# Math 170E Week 3

Aidan Jan

February 5, 2023

## Gaining additional information

**Example:**
In 1958, the United States lost a nuclear submarine (SSN-589) and all its crew. Since the submarine was nuclear powered, they wanted to find the submarine to avoid damaging the environment.

Suppose in an example, there are three regions where the submarine could have been lost. Name these regions 1, 2, and 3.

Let $a_k = \mathbb{P}(\text{The sub was lost in region k})$, and let $b_k = \mathbb{P}(\text{It would be found one day } | a_k)$. Suppose $a_1 = \frac{2}{9}$, $a_2 = \frac{3}{9}$, $a_3 = \frac{4}{9}$, $b_1 = 0.6$, $b_2 = 0.7$, $b_3 = 0.4$.

Each day, the navy can only search one region. Which region is the most worth searching first?

To solve this problem, we are interested in $\mathbb{P}(\text{Find sub in 1 day in region \#k})$, which happens to equal $a_k \cdot b_k$.

$$a_1 \cdot b_1 = \frac{1.2}{9}$$
$$a_2 \cdot b_2 = \frac{2.1}{9}$$
$$a_3 \cdot b_3 = \frac{1.6}{9}$$

Based on this analysis, it is obvious that they should send the team to search the second region first. That is what they did, and they were unable to find the sub. However, at this point, they were unsure what to do next. If it is assumed that the sub was not lost in the second region, which of the remaining two have the larger chance of being missing in? In other words, find $(a_1 | \text{not in } a_2)$, and $(a_3 | \text{not in } a_2)$. From here, the navy began asking the mathematicians for help with solving the problem.

**Example:**
Suppose you have two bags. The first (A) has three red balls, and the second (B) has four red balls. You randomly add a blue ball to one of the bags, with a 40% chance of picking A and a 60% chance of picking B.

Now, you ask a friend to pick a ball randomly out of the bags. What is the chance they pick the blue ball? The first pick is simple to calculate. Basically, find the chance of picking the blue ball given the correct bag was picked.

$$\mathbb{P}(A = b) = \mathbb{P}(b- > A) \cdot \mathbb{P}(\ldots | \ldots)$$

Simplifying, you get $0.4 \times \frac{1}{4} = 10\%$. Similarly, $\mathbb{P}(B = b) = 0.6 \times 0.2 = 12\%$. Thus, the chance of picking the blue ball out of bag A is 10% and the chance of picking the blue ball out of bag B is 12%.

However, this problem becomes interesting if the blue ball is **NOT** picked. The picked ball is replaced. Which bag should be picked next.

$$\mathbb{P}(b- > A | B \neq b) = \frac{\mathbb{P}(B \neq b | b- > A) \cdot \mathbb{P}(b- > A)}{\mathbb{P}(B \neq b)} = \frac{0.4}{\mathbb{P}(B \neq b)} = \frac{0.4}{0.88} = \frac{5}{11}$$

$$\mathbb{P}(b- > B | B \neq b) = 1 - \frac{5}{11} = \frac{6}{11}$$

Now, the new chances of having the blue ball in a given bag are $A = \frac{5}{11}$ and $B = \frac{6}{11}$ instead of 40% and 60%, respectively. The calculation can then be repeated.

## Chapter 2: Discrete Random Variables

**Example:**
Toss two dice

$$S = \{(x_1, x_2) \ : \ x_1, x_2 \in [[1, 6]]\}$$

Suppose you want to know the possible sums of the two dice. If the sum is $X$, this can be written as $X = x_1 + x_2$. This can also be written as a function of $X$, where $X((x_1, x_2)) = x_1 + x_2$, $X \ : \ S \to R$.
For example:

$$X((1, 2)) = 3$$
$$X((2, 4)) = 6$$

$X$ is called a **random variable.**
Because $X$ is defined this way, a $X$ represents a random number based on the sum of the dice, but some numbers have a higher chance of appearing.

