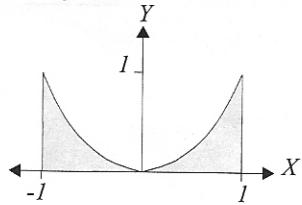


The Reference Solution of Probability Homework #4

助教: 田俊男、詹瑋珊

1. No. We use the 4.7.10 to show the counterexample.

The joint PDF of X and Y and the region of nonzero probability are



$$f_{X,Y}(x,y) = \begin{cases} 5x^2/2 & -1 \leq x \leq 1, 0 \leq y \leq x^2 \\ 0 & \text{otherwise} \end{cases}$$

According the answer of 4.7.10 (c), we know that $\text{Cov}[X, Y] = 0$. So X and Y are uncorrelated by Def. 4.7.

And use Theorem 4.8 to find the marginal PDFs $f_X(x)$ and $f_Y(y)$.

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy \\ &= \int_0^{x^2} \frac{5}{2} x^2 dy \\ &= \frac{5}{2} x^2 y \Big|_0^{x^2} = \frac{5}{2} x^4 \\ f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx \\ &= \int_{-y}^y \frac{5}{2} x^2 dx + \int_{y}^1 \frac{5}{2} x^2 dx \\ &= \frac{5}{6} x^3 \Big|_{-y}^y + \frac{5}{6} x^3 \Big|_y^1 \\ &= -\frac{5}{6} y \sqrt{y} + \frac{5}{6} + \frac{5}{6} - \frac{5}{6} y \sqrt{y} = \frac{5}{3} - \frac{5}{3} y \sqrt{y} \end{aligned}$$

$\therefore f_{X,Y}(x,y) \neq f_X(x) \cdot f_Y(y) \therefore X$ and Y are dependent.

2. According to the appendix 1, we know that $\text{Var}[ax+by] = a^2 \text{Var}[X] + b^2 \text{Var}[Y] + 2ab \text{Cov}[X, Y]$.

$$\therefore \text{Var}\left[\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right] = \frac{1}{\sigma_X^2} \cdot \text{Var}[X] + \frac{1}{\sigma_Y^2} \cdot \text{Var}[Y] - \frac{2}{\sigma_X \sigma_Y} \text{Cov}[X, Y] = 2 - 2\rho_{X,Y}$$

Suppose that $\rho_{X,Y}=1$ and then $\text{Var}\left[\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right] = 0$. So we can know $\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y} = c$ for some constant c .

$$\therefore Y = \frac{\sigma_Y}{\sigma_X} X - C \sigma_Y = ax + b \text{ where } a = \frac{\sigma_Y}{\sigma_X} > 0 \text{ and } b = -C \sigma_Y \#$$

3. According to the Def. 4.6, we know that X and Y are orthogonal if $r_{X,Y} = E[XY] = 0$.

And we know that X and Y are uncorrelated if $\text{Cov}[X, Y] = E[XY] - E[X]E[Y] = 0$ by the Def. 4.7 and Theorem 4.16.

\therefore We can conclude that if $E[X]=0$ or $E[Y]=0$, X and Y are not only orthogonal but also uncorrelated.

4. (Problem 4.1.6) The given function is $F_{X,Y}(x,y) = \begin{cases} 1 - e^{-(x+y)} & x, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$

We find the CDFs $F_X(x)$ and $F_Y(y)$ by Theorem 4.1.

$$F_X(x) = F_{X,Y}(x, \infty) = \begin{cases} 1 & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad F_Y(y) = F_{X,Y}(\infty, y) = \begin{cases} 1 & y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

\therefore for any $x \geq 0$ or $y \geq 0$, $P[X > x] = 0$ or $P[Y > y] = 0$

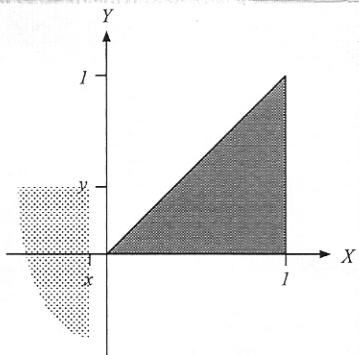
According to Theorem 1.7, we know that $P[\{X > x\} \cup \{Y > y\}] \leq P[X > x] + P[Y > y] = 0$.

However, $P[\{X > x\} \cup \{Y > y\}] = 1 - P[X \leq x, Y \leq y] = 1 - (1 - e^{-(x+y)}) = e^{-(x+y)}$.

Thus, we have the contradiction that $e^{-(x+y)} \leq 0$ for all $x, y \geq 0$. We can conclude that the given function is not a valid CDF.

