Statistics HW01 Solution

Chapter 2

Problem 67

(a)

The density function can be obtained by direct differentiation :

$$f(x) = \frac{d}{dx}F(x) = \frac{\beta}{\alpha}\Big(\frac{x}{\alpha}\Big)^{\beta-1} exp\Big(-(\frac{x}{\alpha})^{\beta}\Big) = \frac{\beta}{\alpha^{\beta}}x^{\beta-1}exp\Big(-(\frac{x}{\alpha})^{\beta}\Big),\ x \geq 0,\ \alpha > 0,\ \beta > 0$$

(b)

Let $X = \left(\frac{W}{\alpha}\right)^{\beta}$, *then the support of* X *is* $\Big\{x \in \mathscr{R} \,\Big|\, 0 \leq x < \infty\Big\}, W = \alpha X^{\frac{1}{\beta}}$, *and* $\frac{dw}{dx} = \frac{\alpha}{\beta}x^{\frac{1}{\beta}-1}$
Thus, we have

$$f_X(x) = f_W(\alpha x^{\frac{1}{\beta}})\left|\frac{dw}{dx}\right| = \frac{\beta}{\alpha}\Big(\frac{1}{\alpha}\times\alpha x^{\frac{1}{\beta}}\Big)^{\beta-1} exp\Big\{-\Big(\frac{1}{\alpha}\times\alpha x^{\frac{1}{\beta}}\Big)^{\beta}\Big\}\Big(\frac{\alpha}{\beta}x^{\frac{1}{\beta}-1}\Big) = e^{-x},\ 0 \leq x < \infty$$

This shows $X \sim Exp(1)$. *Note that the pdf specifies a distribution.*

(c)

We may use the *inverse cdf method (TB p.63 Proposition D)* *to do so.*

Theorem : Let $U \sim U[0,1]$, *and let* $X = F^{-1}(U)$. *Then the cdf of* X *is* F.

**Proof :* $Pr(X \leq x) = Pr(F^{-1}(U) \leq x) = Pr(U \leq F(x)) = F(x)$

**FYI :* (*This proof required that* F *is strictly increasing.*
If F *is not strictly increasing, we may redefine* $F^{-1}(u) := inf\Big\{x : F(x) \geq u\Big\}$.)

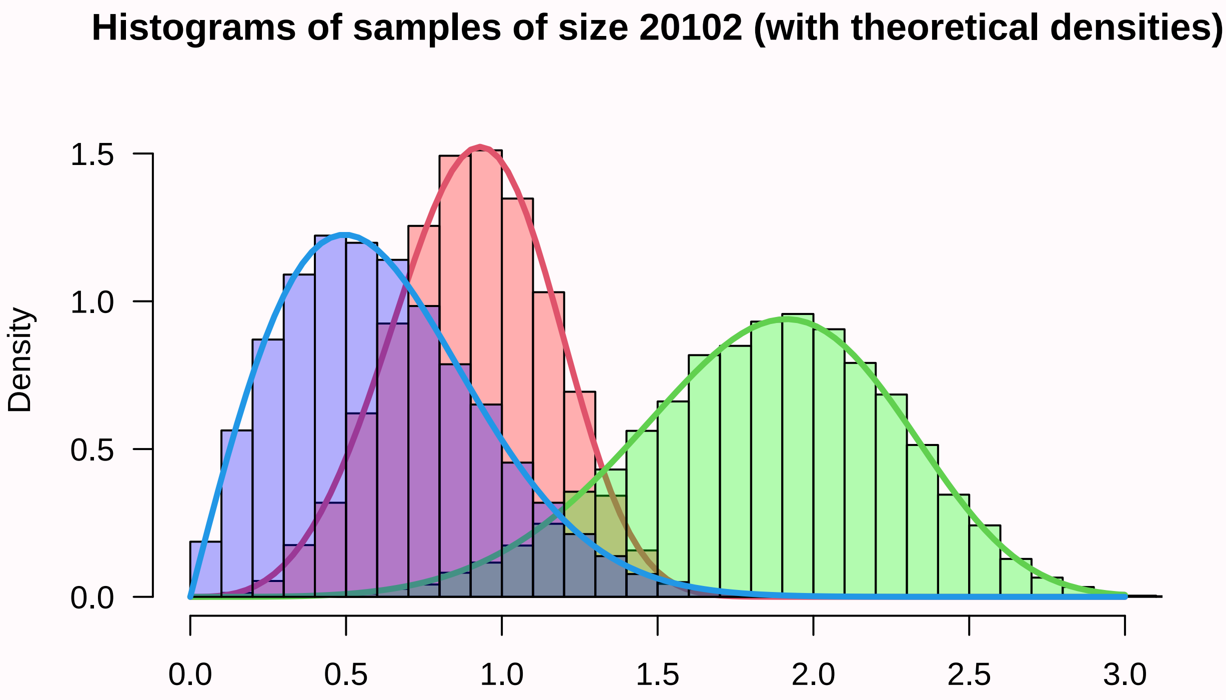
$$Set\ F_X(s) = t = 1 - e^{-(\frac{s}{\alpha})^{\beta}} \implies s = \alpha\Big(-log(1-t)\Big)^{\frac{1}{\beta}} \implies F_X^{-1}(t) = \alpha\Big(-log(1-t)\Big)^{\frac{1}{\beta}}$$

made by 林宸緯，馬翌翔 助教

So we can use the random number generator to generate $U_1, \ldots, U_n \overset{i.i.d.}{\sim} U(0,1)$,
then $F_X^{-1}(U_1), \ldots, F_X^{-1}(U_n)$ *are I.I.D. from the distribution with cdf* F_X $\left(\textbf{Weibull}(\alpha, \beta) - distribution\right)$.

