Midterm 1 Solutions

---

1. **Problem 1** (14 points) Consider a probability space $\mathcal{P} = (\Omega, \mathcal{F}, P)$ and the events $A, B \in \Omega$.

   (a) (5 points) If A and B are independent events, they cannot be disjoint.
   FALSE
   If A and B are independent then, $P[AB] = P[A]P[B]$ and if A and B are disjoint then $AB = \Phi$. Thus A and B can be both independent and disjoint if either $P[A] = 0$ or $P[B] = 0$, or both probabilities are equal to zero.
   Partial credit given for the correct definitions of independence and disjoint.

   (b) (5 points) If $P[A] \geq P[B]$ then for another arbitrary set C, $P[A|C] \geq P[B|C]$
   FALSE
   Let $AC = \Phi$, but $BC \neq \Phi$; furthermore let $P[BC] > 0$. Then if $P[A] \geq P[B]$ we also have: $P[AC] = P[\Phi] = 0$ and $P[BC] > 0$; clearly then we cannot have $P[A|C] \geq P[B|C]$ since $P[A|C] = 0$ and $P[B|C] > 0$.
   Partial credit given for the correct definition of conditional probability.

   (c) (4 points) Given that $A \subset B$ and $P[A] = \frac{1}{4}$ and $P[B] = \frac{1}{3}$, compute the following: $P[A|B]$ and $P[B|A]$.

$$
\begin{aligned}
P[A|B] &= \frac{P[AB]}{P[B]} = \frac{P[A]}{P[B]} = \frac{3}{4} \\
P[B|A] &= \frac{P[AB]}{P[A]} = \frac{P[A]}{P[A]} = 1
\end{aligned}
$$

2. **Problem 2** (22 points total + 4 BONUS)

   (a) (6 points) Given the conditions stated, and letting T be the variable for transmission and let R be the variable for reception, the table of interest is determined as:

$P[T=0] = 1-\alpha-\beta$	$P[R=0\|T=0] = 1-\beta-p$	$P[R=0\|T=G] = \frac{1}{3}$	$P[R=0\|T=1] = 2p$
$P[T=G] = \beta$	$P[R=G\|T=0] = \beta$	$P[R=G\|T=G] = \frac{1}{3}$	$P[R=G\|T=1] = \beta$
$P[T=1] = \alpha$	$P[R=1\|T=0] = p$	$P[R=1\|T=G] = \frac{1}{3}$	$P[R=1\|T=1] = 1-\beta-2p$

   Note that the columns all need to sum to 1. *Also note that if you filled out the table incorrectly, the rest of the problem was graded based on* **your** *table and not the true table. So, if you incorrectly used your own table's values, points were deducted.*

   (b) (6 points) What is the probability of the received signal being a garbage symbol, $P[R = G]$? Can you minimize your probability of a received garbage symbol through choice of α?
   We use total probability:

$$
\begin{aligned}
P[R=G] &= P[R=G|T=0]P[T=0] + P[R=G|T=1]P[T=1] + P[R=G|T=G]P[T=G] \\
&= \beta(1-\alpha-\beta) + \frac{1}{3}\beta + \beta\alpha = \frac{4}{3}\beta - \beta^2
\end{aligned}
$$

   As $P[R = G] \neq g(\alpha)$, we cannot affect $P[R = G]$ through our choice of α.
   Partial credit given for the correct statement of total probability in terms of generic sets.

(c) (4 points) Given a decoded garbage symbol ($R = G$), what is the probability that the garbage symbol was sent ($T = G$)?

We use Bayes rule:

$$P[T = G|R = G] \;=\; \frac{P[R = G|T = G]P[T = G]}{P[R = G]} = \frac{\frac{1}{3}\beta}{\frac{4}{3}\beta - \beta^2} = \frac{1}{4 - 3\beta}$$

Partial credit given for the correct statement of Bayes rule.

(d) (2 points) Evaluate the normalized probability of error: $\frac{P[R=0|T=1]P[T=1]+P[R=1|T=0]P[T=0]}{P[T=1]+P[T=0]}$.

Plugging into the formula from the table we have:

$$P_e \;=\; \frac{2p(\alpha) + p(1 - \alpha - \beta)}{\alpha + (1 - \alpha - \beta)} = \frac{p(1 + \alpha - \beta)}{1 - \beta}$$

(e) (5 points) What is the optimal value for α to minimize the normalized probability of error? Comment on this choice of α.

