Math 350 - Homework 3 - Solutions

1. *The bus will arrive at a time that is uniformly distributed between* 8 *and* 8 : 30 *A.M. If we arrive at* 8 *A.M., what is the probability that we will wait between* 5 *and* 15 *minutes?*

   The probability that we will have to wait between 5 and 15 minutes, having arrived at the bus stop at 8 A.M., is the same as the probability that the bus will arrive between 8 : 05 A.M. and 8 : 15 A.M. This probability is $p = (15 - 5)/30 = 1/3$.

2. *Let X be a binomial random variable with parameters (n, p). Explain why*

$$P\left\{\frac{X - np}{\sqrt{np(1-p)}} \leq x\right\} \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-x^2/2}\, dx.$$

   A binomial random variable X is a sum of i.i.d. random variables $X = X_1 + \cdots + X_n$, where each X_i takes values in $\{0, 1\}$ with probabilities $P(X_i = 1) = p$, $P(X_i = 0) = 1 - p$. The mean and variance of the X_i are

$$\mu = E[X_i] = 0P(X_i = 0) + 1P(X_i = 1) = p,$$

   and

$$\sigma^2 = \mathrm{Var}(X_i) = (0 - p)^2 P(X_i = 0) + (1 - p)^2 P(X_i = 1) = p(1 - p).$$

   Thus the expression under curly brackets is equal to

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

   Therefore, the claim is an immediate consequence of the central limit theorem applied to the i.i.d. $X_1, X_2, \ldots$.

3. *For a Poisson process with rate λ, find $P\{N(s) = k|N(t) = n\}$ when $s < t$.*

   Since a Poisson process $N(t)$ is non-decreasing, we know that

$$P\{N(s) = k|N(t) = n\} = 0 \text{ if } s < t \text{ and } k > n.$$

   Therefore, we will assume that $k \leq n$. From the definition of conditional probability, it follows immediately that for any two events A, B of positive probability

$$P(A|B)P(B) = P(B|A)P(A).$$

Therefore,

$$
\begin{aligned}
P\{N(s)=k|N(t)=n\} &= P\{N(t)=n|N(s)=k\}\frac{P\{N(s)=k\}}{P\{N(t)=n\}} \\
&= P\{N(t-s)=n-k\}\frac{P\{N(s)=k\}}{P\{N(t)=n\}} \\
&= \frac{e^{-\lambda(t-s)}\left(\lambda(t-s)\right)^{n-k}}{(n-k)!}\cdot\frac{e^{-\lambda s}\left(\lambda s\right)^{k}}{k!}\cdot\frac{n!}{e^{-\lambda t}(\lambda t)^n} \\
&= \binom{n}{k}\frac{s^k(t-s)^{n-k}}{t^n}.
\end{aligned}
$$

Therefore,

$$
P\{N(s)=k|N(t)=n\} = \begin{cases} \binom{n}{k}\frac{s^k(t-s)^{n-k}}{t^n} & \text{if } k \leq n \\ 0 & \text{if } k > n. \end{cases}
$$

This result makes sense: notice that, by applying the binomial formula, the sum of these conditional probabilities over k, for $k = 0, \ldots, n$, is 1.

4. *An urn contains four white and six black balls. A random sample of size* 4 *is chosen. Let* X *denote the number of white balls in the sample. An additional ball is now selected from the remaining six balls in the urn. Let* Y *equal* 1 *if this ball is white and* 0 *if it is black. Find*

   (a) $E[Y|X=2]$.

   (b) $E[X|Y=1]$.

   (c) *Var*$(Y|X=0)$.

   (d) *Var*$(X|Y=1)$.

   First observe that the following conditional probabilities are easy to obtain by a simple counting argument:

$$
P(Y=1|X=j) = \frac{4-j}{6}, \quad P(Y=0|X=j) = \frac{2+j}{6},
$$

   where $j = 0, 1, 2, 3, 4$. We also need the probabilities $P(Y = 1, X = j)$. I do this by explicitly counting the elementary outcomes of each event. An elementary outcome can be represented by a string of 0s and 1s of length 5. A moment's thought should convince you that the probability of each string only depends on the number of 0s and 1s, and not on their order.