**Example:**
Tossing 10 coins:

$$S = \{(x_1, ..., x_{10}), x_k \in \{H, T\}\}$$

Now, let $Y$ be a random variable describing the number of heads.

$$Y(x_1, \ldots, x_{10}) = |\{k \ : \ x_k = H, k \leq 10\}|$$

## Probability Mass Functions

Random variables can produce more random variables; a function of a random variable is always a random variable.
Let $X$ be a random variable. $p_X$ is the probability mass function (pmf) of $X$. Let $p_X$ be defined on all real

numbers.

$$p_X : \mathbb{R} \to R$$

implies

$$p_X(k) = \mathbb{P}(X = k)\, \forall\, k \in \mathbb{R}$$

$$p_X(x) = \mathbb{P}(X = x)\, \forall\, x \in \mathbb{R}$$

**For reference:** $k$ is usually used to denote integers, while $x, y, a, b$ are used for real numbers. The two equations above mean the probability that the random variable $X$ is equal to the values $k$ or $x$.

1. $p_X(\mathit{number})$
2. small p, capital x
3. $\mathbb{P}(X = 2) = \mathbb{P}(\{X = 2\})$

   $$\{X = 2\} = \{\omega \in S \,:\, X(\omega) = 2\}$$

**Example:**
You toss two dice. Let $X$ be the sum.

$$S = \{(x_1, x_2) \,:\, x_1, x_2 \in [[1, 6]]\}$$

Suppose you want to find $X = 3$.

$$\{X = 3\} = \{(1, 2), (2, 1)\}$$

**Example:**
Suppose you toss a coin. $X = \{1, 0\} = \{H, T\}$.
$p_X(1) = \mathbb{P}(X = 1) = \frac{1}{2}$, and $p_X(0) = \mathbb{P}(X = 0) = \frac{1}{2}$
$p_X(k) = 0$, for other $k \in \mathbb{R}$
It is important to remember that $p_X$ describes the probability the random variable will equal a value, not the value itself. **Example:**
Toss two dice; let $X$ denote the sum. Let's write $p_X$.

$$p_X(k) = 0 \text{ if } k \notin [[2, 12]]$$

$$p_X(2) = \mathbb{P}(X = 2) = \mathbb{P}(\{(1, 1)\}) = \frac{1}{36}$$

$$p_X(3) = \mathbb{P}(X = 3) = \mathbb{P}(\{(1, 2), (2, 1)\}) = \frac{2}{36}$$

$$\vdots$$

## 1. Bernoulli Random Variable

$$X \sim \text{Bern}(p), 0 \leq p \leq 1$$

$$p_X(k) = \begin{cases} p & k = 1 \\ 1-p & k = 0 \\ 0 & \text{others} \end{cases}$$

**Example:**
Suppose you toss a die. $S = \{1,2,3,4,5,6\}$. $X = 1$ if the number is 6 or 2, otherwise $X = 0$, $X : S \to \mathbb{R}$
Thus,

$$X(\omega) = \begin{cases} 1 & \omega = 2 \text{ or } 6 \\ 0 & \omega = 1,3,4,5 \end{cases}$$

$$p_X(k) = \begin{cases} \frac{1}{3} & k = 1 \\ \frac{2}{3} & k = 0 \\ 0 & \text{others} \end{cases}$$

## 2. Binomial Random Variable

$$X \sim B(n,p), \text{ where } n \in \mathbb{N}, 0 \leq p \leq 1, n, p \text{ are parameters}$$

Then,

$$p_X(k) = \begin{cases} \binom{n}{k} \cdot p^k (1-p)^{n-k} & k \in [[0,n]] \\ 0 & \text{others} \end{cases}$$

An example of where this occurs is when you have $n$ independent experiments, and for each of them, the success rate is $p$. $X$ represents the number of successful experiments among the $n$ experiments. $X \sim B(n,p)$.

**Example:** suppose you have 100 coins. (This acts as your experiments). Let $X$ be the number of heads. $X \sim B(100, \frac{1}{2})$.

$$P(X = 44) = \binom{100}{44} \cdot \left(\frac{1}{2}\right)^{44} \cdot \left(\frac{1}{2}\right)^{56}$$

**Why does the binomial distribution formula work?**
First, consider an example: Let $n$ represent the number of trials, and $p$ be the success rate. $X$ will represent the number of successful trials.