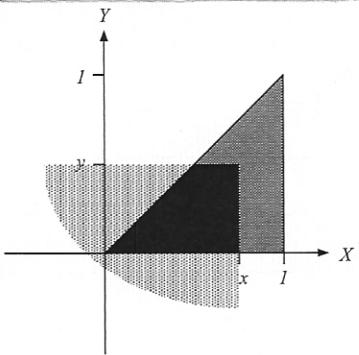
5. (Problem 4.4.4) Just as in Example 4.5, there are five cases. We will use variable u and v as dummy variables for x and y .

case 1 : $x < 0$ or $y < 0$



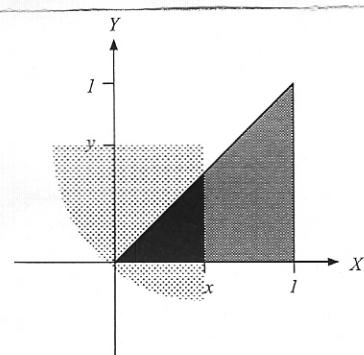
In this case, the region of integration doesn't overlap the region of nonzero probability and $F_{X,Y}(x,y) = \int_{-\infty}^y \int_{-\infty}^x f_{X,Y}(u,v) du dv = 0$

case 2 : $0 < y \leq x \leq 1$



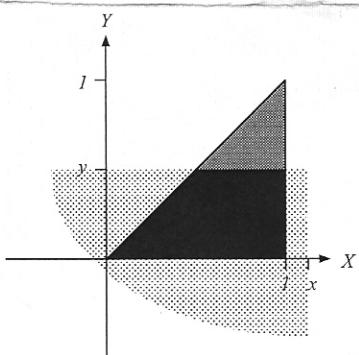
$$\begin{aligned} F_{X,Y}(x,y) &= \int_{-\infty}^y \int_{-\infty}^x f_{X,Y}(u,v) du dv \\ &= \int_0^y \int_v^x 8uv du dv \\ &= \int_0^y 4(x^2 - v^2) v dv \\ &= 2x^2v^2 - v^4 \Big|_{v=0}^{v=y} = 2x^2y^2 - y^4 \end{aligned}$$

case 3 : $0 < x \leq y$ and $0 \leq y \leq 1$



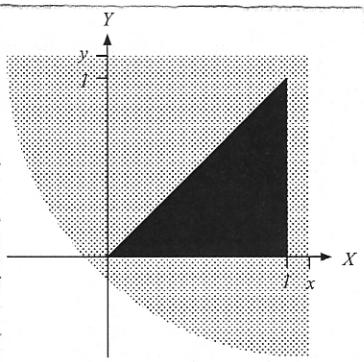
$$\begin{aligned} F_{X,Y}(x,y) &= \int_{-\infty}^y \int_{-\infty}^x f_{X,Y}(u,v) du dv \\ &= \int_0^x \int_0^u 8uv dv du \\ &= \int_0^x 4u^3 du = x^4 \end{aligned}$$

case 4: $0 < y \leq 1$ and $x \geq 1$



$$\begin{aligned} F_{X,Y}(x,y) &= \int_{-\infty}^y \int_{-\infty}^x f_{X,Y}(u,v) du dv \\ &= \int_0^y \int_v^1 8uv du dv \\ &= \int_0^y 4v(1-v^2) dv \\ &= 2y^2 - y^4 \end{aligned}$$

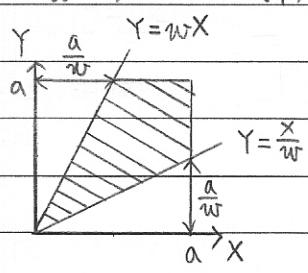
case 5: $x \geq 1$ and $y \geq 1$



In this case, the region of integration completely covers the region of nonzero probability and $F_{X,Y}(x,y) = \int_{-\infty}^y \int_{-\infty}^x f_{X,Y}(u,v) du dv = 1$.

$$\therefore \text{The complete answer for the joint CDF is } F_{X,Y}(x,y) = \begin{cases} 0 & x < 0 \text{ or } y < 0 \\ 2x^2y^2 - y^4 & 0 < y \leq x \leq 1 \\ x^4 & 0 \leq x \leq y, 0 \leq y \leq 1 \\ 2y^2 - y^4 & 0 \leq y \leq 1, x \geq 1 \\ 1 & x \geq 1, y \geq 1 \end{cases}$$

