18.01A Topic 14: Introduction to probability, discrete random variables.
Read: SN: P section 1

Read the supplementary notes

Repeatable experiments: Probability deals with repeatable experiments such as flipping a coin, rolling a die or measuring a distance. Gambling, polling and measuring are typical places where probability is used.

Discrete Random Variables Suppose the number of possible outcomes is finite.
Outcomes: \( \{x_1, x_2, \ldots, x_n\} \).
Probabilities: \( P(x_i) = \text{probability of outcome } x_i \).

Example: Roll 1 die.
Outcomes = \{1, 2, 3, 4, 5, 6\}, with \( P(j) = 1/6 \) for \( j = 1, \ldots, 6 \).
E.g. if we roll a die many times we expect close to 1/6 of the outcomes to be a 2.

Example: Roll 2 die. There are two ways to describe the possible outcomes.
1. Ordered pairs = \{(1,1), (1,2), \ldots, (6,6)\} with \( P(i,j) = 1/36 \) for any of the pairs.
2. Totals = \{2, 3, 4, \ldots, 12\}, here \( P(2) = 1/36, P(7) = 6/36 \) (you will need to compute all the probabilities for the pset).

Example: Toss a fair coin. Outcomes = \{H, T\}, with \( P(H) = P(T) = 1/2 \).

Example: Ask a random voter if they support candidate A.
Outcomes = \{yes, no\}, with \( P(\text{yes}) = p \) and \( P(\text{no})=1-p \).
Note: In all the cases we carefully state how to run the repeatable experiment.

Terms
Sample space = set of possible outcomes = \{x_1, x_2, \ldots, x_n\}.
Probability function, \( P(x_j) = \text{probability of outcome } x_j \).
Trial = one run of the 'experiment'.
Independent: Two trials are independent if their outcomes have no effect on each other. E.g., repeated flips of a coin are independent.

Probability Laws:
Range law: i) \( 0 \leq P(x_j) \leq 1 \), ii) \( P(x_1) + \ldots + P(x_n) = 1 \).
(This says the probability of an outcome is between 0 and 100% and the total probability of every possible outcome is 100%.

Addition law: \( P(x_i \text{ or } x_j) = P(x_i) + P(x_j) \) (provided \( x_i \neq x_j \)).
Example: Roll two dice. Let \( A = \text{‘the total is < 4’} \Rightarrow A = \{2, 3\} \).
\( P(A) = P(2 \text{ or } 3) = P(2) + P(3) = 1/36 + 2/36 = 1/12 \).
Let \( B = \text{‘the total is odd’} = \{3, 5, 7, 9, 11\} \).

Multiplication law: If two trials are independend the \( P(x_i \text{ then } x_j) = P(x_i) \cdot P(x_j) \).
(continued)
Example: Two tosses of a coin. \( P(\text{HH}) = 1/4 = P(\text{HT}) = P(\text{TH}) = P(\text{TT}) \).

Example: Toss a fair die 3 times, what is the probability of getting an odd number each time.

Let \( A = \{1, 3, 5\} \). On any one toss \( P(A) = 1/2 \). Since repeated tosses are independent \( P(A \text{ then } A \text{ then } A) = P(A) \cdot P(A) \cdot P(A) = 1/8 \).

A Finite Random Variable \( X \) consists of

i) A list \( x_1, x_2, \ldots, x_n \) of values \( X \) can take.

ii) A probability function \( P(x_j) \).

Remark: The new requirement is that the outcomes have to be numbers.

Examples: 1. Roll a die, \( X = \) number of spots.
2. Roll a die, \( Y = (\text{number of spots})^2 \).

Expectation (or mean or expected value) of the finite random variable \( X \):

\[
E(X) = x_1 P(x_1) + \ldots + x_n P(x_n) = \sum_{i=1}^{n} x_i P(x_i).
\]

Example: Roll a die, let \( X = \) number of spots.

\[
E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = 3.5.
\]

Interpretation:

1. The expected value is the average over a lot of trials. E.g., roll a die, if I pay you $1 per spot then over a lot of rolls you would average $3.5 per roll.

Note, you would never be paid $3.5 on any one turn, it is the expected average over many turns.

2. Expectation is a weighted average like center of mass.

Concept question: How much would you be willing to pay to roll the die?

Example: Roll a die, you win $5 if it’s a 1 and lose $2 if it’s not. Let \( X = \) your win or loss \( \Rightarrow P(X = 5) = 1/6 \) and \( P(X = -2) = 5/6 \).

\[
E(X) = 5 \cdot \frac{1}{6} - 2 \cdot \frac{5}{6} = -\frac{5}{6}.
\]

Concept question: Is the above bet a good one?

Infinite discrete random variable

i) \( X \) takes values \( x_1, x_2, \ldots \).

ii) Probability function \( \sum_{1}^{\infty} P(x_j) = 1 \).

(continued)
Example: (From supplementary notes) A trial consists of tossing a fair coin until it comes up heads. Let $X = \text{number of tosses}$ ⇒

Total probability $= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = 1.$

Question: If I paid you $n$ for a trial of length $n$, what would you pay for a turn?

answer: Expectation $E(X) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + \ldots = 2.$ (The supplementary notes give a nice method for finding this sum.)

Poisson Random Variable with parameter $m$

i) $X$ takes values 0, 1, 2, 3, \ldots

ii) $P(k) = e^{-m} \frac{m^k}{k!}.$ (The factor $e^{-m}$ is chosen to give total probability 1.)

iii) Poisson random variables model events that occur sparsely (i.e. with low prob.).

Examples:

1. Defects in manufacturing.
2. Errors in data transmission.
3. Number of cars in an hour at a rural tollbooth.
4. Number of chocolate chips in a cookie.

Theorem: $E(X) = m.$

Proof: $E(X) = \sum_{k=0}^{\infty} ke^{-m} \frac{m^k}{k!} = e^{-m}(m + \frac{m^2}{1!} + \frac{m^3}{2!} + \ldots)$

\[= me^{-m}(1 + \frac{m}{1!} + \frac{m^2}{2!} + \ldots) = me^{-m}e^m = m.\]

Example: A manufacturer of widgets knows that 1/200 will be defective. What is the probability that a box of 144 contains no defective widgets?

answer: Since being defective is rare we can model the number of defective widgets in a box as a Poisson R.V. $X$ with mean $m = 144/200$.

We need to find $P(X = 0)$:

$P(X = 0) = e^{-m} \frac{m^0}{0!} = e^{-144/200} = .487.$

What is the probability that more than 2 are defective?

answer: $P(X > 2) = 1 - P(X = 0, 1, 2) = 1 - e^{-m}(1 + m + \frac{m^2}{2!}) = .037.$

Histograms

If we ran many trials and made a histogram of the percentage of each outcome, the result should look like the graph of $P(x_j)$ vs. $x_j$.

Or we make a histogram of the count for each outcome. The histograms at right give examples.