CS3236: Solutions to Tutorial 0
(Probability Review)

1. **[Expectation, Independence, and Variance]**

   Let V and W be discrete random variables defined on some probability space with a joint probability mass function (PMF) $P_{V,W}(v, w)$

   (a) Prove that $\mathbb{E}[V + W] = \mathbb{E}[V] + \mathbb{E}[W]$. Do not assume independence.
   Solution. *We have*

$$\begin{aligned}\mathbb{E}[V+W] &= \sum_{v,w}(v+w)P_{V,W}(v,w) = \sum_{v,w} vP_{V,W}(v,w) + \sum_{v,w} wP_{V,W}(v,w) \\ &= \sum_{v} vP_V(v) + \sum_{w} wP_W(w) = \mathbb{E}[V] + \mathbb{E}[W]\end{aligned}$$

   (b) Prove that if V and W are independent random variables, then $\mathbb{E}[VW] = \mathbb{E}[V]\mathbb{E}[W]$.
   Solution. *We have*

$$\begin{aligned}\mathbb{E}[VW] &= \sum_{v,w} vwP_{V,W}(v,w) = \sum_{v,w} vwP_V(v)P_W(w) \\ &= \sum_{v} vP_V(v) \sum_{w} wP_W(w) = \mathbb{E}[V]\mathbb{E}[W]\end{aligned}$$

   (c) Let V and W be independent rv's and σ_V^2 and σ_W^2 are their respective variances. Find the variance of $Z = V + W$.
   Solution. *We have*

$$\begin{aligned}\sigma_Z^2 &= \mathbb{E}[Z^2] - \mathbb{E}[Z]^2 \\ &= \mathbb{E}[(V+W)^2] - \mathbb{E}[V+W]^2 \\ &= \mathbb{E}[V^2 + W^2 + 2VW] - (\mathbb{E}[V]^2 + 2\mathbb{E}[V]\mathbb{E}[W] + \mathbb{E}[W]^2) \\ &= \mathbb{E}[V^2] - \mathbb{E}[V]^2 + \mathbb{E}[W^2] - \mathbb{E}[W]^2 \\ &= \sigma_V^2 + \sigma_W^2,\end{aligned}$$

   where the second to last equality follows from the fact that $\mathbb{E}[VW] = \mathbb{E}[V]\mathbb{E}[W]$.

2. **[Modulo-2 Sum]**

   Let $X_1, X_2, \ldots, X_n$ be a sequence of n binary independent and identically distributed (i.i.d.) random variables. Assume that $\Pr(X_m = 1) = \Pr(X_m = 0) = 1/2$ for all $1 \leq m \leq n$. Let Z be a parity check on $X_1, \ldots, X_n$, i.e. $Z = X_1 \oplus \ldots \oplus X_n$ (where $0 \oplus 0 = 0, 0 \oplus 1 = 1$ and $1 \oplus 1 = 0$).

   (a) Is Z independent of X_1? (Assume $n > 1$)
   Solution. *See below.*

(b) Are $Z, X_1, \ldots, X_{n-1}$ independent?
Solution. *Yes. This also covers case (a). Note that $P_Z(z) = 1/2$ for $z = 0, 1$. Note that $Z = 0$ if and only if $X_n = x_1 \oplus \ldots \oplus x_{n-1}$. Therefore the conditional PMF is*

$$P_{Z|X_1,\ldots,X_{n-1}}(0|x_1, \ldots, x_{n-1}) = P_{Z|X_1,\ldots,X_{n-1}}(1|x_1, \ldots, x_{n-1}) = \frac{1}{2}$$

which is equal to the unconditional pmf $P_Z(z)$. Thus, Z and $X_1, \ldots, X_{n-1}$ are independent. Since the X_i are all independent by assumption, it follows that all of $(Z, X_1, \ldots, X_{n-1})$ are independent.

(c) Are $Z, X_1, \ldots, X_n$ independent?
Solution. *Clearly not, since $X_1, \ldots, X_n$ determine Z.*

(d) Is Z independent of X_1 if $\Pr(X_m = 1) \neq 1/2$? (You may take $n = 2$ here)
Solution. *Z is NOT statistically independent of X_1 if $\Pr(X_i = 1) \neq 1/2$. Let $\Pr(X_i = 1) = p$. For $n = 2$ and $Z = 0$, the unconditional pmf $P_Z(z)$ is*

$$P_Z(0) = \Pr(X_1 = 1, X_2 = 1) + \Pr(X_1 = 0, X_2 = 0) = p^2 + (1-p)^2$$
$$P_Z(1) = (1-p)p + p(1-p) = 2p(1-p).$$

However, the conditional PMF $P_{Z|X_1}(z|x_1)$ given $x_1 = 0$ is

$$P_{Z|X_1}(0|0) = \Pr(X_2 = 0) = 1 - p$$
$$P_{Z|X_1}(1|0) = \Pr(X_2 = 1) = p$$

Since $P_{Z|X_1}(z|0) \neq P_Z(z)$, we conclude that Z is not independent of X_1.
The purpose of this question is to show that in a group of $n + 1$ random variables, pairwise independence (part a) does not imply statistical independence of the random variables (part c) (even with the statistical independence of groups of n random variables in part b).

