[Homework] Distributions related to normal distribution

1. Compute the PDF of $Y = \log X$, where $X \sim Exp(1)$.
   Sol.
   $$F_Y(y) = P\{Y \leq y\} = P\{logX \leq y\} = P\{X \leq e^y\} = F_X(e^y)$$
   Hence, on differentiation, we obtain
   $$f_Y(y) = F'_Y(y) = f_X(e^y) \cdot e^y = e^{y-e^y}$$

2. Let $Y_1 = X_1 + X_2$ and let $Y_2 = e^{X_1}$, where X_i are iid $X \sim Exp(1)$.

   (a) Identify the distribution of Y_1 and report its mean and variance.
   Sol. Since
   $$X + Y \sim \Gamma(\alpha = 2, \lambda = 1)$$
   $$EY_1 = \frac{\alpha}{\lambda} = 2$$
   $$VarY_1 = \frac{\alpha}{\lambda^2} = 2$$

   (b) Find the joint PDF $f(y_1, y_2)$ of Y_1 and Y_2.
   Sol.
   $$J = \frac{\partial(Y_1, Y_2)}{\partial(X_1, X_2)} = -e^{x_1}$$
   $$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= f_{X_1}(x)f_{X_2}(x_2) \cdot |J|^{-1} \\ &= \lambda^2 e^{-\lambda(x_1+x_2)} \cdot \frac{1}{e^{x_1}} \\ &= \lambda^2 e^{-\lambda y_1} \cdot \frac{1}{y_2} \\ &= e^{-y_1} \cdot \frac{1}{y_2} \end{aligned}$$

3. Let X, Y, and Z be iid $Exp(1)$. Derive the joint distribution of $U = X + Y$, $V = Y + Z$, $W = X + Z$.
   Sol.
   $$f_X(x) = f_Y(y) = f_Z(z) = e^{-x}$$
   $$x = \frac{u + v - w}{2}, y = \frac{u - v + w}{2}, x = \frac{-u + v + w}{2}$$

$$J = \frac{\partial(u,v,w)}{\partial(x,y,z)} = 2$$

$$\begin{aligned} f_{U,V,W}(u,v,w) &= f_X(x)f_Y(y)f_Z(z)\cdot|J|^{-1} \\ &= e^{-x}\cdot e^{-y}\cdot e^{-z}\frac{1}{2} \\ &= \frac{1}{2}e^{-\frac{u+v+w}{2}} \end{aligned}$$

4. Let X_i be iid $Exp(1)$. Compute

   (a) $P(\min\{X_1,\cdots,X_5\}\le a)$.
   Sol.

$$\begin{aligned} P(\min\{X_1,\cdots,X_5\}\le a) &= 1-P(\min\{X_1,\cdots,X_5\}>a) \\ &= 1-P(X_1>a)\cdots P(X_5>a) \\ &= 1-e^{-5\lambda a} = 1-e^{-5a} \end{aligned}$$

   (b) $P(\max\{X_1,\cdots,X_5\}\le a)$.
   Sol.

$$\begin{aligned} P(\max\{X_1,\cdots,X_5\}\le a) &= P(X_1\le a,\cdots,X_5\le a) \\ &= P(X_1\le a)\cdots P(X_5\le a) \\ &= (1-e^{-\lambda a})^5 = (1-e^{-a})^5 \end{aligned}$$

5. Let U_i be iid $U(0,1)$. Let X_n be the number of U_i, $1\le i\le n$ with $U_i\le\frac{1}{n}$ and Y_n be the number of U_i, $1\le i\le n$ with $\frac{1}{n}\le U_i\le\frac{2}{n}$. As $n\to\infty$ find the joint PMF of X_n and Y_n.
   Sol.

