14.30 PROBLEM SET 4 - SUGGESTED ANSWERS

Problem 1

a. The distribution of $Y|X$ is the distribution of $\alpha + \beta X + \varepsilon$, where ε is the only random variable and $\alpha + \beta X$ is fixed, call it C. Then, the distribution of $C + e$ is that of e , shifted by the value C, or $U[C - \frac{1}{2}, C + \frac{1}{2}]$. Thus $Y|X \sim U\left[\alpha + \beta X - \frac{1}{2}, \alpha + \beta X + \frac{1}{2}\right]$.

b. For a random variable Z with a uniform distribution $U[a, b]$, $f_Z(z) = 1/(b-a)$,

$$E[Z] = \int_a^b \frac{z}{b-a} dz = \left[\frac{z^2}{2(b-a)}\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2};$$

$$E[Z^2] = \int_a^b \frac{z^2}{b-a} dz = \left[\frac{z^3}{3(b-a)}\right]_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{b^2+a^2+2ab}{3};$$

$$Var[Z] = E[Z^2] - E[Z]^2 = \frac{b^2+a^2+2ab}{3} - \frac{(b+a)^2}{4} = \frac{(b-a)^2}{12}.$$

Therefore, for $U[\alpha + \beta X - 0.5, \alpha + \beta X + 0.5]$,

$$E[Y|X] = \frac{b+a}{2} = \frac{\alpha+\beta X-0.5+\alpha+\beta X+0.5}{2} = \alpha + \beta X;$$

$$Var[Y|X] = \frac{(b-a)^2}{12} = \frac{[\alpha+\beta X+0.5-(\alpha+\beta X-0.5)]^2}{12} = \frac{1}{12}.$$

c. We can calcualte the expected value directly:

$$\begin{aligned} E[Y] &= E[\alpha + \beta X + e] \\ &= \alpha + \beta E[X] + E[e] \\ &= \alpha + \beta \cdot 0 + 0 \\ &= \alpha \end{aligned}$$

Or, we could use the law of iterated expectations:

$$\begin{aligned} E[Y] &= E_X[E_Y[Y|X]] \\ &= E_X[\alpha + \beta X] \\ &= \alpha + \beta E[X] \\ &= \alpha \end{aligned}$$

For the variance, we can again calculate either directly or using our conditional variance identity.

$$
\begin{aligned}
Var[Y] &= Var[\alpha + \beta X + e] \\
&= \beta^2 Var[X] + Var[e] + 2\beta Cov(X, e) \\
&= \beta^2\sigma^2 + \frac{[0 + 0.5 - (0 - 0.5)]^2}{12} + 2\beta \cdot 0 \\
&= \beta^2\sigma^2 + \frac{1}{12}
\end{aligned}
$$

or

$$
\begin{aligned}
Var[Y] &= E(Var(Y|X)) + Var(E(Y|X)) \\
&= E(\frac{1}{12}) + Var(\alpha + \beta X) \\
&= \frac{1}{12} + \beta^2 Var[X] \\
&= \frac{1}{12} + \beta^2\sigma^2
\end{aligned}
$$

Problem 2

a. The moment generating function is defined as

$$M_X(t) = E\left(e^{tX}\right)$$

Since X is distributed uniformly over $[0, 4]$, we have

$$
\begin{aligned}
M_X(t) &= E\left(e^{tX}\right) \\
&= \int_0^4 \frac{1}{4} e^{tx} dx \\
&= \frac{1}{4t}\left(e^{4t} - 1\right)
\end{aligned}
$$

To find the mean, we need to take the first derivative of the MGF and evaluate it at $t = 0$, using l'Hopital's rule in the fourth line.

$$
\begin{aligned}
\frac{\partial}{\partial t} M_X(t)\,\Big|_{t=0} &= \frac{\partial}{\partial t}\frac{1}{4t}\left(e^{4t} - 1\right)\Big|_{t=0} \\
&= \frac{e^{4t}}{t} - \frac{e^{4t} - 1}{4t^2}\Big|_{t=0} \\
&= \frac{4te^{4t} - e^{4t} + 1}{4t^2}\Big|_{t=0} \\
&= \frac{16te^{4t} + 4e^{4t} - 4e^{4t}}{8t}\Big|_{t=0} \\
&= 2 = E(X)
\end{aligned}
$$

