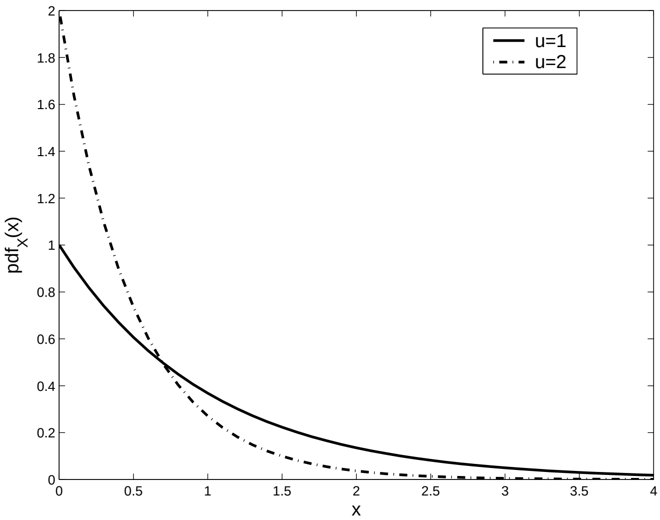
Problem 1

Part a):
i) Since

$$\int_0^\infty f_X(x)dx = \int_0^\infty ae^{-\mu x}dx = 1,$$

we get $a = \mu$.

ii) The pdfs for $\mu = 1, 2$ are shown as following:



Part b):
i) If $c = 0$ then Z takes the only value d with probability 1; otherwise

$$f_Z(z) = \begin{cases} \frac{1}{|c|} f_Y(\frac{z-d}{c}), & a \le \frac{z-d}{c} \le b; \\ 0, & \text{else.} \end{cases}$$

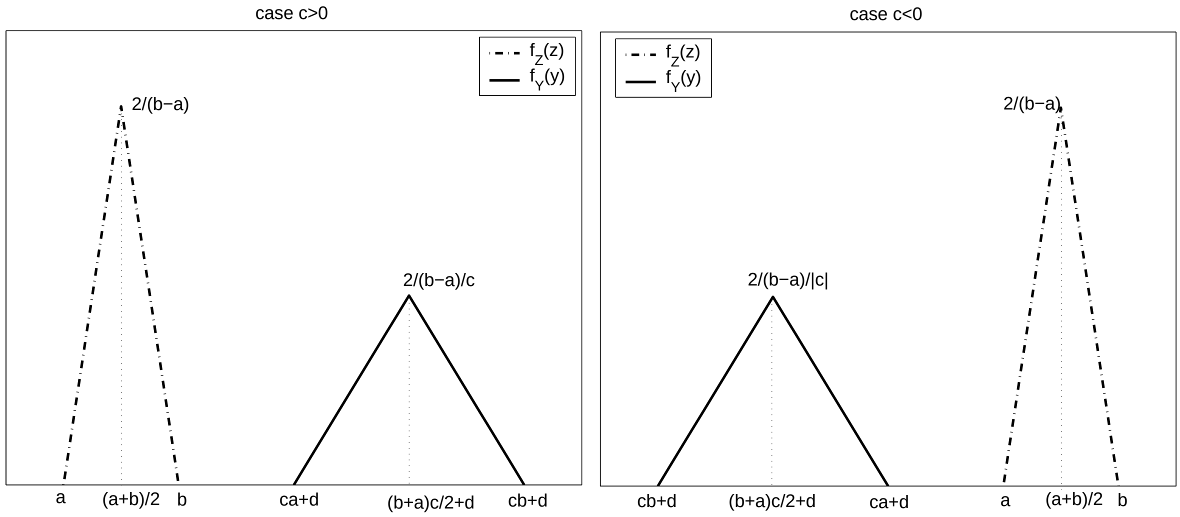
So the triangle for $f_Z(z)$ scales up and narrow if $|c| < 1$, and scales down and flat if $|c| > 1$. The plots for case $c > 0$ and $c < 0$ are shown below:

ii) From figure of $f_Y(y)$ we first notice $f_Y(y)$ is just a convolution of the pdfs of two uniform rvs in $[a/2, b/2]$. So Y is nothing but the sum of these two independent uniform rvs; thus easy to have

$$E[Y] = \frac{a+b}{2}, \quad Var(Y) = \frac{2(b-a)^2}{48} = \frac{(b-a)^2}{24}.$$

