

5.2: **The Normal Distribution as a Limit of Binomial Distributions

The results of the derivation given here may be used to understand the origin of the Normal Distribution as a limit of Binomial Distributions^[1]. A mathematical “trick” using logarithmic differentiation will be used.

First, recall the definition of the Binomial Distribution^[2] as

$$w_n(x) = \binom{n}{x} p^x q^{n-x}$$

where p is the probability of success, $q = 1 - p$ is probability of failure and

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

is the binomial coefficient that counts the number of ways to select x items from n items without caring about the order of selection. Here x is a discrete variable, $x \in \mathbb{Z}$, with $0 \leq x \leq n$.

The trick is to find a way to deal with the fact that $x \in \mathbb{Z}$ (x is a discrete variable) for the Binomial Distribution and $x \in \mathbb{R}$ (x is a continuous variable) for the Normal Distribution^[3]. In other words as we let $n \rightarrow \infty$ we need to come up with a way to let Δx shrink^[4] so that a probability density limit (the Normal Distribution) is reached from a sequence of probability distributions (modified Binomial Distributions). So let $w(x)$ represent the Normal Distribution with mean $\bar{x} = np$ and variance $\sigma^2 = npq$. We will show how $\lim_{n \rightarrow \infty} w_n(x) = w(x)$ where each Binomial Distribution $w_n(x)$ also has mean $\bar{x} = np$ and variance $\sigma^2 = npq$.

The heart of the trick is to notice^[5] that

$$\frac{d}{dx} \ln w(x) = \lim_{\Delta x \rightarrow 0} \frac{w(x + \Delta x) - w(x)}{w(x) \Delta x}.$$

This is perfectly true for the density $w(x)$. The trick is to substitute the distribution $w_n(x)$ for the density $w(x)$ in the RHS of Equation (5.4) to get :

$$\frac{w_n(x + \Delta x) - w_n(x)}{w_n(x) \Delta x} = \frac{w_n(x + 1) - w_n(x)}{w_n(x)}$$

because $\Delta x = 1$. The trick is to now pretend that $w_n(x)$ is a continuous function defined at all $x \in \mathbb{R}$; we just don't know what its values should be for non-integer x . With such a “continuation” of $w_n(x)$ we can write^[6]

$$\frac{d}{dx} \ln w(x) = \lim_{n \rightarrow \infty} \frac{w_n(x + 1) - w_n(x)}{w_n(x)} \quad (5.6)$$

$$= \lim_{n \rightarrow \infty} \frac{\binom{n}{x+1} p^{x+1} q^{n-x-1}}{\binom{n}{x} p^x q^{n-x}} - 1 \quad (5.7)$$

$$= \lim_{n \rightarrow \infty} \frac{n-x}{x+1} \frac{p}{q} - 1. \quad (5.8)$$

Equation (5.8) has no limit; it blows up as $n \rightarrow \infty$. We need to transform x in such a way to gain control on Δx (getting it to shrink as $n \rightarrow \infty$) and to get something that converges. To do that we introduce $h = \frac{1}{\sqrt{n}}$ and a new variable $u = h(x - \bar{x}) = h(x - np)$. With this transformation of variables, the chain rule gives

$$\frac{d}{dx} \ln w(x) = \frac{du}{dx} \frac{d}{du} \ln w(u) = h \frac{d}{du} \ln w(u)$$

and the RHS of Equation (5.8) becomes, using $x = \frac{u}{h} + np$

$$\frac{n-x}{x+1} \frac{p}{q} - 1 = \frac{\left(n - \frac{u}{h} - np\right) p}{\left(\frac{u}{h} + np + 1\right) q} \quad (5.10)$$

$$= \frac{\left(n(1-p) - \frac{u}{h}\right)}{\left(\frac{u}{h} + np + 1\right) \frac{q}{p}} \quad (5.11)$$

$$= \frac{\left(nq - \frac{u}{h}\right)}{\left(\frac{uq}{hp} + nq + \frac{q}{p}\right)} \quad (5.12)$$

$$= \frac{\left(1 - \frac{u}{nhq}\right)}{\left(\frac{u}{nhp} + 1 + \frac{1}{np}\right)} \quad (5.13)$$

$$= \frac{\left(1 - \frac{u}{nhq}\right)}{\left(1 + \frac{u+h}{nhp}\right)} \quad (5.14)$$

Using Equation (5.9), for the LHS, and Equation (5.14), for the RHS, Equation (5.8) becomes

$$h \frac{d}{du} \ln w(u) = \lim_{n \rightarrow \infty} \frac{1 - \frac{u}{nhq}}{1 + \frac{u+h}{nhp}} - 1 \quad (5.15)$$

$$= \lim_{n \rightarrow \infty} \left(1 - \frac{u}{nhq}\right) \left[1 - \frac{u+h}{nhp} + \left(\frac{u+h}{nhp}\right)^2 - \dots\right] - 1 \quad (5.16)$$

$$= \lim_{n \rightarrow \infty} -\frac{1}{np} - \frac{u}{nhq} - \frac{u}{nhp} + O\left(\frac{1}{n}\right) \quad (5.17)$$

$$= \lim_{n \rightarrow \infty} -\frac{1}{np} - \frac{u}{nhpq} + O\left(\frac{1}{n}\right) \quad (5.18)$$

$$= \lim_{n \rightarrow \infty} -\frac{u}{nhpq}. \quad (5.19)$$

where $O\left(\frac{1}{n}\right)$ means terms that will go to zero as $n \rightarrow \infty$, and we have used the relation $\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots$ to get Equation (5.16)) and $p + q = 1$ to go from Equation (5.17) to Equation (5.18). Dividing both sides of Equation (5.19) by h leaves

$$\frac{d}{du} \ln w(u) = \lim_{n \rightarrow \infty} -\frac{u}{nh^2 pq} = -\frac{u}{pq}.$$

