STA 111 (Summer Session I)

Lecture Three – Introduction to Random Variables

Instructor: Olanrewaju (Michael) Akande

Department of Statistical Science, Duke University

May 20, 2016

Outline

- Questions from Last Lecture.
- Introduction
- Random Variables
- The Binomial Distribution
- The Hypergeometric Distribution
- The Poisson Distribution
- The Negative Binomial Distribution
- Recap

Introduction

– In the last lecture we reviewed some counting methods and talked about conditional probability.

– We also talked about independence and Bayes' rule.

– We now know how to calculate the probability of events if we know what the sample space looks like.

– Today we will move on to define real-valued functions on sample spaces and learn how to calculate their probabilities. Specifically, we will learn about real-valued functions of sample spaces, called random variables, and their distributions.

– Lastly, we will learn about some special and very useful discrete distributions.

Random Variables

– A random variable is a function that maps the sample space to the real line. That is, if X is a random variable, then X maps every element of Ω to a real number. If "X" is a random variable, we will write "x" as its observed value.

– *Example 1:* Back to our toy example. Suppose you toss a fair coin twice and let X be the number of heads observed. Then $\Omega = \{HH, HT, TH, TT\}$, and

$$X(HH) = 2,\ X(HT) = 1,\ X(TH),\ X(TT) = 0$$

$$\text{Then,}\quad \mathbb{P}(X = 0) = \mathbb{P}(\{TT\}) = \frac{1}{4},\ \ \mathbb{P}(X = 2) = \mathbb{P}(\{HH\}) = \frac{1}{4s}$$

$$\mathbb{P}(X = 1) = \mathbb{P}(\{HT, TH\}) = \frac{2}{4} = \frac{1}{2}$$

– Clearly, we can define several random variables on the same sample space. Can you think of another random variable defined on Ω?

Examples

Example 2: Suppose you roll a fair die five times and let X be the number of 4's observed. Then writing out the sample space is tedious but clearly, $X \in \{0, 1, 2, 3, 4, 5\}$. So that if for example we observe $\omega = (1, 1, 3, 4, 2) \in \Omega$, $X(\omega) = 1$, but if we observe $\omega = (4, 4, 5, 4, 2) \in \Omega$, $X(\omega) = 3$.

What then is the probability that $X = 2$? Since this is a fair die, $\mathbb{P}$(observing 4 on any roll) $= \frac{1}{6}$ and $\mathbb{P}$(observing any other number on any roll) $= \frac{5}{6}$. For X to be 2, exactly two of the rolls must be 4 and exactly three must be any number besides 4. Lastly, there are $\binom{5}{2}$ ways of getting two rolls with "4" from five rolls.

$$\Rightarrow \mathbb{P}(X = 2) = \binom{5}{2} \times \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} = \binom{5}{2} \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^3 = 0.1608$$

– Turns out there's a name for the collection of the probabilities for this random variable. We'll get to that soon. First, a few definitions.

Some Definitions

– The distribution of a random variable X is the collection of all probabilities of the form $\mathbb{P}(X \in C)$ for all sets C of real numbers such that $\{X \in C\}$ is an event.

– A random variable X has a discrete distribution if X can only take on a finite number of different values or a sequence of countably infinite values. For example, $X = 0, 1, 2$ and 3 implies X is a discrete random variable.

– A random variable X has a continuous distribution if X can take on infinite values. For example, $X \in [0, 1]$ implies X is a continuous random variable.

– Probability Mass Function (pmf): If X has a discrete distribution, the pmf is defined as a function f such that for every real number x,

$$f(x) = \mathbb{P}(X = x); \quad \text{where} \quad \sum_{i=1}^{\infty} f(x_i) = 1$$

– I will continue to use $f(x)$ and $\mathbb{P}(X = x)$ interchangeably. Note that,

$$\mathbb{P}(X \leq c) = \sum_{i:x_i \leq c} \mathbb{P}(X = x_i)$$

Binomial Distribution

– Some distributions of random variables are very useful in statistics. One of them is the binomial distribution.