6. (Problem 4.6.11) Following the hint, we observe that either $Y \geq X$ or $X \geq Y$. In other words, $(\frac{Y}{X}) \geq 1$ or $(\frac{X}{Y}) \geq 1$, so $W = \max(\frac{Y}{X}, \frac{X}{Y}) \geq 1$. To find the CDF $F_W(w)$, we know that $F_W(w) = 0$ for $w < 1$. For $w \geq 1$, we solve $F_W(w) = P[\max(\frac{Y}{X}, \frac{X}{Y}) \leq w] = P[\frac{Y}{X} \leq w, \frac{X}{Y} \leq w] = P[Y \geq \frac{X}{w}, Y \leq wX] = P[\frac{X}{w} \leq Y \leq wX]$



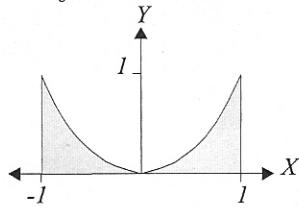
We note that in the middle of the above steps, nonnegativity of X and Y was essential. We can depict the given set $\{\frac{X}{w} \leq Y \leq wX\}$ as the back slash region on the X, Y plane. Because the PDF is uniform over the square, it is easier to use geometry to calculate the probability. In particular, each of the lighter triangles that are not part of the region of interest has area $a^2/2w$.

$$\therefore P\left[\frac{X}{W} \leq Y \leq wX\right] = 1 - \left(\frac{\alpha^2}{2w} \times \frac{1}{\alpha^2}\right) \times 2 = 1 - \frac{1}{w}$$

\therefore The final expression for the CDF of W is $F_W(w) = \begin{cases} 0 & w < 1 \\ 1 - \frac{1}{w} & w \geq 1 \end{cases}$. \therefore The PDF is $f_W(w) = \begin{cases} 0 & w < 1 \\ \frac{1}{w^2} & w \geq 1 \end{cases}$

7. (Problem 4.7.10)

The joint PDF of X and Y and the region of nonzero probability are



$$f_{X,Y}(x,y) = \begin{cases} 5x^2/2 & -1 \leq x \leq 1, 0 \leq y \leq x^2 \\ 0 & \text{otherwise} \end{cases}$$

(a) According to Theorem 4.12, we know that $E[W] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy$ where $W = g(X, Y)$ for random variables X and Y . $\therefore E[X] = \int_{-1}^1 \int_0^{x^2} x \cdot \frac{5}{2} x^2 dy dx = \int_{-1}^1 \frac{5}{2} x^5 dx = \frac{5}{12} x^6 \Big|_{-1}^1 = 0$.

And by Theorem 3.5, we know $\text{Var}[X] = E[X^2] - \text{E}[X]^2$. $\therefore \text{Var}[X] = E[X^2] = \int_{-1}^1 \int_0^{x^2} x^2 \cdot \frac{5}{2} x^2 dy dx = \frac{5}{14} x^7 \Big|_{-1}^1 = \frac{5}{14}$

$$(b) E[Y] = \int_{-1}^1 \int_0^{x^2} y \cdot \frac{5}{2} x^2 dy dx = \int_{-1}^1 \frac{5}{4} x^6 dy dx = \frac{5}{28} x^7 \Big|_{-1}^1 = \frac{5}{14}$$

$$E[Y^2] = \int_{-1}^1 \int_0^{x^2} y^2 \cdot \frac{5}{2} x^2 dy dx = \int_{-1}^1 \frac{5}{6} x^8 dy dx = \frac{5}{54} x^9 \Big|_{-1}^1 = \frac{5}{27}$$

$$\therefore \text{Var}[Y] = E[Y^2] - \text{E}[Y]^2 = \frac{5}{27} - \left(\frac{5}{14}\right)^2 = 0.058$$

(c) Since Theorem 4.16 and $E[X]=0$, $\text{Cov}[X, Y] = E[XY] - E[X]E[Y] = E[XY]$.

$$\therefore \text{Cov}[X, Y] = E[XY] = \int_{-1}^1 \int_0^{x^2} xy \cdot \frac{5}{2} x^2 dy dx = \int_{-1}^1 \frac{5}{4} x^7 dy dx = \frac{5}{32} x^8 \Big|_{-1}^1 = 0$$

(d) According to Theorem 4.14, we know that $E[X+Y] = E[X] + E[Y]$.

$$\therefore E[XY] = E[X] + E[Y] = \frac{5}{14}$$

$$(e) \text{By Theorem 4.15, } \text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y] + 2 \text{Cov}[X, Y] = \frac{5}{14} + 0.058 = 0.712$$

(a)

8. (Problem 4.9.12) By Theorem 4.8, the marginal PDF of X is $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy = 2 \int_0^{\sqrt{r^2-x^2}} \frac{1}{\pi r^2} dy = \frac{2\sqrt{r^2-x^2}}{\pi r^2}$

$$\therefore f_X(x) = \begin{cases} \frac{2\sqrt{r^2-x^2}}{\pi r^2} & -r \leq x \leq r \\ 0 & \text{otherwise} \end{cases} \therefore \text{The conditional PDF of } Y \text{ given } X \text{ is } f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} \frac{1}{2\sqrt{r^2-x^2}} & y \leq \sqrt{r^2-x^2} \\ 0 & \text{otherwise} \end{cases}$$

(b) Given $X=x$, we observe that over the interval $[-\sqrt{r^2-x^2}, \sqrt{r^2-x^2}]$, Y has a uniform PDF.