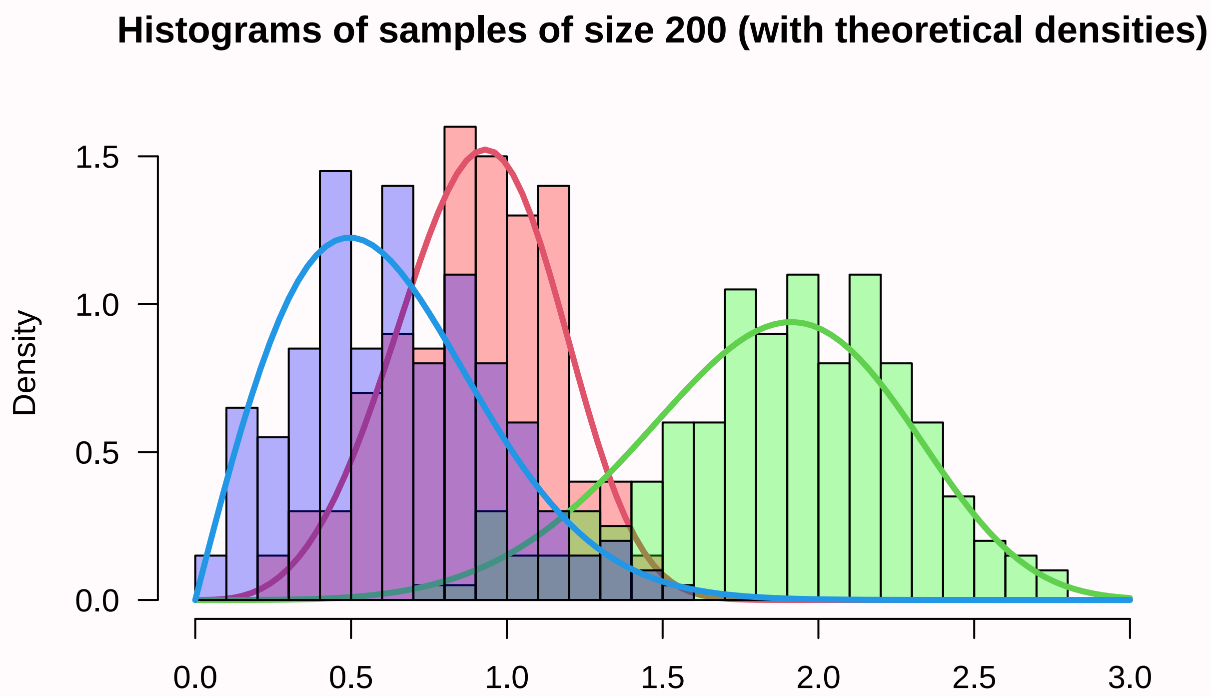
We can use R to verify this method :

```r
set.seed(77) # choose what you like
U <- runif(20102) # so big
S <- 1*(-log(1-U))^(1/4) # alpha = 1, beta = 4
hist(S,breaks = 15,probability = T,xlab="",
     main = "Histograms of samples of size 20102 (with theoretical densities)",las=1,
     col = rgb(1, 0, 0, 0.3),ylim = c(0,1.6),xlim=c(0,3))
curve(dweibull(x,shape = 4,scale = 1),col=2,lwd=3,add=T)
UU <- runif(20102)
SS <- 2*(-log(1-UU))^(1/5) # alpha = 2, beta = 5
hist(SS,breaks = 30,probability = T,main="",add=T, col = rgb(0, 1, 0, 0.3))
curve(dweibull(x,shape = 5,scale = 2),col=3,lwd=3,add=T)
UUU <- runif(20102)
SSS <- (0.7)*(-log(1-UU))^(1/2) # alpha = 0.7, beta = 2
hist(SSS,breaks = 30,probability = T,main="",add=T, col = rgb(0, 0, 1, 0.3))
curve(dweibull(x,shape = 2,scale = 0.7),col=4,lwd=3,add=T)
```



What happens if we choose a small sample size?

(*The code is omitted.*)



What difference do you see? Is this method good?

FYI : There is another method called *rejection sampling*.

