Chapter 3: Probability

Ott & Longnecker Sections: 4.1-4.3

Duzhe Wang

the Department of Statistics, UW-Madison

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

What do we study?

Probability is the branch of mathematics that deals with the laws of chance. Statistics is a discipline that applies probability theory to deal with the variability induced by chance. Thus, to study statistics, one must study probability.

Key Concepts: Probability, random sample, estimate, population parameter, random process, outcome, event, frequentist interpretation of probability, Bayesian interpretation of probability, independent and dependent events.

Outline

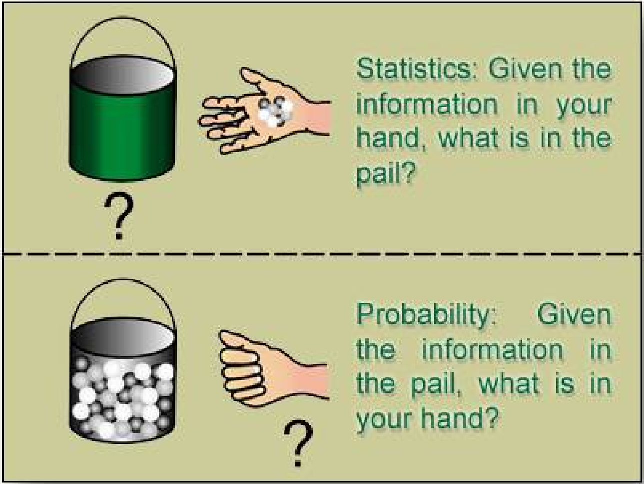
1. Probability vs. Statistics

2. Probability
   - Concepts
   - Properties
   - Example
   - Independent and dependent events
   - Frequentist vs. Bayesian interpretations

Probability vs. Statistics

- A **random sample** is a special kind of sample which results when population members are selected through some kind of random procedure
- Selecting population members randomly is one way to eliminate *bias* in sampling. Thus it is desirable to deliberately introduce randomness in sampling, but the consequence is that we must deal with the resultant uncertainty/variability. Practically everything in Statistics is a trade-off!
- Now we can distinguish between Probability in Statistics. Essentially one is an inversion of the other:
  - **Probability** takes information about a population, and allows us to make statements about what a random sample taken from that population might look like.
  - **Statistics** takes information about a random sample, and seeks to *infer* properties of the population from which the random sample was drawn.

Probability vs. Statistics



Probability vs. Statistics

Example. Suppose little Billy has an ant farm. The farm consists of 100 ants.

1. Suppose Billy knows that 10 of the 100 ants are poisonous, but the poisonous and non-poisonous ants are indistinguishable by sight. Probability would help him answer a question like, "If I select 1 ant at random, what is the probability that it is poisonous?" (Answer: 10%) He is using information about the population to say something about a random sample.
2. Suppose now that Billy doesn't know how many of the 100 ants are poisonous. He takes a random sample of 20 ants, and he determines that 3 of them are poisonous. Statistics would help him answer a question like, "What percentage of the ants in my farm are poisonous?" (Reasonable guess: 15%!) He is using information from a random sample to infer the percentage of ants that are poisonous.

Outline

Concepts

A **random process** is any process that generates results according to chance.

Examples:

1. Shuffling a standard deck of cards and dealing a card off the top is a random process. Sometimes you deal the Ace of Spades and sometimes you deal something else!
2. Another random process is, "select one ant at random from the ant farm and observe whether it is poisonous." Since we are selecting the ant at random and do not know whether it is posionous until we select it, there is random chance involved in the result.

Concepts

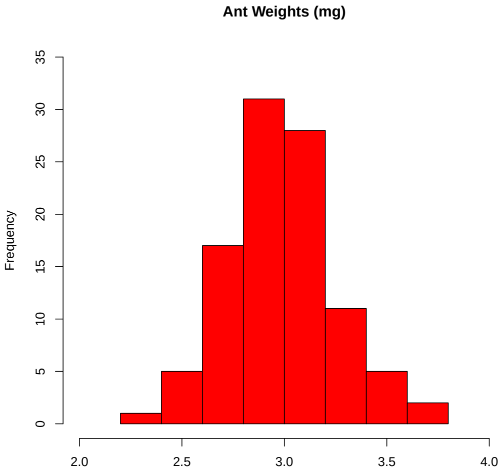
An **outcome** is a **distinct** result of a random process.
Example. If the random process was, "select one ant at random from the ant farm and observe whether it is poisonous," the possible outcomes are "poisonous" or "non-poisonous."
An **event** is a collection of outcomes.
Example. If the random process was, "shuffle a standard deck of cards and deal the top card" a possible event would be "the card is a King." The event "the card is a King" is a collection of four outcomes: King of Spades, King of Hearts, King of Diamonds, and King of Clubs.

Properties

1. For a random process with a finite set of outcomes, the probability of an event is the sum of the probabilities of the outcomes that comprise that event.
   Example: the probability of dealing a King from a well shuffled deck is $4/52 = 0.077$ – there are four Kings in the deck and each has a $1/52$ chance of being dealt.
2. For a random process with a finite set of outcomes, the probability of an event is always between 0 and 1. A probability of exactly 0 indicates the event will never happen no matter how many times the random process is repeated, and a probability of exactly 1 indicates the event will always happen no matter how many times the random process is repeated.
3. The probability that an event does not occur must always be one minus the probability that it does occur.
   Example: the probability of not dealing a King from a well-shuffled deck is 1 - 0.077 = 0.923.

