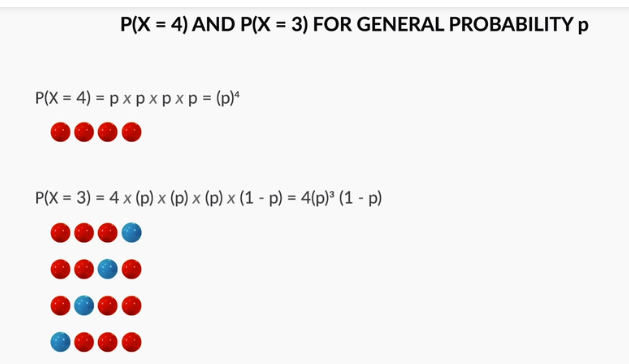
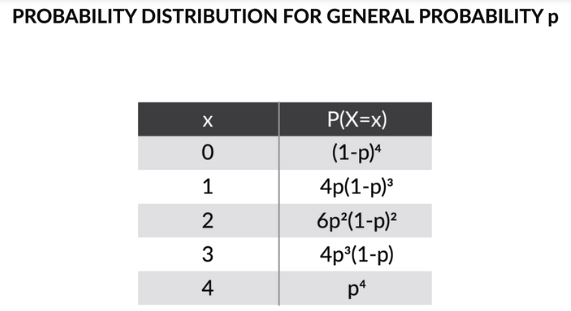
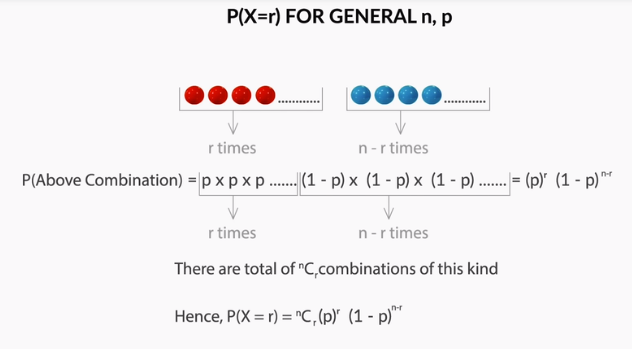
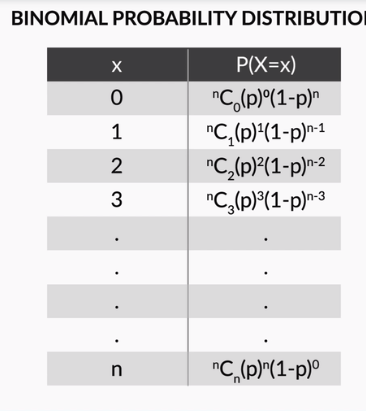
Binomial Distributions

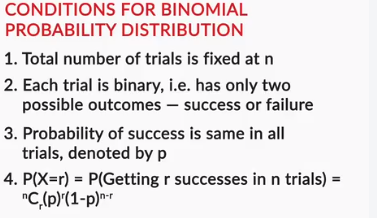
X is the random variable to get red ball

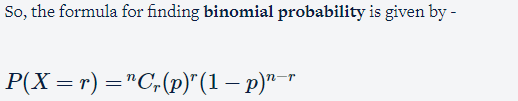












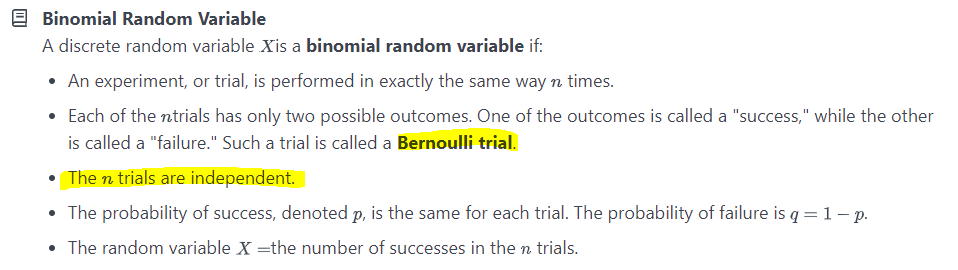
Where **n** is **no. of trials**, **p** is **probability of success** and **r** is **no. of successes after n trials**.

However, as Prof. Tricha said, there are some **conditions** that need to be followed in order for us to be able to apply the formula.

