[SP-9]

1. Beta and Gamma Distribution.

1.

Let $X = 1 - B$ where $B \sim \text{Beta}(a, b)$

(a).

$$f_X(x) = f_B(t) \left| \frac{\partial t}{\partial x} \right|$$
$$= f_B(t) |-1|$$
$$= c\, t^{a-1} (1-t)^{b-1}$$
$$= c\, x^{b-1} (1-x)^{a-1}$$

$\Rightarrow 1 - B \sim \text{Beta}(b, a)$

(b). In Bank-Post Office Story,

$X \sim \text{Gamma}(a, \lambda)$, $Y \sim \text{Gamma}(b, \lambda)$

and X, Y are independent.

We know

$$\frac{X}{X+Y} \sim \text{Beta}(a, b).$$

Then,

$$1 - \frac{X}{X+Y} = \frac{Y}{X+Y} \sim \text{Beta}(b, a)$$

The result makes sense because

$p \sim \text{Beta}(a, b)$ and

$\begin{cases} a = (\#\text{ prior success}) \\ b = (\#\text{ prior failures}) \end{cases}$

$1-p$ is the prob. of failures and we consider failure as "new success"

$\Rightarrow$ it makes sense that $1 - p \sim \text{Beta}(b, a)$.

2.

$X \sim \text{Gamma}(a, \lambda)$, $Y \sim \text{Gamma}(b, \lambda)$
with a, b integers. X, Y are indep.
<u>WTS</u>: $X+Y \sim \text{Gamma}(a+b, \lambda)$.

(1). By convolution.
Let $T = X+Y$.

$$P(T \le t) = \int_x P(X+Y \le t \mid X=x) f_X(x)\, dx$$
$$= \int_x F_Y(t-x) f_X(x)\, dx$$

$$\Rightarrow f_T(t) = \int_x f_Y(t-x) f_X(x)\, dx$$
$$= \int_0^t \frac{1}{\Gamma(b)} [\lambda(t-x)]^b e^{-\lambda(t-x)} \frac{1}{t-x} \cdot \frac{1}{\Gamma(a)} (\lambda x)^a e^{-\lambda x} \frac{1}{x}\, dx$$
$$= \frac{1}{\Gamma(a)\Gamma(b)} \lambda^{a+b} e^{-\lambda t} \underline{\int_0^t x^{a-1} (t-x)^{b-1}\, dx}$$

{ Let $u = \frac{x}{t} \in (0,1)$. Then,

$$\int_0^t x^{a-1} (t-x)^{b-1}\, dx$$
$$= \int_0^1 (tu)^{a-1} [t(1-u)]^{b-1}\, t\, du$$
$$= t^{a+b-1} \int_0^1 u^{a-1} (1-u)^{b-1}\, du$$
$$= t^{a+b-1} \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$ }

$$\Rightarrow f_T(t) = \frac{1}{\Gamma(a+b)} (\lambda t)^{a+b} e^{-\lambda t} \frac{1}{t}.$$

$\Rightarrow T \sim \text{Gamma}(a+b, \lambda)$ ✓

(2). By MGFs.

The MGF of Gamma(a, λ) is

$$M_X(t) = E[e^{tX}]$$

$$= \int_0^\infty e^{tx} \frac{1}{\Gamma(a)} (\lambda x)^a e^{-\lambda x} \frac{1}{x} dx$$

$$= \int_0^\infty \frac{1}{\Gamma(a)} (\lambda x)^a e^{-(\lambda-t)x} \frac{1}{x} dx$$

$$= \left(\frac{\lambda}{\lambda-t}\right)^a \int_0^\infty \frac{1}{\Gamma(a)} [(\lambda-t)x]^a e^{-(\lambda-t)x} \frac{1}{x} dx$$

$$= \left(\frac{\lambda}{\lambda-t}\right)^a$$

Similarly, $M_Y(t) = \left(\frac{\lambda}{\lambda-t}\right)^b$.

