Notes: Bernoulli, Binomial, and Geometric Distributions

CS 3130/ECE 3530: Probability and Statistics for Engineers

September 19, 2017

**Bernoulli distribution:**
Defined by the following pmf:

\[ p_X(1) = p, \quad \text{and} \quad p_X(0) = 1 - p \]

Don’t let the \( p \) confuse you, it is a single number between 0 and 1, not a probability function. If \( X \) is a random variable with this pmf, we say “\( X \) is a Bernoulli random variable with parameter \( p \)”, or we use the notation \( X \sim Ber(p) \). You can think of a Bernoulli trial as flipping a coin where the chance of heads is \( p \) and the chance of tails is \( 1 - p \). Often we call 0 a “failure” and 1 a “success”, so \( p \) is the probability of success.

**Binomial distribution:**
The binomial distribution describes the probabilities for repeated Bernoulli trials – such as flipping a coin ten times in a row. Each trial is assumed to be independent of the others (for example, flipping a coin once does not affect any of the outcomes for future flips). First, we need some definitions.

Remember the definition for factorial:

\[ n! = n \times (n - 1) \times \cdots \times 2 \times 1 \]

This is the number of ways to put \( n \) objects into a specific order.

And the definition for “\( n \) choose \( k \)”:

\[ \binom{n}{k} = \frac{n!}{(n - k)! \, k!} \]

This is the number of ways to select \( k \) objects out of a possible \( n \), where the order does not matter.

The **binomial distribution** with parameters \( n \) and \( p \) is given by the pmf:

\[ p_X(k) = \binom{n}{k} p^k (1 - p)^{n-k}. \]

This is denoted \( X \sim Bin(n, p) \). This distribution is for repeated Bernoulli trials, and it gives the probability that you get \( k \) successes out of \( n \) trials.

**Geometric distribution:**
The **geometric distribution** is also for repeated Bernoulli trials, and it gives the probability that the first \( k - 1 \) trials are failures, while the \( k \)th trial is the first success. Its pmf is

\[ p_X(k) = (1 - p)^{k-1} p. \]

This is denoted \( X \sim Geo(p) \).
In-class Problem: Remember the Monty Hall problem – if we switch doors, we have a $2/3$ chance of winning and $1/3$ chance to lose. If we play the game 4 times, what is the probability that we win exactly once? How about exactly 0, 2, 3, or 4 times? What is the chance that we lose the first three times and finally win on the 4th try?

### Key to variable names

It's important to keep straight what all the variables mean in the above equations. Here is a summary:

- $n$: Number of trials
- $k$: Number of successes in Binomial, OR first success that occurs in Geometric
- $p$: Probability of success