

# STAT 491 - Lecture 3

January 18, 2018

## Ch.3 Probability

### The Set of All Possible Events

Recall the dice from class on Tuesday. 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.

—

—

—

- Most real world events will not have equal probability of occurrence, or in other words, will not be a fair coin.

## What is Probability?

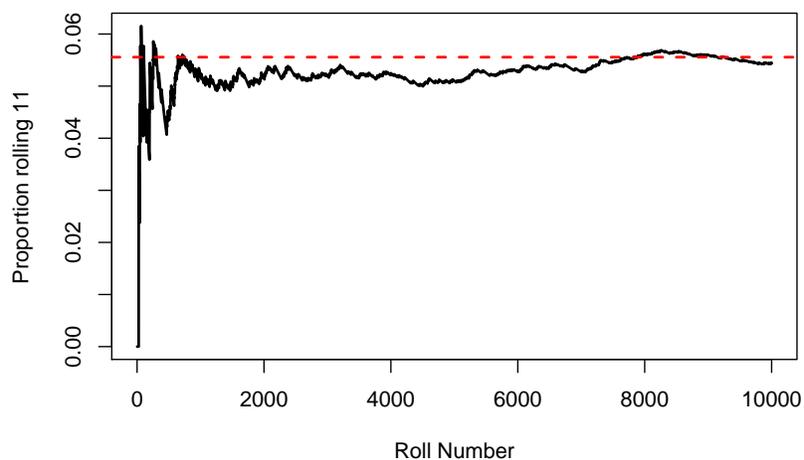
- *Question:* we have used this term several times, can anyone define it?

## Long term frequency

## Simulation

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**

### **Solution: Mathematical Derivation**

### 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.

- 

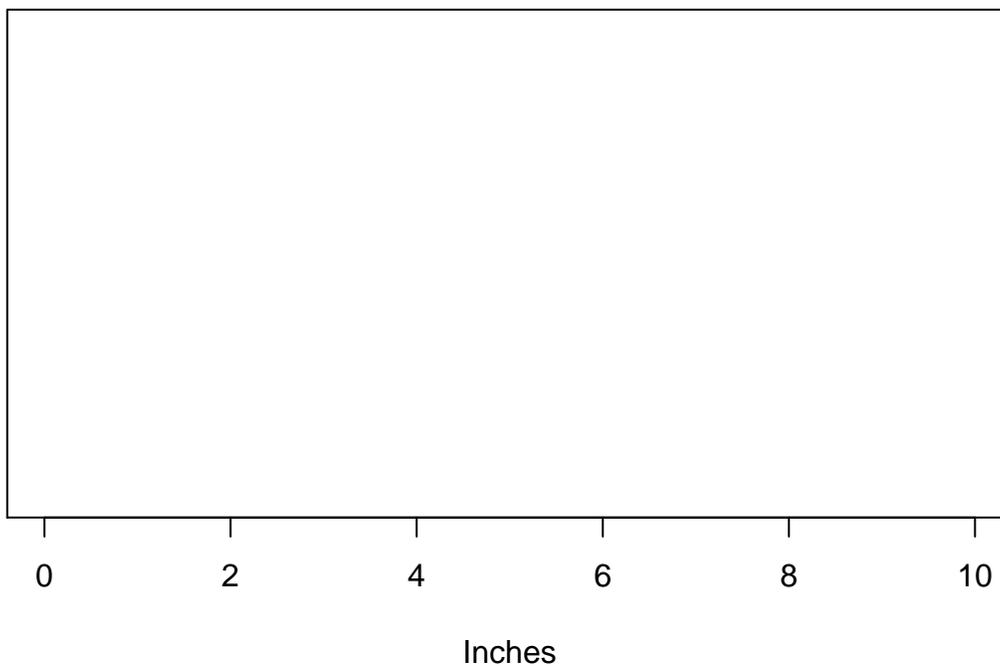
- 

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

### mathematical representation of subjective beliefs

Now assume that the goal is not to model the probability that it will snow tomorrow, but rather the distribution of snowfall in inches.