$$P(X_n=x) = \binom{n}{x}(\frac{1}{n})^x(1-\frac{1}{n})^{n-x}$$

$$P(Y_n=y|X_n=x) = \binom{n-x}{y}(\frac{1}{n-1})^y(1-\frac{1}{n-1})^{n-x-y}$$

$$\begin{aligned} P(X_n=x,Y_n=y) &= P(Y_n=y|X_n=x)\cdot P(X_n=x) \\ &= \frac{(n-x)!}{(n-x-y)!y!}\cdot\frac{n!}{(n-x)!x!}(\frac{1}{n})^x(1-\frac{1}{n})^{n-x}\cdot(\frac{1}{n-1})^y(1-\frac{1}{n-1})^{n-x-y} \\ &= \frac{n!}{(n-x-y)!y!x!}\frac{1}{n^x}\frac{1}{(n-1)^y}(1-\frac{1}{n})^n(\frac{n-1}{n})^x(1-\frac{1}{n-1})^{n-1} \\ &\quad\cdot(\frac{n-2}{n-1})^{-x-y+1} \end{aligned}$$

joint PMF of X_n, Y_n as $n \to \infty$

$$f_{X_n,Y_n}(x,y) = \frac{1}{x!y!}e^{-2}$$

6. The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & \text{for } 0 < x < \infty, 0 < y < \infty, \\ 0 & \text{otherwise.} \end{cases}.$$

   Find the PDF of $\frac{X}{Y}$.
   Sol.

$$\begin{aligned} F_{\frac{X}{Y}}(a) &= P(\frac{X}{Y} \leq a) \\ &= P(X \leq aY) \\ &= \int_0^\infty \int_0^{ay} 2e^{-x}e^{-2y} dxdy \\ &= 1 - \frac{2}{2+a} \end{aligned}$$

$$f_{\frac{X}{Y}}(a) = \frac{d}{da}F_{\frac{X}{Y}}(a) = \frac{2}{(2+a)^2}$$

7. Successive weekly sales, in units of one thousand dollars, have a bivariate normal distribution with common mean 40, common standard deviation 6, and correlation 0.6.

   (a) Find the probability that the total of the next 2 weeks' sales exceeds 90.
   Sol.

$$X_1, X_2 \sim N(40, 6^2)$$

$$\rho = \frac{Cov(X,Y)}{\sqrt{VarX \cdot VarY}}$$

   Therefore, $Cov(X_1, X_2) = 21.6$

$$X_1 + X_2 \sim Normal \quad \text{(Why??)}$$

$$E(X_1 + X_2) = EX_1 + EX_2 = 80$$

$$Var(X_1 + X_2) = VarX_1 + VarX_2 + 2Cov(X_1, X_2) = 115.2$$

$$\begin{aligned} P(X_1 + X_2 > 90) &= P(\frac{(X_1 + X_2) - 80}{\sqrt{115.2}} > \frac{90 - 80}{\sqrt{115.2}}) \\ &= P(Z > \frac{10}{\sqrt{115.2}}) \\ &= 1 - \Phi(\frac{10}{\sqrt{115.2}}) \end{aligned}$$

(b) If the correlation were 0.2 rather than 0.6, do you think that this would increase or decrease the answer to (a)? Explain your reasoning.
Sol.

(c) Repeat (a) with the correlation is 0.2 to check your intuition on (b).
Sol.
$Cov(X_1, X_2) = 7.2$

$$E(X_1 + X_2) = EX_1 + EX_2 = 80$$

$$Var(X_1 + X_2) = VarX_1 + VarX_2 + 2Cov(X_1, X_2) = 86.4$$

$$\begin{aligned} P(X_1 + X_2 > 90) &= P(\frac{(X_1 + X_2) - 80}{\sqrt{86.4}} > \frac{90 - 80}{\sqrt{86.4}}) \\ &= P(Z > \frac{10}{\sqrt{86.4}}) \\ &= 1 - \Phi(\frac{10}{\sqrt{86.4}}) \end{aligned}$$

8. The mean and standard deviation of the midterm scores of the probability corse are 71 and 12 and the mean and standard deviation of the final are 70 and 11. The correlation ρ is 0.6. Suppose the joint distribution of the midterm and final scores is the bivariate normal.

