MAT 105 - Summary of Lecture 14
Thursday, February 25, 2016

How dare we speak of the laws of chance? Is not chance the antithesis of all law? (Joseph Bertrand)

Discrete random variables (cont.)

Some random variables are widely used and knowing that an experiment can be modeled using a certain well known random variable is helpful. For example, we compute the expectation and the variance of a random variable for the general case and then we can apply them in specific examples.

Suppose we roll 3 fair 6-sided dice 432 times. We want to know, on average, how many times we roll 111. We recognize that we can model the number of times we roll 111 with a random variable X that counts them. Since a 111 occurs with probability $\frac{1}{216}$ and we roll 432 times, $X = Binomial(432, 1/216)$. Instead of finding all probabilities $P(X = 0), P(X = 1), \ldots, P(X = 432)$ and then finding the expectation, we already know that $E(X) = 432 \cdot (1/216) = 2$.

1. **Geometric Random Variables**

   Suppose we play darts any I hit the target with probability 0.3. I want to play until I **first** hit the target. Let X count how many times I play. Then I can play anywhere between 1 and infinitely many games before hitting the target.

$$\begin{aligned} P(X=1) &= P(\text{win 1st}) = 0.3 \\ P(X=2) &= P(\text{lose 1st, win 2nd}) = (0.7)(0.3) = 0.21 \\ P(X=3) &= P(\text{lose 1st, 2nd, win 3rd}) = (0.7)(0.7)(0.3) = (.7)^2(0.3) = 1.47 \\ P(X=100) &= P(\text{lose first 99, win 100th}) = (0.7)^{99}(0.3) \end{aligned}$$

   Definition: Let X count the number of trials until the **first** success, with probability of success p in each trial. Then X is a **geometric** random variable with parameter p, denoted by $X = Geom(p)$ and whose distribution is defined for $k \in \{1, 2, 3, \ldots\}$ by

$$P(X = k) = (1 - p)^{k-1} p.$$

   Remark: Note that this is an example of a random variable with infinitely many possible values. One can check that all probabilities add up to 1 by computing the infinite sum

$$\sum_{k=1}^{\infty} P(X = k) = \sum_{k=1}^{\infty} (1 - p)^{k-1} p = 1,$$

   but that is beyond the scope of this class. In fact, the reason why we call this random variable geometric is because the probability distribution involves the infinite sum above, which is known as a "geometric series" in mathematics. Also note that computing the expectation of this random variable involves an infinite sum, which we will not compute, but rather just use the result.

   The expectation of $X = Geom(p)$ is $E[X] = \frac{1}{p}$ and it represents the average number of trials until the first success.

   In our darts example, with $X = Geom(0.3)$, I should play an average of $E[X] = \frac{1}{0.3} \approx 3.33$ games until the first success.

2. **Negative Binomial Random Variables**

Continuing with the darts example, where my probability of hitting the target in each trial is 0.3, suppose I play until I hit the target exactly 10 times. Let X count how many times I play. Then I can play anywhere between 10 and infinitely many games before hitting the target exactly 10 times.

$$\begin{aligned} P(X=10) &= P(\text{win first }10) = (0.3)^{10} \\ P(X=18) &= P(\text{win 18th, win 9 out of the first }17) = \binom{17}{9}(0.7)^8(0.3)^{10} \end{aligned}$$

The binomial coefficient is used to count all possible ways to win 9 out of the first 17 trials; each success has probability 0.3 and each failure has probability 0.7; there are 10 successes and 8 failures in 18 trials; thus, we have the factors $(0.3)^{10}$ and $(0.7)^8$.

Definition: Let X count the number of trials until the r^{th} success, with probability of success p in each trial. Then X is a **negative binomial** random variable with parameters r and p, which we will denote by $X = NBin(r,p)$ and whose distribution is defined for $k \in \{r, r+1, \dots\}$ by

$$P(X=k) = \binom{k-1}{r-1}(1-p)^{k-r}p^r.$$

The expectation of $X = NBin(r,p)$ is $E[X] = \frac{r}{p}$, and it is the average number of trials until I win r times. This makes sense because if we have to wait about $\frac{1}{p}$ for each success, we will wait $r \cdot \frac{1}{p}$ for r successes. This also establishes a strong connection between geometric random variables and negative binomials, namely that a negative binomial $NBin(r,p)$ is a sum of r geometric random variables with parameter p.

If I play darts until I win 10 times, with $X = NBin(10, 0.3)$, I should play an average of $E[X] = \frac{10}{0.3} \approx 33.33$ games until 10 successes occur.

3. **Poisson Random Variables**

Another useful random variable is the Poisson random variable, introduced by S.D.Poisson in 1837, in a book about applications of probability theory to lawsuits and criminal trials. This random variable is used to model quantities that are given in terms of averages. For example, random variables that obey the Poisson probability law are:

- the number of misprints on a page
- the number of accidents in Redmond today
- the number of people in a community who live to be 100 years old
- the number of α-particles discharged in a fixed period of time from a radioactive material
- the number of wrong phone numbers dialed in a day
- the number of DigiPen students fainting in school today

Definition: The **Poisson random variable** with parameter λ (lambda) and denoted by $Poisson(\lambda)$ has distribution

$$P(X=k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!},$$

for $k \in \{0, 1, 2, \dots\}$.