3. **[Coin Flips]**

Flip a fair coin four times. Let X be the number of Heads obtained, and let Y be the position of the first Heads (e.g., if the sequence of coin flips is TTHT, then $Y = 3$; if it is THHH, then $Y = 2$). If there are no heads in the four tosses, then we define $Y = 0$.

(a) Find the joint PMF of X and Y;
Solution. *The underlying sample space is the set*

$$\Omega = \{TTTT, TTTH, \ldots, HHHH\}.$$

Each outcome $\omega \in \Omega$ can be mapped to $X(\omega)$ and $Y(\omega)$, e.g.,

$$X(TTTT) = 0 \qquad Y(TTTT) = 0$$
$$X(THHT) = 2 \qquad Y(THHT) = 2$$
$$X(TTTH) = 1 \qquad Y(TTTH) = 4$$
$$\text{etc.}$$

By listing all 16 elements of Ω, and computing X and Y for each we can see that, e.g.,

$$\{X = 2, Y = 2\} = \{THHT, THTH\}$$

and thus $P_{XY}(2, 2) = \frac{1}{16} + \frac{1}{16} = \frac{1}{8}$. This can be repeated for all feasible values of X and Y as in the table below:

			x		
y	0	1	2	3	4
0	$\frac{1}{16}$				
1		$\frac{1}{16}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{1}{16}$
2		$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	
3		$\frac{1}{16}$	$\frac{1}{16}$		
4		$\frac{1}{16}$			

(b) Using the joint PMF, find the marginal PMF of X

Solution. *Summing over the columns of the table below, we get the PMF of X as*

$$P_X(k) = \begin{cases} \frac{1}{16} & k = 0 \\ \frac{1}{4} & k = 1 \\ \frac{3}{8} & k = 2 \\ \frac{1}{4} & k = 3 \\ \frac{1}{16} & k = 4 \end{cases}$$

4. **[Probability Properties and Bounds]**

(a) For a nonnegative integer-valued rv N, show that $\mathbb{E}[N] = \sum_{n>0} \Pr(N \geq n)$.

Solution. *We have*

$$\begin{aligned} \mathbb{E}[N] &= \sum_{n=0}^{\infty} nP_N(n) = 0 \cdot P_N(0) + 1 \cdot P_N(1) + 2 \cdot P_N(2) + 3 \cdot P_N(3) + \ldots \\ &= [P_N(1) + P_N(2) + P_N(3) + \ldots] \\ &\quad + [P_N(2) + P_N(3) + P_N(4) + \ldots] \\ &\quad + [P_N(3) + P_N(4) + P_N(5) + \ldots] + \ldots \\ &= \Pr(N \geq 1) + \Pr(N \geq 2) + \Pr(N \geq 3) + \ldots = \sum_{n>0} \Pr(N \geq n) \end{aligned}$$

(b) Derive the Markov inequality, which says that for any $a > 0$ and nonnegative X, $\Pr(X \geq a) \leq \mathbb{E}[X]/a$. *(Hint: Let $\mathbf{1}\{\cdot\}$ be the indicator function, which equals one if the condition inside is true and zero otherwise. Use the inequality $\mathbf{1}\{x \geq a\} \leq x/a$ for any $x \geq 0$, which is easily verified by checking the cases $x \geq a$ and $x < a$ separately.)*

Solution. *Using the hint, we know that $\mathbf{1}\{x \geq a\} \leq x/a$ for any $x \geq 0$ and so for any nonnegative random variable X, we have*

$$\mathbf{1}\{X \geq a\} \leq X/a$$

for any ≥ 0. Now take expectation on both sides. Note that $\mathbb{E}[\mathbf{1}\{X \geq a\}] = \Pr(X \geq a)$. This completes the proof.

(c) Derive the Chebyshev inequality, which says that $\Pr(|Y - \mathbb{E}[Y]| \geq b) \leq \sigma_Y^2/b^2$.