We note that P_e is linear in α, thus we have

$$\arg\min_{\alpha} P_e \;=\; \arg\min_{\alpha} \alpha = \min \alpha = 0$$

Note that if $\alpha = 0$, we are not sending any meaningful information over this channel, since we only send 0's. Also observe that this solution is intuitive since $P[R = 0|T = 1] = 2p = 2P[R = 1|T = 0]$. That is we make errors on transmitted 1's twice as often as we make errors on transmitted 0's, thus we should minimize the number of 1's that we send.

(f) (4 points) **BONUS** If all you can determine from the output of your receiver is the normalized probability of error, can you design a strategy to determine β given that you can adjust α and p is known? If so, what is that strategy? And if not, explain why not.

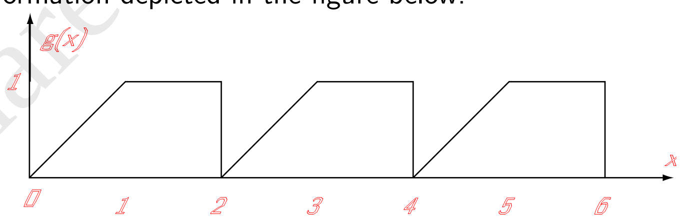
Note the following:

$$\alpha = 1 \;\rightarrow\; P_e = q \;=\; \frac{p(2 - \beta)}{1 - \beta} \;\rightarrow\; \beta = \frac{2p - q}{2 - q}$$
$$\alpha = 0 \;\rightarrow\; P_e = q = p$$

Thus by setting $\alpha = 0$ we can determine p from P_e and similarly, by setting $\alpha = 1$ and knowing p and P_e we can solve for β.

3. **Problem 3** (29 points total + 4 BONUS)

Consider the transformation depicted in the figure below:



Let $Y = g(X)$ where X is a uniform random variable on $[0, 6]$,

(a) (2 point) What kind of random variable is Y? – mixed random variable.

(b) (8 points) Evaluate $f_Y(y)$; simplify your expression as much as possible.

Using the strategy discussed in lecture, we first consider the "flat" parts of $g(x)$. That is

$$\begin{aligned}
P[Y=1] &= P[X \in [1,2] \cup [3,4] \cup [5,6]] = P[X \in [1,2]] + P[X \in [3,4]] + P[X \in [5,6]] \\
&\quad \text{since the intervals above are disjoint} \\
&= \frac{1}{6}(2-1) + \frac{1}{6}(4-3) + \frac{1}{6}(5-6) = \frac{1}{2} \\
&\quad \text{since } X \text{ is } U[0,6]
\end{aligned}$$

The above calculations imply that $f_Y(y)$ has a delta function: $\frac{1}{2}\delta(y-1)$.
We next consider the other regions of the function. Recall the general formula:

$$f_Y(y) = \sum_{l=1}^{n} f_X(x_l) \frac{1}{|g'(x_l)|}$$

For $y \in [0,1)$, there are three roots to the equation: $g(x)-y=0$: $x_1 = y, x_2 = y+2, x_3 = y+4$. Note that $f_X(x_l) = \frac{1}{6}$ for all three of these roots. Furthermore since each root is an affine function of y, we have,$g'(x_l) = 1$ for all three roots. Thus we have:

$$y \in [0,1): \quad f_Y(y) = \sum_{l=1}^{3} \left(\frac{1}{6} \times \frac{1}{1} \right) = \frac{1}{2}$$

Thus the complete pdf is specified by,

$$f_Y(y) = \begin{cases} \frac{1}{2} + \frac{1}{2}\delta(y-1) & y \in [0,1] \\ 0 & \text{else} \end{cases}$$

Note that one could have computed the cdf first and then computed the pdf. That is,

$$\begin{aligned}
y \in [0,1) \ ; \ F_Y(y) = P[Y \leq y] &= P[X \in [0,y] \cup [2, y+2] \cup [4, y+4]] \\
&= P[X \in [0,y]] + P[X \in [2,y+2]] + P[X \in [4,y+4]] \\
&= \frac{1}{6}(y-0) + \frac{1}{6}(y+2-2) + \frac{1}{6}(y+4-4) = \frac{1}{2}y \\
f_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{1}{2} \ \ y \in [0,1)
\end{aligned}$$

And we obtain the same answer as above. A key component of the solution is the mass point/delta function at 1.
Partial credit for the pdf formula of $f_Y(y)$ in terms of $f_X(x)$.

