STAT 325 – Handout 1
Simulation, Sample Space, Events (1.1)

Example 1-1: Random Babies
Suppose that four mothers give birth to baby boys at the same hospital on the same evening. As a very, very sick joke (do not try this at home!), the hospital staff decides to return the babies to the mothers at random! How likely is it (i.e., what is the probability) that at least one mother receives the correct baby?

This is a classic probability problem known as the *matching problem*. If you prefer to think of a different context, suppose that four college students bump into each other and drop their cell phones; then each student picks up one of the four cell phones at random.

We'll take two approaches to answering this question:
1) An approximate analysis via simulation
2) An exact analysis via enumeration

- A process is **random** if individual outcomes are uncertain but there is a regular distribution of outcomes in a large number of repetitions.
- The **probability** of any outcome in a random process can be interpreted as the proportion (**relative frequency**) of times that the outcome would occur in a very large number of repetitions.
- A probability can be approximated by **simulating** (artificially re-creating) the random process a large number of times and determining the relative frequency of occurrences.

a) Describe how you could use four index cards to simulate this random process and approximate the probability of interest.

b) Use four index cards to conduct 10 repetitions of this random process. For each repetition, record the number of mothers who receive the correct baby. Also keep track of how many of your 10 repetitions result in at least one mother receiving the correct baby.

c) Based on your simulation analysis, report the approximate probability that at least one mother receives the correct baby.

d) Suggest how to determine a more accurate approximation for this probability.

e) Now combine your simulation results with those of your classmates to determine another approximation for this probability.

f) Now we turn to technology in order to perform this simulation more quickly and efficiently. We will use an applet found at: www.rossmanchance.com/applets/randomBabies/Babies.html. (Be forewarned that this applet contains rather graphic images that reveal where babies come from!) Conduct 1000 repetitions of this simulation. Report the approximate probability of interest.

- Conducting more repetitions in a simulation analysis generally produces more accurate approximations of probabilities.
- A (rough) rule-of-thumb is that the approximate probability will likely fall within $\pm 1/\sqrt{N}$ of the actual probability, where N represents the number of repetitions.

g) Calculate and interpret the value of $1/\sqrt{N}$ for your applet simulation.

Can we use a more mathematical analysis to calculate this probability exactly? Yes, and we will soon, but first we need to learn some basic probability ideas.

Example 1-2: Ice Cream Prices

Suppose that you have only 50 cents in your pocket and you want to buy an ice cream cone. The owner of the ice cream shop offers a random price determined as follows: You roll a pair of fair, six-sided dice, and the price is the larger number followed by the smaller number (in cents). What is the probability that you'll be able to afford the ice cream cone?

a) Let's use Minitab to simulate this random process and determine an approximate probability that you'll be able to afford the ice cream cone. We'll first do this 100 times and then 10,000 times. Report the approximate probability each time.

b) Now let's use R to conduct this simulation with 10,000 repetitions:

```
d1 = sample(1:6,10000, replace=TRUE)
d2 = sample(1:6,10000, replace=TRUE)
price = 10*pmax(d1,d2) + pmin(d1,d2)
afford=(price<=50)
sum(afford)/10000
```

Report the approximate probability and also a rough estimate of its accuracy.

Conducting an exact mathematical analysis involves listing all possible outcomes of the random process.

- The **sample space** (S) of a random process is a set consisting of all possible outcomes.
- An **event** is a subset of the sample space (often denoted by a capital letter).
- If the outcomes are **equally likely**, then the probability of the event is the number of outcomes in the event divided by the total number of outcomes in the sample space.
- Choosing an object "at random" means that the possible objects are equally likely.
- Two events are **mutually exclusive** if they have no outcomes in common (i.e., if their intersection is the empty set).

c) List the 36 outcomes in the sample space for rolling a pair of fair, six-sided dice.

d) Circle the outcomes that comprise the event that you can afford the ice cream cone.

e) Determine the (exact) probability that you can afford the ice cream cone.

f) Is the exact probability close to our simulation approximations?

Example 1-3: Odd Numbers
Think of a two-digit number, both digits odd, the two digits different from each other.

a) Write down your choice.

b) List the outcomes in the sample space. How many outcomes are there? What is the probability of each outcome, assuming that the number is chosen at random?

I will make a prediction about two numbers that I think will be chosen most often.

c) If the outcomes are really equally likely (i.e., if people are really choosing at random), what is the probability that one of these numbers would be selected?

d) How many people in our class chose one of these numbers? What proportion is this? Does this exceed the theoretical probability, assuming equal likeliness?

Example 1-4: Solitaire
Suppose that every night I play Solitaire on my computer until I win for the first time.

a) Describe the sample space for this random process.

b) Are these outcomes equally likely? Explain why or why not.