1. Total number of trials is fixed at n

2. Each trial is binary, i.e., has only two possible outcomes - success or failure

3. Probability of success is same in all trials, denoted by p



Binomial Distribution Example:

|  |  |
| --- | --- |
| **Binomial Distribution Applicable** | **Binomial Distribution Not Applicable** |
| Tossing a coin 20 times to see how many tails occur | Tossing a coin until a heads occurs |
| Asking 200 randomly selected people if they are older than 21 or not | Asking 200 randomly selected people how old they are |
| Drawing 4 red balls from a bag, putting each ball back after drawing it | Drawing 4 red balls from a bag, not putting each ball back after drawing it |

## Example 10-4



A Quality Control Inspector (QCI) investigates a lot containing 15 skeins of yarn. The QCI randomly samples (without replacement) 5 skeins of yarn from the lot. Let Xequal the number of skeins with acceptable color. Is X a binomial random variable?

#### Answer

No, X is not a binomial random variable, because p, the probability that a randomly selected skein has acceptable color changes from trial to trial. For example, suppose, unknown to the QCI, that 9 of the 15 skeins of yarn in the lot are acceptable. For the first trial, pequals 915. However, for the second trial, pequals either 914 or 814depending on whether an acceptable or unacceptable skein was selected in the first trial. Rather than being a binomial random variable, X is a hypergeometric random variable. If we continue to assume that 9 of the 15 skeins of yarn in the lot are acceptable, then X has the following probability mass function:

f(x)=P(X=x)=(9x)(65−x)(155) for x=0,1,…,5

## Example 10-5



A Gallup Poll of n=1000 random adult Americans is conducted. LetXequal the number in the sample who own a sport utility vehicle (SUV). Is X a binomial random variable?

#### Answer

No, X is technically a hypergeometric random variable, not a binomial random variable, because, just as in the previous example, sampling takes place without replacement. Therefore, p, the probability of selecting an SUV owner, has the potential to change from trial to trial. To make this point concrete, suppose that Americans own a total of N=270,000,000 cars. Suppose too that half (135,000,000) of the cars are SUVs, while the other half (135,000,000) are not. Then, on the first trial, pequals 12 (from 135,000,000 divided by 270,000,000). Suppose an SUV owner was selected on the first trial. Then, on the second trial, p equals 134,999,999 divided by 269,999,999, which equals.... punching into a calculator... 0.499999... Hmmmmm! Isn't that 0.499999... close enough to 12 to just call it 12?Yes...that's what we do!

In general, when the sample size nis small in relation to the population size N, we assume a random variable X, whose value is determined by sampling without replacement, follows (approximately) a binomial distribution. On the other hand, if the sample size nis close to the population size N, then we assume the random variable X follows a hypergeometric distribution.

## Example 10-6

By some estimates, twenty-percent (20%) of Americans have no health insurance. Randomly sample n=15 Americans. Let X denote the number in the sample with no health insurance. What is the probability that exactly 3 of the 15 sampled have no health insurance?

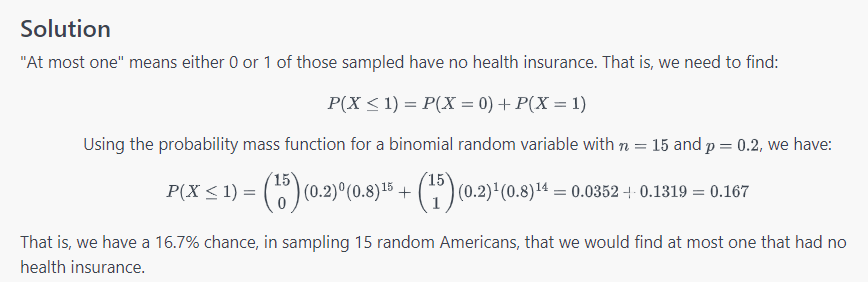
#### Solution

Since n=15 is small relative to the population of N = 300,000,000 Americans, and all of the other criteria pass muster (two possible outcomes, independent trials, ....), the random variable X can be assumed to follow a binomial distribution with n=15 and p=0.2. Using the probability mass function for a binomial random variable, the calculation is then relatively straightforward:

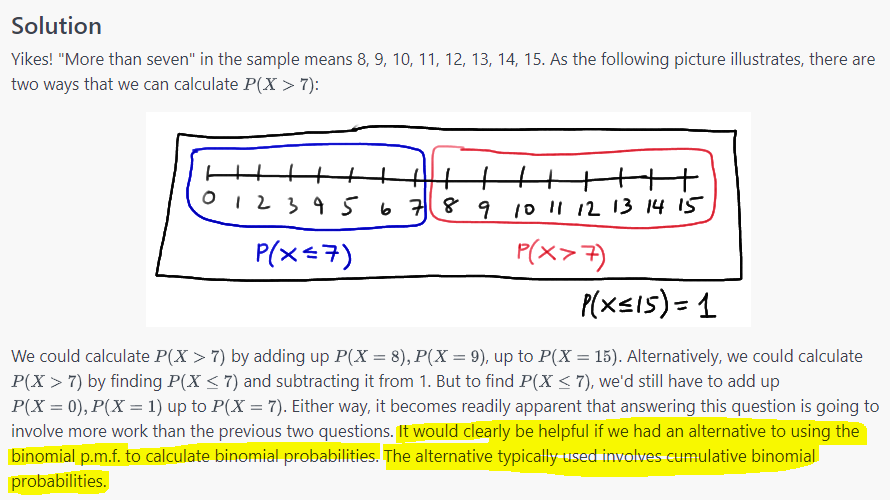
P(X=3)=(153)(0.20)3(0.80)12=0.25

That is, there is a 25% chance, in sampling 15 random Americans, that we would find exactly 3 that had no health insurance.

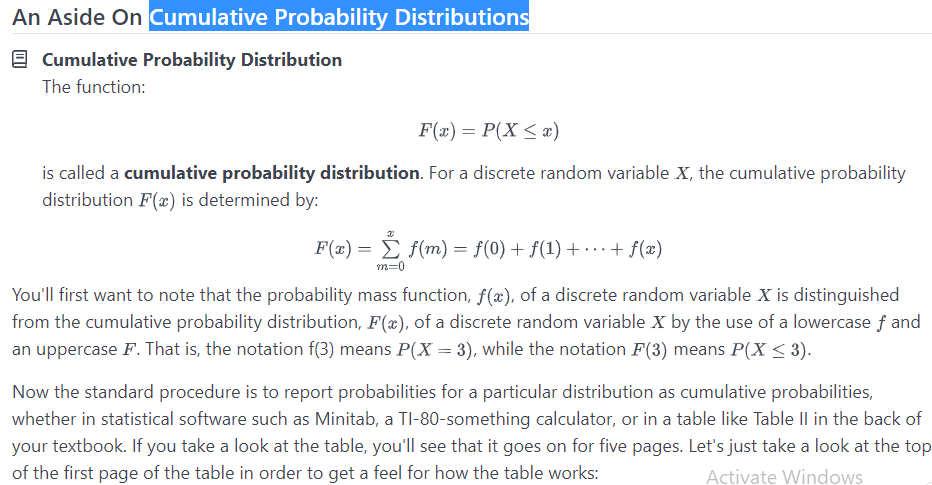
Question: What is the probability that at most one of those sampled has no health insurance?

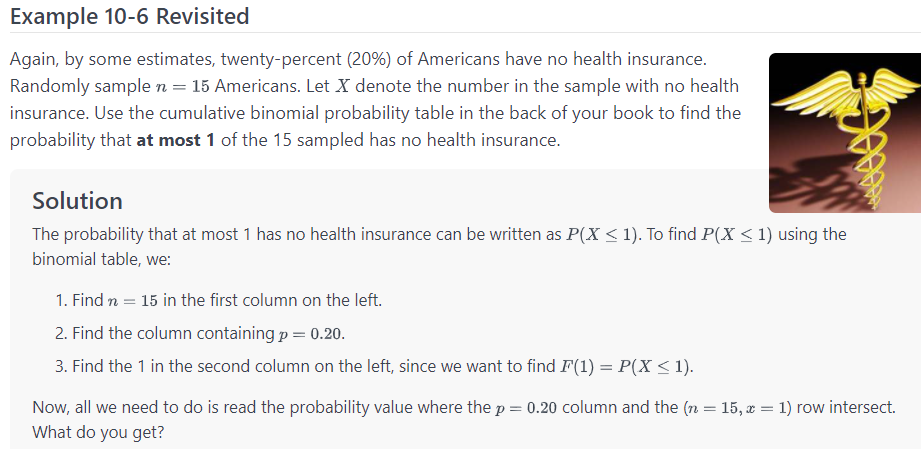


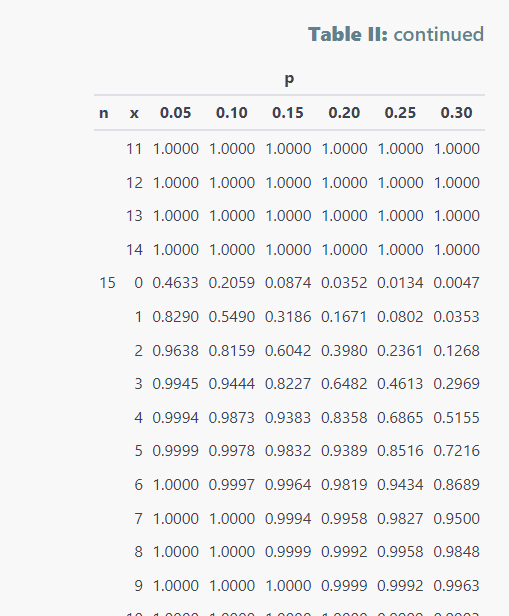
Question: What is the probability that more than seven have no health insurance?

