Math 227C: Introduction to Stochastic Differential Equations

Lecturer: Xiaohui Xie
Scribe: Hannah

Lecture #2
4/2/2104

Probability Generating Functions continued. Recall that a probability function $G(z)$ is a power series that can be used to determine the probability distribution of X, where X is a random variable taking discrete values from a finite set.

- $G(z) \stackrel{\text{def}}{=} E[z^X]$
- Suppose that $X = \{1, 2, ..., n\}$ and $P(X = k) = p_k$.
- Last time, we started to show that we can use the probability generating function to generate probabilities of X taking certain values. We first set z = 0 and computed the following:

$$G(0) = P(X = 0) = p_0$$
$$G^1(z)|_{z=0} = p_1$$
$$\vdots$$
$$G^{(k)}(z)|_{z=0} = p_k \cdot k!$$

- From this, all that remains is to isolate $P(X = k)$ by dividing be $k!$. We then have the following equation which lets us determine the probability distribution of X:

$$P(X = k) = \frac{G^{(k)}(z)|_{z=0}}{k!}$$

Moment Generating Functions. A moment generating function is also a function that can be used to determine the probability distribution of X. A moment generating function may not exist for a given random variable, but when it does exist the moment generating function will be unique.

- $M(t) \stackrel{\text{def}}{=} E[X^{tX}]$
- The *moments* of X are found by taking derviates with respect to time of the moment generating function, evaluated at $t = 0$. The first derivative is the first moment, the second derivative is the second moment, etc. This gives us the expectations of X^k:

$$M^1(t)|_{t=0} = E[Xe^{tX}]|_{t=0} = E[X]$$
$$\vdots$$
$$M^{(k)}(t)|_{t=0} = E[X^k e^{tX}]|_{t=0} = E[X^{(k)}]$$

- For example, for the Binomial random variable, which we discussed last class, the moment generating function is:

$$M(t) = E[e^{tX}] = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k}, \text{ or}$$
$$= E[(e^t)^X] = (e^t p + (1-p))^n$$

- Then, we can set t = 0 and evaluate the derivative of the moment generating function at this point. This gives us the following:

$$M'(k)|_{t=0} = npe^t(e^tp + 1 - p)^{n-1}|_{t=0} = np$$

- which is the expectation of X at the first moment of X, or just the expectation of X.

Coninuous time random variables. We can generalize these ideas to random variables with infinite (but still countable) time.

- To go from discrete to continuous time, start with a line segment divided into intervals of size Δt. Then let the interval size $\Delta t \to 0$. The number of intervals is $n = \frac{t}{\Delta t}$., so we can also think of this as letting the number of intervals $n \to \infty$.
- We can think of a continuous time random variable in the following way: There is a coin toss within each interval. The probability of a "success" (the coin lands on heads) within each interval is $\lambda \Delta t$. Then define X to be the number of sucesses within the interval (0,t). We can then make the time intervals smaller and smaller until we're left with continuous time.
- Let's see how this process works for the Binomial random variable. First, define the probability distribution over the discrete time intervals:

$$P(X = k) = \binom{\frac{t}{\Delta t}}{k} (\lambda \Delta t)^k (1 - \lambda \Delta t)^{\frac{t}{\Delta t} - k}$$

- The probability generating function can then be written as follows:

$$\begin{aligned} G(z) &= (1 + (z-1)\lambda \Delta t)^{\frac{t}{\Delta t}} \\ &= \exp[\frac{t}{\Delta t} ln(1 + (z-1)\lambda \Delta t)] \\ &\approx \exp[\frac{t}{\Delta t} ln(z-1)\lambda \Delta t] \end{aligned}$$

- and taking $\Delta t \to 0$ we get the probability generating function for the continuous random variable X:

$$\begin{aligned} \lim_{\Delta t \to \infty} (1 + (z-1)\lambda \Delta t)^{\frac{t}{\Delta t}} &= e^{(z-1)\lambda \Delta t \frac{t}{\Delta t}} \\ &= e^{\lambda t (z-1)} \end{aligned}$$

- As $\Delta t \to 0$, $p (= \lambda \Delta t) \to 0$ and $n (= \frac{t}{\Delta t}) \to \infty$. For the limiting case of $\Delta t = 0$

$$G(z) = e^{-\lambda} e^{z\lambda}$$

- We can then evaluate the probability generating function at z = 0, as was done in the discrete case:

$$G(0) = P(X = 0) = e^{-\lambda t}$$

$$\vdots$$

$$p(X = k) = \frac{G^{(k)}(z)|_{z=0}}{k!} = \frac{\lambda t^k}{k!} e^{-\lambda t}$$

