

## Lecture 5: Poisson, Hypergeometric, and Geometric Distributions

Sta 111

Colin Rundel

May 20, 2014

## Binomial Approximations

Last time we looked at the normal approximation for the binomial distribution:

- Works well when  $n$  is large
- Continuity correction helps
- Binomial can be skewed but Normal is symmetric
- At a minimum we want  $np \geq 10$  and  $nq \geq 10$

What do we do when  $p$  is close to 0 or 1?

## Alternative Approximation

Let  $X \sim \text{Binom}(n, p)$  which we will reparameterize so that  $p = \lambda/n$  for a fixed value of  $\lambda$ . As such,  $\lambda/n$  is small when  $n$  is large.

We will evaluate the Binomial distribution as  $n \rightarrow \infty$ .

## Alternative Approximation, Cont.

$$A_n = \frac{n!}{n^k(n-k)!}$$

## Alternative Approximation, Cont.

$$B_n = \left(1 - \frac{\lambda}{n}\right)^n$$

## Alternative Approximation, Cont.

$$C_n = \left(1 - \frac{\lambda}{n}\right)^{-k}$$

## Alternative Approximation, cont.

Let  $X \sim \text{Binom}(n, p)$  we will reparameterize such that  $p = \lambda/n$  for a fixed value of  $\lambda$ . As such,  $\lambda/n$  is small when  $n$  is large.

$$P(X = k | n, p = \lambda/n) = \underbrace{\left(\frac{n!}{n^k(n-k)!}\right)}_{A_n} \underbrace{\left(\frac{\lambda^k}{k!}\right)}_{B_n} \underbrace{\left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}}_{C_n}$$

$$\lim_{n \rightarrow \infty} P(X = k | n, p = \lambda/n) = \frac{\lambda^k}{k!} e^{-\lambda}$$

Therefore for large  $n$ ,

$$P(X = k | n, p = \lambda/n) \approx \frac{\lambda^k}{k!} e^{-\lambda}$$

## Poisson Distribution

Let  $X$  be a random variable reflecting the number of events in a given period where the expected number of events in that interval is  $\lambda$  then the probability of  $k$  occurrences ( $k \geq 0$ ) in the interval is given by the Poisson distribution,  $X \sim \text{Pois}(\lambda)$

$$P(X = k | \lambda) = f(k | \lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$$

We use this approximation to the Binomial when  $p$  is very small and  $n$  is very large since  $\lambda = np$  tends to be reasonable.

## Poisson Distribution - Mode

We can use the same approach that we used with the Binomial distribution  
Therefore  $k_{\text{mode}}$  is the smallest integer greater than  $\lambda - 1$

$$k_{\text{mode}} = \begin{cases} \lambda - 1, \lambda & \text{if } \lambda = \lceil \lambda \rceil \\ \lceil \lambda \rceil - 1 & \text{otherwise} \end{cases}$$

## Poisson Distribution - Example

Assume you have a sample of a stable isotope of an element, there are approximately  $10^{20}$  atoms in this sample. If on average one of these atoms will radioactively decay every  $10^{12}$  years ( $\approx 5 \times 10^{19}$  secs).

What is the probability that 4 or fewer atoms decay in the next second?

## Approximation - Mean &amp; Variance

We defined  $p = \lambda/n$  and we know that for a Binomial random variable

$$\begin{aligned} \mu &= np \\ \sigma^2 &= npq \end{aligned}$$

Then for large  $n$  we then get,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mu &= \lim_{n \rightarrow \infty} n \frac{\lambda}{n} = \lambda \\ \lim_{n \rightarrow \infty} \sigma^2 &= \lim_{n \rightarrow \infty} n \frac{\lambda}{n} \left(1 - \frac{\lambda}{n}\right) = \lambda \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right) = \lambda \end{aligned}$$

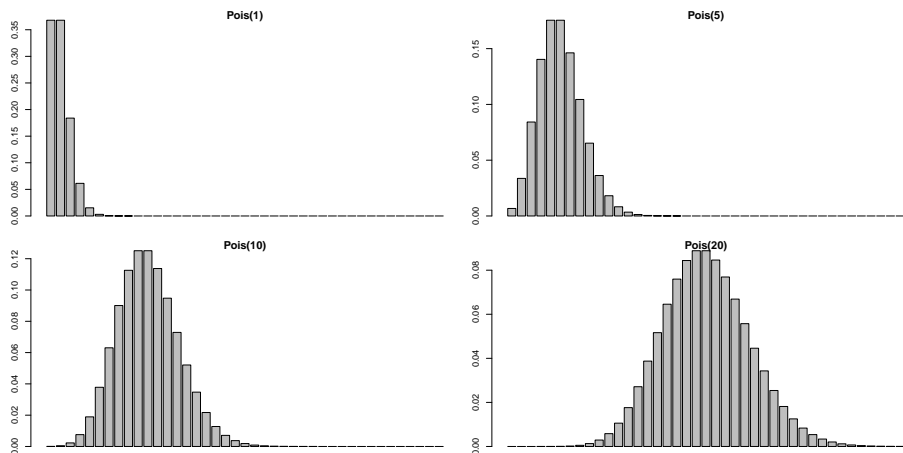
## Poisson and Normal Distributions

Based on the connection between the Binomial and Poisson distributions it intuitively makes sense that we should also be able to approximate the Poisson with a Normal distribution.

For approximation to the binomial we need  $np \geq 10$  and  $nq \geq 10$ .

What is a reasonable requirement for  $\lambda$ ?

# Poisson and Normal Distributions, cont.



# Another way to look at the Binomial

Imagine we have a population that is partitioned into 'good' and 'bad' subsets. Let  $G$  be the number of good elements in the population,  $B$  the number of bad elements, and  $N = B + G$ .

If we sample this population *with replacement* what is the probability that we observe  $g$  good samples and  $b$  bad samples.

This is still the Binomial distribution, but rewritten such that

$$\begin{aligned}
 P(g \text{ good, } b \text{ bad in } n = b + g \text{ tries}) &= \binom{n}{g} \frac{G^g B^b}{N^n} \\
 &= \binom{n}{g} \left(\frac{G}{N}\right)^g \left(\frac{B}{N}\right)^b = \binom{n}{g} \left(\frac{G}{N}\right)^g \left(1 - \frac{G}{N}\right)^{n-g}
 \end{aligned}$$

# Hypergeometric

What would change if we were *sampling without replacement*?

# Hypergeometric Distribution

Let  $X$  be a random variable reflecting the number of successes in  $n$  draws without replacement from a finite population of size  $N$  with  $m$  desired items then the probability of  $k$  successes is given by the Hypergeometric distribution,  $X \sim \text{Hypergeo}(N, m, n)$

$$P(X = k) = f(k|N, m, n) = \frac{\binom{m}{k} \binom{N-m}{n-k}}{\binom{N}{n}}$$

## Hypergeometric Distribution - Example

You are dealt five cards, what is the probability that four of them are aces?

If we use the Hypergeometric distribution then,  $N = 52$ ,  $m = 4$ ,  $n = 5$  and

## Hypergeometric Distribution - Another Way

Let  $X \sim \text{Binom}(m, p)$  and  $Y \sim \text{Binom}(N - m, p)$  be independent Binomial random variables then we can define the Hypergeometric distribution as the conditional probability of  $X = k$  given  $X + Y = n$ .

Note that  $X + Y \sim \text{Binom}(N, p)$

## Geometric Distribution - Version 1

Let  $Y$  be a random variable reflecting the number of failures of independent Bernoulli trials, with probability of success  $p$ , needed before observing the first success. Then the probability of  $k$  failures before the first success is given by the Geometric distribution,  $Y \sim \text{Geo}(p)$

$$P(Y = k) = f(k|p) = p(1 - p)^k$$

## Geometric Distribution - Version 2

Let  $Y$  be a random variable reflecting the number of independent Bernoulli trials, with probability of success  $p$ , needed before observing the first success. Then the probability of  $k$  trials needed to achieve the first success is given by the Geometric distribution,  $Y \sim \text{Geo}(p)$

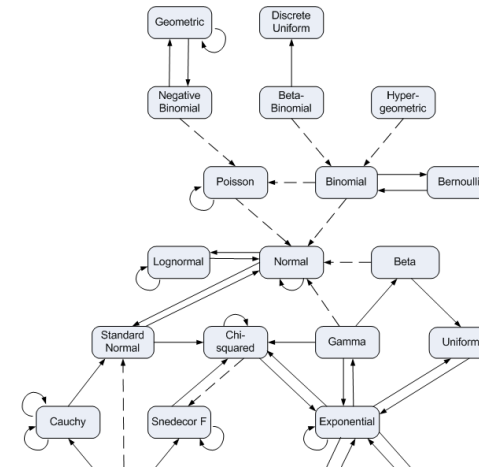
$$P(Y = k) = f(k|p) = p(1 - p)^{k-1}$$

## Negative Binomial Distribution

Let  $X$  be a random variable reflecting the total number of successes before the  $r^{\text{th}}$  failure where each trial is an independent Bernoulli trial with  $p$  probability of success. Then the probability of  $k$  successes is given by the Negative Binomial distribution,  $X \sim \text{NB}(r, p)$

$$P(X = k) = f(k|r, p) = \binom{k+r-1}{k} p^k (1-p)^r$$

## Distribution Relationships



[http://www.johndcook.com/distribution\\_chart.html](http://www.johndcook.com/distribution_chart.html)