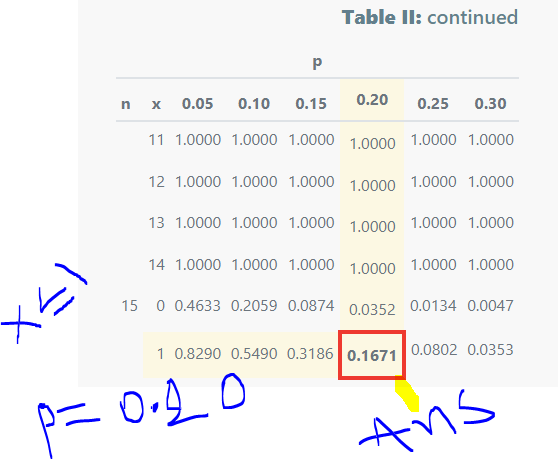


## Cumulative Probability Distributions

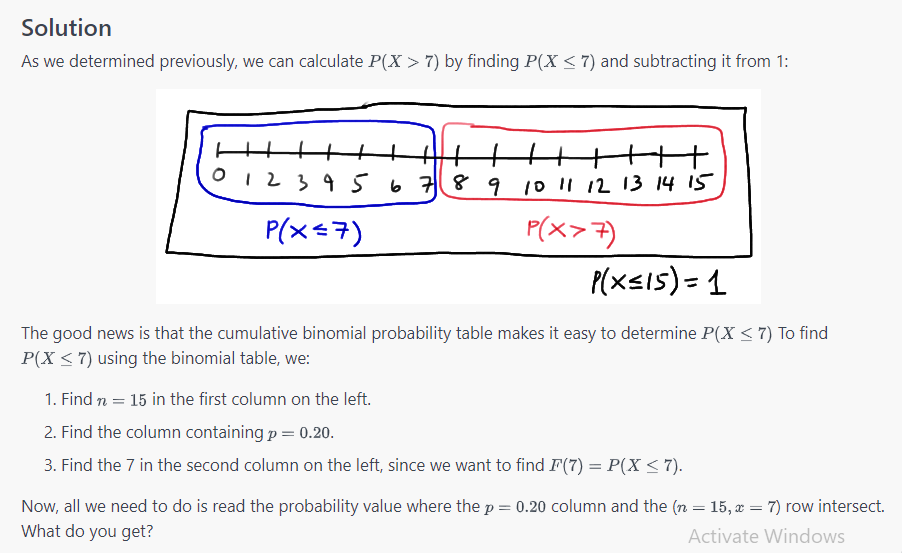


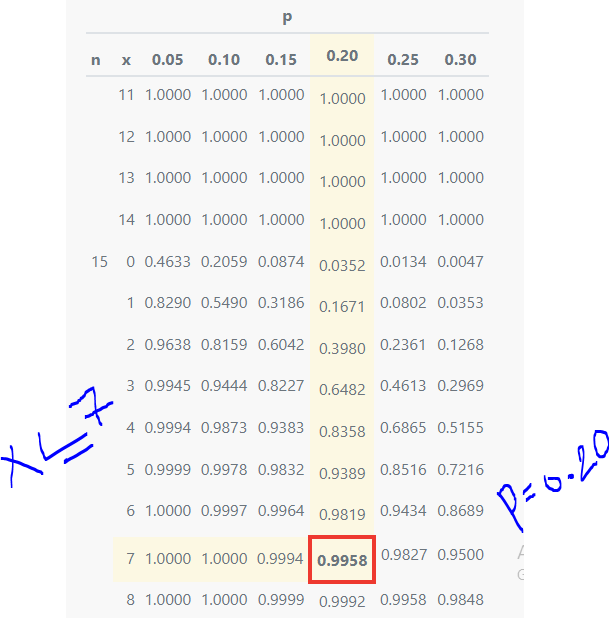


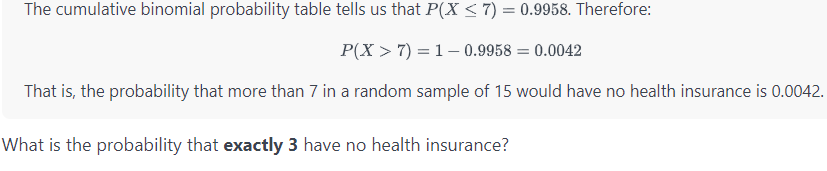


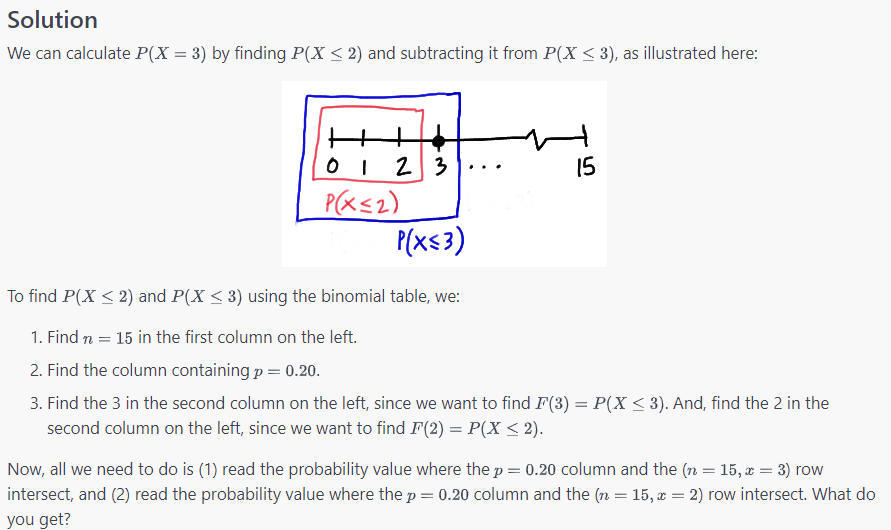


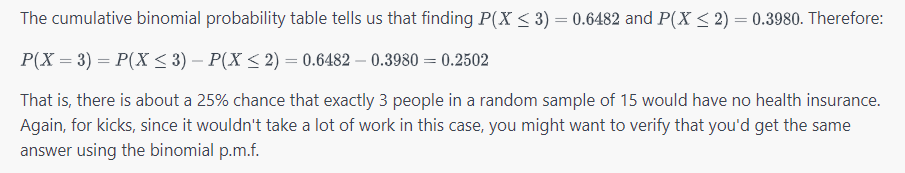
Question: What is the probability that **more than 7** have no health insurance?









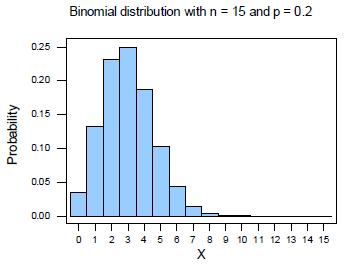




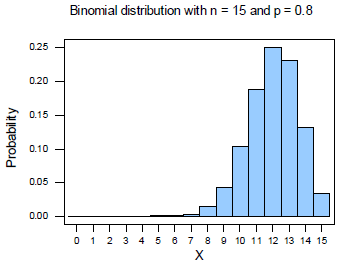
the shape of the binomial distribution is directly related, and not surprisingly, to two things:

1. **n**, the number of independent trials
2. **p**, the probability of success

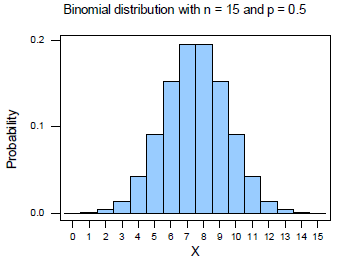
For **small p** and **small n**, the binomial distribution is what we call **skewed right**. That is, the bulk of the probability falls in the smaller numbers 0,1,2,…, and the distribution tails off to the right. For example, here's a picture of the binomial distribution when n=15 and p=0.2:



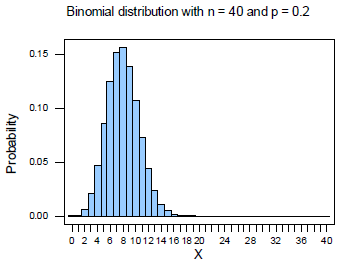
For **large p** and **small n**, the binomial distribution is what we call **skewed left**. That is, the bulk of the probability falls in the larger numbers n,n−1,n−2,… and the distribution tails off to the left. For example, here's a picture of the binomial distribution when n=15 and p=0.8:



For p=0.5 and **large and small n**, the binomial distribution is what we call **symmetric**. That is, the distribution is without skewness. For example, here's a picture of the binomial distribution when n=15 and p=0.5:

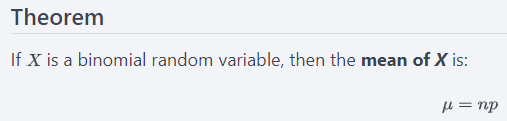


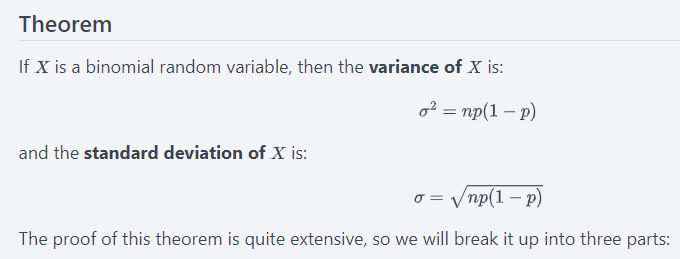
For **small p** and **large n**, the binomial distribution **approaches symmetry**. For example, if p=0.2 and n is small, we'd expect the binomial distribution to be skewed to the right. For large n, however, the distribution is nearly symmetric. For example, here's a picture of the binomial distribution when n=40 and p=0.2:

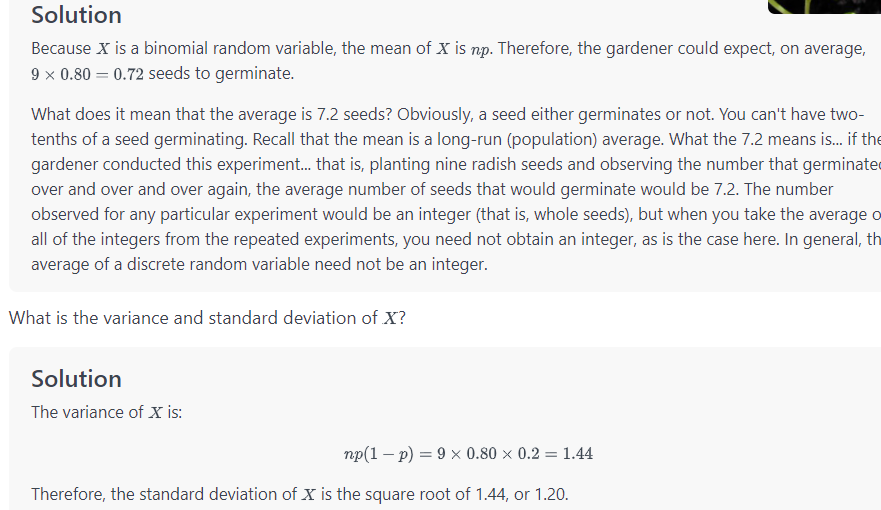


You might find it educational to play around yourself with various values of the n and p parameters to see their effect on the shape of the binomial distribution.

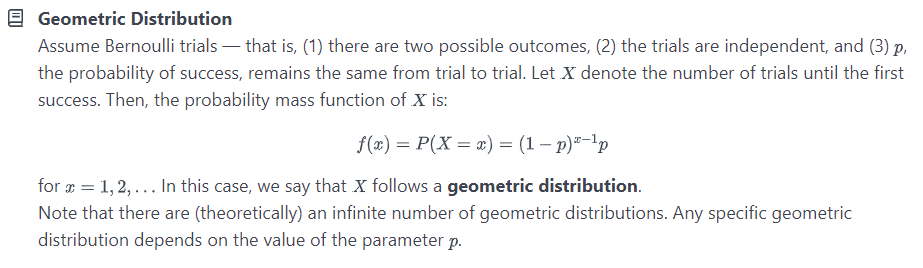






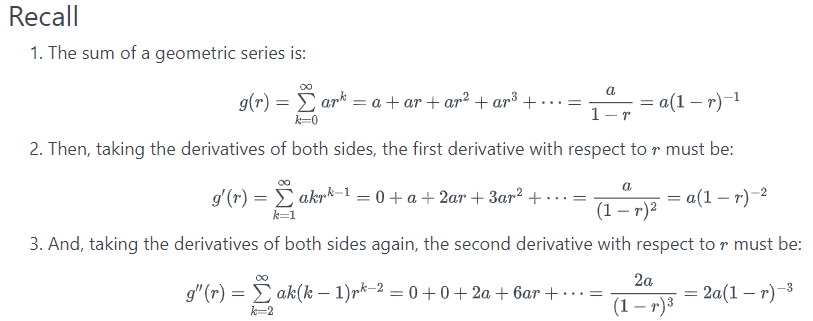


Geometric Distribution:

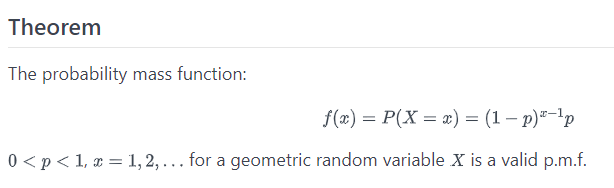


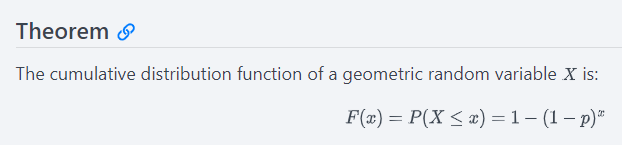
# 11.2 - Key Properties of a Geometric Random Variable

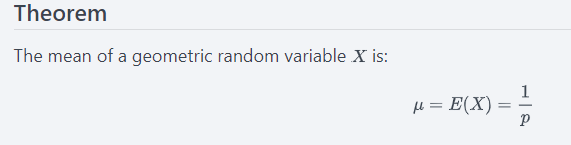
On this page, we state and then prove four properties of a geometric random variable. In order to prove the properties, we need to recall the sum of the geometric series. So, we may as well get that out of the way first.

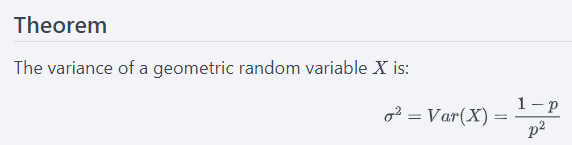


We'll use the sum of the geometric series, first point, in proving the first two of the following four properties. And, we'll use the first derivative, second point, in proving the third property, and the second derivative, third point, in proving the fourth property. Let's jump right in now!





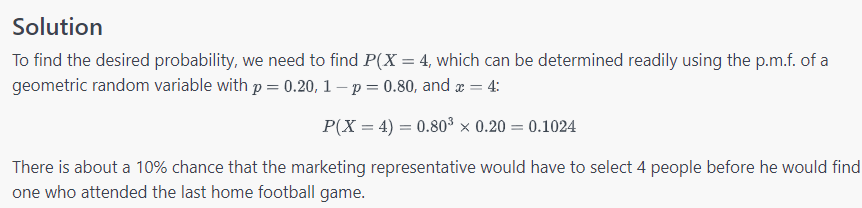




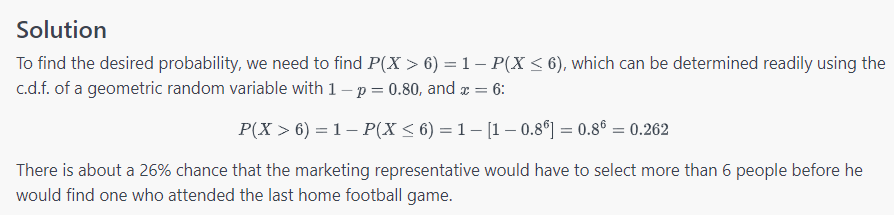
**Example 11-1 Continued**



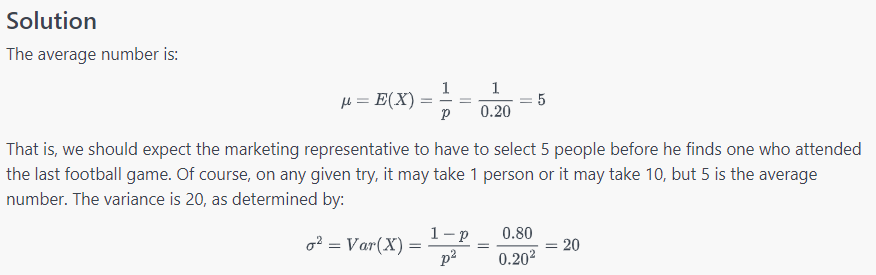
A representative from the National Football League's Marketing Division randomly selects people on a random street in Kansas City, Kansas until he finds a person who attended the last home football game. Let p, the probability that he succeeds in finding such a person, equal 0.20. And, let X denote the number of people he selects until he finds his first success. What is the probability that the marketing representative must select 4 people before he finds one who attended the last home football game?



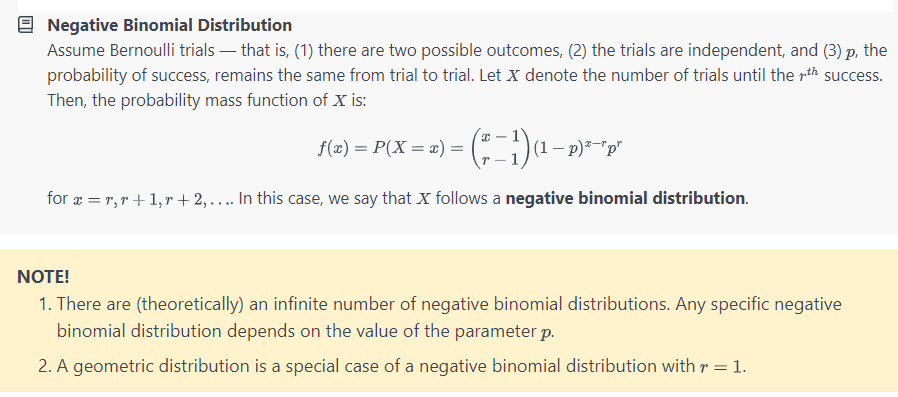
What is the probability that the marketing representative must select more than 6 people before he finds one who attended the last home football game?

