

Week 3 Notes

D. ODay

10/21/2021

Reading Notes

Intro to Random Variables

Can think of a random variable as some sort of Machine that randomly spits out numbers. Unlike a function, which is very deterministic, the output of a random variable is uncertain. Usually referenced with large letters, such as X .

Properties of Random Variables

Distribution

The ‘type’ of random variable. A distribution describes the pattern that the random variable follows. Example, we have a recipe (distribution) that gives instructions to make a specific ‘meal’ (the random variable)

Each distribution has its own properties (expectation, variance, pmf, etc.)

Expectation

Another name for its average., denoted by $E(X)$

Variance

How much spread is inherent in a certain random variable, notated by $Var(X)$

PMFs and CDFs

Random Variables have probability mass functions (PMF) for discrete random variables and probability density functions (PDF) for continuous variables. PMF gives the probability that the random variable takes on a certain value. Denoted by $P(X = x)$.

cumulative density functions is a step further than PMF. Gives the probability $P(X \leq x)$. Notation for CDF is $F(x)$ where X is the random variable.

Support

The set of possible values that the random variable can take on.

Calculating Expectation and Variance

$$E(X) = \sum_i x_i P(X = x_i)$$

$$Var(X) = \sum_i (x_i - E(X))^2 P(X = x_i)$$

Standard deviation is the square root of the variance.

Binomial

The binomial distribution is a discrete distribution. In other words, the support is discrete. We also state that each distribution has a **story**.

Binomial Distribution Story: we perform n independent trials, each with only two outcomes (usually we think of these two outcomes as success or failure) and with a probability of success p that stays constant from trial to trial. The notation for this is:

$$X \sim \text{Bin}(n, p)$$

The \sim means ‘has the distribution of...’ **Example:** Flipping a coin a number of n times in the hopes of getting heads each time.

Parameters: The n is the number of trials and the p is the probability success of each trial

Characterstics/Properties:

$$E(X) = np, Var(X) = npq$$

q is $1-p$. Makes sense the variance grows as n grows: the more trials means more variance and is maximized at $p=1/2$. Again, makes sense since if one has a higher probability, the less uncertainty of a number appearing.

PMF:

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

The reason we have $\binom{n}{x}$ so we can count out all the permutations that represent the same combination. We don't care about the order we get the number of successes. Thus, $\binom{n}{x}$ gives us all of our desirable permutations. Thus, we multiply the probability of each desirable permutations by the total number of desirable permutations to get the overall probability that we find a desirable permutation. In other words, **the $p^x q^{n-x}$ part represents the probability and the $\binom{n}{x}$ is the number of desirable permutations.** The **support** of a binomial distribution is the integers 0 to n .

binomial in R

```
#Find P(X=3) where X ~ Bin(10, 0.5)
dbinom(3,10, 0.5)
```

```
## [1] 0.1171875
```

```
#Find P(X<=6) where X ~ Bin(15, 1/3)
dbinom(6,15, 1/3)
```

```
## [1] 0.1785892
```

```
#Find the value of x such that P(X <= x) = .9, where X ~ Bin(50, 1/5)  
qbinom(0.9,50, 1/5)
```

```
## [1] 14
```

```
#generate 5 random draws from X, where X ~ Bin(30, 1/4)  
rbinom(5, 30, 1/4)
```

```
## [1] 9 7 10 6 7
```

Random Variable Recap

A formal definition of a **random variable** is that it's a function that maps a sample space onto the real line.

Example

We have an experiment where we roll two die. The sample space would be the combinations of the face value of the rolls of the die: $(1,1), (1,2), \dots, (6,6)$. A random variable, or function, that maps this sample space to a real line may be the sum of the two rolls. This maps our sample space S to $\{2,3,4,\dots,12\}$. The 'randomness' comes from the fact that the outcomes are random

Bernoulli Distribution

Story

Basically, it is the Binomial distribution with only one trial with a probability of p success and $1 - p$ failure. Same as the binomial distribution where $n = 1$. Distribution is also written as

Distribution Label

$$X \sim \text{Bern}(p)$$

.

Expectation

The expectation of the distribution is simply p

Variance

The variance is $p(1 - p)$.

PMF

The PMF is simple because this distribution can only have two values: either 1 success or 0. The PMF is:

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

The Bernoulli distribution has a trait called the **Fundamental Bridge**. The idea is we can build a link between probability and expectation. For Bernoulli, it's the probability that the event in question occurs, or $P(X = 1) = E(X)$

Geometric Distributions

Story

Conduction repeated, independent trials. Can be either a success or failure. The trials can be either a success or failure. We want the number of failures before we achieve our first success (we don't count the success as a trial).