Then finally we have

$$E[Z] = cE[Y] + d = \frac{c(b+a)}{2} + d,$$



Note: Not to scale

and

$$Var(Z) = c^2 Var(Y) = \frac{(b-a)^2 c^2}{24}.$$

Note: if you are instead interested in doing integral, here you go:

Easy to see $E[Y] = \frac{b+a}{2}$ since the pdf of Y is symmetric around $\frac{b+a}{2}$, thus

$$E[Z] = E[cY + d] = cE[Y] + d = \frac{(b+a)c}{2} + d.$$

To compute $Var(Z)$, we compute $Var(Y)$ first. Notice

$$f_Y(y) = \begin{cases} \frac{2}{b-a}(1 - \frac{2}{b-a}|y - \frac{b+a}{2}|), & a \le y \le b; \\ 0, & \text{else.} \end{cases}$$

we get

$$\begin{aligned} Var(Y) &= \int_a^b (y - E[Y])^2 f_Y(y) dy \\ &= \int_a^b (y - \frac{b+a}{2})^2 \frac{2}{b-a}(1 - \frac{2}{b-a}|y - \frac{b+a}{2}|) dy \\ (\text{let } z = y - \tfrac{b+a}{2}) \quad &= \frac{2}{b-a} \int_{\frac{-b+a}{2}}^{\frac{b-a}{2}} z^2 (1 - \frac{2}{b-a}|z|) dz \\ &= \frac{4}{b-a} \int_0^{\frac{b-a}{2}} z^2 dz - \frac{8}{(b-a)^2} \int_0^{\frac{b-a}{2}} z^3 dz \\ &= \frac{(b-a)^2}{24}. \end{aligned}$$

Thus finally $Var(Z) = c^2 Var(Y) = \frac{c^2 (b-a)^2}{24}$.

Problem 2

a) Define $U = \max(X, Y)$ and $V = \min(X, Y)$. We first compute cdfs $F_U(u)$ and $F_V(v)$, then take the derivatives to get the pdfs. For U, we have

$$\begin{aligned} F_U(u) &= P(U \le u) \\ &= P(X \le u, Y \le u) \\ &= P(X \le u)P(Y \le u) \quad \text{(since } X \text{ and } Y \text{ are independent.)} \\ &= (1 - e^{-\mu_X u})(1 - e^{-\mu_Y u}), \end{aligned}$$

so

$$f_U(u) = \frac{dF_U(u)}{du} = \begin{cases} \mu_X e^{-\mu_X u} + \mu_Y e^{-\mu_Y u} - (\mu_X + \mu_Y)e^{-(\mu_X+\mu_Y)u}, & u \ge 0; \\ 0, & \text{else.} \end{cases}$$

Similarly for V, we get

$$\begin{aligned} P(V \ge v) &= P(X \ge v, Y \ge v) \\ &= P(X \ge v)P(Y \ge v) \quad \text{(since } X \text{ and } Y \text{ are independent.)} \\ &= e^{-(\mu_X+\mu_Y)v}, \end{aligned}$$

so

$$f_V(v) = \frac{d(1 - P(V \ge v))}{dv} = \begin{cases} (\mu_X + \mu_Y)e^{-(\mu_X+\mu_Y)v}, & v \ge 0; \\ 0, & \text{else.} \end{cases}$$

b) Let $Z = X + Y$. Since X and Y are independent, so f_Z is just the convolution of f_X and f_Y. So, for $z > 0$, we have

