E1 222 Stochastic Models and Applications
Hints for Problem Sheet 3.3

1. Consider a communication system. Let Y denote the bit sent by transmiter. (Y is a binary random variable). The receiver makes a measurement, X, and based on its value decides what is sent. The decision at the receiver can be represented by a function $h : \Re \to \{0,\ 1\}$. For any specific h, let $R_0(h)$ represent the set of all $x \in \Re$ for which $h(x) = 0$ and let $R_1(h)$ represent the set of $x \in \Re$ for which $h(x) = 1$. An error occurs if a wrong decision is made. Argue that the event of error is: $[h(X) = 0, Y = 1] \cup [h(X) = 1, Y = 0]$. Show that probability of error for a decision rule h is

$$\int_{R_0(h)} p_1 f_{X|Y}(x|1)dx \ + \ \int_{R_1(h)} p_0 f_{X|Y}(x|0)dx$$

where $p_i = P[Y = i]$. Now consider a h given by

$$h(x) = 1 \ \text{ if } \ f_{Y|X}(1|x) \geq f_{Y|X}(0|x)$$

(Otherwise $h(x) = 0$). Show that this h would achieve minimum probability of error.

Hint: An error occurs when $h(X)$, which is the decision of the receiver, differs from Y, which is what was sent. Since both are binary, the event representing error is given by $[h(X) = 0, Y = 1] \cup [h(X) = 1, Y = 0]$. Hence we have

$$\begin{aligned}
P[\text{error}] &= P[h(X) = 0, Y = 1] + P[h(X) = 1, Y = 0] \\
&= P[h(X) = 0|Y = 1]P[Y = 1] + P[h(X) = 1|Y = 0]P[Y = 0] \\
&= P[X \in R_0(h)|Y = 1]P[Y = 1] + P[X \in R_1(h)|Y = 0]P[Y = 0] \\
&= \int_{R_0(h)} p_1 \ f_{X|Y}(x|1) \ dx \ + \ \int_{R_1(h)} p_0 \ f_{X|Y}(x|0) \ dx
\end{aligned}$$

because, by definition of $R_0(h)$, $[h(X) = 0]$ is same as $[X \in R_0(h)]$ and similarly for the other term. Also, note that $R_0(h) \cap R_1(h) = \phi$ and $R_0(h) \cup R_1(h) = \Re$. So, every $x \in \Re$ is accounted for in one of the two integrals above.

For the second part of the problem. By Bayes rule, $f_{Y|X}(1|x) \geq f_{Y|X}(0|x)$ is same as $p_1 \ f_{X|Y}(x|1) \geq p_0 \ f_{X|Y}(x|0)$. Hence for the given

h here, in $R_0(h)$ we have $p_1\ f_{X|Y}(x|1) \le p_0\ f_{X|Y}(x|0)$ and the other way in $R_1(h)$. Hence, the probability of error for this h can be written as

$$P[\text{error}] = \int_{\Re} \min(p_1\ f_{X|Y}(x|1), p_0\ f_{X|Y}(x|0))\ dx$$

Comparing this with the expression for probability of error for any other function, we can see that the given h is optimal.

2. Let X, Y have a joint distribution that is uniform over the quadrilateral with vertices at $(-1, 0), (1, 0), (0, -1)$ and $(0, 1)$. Find $P[X > Y]$. Are X, Y independent? (Hint: Can you decide on independence here without calculating the marginal densities?)

Hint: Here, drawing a figure would help. Since X, Y are uniform over some region, probability of any subset would be proportional to the area of that subset. That is the easy way to find $P[X > Y]$. Note that here the joint density of X, Y is uniform over a rhombus and the line $x = y$ divides the rhombus into two equal parts. Of course, you can also get the answer by integrating the joint density.

Similarly, to test independence, you need not find marginal densities and ask whether the joint is a product of marginals. Looking at the figure, you can see, for example, that the range of values that X can take depends on value of Y. Hence you intuitively know they are not independent. In such situations a easy way to show that they are not independent is to identify A, B such that $P[X \in A|Y \in B] = 0$ but $P[X \in A] \ne 0$. For example, you can look at $P[X < -0.75|Y > 0.5]$.

