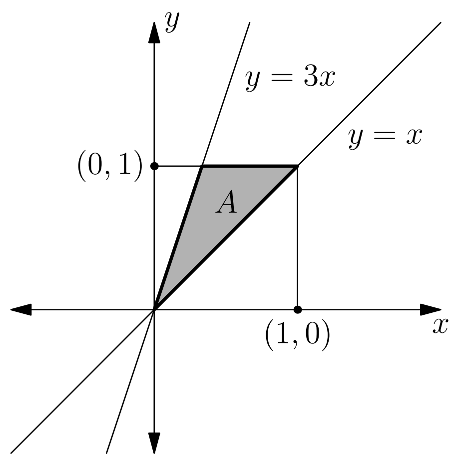

Figure 1: $A = \{(x, y) : 0 < x < y < 1 \text{ and } y < 3x\}$

since $x < y < 3x$ for $x \leq 1/3$ and $x < y < 1$ for $x > 1/3$.

$$\begin{aligned}
P(3X > Y) &= \int_0^{1/3} \int_x^{3x} 2\ dydx + \int_{1/3}^1 \int_x^1 2\ dydx \\
&= \int_0^{1/3} 2y|_x^{3x}\ dx + \int_{1/3}^1 2y|_x^1\ dx \quad = \int_0^{1/3} 4x\ dx + \int_{1/3}^1 2 - 2x\ dx \\
&= 2x^2\Big|_0^{1/3} + \left[2x - x^2\right]_{1/3}^1 = \frac{2}{9} + \frac{4}{9} = \frac{2}{3}.
\end{aligned}$$

□

7. (p.221 #8) Suppose that $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ are random variables such that $\text{Cov}(X_i, Y_j)$ exists for $i = 1, \ldots, m$ and $j = 1, \ldots, n$; and suppose that $a_1, \ldots, a_m$ and $b_1, \ldots, b_n$ are constants. Show that

$$\text{Cov}\left(\sum_{i=1}^m a_i X_i, \sum_{j=1}^n b_j Y_j\right) = \sum_{i=1}^m \sum_{j=1}^n a_i b_j\ \text{Cov}(X_i, Y_j).$$

Solution.

$$\begin{aligned}
\text{Cov}\left(\sum_{i=1}^m a_i X_i, \sum_{j=1}^n b_j Y_j\right) &= E\left(\sum_{i=1}^m a_i X_i\ \sum_{j=1}^n b_j Y_j\right) - E\left(\sum_{i=1}^m a_i X_i\right) E\left(\sum_{j=1}^n b_j Y_j\right) \\
&= E\left(\sum_{i=1}^m \sum_{j=1}^n a_i b_j X_i Y_j\right) - \sum_{i=1}^m a_i E\left(X_i\right) \sum_{j=1}^n b_j E\left(Y_j\right) \\
&= \sum_{i=1}^m \sum_{j=1}^n a_i b_j\ E\left(X_i Y_j\right) - \sum_{i=1}^m \sum_{j=1}^n a_i b_j\ E\left(X_i\right) E\left(Y_j\right) \\
&= \sum_{i=1}^m \sum_{j=1}^n a_i b_j\ \left[E\left(X_i Y_j\right) - E\left(X_i\right) E\left(Y_j\right)\right] \\
&= \sum_{i=1}^m \sum_{j=1}^n a_i b_j\ \text{Cov}(X_i, Y_j).
\end{aligned}$$

□

8. (p.221 #15) Suppose that $X_1, \ldots, X_n$ are random variables such that the variance of each variable is 1 and the correlation between each pair of different variables is 1/4. Determine $\mathrm{Var}(X_1 + \cdots + X_n)$.

