

Week 3: Probability

Ch.3 Probability

The Set of All Possible Events

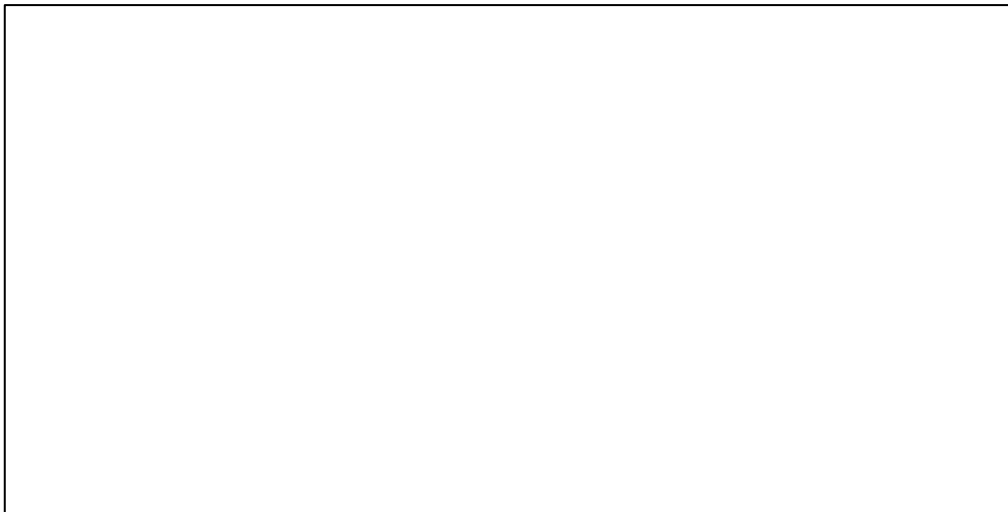
Recall the black die from class. We looked at the probability of rolling a 6, which we deduced was decidedly not $1/6$. However, let's first take a step back.

- *Question:* where does the probability of $1/6$ come from?
- *Answer:*

- *Question:* now assume we have two (fair) dice, what is the probability of rolling dice that sum to 12?
- *Answer:*

- *Question:* finally with two (fair) dice, what is the probability of rolling dice that sum to 11?
- *Answer:*

Enumerating the set of possible outcomes



Outcomes as coin flips

- Rarely do we have a vested interest in the outcome of a single coin flip.
- However the idea and mathematical notation behind coin flips are used to model binary phenomenon.
 - Will the Cats beat the Griz in football (*AGAIN*) next year?
 - Will there be a powder day at Bridger this weekend?
 -
- Most real world events will not have equal probability of occurrence, or in other words, will not be a fair coin.

What is Probability?

- *Question:* we have used this term several times, what does it mean?

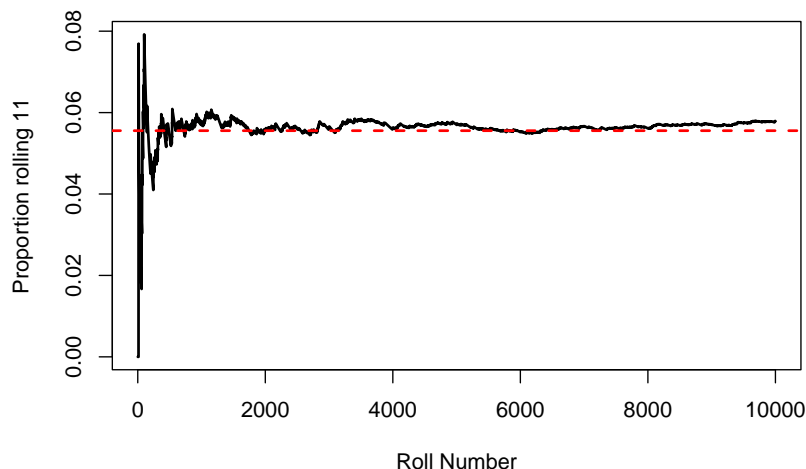
Long term frequency

For some events the probability can easily be formulated as the long-run frequency of an event relative to the other outcomes. The long-run frequency can be usually be simulated or derived mathematically (with minor assumptions).

Simulation (Monte Carlo Techniques) Simulation can be used to estimate probabilities in many situations. This process mimics conducting the process many, many times.

Consider the earlier example for the probability of rolling an eleven with two dice.

```
num.sims <- 10000
die1 <- sample(6, num.sims, replace=T)
die2 <- sample(6, num.sims, replace=T)
sum.dice <- die1 + die2
count.11 <- cumsum(sum.dice == 11)
prob.11 <- count.11 / 1:num.sims
plot(prob.11, type = 'l', lwd=2, ylab='Proportion rolling 11', xlab='Roll Number')
abline(h=1/18, col='red', lwd=2, lty=2)
```



mathematical derivation As we saw earlier, assuming equal probability of each outcome, we can derive the probability mathematically. This generally amounts to enumerating (or counting) all possible outcomes and then computing the proportion of outcomes that satisfy our specified criteria.