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(x) f_Y(z - x)dx \\ &= \int_0^z \mu_X e^{-\mu_X x} \mu_Y e^{-\mu_Y (z-x)} dx \\ &= \\ &= \begin{cases} \frac{\mu_X \mu_Y}{\mu_X - \mu_Y} \left(e^{-\mu_Y z} - e^{-\mu_X z}\right), & \mu_X \ne \mu_Y; \\ \mu_X^2 \, z e^{-\mu_X z}, & \text{else.} \end{cases} \end{aligned}$$

for $z \le 0$, easy to see $f_Z(z) = 0$.

c)

$$\begin{aligned} f_{X|Z}(x|a) &= \frac{f_{X,Z}(x, a)}{f_Z(a)} \\ &= \frac{f_{X,Y}(x, a - x)}{f_Z(a)} \\ &= \frac{f_X(x) f_Y(a - x)}{f_Z(a)} \quad \text{(since } X \text{ and } Y \text{ are independent.)} \\ &= \begin{cases} \frac{\mu_X - \mu_Y}{1 - e^{-(\mu_X - \mu_Y)a}} e^{-(\mu_X - \mu_Y)x}, & 0 \le x \le a, \mu_X \ne \mu_Y; \\ \frac{1}{a}, & 0 \le x \le a, \mu_X = \mu_Y; \\ 0, & \text{else.} \end{cases} \end{aligned}$$

An interesting observation here is that if $\mu_X = \mu_Y$, i.e. X and Y are iid exponentially distributed, then any event of the form $\{X = x, Y = a - x\}$ has the same probability[1] (note: not necessarily true for any distribution), thus it is not surprising that conditional on $X + Y = a$, X is uniformly distributed. This observation is useful later when we deal with Poisson process.

d) We first compute $P(X \geq w|X \geq x)$, then we take the derivative of $F_{X|X\geq x}(w) = 1 - P(X \geq w|X \geq x)$ to get the conditional pdf.

$$\begin{aligned} P(X \geq w|X \geq x) &= \frac{P(X \geq w, X \geq x)}{P(X \geq x)} \\ &= \begin{cases} \frac{P(X\geq w)}{P(X\geq x)}, & w \geq x; \\ \frac{P(X\geq x)}{P(X\geq x)}, & \text{else.} \end{cases} \\ &= \begin{cases} \frac{e^{-\mu_X w}}{e^{-\mu_X x}}, & w \geq x; \\ 1, & \text{else.} \end{cases} \\ &= \begin{cases} e^{-\mu_X (w-x)}, & w \geq x; \\ 1, & \text{else.} \end{cases} \end{aligned}$$

Thus

$$f_{X|X\geq x}() = \begin{cases} \mu_X e^{-\mu_X (w-x)}, & w \geq x; \\ 0, & \text{else.} \end{cases}$$

e) As we seen from part d), given the knowledge of $X \geq x$, the conditional distribution of X is exactly the same as unconditional one[2], i.e. having past knowledge does not has any effect on X, as if it is "memoryless".

In reality, exponential rv is often used to model the packet arrival in router, thus the memoryless property implies no matter how long you have waited for a packet to arrive, e.g. 1 second or 100 hours, the distributions of the time you still need to wait until it finally arrives are the same.

[1] In general, the event $\{X = x, Y = a - x\}$ has probability 0, here for convenience of the discussion, we accept the concept and understand the probability as $f_{X,Y}(x, a - x)dxdy$.

[2] Of course after shifting the origin from 0 to x.

Problem 3

a) It is neither continuous nor discrete, because X takes value among $[0, a]$, 10 and 12. One can use, in strict sense, neither pdf nor pmf to fully describe the distribution.

b) First note $F_X(\infty) = 1$, thus $b = 1 - 0.2 - 0.3 = 0.5$, and $a = \sqrt{b} \approx 0.707$.

c) Combine what we have done in class (for continuous rv) and in homework (for discrete rv), easy to generate X from a uniform rv in $[0, 1]$, defined as Y, as follows:

$$\begin{aligned} x &= 12, \quad \text{if } 0.7 \le y \le 1; \\ x &= 10, \quad \text{if } 0.5 \le y \le 0.7; \\ x &= \sqrt{y}, \quad \text{if } 0 \le y \le 0.5. \end{aligned}$$

