STAT 230 SECTION 2 LECTURE 13

Today's Agenda

Last time:

- Hypergeometric Distribution
- Bernoulli Distribution

Today (Lec 13, 06/03):

- Binomial Distribution
- Negative Binomial Distribution

Definition

A **Bernoulli trial** with probability of success p is an experiment that results in either a success or failure, and the probability of success is p.

In other words, a Bernoulli trial models "either ... or". Moreover, we can build a distribution from it!

Definition

If a random variable X represents the number of successes in a Bernoulli trial with probability of success p, it follows the **Bernoulli distribution**, and we denote it as

$$X \sim Bernoulli(p),$$

where $X \in \{0, 1\}$.

The probability function of X is

$$f(x) = p^x(1-p)^{1-x}, \quad x \in \{0, 1\}.$$

...Alright, Bernoulli is cool, but one trial doesn't sound useful enough for anything.

Binomial Distribution

Definition

- Assume an experiment has two possible outcomes, "S" (success) and "F" (failure).
- Assume that in every try, $P(S) = p$ and $P(\mathit{failure}) = 1 - p$ for $0 < p < 1$.
- Repeat the experiment n times independently and record the number of successes obtained. Then X has a **binomial distribution**, denoted by $X \sim Bin(n, p)$.

Using previous techniques, we find

$$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

for $x \in \{0, 1, 2, \ldots, n\}$.

Exercise: Use the binomial theorem to show that this probability function sums up to 1.

Examples of binomial random variables:

a) A Bernoulli(p) random variable is a Bin($1, p$) random variable.

b) Flip a fair coin independently 20 times, and let X denote the number of heads observed. Then

$$X \sim Bin(20, 0.5)$$

c) Consider drawing numbers from 1-9 **with replacement** to form a 5 number sequence, and let X denote the number of odd digits in the sequence. Then

$$X \sim Bin(5, 5/9)$$

Example

Suppose a tack when flipped has probability 0.6 of landing point up. If the tack is flipped 10 times, what is the probability it lands point up more than twice?

Example

Suppose that in a weekly lottery you have probability 0.02 of winning a prize with a single ticket. If you buy 1 ticket per week for 52 weeks, what is the probability that

a) you win no prizes?

b) you win 2 or more prizes?

Binomial vs Hypergeometric

- Binomial and hypergeometric distributions are fundamentally different!
- In Binomial models, we pick with replacement, in the hypergeometric model without replacement.
- If N is large and n is small, the chance we pick the same object twice is small.
- Thus, letting $r/N = p$, $X \sim Hyp(N, r, n)$ and $Y \sim Bin(n, p)$, then we can approximate

$$P(X \leq k) \approx P(Y \leq k).$$

- See pages 86/87 for more.

Negative Binomial Distribution

Definition

- Assume an experiment has two possible outcomes, "S" (success) and "F" (failure).
- Assume that in every try, $P(S) = p$ and $P(failure) = 1 - p$ for $0 < p < 1$.
- Repeat the experiment independently until a specified number k of successes have been observed.
- The number of failures before the kth success follows a Negative Binomial distribution, $X \sim NegBin(k, p)$.

The probability function of $X \sim NegBin(k, p)$ is

$$f(x) = P(X = x) = \binom{x+k-1}{x} p^k (1-p)^x, \quad x = 0, 1, 2, \ldots$$

Binomial versus Negative Binomial

Binomial distribution: We know number of trials n, but we do not know how many successes.

Negative Binomial distribution: We know the number of successes k, but we do not know how many trials will be needed.

Question

Suppose names are drawn with replacement out of a hat containing 100 names, of which 12 are the name "Kyle", until the name "Kyle" is drawn three times. Let X denote the number of draws required. Then:

A $X \sim Binomial(12, 0.12)$

B $X \sim hyp(100, 12, 7)$

C $X \sim NB(3, 0.12)$

D $X - 3 \sim NB(3, 0.12)$

Question

Suppose 17 names are drawn with replacement out of a hat containing 100 names, of which 12 are the name "Kyle". Let X denote the number of names drawn that are not "Kyle". Then:

A $X \sim Binomial(12, 0.17)$

B $X \sim hyp(100, 88, 17)$

C $X \sim NB(3, 0.12)$

D $X - 3 \sim NB(3, 0.12)$

E $X \sim Binomial(17, 0.88)$

Geometric Distribution

Suppose you flip a coin until HEADS shows up. What is the probability that you need 4 tails before the first head?

$$(1-p)^4 p$$

What is the probability you need k tails until the first head?

$$(1-p)^k p$$

Same would have been obtained with a NegBin distribution!

Geometric Distribution

Definition

- Assume an experiment has two possible outcomes, "S" (success) and "F" (failure).
- Assume that in every try, $P(S) = p$ and $P(failure) = 1 - p$ for $0 < p < 1$.
- Repeat the experiment independently until the first success has been observed.
- The number of failures before the 1st success follows a Geometric Distribution, $X \sim Geo(p)$.

The geometric distribution has pf

$$f(x) = p(1-p)^x, \quad x = 0, 1, 2, \ldots.$$

Note that the geometric distribution is a special case of the negative binomial distribution.

Example

Suppose a fair coin is to be flipped until it comes up heads. What is the probability it will take more than 3 flips?