Lecture 12, Feb 6, 2023

Hypergeometric Distribution

Definition

Given N total objects, where K of the N are successes, and sampling n times without replacement, the *hypergeometric distribution* describes the probability of x successes and $n - x$ failures:

$$h(x; N, n, K) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

- Statistics:
  - $\mu = \dfrac{nK}{N}$
  - $\sigma^2 = \dfrac{Kn(N-n)}{N(N-1)}\left(1 - \dfrac{K}{N}\right)$

Negative Binomial Distribution and Geometric Distribution

Definition

Given repeated trials with probability p of success and $1-p$ of failure, the *negative binomial distribution* describes the probability of observing the k-th success on trial number x:

$$b^*(x, k, p) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$$

- i.e. x is how many trials it takes to get k successes
- Intuition:
  - If we are on trial $x<$ then the chance of $k-1$ successes in the last $x-1$ trials is $b(k-1; x-1, p) = \binom{x-1}{k-1} p^{k-1}(1-p)^{x-1-k+1}$
  - Getting the next success is just multiplication by p: $b^*(x; k, p) = pb(k-1; x-1, p) = p\binom{x-1}{k-1} p^{k-1}(1-p)^{x-1-k+1} = \binom{x-1}{k-1} p^k (1-p)^{x-k}$
- The *geometric distribution* is just the negative binomial distribution with $k = 1$

Definition

Given repeated trials with probability p of success, the *geometric distribution* describes the probability of the first success occurring on trial x:

$$g(x; p) = b^*(x; 1, p) = p(1-p)^{x-1}$$

- Statistics of the geometric distribution:
  - $\mu = \dfrac{1}{p}$
  - $\sigma^2 = \dfrac{(1-p)}{p^2}$