The expectation and variance of X are $E[X] = \lambda$ and $Var(X) = \lambda$.

Examples:

(a) The number of typographical errors in a textbook from a certain publisher has an average of $\lambda = 0.5$ errors per page. Find the probability that there is at least one error on the page you are reading.

Let X count the number of errors on the page. We want to find $P(X \geq 1)$. Note that if we knew how many words were on the page, we could use a binomial random variable to model X. But we do not know that information. Because the random variable is given in terms of its average, we recognize that in fact $X = Poisson(0.5)$. Then

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \left(e^{-0.5} \cdot \frac{(0.5)^0}{0!}\right) = 1 - e^{-0.5} \approx 0.39.$$

(b) Suppose that earthquakes occur in western US at an average rate of 2 per week. Find the probability that exactly 3 earthquakes occur this week.

Let X = number of earthquakes. Then $X = Poisson(2)$ and $P(X = 3) = e^{-2} \cdot \dfrac{2^3}{3!} = \dfrac{4e^{-3}}{3} = 0.18$

The Poisson random variable can be used to approximate large binomials. More precisely, if $X = Binomial(n, p)$, with n large, p small and np moderate (np between .2 and 20), then

$$P(X = k) \approx e^{-np} \cdot \frac{(np)^k}{k!}.$$

Another way to write this is $Binomial(n, p) \approx Poisson(np)$. Here are two examples on how the approximation works.

3. Suppose the probability that a driver is in an accident is 0.0004 on any given day. If about 10000 drivers pass through Redmond today, approximate

   (a) the probability of no accidents;

   (b) the probability of exactly 2 accidents;

   (c) on average, how many accidents will be in Redmond today?

   Let X = number of accidents, so $X = Binomial(10000, 0.0004)$. First we find **exact probabilities**

   (a) $P(X = 0) = \binom{10000}{0}(.0004)^0(.9996)^{10000} = 0.01830$

   (b) $P(X = 2) = \binom{10000}{2}(.0004)^2(.9996)^{9998} = 0.14651$

   (c) $E[X] = (10000)(.0004) = 4$.

   Approximating with the Poisson, we let $\lambda = (10000)(0.0004) = 4$ and then

   (a) $P(X = 0) \approx e^{-4} \cdot \frac{4^0}{0!} = e^{-4} = 0.01831$

   (b) $P(X = 2) \approx e^{-4} \cdot \frac{4^2}{2!} = 8e^{-4} = 0.14652$

   (c) $E[X] = \lambda = 4$. Note that the mean of the binomial always agrees with the mean of the Poisson used to approximate!

   Remark that the approximation is quite accurate!

4. Suppose that the probability that an item produced by a certain machine will be defective is 0.1. In a sample of 10 items,

   (a) find the exact probability of at most one defective item;

(b) approximate, using a Poisson random variable, the probability of at most one defective item.

(a) Let X count the number of defective items, in a set of 10. Then $X = Binomial(10, 0.1)$, so

$$P(X \leq 1) = P(X = 0)+P(X = 1) = \binom{10}{0}(0.1)^0(0.9)^{10}+\binom{10}{1}(0.1)^1(0.9)^9 = (0.9)^{10}+(0.9)^9 = 0.7361$$

(b) Let $\lambda = (10)(0.1) = 1$ and approximate X with a $Poisson(1)$. Then

$$P(X \leq 1) = P(X = 0) + P(X = 1) = e^{-1} \cdot \frac{1^0}{0!} + e^{-1} \cdot \frac{1^1}{1!} = 2e^{-1} = 0.7358.$$

NAMES:

MAT 105 - Group Work - February 25, 2016

Hand in ONE solution per group. Pay special attention to which random variables you use.

1. On a multiple choice exam, with 3 possible answers for each of the 7 questions, what is the probability that a student would get 5 or more correct just by guessing?

2. A certain typing agency employs 2 typists. The average number of errors per article is 3 when typed by the first typist and 2 when typed by the second.
   (a) If your article is typed by the first typist, find the probability that it will have no errors.

   (b) If your article is typed by the second typist, find the probability that it will have no errors.

   (c) If your article is equally likely to be typed by either typist, find the probability that it will have no errors.

OVER $-->$

3. We play a game for which we have a 1/20 chance of winning. We play the game 100 times.
   (a) What is the *expected* number of times we win?

   (b) What it the *expected* number of times we play until the **first** win?

   (c) What it the *expected* number of times we play until we win 10 times?

   (d) Find the probability that the **first** win occurs on the 8 trial.

   (e) Find the exact probability that we win 10 times out of 100 (leave answer in terms of factorials).

   (f) Use a Poisson random variable to approximate the probability we win 10 times out of 100.