■

(*INTENTIONALLY LEFT BLANK*)

Chapter 3

Problem 7

We can find the desired densities by direct differentiation :

$$\begin{cases} f_{X,Y}(x,y) = \dfrac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) = \alpha\beta e^{-\alpha x - \beta y} & , x \geq 0,\ y \geq 0,\ \alpha > 0,\ \beta > 0 \\ f_X(x) = \displaystyle\int_0^\infty f_{X,Y}(x,y)dy = \alpha e^{-\alpha x}\Big(\int_0^\infty \beta e^{-\beta y} dy\Big) = \alpha e^{-\alpha x} & , x \geq 0,\ \alpha > 0 \\ f_Y(y) = \displaystyle\int_0^\infty f_{X,Y}(x,y)dx = \beta e^{-\beta y}\Big(\int_0^\infty \alpha e^{-\alpha x} dx\Big) = \beta e^{-\beta y} & , y \geq 0,\ \beta > 0 \end{cases}$$

Alternative method :

$$\begin{cases} F_X(x) = \lim_{y \to \infty} F_{X,Y}(x,y) = 1 - e^{-\alpha x} \\ F_Y(y) = \lim_{x \to \infty} F_{X,Y}(x,y) = 1 - e^{-\beta y} \end{cases}$$

$$\Longrightarrow \begin{cases} f_X(x) = \dfrac{d}{dx} F_X(x) = \alpha e^{-\alpha x} & , x \geq 0 \\ f_Y(y) = \dfrac{d}{dy} F_Y(y) = \beta e^{-\beta y} & , y \geq 0 \end{cases}$$

$\Longrightarrow$ *This yields the same results!*

■

(*INTENTIONALLY LEFT BLANK*)

Problem 10

(a)

First find the marginal densities :

$$
\begin{aligned}
f_X(x) &= \int_0^\infty x e^{-x(y+1)} dy \\
&= \left(- e^{-x(y+1)} \right)\Big|_0^\infty \\
&= e^{-x}, \ x \in [0, \infty) \\
f_Y(y) &= \int_0^\infty x e^{-x(y+1)} dx \\
&\overset{(\star)}{=} \left(\left(\left(\frac{-x}{y+1} \right) e^{-x(y+1)} \right)\Big|_0^\infty \right) + \left(\frac{1}{y+1} \int_0^\infty e^{-x(y+1)} dx \right) \\
&= \left(\frac{1}{y} \right) \left(\frac{-1}{y+1} e^{-x(y+1)} \right)\Big|_0^\infty \\
&= \frac{1}{(y+1)^2}, \ y \ \in [0, \infty)
\end{aligned}
$$

$(\star)$: *Do integration by parts with* $\begin{pmatrix} u = x & du = dx \\ v = \dfrac{-1}{y+1} e^{-x(y+1)} & dv = e^{-x(y+1)} dx \end{pmatrix}$ *over* $\left\{ x \in \mathscr{R} \,\middle|\, 0 \le x < \infty \right\}$

and recall that $\int_0^\infty u dv = \ uv \Big|_0^\infty - \int_0^\infty v du$

Since $f_{X,Y}(x,y) \neq f_X(x) f_Y(y)$, *we conclude that* X *and* Y *are not independent.*

(b)

With the help of both of the joint density and the marginal densities, we can find :

$$
\begin{cases}
f_{X|Y}(x|y) = \dfrac{f(x,y)}{f_Y(y)} = (y+1)^2 x e^{-x(y+1)} & , \ x \in [0, \infty) \text{ and } y \in [0, \infty) \\
f_{Y|X}(y|x) = \dfrac{f(x,y)}{f_X(x)} = x e^{-xy} & , \ x \in [0, \infty) \text{ and } y \in [0, \infty)
\end{cases}
$$

■

Problem 21

$$Define\ \mathbf{I} := \begin{cases} 1 & , if\ X\ is\ detected. \\ 0 & , otherwise. \end{cases} \implies Pr\Big(\mathbf{I}=1\Big|X=x\Big) = R(x)$$

$(\star)$: By the multiplicative law, the joint distribution of $\Big(X,\mathbf{I}\Big)$ when $\mathbf{I}=1$ is :

$$f_{X,\mathbf{I}}(x,1) = f_X(x) \times Pr\Big(\mathbf{I}=1\Big|X=x\Big) = f(x)R(x)$$

$$\implies G_Y(y) = Pr\Big(Y \le y\Big) = Pr\Big(X \le y \;\Big|\; \mathbf{I}=1\Big) = \frac{Pr(X\le y, \mathbf{I}=1)}{Pr(\mathbf{I}=1)} = \frac{\int_0^y f_{X,\mathbf{I}}(x,1)dx}{\int_{-\infty}^{\infty} f_{X,\mathbf{I}}dx(x,1)} \overset{(\star)}{=} \frac{\int_0^y R(x)f(x)dx}{\int_0^{\infty} R(x)f(x)dx}$$

$$\implies g_Y(y) = \boxed{\frac{d}{dy}G_Y(y) = \frac{R(y)f(y)}{\int_0^{\infty} R(x)f(x)dx}}$$

Holds by the fundamental theorem of calculus.

■

Problem 64

Since X and Y are independent, we have :

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) = \lambda^2 e^{-\lambda(x+y)},\ x>0,\ y>0$$

Now, consider the transformation :

$$Define \begin{cases} V = X+Y \\ W = X/Y \end{cases} \implies \begin{cases} X = VW/(W+1) \\ Y = V/(W+1) \end{cases}$$

This transformation corresponds to a Jacobian $\mathbf{J}$, *which satisfies :*

$$det(\mathbf{J}) = \begin{vmatrix} \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \end{vmatrix} = \frac{-v}{(w+1)^2}$$

So the joint pdf of V and W is :

