Lsn11

Clark

Review

We began this course by reviewing some basic concepts of probability theory that attempted to expand on what we had learned in MA206.

Recall that a *Sample Space* is the set consisting of all possible sample points and an *event* in a discrete sample space is a collection of sample points and in a continuous sample space is a subset of $\mathbb{R}$ (Actually a bit more complex that that, but for our purposes this is probably sufficient).

Probabilities then are measures of the likelihood of events that must obey three axioms:

In order to find these probabilities, if we have a discrete sample space we can use the sample-point method which:

Oftentimes counting techniques can be helpful. For instance, let's consider the homework problem from today. We have six students that we randomly assign grades to, what is the probability of getting two As, two Bs, and two Cs?

In order to answer this, we first need to know the size of the sample space. How many *simple events* possible?

Now how can we find the number of *simple events* in our event of interest?

In determining probabilities we next focused on some laws of probability that help us find probabilities when it is too difficult or cumbersome to actually count the simple events (or we are in a continuous world and cannot do so). The first tool that helps is the definition of *conditional probability*:

The next is the definition of *independence.*

We also have the *additive law of probability* and the *multiplicative law of probability*

And all of this work seemed to pinnacle at *Law of Total Probability* and *Bayes Rule*

While working with the actual sample space is intuitive, oftentimes we don't work with the events defined on the sample space, but rather we work with *Random Variables.* Recall that a random variable:

The real benefit of working with random variables is that it allows us to work with numbers rather than events. For example, we might have discrete random variables that take on the values:

Or Continuous Random variables that are defined on:

If we have discrete random variables, we can think of the *probability mass function* as the function that assigns probabilities for each possible outcome. Which sometimes we can find directly, like if we have five balls, numbered 1,2,3,4, and 5, placed in an urn and we draw two of them, we can let the random variable Y be the largest of the two sampled numbers. Using this information for Y we can write down the pmf as:

Or we can recognize the situation we are in and used a named distribution. For example, if a new surgical procedure is successful with probability $p = .8$ and the operation is performed five times and we let Y be the number of successful outcomes, we know the distribution of Y is:

Either way, we can use the pmf to find *expected values* or functions of expected values by

If our random variable is continuous we cannot define a pmf but we can define a *probability density function*. The pdf can be used to find probabilities but is not a probability in of itself. For example, if our pdf is denoted as $f(y)$ we can find the probability that our random variable, $Y \in (3, 6)$ through:

Our PDFs can be used similarly to our PMFs to find expected values or functiosn of expected values by replacing our summation above by

Regardless if we have a discrete random variable or a continuous random variable we can define our *cumulative distribution function*, $F(y)$ by finding $P(Y \leq y)$

Similar to our discrete random variables we can either give the function for a PDF or we can recognize what situation we are in and use a known/named PDF. The PDFs we had in our reading were Uniform, Normal, and Gamma of which the Exponential is a special case. If we are asked the expected value of our random variable and we *know* it is distribued Uniformly, well then there's no need to actually do any integration!

We extended these concepts to jointly defined random variables. In the discrete case we defined the joint pmf $p(y_1, y_2)$ as $P(Y_1 = y_1, Y_2 = y_2)$ and clearly if Y_1 and Y_2 are independent we can use properties of independence to say that $P(Y_1 = y_1, Y_2 = y_2)=$

Similar to our joint PMF we also have a joint pdf that we define through the joint CDF. Again, recall that for either discrete or continuous RVs our joint CDF, $F(y_1, y_2) = P(Y_1 \leq y_1, Y_2 \leq y_2)$. Using this, our joint PDF is defined as:

Joint distributions also can be averaged over one of the axis to obtain the marginal PDF. Recall that our marginal PDF can be found by:

and similarly our marginal PMF can be found by:

We also have the *conditional* PDF and PMF that are found by:

We can put all these pieces together to answer questions like. A soft-drink machine has a random about Y_2 in supply at the beginning of a given day and dispenses a random amount, Y_1 during the day. It is nto resupplied during the day and hence $Y_1 \leq Y_2$. The joint pdf is given by $f(y_1, y_2) = 1/2\mathbb{1}(0 \leq y_1 \leq y_2)\mathbb{1}(y_1 \leq y_2 \leq 2)$. What is the probability that less than 1/2 gallon will be sold, given that the machien contains 1.5 gallons at

the start of the day?

While joint PDFs/PMFs are often given to us, we can also use independence to derive a joint PDF. if X_1 and X_2 are independent, it follows that $f(x_1, x_2) = f(x_1)f(x_2)$. Which simpifies our work quite a bit. We can also note that this is a way to show that our random variables are independent if we can factor our joint PDF into our marginals. In fact we can go a step further if our support does not depend on our random variables (see Theorem 5.5).

We finally discussed expected values of functions of our jointly defined random variables, which we noted behave very much the same as univariate random variables. A special function of expected values is called the Covariance of X, Y and is found by:

And we note that if X and Y are independent then our covariance is zero. However, hopefully we saw in our homework that covariance being equal to zero does not necessarily mean that our random variables are independent!

Finally we talked about MGFs. Recall that an MGF is a special expected value

and if we have an MGF we can use it to find $E[X]$ by:

or $E[X^2]$ by:

or in general $E[X^k]$ by: