Probability 1 - Hw 1

Daniel Yu

September 13,2024

1. Prove $P(\cup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i)$.

   Proof. Let each event $A_i \subseteq \Omega$ and the set $\{A_i\}_{i\in 1,\ldots,n}$ be a finite collection of subsets of Ω. If each A_i is disjoint, then we are done from the definition of a probability measure property 3. If A_i is not disjoint, then there exists $x \in A_i \cap A_j$ for two distinct subsets A_i, A_j. Let the set $X = A_i \cap A_j$. Start with the base case where $n = 2$ and there are two events A_1, A_2. Then, $P(A_1 \cup A_2) = P(A_1 \setminus X \cup A_2 \setminus X \cup X) = P(A_1 \setminus X \cup A_2) = P(A_1 \setminus X) + P(A_2) \leq P(A_1) + P(A_2)$ since each of $A_1 \setminus X, A_2 \setminus X, X$ are disjoint sets by construction. Assume that the hypothesis holds for some k, $P(\cup_{i=1}^k A_i) \leq \sum_{i=1}^k P(A_i)$. Then for $k+1$, $P(\cup_{i=1}^{k+1} A_i) = P(\cup_{i=1}^k A_i \cup A_{k+1})$. If A_{k+1} is disjoint from $\cup_{i=1}^k A_i$ then $P(\cup_{i=1}^k A_i \cup A_{k+1}) = P(\cup_{i=1}^k A_i) + P(A_{k+1}) \leq \sum_{i=1}^k P(A_i) + P(A_{k+1}) = \sum_{i=1}^{k+1} P(A_i)$ and we are done. If A_{k+1} is not disjoint, then a similar argument as from the base case holds. Take $X = A_{k+1} \cap \cup_{i=1}^k A_i$. Then, $P(\cup_{i=1}^k A_i \cup A_{k+1}) = P(\cup_{i=1}^k A_i \setminus X \cup A_{k+1} \setminus X \cup X) = P(\cup_{i=1}^k A_i \cup A_{k+1} \setminus X) = P(\cup_{i=1}^k A_i) + P(A_{k+1} \setminus X) \leq \sum_{i=1}^k P(A_i) + P(A_{k+1} \setminus X) \leq \sum_{i=1}^k P(A_i) + P(A_{k+1}) = \sum_{i=1}^{k+1} P(A_i)$. QED $\square$

2. What is the value of c_1, c_2?

   Note.

$$\begin{aligned}\int_{-\infty}^{\infty} f_X(s)ds &= \int_0^4 \frac{c_1}{\sqrt{s}} + c_2 s ds \\ &= [2c_1 s^{\frac{1}{2}} + \frac{1}{2}c_2 s^2]_0^4 \\ &= 2 \cdot 2 \cdot c_1 + 8 \cdot c_2 \\ &= 1.\end{aligned}$$

   We are also given:

$$P[X \leq 1] = \frac{1}{3}.$$

   So,

$$\begin{aligned}P[X \leq 1] = \int_0^1 f_X(s)ds &= [2c_1 s^{\frac{1}{2}} + \frac{1}{2}c_2 s^2]_0^1 \\ &= 2 \cdot c_1 + \frac{1}{2} \cdot c_2 \\ &= \frac{1}{3}.\end{aligned}$$

   Solving the system and substituting $c_2 = \frac{2}{3} - 4c_1$:

$$\begin{aligned}4 \cdot c_1 + 8 \cdot c_2 &= 4 \cdot c_1 + 8 \cdot (\frac{2}{3} - 4c_1) \\ &= 4c_1 + \frac{16}{3} - 32c_1 \\ &= -28c_1 + \frac{16}{3} \\ &= 1. \\ c_1 &= \frac{13}{84}.\end{aligned}$$

   and,

$$c_2 = \frac{1}{21}.$$

3. Let N be a random variable that takes non-negative integer values. Show

$$E[N] = \sum_{n=1}^{\infty} P[N \geq n].$$

Proof.

$$\sum_{n=1}^{\infty} P[N \geq n] = P[N \geq 1] + P[N \geq 2] + P[N \geq 3] + \ldots$$
$$= (P[N=1] + P[N=2] + P[N=3] + \ldots) + (P[N=2] + P[N=3] + \ldots) + (P[N=3] + \ldots) + \ldots$$
$$= P[N=1] + 2P[N=2] + 3P[N=3] + \ldots + \sum_{n=1}^{k} P[n=k] + \ldots$$
$$= \sum_{k=1}^{\infty} \sum_{n=1}^{k} P[N=k]$$
$$= \sum_{k \in range(N)} kP[n=k]$$
$$= E[N].$$

□

4. Consider two random variables X, Y:

   (a) What is the probability that $X = Y$?