Since, X, Y are indep.,

$$M_{X+Y}(t) = M_X(t) M_Y(t)$$

$$= \left(\frac{\lambda}{\lambda-t}\right)^{a+b}$$

$\Rightarrow X+Y \sim \text{Gamma}(a+b)$.

(3). By story.

Since a, b are integers,

X is sum of a i.i.d. Expo(λ),

Y is sum of b i.i.d. Expo(λ).

Therefore,

$X+Y$ is sum of $a+b$ i.i.d. Expo(λ)

$\Rightarrow X+Y \sim \text{Gamma}(a+b, \lambda)$.

3.

Let $T = X+Y$, $W = \frac{X}{Y}$.

$$\begin{cases} t = x+y \\ w = x/y \end{cases} \Rightarrow \begin{cases} x = \frac{w}{1+w} t \\ y = \frac{1}{1+w} t \end{cases}.$$

The Jacobian matrix is

$$J = \frac{\partial(x,y)}{\partial(t,w)} = \begin{pmatrix} \frac{w}{1+w} & \frac{t}{(1+w)^2} \\ \frac{1}{1+w} & -\frac{t}{(1+w)^2} \end{pmatrix}$$

$$\Rightarrow |J| = -\frac{t}{(1+w)^2}$$

$$\Rightarrow f_{T,W}(t,w) = f_{X,Y}(x,y) \left| -\frac{t}{(1+w)^2} \right|$$

$$= \left(\frac{1}{\Gamma(a)} (\lambda x)^a e^{-\lambda x} \frac{1}{x}\right) \left(\frac{1}{\Gamma(b)} (\lambda y)^b e^{-\lambda y} \frac{1}{y}\right) \frac{t}{(1+w)^2}$$

$$= \left(\frac{1}{\Gamma(a+b)} (\lambda t)^{a+b} e^{-\lambda t} \frac{1}{t}\right) \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \left(\frac{w}{1+w}\right)^{a-1} \left(\frac{1}{1+w}\right)^{b+1}\right)$$

$\Rightarrow X+Y$, X/Y are indep. ✓

{ A neater way is:

We already know $X+Y$ and $W = \frac{X}{X+Y}$ are independent $\Rightarrow$ $X+Y$ and any fcn of W are indep. Since

$$\frac{X}{Y} = \frac{\frac{X}{X+Y}}{\frac{Y}{X+Y}} = \frac{W}{1-W}$$

$\Rightarrow X+Y$ and X/Y are indep. }

4. Find distribution of $mV/(n+mV)$ for $V \sim F(m,n)$

Given $X \sim \text{Gamma}(\frac{m}{2}, \frac{1}{2})$

$Y \sim \text{Gamma}(\frac{n}{2}, \frac{1}{2})$

$$V = \frac{X/m}{Y/n} = \frac{nX}{mY}.$$

$$\frac{mV}{n+mV} = \frac{\frac{nX}{Y}}{n+\frac{nX}{Y}} = \frac{X}{X+Y} \sim \text{Beta}(\frac{m}{2}, \frac{n}{2})$$

{We need that X, Y are independent}

II. Order Statistics

1.

$$f_{U_{(j)}}(x) = n\binom{n-1}{j-1} x^{j-1}(1-x)^{n-j}$$

$$= \frac{\Gamma(n+1)}{\Gamma(j)\Gamma(n-j+1)} x^{j-1}(1-x)^{n-j}$$

$\Rightarrow U_{(j)} \sim \text{Beta}(j, n-j+1)$

$\Rightarrow E[U_{(j)}] = \frac{j}{n+1}$

$$E[U_{(j)}^2] = \int_0^1 \frac{\Gamma(n+1)}{\Gamma(j)\Gamma(n-j+1)} x^{j+1}(1-x)^{n-j}\,dx$$

$$= \frac{(j+1)j}{(n+2)(n+1)}$$

$$\Rightarrow \text{Var}(U_{(j)}) = E[U_{(j)}^2] - (E[U_{(j)}])^2$$

$$= \frac{j(n-j+1)}{(n+2)(n+1)^2}.$$

2.

