Chapter 1 (Not a good idea to give the Big picture
Definition of statistics. without knowing anything.)

Statistics is the science of data, how to interpret data, analyze data, and design studies to collect data.

Chapter 2.

2.1 Sample space & events

$\boxed{\text{Probability}}$ is a measure of Likelihood that an ~~occurrence of~~ occur future event. Examples include:

- Torrow's temperature exceeds 80 degrees.
- manufacturing a defective part.

$\boxed{\text{Sample space}}$.

The set of All possible outcomes for a given random experiment is called the sample space, denoted by S. The number of outcomes in S is denoted by nS.

Example 1. Find the sample space S (a set essentially) and the number of outcomes n_S.

The Michigan state lottery calls for a three-digit integer to be selected.

$$S = \{000, 001, \ldots, 999\}.$$

The size of the set of all possible outcomes is 1000

Example 2.

Four equally qualified applicants (a, b, c, d) are competing for 2 positions.

A. if the two positions are identical.

$$S = \{ab, ac, ad, bc, bd, cd\}.$$

(Note identical means $ab \Leftrightarrow ba$) $n_S = 6$.

(How to do it? $\binom{4}{2} = \frac{4\times3\times2\times1}{2\times1} = 12$,

$12/2 = 6$ because ab and ba are identical)

B. if the two positions are different.

$$S = \{ab, ac, ad, bc, bd, cd, ba, ca, da, cb, db, dc\}. \quad n_S = 12.$$

(an alternative way to do this is

a, b, c, d → And you got 6.

Event An event is a ~~set~~ subset of possible outcomes that is of interest. E.g.

sample space: $S = \{1, 2, 3, 4, 5, 6\}$. dice.

Event. $A = \{2, 4, 6\}$.

(seeing even outcomes).

Task. We will develop a mathematical framework so that we can assign probability to an event A, which will quantify how likely this event is.

The probability that A occurs will be denoted as $P(A)$.

Equiprobability Model.

Suppose that a sample space S contains $n_S < \infty$ outcomes, each of which is equally likely.

If event A contains n_A outcomes, then

$$P(A) = \frac{n_S}{n_A}.$$

This is called an equiprobability Model. Its main requirement is that all outcomes in S are equally likely.

(This result is no longer applicable if out-comes in S are not equally likely)

Example 2.1 (Lottery example).

$$S = \{000, 001, \cdots 999\}, \quad n_S = 1000$$

Define event A as

$$A = \{000, 005, \ldots\ldots 990, 995\}.$$

$$= \{\text{winning number is a multiple of } 5\}$$

Then there are $200 = n_A$ outcomes in A.

Since it's reasonable to assume that all possible outcomes are equally likely, therefore

$$P(A) = \frac{200}{1000} = 0.2.$$

Example 2.2.. (Job application).

A. if positions are identical.

$S = \{ab, ac, ad, bc, bd, cd\}$. $n_S = 6$.

Define event A as applicant d got selected for one of the two positions.

$A = \{ad, bd, cd\}$ [identical, Don't need to flip the order]

$n_A = 3$.

If each applicant has the chance of Being selected, each of the $n_A = 3$ in S is equally likely. Hence $P(A) = \frac{n_A}{n_S} = \frac{3}{6} = 0.5$.

Interpretation of $P(A)$.

- $P(A)$ measures the likelihood that A will occur on any given experiment
- "relative frequency interpretation

If the experiment is performed many times, then $P(A)$ can be interpreted as

"the percentage of times that A will occur over the long run".

This is called the relative frequency interpretation.

E.g. flip a coin ~~$P(A)$~~.

$S = \{1, 0\}$.

$n_S = 2$,

see a head (1)

$P(A) = \frac{1}{2} = 0.5$.

Trial 1.	Head	1/1 .	Percentage 100%.
Trial 2.	Head	2/2 .	100%.
Trial 3	Tail	2/3	66%
Trial 4	Head	3/4	75%
Trial 5	Tail.	3/5	60%
⋮			
Trial 10,000	Tail.	5000/10,000	50%

) This process is simulation / Show the R demo.