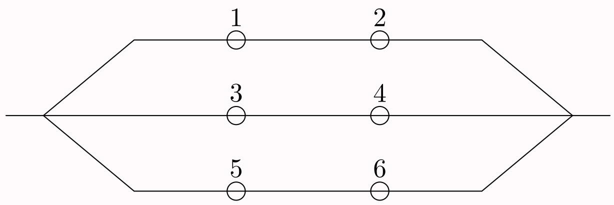
$$f_{V,W}(v,w) = f_{X,Y}\Big(\frac{vw}{w+1}, \frac{v}{w+1}\Big) \times |\,\mathbf{J}\,| = \underbrace{\boxed{\Big(ve^{-\lambda v}\mathbf{I}_{(0,\infty)}(v)\Big)}}_{:=\ g_1(v)} \times \underbrace{\boxed{\Big(\frac{\lambda^2}{(w+1)^2}\mathbf{I}_{(0,\infty)}(w)\Big)}}_{:=\ g_2(w)}$$

Since the joint pdf of (V,W) *can be written as a product of the marginal pdf of V and the marginal pdf of W and their supports are not related,* $(X+Y)$ *and* (X/Y) *are independent.*

■

Problem 66

$First\ label\ the\ points\ as\ the\ following\ graph:$



$Let\ \Big\{T_1, T_2\Big\},\ \Big\{T_3, T_4\Big\}\ and\ \Big\{T_5, T_6\Big\}\ be\ the\ lifetimes\ of\ the\ top\ 2,\ middle\ 2,\ and\ the\ bottom\ 2\ components,\ respectively.$

$Let\ T^*\ be\ the\ lifetime\ of\ the\ system.$

$Follow\ the\ hint,\ we\ know\ that:$

$$T^* = max\Big\{W_1, W_2, W_3\Big\},\ where\ W_1 = min\Big\{T_1, T_2\Big\},\ W_2 = min\Big\{T_3, T_4\Big\},\ W_3 = min\Big\{T_5, T_6\Big\}.$$

$Since\ T_1, \ldots, T_6\ are\ IID,\ W_1, W_2, W_3\ are\ also\ IID,\ hence\ it\ suffices\ to\ find\ the\ distribution\ of\ W_1:$

$$\begin{aligned} Pr\Big(W_1 \le w_1\Big) &= Pr\Big(min\Big\{T_1, T_2\Big\} \le w\Big) \\ &= 1 - Pr\Big(min\Big\{T_1, T_2\Big\} > w\Big) \\ &= 1 - Pr\Big(T_1 > w_1, T_2 > w_1\Big) \\ &= 1 - Pr\Big(T_1 > w_1\Big)Pr\Big(T_2 > w_1\Big) \qquad \boxed{By\ the\ independence\ between\ T_1\ and\ T_2} \\ &= 1 - e^{-2\lambda w_1},\ w_1 > 0 \end{aligned}$$

$Thus,\ we\ have:$

$$(\square)\begin{cases} Pr\Big(W_1 \le w_1\Big) = 1 - e^{-2\lambda w_1} & ,\ w_1 > 0. \\ Pr\Big(W_2 \le w_2\Big) = 1 - e^{-2\lambda w_2} & ,\ w_2 > 0. \\ Pr\Big(W_3 \le w_3\Big) = 1 - e^{-2\lambda w_3} & ,\ w_3 > 0. \end{cases}$$

$The\ cdf\ of\ the\ system's\ lifetime\ is:$

$$\begin{aligned} F_{T^*}(t) &= Pr\Big(T^* \le t\Big) \\ &= Pr\Big(max\Big\{W_1, W_2, W_3\Big\} \le t\Big) \\ &= Pr\Big(W_1 \le t, W_2 \le t, W_3 \le t\Big) \\ &\overset{(\triangle)}{=} Pr\Big(W_1 \le t\Big)Pr\Big(W_2 \le t\Big)Pr\Big(W_3 \le t\Big) \qquad (\triangle):\ \boxed{By\ the\ independence\ of\ W_1,\ W_2\ and\ W_3} \\ &\overset{(\square)}{=} \Big(1 - e^{-2\lambda t}\Big)^3,\ t \ge 0 \end{aligned}$$

The density of the system's lifetime is :

$$f_{T^*}(t) = \frac{d}{dt}F_{T^*}(t) = 6\lambda e^{-2\lambda t}\left(1 - e^{-2\lambda t}\right)^2,\ t \geq 0$$

■

Problem 70

$$\begin{aligned}
F(x,y) =& Pr\left(X_{(1)} \leq x, X_{(n)} \leq y\right) \\
=& Pr\left(X_{(n)} \leq y\right) - Pr\left(X_{(1)} > x, X_{(n)} \leq y\right) \\
=& Pr\left(X_1 \leq y, \dots, X_n \leq y\right) - Pr\left(X_1 \in (x,y], \dots, X_n \in (x,y]\right) \\
=& \prod_{i=1}^{n} F_{X_i}(y) - \prod_{i=1}^{n}\left(F_{X_i}(y) - F_{X_i}(x)\right) \quad \boxed{\text{By the independence between each pair of } X_1, \dots, X_n} \\
=& \left(F(y)\right)^n - \left(F(y) - F(x)\right)^n,\ x \leq y \quad \boxed{\text{By the identically distributed property of } X_1, \dots, X_n}
\end{aligned}$$