Given $X, Y \overset{i.i.d.}{\sim} \text{Expo}(\lambda)$, $M = \max(X, Y)$.

<u>WTS</u>: M and $X + \frac{1}{2}Y$ have the same distribution.

(1). $$P(M \le m) = P(X \le m, Y \le m)$$
$$= P(X \le m) P(Y \le m)$$
$$= (1 - e^{-\lambda m})^2$$

Let $T = X + \frac{1}{2}Y$.

$$P(T \le t) = \int P(X \le t - \tfrac{1}{2}y \mid Y = y) f_Y(y) \, dy$$

$$= \int P(X \le t - \tfrac{1}{2}y) f_Y(y) \, dy$$

$\frac{1}{2}y \le t \downarrow$
$$= \int_0^{2t} \left(1 - e^{-\lambda(t - \frac{1}{2}y)}\right) \lambda e^{-\lambda y} \, dy$$

$$= \int_0^{2t} \lambda e^{-\lambda y} \, dy - \int_0^{2t} \lambda e^{-\lambda(t + \frac{1}{2}y)} \, dy$$

$$= -e^{-\lambda y} \Big|_0^{2t} - e^{-\lambda t} \left[-2 e^{-\lambda \frac{1}{2} y} \Big|_0^{2t} \right]$$

$$= 1 - 2e^{-\lambda t} + e^{-2\lambda t}$$

$$= (1 - e^{-\lambda t})^2$$

$\Rightarrow$ M has the same distribution as $X + \frac{1}{2}Y$.

(2). { By using memoryless and other properties of Expo. }

The story is that 2 students start to do their homeworks at the same time, i.e., 1p.m. Each student takes an Exponential time with mean $\frac{1}{\lambda}$, represented by X, Y. $M = \max(X, Y)$ is the time when both students finish their homeworks. We can rewrite

$$M = L + (M - L)$$

where $L = \min(X, Y)$.

By memoryless property, $M - L$ is indep. of L. Moreover,

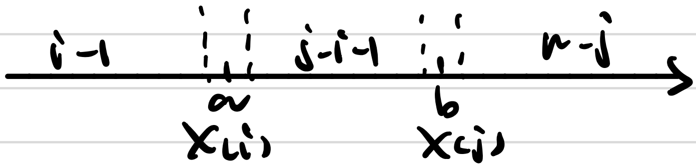
$L \sim \text{Expo}(2\lambda)$, $M - L \sim \text{Expo}(\lambda)$

On the other hand, $\frac{1}{2}Y \sim \text{Expo}(2\lambda)$, $X \sim \text{Expo}(\lambda)$, $X, \frac{1}{2}Y$ are indep.

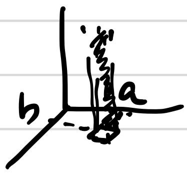
$\Rightarrow$ M has the same distribution as $X + \frac{1}{2}Y$.

3. WTF: joint PDF of $X_{(i)}$ and $X_{(j)}$, $1 \le i \le j \le n$.
By definition of order statistics,

$$X_{(i)} \le X_{(j)}.$$



"$f_{X_{(i)}, X_{(j)}}(a, b)\, da\, db$" is the probability that $X_{(i)}$ and $X_{(j)}$ falls into infinitesimal intervals of length da, db around a, b.



It is equivalent that we independently put n objects into 5 categories.

Therefore,

$f_{X_{(i)},X_{(j)}}(a,b)\, da\, db$

$$= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} F(a)^{i-1} f(a)da \left(F(b)-F(a)\right)^{j-i-1} f(b)\,db\,(1-F(b))^{n-j}$$

We then drop da, db on both sides and get

$$f_{X_{(i)},X_{(j)}}(a,b) = F(a)^{i-1} f(a) (F(b)-F(a))^{j-i-1} f(b)\,(1-F(b))^{n-j},$$

where $a<b$.