Solution. *Take $X = |Y - \mathbb{E}[Y]|^2$ and $a = b^2$ in the Markov inequality. Note that $\sigma_Y^2 = \mathbb{E}[X^2] = \mathbb{E}[|Y - \mathbb{E}[Y]|^2]$.*

(d) **Weak Law of Large Numbers**: Let $X_1, \ldots, X_n$ be a sequence of i.i.d. rvs with zero-mean and finite variance σ^2. Show that for any $\epsilon > 0$, $\Pr\left[\frac{1}{n}(X_1 + \ldots + X_n) > \epsilon\right] \leq \sigma^2/(n\epsilon^2)$.

(Hint: What can we say about the variance of the random variable $\frac{1}{n}(X_1 + \ldots + X_n)$ as n becomes large?)

Solution. *Let $Y = \frac{1}{n}(X_1 + \ldots + X_n)$. The mean of Y is zero and the variance is $\frac{1}{n}\sigma^2$. Apply Chebyshev's inequality.*

5. **[Modulo-M Sum]**

Let X be a random variable uniformly distributed over $\{0, 1, 2, \ldots, M-1\}$. Let Y be a random variable arbitrarily distributed over $\{0, 1, 2, \ldots, M-1\}$, and independent of X. Show that the sum of X and Y modulo M is uniformly distributed and independent of Y.

Solution. *Let $Z = X + Y \ mod \ M$ and $z \in \{0, 1, 2, \ldots, M-1\}$.*

$$\begin{aligned} P(Z = z) &= \sum_{x=0}^{M-1} P(X = x,\, Y = z - x \ mod \ M) \\ &\overset{(a)}{=} \sum_{x=0}^{M-1} P(X = x) \cdot P(Y = z - x \ mod \ M) \\ &= \sum_{x=0}^{M-1} \frac{1}{M} \cdot P(Y = z - x \ mod \ M) \\ &= \frac{1}{M} \sum_{x=0}^{M-1} P(Y = z - x \ mod \ M) \\ &\overset{(b)}{=} \frac{1}{M}\,. \end{aligned}$$

where (a) uses the independence of X and Y, and (b) uses the fact that the sum is just summing $P(Y = y)$ over all possible values of y. Therefore, Z is uniform.

To prove the independence property, observe that for $y \in \{0, 1, 2, \ldots, M-1\}$, we have

$$\begin{aligned} P(Z = z,\, Y = y) &= P(X = z - y \ mod \ M,\, Y = y) \\ &= P(X = z - y \ mod \ M) \cdot P(Y = y) \\ &= \frac{1}{M} \cdot P(Y = y) \\ &= P(Z = z) \cdot P(Y = y)\,. \end{aligned}$$

Therefore, Z and Y are independent.

6. **[Presidential Poll]**

A couple of weeks before the presidential election, ECC News conducts a poll on the three candidates h, t, and j, based on two criteria: winning the election E via the electoral college, and winning the popular votes V . The joint probability distribution of the result is given below:

$$\begin{aligned} P_{EV}(E = h, V = h) &= 0.154 \\ P_{EV}(E = h, V = t) &= 0.179 \\ P_{EV}(E = h, V = j) &= 0.092 \\ P_{EV}(E = t, V = h) &= 0.254 \\ P_{EV}(E = t, V = t) &= 0.134 \\ P_{EV}(E = t, V = j) &= 0.016 \\ P_{EV}(E = j, V = h) &= 0.028 \\ P_{EV}(E = j, V = t) &= 0.001 \\ P_{EV}(E = j, V = j) &= 0.142 \end{aligned}$$

where E and V are the random variables; and h, t and j are the outcomes.

(a) Determine the marginal probability P_V.

Solution. *Summing up the suitable probabilities, we get*

P_V	h	t	j
	0.436	0.314	0.25

(b) Are the events candidate j winning the election ($E = j$), but candidate h winning the popular votes ($V = h$) independent? Give the reason

Solution. *No, they are not independent because $P(E = j, V = h) \neq P(E = j) \times P(V = h)$:*

$$\begin{aligned} P(E = j, V = h) &= 0.028\,, \\ P(E = j) \cdot P(V = h) &= 0.171 \times 0.436 = 0.074556\,. \end{aligned}$$

(c) Are the random variables E and V independent?

Solution. *No, they are not independent because $P(E = j, V = h) \neq P(E = j) \times P(V = h)$.*

(d) What is the probability that the candidate t wins the election, given that he does not win the popular vote?