3. Let X, Y be *iid* uniform over $(0,\ 1)$. Let $Z = \max(X, Y)$ and $W = \min(X, Y)$. Find the density of $Z - W$.

Hint: We found the expression for f_{ZW} in class. You also know how to find density of $X - Y$ if you know joint density of X, Y. So, it is just a matter of combining the two. The final answer for density of $Z - W$ is

$$f(u) = 2(1 - u),\ 0 < u < 1$$

4. Let X, Y be iid exponential random variables with mean 1. Let $Z = X + Y$ and $W = X - Y$. Find the conditional density $f_{W|Z}$

Hint: You have a straight formula for joint density of Z, W here. Since sum of two independent exponential rv would be a gamma rv, you know f_Z. Hence, you can write down the expression for the conditional density. The final answers are

$$\begin{aligned} f_{ZW}(z, w) &= \frac{1}{2}e^{-z}, \quad -z < w < z, \ z > 0 \\ f_Z(z) &= ze^{-z}, \quad z > 0 \\ f_{Z|W}(z|w) &= \frac{1}{2z}, \quad -z < w < z, \ z > 0 \end{aligned}$$

5. Let X, Y be independent Gaussian random variables with mean zero and variance unity. Define random variables D and θ by

$$D = X^2 + Y^2; \qquad \theta = \tan^{-1}(Y/X)$$

(where, by convention, we assume θ takes values in $[0, 2\pi]$; for this we first calculate $\tan^{-1}(|Y|/|X|)$ in the range $[0, \pi/2]$ and then put that angle in the appropriate quadrant based on signs of Y and X). Find the joint density of D and θ and their marginal densities. Are D and θ independent? (Hint: Note that the (D, θ) to (X, Y) mapping is invertible. Hence you can use the formula).

Hint: The inverse of the given transformation is $X = \sqrt{D}\ \cos(\theta)$ and $Y = \sqrt{D}\ \sin(\theta)$. The jacobian is 1/2 and by applying the formula you see that D is exponential with $\lambda = \frac{1}{2}$ and θ is uniform over $[0,\ 2\pi]$ and they are independent.

Note that if you start with independent D, θ with these densities and define X, Y by $X = \sqrt{D}\ \cos(\theta)$ and $Y = \sqrt{D}\ \sin(\theta)$, then the same calculation would show you that X, Y are ind standard Gaussian.

6. Consider the following algorithm for generating random numbers X and Y :

   1. Generate U_1 and U_2 uniform over $[0,\ 1]$.
   2. Set $X = \sqrt{-2\log(U_1)}\cos(2\pi U_2)$ and $Y = \sqrt{-2\log(U_1)}\sin(2\pi U_2)$.

   What would be the joint distribution of X and Y ? (Hint: Recall that when U is uniform over $[0,\ 1]$, we know that $-a\log(U)$ is exponential with parameter $1/a$ and $2\pi U$ is uniform over $[0, 2\pi]$).

Hint: From the hint in the problem, you know that $-2\log(U_1)$ is exponential with $\lambda = \frac{1}{2}$. When $U_2 \sim U[0,\ 1]$, $2\pi U_2$ is uniform over $[0,\ 2\pi]$. Now, from the previous problem, you can easily see, X, Y are ind Gaussian with mean zero variance 1.

This tells you a simple algorithm whereby using two uniform random numbers, you can generate Gaussian random numbers. This is a very popular method for generating Gaussian random numbers and these equations are known as Box-Muller transform.

7. Consider the following algorithm for generating random variables V_1 and V_2 :

    1. Generate X_1 and X_2 uniform over $[-1,\ 1]$.
    2. If $X_1^2 + X_2^2 > 1$ then go to step 1; else set $V_1 = X_1,\ V_2 = X_2$ and exit.

    What would be the joint distribution of V_1 and V_2 ?

Hint: This process always stops with a point on or inside the unit circle. So, V_1, V_2 would be uniform over the unit disc, The way to see this is as follows. Take any subset of the unit disc. Call it A. Now suppose you repeat the above till (V_1, V_2) is either in A or in A^c. The probability that you end in A is now proportional to its area.

This is a generic method to generate random vectors uniform over some region (which is not a cylindrical set). In $\Re^2$ if X, Y are uniform over a cylindrical set then they are independent and each is uniform over the corresponding interval. We can easily generate them. Thus generating X_1, X_2 uniform over $[-1,\ 1] \times [-1,\ 1]$ is easy. But to generate them uniform over the interior of unit circle is difficult because they are no longer independent. The above gives a method of handling this.

You can also find the expected time complexity of this method. Find the expected number of random variables uniform over $[-1,\ 1]$ that you need to generate to get one random vector uniform over the interior of unit circle.