II. Conditional Expectation

1.

WLOG, we suppose one envelop contains x and the other contains y. Let $x<y$. Let W be the event that we get more money (y). Let X be the event that we choose envelop containing less money (x). Then,

$$\begin{aligned} P(W) &= P(W|X)P(X) + P(W|X^c)P(X^c) \\ &= P(T>x)\tfrac{1}{2} + P(T<y)\tfrac{1}{2} \\ &= [1-(1-e^{-x})]\tfrac{1}{2} + (1-e^{-y})\tfrac{1}{2} \\ &= \tfrac{1}{2} + \tfrac{1}{2}(e^{-x}-e^{-y}) \end{aligned}$$

Since $0<x<y \Rightarrow e^{-x}-e^{-y}>0$

$\Rightarrow P(W) > \frac{1}{2}.$

2.

pf:

We condition on the first step.

Let A be the event that first toss is "H".

Let $X \sim \text{Geom}(p)$. $\Rightarrow P(A) = p$. Let $q = 1-p$.

$$E[X] = E[X \mid A]P(A) + E[X \mid A^c]P(A^c)$$
$$= 0 \cdot p + E[X \mid X \geq 1] \cdot q$$

Memoryless
$$= (1 + E[X])q$$

$$\Rightarrow E[X] = \frac{q}{1-q} = \frac{q}{p}.$$

3.

for $5 \leq j \leq 35$

$$a_j = 1 + \frac{\binom{35-j}{5}}{\binom{35}{5}} a_j + \frac{\binom{j}{1}\binom{35-j}{4}}{\binom{35}{5}} a_{j-1} + \frac{\binom{j}{2}\binom{35-j}{3}}{\binom{35}{5}} a_{j-2}$$
$$+ \cdots + \frac{\binom{j}{5}}{\binom{35}{5}} a_{j-5}$$

for $0 < j < 5$,

$$a_j = 1 + \frac{\binom{35-j}{5}}{\binom{35}{5}} a_j + \frac{\binom{j}{1}\binom{35-j}{4}}{\binom{35}{5}} a_{j-1}$$
$$+ \cdots + \frac{\binom{j}{j}\binom{35-j}{5-j}}{\binom{35}{5}} a_0$$

for $j = 0$, $a_j = 0$.

{ We can let $a_k = 0$ when $k < 0$, then we can combine all conditions. }

4.

Given $V \sim \text{Unif}(0,1)$

$$E[V-b] = \overset{= 0}{E[V-b \mid b < \tfrac{2}{3}V]}\, P(b < \tfrac{2}{3}V)$$
$$+ E[V-b \mid b \geq \tfrac{2}{3}V]\, P(b \geq \tfrac{2}{3}V)$$
$$= E[V-b \mid V \leq \tfrac{3}{2}b]\, P(V \leq \tfrac{3}{2}b)$$
$$= \{E[V \mid V \leq \tfrac{3}{2}b] - b\}\, P(V \leq \tfrac{3}{2}b)$$

If $b \geq \frac{2}{3}$, the bid must be accepted but we lose money on average.

why lose money on average?

Assuming $b < \frac{2}{3}$. Since

$$V \mid V \leq \tfrac{3}{2}b \sim \text{Unif}(0, \tfrac{3}{2}b),$$

$$E[V-b] = \{\tfrac{3}{4}b - b\}\, \tfrac{3}{2}b$$
$$= -\tfrac{3}{8}b^2.$$

$\Rightarrow$ optimal bid is $b = 0$.

This story tells us:

1). Investing in an asset without any information about its value is a bad idea.

2). Conditioning on any information!

It is crucial in the above calculation to use $E[V \mid V \leq \frac{3}{2}b]$ rather than $E[V] = \frac{1}{2}$. Knowing the bid was accepted gives information about how much the mystery prize is worth!