(c) (6 points) Determine the characteristic function of Y; simplify your expression as much as possible.

$$\begin{aligned}
\Phi_Y(\omega) &= \mathbf{E}\left[e^{j\omega Y}\right] = \int f_Y(y) e^{j\omega y} dy = \int_0^1 \left(\frac{1}{2} + \frac{1}{2}\delta(y-1) \right) e^{j\omega y} dy \\
&= \left. \frac{e^{j\omega y}}{2j\omega} \right|_0^1 + \left. \frac{e^{j\omega y}}{2} \right|_{y=1} = \frac{1}{2j\omega}\left(e^{j\omega} - 1\right) + \frac{1}{2}e^{j\omega}
\end{aligned}$$

NOTE: Recall the sifting property for delta functions: $\int f(y)\delta(y-y_0)dy = f(y_0)$ – assuming that the interval of integration includes the point y_0. *Partial credit given for the expectation and/or integral definition of the characteristic function alone.*

(d) (5 points)Determine $P[Y > b]$ exactly for $b \in (0,1)$.

$$\begin{aligned} P[Y > b] &= 1 - F_Y(b) = \int_b^\infty f_Y(y)dy = \int_b^1 \left(\frac{1}{2} + \frac{1}{2}\delta(y-1)\right) dy \\ &= \frac{1}{2}\, y|_b^1 + \frac{1}{2} = \frac{1}{2}(1-b) + \frac{1}{2} = 1 - \frac{b}{2} \end{aligned}$$

Partial credit given for recognizing the relationship between the desired probability and the cdf.

(e) (4 points) Determine $\mathbf{E}[Y^3]$. Simplify your answer as much as possible.

$$\mathbf{E}\left[Y^3\right] = \int f_Y(y) y^3 dy = \int f_X(x) g^3(x) dx$$

Looking ahead, we consider the calculation of $\mathbf{E}[Y^n]$,

$$\begin{aligned} \mathbf{E}[Y^n] &= \int f_Y(y) y^n dy = \int_0^1 \left(\frac{1}{2} + \frac{1}{2}\delta(y-1)\right) y^n dy \\ &= \frac{1}{2}\, \frac{y^{n+1}}{n+1}\bigg|_0^1 + \frac{1}{2}\, y^n|_{y=1} = \frac{1}{2(n+1)} + \frac{1}{2} = \frac{n+2}{2n+2} \\ n = 3 \;\; ;\mathbf{E}\left[Y^3\right] &= \frac{5}{8} \\ n = 2 \;\; ;\mathbf{E}\left[Y^2\right] &= \frac{2}{3} \end{aligned}$$

Partial credit given for the definition of expectation of a function, the desired expectation in terms of X, and/or the relationship between the third moment and the characteristic function.

(f) (4 points) Determine a bound on $P[Y > b]$ using $\mathbf{E}[Y^3]$. Simplify your answer as much as possible.

Recall the generalization of the Markov inequality, for positive, non-decreasing functions of positive random variables we have:

$$\begin{aligned} P[Y > b] &\leq \frac{\mathbf{E}[g(Y)]}{g(b)} \\ \text{let } g(x) = x^3 \text{ then } P[Y > b] &\leq \frac{\mathbf{E}\left[Y^3\right]}{b^3} = \frac{5}{8b^3} \end{aligned}$$

(g) (4 points) **BONUS** For what values of b is your bound based on $\mathbf{E}[Y^3]$ tighter than that provided by the non-central Chebyshev bound?

For the non-central Chebyshev bound we have $g(x) = x^2$ and the bound is given by

$$P[Y > b] \leq \frac{\mathbf{E}\left[Y^2\right]}{b^2} = \frac{2}{3b^2}$$

For our third moment bound to be tighter, it must be **smaller** than the second moment bound, thus we seek the range of b such that:

$$\frac{5}{8b^3} < \frac{2}{3b^2} \quad \rightarrow \quad b > \frac{15}{16}$$

4. **Problem 4** (16 points total)

   Consider two cases of coins in pockets. Wenyi has four coins in his pocket: 3 pennies (one cent each) and one quarter (25 cents each). Cecilia has four coins in her pocket: 3 dimes (10 cents each) and 1 nickel (5 cents each).