Then, to find the variance, we will take the second derivative of the MGF at $t = 0$ (again making use of l'Hopital's rule), and then subtract the square of the expected value.

$$\begin{aligned}
\frac{\partial^2}{\partial t^2} M_X(t)\Big|_{t=0} &= \frac{4te^{4t}-e^{4t}}{t^2} - \frac{16t^2e^{4t}-8t\left(e^{4t}-1\right)}{16t^4}\Big|_{t=0} \\
&= \frac{4te^{4t}-e^{4t}}{t^2} - \frac{2te^{4t}-e^{4t}+1}{2t^3}\Big|_{t=0} \\
&= \frac{8t^2e^{4t}-4te^{4t}+e^{4t}-1}{2t^3}\Big|_{t=0} \\
&= \frac{32t^2e^{4t}+16te^{4t}-16te^{4t}-4e^{4t}+4e^{4t}}{6t^2}\Big|_{t=0} \\
&= \frac{16e^{4t}}{3}\Big|_{t=0} \\
&= \frac{16}{3} = E\left(X^2\right) \\
V(X) &= \frac{16}{3} - 4 = \frac{4}{3}
\end{aligned}$$

b. The Chebyshev inequality states that $\Pr\left(|X-E[X]| \geq t\right) \leq \frac{Var[X]}{t^2}$ for $t > 0$. Here: $E[X] = 2, Var[X] = \frac{4}{3}, t = \frac{3}{2}$ so: $\Pr\left(X \notin (0.5, 3.5)\right) = \Pr\left(|X-2| \geq \frac{3}{2}\right) \leq \frac{(4/3)}{(3/2)^2} = \frac{16}{27}$.

c. $\Pr\left(X \notin (0.5, 3.5)\right) = \int_0^{0.5} \frac{1}{4}dx + \int_{3.5}^{4} \frac{1}{4}dx = \frac{1}{4} < \frac{16}{27}$ so we get a lower result than in part a.

d. The Chebyshev inequality is very useful for evaluating distributions for which you only know the mean and the variance, but not the actual distribution. If you know the actual distribution you can get a more precise answer. But this is only because you are using additional information.

Problem 3

a. As instructed, we will find the pdf of Y using both the 1-step and 2-step methods. Note that Y will take on values in $[0,1]$. We begin with the 2-step method by calculating the CDF:

$$\begin{aligned}
F_Y(y) &= \Pr(Y \leq y) = \Pr\left(X^2 \leq y\right) \\
&= \Pr\left(-\sqrt{y} \leq X \leq \sqrt{y}\right) \\
&= \Pr\left(0 \leq X \leq \sqrt{y}\right) \\
&= F_X\left(\sqrt{y}\right) = \sqrt{y} \\
f_Y(y) &= \frac{1}{2}y^{-\frac{1}{2}} \text{ for } y \in [0,1] \\
&= 0 \text{ elsewhere}
\end{aligned}$$

Then we use the 1-step method. Because the transformation is already monotonic on the relevant range, we do not have to worry about dividing the range into monotonic pieces.

$$\begin{aligned} f_Y(y) &= f_X\left(r^{-1}(y)\right)\left|\frac{\partial r^{-1}(y)}{\partial y}\right| \text{ for } y \in [0,1] \\ f_Y(y) &= f_X\left(\sqrt{y}\right)\left|\frac{1}{2}y^{-\frac{1}{2}}\right| \text{ for } y \in [0,1] \\ &= \frac{1}{2}y^{-\frac{1}{2}} \text{ for } y \in [0,1] \\ &= 0 \text{ elsewhere} \end{aligned}$$

b. Now we have $Z = -\lambda \ln X$. So the range of Z is $[0,\infty)$. We again have a monotonic transformation, so we will use the 1-step method. Note that $r^{-1}(z) = e^{-\frac{z}{\lambda}}$.

