C.I Discrete Random Variables

PROBLEM 1. The Im Z is discrete because Im X and Im Y are discrete sets.

Let $z \in \mathbb{R}$. If $\{Z = z\} = \varnothing$, then $\{Z = z\}$ is an event because the set $\varnothing$ is always an event. Assume that $\{Z = z\} \neq \varnothing$. We have to consider two cases.

i) $\underline{z = 0.}$ In this case, the only way that $Z(s) = 0$ is if $X(s) = 0$ or $Y(x) = 0$. Therefore,

$$\{Z = 0\} = \{X = 0\} \cup \{Y = 0\}.$$

Since $\{X = 0\}$ and $\{Y = 0\}$ are events, we conclude that $\{Z = 0\}$ are events (recall that, by assumption, X and Y are discrete random variables).

ii) $\underline{z \neq 0.}$ In this case, the functions X and Y can't take the value 0. If $s \in \{Z = z\}$, then $X(s)Y(s) = Z(s) = z$. Therefore $X(s) = z/Y(s)$. Let $y = Y(s)$. Then $X(s) = z/y$ and $Y(s) = y$. In other words, $s \in \{X = z/y\} \cap \{Y = y\}$. On the other hand, if $s \in \{X = z/y\} \cap \{Y = y\}$, then $X(s) = z/y$ and $Y(s) = y$. Therefore, $Z(s) = X(s)Y(s) = (z/y)y = z$ and then $s \in \{Z = z\}$. In summary, we have just proved that

$$\{Z = z\} = \bigcup_{y \in \operatorname{Im} Y, y \neq 0} \Big(\{X = z/y\} \cap \{Y = y\} \Big).$$

For a given $z \in \mathbb{R}$, $z \neq 0$ and $y \in \operatorname{Im} Y$, the event $\{X = z/y\} \cap \{Y = y\}$ is an event because X and Y are discrete random variable and $\mathcal{A}$ is an event space. Thus, a countable union of these events will remain an event. Hence, $\{Z = z\}$ is an event.

In each case, $\{Z = z\}$ is an event. The map Z satisfies condition (a) and (b) in Definition 1 and therefore Z is a discrete random variable. △

PROBLEM 2. We have $\operatorname{Im}(1_A) = \{0, 1\}$, which is a finite set (therefore discrete).

Let $x \in \mathbb{R}$. We have three cases to consider.

1. $\underline{x = 0.}$ In this case, $\{1_A = 0\} = \overline{A}$. Since A is an event, we know that $\overline{A}$ is also an event. Therefore, $\{1_A = 0\}$ is an event.
2. $\underline{x = 1.}$ In this case, $\{1_A = 1\} = A$ and A is an event. Hence, $\{1_A = 1\}$ is an event.
3. $\underline{x \neq 0 \text{ and } x \neq 1.}$ In this case, $\{1_A = x\} = \varnothing$ because there is no s such that $1_A(x) = x$ (the only possible values are 0 and 1 for 1_A). Since $\varnothing$ is an event, $\{1_A = x\}$ is an event. △

PROBLEM 3.

a) Let $x \in \mathbb{R}$. If $\{X \leq x\} = \varnothing$, then $\{X \leq x\}$ is an event. Assume that $\{X \leq x\} \neq \varnothing$. Since X is a discrete random variable, the set Im X is discrete. This means there are only a countable values of Im X that can be smaller than the number x. List them in decreasing order, say x_1, x_2, x_3, ..., with $x_i \geq x_j$, when $i \leq j$ and $x_j \leq x$ for any j. Therefore, we can write

$$\{X \leq x\} = \bigcup_{j=1}^{\infty} \{X = x_j\}.$$

The map X is a discrete random variable. Therefore, each set $\{X = x_j\}$ is an event and this implies that $\cup_{j=1}^{\infty}\{X = x_j\}$ is an event. Hence, $\{X \leq x\}$ is an event.

b) Let $x \in \mathbb{R}$. We can write

$$\{X < x\} = \{X \leq x\} \cap \overline{\{X = x\}}.$$

In other words, the set $\{X < x\}$ is the set of $s \in S$ that belong to $\{X \leq x\}$ but are not in $\{X = x\}$. From part a), the set $\{X \leq x\}$ is an event and from the fact that X is assumed to be a discrete random variable, $\{X = x\}$ is an event. Therefore, $\{X \leq x\} \cap \overline{\{X = x\}}$ is an event and hence $\{X < x\}$ is an event.

c) We have

$$\{X \geq x\} = \overline{\{X < x\}}.$$

From part b), we know that $\{X < x\}$ is an event, hence $\{X \geq x\}$ is also an event.

d) We have

$$\{X > x\} = \overline{\{X \leq x\}}.$$

From part c), we know that $\{X \leq x\}$ is an event, hence $\{X > x\}$ is also an event.

PROBLEM 4. Assume that X is a discrete random variable. Then $\operatorname{Im} X$ is discrete and $\{X = x\}$ is an event for every $x \in \mathbb{R}$. From Problem 3, part a), the set $\{X \leq x\}$ is an event. Therefore, conditions a) and b) in the statement are satisfied.

Assume that the two conditions in the statement are satisfied. Then, in particular, $\operatorname{Im} X$ is discrete. Also, for an $x \in \mathbb{R}$, the set $\{X > x\}$ is an event because it is the complement of the event $\{X \leq x\}$. Also, for an $x \in \mathbb{R}$, we have

$$\{X < x\} = \bigcup_{j=1}^{\infty} \Big\{X \leq x - \frac{1}{j}\Big\}.$$

This is a countable unions of the events $\{X \leq x - \frac{1}{j}\}$ and therefore $\{X < x\}$ is an event. But also $\{X \geq x\}$ is also an event because it is the complement of $\{X < x\}$. Let $x \in \mathbb{R}$. We can write

$$\{X = x\} = \{X \leq x\} \cap \{X \geq x\},$$

the intersection of two events! So $\{X = x\}$ is also an event. Hence X is a discrete random variable.

C.II Probability Mass Functions

Problem 5. The probability measure P on S is given by

$$P(\{r\}) = \frac{2}{5}, \; P(\{b\}) = \frac{2}{5}, \quad P(\{y\}) = \frac{1}{5}.$$

The function $X : S \to \mathbb{R}$ is given by $X(\{r\}) = -10$, $X(\{b\}) = 10$, and $X(\{y\}) = 20$. Therefore, we have

- $p_X(-10) = P(X = -10) = P(\{r\}) = \frac{2}{5}$.
- $p_X(10) = P(X = 10) = P(\{b\}) = \frac{2}{5}$.
- $p_X(20) = P(X = 20) = P(\{y\}) = \frac{1}{5}$.
- $p_X(x) = 0$, for $x \neq -10, 10, 20$.

Problem 6. A child may or may not identify properly the picture. Let w_1, w_2, w_3 be the words corresponding to the animal in picture p_1, p_2, p_3. A child will identify correctly a picture if the word w_i put under the picture p_i. Therefore, we can identify an outcome as an ordered list of three symbols from $\{w_1, w_2, w_3\}$. For example, $w_1w_2w_3$ means that the child identified the animal in picture p_1 as w_1, in picture p_2 as w_2, and in picture p_3 as w_3. Therefore, the sample space is

$$S = \{w_1w_2w_3, w_1w_3w_2, w_2w_1w_3, w_3w_2w_1, w_3w_1w_2, w_2w_1w_3\}.$$

If we just keep the numbers

$$S = \{123, 132, 213, 321, 312, 231\}.$$

Each outcome are equally likely to happen, so with $1/6$ chance.

Let $Y : S \to \mathbb{R}$. Notice that, if the child successfully matches 2 pictures with their words, then the third picture will be also successfully matched. Therefore, the child may correctly identify 0, 1, or 3 of the pictures presented and $\operatorname{Im} Y = \{0, 1, 3\}$. Then, we have $p_Y(y) = 0$ for any $y \neq 0, 1, 3$. For the other values of y:

- $p_Y(0) = P(Y = 0) = P(\{312, 231\}) = 1/3$.
- $p_Y(1) = P(Y = 1) = P(\{132, 213, 312\}) = \frac{1}{2}$.
- $p_Y(3) = P(Y = 3) = P(\{123\}) = \frac{1}{6}$.

Problem 7. The sample space is all distinct subsets of two numbers from $\{1, 2, 3, 4, 5\}$. There are $\binom{5}{2} = 10$ possible outcomes and all of the outcome are equally likely to occur.

1. We have $\operatorname{Im} X = \{2, 3, 4, 5\}$. The number 1 is missing because in the two balls selected, if ball #1 is selected, then the other ball's number is automatically one of 2, 3, 4, 5. We will present the pmf of X in a table.

x	2	3	4	5
$p_X(x)$	1/10	1/5	3/10	2/5

   To compute $p_X(2)$, we first notice that $\{X = 2\} = \{\{1, 2\}\}$ and therefore $P(X = 2) = 1/10$. To compute $p_X(3)$, we first notice that $\{X = 3\} = \{\{1, 3\}, \{2, 3\}\}$ and therefore $P(X = 3) = 2/10 = 1/5$. Similar calculations lead to the values of $p_X(4)$ and $p_X(5)$. Notice that $1/10 + 1/5 + 3/10 + 2/5 = 1$.

2. Removing the parenthesis in the set and considering them as unordered list, the outcomes of S can be explicitly enumerated:

$$S = \{12, 13, 14, 15, 23, 24, 25, 34, 35, 45\}.$$

Therefore, considering all the outcomes and adding the numbers, we see that $\operatorname{Im} X = \{3, 4, 5, 6, 7, 8, 9\}$. We have, more precisely, $X(12) = 3$, $X(13) = 4$, $X(14) = X(23) = 5$, $X(24) = X(15) = 6$, $X(25) = X(34) = 7$, $X(35) = 8$, $X(45) = 9$. Using the same strategy as in a), we find the following values for p_X.

x	3	4	5	6	7	8	9
$p_X(x)$	1/10	1/10	1/5	1/5	1/5	1/10	1/10

Problem 8. If p is a probability mass function, then we know it should satisfy $\sum_{k=1}^{\infty} p(k) = 1$. This gives the following condition:

$$\sum_{k=1}^{\infty} \frac{c}{k(k+1)} = 1.$$

Now, using the trick $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$, we see that the series $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$ is convergent and

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \lim_{N\to\infty} \sum_{k=1}^{N} \left(\frac{1}{k} - \frac{1}{k+1}\right) = \lim_{N\to\infty} 1 - \frac{1}{N+1} = 1.$$

Therefore, using the properties of series, we see that

$$\sum_{k=1}^{\infty} \frac{c}{k(k+1)} = 1 \iff c\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1 \iff c \cdot 1 = 1 \iff c = 1.$$

This means the function p is a pmf if and only if $c = 1$.

C.III Functions of Discrete Random Variables

Problem 9. Setting $X = 1$, $X = 2$, $X = 3$, and $X = 4$ in the expression of Y, we get $Y = 0, 3, 8, 15$. Therefore, $\operatorname{Im} Y = \{0, 3, 8, 15\}$.

Using Theorem 3, with $g(x) = x^2 - 1$, we have

$$p_Y(y) = \sum_{x\in g^{-1}(y)} P(X = x).$$

For

- $y = 0$, we have $g^{-1}(0) = \{x \in \operatorname{Im} X \,:\, g(x) = 0\} = \{1\}$;
- $y = 1$, we have $g^{-1}(3) = \{x \in \operatorname{Im} X \,:\, g(x) = 3\} = \{2\}$;
- $y = 8$, we have $g^{-1}(8) = \{x \in \operatorname{Im} X \,:\, g(x) = 8\} = \{3\}$;
- $y = 15$, we have $g^{-1}(15) = \{x \in \operatorname{Im} X \,:\, g(x) = 15\} = \{4\}$.

Therefore,

- $p_Y(0) = P(X = 1) = 0.4$.
- $p_Y(3) = P(X = 2) = 0.3$.
- $p_Y(8) = P(X = 3) = 0.2$.
- $p_Y(15) = P(X = 4) = 0.1$.
- $p_Y(y) = 0$ for any other values y different from $0, 3, 8, 15$. △

PROBLEM 10. Setting $X = 1, 2, 3, 4$ in the expression of Y, we get $Y = 1, 0, -1, 0$ respectively. Therefore, $\operatorname{Im} Y = \{-1, 0, 1\}$.

Using Theorem 3 again, but with $g(x) = \sin(\frac{\pi}{2}x)$, we have

$$p_Y(y) = \sum_{x \in g^{-1}(y)} P(X = x).$$

For

- $y = -1$, $g^{-1}(-1) = \{3\}$;
- $y = 0$, $g^{-1}(0) = \{2, 4\}$;
- $y = 1$, $g^{-1}(1) = \{1\}$.

Therefore,

- $p_Y(-1) = P(X = 3) = 0.2$.
- $p_Y(0) = P(X = 2) + P(X = 4) = 0.3 + 0.1 = 0.4$.
- $p_Y(1) = P(X = 1) = 0.4$. △

C.IV Expectation and Variance

PROBLEM 11. Let t_1 be the label "the dimensions of the trailer are $8 \times 10 \times 30$. and let t_2 be the label "the dimensions of the trailer are $8 \times 10 \times 40$. Given a trailer, the possible outcome is the trailer is of type t_1 or of type t_2. Therefore, $S = \{t_1, t_2\}$ with $P(\{t_1\}) = 0.3$ and $P(\{t_2\}) = 0.7$.

Let X be the map given the volume of a trailer. We have

$$X(t_1) = 8 \cdot 10 \cdot 30 = 2400 \quad \text{and} \quad X(t_2) = 8 \cdot 10 \cdot 40 = 3200.$$

Therefore, we get

$$\operatorname{Exp}(X) = X(t_1)P(X = 2400) + X(t_2)P(X = 3200) = (2400)(0.3) + (3200)(0.7) = 2960.$$

The average volume shipped per trailer load is 2960ft^3. △

PROBLEM 12. A firm can be assign one or the two contracts. Therefore, we can generate the set of outcomes as couple of letters taken from $\{a, b, c\}$. For example AA means A was assigned to the two contracts, but AB or BA means that A and B was assigned to one of the contracts. The sample space S is

$$S = \{aa, ab, ba, ac, ca, bb, bc, cb, cc\}.$$

Since the firms are assigned a contract at random, each outcome are equally likely to occur, so with $1/9$.

a) In the first scenario, assume that X is the possible profit made by firm A after the contracts were assigned. Therefore, this means

- $X(aa) = 180,000$.
- $X(ab) = X(ba) = X(ac) = X(ca) = 90,000$.
- $X(bb) = X(bc) = X(cb) = X(cc) = 0$.

The expectation is then calculated as followed:

$$\begin{aligned}
\mathrm{Exp}(X) &= 180,000P(X = 180,000) + 90,000P(X = 90,000) + 0P(X = 0) \\
&= 180,000P(\{aa\}) + 90,000(P(\{ab, ba, ac, ca\}) + 0 \\
&= \frac{180,000}{9} + \frac{90,000 \cdot 4}{9} \\
&= 20,000 + 40,000 \\
&= 60,000
\end{aligned}$$

b) Let Y be the possible possible made by firms A and B after the contrasts were assigned. Therefore, this means

- $X(aa) = X(bb) = X(ab) = X(ba) = 180,000$.
- $X(ac) = X(ca) = X(bc) = X(cb) = 90,000$.
- $X(cc) = 0$.

The expectation is then calculated as followed:

$$\begin{aligned}
\mathrm{Exp}(X) &= 180,000P(X = 180,000) + 90,000P(X = 90,000) + 0P(X = 0) \\
&= 180,000P(\{aa, bb, ab, va\}) + 90,000P(\{ac, ca, bc, cb\}) \\
&= \frac{(180,000)(4)}{9} + \frac{(90,000)(4)}{9} \\
&= 80,000 + 40,000 \\
&= 120,000.
\end{aligned}$$

△

PROBLEM 13. We have $\mathrm{Im}\, X = \{1, 2, 3, 4, 5, 6\}$ and each value of X has a chance of $1/6$ to occur. Therefore,

$$\mathrm{Exp}(X) = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = \frac{21}{6} = 3\tfrac{1}{2}.$$

The variance is calculated using formula in the Theorem 6. We first have $\mathrm{Im}\, X^2 = \{1, 4, 9, 16, 25, 36\}$ and

$$\mathrm{Exp}(X) = \frac{1}{6} + \frac{4}{6} + \frac{9}{6} + \frac{16}{6} + \frac{25}{6} + \frac{36}{6} = \frac{91}{6} = 30\tfrac{1}{6}.$$

Therefore,

$$\mathrm{Var}(X) = \mathrm{Exp}(X^2) - (\mathrm{Exp}(X))^2 = \frac{91}{6} - \frac{21}{6} = \frac{70}{6} = 11\tfrac{2}{3}.$$

Thus,

$$\sigma = \sqrt{\mathrm{Var}(X)} = \sqrt{70/6} \approx 3.4157.$$

△

Problem 14. By the formula in Theorem 6, we have

$$\mathrm{Var}(aX+b) = \mathrm{Exp}((aX+b)^2) - (\mathrm{Exp}(aX+b))^2.$$

We have $(aX+b)^2 = a^2X^2 + 2abX + b^2$ and $\mathrm{Exp}(aX+b) = a\mathrm{Exp}(X)+b$. Therefore,

$$\begin{aligned}\mathrm{Var}(aX+b) &= a^2\mathrm{Exp}(X^2) + 2ab\mathrm{Exp}(X) + b^2 - a^2(\mathrm{Exp}(X))^2 - 2ab\mathrm{Exp}(X) - b^2\\ &= a^2\mathrm{Exp}(X^2) - a^2\mathrm{Exp}(X)\\ &= a^2\mathrm{Var}(X).\end{aligned}$$

