ECS315 2014/1 Part I.1 Dr.Prapun

1 Probability and You

Whether you like it or not, probabilities rule your life. If you have ever tried to make a living as a gambler, you are painfully aware of this, but even those of us with more mundane life stories are constantly affected by these little numbers.

Example 1.1. Some examples from daily life where probability calculations are involved are the determination of insurance premiums, the introduction of new medications on the market, opinion polls, weather forecasts, and DNA evidence in courts. Probabilities also rule who you are. Did daddy pass you the X or the Y chromosome? Did you inherit grandma's big nose?

Meanwhile, in everyday life, many of us use probabilities in our language and say things like "I'm 99% certain" or "There is a one-in-a-million chance" or, when something unusual happens, ask the rhetorical question "What are the odds?". [17, p 1]

1.1 Randomness is everywhere.

1.2. Many clever people have thought about and debated what randomness really is, and we could get into a long philosophical discussion that could fill up a whole book. Let's not. The French mathematician Laplace (1749–1827) put it nicely:

> "Probability is composed partly of our ignorance, partly of our knowledge."

Inspired by Laplace, let us agree that you can use probabilities whenever you are faced with uncertainty. [17, p 2]

1.3. Random phenomena arise because of [13]:

(a) our partial ignorance of the generating mechanism

(b) the laws governing the phenomena may be fundamentally random (as in quantum mechanics; see also Ex. 1.7.)

Laziness

(c) our unwillingness to carry out exact analysis because it is not worth the trouble

Example 1.4. Communication Systems [23]: The essence of communication is randomness.

(a) **Random Source**: The transmitter is connected to a random source, the output of which the receiver cannot predict with certainty.

- If a listener knew in advance exactly what a speaker would say, and with what intonation he would say it, there would be no need to listen!

(b) **Noise**: There is no communication problem unless the transmitted signal is disturbed during propagation or reception in a random way.

(c) Probability theory is used to *evaluate the performance* of communication systems.

Example 1.5. Random numbers are used directly in the transmission and security of data over the airwaves or along the Internet.

(a) A radio transmitter and receiver could switch transmission frequencies from moment to moment, seemingly at random, but nevertheless in synchrony with each other.

(b) The Internet data could be credit-card information for a consumer purchase, or a stock or banking transaction secured by the clever application of random numbers.

Example 1.6. Randomness is an essential ingredient in games of all sorts, computer or otherwise, to make for unexpected action and keen interest.

Example 1.7. On a more profound level, quantum physicists teach us that everything is governed by the laws of probability. They toss around terms like the Schrödinger wave equation and Heisenberg's uncertainty principle, which are much too difficult for most of us to understand, but one thing they do mean is that the fundamental laws of physics can only be stated in terms of probabilities. And the fact that Newton's deterministic laws of physics are still useful can also be attributed to results from the theory of probabilities. [17, p 2]

1.8. Most people have preconceived notions of randomness that often differ substantially from true randomness. Truly random data sets often have unexpected properties that go against intuitive thinking. These properties can be used to test whether data sets have been tampered with when suspicion arises. [21, p 191]

- [14, p 174]: "people have a very poor conception of randomness; they do not recognize it when they see it and they cannot produce it when they try"

Example 1.9. Apple ran into an issue with the random shuffling method it initially employed in its iPod music players: true randomness sometimes produces repetition, but when users heard the same song or songs by the same artist played back-to-back, they believed the shuffling wasn't random. And so the company made the feature "less random to make it feel more random," said Apple founder Steve Jobs. [14, p 175]

Reading assignment.

1.2 Background on Some Frequently Used Examples

Probabilists love to play with coins and dice. We like the idea of tossing coins, rolling dice, and drawing cards as experiments that have equally likely outcomes.

1.10. ***Coin flipping*** or ***coin tossing*** is the practice of throwing a coin in the air to observe the outcome.

When a **coin** is tossed, it does not necessarily fall heads or tails; it can roll away or stand on its edge. Nevertheless, we shall agree to regard "**heads**" (**H**) and "**tails**" (**T**) as the only possible outcomes of the experiment. [4, p 7]

- Typical experiment includes
  - "Flip a coin N times. Observe the sequence of heads and tails" or "Observe the number of heads."

1.11. Historically, ***dice*** is the plural of ***die***, but in modern standard English dice is used as both the singular and the plural. [Excerpted from Compact Oxford English Dictionary.]

- Usually assume six-sided dice
- Usually observe the number of dots on the side facing upwards.

1.12. A complete set of **cards** is called a pack or **deck**.

(a) The subset of cards held at one time by a player during a game is commonly called a **hand**.

(b) For most games, the cards are assembled into a deck, and their order is randomized by **shuffling**.

(c) A standard deck of 52 cards in use today includes thirteen ranks of each of the four French suits.

- The four suits are called spades ($\spadesuit$), clubs ($\clubsuit$), hearts ($\heartsuit$), and diamonds ($\diamondsuit$). The last two are red, the first two black.

(d) There are thirteen face values $(2, 3, \ldots, 10$, jack, queen, king, ace) in each suit.

- Cards of the same face value are called of the same **kind**.
- "court" or face card: a king, queen, or jack of any suit.

1.3 A Glimpse at Probability Theory

1.13. Probabilities are used in situations that involve ***randomness***. A ***probability*** is a number used to describe how likely something is to occur, and ***probability*** (without indefinite article) is the study of probabilities. It is "the art of ***being certain of how uncertain you are***." [17, p 2–4] If an event is certain to happen, it is given a probability of 1. If it is certain not to happen, it has a probability of 0. [7, p 66]

"event"

1.14. Probabilities can be expressed as fractions, as decimal numbers, or as percentages. If you toss a coin, the probability to get heads is 1/2, which is the same as 0.5, which is the same as 50%. There are no explicit rules for when to use which notation.