- So when $\Delta t \to 0$, $P(X = k) = \frac{\lambda t^k}{k!} e^{-\lambda t}$ with $X \in \{0, 1, 2, ...\}$ (discrete but no upper limit). This is the Poisson distribution!
- For the Poisson distribution, there is one parameter, λ, which tells the rate of sucess (usually very small). Two things to note:
  1. Events are independent of each other
  2. Within a small interval Δt, the probability of seeing one event is $\lambda \Delta t$
- Some things to know about the Poisson distribution:
  - The sum of all the probabilities is 1:

$$\sum_{k=0}^{\infty} P(X = k) = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k} e^{-\lambda t} = e^{\lambda t} e^{-\lambda t} = 1$$

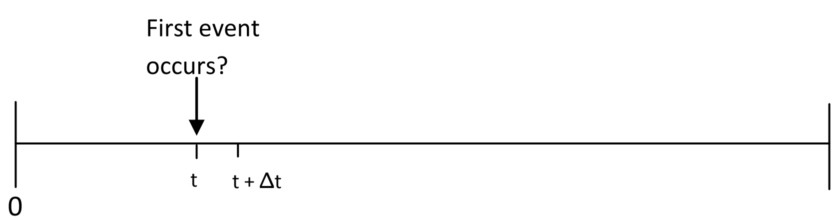
  - A special property of the Poisson distribution is that the expectation and variance are the same.
  - The expectation is:

$$E[X] = np = \frac{t}{\Delta t} \lambda \Delta t = \lambda t$$

  - and the variance is:

$$Var[X] = np(1 - p) = \frac{t}{\Delta t} \lambda \Delta t (1 - \lambda \Delta t) = \lambda t (1 - \lambda t)$$

  - We might want to determine the probability of the next event occuring at some specified time. Or, suppose we start at time $t = 0$ and want to know the probability of the first event occuring at a specifed time. In continuous time, we can determine this probability by considering the probability that an event will occur in some small interval starting at our specified time (and that no event occurs before this time interval).



  - Let T = the time to the next event. Then,

$$P(T \in [t, t + \Delta t]) = e^{-\lambda t} \lambda \Delta t$$

  - where $e^{-\lambda t}$ is the probability that no event occurs before t and $\lambda \Delta t$ is the probability that the event occurs in the $(t, t + \Delta t)$ time window.
  - This is an exponential distribution in continuous time:

$$f(t) = \lambda e^{-\lambda t}$$

Probability Spaces and Sigma-algebras. A probability space consists of the following three items: $(S, \mathcal{P}, \mu)$: a set of possible outcomes, a collection of subsets of these outcomes, and a probability measure on this collection of subsets. When the number of possible outcomes is infinite, we use a σ-algebra.

- There are many possibilities for what kind of sample space, S, we may want to consider:
  1. S can be finite.
  2. S can be infinite
     (a) countable
     (b) or uncountable
  3. or we can look at subsets of S. In particular, we might want to consider the *powerset* of S, which is composed of all possible subsets of S
     (a) If S is finite: the number of subsets is $2^{|S|}$
     (b) If S is infinite: the powerset of S is uncountable. We then want to use a σ-algebra.
- A probability space consists of the following three items: $(S, \mathcal{P}, \mu)$
  - S is a set.
  - $\mathcal{P}$ is a collection of subsets meeting the following four conditions:
    1. $\emptyset \in \mathcal{P}$
    2. $S \in \mathcal{P}$
    3. Any countable union of sets is contained in $\mathcal{P}$
    4. If $A \in \mathcal{P}$, $S \backslash A \in \mathcal{P}$, where $S \backslash A = S - A = \{x \in S | x \notin A\}$ is the complement of A.
  - μ is a map $\mathcal{P}$ in $[0, 1]$ s.t.

$$\mu(\bigcap_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i), \text{ if } \{A_i\} \text{ is a disjoint collection, and}$$

$$\mu(S) = 1$$

- Suppose $S = \mathbb{R} = (-\infty, \infty)$ and $(a, b) \in \mathcal{P}$. The *Borel Set* is the smallest σ-algebra containing open intervals. We almost always assume this set when working with σ-algebras.

Continuous Variables. So far we have been talking about discrete random variables, first in deiscrete time and then in continuous time. We can also talk about continuous random variables.

- For example, consider the one-dimensional Gaussian density function:
  - $X \in \mathbb{R}$ That is, X can take an infinite number of values.
  - The probability density function is given by $\rho(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}$
  - The probability that X will take a value within a certain interval is given by $P\{x \in [x, x + dx]\} = \rho(x) \cdot dx$

– Some properties of the one-dimensional Gaussian (these will be elaborated on next time):

1. $\int_{-\infty}^{\infty} \rho(x)dx = 1$
2. $E[X] = \int_{-\infty}^{\infty} x\rho(x)dx = m$
3. $Var[X] = \int_{-\infty}^{\infty} (x-m)^2\rho(x)dx = \sigma^2$