\therefore the conditional PDF $f_{Y|X}(y|x)$ is symmetric about $y=0$. $\therefore E[Y|X=x] = 0$.

9. (Problem 4.9.14)

(a.) The number of buses, N , must be greater than zero. Also, the number of minutes that pass cannot be less than the number of buses. Thus, $P[N=n, T=t] \geq 0$ for integers n, t satisfying $1 \leq n \leq t$.

(b.) First, we find the joint PMF of N and T by carefully considering the possible sample paths. In particular,

$$P_{N,T}(n,t) = P[A \cap B \cap C] = P[A]P[B]P[C]$$

$A = \{n-1 \text{ buses arrive in the first } t-1 \text{ minutes}\}$

$B = \{\text{none of the first } n-1 \text{ buses are boarded}\}$

$C = \{\text{at time } t \text{ a bus arrives and is boarded}\}$

These events are independent since each trial to board a bus is independent of when the buses arrive.

$$\text{These events have probabilities: } P[A] = \binom{t-1}{n-1} p^{n-1} (1-p)^{t-1-(n-1)},$$

$$P[B] = (1-q)^{n-1}, \quad P[C] = pq$$

$$\text{Consequently, the joint PMF of } N \text{ and } T \text{ is } P_{N,T}(n,t) = \begin{cases} \binom{t-1}{n-1} p^{n-1} (1-p)^{t-n} (1-q)^{n-1} pq & n \geq 1, t \geq n \\ 0 & \text{otherwise} \end{cases}$$

(c.) Considering this experiment, we find that it is boarded with probability q when a bus arrives.

Moreover, the experiment ends when a bus is boarded. By viewing whether each arriving bus is boarded as an independent trial, N is the number of trials until the first success. Thus, N has the geometric PMF

$$P_N(n) = \begin{cases} (1-q)^{n-1} q & n = 1, 2, \dots \\ 0 & \text{otherwise} \end{cases}$$

To find the PMF of T , suppose we regard each minute as an independent trial in which a success occurs if a bus arrives and that bus is boarded. In this case, the success probability is pq and T is the number of minutes up to and including the first success. The PMF of T is also geometric.

$$\therefore P_T(t) = \begin{cases} (1-pq)^{t-1} pq & t = 1, 2, \dots \\ 0 & \text{otherwise} \end{cases}$$

(d.) Once we have the marginal PMFs, the conditional PMFs are easy to find.

$$P_{N|T}(n|t) = \frac{P_{N,T}(n,t)}{P_T(t)} = \begin{cases} \binom{t-1}{n-1} \left(\frac{p(1-q)}{1-pq}\right)^{n-1} \left(\frac{1-p}{1-pq}\right)^{t-n} & n = 1, 2, \dots, t \\ 0 & \text{otherwise} \end{cases}$$

That is, given you depart at time $T=t$, the number of buses that arrive during minutes $1, \dots, t-1$ has a binomial PMF ($P_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$) since in each minute a bus arrives with probability p .

$$\text{Similarly, the conditional PMF of } T \text{ given } N \text{ is } P_{T|N}(t|n) = \frac{P_{N,T}(n,t)}{P_N(n)} = \begin{cases} \frac{(t-1)}{(n-1)} p^n (1-p)^{t-n} & t=n, n+1, \dots \\ 0 & \text{otherwise} \end{cases}$$

This result can be explained. Given that you board bus $N=n$, the time T when you leave is the time for n buses to arrive. If we view each bus arrival as a success of an independent trial, the time for n buses to arrive has the above Pascal PMF ($P_X(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$)

10. (Problem 4.9.15)

If you construct a tree describing what type of call (if any) that arrived in any 1 millisecond period, it will be apparent that a fax call arrives with probability $\alpha = p$ or no fax arrives with probability $1-\alpha$. That is, whether a fax message arrives each millisecond is a Bernoulli trial with success probability α . Thus, the time required for the first success has the geometric PMF $P_T(t) = \begin{cases} (1-\alpha)^{t-1} \alpha & t=1, 2, \dots \\ 0 & \text{otherwise} \end{cases}$

Note that N is the number of trials required to observe 100 successes. Moreover, the number of trials needed to observe 100 successes is $N=T+N'$ where N' is the number of trials needed to observe successes 2 through 100.