Then easy to verify X has the desire pdf as follows:

$$f_X(x) = \begin{cases} f_Y(x^2)/\frac{1}{2x} = 2x, & \text{if } 0 \le x \le \sqrt{0.5}; \\ 0.2\delta(x-10) + 0.3\delta(x-12), & \text{else.} \end{cases}$$

d) Two methods. First is to use $Bern(1/2)$ to generate binary sequences with all the sequences having equal probabilities. Then map the binary sequence approximately to a real number in $[0, 1]$, by this way we generate an uniform rv in $[0, 1]$, then we can follow the process in part c) to generate X. Second method is to generate a bunch Bernoulli rvs and sum them up to generate an approximate Gaussian rv, say Z. We can also have $F_Z(z)$ since $E[Z]$ and $Var(Z)$ can be computed. Then we generate $F_Z(Z)$, which then follow uniform distribution in $[0, 1]$. Thus in this way we can generate an uniform rv, and then use it to generate X.

Problem 4

Part I

a) MAP rule is just to compare the posterior probabilities, i.e. the conditional probability of input given observation(s):

$$
\begin{aligned}
& P(X=a|Y_a=y_a) \underset{\hat{X}=-a}{\overset{\hat{X}=a}{\gtrless}} P(X=-a|Y_a=y_a) \\
\Leftrightarrow \quad & f_{Y_a|X}(y_a|a) \underset{\hat{X}=-a}{\overset{\hat{X}=a}{\gtrless}} f_{Y_a|X}(y_a|-a) \quad \text{(Note } X \text{ is equally likely to be } \pm a\text{)}.
\end{aligned}
$$

Conditional on $X = \pm a$, Y_a is just gaussian with mean $\pm h_a a$ and variance σ^2, thus the above rule becomes

$$
\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_a - h_a a)^2}{2\sigma^2}} \underset{\hat{X}=-a}{\overset{\hat{X}=a}{\gtrless}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_a + h_a a)^2}{2\sigma^2}};
$$

further simplification produces:

$$
h_a y_a \underset{\hat{X}=-a}{\overset{\hat{X}=a}{\gtrless}} 0.
$$

(Note here you might not want to remove coefficient h_a because the sign of $h_a 0$ then might change the decision from $\hat{X} = a$ to $\hat{X} = -a$.)

Similar to what we did in class and homework, noticing given $X = \pm a$, $h_a Y_a$ follows $N(\pm h_a^2 a, h_a^2 \sigma^2$, thus the error probability is

$$
\begin{aligned}
Pr(e) &= Pr(e|X=a)P(X=a) + Pr(e|X=-a)P(X=-a) \\
&= P(h_a Y_a \le 0|X=a)\frac{1}{2} + P(h_a Y_a \ge 0|X=-a)\frac{1}{2} \\
&= Q\left(\sqrt{\frac{h_a^2 a^2}{\sigma^2}}\right) = Q\left(\sqrt{h_a^2 SNR}\right).
\end{aligned}
$$

b) Still, we just need to compare the posterior probabilities, similar to part a) we get:

$$
f_{Y_a,Y_b|X}(y_a, y_b|a) \underset{\hat{X}=-a}{\overset{\hat{X}=a}{\gtrless}} f_{Y_a,Y_b|X}(y_a, y_b|-a),
$$

Conditional on $X = \pm a$, Y_a and Y_b are just gaussian rvs with mean being $\pm h_a a$, $\pm h_b a$ and variance all being σ^2, thus the above rule can be further simplified to

$$
h_a y_a + h_b y_b \underset{\hat{X}=-a}{\overset{\hat{X}=a}{\gtrless}} 0.
$$

The intuition here is the observation from a less strong channel play a less significant role in making the decision.

