Sept. 30 - Oct. 4

Stat 330 Tutorial 4 - Midterm Review

Tutorial, and these notes, are meant to supplement, not replace, your own studying!!!

Midterm is next Monday. Covers ~~a~~ everything up to and including 2.5. This is a course about problem solving, not memorization. That is, test questions will require that you apply similar techniques as on the homework, but you can be confident that the questions will be at least a little bit different.

Probability Model: Has 3 components

(S) i) A sample space. This can be any set, containing whatever elements you want. In practice, it lists all possible outcomes of some random phenomenon.

ii) A collection of events. These are subsets of the sample space.

(P) iii) A probability measure. This is a function ~~with~~ who's domain ~~equal to~~ is the collection of events ~~#~~ and range is ~~the i~~ in the interval $[0,1]$. A probability measure must satisfy Kolmogorov's axioms.

~~Ideas~~ When working with complex events, it is often helpful to draw a Venn-diagram

(1.2.14) Let S, the sample space, be ~~ca~~ finite or countable. Is it possible to have $P(\{x\})=0$ for all $x \in S$? Why or why not?

(1.2.15) Let S be uncountable. Is it possible to have $P(\{x\})=0$ for all $x \in S$? Why or why not?

Properties of probability models: This is how you prove things about probability models. A Venn-diagram is not, in general, a proof. Questions from this section are kind of like trig. identities. You have a few formal rules for manipulating symbols that you have to apply creatively to get from one expression to another.

Uniform probability on finite spaces: In general terms, this section consists of finite S and P just counts the number of elements in its argument then divides by the size of S. More practically, this section is about combinatorics. There is no magic formula for solving combinatorics problems (That I'm aware of), just practice. If you run out of practice questions in the textbook, try going back to your 270 or 285 book, or you can ask me (or ~~thou~~ you could come up with your own).

Conditional probability and independence: This section introduces some new properties of prob. measures. This expands your toolkit for calculating probabilities, which means you can calculate probabilities of more ~~ca~~ types of events and that you can be asked ~~a~~ more types of questions. In particular, Bayes theorem is very powerfull. It allows you to calculate probabilities that would be impossible any other way.

Continuity of probability measures: This section allows you to calculate probabilities of complicated events by taking limits of probabilities of more simple events. It's not always obvious what the simple events should be, but this strategy is extremely common in probability ~~theory~~ theory and mathematical analysis in general.

Random Variables (RVs): A random variable is a function from the sample space to $\mathbb{R}$, the real numbers. This seems like a weird definition, and it kind of is because ~~in~~ this course hides some ~~with~~ measure theoretic details that encourage us to define random variables in this way (you do not need to know about the measure theory part!).

The examples in the book illustrate this concept pretty well, ~~excs~~ but let's look at one anyway. Let S be the outcome of flipping 3 coins. Let A be the event that we get at ~~lease~~ least 2 heads. Define X to be the number of heads and $Y = I_A$ be 1 if we get at least 2 heads and 0 otherwise. Both X and Y are random variables (why?). Without our definition of random variable, we would have to define new sample spaces and probability measures for X and Y. With our definition, we can explicitly maintain the relationship between X and Y, and the underlying sample space of coin flips.

Distributions of RVs: Remember that a probability measure is defined as a function on the collection of events (subsets of S). ~~This means that, without another defin~~ That is, probabilities are only defined for subsets of S. But random variables take values in $\mathbb{R}$, so our theory doesn't yet tell us how to talk about things like $P(X=2)$, where X is a RV. Fortunately, it's not hard to solve this. We define $P(X=2) := P(\{x \in S : X(x) = 2\})$ or, more generally, if $A \subseteq \mathbb{R}$ then $P(X \in A) := P(\{x \in S : X(x) \in A\})$. People often say that X "induces" a probability measure on $\mathbb{R}$. This measure on $\mathbb{R}$ is called the distribution of X. Having this induced probability measure on $\mathbb{R}$ often allows us to forget about the underlying sample space.

Discrete distributions: A random variable is called discrete if it takes finitely or countably many values. The corresponding distribution on $\mathbb{R}$ is called a discrete distribution. Popular discrete distributions include Bernoulli, binomial, geometric, and Poisson.

Continuous distributions: A distribution is called continuous if it assigns zero probability to every individual point in $\mathbb{R}$. Note that this definition implies that the corresponding random variable takes uncountably many values (why? hint: think about countable additivity). A distribution is called absolutely continuous if it has a density function. For our purposes, these definitions are equivalent. In general, absolute continuity implies continuity (check this), but not the other way. Even the easiest counterexample is quite technical and doesn't reflect any realistic data generating process. Important continuous distributions include the Normal, exponential, gamma, chi-squared, T, F and uniform.

Cumulative distribution functions (CDFs): The CDF of a distribution is a function from ~~R to~~ a point $x \in \mathbb{R}$ to the probability of not exceeding x. I.e. $F_X(x) := P(X \leq x)$, where X is the corresponding R.V. It turns out that the CDF of a distribution completely characterizes the distribution (i.e. using just F, it is possible to calculate the probability of any event in $\mathbb{R}$). ~~[illegible]~~ CDFs of discrete distributions are calculated by summing over all values up to x and absolutely continuous CDFs are calculated by integrating up to x. With absolutely continuous distributions, you can also obtain the density function by differentiating the CDF.

Some distributions are neither discrete nor continuous. The easiest example is a mixture distribution of a discrete and a continuous ~~dis~~ distribution. ¶ Such a distribution is harder to describe because neither a ~~[illegible]~~ probability mass function nor a density ~~[illegible]~~ tell the whole story. ¶ You can however, completely characterize such a distribution ~~[illegible]~~ by its CDF.