■

$\left(\textit{INTENTIONALLY LEFT BLANK}\right)$

Chapter 4

Problem 54

Since X, Y, Z *are uncorrelated, we have* $Cov(X,Y) = Cov(Y,Z) = Cov(Z,X) = 0$, *so*

$$\begin{aligned} & Cov(U,V) \\ &= Cov(Z+X, Z+Y) \\ &= Cov(Z,Z) + \cancel{Cov(Z,Y)} + \cancel{Cov(X,Z)} + \cancel{Cov(X,Y)} \\ &= Cov(Z,Z) = Var(Z) = \sigma_Z^2 \end{aligned}$$

$$\begin{cases} Var(U) = Var(Z+X) & = Var(Z) + Var(X) + 2\cancel{Cov(Z,X)} & = \sigma_Z^2 + \sigma_X^2 \\ Var(V) = Var(Z+Y) & = Var(Z) + Var(Y) + 2\cancel{Cov(Z,Y)} & = \sigma_Z^2 + \sigma_Y^2 \end{cases}$$

$$\implies \rho_{UV} = \frac{Cov(U,V)}{\sqrt{Var(U)Var(V)}} = \frac{\sigma_Z^2}{\sqrt{(\sigma_Z^2+\sigma_X^2)(\sigma_Z^2+\sigma_Y^2)}}$$

■

Problem 60

First note that $E(S) = 1 \times Pr(S=1) + (-1) \times Pr(S=-1) = 1 \times \frac{1}{2} - 1 \times \frac{1}{2} = 0$, *so*

$$\begin{aligned} Cov(X,Y) &= E(XY) - E(X)E(Y) \\ &= \boxed{E(SY^2)} - E(SY)E(Y) \\ &= \boxed{E(S)E(Y^2)} - E(S)E(Y)E(Y) \\ &= 0 \quad \boxed{\text{Since } S \text{ and } Y \text{ are independent.}} \end{aligned}$$

$$\begin{aligned} F_X(x) &= Pr(X \le x) \\ &= Pr(X \le, S=1) + Pr(X \le x, S=-1) \\ &= Pr(X \le x | S=1)Pr(S=1) + Pr(X \le x | S=-1)Pr(S=-1) \\ &= Pr(Y \le x) \times \frac{1}{2} + \boxed{Pr(-Y \le x)} \times \frac{1}{2} \\ &= Pr(Y \le x) \times \frac{1}{2} + \boxed{Pr(Y \le x)} \times \frac{1}{2} \\ &= Pr(Y \le x) \\ &= F_Y(x) \end{aligned}$$

$$\implies f_X(x) = f_Y(x), i.e.^* \ \boxed{X \text{ and } Y \text{ have the same pdf.}}$$

To show $Pr(-Y \leq x) = Pr(Y \leq x)$:

$$\begin{aligned} Pr(-Y \leq x) &= Pr(Y \geq -x) \\ &= \int_{-x}^{\infty} f_Y(y)dy \\ &= -\int_{x}^{-\infty} f_Y(-w)dw \qquad \Big(\text{Let } w = -y\Big) \\ &= \int_{-\infty}^{x} f_Y(w)dw \qquad \left(\begin{array}{l}\text{Since } f_Y \text{ is symmetric about } 0, \\ \text{we have } f_Y(w) = f_Y(-w)\ \forall\ w \in \mathscr{R}.\end{array}\right) \\ &= Pr(Y \leq x) \end{aligned}$$

$$\begin{aligned} F_{X,Y}(x,x) &= Pr(X \leq x, Y \leq x) \\ &= Pr(X \leq x, Y \leq x, S = 1)Pr(S = 1) + Pr(X \leq x, Y \leq x, S = -1)Pr(S = -1) \\ &= \frac{1}{2}Pr(Y \leq x, X \leq x) + \frac{1}{2}Pr(-Y \leq x, X \leq x) \\ &= \frac{1}{2}Pr(Y \leq x) + \frac{1}{2}Pr(-x \leq Y \leq x) \\ &= \frac{1}{2}Pr(Y \leq x) + \frac{1}{2}\Big(Pr(Y \leq x) - Pr(Y \leq -x)\Big) \\ &= Pr(Y \leq x) - \frac{1}{2}Pr(Y \leq -x) \\ &\neq F_X(x)F_Y(y), \text{ in general} \qquad \Longrightarrow \quad X \text{ and } Y \text{ are not independent!} \end{aligned}$$

**Another simpler method* :

The conditional distribution of $\Big(X \,\Big|\, Y\Big)$ *is*

$$f_{X|Y}(x|y) = \begin{cases} Pr\Big(S = 1\Big) = \frac{1}{2} & , \text{ if } x = y \\ Pr\Big(S = -1\Big) = \frac{1}{2} & , \text{ if } x = -y \\ 0 & , \text{ otherwise} \end{cases}$$

Because the conditional distribution of $\Big(X \,\Big|\, Y\Big)$ *is not the marginal distribution of* X $\Big(i.e.,\ f_{X|Y}(x|y) \neq f_X(x)\Big)$, X *and* Y *are not independent!*

Another even simpler method :

Note that $Pr\Big(X \in [-\delta, \delta] \,\Big|\, Y = \delta\Big) = 1 \neq Pr\Big(X \in [-\delta, \delta]\Big)$, *for some suitably chosen* $\delta > 0$.