Since N' is just the number of trials needed to observe 99 successes, it has the Pascal ($k=99, p$) PMF $P_{N'}(n) = \binom{n-1}{98} \alpha^{99} (1-\alpha)^{n-99}$. Since the trials needed to generate successes 2 through 100 are independent of the trials that yield the first success, N' and T are independent. Hence, $P_{N|T}(n|t) = P_{N'|T}(n-t|t) = P_{N'}(n-t)$.

Applying the PMF of N' found above, we have $P_{N|T}(n|t) = \binom{n-t-1}{98} \alpha^{99} (1-\alpha)^{n-t-99}$

Finally the joint PMF of N and T is $P_{N,T}(n,t) = P_{N|T}(n|t) P_T(t) = \begin{cases} \binom{n-t-1}{98} \alpha^{99} (1-\alpha)^{n-t-99} & t=1, 2, \dots; n=99+t, 100+t, \dots \\ 0 & \text{otherwise} \end{cases}$

This solution can also be found a consideration of the sample sequence of Bernoulli trials in which we either observe or do not observe a fax message.

To find the conditional PMF $P_{T|N}(t|n)$, we first must recognize that N is simply the number of trials needed to observe 100 successes and thus has the Pascal PMF $P_N(n) = \binom{n-1}{99} \alpha^{100} (1-\alpha)^{n-100}$.

Hence for any integer $n \geq 100$, the conditional PMF is $P_{T|N}(t|n) = \frac{P_{N,T}(n,t)}{P_N(n)} = \begin{cases} \frac{\binom{n-t-1}{98}}{\binom{n-1}{99}} & t=1, 2, \dots, n-99 \\ 0 & \text{otherwise} \end{cases}$

11. (Problem 4.10.10)

(a.) According to the table of the problem description, we can get the following joint PMF of N and D .

$P_{N,D}(n,d)$	$d=20$	$d=100$	$d=300$
$n=1$	0.2	0.2	0.2
$n=2$	0.1	0.2	0.1

$$(b.) \text{ By Theorem 4.12, } E[D] = \sum_{n \in N} \sum_{d \in D} d P_{N,D}(n,d) = 20 \times 0.2 + 100 \times 0.2 + 300 \times 0.2 + 20 \times 0.1 + 100 \times 0.2 + 300 \times 0.1 \\ = 4 + 20 + 60 + 2 + 20 + 30 = 136$$

(c.) To find the conditional PMF $P_{D|N}(d|2)$, we first need to find the probability of the conditioning event

$$P_N(2) = P_{N,D}(2,20) + P_{N,D}(2,100) + P_{N,D}(2,300) = 0.4$$

$$\text{The conditional PMF of } D \text{ given } N=2 \text{ is } P_{D|N}(d|2) = \frac{P_{N,D}(2,d)}{P_N(2)} = \begin{cases} \frac{1}{4} & d=20 \\ \frac{1}{2} & d=100 \\ \frac{1}{4} & d=300 \\ 0 & \text{otherwise} \end{cases}$$

(d.) By Theorem 4.23 the conditional expectation of D given $N=2$ is

$$E[D|N=2] = \sum_{d \in D} d P_{D|N}(d|2) = 20 \cdot \frac{1}{4} + 100 \cdot \frac{1}{2} + 300 \cdot \frac{1}{4} = 130$$

(e.) To check independence, we could calculate the marginal PMFs of N and D . In this case, however, it is simpler to observe that $P_D(d) \neq P_{D|N}(d|2)$. Hence N and D are dependent.

$$(f.) \text{ In terms of } N \text{ and } D, \text{ the cost (in cents) of a fax is } C = ND. \text{ By Theorem 4.12, the expected value of } C \text{ is } E[C] = \sum_{n \in N} \sum_{d \in D} nd P_{N,D}(n,d) \\ = 1 \cdot (20) \cdot (0.2) + 1 \cdot (100) \cdot (0.2) + 1 \cdot (300) \cdot (0.2) + 2 \cdot (20) \cdot (0.1) + 2 \cdot (100) \cdot (0.2) + 2 \cdot (300) \cdot (0.1) \\ = 4 + 20 + 60 + 4 + 40 + 60 = 188$$

12. (Problem 4.10.14)

\Leftarrow According to Theorem 4.4, we know that $f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$.

