Math 217 Probability and Statistics

Prof. D. Joyce, Clark University

Friday, 19 Oct 2007

Due Monday. From Chapter 5.1: 1, 2, 4, 8.

Quiz Wednesday.

Due Friday. From Chapter 5.1: 7, 9, 16, 18, 25.

Last time. Discussion on uniform distributions (both discrete and continuous) and distributions related to the Bernoulli trial process, namely the Bernoulli distribution (either success or failure), binomial distribution (sampling with replacement), the geometric distribution, the negative binomial distribution, and the hypergeometric distribution (sampling without replacement).

Today. The Poisson process and related distributions. The Poisson process is a mathematical model for the idea of random events occurring over time at a constant rate. It's easy to understand what this means when the events aren't random, but deterministic. For instance a clock that ticks every second has 60 events/minute, and they're evenly spaced, exactly 1 second apart. But when the events are random, it's a more difficult to pin down exactly what it means for events to occur at a constant rate. Poisson did that.

Axioms for the Poisson process. We assume that events occur over time subject to the following axioms.

(1). The number of events in nonoverlapping regions are independent.

(2). The probability of an event occurring in an interval $[a, a+h]$ of short length h is approximately proportional to h, with proportionality ratio λ, called the rate of events. More precisely,

$$\lim_{h\to 0} \frac{P(\text{an event in } [a, a+h])}{h} = \lambda.$$

(3). The probability of two events occurring in a short interval $[t, t+h]$ is much smaller than the probability of one event. More precisely,

$$\lim_{h\to 0} \frac{P(\text{two or more events in } [a, a+h])}{P(\text{an event in } [a, a+h])} = 0.$$

Development from the axioms. We'll use these axioms to develop various related probability distributions. Let's use the notation $P_0(t)$ to stand for the probability that no events take place in an interval of length t. We'll figure out the derivative P_0' in terms of P_0.

First, note that

$$P_0(t+h) = P_0(t)\, P_0(h).$$

That's because if there are no events in the interval $[0, t+h]$, then there are none in the two intervals $[0, t]$ and $[t, t+h]$, and since those are nonoverlapping regions, the probabilities are independent, and so multiply. Therefore,

$$\begin{aligned} P_0'(t) &= \lim_{h\to 0} \frac{P_0(t+h) - P_0(t)}{h} \\ &= \lim_{h\to 0} \frac{P_0(t)\, P_0(h) - P_0(t)}{h} \\ &= P_0(t) \lim_{h\to 0} \frac{P_0(h) - 1}{h} \end{aligned}$$

Let's pause for a moment. The probability $P_0(h)$ that no events occur in $[t, t+h]$ equals $1 - P(\text{at least one event in } [t, t+h])$. That gives us

$$\begin{aligned} & P_0'(t) \\ &= P_0(t) \lim_{h\to 0} \frac{-P(\text{at least one event in } [t, t+h])}{h} \\ &= -\lambda P_0(t) \end{aligned}$$

We now have a differential equation

$$P_0'(t) = -\lambda P_0(t).$$

We'll see in class why this differential equation, along with the extra information $P_0(0) = 1$, has the solution

$$P_0(t) = e^{-\lambda t}.$$

Next, let $P_1(t)$ be the probability that exactly one event takes place in an interval of length t. Then

$$\begin{array}{rl} & P_1'(t) \\ = & \lim\limits_{h\to 0} \dfrac{P_1(t+h) - P_1(t)}{h} \\ = & \lim\limits_{h\to 0} \dfrac{P_1(t)\,P_0(h) + P_0(t)\,P_1(h) - P_1(t)}{h} \\ = & \lim\limits_{h\to 0} \dfrac{P_1(t)\,P_0(h) - P_1(t)}{h} + \lim\limits_{h\to 0} \dfrac{P_0(t)\,P_1(h)}{h} \\ = & P_1(t)\,\lim\limits_{h\to 0} \dfrac{P_0(h) - 1}{h} + P_0(t)\,\lim\limits_{h\to 0} \dfrac{P_1(h)}{h} \\ = & -\lambda P_1(t) + \lambda P_0(t). \end{array}$$

Note that axiom 3 was used in the last equation. We now have the differential equation

$$P_1'(t) = -\lambda P_1(t) + \lambda e^{-\lambda t}.$$

That's not such an elementary equation as the first one, but can be solved, along with the extra information $P_1(0) = 0$, to get

$$P_1(t) = \lambda t e^{-t}.$$

In general, if we let $P_n(t)$ be the probability that exactly n events take place in an interval of length t, we can derive a differential equation

$$P_n'(t) = -\lambda P_n(t) + \lambda P_{n-1}(t)$$

which has the solution

$$P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

This is the Poisson distribution. If X is the number of events in an interval of length t, then

$$P(X{=}n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

The exponential distribution. We've already studied the exponential distribution, but now we can justify the claims we made for it. Let T be the time of the first event after $t = 0$. Then T is what we've called the exponential distribution. It's very closely related to the function $P_0(t)$ we found above. Let $F_T(t)$ denote the cumulative distribution function for the exponential distribution. Then

$$\begin{array}{rcl} F_T(t) & = & P(T \le t) \\ & = & 1 - P(T \ge t) \\ & = & 1 - P_0(t) = 1 - e^{-\lambda t} \end{array}$$

That gives us the justification for the formula we've been using for the exponential distribution. Differentiating it gives us the density function

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