$$\begin{aligned} f_Z(z) &= f_X\left(r^{-1}(z)\right)\left|\frac{\partial r^{-1}(z)}{\partial z}\right| \text{ for } z \in [0,\infty) \\ &= f_X\left(e^{-\frac{z}{\lambda}}\right)\left|-\frac{e^{-\frac{z}{\lambda}}}{\lambda}\right| \text{ for } z \in [0,\infty) \\ &= \frac{e^{-\frac{z}{\lambda}}}{\lambda} \text{ for } z \in [0,\infty) \\ &= 0 \text{ elsewhere} \end{aligned}$$

Note that this is the exponential distribution. It turns out that most distributions can be constructed as a transformation of a $U[0,1]$ random variable.

Problem 4

Because $Y = \frac{X}{X+1}$ is a one-to-one transformation, we know that there will be only one x value that corresponds to each valid y value.

$$\begin{aligned} f_Y(y) &= \Pr(Y = y) = \Pr\left(\frac{X}{X+1} = y\right) \\ &= \Pr\left(X = \frac{y}{1-y}\right) = f_X\left(\frac{y}{1-y}\right) \\ &= \frac{1}{3}\left(\frac{2}{3}\right)^{\frac{y}{1-y}} \end{aligned}$$

for valid values of y. What is the support of Y? We can substitute in for the first few potential values of X to see the pattern:

$$\begin{aligned}\frac{0}{0+1} &= 0\\ \frac{1}{1+1} &= \frac{1}{2}\\ \frac{2}{2+1} &= \frac{2}{3}\\ \frac{3}{3+1} &= \frac{3}{4}\end{aligned}$$

so $Y \in \{0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, ...\}$, or $Y \in \{y|y = \frac{x}{x+1}$ for whole numbers $x\}$

Problem 5

a. We will use the 1-step method, and so we must first divide the range of X into segments for which the transformation function is monotone, $(-1,0)$ and $(0,1)$ (we can ignore the endpoints of these intervals because the probability that X equals a particular point is zero). We will use the 1-step method on each segment and then sum our results.

For the segment $(-1,0)$, our transformation function is $r(x) = x^2$, and the inverse function is $r^{-1}(y) = -\sqrt{y}$. The range of Y is $(0,1)$. So we have

$$\begin{aligned}f_Y(y) &= f_X(-\sqrt{y})\left|-\frac{1}{2}y^{-\frac{1}{2}}\right| \text{ for } y \in (0,1)\\ &= \frac{1}{4}(1-\sqrt{y})\,y^{-\frac{1}{2}} \text{ for } y \in (0,1)\\ &= \frac{1}{4}\left(y^{-\frac{1}{2}}-1\right) \text{ for } y \in (0,1)\\ &= 0 \text{ elsewhere}\end{aligned}$$

Then, for the segment $(0,1)$, our transformation function is $r(x) = x^2$, and the inverse function is $r^{-1}(y) = \sqrt{y}$. The range of Y is again $(0,1)$. So we have

$$\begin{aligned}f_Y(y) &= f_X(\sqrt{y})\left|\frac{1}{2}y^{-\frac{1}{2}}\right| \text{ for } y \in (0,1)\\ &= \frac{1}{4}(1+\sqrt{y})\,y^{-\frac{1}{2}} \text{ for } y \in (0,1)\\ &= \frac{1}{4}\left(y^{-\frac{1}{2}}+1\right) \text{ for } y \in (0,1)\\ &= 0 \text{ elsewhere}\end{aligned}$$

So adding these two functions together we get

$$\begin{aligned} f_Y(y) &= \frac{1}{4}\left(y^{-\frac{1}{2}}-1\right)+\frac{1}{4}\left(y^{-\frac{1}{2}}+1\right) \text{ for } y \in (0,1) \\ &= \frac{1}{2}y^{-\frac{1}{2}} \text{ for } y \in (0,1) \\ &= 0 \text{ elsewhere} \end{aligned}$$

b. To get the moment generating function of Y, we find $E\left(e^{tY}\right)$.