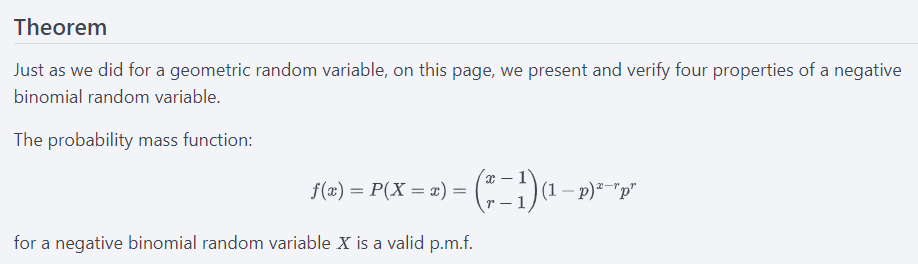


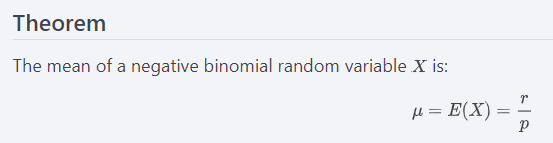
How many people should we expect (that is, what is the average number) the marketing representative needs to select before he finds one who attended the last home football game? And, while we're at it, what is the variance?

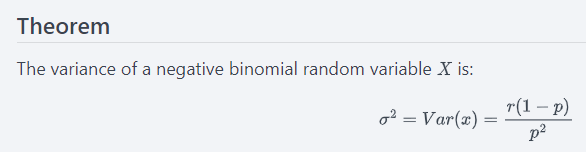


Negative Binomial Distribution:









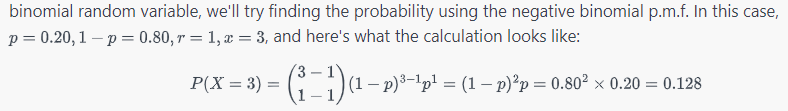
## Example 11-2



An oil company conducts a geological study that indicates that an exploratory oil well should have a 20% chance of striking oil. What is the probability that the first strike comes on the third well drilled?

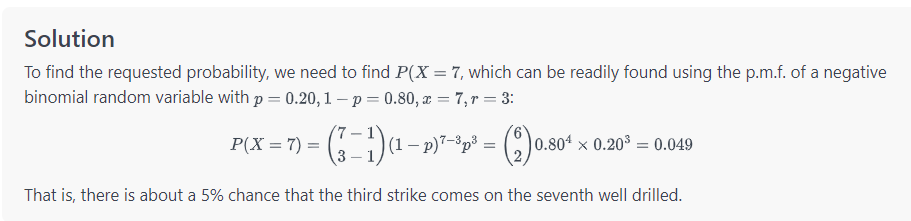
#### Solution

To find the requested probability, we need to find P(X=3. Note that Xis technically a geometric random variable, since we are only looking for one success. Since a geometric random variable is just a special case of a negative binomial random variable, we'll try finding the probability using the negative binomial p.m.f. In this case, p=0.20,1−p=0.80,r=1,x=3, and here's what the calculation looks like:

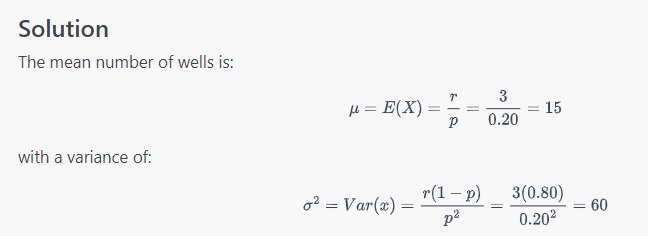


It is at the second equal sign that you can see how the general negative binomial problem reduces to a geometric random variable problem. In any case, there is about a 13% chance thathe first strike comes on the third well drilled.

Question: What is the probability that the third strike comes on the seventh well drilled?



Question: What is the mean and variance of the number of wells that must be drilled if the oil company wants to set up three producing wells?



REF: https://online.stat.psu.edu/stat414/lesson/10/10.1