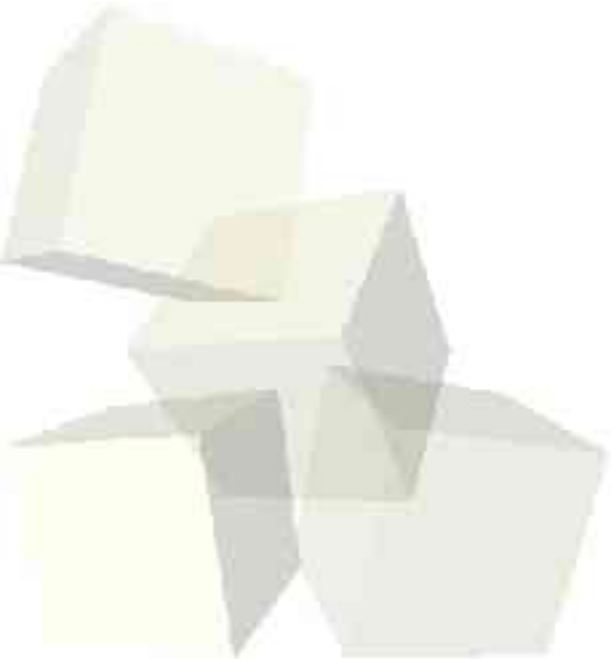
Observables and Stats

1-18-2005



Opening Discussion

- Do you have any questions about the syllabus or the way the course will be run?
- What do you need to do before you can write a simulation? If you have a hard time with that, what do you need to do before you write any type of program?

Model Design

- Before you can code up a simulation, you have to design the model that you will be using for the simulation. This is a high level task and something we will spend a lot of time talking about in this course.
- Last time we discussed different models we could use for some simple queue problems. We avoided talking about some of the basic inputs for any of those models, how to pick when people get in a queue or when they leave.

Random Variables

- You likely talked about random number generation in PAD1 with the linear congruential generator method. There are many different ways to generate pseudo-random numbers on a computer, but typically they pull a new number uniformly from some particular range.
- We use these random numbers a lot in doing simulations because there are typically aspect outside the simulation or even inside it that we aren't modeling exactly.

Probability Distributions

- I said random numbers typically come from a uniform distribution. There are other distribution though. We can draw them as plots where each possible value for something has a probability of occurring.
- Some common distributions between a and b are
  - Uniform : $f(x)=1/(b-a)$ in $[a,b]$ else 0
  - Triangular
  - Exponential : $f(x)=d*e^{-d*x}$
  - Normal : $f(x)=1/(s*\sqrt{2*PI})*\exp(-0.5*((x-u)/s)^2)$

Discrete vs. Continuous

- Distributions can be discrete, like the odds of rolling a number on a die, or continuous, like the odds of a person throwing a rock a certain distance.
- When dealing with discrete distributions we don't have connected curves and we look at summations for things like normalization.
- When dealing with continuous distributions we have connected curves and need integrals instead of sums.

Differential Distributions

- Most of the time when talking about distributions we look at differential distributions. These are the odds that we get a value in some range. For continuous it would be between x and x+dx.
- When we draw distributions this is typically what we thing of. When there is a greater magnitude the odds are higher.
- While this method is nice conceptually, it isn't always ideal numerically.
- Think of binning values.

Cumulative Distributions

- An alternate way of looking at distributions is the cumulative distribution. In a cumulative distribution our function, f(x), returns to odds you get something with value x or less.
- Basically, a cumulative distribution is the integral of a differential distribution.
- These distributions are monotonically increasing.

Picking with Randoms

- So if I give you some probability distribution, how do you use it in a simulation to give me numbers selected from that distribution?
- One way is to pick two numbers from uniform distributions, r1 and r2, then see if f (r1)>r2. If so, you return r1, otherwise you generate a new pair.
- Let's look at how this works graphically.
- What do you see as drawbacks of this method? What are the strengths?

Picking with Cumulative Dist.

- An alternate way of picking numbers from a distribution can be used if you have a cumulative distribution and you can take the inverse of it.
- The cumulative distribution should go from zero up to one monotonically. If you generate uniform a random number in that range it will map directly to a value in the domain of your distribution.
- If f(x) is the distribution and $f^{-1}(x)$ is the inverse, then $f^{-1}(rand)$ is your value.

What Distribution is Right?

- The real question with model design is what distribution to use.
- For some problems you can make logical arguments for why you should use a standard distribution and you just need to measure some values to get the details.
- In other situations you need to get enough data to figure out exactly what the distribution is. You either fit it to some type of curve, or you use a table for your distribution.

Monte Carlo Methods

- There is an entire branch of simulation called Monte Carlo simulations that are basically nothing other than repeated intelligent drawing of random numbers.
- The method we just mentioned before for picking numbers from a distribution is basically a Monte Carlo simulation of a very simple nature.
- For many processes where events are unrelated this is a great way to do things.

Fuzzy Data

- Another use of probabilities is in the use of fuzzy data and fuzzy systems. These took the world by storm in the late 80s and early 90s. You see less about them today, but they can still be very powerful.
- Fuzzy data is a method we can use to describe things when the nature of something isn't clean cut, but instead their might be ambiguity.
- For example, take a temperature and convert it to being cold, cool, warm, or hot.

Minute Essay

- When would you use different methods of picking random numbers from distributions?
- A week from today you will take your first quiz. A week after that the first assignment will be due.