Solution. Recall that

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum\sum_{i<j} \mathrm{Cov}(X_i, X_j).$$

Applying the above identity to the current problem, by noting that the correlation

$$\rho(X_i, X_j) = \frac{\mathrm{Cov}(X_i, X_j)}{\sqrt{\mathrm{Var}(X_i)\mathrm{Var}(X_j)}} = \frac{1}{4} \quad \Longrightarrow \mathrm{Cov}(X_i, X_j) = \frac{1}{4},$$

we have

$$\mathrm{Var}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} 1 + 2\sum\sum_{i<j} \frac{1}{4} = n + \binom{n}{2}\left(\frac{1}{2}\right) = \frac{n^2 + 3n}{4}.$$

Note: the number of terms in the double sum is n choose 2, as it is the number of distinct unordered pairs which can be formed from n elements. □

9. (p.244 #14) Suppose that $X_0, X_1, \ldots, X_n$ are independent random variables, each having the same variance σ^2. Let $Y_j = X_j - X_{j-1}$ for $j = 1, \ldots, n$, and let $\overline{Y}_n = \frac{1}{n}\sum_{j=1}^{n} Y_j$. Determine the value of $\mathrm{Var}\left(\overline{Y}_n\right)$.

Solution. First expand the sum

$$\overline{Y}_n = \frac{1}{n}\sum_{j=1}^{n} Y_j = \frac{1}{n}\Big(X_1 - X_0 + X_2 - X_1 + \cdots + X_n - X_{n-1}\Big) = \frac{1}{n}(X_n - X_0),$$

then

$$\mathrm{Var}\left(\overline{Y}_n\right) = \mathrm{Var}\left(\frac{X_n}{n} - \frac{X_0}{n}\right) = \frac{1}{n^2}\mathrm{Var}(X_n) + \frac{1}{n^2}\mathrm{Var}(X_0) = \frac{2\sigma^2}{n^2}.$$

□

10. (p.251 #9) Suppose that the random variables $X_1, \ldots, X_n$ form n Bernoulli trials with parameter p. Determine the conditional probability that $X_1 = 1$, given that $\sum_{i=1}^{n} X_i = k$ for any fixed but arbitrary $k = 1, \ldots, n$.

Solution. Given that there have been k successes in n trials, the probability that $X_i = 1$ (i^{th} trial is a success) is k/n. Alternatively, we can use the mathematical definition of conditional probability.

First note that $\sum_{i=1}^{n} X_i \sim \text{Binomial}(n, p)$. Then

$$\begin{aligned}
P(X_1 = 1 | \sum_{i=1}^{n} X_i = k) &= \frac{P(X_1 = 1, \sum_{i=1}^{n} X_i = k)}{P(\sum_{i=1}^{n} X_i = k)} \\
&= \frac{P(X_1 = 1, \sum_{i=2}^{n} X_i = k-1)}{P(\sum_{i=1}^{n} X_i = k)} \\
&= \frac{P(X_1 = 1)P(\sum_{i=2}^{n} X_i = k-1)}{P(\sum_{i=1}^{n} X_i = k)} \\
&= \frac{p\binom{n-1}{k-1}p^{k-1}(1-p)^{n-1-(k-1)}}{\binom{n}{k}p^k(1-p)^{n-k}} \\
&= \frac{p\binom{n-1}{k-1}p^{k-1}(1-p)^{n-k}}{\frac{n}{k}\binom{n-1}{k-1}p^k(1-p)^{n-k}} \\
&= \frac{k}{n}.
\end{aligned}$$

□

11. (p.267 #9) Suppose that an electronic system contains n components that function independently of each other, and suppose that these components are connected in series, as defined in Exercise 5 of Section 3.7 (page 157). Suppose also that each component will function properly for a certain number of periods and then will fail. Finally, suppose that for $i = 1, \ldots, n$, the number of periods for which component i will function properly is a discrete random variable having a geometric distribution with parameter p_i. Determine the distribution of the number of periods for which the system will function properly.

    Solution. Connected in series, if any component fails then the system will fail; that is, the system will function properly only when all components are functioning. In this case the distribution parameter p_i represents the probability of a component failure. As the components function independently the probability that all components function is $\prod_{i=1}^{n}(1 - p_i)$, hence the probability of a system failure is $1 - \prod_{i=1}^{n}(1 - p_i)$. Like the components, the system will only fail once; that is, the distribution of the number of periods for which the system will function properly is also geometric but with parameter $p = 1 - \prod_{i=1}^{n}(1 - p_i)$. □

12. (p.281 #15) Suppose that 10 percent of the people in a certain population have the eye disease glaucoma. For persons who have glaucoma, measurements of eye pressure X will be normally distributed with a mean of 25 and a variance of 1. For persons who do not have glaucoma, the pressure X will be normally distributed with a mean of 20 and a variance of 1. Suppose that a person is selected at random from the population and her eye pressure X is measured.