If $n = 3$, then

$$S = \{(x_1, x_2, x_3), x_{1,2,3} \in \{S, F\}\}$$

$$A_1 = \{x_1 = s\},\ A_2 = \{x_2 = s\},\ \ldots$$

Since all the events are independent, we know that

$$\mathbb{P}(A_1) = \mathbb{P}(A_2) = \mathbb{P}(A_3) = \cdots = p$$

The event $\{X = 0\}$ can be written as $(A_1^c \cap A_2^c \cap A_3^c)$. Thus, $\mathbb{P}(\{X = 0\}) = \mathbb{P}(A_!^c \cap A_2^c \cap A_3^c)$ This can be simplified:

$$\mathbb{P}(A_1^c) \cdot \mathbb{P}(A_2^c) \cdot \mathbb{P}(A_3^c) = (1-p)^3$$

Now, consider the case where $\{X = 3\}$. This can be written as $(A_1 \cap A_2 \cap A_3) = p^3$. The case $\{X = 1\}$ would include $\{(S, F, F), (F, S, F), (F, F, S)\} = (A_1 \cap A_2^c \cap A_3^c) \cup (A_1^c \cap A_2 \cap A_3^c) \cup (A_1^c \cap A_2^c \cap A_3) = 3 \cdot (p)(1-p)^2$. The case $\{X = 2\}$ would include $\{(S, S, F), (S, F, S), (F, S, S)\} = (A_1 \cap A_2 \cap A_3^c) \cup (A_1 \cap A_2^c \cap A_3) \cup (A_1^c \cap A_2 \cap A_3) = 3 \cdot (p)^2(p-1)$

Now, extend this example to $n = 5$ and suppose we want to know the number of cases for $k = 2$, or having two successes. This can be written as

$$\{X = 2\} = (A_1 \cap A_2 \cap A_3^c \cap A_4^c \cap A_5^c) \cup (A_1 \cap A_2^c \cap A_3 \cap A_4^c \cap A_5^c) \cup (A_1 \cap A_2^c \cap A_3^c \cap A_4 \cap A_5^c) \cup \dots$$

There are exactly $\binom{5}{2} = 10$ different combinations of them, and also note that each case has the same probability of occurring.

From this, we derive the formula for any $n$ and $k$ that the binomial theorem would be $\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$.
**Example:**
Suppose you toss 100 dice. Let $X$ represent the number of dice that land on 6. Find the probability that at least 20 dice land on 6.

This can be written as

$$\mathbb{P}(X \geq 20) = \sum_{k \geq 20}^{100} \mathbb{P}(X = k) = \sum_{k \geq 20}^{100} \binom{100}{k} \cdot \left(\frac{1}{6}\right)^k \cdot \left(\frac{5}{6}\right)^{100-k}$$

There is no easy to solve this equation, the math is not difficult, but is extremely tedious. Use an calculator or some tool to solve problems like this.

## 3. Poisson Random Variable

Written as $X \sim \text{Poisson}(\lambda),\ \ \lambda \in \mathbb{R}$. The distribution follows the equation:

$$P_X(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!},\ k \in \mathbb{Z}, k \geq 0$$

The $\lambda$ represents the average number of successes per trial.
**Example:**
A while back, when quarks were being discovered, a physics laboratory was measuring detection of muons. After setting up their equipment, they counted the number of muons it detected every minute. If they graphed their data, with the y-axis representing number of minutes, and the x-axis represent number of detections in a given minute, they noticed that the distribution seemed to follow a curve. In fact, it was very close to the Poisson curve.

**Example:**
Suppose you have 80 balls and 40 boxes, Let $X$ represent the number of balls in the first box. If many, many trials were done, then the number of balls in the first box for each trial would follow the Poisson curve for $\lambda = 2$. In fact, if you had $m$ balls and $n$ boxes, where both $m$ and $n$ are large numbers, then the distribution will follow the Poisson curve of $\lambda = \frac{m}{n}$. To prove this, consider a single ball in one trial. Let $X$ represent

the number of balls in box 1, and $A_k$ represent the kth ball in box #1. Thus, $\mathbb{P}(A_k) = \frac{1}{n}$. The number of balls in the first box at the end of the trial can be represented by

$$X = \sum_{k=1}^{m} 1_{A_k}$$

**Lemma:**

$$X = \sum_{k=1}^{m} 1_{B_k}, \text{ where } B_k\text{'s are independent}$$

Then,

$$X \approx \text{Poi}(\sum_{k=1}^{m} P(B_k))$$