   (a)

$$
P\{Y=1, X=0\} = P\{(0,0,0,0,1)\} = \frac{6\cdot5\cdot4\cdot3\cdot4}{10\cdot9\cdot8\cdot7\cdot6} = \frac{1}{21}.
$$

   (b)

$$
P\{Y=1, X=1\} = P\{(1,0,0,0,1),(0,1,0,0,1),(0,0,1,0,1),(0,0,0,1,1)\} = 4\frac{6\cdot5\cdot4\cdot4\cdot3}{10\cdot9\cdot8\cdot7\cdot6} = \frac{4}{21}.
$$

   (c)

$$
\begin{aligned}
P\{Y=1, X=2\} &= P\{(1,1,0,0,1),(1,0,1,0,1),(1,0,0,1,1),(0,1,1,0,1),(0,1,0,1,1),(0,0,1,1,1)\} \\
&= 6\frac{6\cdot5\cdot4\cdot3\cdot2}{10\cdot9\cdot8\cdot7\cdot6} = \frac{1}{7}.
\end{aligned}
$$

(d)

$$P\{Y=1, X=3\} = P\{(1,1,1,0,1),(1,1,0,1,1),(1,0,1,1,1),(0,1,1,1,1)\} = 4\frac{4\cdot 3\cdot 2\cdot 6\cdot 1}{10\cdot 9\cdot 8\cdot 7\cdot 6} = \frac{2}{105}.$$

(e)

$$P\{Y=1, X=4\} = P(\emptyset) = 0.$$

From the above it follows, for example, that

$$P\{Y=1\} = \sum_{j=0}^{4} P\{Y=1, X=j\} = \frac{1}{21} + \frac{4}{21} + \frac{1}{7} + \frac{2}{5\times 21} = \frac{42}{105} = \frac{2}{5}.$$

It is interesting to observe here that we could have obtained this last probability much more easily by the following argument: the probability of each elementary outcome $(a_1, \dots, a_5)$ only depends on the number of 0s and 1s, and not on their order. Therefore, the probability that the last entry is 1 is the same as the probability that the first entry is one. But this probability is clearly $4/10 = 2/5$.

We now calculate the asked for values:

(a) $E[Y|X=2] = 0P(Y=0|X=2) + 1P(Y=1|X=2) = P(Y=1|X=2) = \frac{1}{3}$.

(b) $E[X|Y=1] = \sum_{j=0}^{4} jP\{X=j|Y=1\} = \sum_{j=0}^{4} j\frac{P\{Y=1,X=j\}}{P\{Y=1\}} = \frac{\frac{4}{21}+\frac{2}{7}+\frac{3\times 2}{105}}{\frac{2}{5}} = \frac{4}{3}$.

(c) We know that $\mathrm{Var}(Y|X=0) = E[Y^2|X=0] - E[Y|X=0]^2$. Since Y only takes values $0,1$, we also have $Y^2 = Y$. Now, $E[Y|X=0] = 0P\{Y=0|X=0\} + 1P\{Y=1|X=0\} = 4/6 = 2/3$. Therefore,

$$\mathrm{Var}(Y|X=0) = \frac{2}{3} - \left(\frac{2}{3}\right) = \frac{2}{9}.$$

(d) We can calculate $E[X^2|Y=1]$ as we did for $E[X|Y=1]$ in part (b). The result is

$$E[X^2|Y=1] = \frac{\frac{4}{21}+\frac{2^2}{7}+\frac{3^2\times 2}{105}}{\frac{2}{5}} = \frac{7}{3}.$$

Therefore,

$$\mathrm{Var}(X|Y=1) = E[X^2|Y=1] - E[X|Y=1]^2 = \frac{7}{3} - \left(\frac{4}{3}\right)^2 = \frac{5}{9}.$$

We can gain some confidence that the above results are correct by running a simulation of the process. The following computes the conditional expectation $E[X|Y=1]$ by simulating 500000 trials of the experiment. The conditional expectation is obtained by averaging the number of white balls over only the trials having white 5th ball.

```
m=500000;     %Number of trials
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
p=4/10;       %Probability that the first ball is white
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
O=zeros(m,5); %Each row of matrix O is a string of five
              %number from {0,1}, where 1 stands for white
              %and 0 for black.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
R=rand(m,5);  %Matrix R contains all the random numbers we will
```

```
                %need below.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
O(:,1)=(R(:,1)<=p); %Fills the first column of O with 1s with probability
                   %p=4/10 and 0 with probability 6/10.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
W=O(:,1); %W counts the number of white balls at each step of each trial.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
for j=2:5 %We now draw the second, third, ..., fith balls.
    p = (4*ones(m,1)-W)/(11-j); %Probability of picking a white ball
                                %taking into account the number of white
                                %balls already drawn.
    O(:,j) = (R(:,j)<=p);%Draw new balls for position j
    W = sum(O(:,1:j),2); %Count number of white balls already gotten.
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
a=find(O(:,5)==1); %Find the indices of the rows of O corresponding
                   %to trials that have white 5th ball.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
C=O(a,1:4);%C is a matrix containing only the outcomes of the trials
           %for which the 5th ball is white. It registers only the first
           %four balls
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
N=sum(C,2);%N is the number of white balls among the first 4 in each trial
           %trial for which the 5th ball is white.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
[n,k]=size(N);%We need the number n of trials for which the 5th ball is white.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
f=sum(N)/n;%f is the average number of white balls among the first 4 drawn
           %given that the 5th is white. This is an approximation of
           %the conditional expectation E[X|Y=1].
```