$$\therefore f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} = \frac{\partial^2 F_X(x) F_Y(y)}{\partial x \partial y} = \frac{\partial [F_X(x) \cdot F_Y(y)]}{\partial y} = f_X(x) \cdot f_Y(y), \therefore X \text{ and } Y \text{ are independent by Def. 4.16.}$$

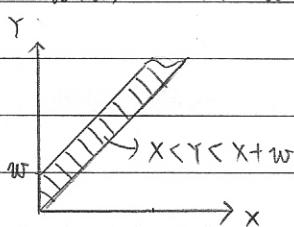
\Rightarrow If X and Y are independent, then $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$.

$$\text{By Def. 4.3, } F_{X,Y}(x,y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u,v) du dv \\ = \left(\int_{-\infty}^x f_X(u) du \right) \left(\int_{-\infty}^y f_Y(v) dv \right) \\ = F_X(x) \cdot F_Y(y)$$

13. (Problem 4.10.15)

The joint PDF of X and Y is $f_{X,Y}(x,y) = \begin{cases} \lambda^2 e^{-\lambda y} & 0 \leq x < y \\ 0 & \text{otherwise} \end{cases}$

For $W = Y - X$ we can find $f_W(w)$ by integrating over the region indicated in the figure below to get $F_W(w)$ then taking the derivative with respect to w . Since $Y \geq X$, $W = Y - X$ is nonnegative. Hence $F_W(w) = 0$ for $w < 0$. For $w \geq 0$, $F_W(w) = 1 - P[W > w] = 1 - P[Y > X + w]$



$$\begin{aligned} F_W(w) &= 1 - \int_0^\infty \int_{x+w}^\infty \lambda^2 e^{-\lambda y} dy dx \\ &= 1 - \int_0^\infty -\lambda e^{-\lambda y} \Big|_{x+w}^\infty dx \\ &= 1 - \int_0^\infty \lambda e^{-\lambda(x+w)} dx \\ &= 1 + e^{-\lambda(x+w)} \Big|_0^\infty = 1 + e^{-\lambda w} \end{aligned}$$

The complete expressions for the joint CDF and corresponding joint PDF are

$$F_W(w) = \begin{cases} 0 & w < 0 \\ 1 - e^{-\lambda w} & w \geq 0 \end{cases} \quad f_W(w) = \begin{cases} 0 & w < 0 \\ \lambda e^{-\lambda w} & w \geq 0 \end{cases}$$

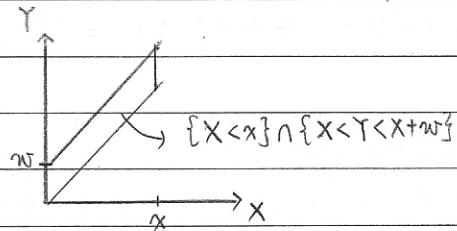
14. (Problem 4.10.16)

(a.) To check if W and X are independent, we must be able to factor the joint density function $f_{X,W}(x,w)$ into the product $f_X(x)f_W(w)$ of marginal density functions. First we must find the joint CDF of X and W and then find the PDF of X and W .

$$F_{X,W}(x,w) = P[X \leq x, W \leq w] = P[X \leq x, Y - X \leq w] = P[X \leq x, Y \leq X + w]$$

Since $Y \geq X$, the CDF of W satisfies $F_{X,W}(x,w) = P[X \leq x, X \leq Y \leq X + w]$. Thus, for $x \geq 0$ and $w \geq 0$,

$$\begin{aligned} F_{X,W}(x,w) &= \int_0^x \int_{x'}^{x+w} \lambda^2 e^{-\lambda y} dy dx' \\ &= \int_0^x -\lambda e^{-\lambda y} \Big|_{x'}^{x+w} dx' \\ &= \int_0^x -\lambda e^{-\lambda(x'+w)} + \lambda e^{-\lambda x'} dx' \\ &= e^{-\lambda(x'+w)} - e^{-\lambda x'} \Big|_0^x \\ &= e^{-\lambda(x+w)} - e^{-\lambda x} - e^{-\lambda w} + e^{-\lambda \cdot 0} \\ &= -e^{-\lambda x} (1 - e^{-\lambda w}) + (1 - e^{-\lambda w}) \\ &= (1 - e^{-\lambda w})(1 - e^{-\lambda x}) \end{aligned}$$



$$\text{By Theorem 4.4, } f_{X,W}(x,w) = \frac{\partial^2 F_{X,W}(x,w)}{\partial x \partial w} = \frac{\partial [1 - e^{-\lambda x} (1 - e^{-\lambda w})]}{\partial w} = \lambda e^{-\lambda x} \lambda e^{-\lambda w} = \lambda^2 e^{-\lambda x} e^{-\lambda w}$$