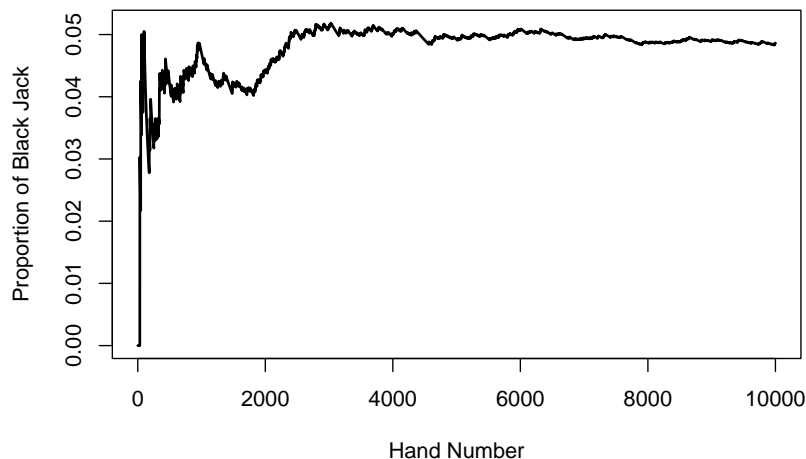
Exercise: Black Jack

Either write pseudocode or a mathematical derivation for being dealt black jack in a hand of two cards. Black Jack is two cards that add up to 21, where aces are worth 11 and 10's and face cards are equal to 10.

Solution: Simulation

```
num.sims <- 10000
cards.df <- data.frame(card.number = rep(c('ace',2:10,'jack','queen','king'),4),
                      suit = rep(c('hearts','spades','diamonds','clubs'), each = 13),
                      value = rep(c(11,2:10,10,10,10),4))
card.value <- rep(0,num.sims)
for (i in 1:num.sims){
  card.value[i] <- sum(sample(cards.df$value, 2))
}

count.21 <- cumsum(card.value == 21)
prob.21 <- count.21 / 1:num.sims
plot(prob.21, type = 'l', lwd=2, ylab='Proportion of Black Jack', xlab='Hand Number')
```



Based on 10^4 Monte Carlo simulations the proportion of hands resulting in Black Jack is 0.0486.

Solution: Mathematical Derivation

In this case there are $\binom{52}{2}$ possible outcomes where two cards are selected from a deck of 52 cards. This notation is read 52 *choose* 2 and can be written as $\binom{52}{2} = \frac{52!}{(52-2)!2!}$.

The next step is to count how many of these hands result in black jack. Similarly this can be written as $\binom{16}{1} \times \binom{4}{1}$, that is we choose one of the 16 cards worth ten points *and* one of the ace cards.

Thus the solution is $\frac{16 \cdot 4}{\binom{52}{2}} = 0.0482655$.

subjective beliefs

An alternative way to think about probabilities is as subjective belief. There is a subtle difference, but this is not the true long-run frequency, but rather the degree of belief in each possible probability. As a *subjective* belief these will vary from person-to-person.

These beliefs can be calibrated in hypothetical betting scenarios. For example, consider a subjective belief on the probability that it will snow tomorrow in Bozeman.

- Gamble 1: You get \$100 if it snows in Bozeman tomorrow.
- Gamble 2: You get \$100 if a coin toss results in heads.

These comparisons can become more detailed to refine the prior belief on the probability.

Probabilities assign numbers to possibilities

Probability is way to of assigning numbers based on the likelihood of occurrence to a set of mutually exclusive possibilities. These numbers are called probabilities and must adhere to three properties known as Kolmogorov's Axioms:

1. A probability must be nonnegative.
2. The sum of probabilities across all events in the entire sample space must be 1.
3. For any two mutually exclusive events the probability that one *or* the other occurs is the sum of their individual probabilities.

Probability Distributions

A distribution is a list of all possible outcomes and their corresponding probabilities. In the coin flip setting this is a fairly trivial notion, we have the probability of heads (or one outcome) denoted as θ and due to the axioms of probability the other outcome has probability of $1 - \theta$. These distributions are more complex in countably, finite situations such as the magpies in my tree or even continuous data.

Discrete Distributions: Probability Mass

With discrete outcomes, each physical outcome can assigned a probability of occurrence.

Distribution of



Continuous Distributions: Probability Density

Now consider a continuous quantity, like height that can theoretically be measured to an arbitrarily precision.

Distribution of