   (a) (3 points) Wenyi takes one coin out, notes its value, returns it and then takes out a second coin. What is the probability that he selects the quarter **twice**?

   Since we have replacement, the total number of combinations is $4^2 = 16$. For each experiment (taking out a coin), there is only one way to get a quarter. Thus the probability of this event is $\frac{1}{16}$.

   (b) (4 points) Let C_1 be the coin value of the first coin taken out and then put back in the pocket and let C_2 be the value of the second coin taken out. What is $J_w \doteq \mathbb{E}[C_1 + C_2]$? We note that the probability mass function for each C_i is the same, thus $J_w = 2\mathbb{E}[C_1]$. There are only two values for each coin: $C_i \in 1, 25$. From our reasoning above the probability of a quarter is $\frac{1}{4}$, thus the probability of a penny must be $1 - \frac{1}{4} = \frac{3}{4}$.

$$\mathbb{E}[C_1] = \frac{1}{4} \times 25 + \frac{3}{4} \times 1 = \frac{28}{4} = 7$$

   (c) (5 points) Cecilia takes out two coins at one time. What is the probability that she selects two dimes?

   In this case, we have sampling without replacement and without order and thus the choose function is the relevant function:

$$\begin{aligned}
\#\text{ selecting two coins from four} &= \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \frac{n!}{k!(n-k)!} = \frac{4!}{2!2!} = 6 \\
\#\text{ selecting two coins from three} &= \frac{3!}{2!1!} = 3 \\
\text{probability of two dimes} &= \frac{3}{6} = \frac{1}{2}
\end{aligned}$$

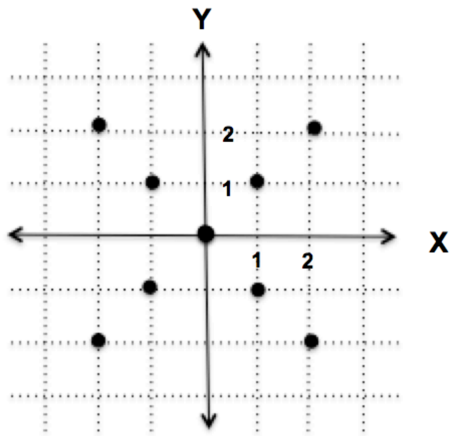
   (d) (3 points) Let C_3 be value of one of the coins Cecilia takes out and C_4 the value of the other coin. What is $J_c \doteq \mathbb{E}[C_3 + C_4]$ for the method specified in (c)?

   We solve this problem by recognizing that there are only two cases here: either we select two dimes or we select one dime and one nickel. Thus, if the probability of selecting two dimes, event D, is $\frac{1}{2}$, then $\mathbb{P}[D^c] = 1 - \mathbb{P}[D] = \frac{1}{2}$. Thus we have that $J_c = \frac{1}{2}[10 + 10] + \frac{1}{2}[5 + 10] = \frac{35}{2} = 17.5$.

   (e) (1 points) Clearly Cecilia has the higher average coin value. Thus even though Wenyi has one quarter that he can possibly select twice, the probability is not high enough to outweigh the presence of the three pennies.

5. **Problem 5**(25 points total + 4 BONUS)

   Consider the figure showing the location of mass points for a bivariate random variable (X, Y).



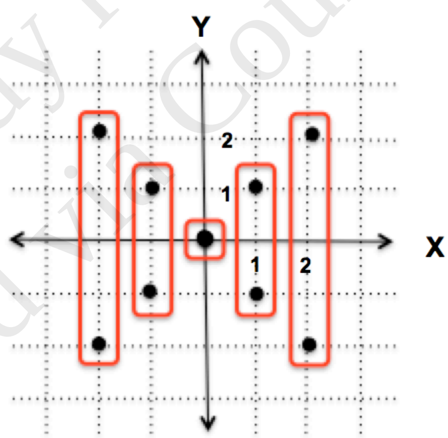
$$\mathbb{P}[X = x_i, Y = y_j] \quad = \quad \begin{cases} 4a & x_i = 0, y_j = 0 \\ a & \text{else} \end{cases}$$

(a) (3 points) The valid value of a is determined as,

$$\sum_{i,j} \mathbb{P}[X = x_i, Y = y_j] \quad = \quad 4a + \sum_{i=1}^{8} a = 12a = 1$$
$$\rightarrow \qquad a = \frac{1}{12}$$