Solution. *We have*

$$\begin{aligned} P(E = t \,|\, V \neq t) &= \frac{P(E = t, V \neq t)}{P(V \neq t)} \\ &= \frac{0.254 + 0.016}{0.436 + 0.25} \\ &\approx 0.3935\,. \end{aligned}$$

7. **[Sending Files Over a Noisy Channel]**

You are given a huge set of files. Your job is to send those to the ISS (International Space Station). Each of the files is encoded using one of the three different encoding algorithms: A, B, and C. Unfortunately, your lab partner forgot to annotate those encoded files based on the algorithms used after encoding them. All you know is one-third of the files are encoded using algorithm A, half using B, and one-sixth using C. It is known that using algorithm A, the chance that a file is received corrupted is 0.35. Similarly, if algorithm B or C is used, the probability is 0.32 and 0.13, respectively.

(a) You pick a file and send it to the ISS. What is the probability that the file is received uncorrupted?

Solution. *Let A, B, C denote the event the file is encoded using algorithm A, B, and C, respectively. Letting U be the event that a file is received uncorrupted, we have*

$$\begin{aligned} P(U) &= P(A) \cdot P(U|A) + P(B) \cdot P(U|B) + P(C) \cdot P(U|C) \\ &= \frac{1}{3} \cdot 0.65 + \frac{1}{2} \cdot 0.68 + \frac{1}{6} \cdot 0.87 \\ &= 0.216 + 0.34 + 0.145 \\ &= 0.701. \end{aligned}$$

(b) ISS reports that they successfully received an uncorrupted file. What is the probability that the file is encoded using algorithm B?

Solution. *Using Bayes' rule:*

$$\begin{aligned} P(B|U) &= \frac{P(B)\cdot P(U|B)}{P(A)\cdot P(U|A)+P(B)\cdot P(U|B)+P(C)\cdot P(U|C)} \\ &= \frac{\frac{1}{2}\cdot 0.68}{\frac{1}{3}\cdot 0.65+\frac{1}{2}\cdot 0.68+\frac{1}{6}\cdot 0.87} \\ &= \frac{0.34}{0.701} \\ &\approx 0.485. \end{aligned}$$

8. **(Advanced) [Proving Existence Properties Using Probability]**

(a) Let $\mathcal{M} = \{0, 1, \ldots, N^2 - 1\}$ for some positive integer N, and let $\mathcal{A}$ be a subset of $\mathcal{M}$ of size N. Show that there exists a subset $\mathcal{B}$ of $\mathcal{M}$ of size at most N such that the set $\mathcal{C} = \{a+b \mod N^2 | a \in \mathcal{A}, b \in \mathcal{B}\}$ has cardinality at least $\frac{N^2}{2}$.

(Hint: Place N items in the set $\mathcal{B}$ uniformly at random with replacement, and study the average number of integers in $\mathcal{M}$ but not $\mathcal{C}$. The inequality $(1-1/n)^n < \frac{1}{e}$ is also useful, as is linearity of expectation.)

Solution. *Place N items the set $\mathcal{B}$ uniformly at random with replacement to obtain the set $\mathcal{C}$. This means that during each of the N placements, any given $j \in \mathcal{M}$ is placed with probability $\frac{1}{N^2}$. There may be duplicates (since we are sampling* with *replacement), but nevertheless $\mathcal{B}$ clearly has size at most N as required.*

Let the random variable X be the number of integers in $\mathcal{M}$ but not in $\mathcal{C}$; and X_j be the indicator random variable for the event that integer j is not in $\mathcal{C}$. That is,

$$X_j = \begin{cases} 1, & \text{if } j \notin \mathcal{C} \\ 0, & \text{otherwise.} \end{cases}$$

Notice that for fixed j, there are N possible values of b that lead to j being in $\mathcal{C}$ (one for each value of $a \in \mathcal{A}$). The N trials are done independently, and the probability that $X_j = 1$ is the probability that none of those trials include any of those N possible values, which implies

$$\mathbb{E}[X_j] = \mathbb{P}[X_j = 1] = \left(1 - \frac{N}{N^2}\right)^N = \left(1 - \frac{1}{N}\right)^N < \frac{1}{e}\,.$$

By linearity of expectation,

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}\left[\sum_{j=1}^{N^2} X_j\right] \\ &= \sum_{j=1}^{N^2} \mathbb{E}\left[X_j\right] \\ &< \sum_{j=1}^{N^2} \frac{1}{e} \\ &= \frac{N^2}{e} \\ &\approx 0.368N^2\,. \end{aligned}$$

Therefore, choosing $\mathcal{B}$ at random, the number of integers in $\mathcal{C}$ on average is at least $(1-0.368)N^2 = 0.632N^2$. Since the average is more than half the total size, we can conclude that there exist a set $\mathcal{B}$ such that the set $\mathcal{C}$ has cardinality at least $\frac{N^2}{2}$.