    (a) Determine the conditional probability that the person has glaucoma given that $X = x$.

    (b) For what values of x is the conditional probability in part (a) greater than 1/2?

Solution. (a) Let G be the event that a person has glaucoma, since G and G^c form a partition of the sample space we are able to apply Bayes' theorem.

$$\begin{aligned} P(G|X=x) &= \frac{P(G)f_X(x|G)}{P(G)f_X(x|G)+P(G^c)f_X(x|G^c)} \\ &= \frac{\frac{1}{10}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}(x-25)^2\right]}{\frac{1}{10}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}(x-25)^2\right]+\frac{9}{10}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}(x-20)^2\right]} \\ &= \frac{\exp\left[-\frac{1}{2}(x-25)^2\right]}{\exp\left[-\frac{1}{2}(x-25)^2\right]+9\exp\left[-\frac{1}{2}(x-20)^2\right]}. \end{aligned}$$

(b) Solve for x so that $P(G|X=x) > 1/2$.

$$\begin{aligned} & P(G|X=x) > \frac{1}{2} \\ \iff\quad & \exp\left[-\frac{1}{2}(x-25)^2\right] > 9\exp\left[-\frac{1}{2}(x-20)^2\right] \\ \iff\quad & -\frac{1}{2}(x-25)^2 > \ln(9) - \frac{1}{2}(x-20)^2 \\ \iff\quad & x^2-50x+625 < -2\ln(9)+x^2-40x+400 \\ \iff\quad & 10x > 225+2\ln(9) \\ \iff\quad & x > 22.5+\frac{1}{5}\ln(9). \end{aligned}$$

The conditional probability in part (a) is greater than $1/2$ for $x > 22.5+\frac{1}{5}\ln(9)$.

□

13. (p.302 #6) Suppose that $X_1,\ldots,X_n$ form a random sample of size n from an exponential distribution with parameter β. Determine the distribution of the sample mean $\overline{X}_n$.

Solution. Let $M_{X_i}(t/n)$ denote the moment generating function of X_i/n for $i=1,\ldots,n$, and let $M_{\overline{X}_n}(t)$ denote the moment generating function of $\overline{X}_n$. Since the variables $X_1,\ldots,X_n$ are independent and identically distributed (random sample), then

$$M_{\overline{X}_n}(t) = \prod_{i=1}^{n} M_{X_i}(t/n) \;= \prod_{i=1}^{n}\frac{\beta}{\beta-t/n} \;= \left(\frac{\beta}{\beta-t/n}\right)^n.$$

$M_{\overline{X}_n}(t)$ is the moment generating function of a gamma distribution. Hence, the sample mean $\overline{X}_n$ has a gamma distribution with parameters $\alpha=n,\ \beta=\beta$. □

14. (p.302 #8) Suppose that the random variable $X_1,\ldots,X_k$ are independent and X_i has an exponential distribution with parameter β_i $(i=1,\ldots,k)$. Let $Y=\min\{X_1,\ldots,X_k\}$. Show that Y has an exponential distribution with parameter $\beta_1+\cdots+\beta_k$.

Proof. Recall that if X has an exponential distribution with parameter β, then for every number $t>0$,

$$1-F_X(t) \;= P(X\geq t) = \int_t^\infty \beta e^{-\beta x}\,dx \;= e^{-\beta t}.$$

Because $Y = \min\{X_1, X_2, \ldots, X_k\}$ for every number $t > 0$,

$$
\begin{aligned}
P(Y > t) &= P(X_1 > t, X_2 > t, \ldots, X_k > t) \ = P(X_1 > t)P(X_2 > t) \cdots P(X_k > t) \\
&= e^{-\beta_1 t} e^{-\beta_2 t} \cdots e^{-\beta_k t} \ = e^{-t(\beta_1 + \beta_2 + \cdots + \beta_k)}.
\end{aligned}
$$

Therefore, Y has an exponential distribution with parameter $\beta = \beta_1 + \beta_2 + \cdots + \beta_k$. $\square$