Homework # 2 Answer key

1)

x	0.5	1	1.75	3
$P(X=x)$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$

These values can be calculated by measuring the height of the "jumps" in the cdf $F_X(x)$.

2) We can use the inverse cdf method as follows:

$$F_Y(y) = P(Y \le y) = P(e^X \le y)$$

$$= P(X \le \log(y)) = \int_0^{\log(y)} 1 \, dx$$

$$= \log(y), \quad \text{for } 1 < y < e$$

It follows that the pdf of Y is

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy}(\log(y))$$

$$= \frac{1}{y}, \quad 1 < y < e$$

So, $f_Y(y) = \frac{1}{y} I\{1 < y < e\}$.

3) a) We need to find C so that the density integrates to 1. So, we solve

$$\int_0^2 C(4x - 2x^2)\,dx = 1$$

$$\Rightarrow C\left[2x^2\Big|_0^2 - \frac{2}{3}x^3\Big|_0^2\right] = 1$$

$$\Rightarrow 8 - \frac{16}{3} = \frac{1}{C} \Rightarrow \frac{8}{3} = \frac{1}{C} \Rightarrow C = \frac{3}{8}$$

b) $$P(X > 1) = \int_1^2 \frac{3}{8}(4x - 2x^2)\,dx$$

$$= \frac{3}{8}\left[2x^2\Big|_1^2 - \frac{2}{3}x^3\Big|_1^2\right]$$

$$= \frac{3}{8}\left[8 - 2 - \frac{16}{3} + \frac{2}{3}\right] = \frac{3}{8}\left[6 - \frac{14}{3}\right]$$

$$= \frac{3}{8}\left(\frac{4}{3}\right) = \frac{1}{2}$$

4) We need to think about two possibilities for each fixed p: either $U < p$ or $U > p$. When $U > p$, the p will be on the left stick, which will have ~~densit~~ length U. However, when $U < p$, p will be on the right stick, which will have length $1-U$. So, let $L_p(U)$ = length of the stick on which p lies. We have:

$$L_p(u) = \begin{cases} U, & U > p \\ 1-U, & U < p \end{cases}$$

4) continued...

So, we just need $E[L_p(U)]$:

$$E[L_p(U)] = \int_0^1 L_p(U)\, du$$

$$= \int_0^p 1-u\, du + \int_p^1 u\, du$$

$$= p - \frac{1}{2}p^2 + \frac{1}{2} - \frac{1}{2}p^2$$

$$= \frac{1}{2} + p - p^2 = \frac{1}{2} + p(1-p)$$

5) First we calculate the marginal pdfs $f_X(x)$ & $f_Y(y)$:

$$f_X(x) = \int_0^\infty 2e^{-x} e^{-2y}\, dy = -e^{-x} e^{-2y} \Big|_{y=0}^{y=\infty}$$

$$= e^{-x}, \quad x \in (0, \infty)$$

$$f_Y(y) = \int_0^\infty 2e^{-x} e^{-2y}\, dx = -2e^{-x} e^{-2y} \Big|_{x=0}^{x=\infty}$$

$$= 2e^{-2y}, \quad y \in (0, \infty)$$

Then, we use the law of total probability

$$P(X < Y) = \int_y P(X < y \mid Y = y)\, f_Y(y)\, dy$$

$$= 2\int_0^\infty \left(\int_0^y f_X(x)\, dx \right) e^{-2y}\, dy$$

5) continued...

$$= 2\int_0^\infty \left(\int_0^y e^{-x}\,dx\right) e^{-2y}\,dy$$

$$= 2\int_0^\infty \left(-e^{-x}\Big|_{x=0}^{x=y}\right) e^{-2y}\,dy$$

$$= 2\int_0^\infty (1-e^{-y})\,e^{-2y}\,dy$$

$$= 2\int_0^\infty e^{-2y} - e^{-3y}\,dy$$

$$= 2\left[-\frac{1}{2}e^{-2y}\Big|_0^\infty + \frac{1}{3}e^{-3y}\Big|_0^\infty\right]$$

$$= 2\left[\frac{1}{2}-\frac{1}{3}\right] = 2\left(\frac{1}{6}\right) = \frac{1}{3}$$

So, $P(X<Y) = \frac{1}{3}$

6) We can use the convolution formula for each of these calculations.

First, note that $f_X(x) = 1, \quad x \in (0,1)$

$$F_X(x) = x$$

$$f_Y(y) = e^{-y}, \quad y > 0$$

$$F_Y(y) = 1 - e^{-y}$$

Thus, $F_Z(a) = F_{X+Y}(a) = \int_y F_X(a-y)\,f_Y(y)\,dy$

$$= \int_y (a-y)\,e^{-y}\,dy \qquad (*)$$

We have to be smart about the limits of integration here. Note that

6) continued...

$$x+y=a \Rightarrow y=a-x$$

Since $x \in (0,1)$, it follows that y must be between $a-1$ & a. Moreover, since $y>0$ it follows that $y \in (\max\{0,a-1\}, a)$

So, we can calculate $(*)$ as:

$$= \int_{\max\{0,a-1\}}^{a} (a-y)e^{-y}\,dy = ae^{-\max\{0,a-1\}} - ae^{-a} - \int_{\max\{0,a-1\}}^{a} ye^{-y}\,dy.$$

Then, the density can be calculated by taking derivatives. (Not all work is shown here...)

7) $$P(X=x \mid X+Y=n) = \frac{P(X=x,\ X+Y=n)}{P(X+Y=n)}$$

$$= \frac{P(X=x,\ Y=n-x)}{P(X+Y=n)}$$

$$= \frac{P(X=x)\,P(Y=n-x)}{P(X+Y=n)} \quad \text{by independe}$$

Now, it is easy to show that if $X \sim Po(\lambda_1)$ and $Y \sim Po(\lambda_2)$ and they are independent,

7) continued...

then $X + Y \sim \underline{Po}(\lambda_1 + \lambda_2)$.

It follows that

$$\underline{P}(X = x \mid X + Y = n) = \frac{\lambda_1^x e^{-\lambda_1} \lambda_2^{n-x} e^{-\lambda_2} n!}{(\lambda_1 + \lambda_2)^n e^{-\lambda_1 - \lambda_2} x!(n-x)!}$$

$$= \left(\frac{\lambda_1}{\lambda_2}\right)^x \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^n \binom{n}{x}$$

$$= \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n-x} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^x \binom{n}{x}$$

which is the form of a $\text{Bin}\left(n, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$

random variable! Hence, $X \mid X + Y = n \sim \text{Bin}\left(n, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$

8) Note: One can calculate these probabilities by investigating each of 36 total outcomes. I don't do this here, but the calculation (though long) is straightforward.

<u>Notes on computation</u>:

① $X + Y$ is just a Gamma(1, 2) random variable. This can be found using the convolution formula.

But the point of this exercise is to show that simulation can be used to obtain the distribution of a convolution and can be done pretty quickly.

(2) Strategy 2 turns out to be more efficient.