■

made by 林宸緯，馬翌翔 助教

Problem 61

(a)

Let $U_1, U_2 \overset{i.i.d.}{\sim} U[0,1]$ *and let* $X = min(U_1, U_2),\ Y = max(U_1, U_2),$ *then* $f_{X,Y}(x,y) = 2,\ 0 \le x \le y \le 1$

$$\Longrightarrow \begin{cases} f_X(x) = \int_x^1 2dy = 2(1-x) & ,\ 0 \le x \le 1 \\ f_Y(y) = \int_0^y 2dx = 2y & ,\ 0 \le y \le 1 \end{cases}$$

$$\Longrightarrow \begin{cases} E(X) = \int_0^1 xf_X(x)dx & = \int_0^1 (2x - 2x^2)dx & = \left(x^2 - \frac{2}{3}x^3\right)\Big|_0^1 & = \frac{1}{3} \\ E(Y) = \int_0^1 yf_Y(y)dy & = \int_0^1 2y^2dy & = \left(\frac{2}{3}y^3\right)\Big|_0^1 & = \frac{2}{3} \\ E(X^2) = \int_0^1 x^2 f_X(x)dx & = \int_0^1 \left(2x^2 - 2x^3\right)dx & = \left(\frac{2}{3}x^3 - \frac{1}{2}x^4\right)\Big|_0^1 & = \frac{1}{6} \\ E(Y^2) = \int_0^1 y^2 f_Y(y)dy & = \int_0^1 2y^3dy & = \left(\frac{1}{2}y^4\right)\Big|_0^1 & = \frac{1}{2} \\ E(XY) = \int_0^1 \int_0^y 2xydxdy & = \int_0^1 y^3dy & = \left(\frac{1}{4}y^4\right)\Big|_0^1 & = \frac{1}{4} \end{cases}$$

$$\Longrightarrow \begin{cases} Var(X) = E(X^2) - \left(E(X)^2\right) & = \frac{1}{6} - \left(\frac{1}{3}\right)^2 & = \frac{1}{18} \\ Var(Y) = E(Y^2) - \left(E(Y)^2\right) & = \frac{1}{2} - \left(\frac{2}{3}\right)^2 & = \frac{1}{18} \end{cases}$$

$$\Longrightarrow \begin{cases} Cov(X,Y) = E(XY) - E(X)E(Y) & = \frac{1}{4} - \frac{1}{3} \times \frac{2}{3} & = \frac{1}{36} \\ Cor(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} & = \frac{\frac{1}{36}}{\frac{1}{18}} & = \frac{1}{2} \end{cases}$$

The sign of the correlation makes sense intuitively, because $Y > X$*. So :*

$$\begin{cases} if\ X\ is\ large & ,\ then\ so\ is\ Y. \\ if\ Y\ is\ small & ,\ then\ so\ is\ X. \end{cases}$$

$\left(INTENTIONALLY\ LEFT\ BLANK\right)$

(b)

$$\begin{cases} f_{X|Y}(x|y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)} = \dfrac{1}{y} & , 0 \le x \le y \le 1 \\ f_{Y|X}(y|x) = \dfrac{f_{X,Y}(x,y)}{f_X(x)} = \dfrac{1}{1-x} & , 0 \le x \le y \le 1 \end{cases}$$

$$\Longrightarrow \begin{cases} E(X|Y=y) = \displaystyle\int_0^y \frac{x}{y}dx = \left(\frac{x^2}{2y}\right)\Big|_0^y = \frac{y}{2} & , 0 \le y \le 1 \\ E(Y|X=x) = \displaystyle\int_x^1 \frac{y}{1-x}dx = \left(\frac{y^2}{2(1-x)}\right)\Big|_x^1 = \frac{1+x}{2} & , 0 \le x \le 1 \end{cases}$$

$$\Longrightarrow \boxed{\textit{These results make sense, because we have } \begin{cases} \left(X \,\middle|\, Y=y\right) & \sim U[0,y] \\ \left(Y \,\middle|\, X=x\right) & \sim U[x,1] \end{cases}}$$

Does this match your intuition?

(c)

Fisrt note that $E(X|Y)$ is a function of Y, $E(Y|X)$ is a function of X!

$$Define \begin{cases} Z := E(X|Y) = \boxed{\dfrac{Y}{2}} & \boxed{\textit{By part (b)}} \\ W := E(Y|X) = \boxed{\dfrac{1+X}{2}} & \boxed{\textit{By part (b)}} \end{cases}$$

$$\Longrightarrow \begin{cases} F_Z(z) = Pr(Z \le z) = Pr(Y \le 2z) = \begin{cases} \displaystyle\int_0^{2x} 2ydy = 4z^2 & , 0 \le z < \frac{1}{2} \\ 1 & , \frac{1}{2} \le z \end{cases} \\ F_W(w) = Pr(W \le w) = Pr(X \le 2w-1) = \begin{cases} \displaystyle\int_0^{2w-1} 2(1-x)dx = -4w^2 + 8w - 3 & , \frac{1}{2} \le w < 1 \\ 1 & , 1 \le w \end{cases} \end{cases}$$

$$\Longrightarrow \begin{cases} f_Z(z) = \dfrac{dF_Z(z)}{dz} = \Big(8z\Big) \times \Big(\mathbf{I}_{(0,\frac{1}{2})}(z)\Big) \\ f_W(w) = \dfrac{dF_W(w)}{dw} = \Big(8(1-w)\Big) \times \Big(\mathbf{I}_{(\frac{1}{2},1)}(w)\Big) \end{cases}$$