$$\begin{aligned} M_Y(t) &= E\left(e^{tY}\right) \\ &= \int_0^1 e^{ty}\frac{1}{2}y^{-\frac{1}{2}}\partial y \end{aligned}$$

which does not have a closed form.

c. We can still use the MGF to get $E(Y)$:

$$\begin{aligned} E(Y) &= \frac{\partial}{\partial t}M_Y(t)\Big|_{t=0} \\ &= \int_0^1 \frac{\partial}{\partial t}\left(e^{ty}\frac{1}{2}y^{-\frac{1}{2}}\right)\Big|_{t=0}\partial y \\ &= \frac{1}{2}\int_0^1\left(ye^{ty}y^{-\frac{1}{2}}\right)\Big|_{t=0}\partial y \\ &= \frac{1}{2}\int_0^1 y^{\frac{1}{2}}\partial y \\ &= \frac{1}{3}y^{\frac{3}{2}}\Big|_{y=0}^{1} = \frac{1}{3} \end{aligned}$$

And similarly, we can get $E\left(Y^2\right)$:

$$\begin{aligned} E\left(Y^2\right) &= \frac{\partial^2}{\partial t^2}M_Y(t)\Big|_{t=0} \\ &= \int_0^1 \frac{\partial^2}{\partial t^2}\left(e^{ty}\frac{1}{2}y^{-\frac{1}{2}}\right)\Big|_{t=0}\partial y \\ &= \frac{1}{2}\int_0^1\left(e^{ty}y^{\frac{3}{2}}\right)\Big|_{t=0}\partial y \\ &= \frac{1}{2}\int_0^1 y^{\frac{3}{2}}\partial y \\ &= \frac{1}{5}y^{\frac{5}{2}}\Big|_{y=0}^{1} = \frac{1}{5} \end{aligned}$$

So $Var(Y) = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}$.

d. We start by finding the pdf of $Y|X \geq \frac{1}{2}$. This will be the similar to part a., except that our transformation function is monotonic over the relevant range, and we will use the pdf of $X|X \geq \frac{1}{2}$, which is just the pdf of X over the relevant range scaled up by one over the probability that X is in this range:

$$\begin{aligned} f_{X|X\geq\frac{1}{2}}(x) &= \frac{\frac{1}{2}(x+1)}{\int_{\frac{1}{2}}^{1}\frac{1}{2}(x+1)\,dx} \text{ for } \frac{1}{2} \leq x \leq 1 \text{ (and 0 elsewhere)} \\ &= \frac{x+1}{\left(\frac{x^2}{2}+x\right)\Big|_{x=\frac{1}{2}}^{1}} \\ &= \frac{8}{7}(x+1) \end{aligned}$$

Thus we can calculate

$$\begin{aligned} f_{Y|X\geq\frac{1}{2}}(y) &= f_{X|X\geq\frac{1}{2}}(\sqrt{y})\left|\frac{1}{2}y^{-\frac{1}{2}}\right| \text{ for } y \in \left(\frac{1}{4},1\right) \text{ (and 0 elsewhere)} \\ &= \frac{4}{7}\left(1+y^{-\frac{1}{2}}\right) \end{aligned}$$

It is then straightforward to calculate $E\left(Y|X \geq \frac{1}{2}\right)$:

$$\begin{aligned} E\left(Y|X\geq\frac{1}{2}\right) &= \int_{\frac{1}{4}}^{1}\frac{4y}{7}\left(1+y^{-\frac{1}{2}}\right)dy \\ &= \frac{4}{7}\int_{\frac{1}{4}}^{1} y + y^{\frac{1}{2}}\,dy \\ &= \frac{4}{7}\left(\frac{y^2}{2}+\frac{2}{3}y^{\frac{3}{2}}\right)\Big|_{y=\frac{1}{4}}^{1} \\ &= \frac{101}{168} \approx 0.601 \end{aligned}$$