Second, we must find the marginal PDF $f_X(x)$ and $f_W(w)$. And the $f_W(w) = \lambda e^{-\lambda w}$ when $w \geq 0$ according to the answer of Problem 4.10.15. So we just find the $f_X(x)$ to check this.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,W}(x,w) dw = \int_0^{\infty} \lambda^2 e^{-\lambda x} e^{-\lambda w} dw = -\lambda e^{-\lambda x} e^{-\lambda w} \Big|_0^{\infty} = 0 + \lambda e^{-\lambda x} e^{-\lambda \cdot 0} = \lambda e^{-\lambda x}$$

$$\therefore f_{X,W}(x,w) = \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda w} = f_X(x) \cdot f_W(w) \quad \therefore W \text{ and } X \text{ are independent.}$$

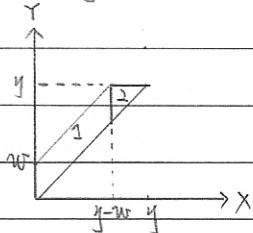
(b.) Following the same procedure, we find the joint CDF of Y and W .

$$F_{Y,W}(y, w) = P[W \leq w, Y \leq y] = P[Y - X \leq w, Y \leq y] = P[Y \leq X + w, Y \leq y]$$

The region of integration corresponding to the event $\{Y \leq X + w, Y \leq y\}$ depends on whether $y < w$ or $y \geq w$.

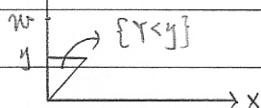
Keep in mind that although $W = Y - X \leq Y$, the dummy arguments u and v of $f_{W,Y}(u,v)$ need not obey the same constraints. In any case, we must consider each case separately.

For $y > w$, the integration is $F_{W,Y}(w, y) = \int_0^{y-w} \int_u^{u+w} \lambda^2 e^{-\lambda v} dv du + \int_{y-w}^y \int_u^y \lambda^2 e^{-\lambda v} dv du$



$$\begin{aligned} &= \int_0^{y-w} -\lambda e^{-\lambda v} \int_u^{u+w} du + \int_{y-w}^y -\lambda e^{-\lambda v} \int_u^y du \\ &= \int_0^{y-w} \lambda e^{-\lambda u} - \lambda e^{-\lambda(u+w)} du + \int_{y-w}^y \lambda e^{-\lambda u} - \lambda e^{-\lambda y} du \\ &= -e^{-\lambda u} + e^{-\lambda(u+w)} \Big|_0^{y-w} + (-e^{-\lambda u} - \lambda e^{-\lambda y}) \Big|_{y-w}^y \\ &= -e^{-\lambda(y-w)} + e^{-\lambda y} + 1 - e^{-\lambda w} - \lambda e^{-\lambda y} - e^{-\lambda y} y + e^{-\lambda(y-w)} + \lambda e^{-\lambda y} (y-w) \\ &= 1 - e^{-\lambda w} - w \lambda e^{-\lambda y} \end{aligned}$$

$$\text{For } y \leq w, F_{W,Y}(w, y) = \int_0^y \int_u^y \lambda^2 e^{-\lambda v} dv du = \int_0^y -\lambda e^{-\lambda v} \Big|_u^y du = \int_0^y \lambda e^{-\lambda u} - \lambda e^{-\lambda y} du = -e^{-\lambda u} - \lambda e^{-\lambda y} \Big|_0^y \\ = -e^{-\lambda y} - \lambda e^{-\lambda y} y + 1 \\ = 1 - (1 + \lambda y) e^{-\lambda y}$$



$$\text{The complete expression for joint CDF is } F_{W,Y}(w, y) = \begin{cases} 1 - e^{-\lambda w} - \lambda w e^{-\lambda y} & 0 \leq w \leq y \\ 1 - (1 + \lambda y) e^{-\lambda y} & 0 \leq y \leq w \\ 0 & \text{otherwise} \end{cases}$$

$$\text{By applying Theorem 4.4, } f_{W,Y}(w, y) = \frac{\partial^2 F_{W,Y}(w, y)}{\partial w \partial y} = \frac{\partial \left[\lambda e^{-\lambda w} - \lambda e^{-\lambda y} \right]}{\partial y} = \lambda^2 e^{-\lambda y}.$$

$$\text{The complete expression for joint PDF is } f_{W,Y}(w, y) = \begin{cases} \lambda^2 e^{-\lambda y} & 0 \leq w \leq y \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{W,Y}(w, y) dw = \int_0^y \lambda^2 e^{-\lambda y} dw = \lambda^2 y e^{-\lambda y} \quad \text{and} \quad f_W(w) = \lambda e^{-\lambda w}$$

$$\therefore f_{W,Y}(w, y) = \lambda^2 e^{-\lambda y} \neq \lambda e^{-\lambda w} \cdot \lambda^2 y e^{-\lambda y} = f_W(w) \cdot f_Y(y) \therefore W \text{ and } Y \text{ are dependent.}$$