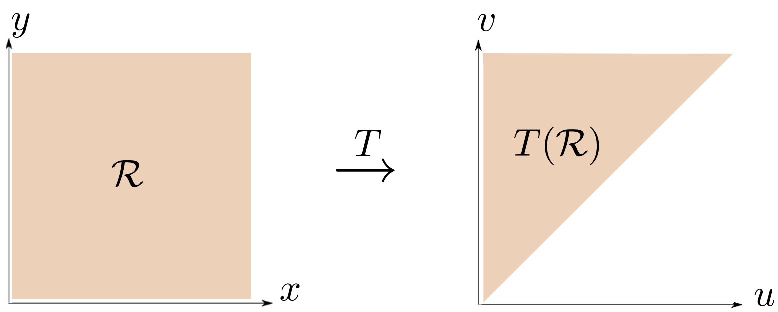
One run of this program gives the typical value $f = 1.3366$. This is to be compared with the theoretical solution above (part (b)), which gives $4/3 = 1.333...$

5. *If X and Y are independent and identically distributed exponential random variables, show that the conditional distribution of X, given that $X + Y = t$, is the uniform distribution on $(0, t)$.*

   We first need to discuss some preliminaries. Let X, Y be two random variables with joint probability density equal to $f(x, y)$. Suppose we define new random variables U, V in terms of X, Y by $U = T_1(X, Y)$ and $V = T_2(X, Y)$, or $(U, V) = T(X, Y)$ for short. For example, let $(U, V) = (X, X + Y)$. This transformation maps the positive octant of $\mathbb{R}^2$ (denoted by $\mathcal{R}$ in the figure) to the wedge shaped region $T(\mathcal{R})$.

   A preliminary question we wish to solve is: what is the joint probability density for the new random variables U, V? Let us call this density $g(u, v)$. The fundamental observation is that if A denotes an event defined in terms of X, Y (that is, a subset of the region $\mathcal{R}$), then

$$P((X, Y) \in A) = P((U, V) \in T(A)),$$



where $T(A)$ is the image of A under the transformation T (an event defined in terms of U, V, hence a subset of $T(\mathcal{R})$). It is not hard to show that this equality implies, any function $F(U,V)$, an equality of expectations:

$$E[F(U,V)] = E[F(T(U,V))],$$

where the first expectation is calculated using the density $g(u,v)$ and the second using the density $f(x,y)$. This means that

$$\iint_{T(\mathcal{R})} F(u,v)g(u,v)\,dudv = \iint_{\mathcal{R}} F(T(x,y))f(x,y)\,dxdy.$$

From the general formula for change of coordinates in multiple integrals, the integral on the right satisfies:

$$\iint_{\mathcal{R}} F(T(x,y))f(x,y)\,dxdy = \iint_{T(\mathcal{R})} F(u,v)f(T^{-1}(u,v))|D(u,v)|\,dudv,$$

where $|D(u,v)|$ indicates the absolute value of the Jacobian determinant of the inverse transformation T^{-1}, namely

$$D(u,v) = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}.$$

Since the equality of integrals holds for all functions F, we conclude that

$$g(u,v) = f(T^{-1}(u,v))|D(u,v)|.$$

We can now return to the specific situation of the problem. Here the transformation is $T(x,y) = (x, x+y)$, which has inverse $T^{-1}(u,v) = (u, v-u)$. Notice that $v = z$, in the notation of the exercise. The Jacobian determinant is the absolute value of the determinant of the matrix $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, which is 1. So $g(u,v) = f(u, v-u)$. The function $f(x,y)$ is the joint density of two independent, identically distributed exponential random variables X, Y, so $f(x,y)$ is written as the product of the densities of X and Y:

$$f(x,y) = \begin{cases} \left(\lambda e^{-\lambda x}\right)\left(e^{-\lambda y}\right) = \lambda^2 e^{-\lambda(x+y)} & \text{if } x, y \geq 0 \\ 0 & \text{if } x < 0 \text{ or } y < 0. \end{cases}$$

Therefore,

$$g(u,v) = \begin{cases} \lambda^2 e^{-\lambda v} & \text{if } 0 \leq u \leq v \\ 0 & \text{if } u > v. \end{cases}$$

This shows that the conditional probability density for X given $Z = t$ is

$$g(x|z=t) = \begin{cases} \frac{\lambda^2 e^{-\lambda t}}{\int_0^t \lambda^2 e^{-\lambda t}\, dx} = \frac{1}{t} & \text{if } 0 \leq x \leq t \\ 0 & \text{if } x > t. \end{cases}$$

But this is exactly the probability density function for a uniform random variable taking values in the interval $[0, t]$.