Our transformation, with its \sqrt{n} , has given us the exact control we need to keep the limit from disappearing or blowing up. Integrating Equation (5.20) gives

$$w(u) = C e^{-\frac{u^2}{2pq}}$$

where C is the a constant of integration. Switching back to the x variable

$$w(x) = C e^{-\frac{(h[x-\bar{x}])^2}{2pq}} \quad (5.22)$$

$$= C e^{-\frac{(x-\bar{x})^2}{2npq}} \quad (5.23)$$

$$= C e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}. \quad (5.24)$$

To evaluate the constant of integration, C , we impose $\int_{-\infty}^{\infty} w(x) dx = 1$ because we want $w(x)$ to be a probability distribution. So

$$C \int_{-\infty}^{\infty} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} dx = C \sqrt{2\pi\sigma^2} = 1$$

so

$$C = \frac{1}{\sqrt{2\pi\sigma^2}}$$

and

$$w(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

which is the Normal Distribution that approximates Binomial Distributions with the same mean and variance as n gets large.

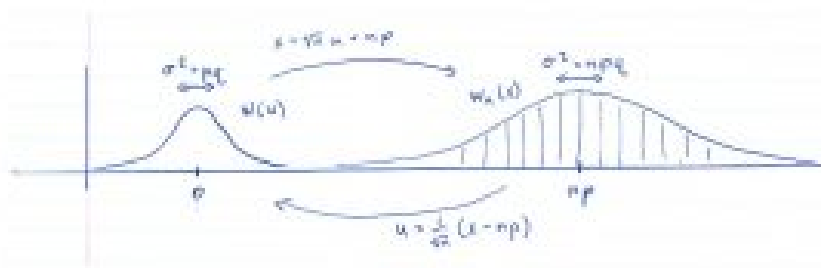


Figure 5.1 : The transformation $u = \frac{(x-np)}{\sqrt{n}}$ effectively shrinks the Δx of the Binomial Distribution with mean $\bar{x} = np$ and variance $\sigma^2 = npq$ by pulling a continuous version $w_n(x)$ back to the constant Normal Distribution $w(u)$. Another way of thinking about it is that the transformation $x = \sqrt{n}u + np$ takes the fixed Normal Distribution $w(u)$ to the Normal Distribution $w(x)$ that provides a better and better approximation of $w_n(x)$ as $n \rightarrow \infty$.

You may be wondering why that transformation $u = \frac{1}{\sqrt{n}}(x - np)$ worked because it seems to have been pulled from the air. According to Lindsay & Margenau, it was Laplace who first used this transformation and derivation in 1812. What this transformation does is pull the Binomial Distribution $w_n(x)$ back to have a mean of zero (by subtracting $\bar{x} = np$) which keeps x from running off to infinity and, more importantly, allows us to define a function $w(u)$ with $u \in \mathbb{R}$ that has a constant variance of pq that we can match to npq when we transform back to x at each n , see Figure 5.1. Looking at it the other way around, the Normal Distribution^[7] $w(x)$ with $x = \sqrt{n}u + np$ is an approximation for Binomial Distribution $w_n(x)$ that “asymptotically” approaches $w_n(x)$ as $n \rightarrow \infty$.

This is not the only way to form a probability density limit from a sequence of Binomial distributions. It is one that gives a good approximation of the Binomial Distribution when n is fairly small if the term $\frac{1}{np}$ in Equation (5.18) becomes small quickly. If p is very small, this does not happen and another limit of Binomial Distributions that leads to the Poisson Distribution is more appropriate. When p and q are close to 0.5 or more generally when $np \geq 5$ and $nq \geq 5$ then the Normal approximation is a good one. Either way, the density limit is a mathematical idealization, a convenience really, that is based on a discrete probability distribution that just summarizes the result of counting outcomes. Counting gives the foundation for probability theory.

1. The formula for the Binomial Distribution was apparently derived by Newton according to: Lindsay RB, Margenau. Foundations of Physics. Dover, New York, 1957 (originally published 1936). For that claim, Lindsay & Margenau quote: von Mises R. Probability, Statistics, and Truth. Macmillan, New York, 1939 (originally published 1928). The derivation of the Normal Distribution presented here largely follows that given in Lindsay & Margenau's book. ↩
2. In class we denoted the Binomial distribution as $P(x | n)$. Here we use $w_n(x) = P(x | n)$ to avoid using too many P's and p's. ↩
3. Remember that the Normal Distribution is technically a probability density but we slur the use of the word distribution between probability distribution (discrete x) and probability density (continuous x) like everyone else. ↩
4. $\Delta x = 1$ for the Binomial Distribution. ↩
5. Remember that $\frac{d}{dx} \ln(x) = \frac{1}{x}$ and use the chain rule to notice this. ↩
6. You can probably imagine many ways to continue the Binomial Distribution from $x \in \mathbb{Z}$ to $x \in \mathbb{R}$. It doesn't matter which one you pick as long as the behaviour of your new function is not too crazy between the integers; that is, $\lim_{n \rightarrow \infty} w_n(x)$ should exist at all $x \in \mathbb{R}$. ↩
7. Our symbols here are not mathematically clean; we should write something like $w(u(x))$ instead of $w(x)$ or w composed with u at x , $w \circ u_n(x)$, instead of $w(x)$. But to emphasize the intuition we use $w(x)$. In clean symbols, the function $w \circ u_n(x)$

asymptotically approaches $w_n(x)$ where $u_n(x) = \frac{(x-np)}{\sqrt{npq}}$. ↵

This page titled [5.2: **The Normal Distribution as a Limit of Binomial Distributions](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Gordon E. Sarty](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.