Distribution Label

$$X \sim \text{Geom}(p)$$

Expectation

$$\frac{1-p}{p} = \frac{1}{p} - 1$$

Think of it with this example. When we see something occurs 5% of the time we expect it to occur 1 out of 20 times. Thus, we would expect 19 failures until we get the success. Dividing by p just gives us the total # of trials including the success. Then we subtract that one success away.

Variance

$$\frac{1-p}{p^2}$$

PMF

$$P(X = x) = (1 - p)^x(p)$$

The Intuition behind this is pretty self explanatory. A failure has to be rolled x times until we get to the success with a probability of p .

First Success

Story

Same as Geometric except we count the first success

Distribution Label

$$Y \sim FS(p)$$

Expectation

$$1/p$$

Variance

$$\frac{1-p}{p^2}$$

PMF

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

Negative Binomial**Story**

Basically, an extension of the Geometric distribution. Basically, we count the number of failures before our r^{th} success, where each trial has a probability of p success.

Distribution Label

$$X \sim NBin(r, p)$$

Expectation

Same as Geometric but multiplied by r .

$$\frac{r(1-p)}{p}$$

Variance

Same as Geometric but multiply by r .

$$\frac{r(1-p)}{p^2}$$

PMF

$$P(X = x) = \binom{x+r-1}{r-1} p^r (1-p)^{x-r}$$

Regarding p and $1-p$ we are just multiplying by the number of of success we find for p and the number of trial ran until we get to the number of success for $1-p$. For the $\binom{x+r-1}{r-1}$, we know the last event will have to be a success since the experiment ends needs to end with a success. Also, we need to find a way to order all the $r-1$ other successes within the trial with the x failures

Poisson Distribution

Story

If we have many , with very small probabilities, chances at success then we can use the Poisson Distribution to model the number of occurrences of the event. An example would be lottery tickets.

Distribution Label

$$X \sim Pois(\lambda)$$

λ is the number of occurrences of the rare events/successes. Also labeled as the rate of occurrence. Example is if we say we expect 5 lottery winners out of all the tickets out there. $\lambda = 5$.

Expectation and Variance

$$\lambda$$

PMF

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

A binomial distribution converges to a poisson distribution when $n \rightarrow \infty$ and $p \rightarrow 0$.

HyperGeometric

Story

Counts the number of success in n draws from a population of b undesired objects and w desired objects without replacement. w is marked as a success and b as a failure. This is all done without replacement.

Distribution Label

$$X \sim HGeom(w, b, n)$$

Expectation

$$\frac{nw}{w + b}$$

Think of as if you want to draw white balls from a bag of balls. w is the white balls and b is all the other. You draw n times. Thus the chance of pulling a white ball from the bag is the number of white balls divided by the number of total balls.

Variance

not given

PMF

$$P(X = x) = \frac{\binom{w}{x} \binom{b}{n-x}}{\binom{w+b}{n}}$$

Expectation, Indicators and Memorylessness ## Expectation ### Linearity of Expectation

$$E(X + Y) = E(X) + E(Y), E(aX) = aE(X)$$

Also, we have another way of calculating variance:

$$\text{Var}(X) = E(X^2) - (E(X))^2$$

Indicators

Also, know as **indicator random variables**

I_A is an indicator random variable that takes on the value of 1 if event A occurs and 0 if event A doesn't occur. Therefore, A occurs with probability p .

Remember the **Fundamental Bridge** concept: the probability that an event occurs is also the expectation of the indicator random variable.

Memorylessness

A property of random variables. Example, say that we're waiting for food at a restaurant. the expected value is 10 minutes. It doesn't matter how long you have been waiting for food, you should expect to wait 10 minutes, going forward. The Geometric distribution is good in this regard.

Reading Notes

Example of PMf

We roll a six sided, let X be the sum of the two rolls

$$P(X = 2) = P(11) = \frac{1}{36} P(X = 3) = P(12, 21) = \frac{2}{36} \dots$$

Expectation of a Discrete Random Variable

$$E(X) = \sum_k kP(X = k)$$

The probability acts as the weight on value k

If c is a constant:

$$E(c) = c$$

If a and b are constants:

$$E(aX + b) = aE(X) + b$$

When working with data, this helps us calculate data after it has been transformed.

Variance

$$\sigma^2 = V(X) = E[(X - E(X))^2] = E(X^2) - (E(X))^2$$