   Note. We know that $P[X=2,Y=3]+P[X=2,Y=1]+P[X=2,Y=2] = P[X=2] = \frac{1}{3}$ and $P[X=2,Y=2]+P[X=1,Y=2]+P[X=3,Y=2] = P[Y=2] = \frac{1}{3}$. Substituting in the give values:

$$\frac{1}{12} + P[X=2,Y=1] + P[X=2,Y=2] = \frac{1}{3}.$$
$$P[X=2,Y=2] + P[X=1,Y=2] + \frac{1}{4} = \frac{1}{3}.$$

   We also know $P[X=1,Y=1]+P[X=2,Y=1]+P[X=3,Y=1] = P[Y=1] = \frac{1}{3}$ and $P[X=1,Y=1]+P[X=1,Y=2]+P[X=1,Y=3] = P[X=1] = \frac{1}{3}$.

$$\frac{1}{12} + P[X=2,Y=1] + P[X=3,Y=1] = \frac{1}{3}.$$
$$\frac{1}{12} + P[X=1,Y=2] + P[X=1,Y=3] = \frac{1}{3}.$$

   . Finally, we know:

$$P[X=3,Y=1] + \frac{1}{4} + P[X=3,Y=3] = \frac{1}{3}.$$
$$P[X=1,Y=3] + \frac{1}{12} + P[X=3,Y=3] = \frac{1}{3}.$$

   There are 6 equations and 6 unknowns, so we can solve this linear system.

$$P[X=2, Y=1] = a$$

$$\boxed{P[X=2, Y=2] = b}$$

$$P[X=1, Y=2] = c$$

$$P[X=3, Y=1] = d$$

$$P[X=1, Y=3] = e$$

$$\boxed{P[X=3, Y=3] = f}$$

$$\begin{array}{c|c}
 & \begin{array}{cccccc} a & b & c & d & e & f \end{array} \\
\begin{array}{c} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \end{array} &
\left[\begin{array}{cccccc}
1 & 1 & & & & \\
 & 1 & 1 & & & \\
1 & & & 1 & & \\
 & & 1 & & 1 & \\
 & & & 1 & & 1 \\
 & & & & 1 & 1
\end{array}\right]
\end{array}
\left|\begin{array}{c} 1/4 \\ 1/12 \\ 1/4 \\ 1/4 \\ 1/12 \\ 1/4 \end{array}\right]
\quad r_2 - r_4 \Longrightarrow$$