## Subjective Belief on Tomorrow's Snow Fall



### **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.

2.

3.

### **Exercise: Magpies**

I have a 50 foot spruce tree in my front yard and am interested in learning how many magpies live in the tree.

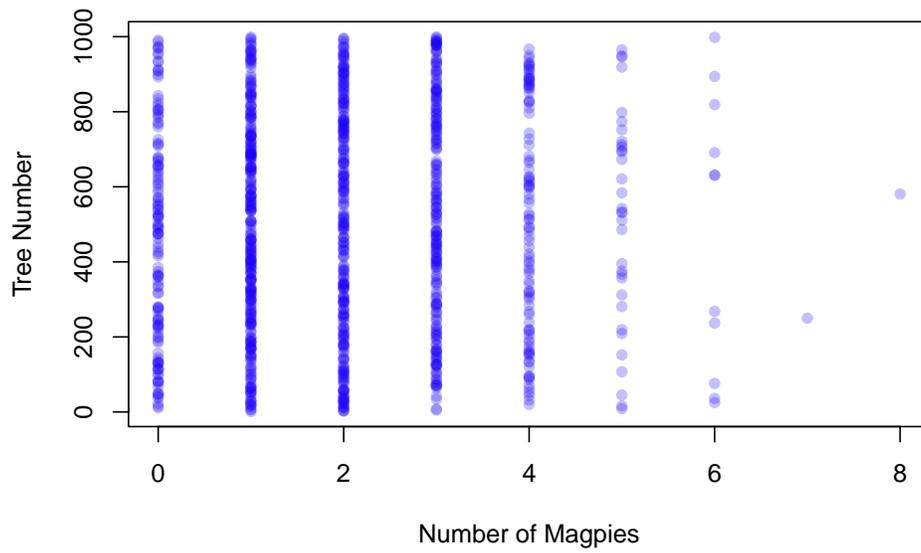
1. What are the set of possible outcomes?

2. Assign a set of probabilities to each of these outcomes and sketch the probabilities.

3. Show that these probabilities satisfy the Kolmogorov Axioms.

# Probability Distributions

## Discrete Distributions: Probability Mass



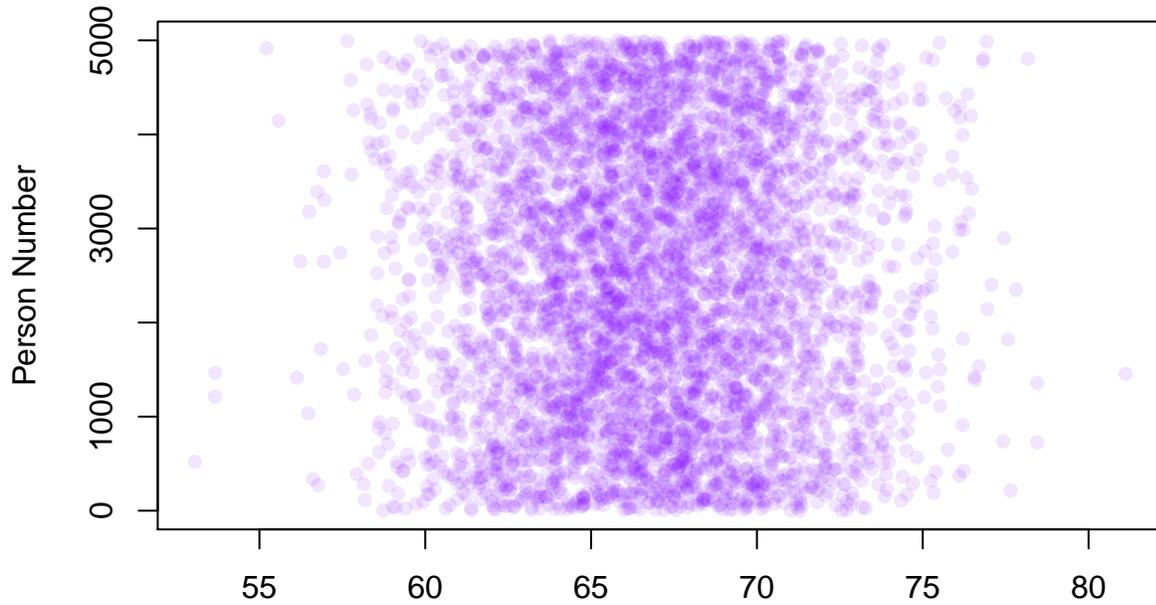
```
## num.magpies
## 0 1 2 3 4 5 6 7 8
## 131 274 247 204 98 33 11 1 1
```