(b) (5 points) Determine the marginal probability mass functions for X and Y.
We average over the other variable:

$$\mathbb{P}[X = x_i] \quad = \quad \sum_{j} \mathbb{P}[X = x_i, Y = y_j]$$



Consider the rectangles drawn over the joint pmf mass points for computing $\mathbb{P}[X = x_i]$, as in the figure above, we thus have

$$\mathbb{P}[X = -2] \quad = \quad \mathbb{P}[X = -2, Y = 2] + \mathbb{P}[X = -2, Y = -2] = \frac{1}{12} + \frac{1}{12} = \frac{1}{6}$$

Similarly we have, $\mathbb{P}[X = -1] = \mathbb{P}[X = 1] = \mathbb{P}[X = 2] = \frac{1}{6}$. The final mass point has the remaining probability, $\mathbb{P}[X = 0] = \frac{1}{3}$. Summarizing we have,

$$\mathbb{P}[X = x_i] \quad = \quad \begin{cases} \frac{1}{6} & x_i = \pm 1, \pm 2 \\ \frac{1}{3} & x_i = 0 \end{cases}$$

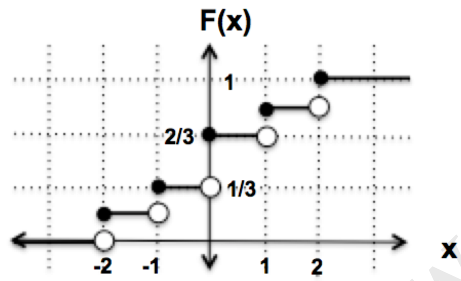
Note that $\sum \mathbb{P}[X = x_i] = 1$ as we expect for a valid probability mass function. From symmetry, Y has the same probability mass function.

(c) (5 points) Determine the cumulative distribution function for X and draw the function.

The cdf is given by

$$F_X(x) \;=\; \sum_{i:x_i \le x} \mathbb{P}[X = x_i]$$

The figure is below. A critical component to have drawn was the right continuity (*e.g.*) the open circles versus the closed circles ending the horizontal lines.



(d) (5 points) Determine the mean and variance of Y.

We observe that X and Y have the same distributions. Thus

$$\begin{aligned}
\mathbb{E}[Y] &= \sum_j y_j \mathbb{P}[Y = y_j] = \frac{1}{6}\Big(-2+2) + \frac{1}{6}(-1+1) + \frac{1}{3}(0) = 0 \\
\mathbf{Var}[Y] &= \mathbb{E}\left[Y^2\right] - (\mathbb{E}[Y])^2 = \mathbb{E}\left[Y^2\right] = \sum_j y_j^2 \mathbb{P}[Y = y_j] \\
&= \frac{1}{6}\left((-2)^2 + 2^2\right) + \frac{1}{6}\left((-1)^2 + 1^2\right) + \frac{1}{3}(0^2) = \frac{10}{6} = \frac{5}{3}
\end{aligned}$$

(e) (4 points) Are X and Y uncorrelated?

If X and Y are uncorrelated we have that $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$,

$$\begin{aligned}
\mathbb{E}[XY] &= \sum_i \sum_j x_i y_j \mathbb{P}[X = x_i, Y = y_j] \\
&= \frac{1}{12}((-2)(-2) + (2)(2)) + \frac{1}{12}((-1)(-1) + (1)(1)) \\
&\quad + \frac{1}{12}((1)(-1) + (-1)(1)) + \frac{1}{12}((2)(-2) + (2)(-2)) + \frac{1}{3}(0 \times 0) = 0
\end{aligned}$$

We know from (d) that both X and Y are zero mean. Thus, X and Y are uncorrelated.

(f) (3 points) Are X and Y statistically independent?

Since $\mathbb{P}[X = 0, Y = 0] = \frac{1}{3} \neq \frac{1}{9} = \mathbb{P}[X = 0]\mathbb{P}[Y = 0]$, they cannot be independent.

(g) (4 points) **Bonus:** If X and Y are not statistically independent, determine a third random variable Z such that Y is a function of X and Z. If they **are** statistically independent, compute the probability generating function for X and Y.

As determined above, X and Y are not statistically independent. Consider $Z = \pm 1$ with probability $\frac{1}{2}$. Let $Y = ZX$.