△

C.V Conditional Expectation and the Partition Theorem

Problem 15. Let $B = \{X = x\}$. Then, we have

$$E(g(X)|B) = \sum_{y\in \mathrm{Im}\, g(X)} yP(g(X) = y|B).$$

However, if B has occurred, then $X = x$ and the sum over $\mathrm{Im}\, g(X)$ is restricted to the value $y = g(x)$. Hence,

$$E(g(X)|B) = g(x)P(g(X) = g(x)) = g(x)P(X = x).$$

△

C.VI Examples of Discrete Random Variables

Problem 16. The map X has a discrete range, that is $\mathrm{Im}\, X = \{0, 1, 2, 3, \ldots, 30\}$. However it does not have a binomial distribution because the probability that there is rain on a given day varies from day to day. Therefore, the parameter p is not fixed. △

Problem 17.

a) In this case, if it was explicitly mentioned "The number of students in a sample of X students who took the SAT", then we could model the distribution of X on the binomial distribution with $n = 100$ and $q = 0.45$. Unfortunately, it is not mentioned and therefore we can't model the distribution with a binomial distribution.

b) It can't be model by a binomial distribution because there is not enough information to find the parameter q. We will see later that the distribution of the scores of the 100 students can be model by a normal distribution.

c) If X_j is the random variable "The student labeled j scored above average on the SAT", then X: "the number of students in the sample who scored above average on the SAT", which is equal to $X_1 + X_2 + \ldots + X_{100}$, has a binomial distribution. In this case, $n = 100$ and the value of q is not possible to find. We would need more information to compute an approximate value for q. For example, with the additional assumption that the distribution of the student's scores is a Normal distribution, then we can assume that $q = 0.5$, because $P(X > \mu) = 0.5$ for any normal distribution.

d) It can't be model by a binomial distribution because there is not enough information to find the parameter q. We would need additional information on the average time of a student to complete the test. We will see later that the distribution of the random variable will be modeled by a Normal distribution. △

PROBLEM 18. By Definition 3, we have

$$\mathrm{Exp}(X) = \sum_{k=0}^{n} kP(X = k) = \sum_{k=0}^{n} k\frac{n!}{k!(n-k)!}q^k(1-q)^{n-k}.$$

The term with $k = 0$ disappears and we enter into the following chain of equalities:

$$\begin{aligned}
\mathrm{Exp}(X) &= \sum_{k=1}^{n} \frac{kn!}{k!(n-k)!}q^k(1-q)^{n-k} \\
&= \sum_{k=0}^{n-1} \frac{(k+1)n!}{(k+1)!(n-k-1)!}q^{k+1}(1-q)^{n-k-1} \\
&= nq\sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-1-k)!}q^k(1-q)^{n-k-1} \\
&= nq(q+(1-q))^{n-1} \\
&= nq.
\end{aligned}$$

To compute the $\mathrm{Var}(X)$, we use the formula Theorem 6. We have $\mathrm{Exp}(X) = nq$ from the previous calculations. We need to compute $\mathrm{Exp}(X^2)$:

$$\mathrm{Exp}(X^2) = \sum_{k=0}^{n} k^2 P(X^2 = k^2) = \sum_{k=0}^{n} k^2 P(X = k),$$

where $\{X^2 = k^2\} = \{X = k\}$ because X assumes only non-negative integer values. Therefore,

$$\begin{aligned}
\mathrm{Exp}(X^2) &= \sum_{k=0}^{n} \frac{k^2 n!}{k!(n-k)!}q^k(1-q)^{n-k} \\
&= \sum_{k=1}^{n} \frac{k^2 n!}{k!(n-k)!}q^k(1-q)^{n-k} \\
&= nq\sum_{k=0}^{n-1} \frac{(k+1)(n-1)!}{k!(n-1-k)!}q^k(1-q)^{n-1-k} \\
&= nq\Big(\sum_{k=0}^{n-1} \frac{k(n-1)!}{k!(n-1-k)!}q^k(1-q)^{n-1-k} + \sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-1-k)!}q^k(1-q)^{n-1-k}\Big) \\
&= nq\Big(\sum_{k=1}^{n-1} \frac{k(n-1)!}{k!(n-1-k)!}q^k(1-q)^{n-1-k} + 1\Big) \\
&= nq\Big((n-1)q\sum_{k=0}^{n-2} \frac{(n-2)!}{k!(n-2-k)!}q^k(1-q)^{n-2-k} + 1\Big) \\
&= nq\Big((n-1)q + 1\Big) \\
&= nq(nq - q + 1) \\
&= n^2q^2 + nq(1-q).
\end{aligned}$$

Hence,

$$\mathrm{Var}(X) = n^2q^2 + nq(1-q) - n^2q^2 = nq(1-q). \qquad \triangle$$