– The binomial formula gives the probability of exactly x successes in n tries, where each try has the same probability of success p and each try is independent.

$$\mathbb{P}(\text{exactly } x \text{ successes}) = \mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x};\ x = 0, 1, \ldots, n.$$

– This is exactly the situation one has when trying to find the probability of x heads in n tosses of a coin that has probability p of coming up heads. The random variable X is said to have the binomial distribution $\text{Bin}(n, p)$.

– When $n = 1$, the distribution is called a Bernoulli distribution. Thus, the binomial distribution results from independent Bernoulli trials.

The Intuition behind the Binomial Formula

– Why does this formula work? We already discussed this idea in example 2, but we need to make the idea general.

- How many arrangements are there that give x heads in n tries? From the previous class, we know the answer is $\binom{n}{x}$.
- Each arrangement is different from the other arrangements. Thus the probability of exactly x successes is the sum over all possible arrangements, and there are $\binom{n}{x}$ of those.
- Each arrangement has been the same probability: $p^x(1-p)^{n-x}$. To see this, consider the sequences HTHT and THTH from four independent coin tosses. The first has probability $p \times (1-p) \times p \times (1-p) = p^2(1-p)^2$. The second arrangement has probability $(1-p) \times p \times (1-p) \times p$ $= p^2(1-p)^2$.
- Combining all pieces gives $\mathbb{P}(\text{exactly } x \text{ successes}) = \binom{n}{x}p^x(1-p)^{n-x}$

Examples

Example 3: What is the probability of exactly three fives in six rolls of a fair die?

$$\begin{aligned}\mathbb{P}(\text{exactly 3 successes}) &= \binom{n}{x} p^x (1-p)^{n-x} \\ &= \binom{6}{3} (\frac{1}{6})^3 (1-\frac{1}{6})^{6-3} \\ &= 20 \times (\frac{1}{6})^3 \times (\frac{5}{6})^3 = 0.0536\end{aligned}$$

– *Example 2 again:* Suppose you roll a fair die five times and let X be the number of fours observed. Then X has the binomial distribution $\text{Bin}(5, \frac{1}{6})$. Using the formula,,

$$\mathbb{P}(X=2) = \binom{5}{2} \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^3 = 0.1608 \quad \text{Same as before.}$$

Examples

Example 4 (to be done in class): – Suppose you are on a spin wheel game show and the spin wheel has numbers 1 to 10.

– You will win $2 every time the spin wheel stops at either numbers 2, 6 or 7.

– If even numbers are twice as likely to show up as odd numbers, but all even numbers are equally likely between them and all odd numbers are also equally likely between them, what is the probability that you will win $6 in 6 independent spins? What about the probability that you will win at least $4?

Poisson Distribution

– Another important distribution is the Poisson distribution.

– The Poisson distribution gives the probability of exactly x occurrences in a fixed period of time or a fixed area. For example,

- The number of calls to a telephone switchboard in a minute.
- The number of accidents on the highway in one month.
- The number of thefts on Duke campus between 2001 and 2016.

– The Poisson is used when each event is independent, the probability of each is fixed, and the number of potential events is unbounded.

Poisson Distribution

– Suppose the mean number of events is λ. Then X has the Poisson distribution Pois(λ) and the probability of exactly x events is:

$$\mathbb{P}[\text{ exactly } x \text{ events }] = \frac{\lambda^x}{x!} e^{-\lambda}; \quad x = 0, 1, 2, \ldots$$

– *Example 5:* Military analysts divided London into 576 areas and studied 500 V-2 attacks. So the average number of attacks per block was $500/576 = 0.868$.