(d)

To find $\underset{(a,b)}{argmin}\, E\Big(\big(Y-(a+bX)\big)^2\Big) = \underset{(a,b)}{argmin}\, \mathscr{G}(a,b)$, *where*

$$\mathscr{G}(a,b) := E\Big((Y-\hat{Y})^2\Big) = E\Big(\big(Y-(a+bX)\big)^2\Big) = E\Big(Y^2 - 2aY - 2bXY + a^2 + b^2X^2 + 2abX\Big)$$

$$Set \begin{cases} \dfrac{\partial}{\partial a}\mathscr{G}(a,b) = 0 \\ \dfrac{\partial}{\partial b}\mathscr{G}(a,b) = 0 \end{cases} \Longrightarrow \begin{cases} a = E(Y) - bE(X) = \dfrac{1}{2} \\ b = -2E(XY) + 2bE(X^2) + 2E(Y)E(X) - 2b\Big(E(X)\Big)^2 = 0 \end{cases}$$

$$\Longrightarrow\ b\Big(E(X^2) - \big(E(X)\big)^2\Big) = E(XY) - E(X)E(Y) \Longrightarrow\ b = \frac{Cov(X,Y)}{Var(X)} = \frac{1}{2}$$

So $\hat{Y} = \frac{1+X}{2}$ *is the linear predictor of* Y *in terms of* X *that has minimal MSE,*
and its corresponding mean square prediction error is

$$\begin{aligned} E\Big(\big(Y - \frac{1+X}{2}\big)^2\Big) &= Var\Big(Y - \frac{1+X}{2}\Big) + \Big(E\Big(Y - \frac{1+X}{2}\Big)\Big)^2 \\ &= Var(Y) + Var\Big(\frac{1+X}{2}\Big) - 2Cov\Big(Y, \frac{1+X}{2}\Big) + \cancel{\Big(\frac{2}{3} - \frac{1+\frac{1}{3}}{2}\Big)^2} \\ &= Var(Y) + Var\Big(\frac{X}{2}\Big) - 2Cov\Big(Y, \frac{X}{2}\Big) \\ &= Var(Y) + \Big(\frac{1}{2}\Big)^2 Var(X) - 2Cov\Big(Y, X\Big) \times \frac{1}{2} \\ &= \frac{1}{18} + \frac{1}{4} \times \frac{1}{18} - 2 \times \frac{1}{36} \times \frac{1}{2} \\ &= \frac{1}{24} \end{aligned}$$

(d*)

We can also use [*Example 3.4 at LN p.2 − 54*] *to solve this problem.*

> *BEST linear prediction of Y using X :*
>
> $$E_{X,Y}\Big[Y - (a+bX)\Big]^2 \ge E_{X,Y}\Big\{Y - \Big[\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}\Big(X - \mu_X\Big)\Big]\Big\}^2 = \sigma_Y^2(1-\rho^2)$$
>
> *The equality holds if and only if* $a = \mu_Y - b\mu_X$ *and* $b = \rho\frac{\sigma_Y}{\sigma_X}$.

Plug $\mu_X = \frac{1}{3}, \mu_Y = \frac{2}{3}, \sigma_X = \sqrt{\frac{1}{18}}, \sigma_Y = \sqrt{\frac{1}{18}}, \rho = \frac{1}{2}$ *into the above equality.*

Then we also get $\hat{Y} = \frac{1+X}{2}$ *and the* $PMSE = \frac{1}{24}$ *can be calculated in the same way.*

made by 林宸緯，馬翌翔 助教

(e)

$To\ find\ \underset{h(X)}{argmin}\, E\Big(\big(Y-h(X)\big)^2\Big),\ consider\ the\ following\ decomposition:$

$$
\begin{aligned}
&E\Big\{\big(Y-h(X)\big)^2\Big\}\\
&= E\Big\{E\Big(\big(Y-h(X)\big)^2\Big|X\Big)\Big\}\\
&= E\Big\{E\Big(\big(Y-E(Y|X)+E(Y|X)-h(X)\big)^2\Big|X\Big)\Big\}\\
&= E\Big\{E\Big(\big(Y-E(Y|X)\big)^2+\big(E(Y|X)-h(X)\big)^2+2\big(Y-E(Y|X)\big)\big(E(Y|X)-h(X)\big)\Big|X\Big)\Big\}\\
&= E\Big\{E\Big(\big(Y-E(Y|X)\big)^2\Big|X\Big)\Big\}+E\Big\{E\Big(\big(E(Y|X)-h(X)\big)^2\Big|X\Big)\Big\}\\
&\qquad \boxed{+2E\Big\{\big(E(Y|X)-h(X)\big)\times\Big(E\Big(\big(Y-E(Y|X)\big)\Big|X\Big)\Big)\Big\}}\ \boxed{=0}\\
&= E\Big\{Var\Big(Y\ \Big|\ X\Big)\Big\}+\boxed{E\Big\{E\Big(\big(E(Y|X)-h(X)\big)^2\Big|X\Big)\Big\}}\ \boxed{\geq 0,\ equality\ holds\iff h(X)=E(Y|X).}
\end{aligned}
$$

$$\Longrightarrow\ h\big(X\big)=E\big(Y\big|X\big)\ =\frac{1+X}{2}=\underset{h(X)}{argmin}\, E\Big(\big(Y-h(X)\big)^2\Big),\ PMSE\ of\ \frac{1+X}{2}\ is\ \frac{1}{24}.$$