Example: probability and histograms

Suppose the frequency histogram below represents the weights of the 100 ants in Billy's ant farm. It is very important in this example that we remember that these 100 ants comprise the entire population:



Example: probability and histograms

The bin cut points are in jumps of 0.2, so the first bin goes from 2.2 - 2.4, etc. The counts, left to right, are: 1, 5, 17, 31, 28, 11, 5, 2.
Question 1: If one ant is selected at random, what is the probability that the selected ant will have a weight that is less than 2.8 mg?
We assume that no ant weighs exactly 2.8 mg.

Example: probability and histograms

The bin cut points are in jumps of 0.2, so the first bin goes from 2.2 - 2.4, etc. The counts, left to right, are: 1, 5, 17, 31, 28, 11, 5, 2.

Question 1: If one ant is selected at random, what is the probability that the selected ant will have a weight that is less than 2.8 mg?

We assume that no ant weighs exactly 2.8 mg.

Each ant represents a single outcome, and since we are drawing the ant at random, every ant has an equal probability of being drawn, namely, $1/100$.

The event that the selected ant weighs less than 2.8 mg is a collection of all the outcomes where the selected ant weighs less than 2.8 mg. Thus we can add together the probabilities for each ant that weighs less than 2.8 mg.

From the histogram, we see that $1 + 5 + 17 = 23$ of them have weights less than 2.8 mg. Therefore the probability that a single randomly selected ant weighs less than 2.8 mg is $1/100 + 1/100 + ... + 1/100$, 23 times, which is simply $23/100$ or 0.23.

Example: probability and histograms

Question 2: What is the probability that a randomly selected ant weighs more than 2.8 mg?

Example: probability and histograms

Question 2: What is the probability that a randomly selected ant weighs more than 2.8 mg?
This is easy using the solution to the first example. In the previous example, we computed the probability that a single chosen ant weighed less than 2.8 mg. If an ant weighs more than 2.8 mg, it cannot weigh less than 2.8 mg. Therefore,

$$\text{P(One randomly selected ant weighs} > 2.8 \text{ mg)} = 1 - \text{P(One randomly selected ant weighs} < 2.8 \text{ mg)} = 1 - 0.23 = 0.77$$

Independent and dependent events

Two events are said to be **independent** if one event occurs does not affect the probability of the other event occurring. If two events are not independent, we call them **dependent**.

Example. Think back to when Billy was considering poisonous and non-poisonous ants. Suppose he has 20 ants in his farm, and 5 are poisonous. If one ant is selected at random, the probability of selecting a poisonous ant is $5/20 = 0.25$. Now consider two scenarios:

1. Suppose that the first ant selected is not returned to the ant farm. If the first ant selected is poisonous, the probability that the second ant selected is poisonous is now $4/19 = 0.21$. However, if the first ant selected was non-poisonous, the probability that the second ant selected is poisonous is now $5/19 = 0.26$. The probability the second ant is poisonous differs depending on whether the first ant is observed to be poisonous or not. Thus these events are dependent.
2. Suppose instead that the first ant selected is returned to the ant farm. If the first ant selected is poisonous, the probability that the second ant selected is poisonous is $5/20 = 0.25$. If the first ant selected was non-poisonous, the probability that the second ant selected is poisonous is still $5/20 = 0.25$. In this scenario, the probability of the second event is exactly the same regardless of the outcome of the first event. Thus these events are independent.

The definition of independence implies another important fact: **If two events are independent, then the probability that they both occur is the product of the individual probabilities of the events.**

Example. If Billy selects two ants with replacement, what is the probability that both ants are poisonous?

Independent and dependent events

The definition of independence implies another important fact: **If two events are independent, then the probability that they both occur is the product of the individual probabilities of the events.**

Example. If Billy selects two ants with replacement, what is the probability that both ants are poisonous?

$0.25 \times 0.25 = 0.0625$

Frequentist vs. Bayesian interpretations

Let's return to the basics of probability. In particular, what does probability mean?

Frequentist vs. Bayesian interpretations

- The **long-run frequency** or **frequentist** interpretation of probability says that if a random process is replicated over and over, the probability of an event is the proportion or percent of times that the event occurs. This interpretation is quite valuable in situations where replications are hypothetically possible, for example, drawing 10 ants from the ant farm, or flipping a coin.
- The **subjective** or **Bayesian** interpretation says that the probability of an event represents a degree of belief whether the event will occur. This interpretation makes more sense in situations where a random process can never be exactly repeated, for example, the chance that it rains today. Despite seeming very informal, this interpretation is still quite valuable in many applications.

Frequentist vs. Bayesian interpretations

While the Bayesian interpretation seems to be more consistent with how humans perceive probability, in this class we will focus on the frequentist interpretation. Note:

1. The mathematics of Bayesian statistics are quite advanced, and to learn it it is useful to learn Frequentist statistics first.
2. In the presence of a sufficient amount of data, Bayesian and frequentist statistics usually agree very well (a huge amount of data speaks for itself!).

What's the next?

In the next chapter we will discuss random variables and probability distributions.