Noticing given $X = \pm a$, $h_a Y_a + h_b Y_b$ still follows gaussian distribution $N(\pm h_a^2 a \pm h_b^2 a, h_a^2\sigma^2 + h_b^2\sigma^2)$, the probability error is

$$Pr(e) = Pr(e|X = a)P(X = a) + Pr(e|X = -a)P(X = -a) = Q\left(\sqrt{h_a^2 SNR + h_b^2 SNR}\right).$$

c) From part a), we see that if the channel is in bad state, i.e. $|h_a|$ is very small, then the system performs poorly. Thus we really want to reduce the chance of the channel in bad state, but it is out of our control. From part b), however, we see we can play with independent channels to achieve good performance. As long as one of the channels is good, i.e. $|h_a|$ or $|h_b|$ is large, then the system performance will be ok. Since the chance for two independent channel to be both in bad state is small, the chance for getting a small RECEIVING SNR $\frac{h_a^2 a^2}{\sigma^2} + \frac{h_b^2 a^2}{\sigma^2}$ is small, thus we can retain good performance most of the time. This is the diversity technique used in many wireless communication systems. (Q: what about we have more and more receiving antennas?)

Part II:

a) Since now h_a and h_b is rvs following $N(0, 1)$, then given input X, the conditional (joint or individual) distributions of Y_a and Y_b do not change, i.e. they are just the same as the unconditional ones. In other words observations Y_a, Y_b are independent of input X, i.e. no correlation between input and observations. Thus observing Y_a, Y_b won't give you any information about X. Thus the system can not convey any information from sender to receiver.

b) Answer is included in a).

c) Now the conditional distributions of Y_1 and Y_2 changes as the input H changes, i.e. input and observations are correlated. Thus one can expect to infer some knowledge about H by observing Y_1 and Y_2. So the system should be able to convey some information from sender to receiver.

Again, the MAP rule is just to compare the conditional probability of input given observation(s):

$$P(H = 0|Y_{1a} = y_{1a}, Y_{2a} = y_{2a}) \underset{\hat{H}=1}{\overset{\hat{H}=0}{\gtrless}} P(H = 1|Y_{1a} = y_{1a}, Y_{2a} = y_{2a})$$

$$\Leftrightarrow \quad f_{Y_{1a},Y_{2a}|H}(y_{1a}, y_{2a}|0) \underset{\hat{H}=1}{\overset{\hat{H}=0}{\gtrless}} f_{Y_{1a},Y_{2a}|H}(y_{1a}, y_{2a}|1) \quad \text{(Note } H \text{ is equally likely to be 1 or 0).}$$

Conditional on $H = 0$, $Y_{1a} \sim N(0, 1 + \sigma^2)$, $Y_{2a} \sim N(0, \sigma^2)$, and they are independent; conditional on $H = 1$, $Y_{1a} \sim N(0, \sigma^2)$, $Y_{2a} \sim N(0, 1 + \sigma^2)$, and they are independent. Thus the above rule becomes

$$e^{-\frac{y_{1a}^2}{2(1+\sigma^2)}} e^{-\frac{y_{2a}^2}{2\sigma^2}} \underset{\hat{H}=1}{\overset{\hat{H}=0}{\gtrless}} e^{-\frac{y_{1a}^2}{2\sigma^2}} e^{-\frac{y_{2a}^2}{2(1+\sigma^2)}};$$

further simplification produces:

$$y_{1a}^2 \underset{\hat{H}=1}{\overset{\hat{H}=0}{\gtrless}} y_{2a}^2.$$

d) In the two antennas case, we have two observations each time slot, and that is the ONLY change to the system. Also note given H, Y_{1a}, Y_{1b} Y_{2a} and Y_{2b} are just independent gaussian rvs. Follow the similar procedure, we can have the MAP rule as

$$f_{Y_{1a},Y_{2a},Y_{1b},Y_{2b}|H}(y_{1a},y_{2a},y_{1b},y_{2b}|0) \underset{\hat{H}=1}{\overset{\hat{H}=0}{\gtrless}} f_{Y_{1a},Y_{2a},Y_{1b},Y_{2b}|H}(y_{1a},y_{2a},y_{1b},y_{2b}|1)$$

$$\Leftrightarrow \quad y_{1a}^2 + y_{1b}^2 \underset{\hat{H}=1}{\overset{\hat{H}=0}{\gtrless}} y_{2a}^2 + y_{2b}^2.$$