(e*)

$We\ can\ also\ use$ $\boxed{Example\ 3.3\ at\ LN\ p.2-54}$ $to\ solve\ this\ problem.$

$\underline{BEST\ prediction\ of\ Y\ using\ X:}$

$$E_{X,Y}\Big[Y-h(X)\Big]^2\geq E_{X,Y}\Big[Y-E_{X,Y}(Y|X)\Big]^2=E_X\Big[Var_{Y|X}(Y|X)\Big]$$

$The\ equality\ holds\ if\ and\ only\ if\ h(x)=E_{Y|X}(Y|x).$

$By\ the\ result\ of\ part\ (b),\ the\ equality\ holds\ if\ and\ only\ if\ h(x)=\dfrac{1+x}{2}.$

$So\ the\ desired\ predictor\ is\ \hat{Y}=\dfrac{1+X}{2},\ and\ the\ PMSE\ is\ also\ \dfrac{1}{24}.$

■

(supplementary information)

FYI :

The $\hat{Y}$ in problem 61 should be pronounced as "Y hat".
In statistics, we usually put a hat on a parameter to represent a statistic, which is used to estimate that (unknown but fixed) parameter.

$\hat{Y}$ in this problem is used to estimate μ_Y.

Problem 67

We may use the $\overbrace{\text{law of total expectation}}^{(\star)}$ and $\overbrace{\text{the following equation}}^{(\star\star)}$ to find the desired quantities.

To show $\overbrace{E(XY) = E\Big(XE(Y|X)\Big) = E\Big(E(XY|X)\Big)}^{(\star\star)}$:

$$\begin{aligned} E(XY) &= \int\int xyf(x,y)dxdy \\ &= \int\int xyf_X(x)f_{Y|X}(y|x)dxdy \\ &= \boxed{\int x\Big(\int yf_{Y|X}(y|x)dy\Big)f_X(x)dx} \qquad \boxed{E\Big(XE(Y|X)\Big)} \\ &= \boxed{\int \Big(\int xyf_{Y|X}(y|x)dy\Big)f_X(x)dx} \qquad \boxed{E\Big(E(XY|X)\Big)} \end{aligned}$$

$$\text{Also note that} \begin{cases} X \sim U(0,1) \implies E(X) = \int_0^1 xdx = \Big(\frac{x^2}{2}\Big)\Big|_0^1 = \frac{1}{2} \\ \Big(Y\Big|X\Big) \sim U(0,X) \implies E(Y|X=x) = \int_0^x \frac{y}{x}dy = \Big(\frac{y^2}{2x}\Big)\Big|_0^x = \frac{x}{2} \end{cases}$$

The expected circumference is

$$E(2X+2Y) = 2E(X) + 2E(Y) \overset{(\star)}{=} 1 + 2E\Big(E(Y|X)\Big) = 1 + 2\,E\Big(\frac{X}{2}\Big) = 1 + E(X) = \frac{3}{2}$$

The expected area is

$$E(XY) \overset{(\star\star)}{=} E\Big(E(XY|X)\Big) \overset{(\star\star)}{=} E\Big(XE(Y|X)\Big) = E\Big(\frac{X^2}{2}\Big) = \int_0^1 \frac{x^2}{2}dx = \Big(\frac{x^3}{6}\Big)\Big|_0^1 = \frac{1}{6}$$

■

Problem 75

$$\boxed{Suppose\ T\ \sim\ Exp(\lambda)\ and\ E(T) = \frac{1}{\lambda}.}$$

$We\ may\ use\ the\ \overbrace{law\ of\ total\ expectation}^{(\star)}\ and\ the\ \overbrace{law\ of\ total\ variance}^{(\star\star)}\ to\ find\ E(U)\ and\ Var(U),$
$to\ this\ end,\ we\ need\ to\ find\ E(U|T)\ and\ Var(U|T):$

$$\Big(U|T\Big) \sim Uniform\Big(0, T\Big) \implies \left\{ \begin{array}{llll} E(U|T=t) = \displaystyle\int_0^t \frac{u}{t} du & = \left(\dfrac{u^2}{2t}\right)\Bigg|_0^t & = \dfrac{t}{2} \\ E(U^2|T=t) = \displaystyle\int_0^t \frac{u^2}{t} du & = \left(\dfrac{u^3}{3t}\right)\Bigg|_0^t & = \dfrac{t^2}{3} \end{array} \right\}$$

$$\implies Var(U|T) = E(U^2|T=t) - \Big(E(U|T=t)\Big)^2 = \frac{T^2}{3} - \frac{T^2}{4} = \frac{T^2}{12}$$

$So\ we\ have:$

$$\left\{ \begin{array}{l} E(U) \overset{(\star)}{=} E\Big(E(U|T)\Big) = E\left(\dfrac{T}{2}\right) = \dfrac{1}{2\lambda} \\ Var(U) \overset{(\star\star)}{=} E\Big(Var(U|T)\Big) + Var\Big(E(U|T)\Big) = E\left(\dfrac{T^2}{12}\right) + Var\left(\dfrac{T}{2}\right) = \dfrac{5}{12\lambda^2} \end{array} \right.$$

■