(b) Let $x_1, x_2, \ldots, x_n$ be real numbers (not all zero) such that $x_1 + \cdots + x_n = 0$. Show that there is a permutation $y_1, \ldots, y_n$ of $x_1, \ldots, x_n$ such that $y_1 \cdot y_2 + y_2 \cdot y_3 + \cdots + y_n \cdot y_1 < 0$.
(Hint: Choose a random permutation and consider the expectation $\mathbb{E}[y_k \cdot y_{k+1}]$ of the product of a consecutive pair. Again use linearity of expectation.)

Solution. *Let $y = y_1, \ldots, y_n$ be a uniformly random permutation of $x_1, x_2, \ldots, x_n$ (i.e., all permutations are equally likely); and the random variable Y be the value of $y_1 \cdot y_2 + y_2 \cdot y_3 + \cdots + y_n \cdot y_1$. Take any pair of consecutive y's. Their product expectation,*

$$\begin{aligned}
\mathbb{E}[y_k \cdot y_{k+1}] &= \sum_{i \neq j} x_i \cdot x_j \; P(y_k = x_i,\, y_{k+1} = x_j) \\
&\overset{(a)}{=} \frac{1}{n(n-1)} \sum_{i \neq j} x_i \cdot x_j \\
&= \frac{1}{n(n-1)} \sum_i x_i \sum_{j \neq i} x_j \\
&\overset{(b)}{=} \frac{1}{n(n-1)} \sum_i x_i(-x_i) \\
&= -\frac{1}{n(n-1)} \sum_i x_i^2 \,.
\end{aligned}$$

where (a) follows since any two given values in a uniformly random permutation are uniformly distributed over the $n(n-1)$ possible (ordered) pairs, and (b) uses the assumption $x_1 + \ldots + x_n = 0$. By linearity of expectation, (and using the notation $y_{n+1} \equiv y_1$ to simplify the presentation a little)

$$\begin{aligned}
\mathbb{E}[Y] &= \mathbb{E}\left[\sum_{k=1}^{n} y_k \cdot y_{k+1}\right] \\
&= \sum_{k=1}^{n} \mathbb{E}[\, y_k \cdot y_{k+1}\,] \\
&= \sum_{k=1}^{n} -\frac{1}{n(n-1)} \sum_i x_i^2 \\
&= n\left(-\frac{1}{n(n-1)} \sum_i x_i^2\right) \\
&= -\frac{1}{(n-1)} \sum_i x_i^2 \\
&< 0\,.
\end{aligned}$$

Note that the last inequality is strict, because we assumed that not all of the x_i values are zero. Since the average Y value of random permutations y is negative, we can conclude that there must exist one specific permutation giving a negative value.

9. [**Repetition Repetition Repetition Code**]

Consider the repetition code R_4 over the alphabet $\mathcal{A} = \{1, 2, 3\}$, and the majority-vote decoder (maximum likelihood decoder).

(a) List all the received codewords that will be uniquely decoded as 2222. (Here the "unique" requirement means that neither 1111 nor 3333 are equally far from the received sequence compared to 2222)

Solution.

2222,

1222, 3222, 2122, 2322, 2212, 2232, 2221, 2223,

1322, 3122, 1232, 3212, 1223, 3221, 2132, 2312, 2123, 2321, 2213, 2231.

(b) Suppose the probability of each symbol being flipped to another symbol is f, equally likely for each symbol. An error is said to occur when a codeword is incorrectly decoded. If we adopt a pessimistic viewpoint and assume that an error always occurs when there is a tie in decoding (e.g., 1122 is received, which is equally close to 1111 and 2222), then what is the probability of error p_b of R_4?

Solution. *Without lost of generality, suppose the symbol 1 is transmitted, i.e.* $\mathbf{s} = 1$, $\mathbf{t} = 1111$.

By assumption, we have P(1 is received) $= 1 - f$ *and P(3 or 2 is received)* $= f$*. Therefore, P(3 is received)* $=$ *P(2 is received)* $= \frac{f}{2}$.

The probability of correctly decoding the codeword is therefore

$$\begin{aligned}(1-f)^4 \;+\; 8\,\frac{f}{2}(1-f)^3 \;+\; 12\left(\frac{f}{2}\right)^2(1-f)^2 \\ =\; (1-f)^2\,(1+2f)\,.\end{aligned}$$

where the numbers 8 and 12 come from counting the sequences in part (a).

Therefore, the probability of error is

$$\begin{aligned}p_b \;&=\; 1-(1-f)^2(1+2f) \\ &=\; f^2\,(3-2f)\,.\end{aligned}$$