October 14, 2012 Debdeep Pati

1 Random variables

1.1 Expectation of a function of a random variable

1. Expectation of a function of a random variable
2. We know $E(X) = \sum_x xp(x)$
3. What is E(g(X))? (expectation of a transformed r.v.)
4. Note that g(X) is also a random variable.
5. In fact, any function of a random variable is also a random variable
6. Let $Y = g(X)$. Then $E(g(X)) = E(Y)$.
7. $P(X = -1) = 0.2, P(X = 0) = 0.5, P(X = 1) = 0.3$
8. What is $E(X^2)$?
9. Let $Y = X^2$, still a r.v. since $Y(s) = X(s)X(s), s \in S$.
10. What is the pmf of Y ?
11. $P(Y = 0) = P(X^2 = 0) = P(X = 0) = 0.5, P(Y = 1) = P(X^2 = 1) = P(X = 1 or - 1) = 0.2 + 0.3 = 0.5$
12. $E(Y) = 1 \times 0.5 + 0 \times 0.5 = 0.5$
13. Therefore, $E(X^2) = 0.5$.
14. Note that $E(X^2) \neq E(X)E(X)$.
15. Proposition. If X is a discrete r.v. that takes on one of the values $x_i, , i \geq 1$, with respective probabilities $p(x_i)$ (so p() is a pmf), then for any real-valued function g $E(g(X)) = \sum_i g(x_i)p(x_i)$.
16. For example, $E(X^2) = (-1)^2 \times 0.2 + 1^2 \times 0.3 = 0.5$.
17. Corollary: Corollary. $E(aX + b) = aE(X) + b$, where a and b are constants.
18. Moments and Variance (raw moments and central moments) kth raw moment: $E(X^k)$, kth central moment: $E((X - E(X))^k)$

19. Some propositions. $Var(X) = E(X^2) - (EX)^2$, $Var(b + aX) = a^2Var(X)$, where a, b are constants.

20. What is Var(X) for the former r.v. X? ($P(X = -1) = 0.2, P(X = 0) = 0.5, P(X = 1) = 0.3$)

21. What is $Var(2X)$?

22. $E(X)$: a measure of the location of X

23. $Var(X)$: a measure of the spread (scale) of X

24. $SD(X) = \sqrt{Var(X)}$: the standard deviation of X

2 Commonly used random variables

2.1 Bernoulli Random variable

1. Example. Flip a coin. If the outcome is heads, $X = 1$, otherwise $X = 0$.

2. Then X is a Bernoulli r.v., or X follows a Bernoulli distribution.

3. Denition. $P(X = 1) = p, P(X = 0) = 1?p$, where $p \in (0, 1)$.

4. Then $X \sim Bernoulli(p)$

5. $E(X) = p$, $Var(X) = p(1 - p)$

2.2 Binomial Random Variable

1. Example. Flip five coins independently. X : the number of heads. Find the p.m.f. of X.

2. $P(X = 0)$

3. $P(X = k)$

4. Consider a general problem: Perform n independent trials each of which results in a success w.p. p and in a failure w.p. 1 - p. Let X represent the number of successes that occur in these n trials. What is the pmf of X?

5. X is a binomial random variable with parameters (n, p).

2.2.1 Properties

1. Let $X \sim Bin(n,p)$ with $0 < p < 1$. Then as k goes from $0 \to n$, $P(X = k)$ rst increases monotonically then decreases monotonically, reaching its largest value when k is the largest integer $\leq (n+1)p$. $\frac{p(k)}{p(k-1)} = \frac{(n-k+1)p}{k(1-p)}$

2. Discuss skewed binomial distribution

3. Let $X \sim B(n,p)$. $E(X) = np, Var(X) = np(1-p)$ Use the binomial theorem! See Ross (P139).

Problem

It is known that screws produced by a certain company will be defective with probability .01 independently of each other. The company sells the screws in packages of 10 and offers a money-back guarantee that at most 1 of the 10 screws is defective. What proportion of packages sold must the company replace?

If X is the number of defective screws, then X is a binomial random variable with parameters $(10, .01)$. The probability that a package will have to be replaced is $1 - P(X = 0) - P(X = 1)$

Problem

In a gambling game, a player bets on one of the numbers 1 through 6. Three dice are then rolled, and if the number bet by the player appears i times, i = 1,2,3, then the player wins i units; on the other hand, if the number bet by the player does not appear on any of the dice, then the player loses 1 unit, Is this game fair to the player? If we assume that the dice are fair and act independently of each other, then the number of times that the number bet appears is a binomial random variable with parameter (3, 1/6) Let X denote the players winnings, we have

$$\begin{aligned} P(X=-1) &= bin(3,1/6,0) = 125/216, P(X=1) = bin(3,1/6,1) = 75/216, \\ P(X=2) &= bin(3,1/6,2) = 15/216, P(X=3) = bin(3,1/6,3) = 1/216, \end{aligned}$$

$E(X) = -17/216$.

3 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, phone calls 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 $\lambda = r \times t$.

1. Parameter r represents expected number of events per unit time.
2. Parameter λ represents expected number of events over time period t.
3. 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
4. The event can occur in a period of time or in a particular area
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 $np \leq 10$, $\lambda = np$.

3.1 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

3.2 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 $P(X \geq 1) = 1 - P(X = 0) = 1 - \exp\{-1/2\}$)

3.3 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. If n independent trials each of which results in a success w.p. p (so each ? Bernoulli(p)), then if $np \to \lambda$ as $n \to \infty$, then the number of successes occurring (? B(n, p)) is approximately $Poi(\lambda)$ when n is large.
5. That is, $B(n,p) \sim Poi(\lambda)$ for large n, provided $np \to \lambda$.
6. A nonrigorous argument (not required)
7. It is interesting to note that this can also be true when the independency fails or the success probabilities are not equal to each other
8. $E(X) = \lambda, V(X) = \lambda$