- In daily language, proper fractions are often used and often expressed, for example, as "one in ten" instead of 1/10 ("one tenth"). This is also natural when you deal with equally likely outcomes.
- **Decimal numbers** are more common in technical and scientific reporting when probabilities are calculated from data. Percentages are also common in daily language and often with "chance" replacing "probability."
- Meteorologists, for example, typically say things like "there is a 20% chance of rain." The phrase "the probability of rain is 0.2" means the same thing.
- When we deal with probabilities from a theoretical viewpoint, we always think of them as numbers between 0 and 1, not as percentages.
- See also 3.5.

[17, p 10]

Definition 1.15. Important terms [13]: Ingredients of Probability Theory

(a) An activity or procedure or observation is called a **random experiment** if its outcome cannot be predicted precisely because the conditions under which it is performed cannot be predetermined with sufficient accuracy and completeness.

- The term "experiment" is to be construed loosely. We do not intend a laboratory situation with beakers and test tubes.
- Tossing/flipping a coin, rolling a dice, and drawing a card from a deck are some examples of random experiments.

(b) A random experiment may have several separately identifiable **outcomes**. We define the **sample space** Ω as a collection of all possible (separately identifiable) outcomes/results/measurements of a random experiment. Each outcome (ω) is an element, or sample point, of this space.

- Rolling a dice has six possible identifiable outcomes ($1, 2, 3, 4, 5$, and 6).

(c) **Events** are sets (or classes) of outcomes meeting some specifications.

- Any[1] event is a subset of Ω.
- Intuitively, an event is a statement about the outcome(s) of an experiment.

The goal of probability theory is to compute the probability of various events of interest. Hence, we are talking about a *set function* which is defined on subsets of Ω.

Example 1.16. The statement "when a coin is tossed, the probability to get heads is l/2 (50%)" is a *precise* statement.

(a) It tells you that you are as likely to get heads as you are to get tails.

(b) Another way to think about probabilities is in terms of **average long-term behavior**. In this case, if you toss the coin repeatedly, in the long run you will get *roughly* 50% heads and 50% tails.

[1]For our class, it may be less confusing to allow event A to be any collection of outcomes (, i.e. any subset of Ω).

In more advanced courses, when we deal with uncountable Ω, we limit our interest to only some subsets of Ω. Technically, the collection of these subsets must form a σ-algebra.

Although the outcome of a random experiment is unpredictable, there is a **statistical regularity** about the outcomes. What you cannot be certain of is how the next toss will come up. [17, p 4]

1.17. ***Long-run frequency interpretation***: If the probability of an event A in some actual physical experiment is p, then we believe that if the experiment is *repeated independently* over and over again, then a theorem called the ***law of large numbers*** (LLN) states that, in the long run, the event A will happen approximately $100p\%$ of the time. In other words, if we repeat an experiment a large number of times then the fraction of times the event A occurs will be close to $P(A)$.

Example 1.18. Return to the coin tossing experiment in Ex. 1.16:

Toss a coin

	Result
1st time $\rightarrow$	H
2nd time $\rightarrow$	H
3rd time $\rightarrow$	T
4th time $\rightarrow$	H
$\vdots$	

$\frac{N(A,n)}{n}$ Relative freq.

$\frac{1}{1} = 1$

$\frac{2}{2} = 1$

$\frac{2}{3} \approx 0.67$

$\frac{3}{4} = 0.75$

Definition 1.19. Let A $= [\text{result} = H]$ be one of the events of a random experiment. If we conduct a sequence of n independent trials of this experiment, and if the event A occurs in $N(A, n)$ out of these n trials, then the fraction

$$\frac{N(A,n)}{n}$$

is called the **relative frequency** of the event A in these n trials.

1.20. The long-run frequency interpretation mentioned in 1.17 can be restated as

$$P(A) \text{ ``}{=}\text{'' } \lim_{n\to\infty} \frac{N(A,n)}{n}.$$

LLN (law of large numbers)

1.21. Another interpretation: The probability of an outcome can be interpreted as our subjective probability, or degree of belief, that the outcome will occur. Different individuals will no doubt assign different probabilities to the same outcomes.

1.22. In terms of practical range, probability theory is comparable with ***geometry***; both are branches of applied mathematics that are directly linked with the problems of daily life. But while pretty much anyone can call up a natural feel for geometry to some extent, many people clearly have trouble with the development of a good intuition for probability.

- Probability and intuition do not always agree. ***In no other branch of mathematics is it so easy to make mistakes as in probability theory.***
- Students facing difficulties in grasping the concepts of probability theory might find comfort in the idea that even the genius Leibniz, the inventor of differential and integral calculus along with Newton, had difficulties in calculating the probability of throwing 11 with one throw of two dice. (See Ex. 3.4.)

[21, p 4]

$P(A)$ = probability of event A

Let's consider a random experiment and a specific event A.

For example, toss two (fair) dice

Let A be the event that the sum is 11.

When the experiment has been performed, the event A may $\begin{cases} \text{occur} \\ \text{not occur.} \end{cases}$

The probability that it occurs is denoted by $P(A)$.

Q: How to interpret the value of probability?

What does the value of $P(A)$ tell us about event A?

A: "Long-run frequency interpretation".

Repeat the experiment n times.

Count the "fraction of times that A occurs"

↳ relative freq. of event A.

LLN: As $n \to \infty$, rel. freq. $\longrightarrow P(A)$