

Normal approximation to binomial distribution

Alex Tse
Department of Statistics
The University of Warwick

February 24, 2014

1 Overview

This is a short supplementary note to the ST111/112 tutorial class of week 7. When working with a random variable with binomial distribution $X \sim \text{Bin}(n, p)$, an important technique is to apply normal approximation to X to compute probability of events like $\mathbb{P}(X \leq k)$ or $\mathbb{P}(X \geq k)$. I would like to give more colours regarding the essential characteristics of the underlying distribution in order to give a “good” approximation.

2 Is large n sufficient to give a good normal approximation?

A common misconception with normal approximation is that a large n alone is sufficient to give a good fit (answer like “ $n > 30$ ” is quite common among students’ work). You will probably think $n = 100$ is “large”, but does it guarantee a good fit? Consider examples of $\text{Bin}(100, 0.01)$ and $\text{Bin}(100, 0.99)$ where the associated probability mass plots are given in figure 1. You can see that the probability mass are highly skewed to either side and this asymmetry casts doubt over the accuracy of normal approximation.

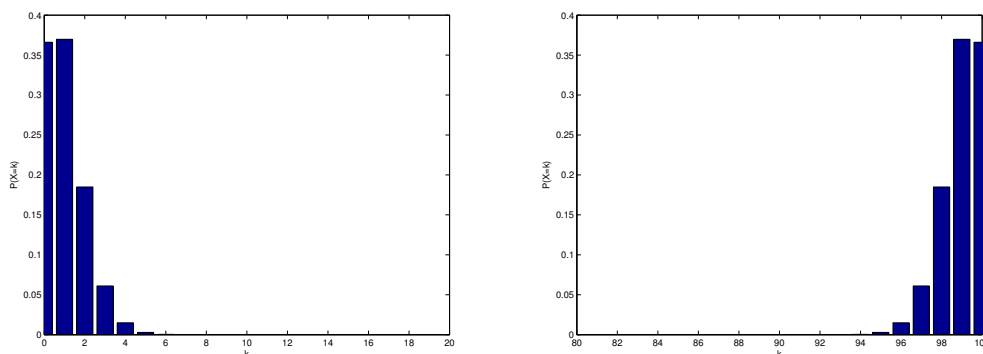


Figure 1: Probability mass plots of $\text{Bin}(100, 0.01)$ (left) and $\text{Bin}(100, 0.99)$ (right).

3 What have failed?

The most distinguished feature of a normal distribution is that it has a symmetric shape along the entire real number axis. However, binomial and a handful of other common discrete-outcome distributions are typically asymmetric on a limited range of outcomes. For example, the peak of a binomial distribution occurs at $\lfloor (n + 1)p \rfloor$ and thus it could be highly skewed to either end if $(n + 1)p$ is too small or too big, which are exactly the scenarios shown by figure 1. To

summarize, on one hand you want an n which is reasonably large so that the normality scaling kicks in. On the other hand, you also want the peak of the distribution not staying too close to the boundary values of the outcomes to ensure symmetry.

Deciding whether normal approximation is suitable for a certain distribution is more an art rather than science. But motivated by the above examples, a good gauge appears to be the skewness of the distribution. In the case of a binomial random variable $X \sim \text{Bin}(n, p)$, the skewness measure is given (can you prove this?) by

$$\mathbb{E} \left(\frac{X - \mu}{\sigma} \right)^3 = \frac{1 - 2p}{\sqrt{np(1-p)}}$$

where $\mu = np$ and $\sigma = \sqrt{np(1-p)}$. This measure should be small in magnitude to yield a good normal approximation. But as you can see from the above expression, it will tend to positive or negative infinity when p approaches 0 or 1.

4 Connection to Poisson distribution

As you would recall from lectures, a Poisson distribution with parameter λ is actually a limiting result of $X \sim \text{Bin}(n, p)$ when one fixes $np = \lambda$ and lets $n \rightarrow \infty$. If λ is small, it can be thought as a binomial distribution with large n but small np which is the case of the left panel in figure 1, and you would expect normal approximation to give a poor fit. In the tutorial class of week 7 we have sketched the probability mass function of Poisson distribution with $\lambda < 1$ which is strictly decreasing and in no way resembles a normal distribution. In short, you need λ to be large if you want to approximate it by a normal distribution. See figure 2 for some illustrations.

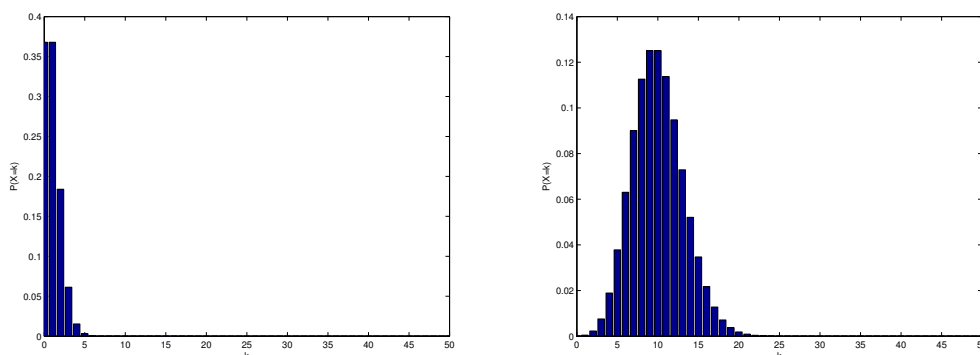


Figure 2: Probability mass plots of $\text{Poisson}(\lambda = 1)$ (left) and $\text{Poisson}(\lambda = 10)$ (right).