– What is the probability that Buckingham Palace was undamaged?

$$\mathbb{P}[\text{ 0 hits on BP }] = \frac{0.868^0}{0!} e^{-0.868} = 0.419.$$

– What is the probability of at most one hit on Buckingham Palace?

$$\mathbb{P}[\text{at most one hit on BP }] = \frac{0.868^0}{0!} e^{-0.868} + \frac{0.868^1}{1!} e^{-0.868} = 0.784.$$

Examples

– *Example 6 (D.S. 5.4.1 – to be done in class):* A store owner believes that customers arrive at his store at a rate of 4.5 customers per hour on average. What is the probability that 5 customers would arrive in the next hour?

– There is an interesting relationship between the binomial and Poisson distributions. Turns out the $\text{Bin}(n, p)$ can be approximated by $\text{Po}(\lambda = np)$ as $n \rightarrow \infty$ and $p \rightarrow 0$. The larger the n and the smaller the p, the better the approximation. The mathematical development is provided in the textbook for the energetic student.

Hypergeometric Distribution

– The hypergeometric distribution describes the number of successes in n draws from a finite population of size N in which M of the N elements are successes.

$$\mathbb{P}[\text{ exactly } x \text{ successes }] = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}};\ \max\{0, n-(N-M)\} \leq x \leq \min\{n, M\}$$

This distribution often arises in quality assurance and lot inspection.

– This pmf is more intuitive than it looks at first. First, there are $\binom{N}{n}$ ways to pick n elements from a population of size N. This is the total number of elements in our sample space. All we need to do next is figure out how many of those will give us x successes.

– Well, if there are M total successes, there are $\binom{M}{x}$ ways to choose x successes from M possible successes. Lastly, to have x successes, there must be $n-x$ failures as well with $\binom{N-m}{N-n}$ ways to choose those.

Examples

– *Example 7:* Suppose a shipment of 100 lightbulbs contains 10 defective bulbs. You draw five at random, and test them. What is the probability that all five work?

$$\mathbb{P}[X = x] = \frac{\binom{90}{5}\binom{100-90}{5-5}}{\binom{100}{5}} = 0.5838$$

– *Example 8:* You are dealt five card from a standard deck. What is the probability of getting two or more clubs?

$$\begin{aligned}\mathbb{P}[X \geq 2] &= 1 - \mathbb{P}[X = 0] - \mathbb{P}[X = 1] \\ &= 1 - \frac{\binom{13}{0}\binom{39}{5}}{\binom{52}{5}} - \frac{\binom{13}{1}\binom{39}{4}}{\binom{52}{5}} \\ &= 1 - 0.2215 - 0.4115 = 0.3671.\end{aligned}$$

Negative Binomial Distribution

– Recall that for a binomial distribution, we are interested in the number of successes x in n trials where p is the probability of success on any trial. What if we are instead interested in the number of trials or equivalently, the number of failures needed to observe r successes?

– The negative binomial is just like it sounds; it describes an experiment that negates the binomial. The negative binomial distribution describes the number of failures needed to observe r successes, where each try has the same probability of success p and each try is independent.

$$\mathbb{P}(X = x) = \binom{r+x-1}{x} p^r (1-p)^x, \; x = 0, 1, 2, \ldots$$

– When $r = 1$, the distribution is called a geometric distribution, which would be the number of failures until the first success.

Examples

– *Example 9 (D. S. 5.5.1):* Suppose that a machine produces parts that can be either good or defective. Assume that the parts are good or defective independently of each other with probability 0.2 of being defective. An inspector observes the parts produced by this machine until she sees four defectives. Let X be the number of good parts observed by the time that the fourth defective is observed. How likely is she to observe six good parts?

– Note that this question is worded so that good part are failures and defective parts are successes.

$$\begin{aligned}\mathbb{P}(X=6) &= \binom{r+x-1}{x} p^r (1-p)^x \\ &= \binom{4+6-1}{6} 0.2^4 (1-0.2)^6 \\ &= 84 \times 0.0016 \times 0.2621 = 0.0352\end{aligned}$$

Recap

We discussed the following:

- What random variables are
- What distributions and pmfs are
- Some important distributions