Random Variables - Class 14

October 10, 2013 Debdeep Pati

1 Random variables

1.1 Poisson random variable

In probability theory and statistics, the Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate and independently of the time since the last event. (The Poisson distribution can also be used for the number of events in other specified intervals such as distance, area or volume). Suppose someone typically gets on the average 4 pieces of mail per day. There will be however a certain spread: sometimes a little more, sometimes a little less, once in a while nothing at all. Given only the average rate, for a certain period of observation (pieces of mail per day, phonecalls per hour, etc.), and assuming that the process, or mix of processes, that produce the event flow are essentially random, the Poisson distribution specifies how likely it is that the count will be 3, or 5, or 11, or any other number, during one period of observation. That is, it predicts the degree of spread around a known average rate of occurrence.

The probability of i events in a time period t for a Poisson random variable with parameter λ (μ is also commonly used) is

$$P(X = i) = e^{-\lambda}\frac{\lambda^i}{i!}, i = 0, 1, 2, \ldots, \infty$$

where $\lambda = r \times t$, where r represents expected number of events per unit time.

1. Parameter λ represents expected number of events over time period t.
2. Difference between Binomial and Poisson distribution: There are a finite number of trials n in Binomial distribution The number of events can be infinite for Poisson distribution
3. The event can occur in a period of time or in a particular area

1.2 Example of Poisson random variable

1. The number of misprints on a page of a book
2. The number of people in a community living to 100 years of age

3. The number of wrong telephone numbers that are dialed in a day
4. The number of packages of dog biscuits sold in a particular store each day
5. The number of customers entering a post office on a given day
6. The number of vacancies occurring during a year in the federal judicial system

1.3 Problem

Suppose that the number of typographical errors on a single page of a book has a Poisson distribution with parameter $\lambda = 1/2$. Calculate the probability that there is at least one error on one page. (Let X denote the number of errors on this page, then $P(X \geq 1) = 1 - P(X = 0) = 1 - \exp\{-1/2\}$)

1.4 Poisson approximation of the Binomial distribution

Suppose that the probability that an item produced by a certain machine will be defective is 0.1. Find the probability that a sample of 10 items will contain at most 1 defective item.

1. Binomial distribution with parameters (10, 0.1)

$$\binom{10}{0}(0.1)^0(0.9)^{10} + \binom{10}{1}(0.1)^1(0.9)^{10-1} = 0.7361$$

2. Poisson distribution with parameter 1. $e^{-1} + e^{-1} = 0.7358$
3. Poisson can be considered as the limit of binomial
4. $B(n,p) \sim Poi(\lambda)$ for large n, provided $np \to \lambda$.
5. At a certain place Poisson random variable can be used as an approximation for a binomial random variable with parameters (n,p) when n is large and p is small so that np is a moderate size.
6. Poisson distribution is a good approximation of the binomial distribution if n is at least 20 and p is smaller than or equal to 0.05, and an excellent approximation if $n \geq 100$ and $np \leq 10$, $\lambda = np$.
7. $E(X) = \lambda, V(X) = \lambda$
8. Let $X \sim Poisson(\lambda)$. What value of k maximize $P(X = k)$?
9. $P(X = k)/P(X = k-1) = \lambda/k$

10. When $k = \lfloor \lambda \rfloor$ (the largest integer $\leq \lambda$), $P(X = k)$ reaches its max

11. Note that $k_{max} \approx EX$

12. Therefore, for both Poisson and Binomial, $p(k)$ first increases and then decreases

1.5 Examples

1. Suppose the # of accidents occurring on a highway each day is a Poisson r.v. with parameter $\lambda = 3$. Find the probability that 3 or more accidents occur today given that at least 1 accident occurs today. $P(X \geq 3 \mid X \geq 1)$

2 Other Discrete Probability Distributions

2.1 Geometric distribution:

Independent trials, each having a probability p of being a success, are performed until a success occurs. If X is the number of trials required, then $P(X = n) = (1 - p)^{n-1}p, n = 1, 2, \cdots \infty$ We say X is a geometric variable with parameter p.

1. $E(X) = 1/p, V(X) = (1 - p)/p^2$

2. Memoryless property: $P(X = n + k \mid X > n) = P(X = k)$