We want $b + f = ?$

Figure 1

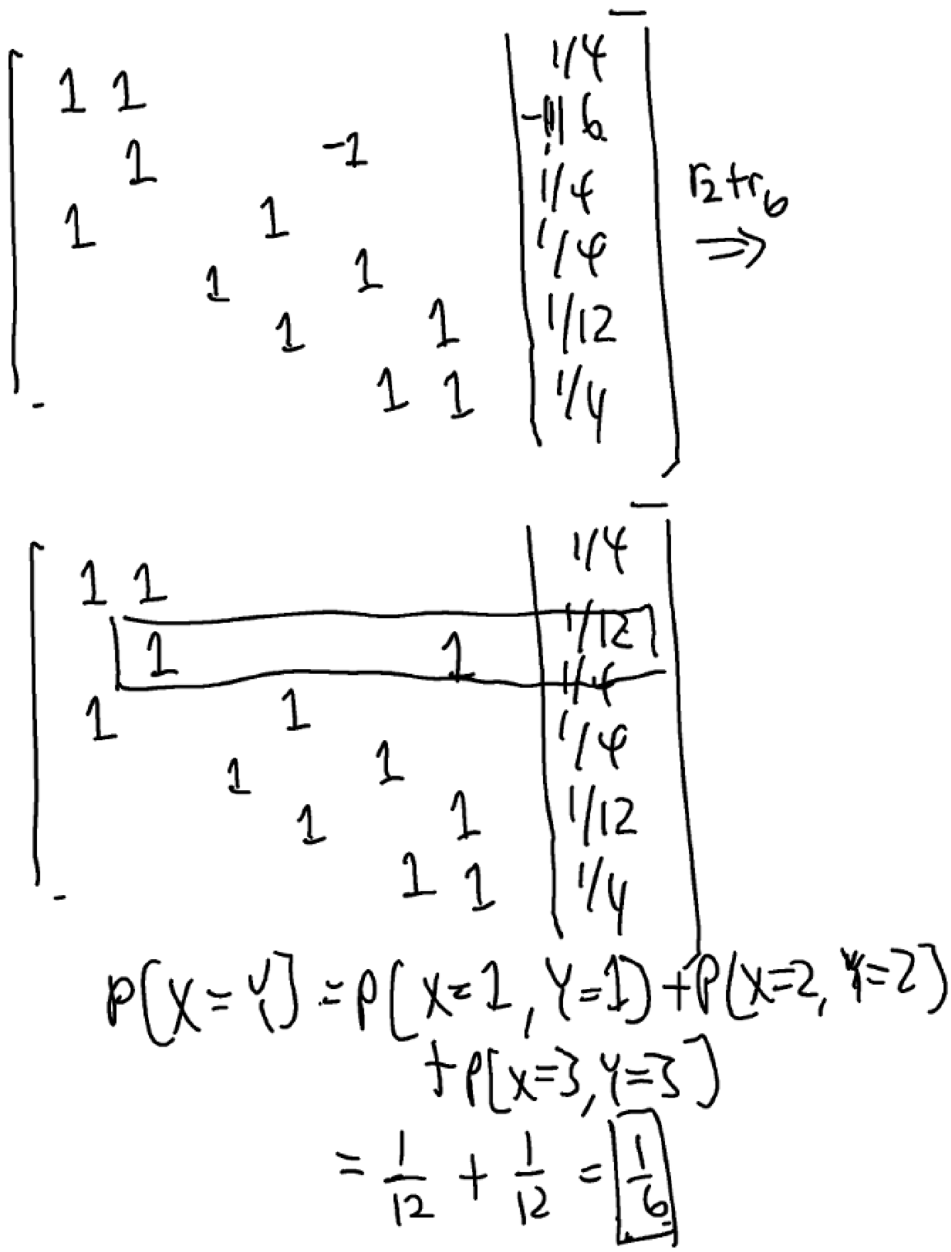

Figure 2

Note that the matrix is not full rank and the rows and columns of the matrix are not linearly independent!

(b) What is the smallest possible value of $P[X = 2, Y = 1]$.

Note. From the matrix we can extract the following equations where $P[x = 2, Y = 1] = a$

$$a + b = \frac{1}{4}.$$

$$b + f = \frac{1}{12}.$$

We know that probabilities have to be non-negative so $a, b, c, d, e, f \geq 0$. Given this constraint, let $f = 0$, then maximum possible value for b is:

$$b + 0 = \frac{1}{12} \rightarrow b = \frac{1}{12}.$$

and

$$a + \frac{1}{12} = \frac{1}{4} \to a = \frac{1}{6}.$$

Thus, the smallest possible value of $a = \frac{1}{6}$.

5. Consider stick of length 1. We break along a uniformly chosen point on the stick. Let L be the longer piece's length:

   (a) What is the PDF $f_L(s)$?

   Note. Start by considering the CDF: $F_L(s) = P[L \leq s]$ where s could be any value for the length of the longer piece. First, notice that since it is the length of the longer piece, $L \geq .5$ and $s \geq .5$. So,

$$F_L(s) = \begin{cases} P[L \leq s], 1 \geq s \geq .5 \\ 0, \text{otherwise} \end{cases}.$$

   Imagine the stick as the interval $[0, 1]$. If we choose the point at the distance .1 as the cut or the point at the distance .9 as the cut, the length of the longer piece would still be .9. In fact over the entire interval $[.1, .9]$, the length of the longer piece $L \leq .9$. Since we know that the stick is distributed uniformly along the interval $[0, 1]$,

$$F_L(s) = \begin{cases} \frac{s-(1-s)}{1} = 2s - 1, 1 \geq s \geq .5 \\ 0, \text{ otherwise} \end{cases}.$$

   Then to take the pdf, we would simply take the first derivative:

$$\frac{d}{ds} F_L(s) = \frac{d}{ds}(2s - 1) = 2.$$

   . Thus, the pdf would simply be

$$f_L(s) = \begin{cases} 2, 1 \geq s \geq .5 \\ 0, \text{otherwise} \end{cases}.$$

   (b) What is the expected value of L?

$$E[L] = \int_{-\infty}^{\infty} s \cdot f_L(s) ds = \int_{.5}^{1} s \cdot 2 ds = [s^2]_{.5}^{1}.$$

$$= .75.$$

   (c) What is the variance of L?

$$\begin{aligned} var(L) &= E[L^2] - (E[L])^2 \\ &= \int_{-\infty}^{\infty} s^2 \cdot f_L(s) ds - (.75)^2 \\ &= \int_{.5}^{1} s^2 \cdot 2 - .5625 \\ &= [\frac{2}{3} s^3]_{.5}^{1} - .5625 \\ &= [\frac{2}{3} - \frac{1}{12}] - .5625 \\ &= \frac{1}{48}. \end{aligned}$$