## Distribution of Magpies per tree

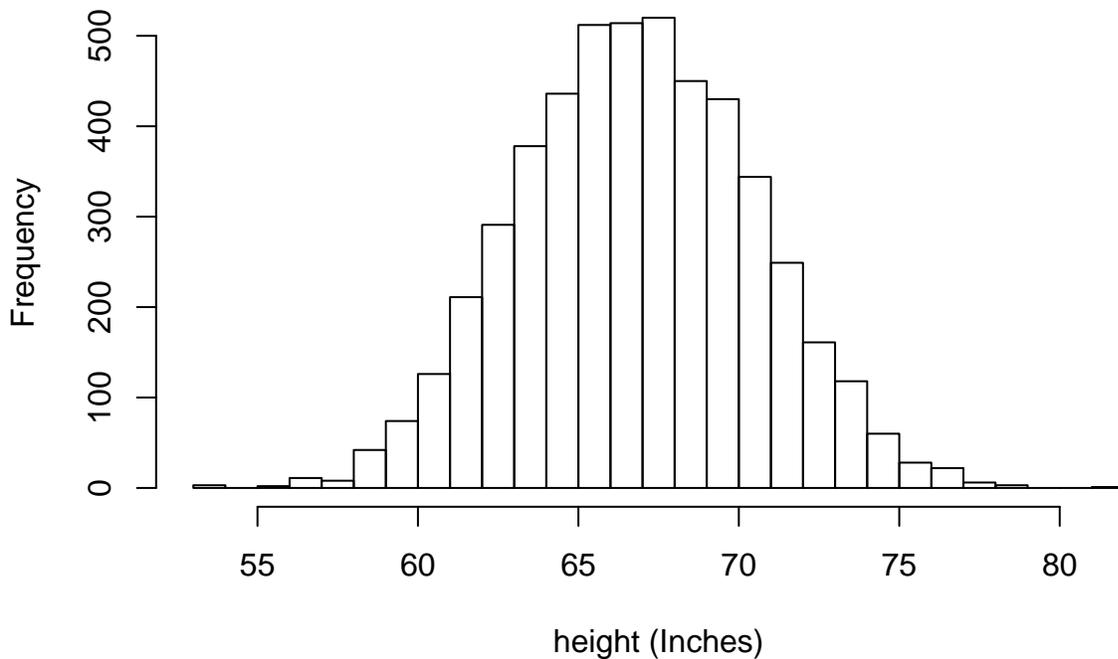


## Continuous Distributions: Probability Density

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



Height (inches)  
**histogram of height**



How do we think about the probability distribution in this situation using heights of 5000 people?

- **Question:** Does the histogram represent a probability density that obeys the Kolomogorov Axioms?

- *Answer:*

- **Question:** What is the approximate probability that a person is exactly 66 inches tall?

- *Answer:*

- **Question:** What is the approximate probability that a person is 67 inches tall ( $\pm$  half an inch)?

- *Answer:*

## Properties and Notation of Probability Density Functions

Assume a probability density function is split into intervals, then

### A bit about integration

- Integration is the continuous analog of a summation. In the above example, the interval width become infinitesimally small such that integration was required. ( $\sum_x \rightarrow \int dx$ )
  
- There will be a fair amount of calculus in this class; however, we will the actual calculations will generally use  $\int p(x)dx = 1$  and amount to algebraic manipulations.