15 (Problem 4.10.17)

We need to define the events $A = \{U \leq u\}$ and $B = \{V \leq v\}$. In this case,

$$F_{U,V}(u, v) = P[AB] = P[B] - P[A^c B] = P[V \leq v] - P[U > u, V \leq v] \quad (\because A \cap B = B - (A^c \cap B))$$

Note that $U = \min(X, Y) > u$ if and only if $X > u$ and $Y > u$. In the same way, since $V = \max(X, Y)$, $V \leq v$ if and only if $X \leq v$ and $Y \leq v$. Thus $P[U > u, V \leq v] = P[X > u, Y > u, X \leq v, Y \leq v] = P[u < X \leq v, u < Y \leq v]$.

Thus, the joint CDF of U and V satisfies $F_{U,V}(u, v) = P[V \leq v] - P[U > u, V \leq v]$

$$= P[X \leq v, Y \leq v] - P[u < X \leq v, u < Y \leq v]$$

Since X and Y are independent random variables,

$$\begin{aligned} F_{U,V}(u,v) &= P[X \leq u] P[Y \leq v] - P[u < X \leq v] P[u < Y \leq v] \\ &= F_X(v) F_Y(v) - (F_X(v) - F_X(u))(F_Y(v) - F_Y(u)) \\ &= F_X(v) F_Y(u) + F_X(u) F_Y(v) - F_X(u) F_Y(u) \end{aligned}$$

$$\text{The joint PDF is } f_{U,V} = \frac{\partial^2 F_{U,V}(u,v)}{\partial u \partial v} = \frac{\partial}{\partial v} [F_X(v) f_Y(u) + f_X(u) F_Y(v) - f_X(u) f_Y(u)] \\ = f_X(v) f_Y(u) + f_X(u) f_Y(v)$$

16. (Problem 4.11.6)

The given joint PDF is $f_{X,Y}(x,y) = d e^{-(a^2x^2 + bxy + cy^2)}$. In order to be an example of the bivariate Gaussian PDF given in Def. 4.17, we must let $\mu_1 = \mu_2 = 0$ and have

$$a^2 = 1/2\sigma_x^2(1-\rho^2), \quad b = -\rho/\sigma_x \sigma_y (1-\rho^2), \quad c^2 = 1/2\sigma_y^2(1-\rho^2), \quad \text{and } d = 1/(2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}).$$

We can solve for σ_x and σ_y , yielding $\sigma_x = 1/\sqrt{2(1-\rho^2)}$ and $\sigma_y = 1/\sqrt{2(1-\rho^2)}$.

Plugging these values into the equation for b , it follows that $b = -2ac\rho$, or, equivalently, $\rho = -b/(ac)$.

$$\text{This implies } d^2 = 1/(4\pi^2 \sigma_x^2 \sigma_y^2 (1-\rho^2)) = (1-\rho^2) a^2 c^2 / \pi^2 = \frac{a^2 c^2}{\pi^2} - \frac{b^2}{4\pi^2}$$

Since $|P| \leq 1$, we see that $|b/(ac)| \leq 1$. Further, for any choice of a, b and c that meets this constraint, choosing $d = \frac{1}{\pi} \sqrt{a^2 c^2 - \frac{b^2}{4}}$ yields a valid bivariate Gaussian PDF.

$$\text{Appendix 1. } \text{Var}[aX+bY] = a^2 \text{Var}[X] + b^2 \text{Var}[Y] + 2ab \text{Cov}[X,Y]$$

$$\begin{aligned} \text{pf: } \text{Var}[aX+bY] &= E[(aX+bY) - E[aX+bY]]^2 \\ &= E[(aX+bY) - aE[X] - bE[Y]]^2 \\ &= E[(a(X-E[X]) + b(Y-E[Y]))]^2 \\ &= E[a^2(X-E[X])^2 + b^2(Y-E[Y])^2 + 2ab(X-E[X])(Y-E[Y])] \\ &= a^2 E[(X-E[X])^2] + b^2 E[(Y-E[Y])^2] + 2ab E[(X-E[X])(Y-E[Y])] \\ &= a^2 \text{Var}[X] + b^2 \text{Var}[Y] + 2ab \text{Cov}[X,Y] \# \end{aligned}$$