8. Suppose we have access to a random number generator that can generate random numbers uniformly distributed over $(0, 1)$. Using the results of the previous problems, suggest a method for generating samples of X when X has Gaussian density with mean zero and variance unity.

Hint: We already saw the Box-Muller transform for this. One difficulty with equations you had there is that you need to compute $\cos(\theta)$ and $\sin(\theta)$ which have to be computed through Taylor series and hence are computationally expensive.

Suppose we can generate (V_1, V_2) uniform over the interior of the unit circle. Let θ be the angle that the line joining origin to (V_1, V_2) makes with X-axis. Then, that θ would be uniform over $[0,\ 2\pi]$. Hence, $\frac{V_1}{\sqrt{V_1^2+V_2^2}}$ would have the same distribution as that of $\cos(\theta)$ with θ being uniform over $[0,\ 2\pi]$. Hence we can replace $\cos(\theta)$ in Box-Muller transform by $\frac{V_1}{\sqrt{V_1^2+V_2^2}}$ with V_1, V_2 generated as in the previous problem. This improves computational efficiency of Box-Muller transform

9. Let $p_i, q_i,\ i = 1, \cdots, N$, be positive numbers such that $\sum_{i=1}^{N} p_i = \sum_{i=1}^{N} q_i = 1$ and $p_i \leq Cq_i,\ \forall i$ for some positive constant C. Consider the following algorithm to simulate a random variable, X:

    1. Generate a random number Y such that $P[Y = j] = q_j,\ j = 1, \cdots, N$. (That is, the mass function of Y is $f_Y(j) = q_j$).
    2. Generate U uniform over $[0,\ 1]$.
    3. Suppose the value generated for Y in step-1 is j. If $U < (p_j/Cq_j)$, then set $X = Y$ and exit; else go to step-1.

    On any iteration of the above algorithm, if condition in step-3 becomes true, we say the generated Y is accepted. Find the value of $P[Y \text{ is accepted} \mid Y = j]$. Show that $P[Y \text{ is accepted},\ Y = j] = p_j/C$. Now calculate $P[Y \text{ is accepted}]$. Use these to calculate the mass function of X.

Hint: In the third step of the algorithm, if $Y = j$ then it is accepted if $U < (p_j/Cq_j)$. U is uniform and by the condition on C, $(p_j/Cq_j) \leq 1$. Hence, $P[Y \text{ is accepted} \mid Y = j] = (p_j/Cq_j)$. Now you get the second part of the question by noting that $p[Y = j] = q_j$. Summing $P[Y \text{ is accepted},\ Y = j]$ over j you get $P[Y \text{ is accepted}] = 1/C$. $X = j$ can happen by exiting the loop with $Y = j$. Exiting the loop on the n^{th} time with $Y = j$, for different n, constitute mutually exclusive events and the union of all these is the event of $X = i$ and thus you get the mass function of X as $f_X(i) = p_i$

Comment: This is known as rejection-sampling method to generate a sample of the random variable X. Here, the distribution of Y is called the proposal distribution. If, generating Y is much simpler than generating X, this would be a useful method.

10. Suppose X is a discrete rv taking values $\{x_1, x_2, \cdots, x_m\}$ with probabilities $p_1, \cdots p_m$. The usual method of simulating such a rv is as follows. We divide the $[0,\ 1]$ interval into bins of length p_1, p_2 etc. Then we generate a rv, uniform over $[0,\ 1]$ and depending on the bin it falls in, we decide on the value for X. That is, if $U \leq p_1$ we assign $X = x_1$; if $p_1 < U \leq p_1 + p_2$ then we assgn $X = x_2$ and so on.
    Suppose X is a discrete random variable taking values $1, 2, \cdots, 10$. Its mass function is: $f_X(1) = 0.08, f_X(2) = 0.13, f_X(3) = 0.07, f_X(4) = 0.15, f_X(5) = 0.1, f_X(6) = 0.06, f_X(7) = 0.11, f_X(8) = 0.1, f_X(9) = 0.1, f_X(10) = 0.1$. Can you use the result of previous problem to suggest an efficient method for simulating X.

Hint: Here take the given distribution of X as the p_i's in the previous problem. Take $q_i = 0.1,\ 1 \leq i \leq 10$. Now $p_i \leq Cq_i,\ \forall i$ is satisfied with $C = 1.5$. Generating Y is very simple here: take Y as the integer part of $10U$ where U is uniform over $[0,\ 1]$. What role does C have in determining the efficiency of this method?