(a) Predict the final score for someone who got 83 on midterm.
Sol.

(b) Predict the midterm score for someone who got 76.6 on final.
Sol.

(c) Predict the final percentile for someone who got 31st percentile on midterm.
Sol.

(d) Predict the midterm percentile for someone who got 38th percentile on final.
Sol.

9. The random variables X and Y are described by a joint PDF of the form

$$f_{X,Y}(x,y) = ce^{-8x^2-6xy-18y^2}$$

Find the means, variances, and the correlation coefficient of X and Y. Also, and the value of the constant c.

Sol. Since

$$\int_{\infty}^{\infty}\int_{\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$$

$$\begin{aligned}
\int_{\infty}^{\infty}\int_{\infty}^{\infty} f_{X,Y}(x,y)dxdy &= \int_{\infty}^{\infty} ce^{-\frac{135}{8}y^2}\int_{\infty}^{\infty} e^{-8(x+\frac{3}{8}y)^2}dxdy \\
&= \int_{\infty}^{\infty} ce^{-\frac{135}{8}y^2}\int_{\infty}^{\infty} \frac{1}{4}e^{-\frac{1}{2}u^2}dudy \quad (u = 4(x+\frac{3}{8}y)) \\
&= \int_{\infty}^{\infty} ce^{-\frac{135}{8}y^2}\frac{1}{4}\sqrt{2\pi}dy \\
&= \frac{\sqrt{2\pi}c}{4}\int_{\infty}^{\infty} e^{-\frac{135}{8}y^2}dy \quad (v = \frac{\sqrt{135}}{2}y) \\
&= \frac{\sqrt{2\pi}c}{2\sqrt{135}}\int_{\infty}^{\infty} e^{-\frac{1}{2}v^2}dv \\
&= \frac{\pi}{\sqrt{135}}
\end{aligned} \tag{0.1}$$

Therefore $c = \frac{\sqrt{135}}{\pi}$

Similarly,

$$\rho = -\frac{1}{4}, EX = 0, VarX = \frac{9}{135}, EY = 0, VarY = \frac{4}{135}$$

(Why? Explain!Find bivariate normal distribution.)

10. Let X_1 and X_2 be independent standard normal random variables. Define the random variables Y_1 and Y_2 by

$$Y_1 = 2X_1 + X_2, \quad Y_2 = X_1 - X_2$$

Find $E[Y_1]$, $E[Y_2]$, $Var(Y_1)$, $Var(Y_2)$ $Cov(Y_1, Y_2)$, and the joint PDF f_{Y_1,Y_2}.

Sol. Since $X_1, X_2 \sim N(0, 1^2)$

$$EY_1 = 2EX_1 + EX_2 = 0, EY_2 = 0$$

$$VarY_1 = 4VarX_1 + VarX_2 = 5, VarY_2 = VarX_1 + VarX_2 = 2$$

$$Cov(Y_1, Y_2) = Cov(2X_1 + X_2, X_1 - X_2) = 1$$

$$J = \frac{\partial(Y_1, Y_2)}{\partial(X_1, X_2)} = -3 \text{ (why?)}$$

Since $x_1 = \frac{y_1+y_2}{3}, x_2 = \frac{y_1-2y_2}{3}$

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= f_{X_1,X_2}(x_1, x_2) \cdot |J|^{-1} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{x_1^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x_2^2}{2}} \cdot \frac{1}{3} \\ &= \frac{1}{6\pi} e^{-\frac{1}{2}\{(\frac{y_1+y_2}{3})^2 + (\frac{y_1-2y_2}{3})^2\}} \end{aligned}$$

[Extra]

1. Suppose that X and Y are independent normal random variables with the same variance. Show that $X - Y$ and $X + Y$ are independent.