# Lecture 9
# Central Limit Theorem

**Chao Song**

College of Ecology

Lanzhou University

October 14, 2024

# Central limit theorem

From the previous lecture, we know that if $X_1, X_2, \ldots, X_n$ are a random sample from a normal distribution $N(\mu, \sigma^2)$, then the sample mean

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n}) \quad or \quad \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

**Central Limit Theorem**: If $\overline{X}$ is the mean of a random sample $X_1, X_2, \ldots, X_n$ of size $n$ from a distribution with a finite mean $\mu$ and a finite positive variance $\sigma^2$, then the distribution of

$$X = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}$$

is $N(0,1)$ in the limit as $n \to \infty$

# Proof of central limit theorem

If a sequence of MGFs approaches a certain MGF, say $M(t)$, for $t$ in an open interval around 0, then the limit of the corresponding distributions must be the distribution corresponding to $M(t)$.

We first consider the MGF of $W$,

$$\begin{aligned} m_W(t) = E(e^{tW}) &= E\left[\exp\left(t\frac{(\sum_{i=1}^{n} X_i - n\mu)}{\sqrt{n}\sigma}\right)\right] \\ &= E\left[\exp\left[\left(\frac{t}{\sqrt{n}}\right)\left(\frac{X_1-\mu}{\sigma}\right)\right]\cdots\exp\left[\left(\frac{t}{\sqrt{n}}\right)\left(\frac{X_n-\mu}{\sigma}\right)\right]\right] \\ &= E\left[\exp\left[\left(\frac{t}{\sqrt{n}}\right)\left(\frac{X_1-\mu}{\sigma}\right)\right]\right]\cdots E\left[\exp\left[\left(\frac{t}{\sqrt{n}}\right)\left(\frac{X_n-\mu}{\sigma}\right)\right]\right] \end{aligned}$$

which follows from the independence of $X_1, X_2, \ldots, X_n$. Then

$$E(e^{tW}) = \left[m\left(\frac{t}{\sqrt{n}}\right)\right]^n$$

where $m(t)$ is the common MGF of each $Y_i = (X_i - \mu)/\sigma$.

# Proof of central limit theorem

We know $E(Y_i) = 0$ and $E(Y_i^2) = 1$, thus,

$$m(0) = 1, \quad m'(0) = 0, \quad m''(0) = 1$$

Hence, using Taylor's formula with a remainder, we know that there exist a number $t_1$ between 0 and $t$ such that

$$m(t) = m(0) + m'(0)t + \frac{m''(t_1)t^2}{2} = 1 + \frac{m''(t_1)t^2}{2}$$

Using this expression of MGF, we have

$$m_W(t) = \left[m\Big(\frac{t}{\sqrt{n}}\Big)\right]^n = \left(1 + \frac{m''(t_1)t^2}{2n}\right)^n$$

where $t_1$ is between 0 and $t/\sqrt{n}$. Here, we see that $t_1 \to 0$ and $m''(t_1) \to 1$ as $n \to \infty$.

# Proof of central limit theorem

Thus, we obtain the MGF of $W$ as $n \to \infty$

$$\lim_{n\to\infty} m_W(t) = \lim_{n\to\infty} \left(1 + \frac{m''(t_1)t^2}{2n}\right)^n = e^{\frac{t^2}{2}}$$

Here, $e^{t^2/2}$ is the MGF of a standard normal distribution. It follows that the limiting distribution of

$$W = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

is a standard normal distribution, i.e., $N(0,1)$.

# Approximation for discrete distributions

From central limit theorem, we see that the distribution of any random variable that is the sum of independent and identically distributed random variables can be approximated by a normal distribution.

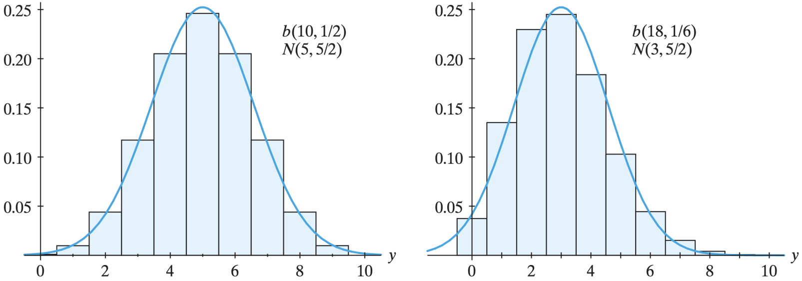
Recall that a binomial random variables can be described as the sum of Bernoulli distributions. If $Y$ has a binomial distribution, central limit theorem states that the distribution of

$$W = \frac{Y - np}{\sqrt{np(1-p)}}$$

is $N(0, 1)$ in the limit as $n \to \infty$. Thus, if $n$ is "sufficiently large", the distribution of $Y$ is approximately $N\left[np, np(1-p)\right]$

# Approximation for discrete distributions

If $n$ is "sufficiently large", the distribution of $Y$ is approximately $N\big(np, np(1-p)\big)$. A rule often stated is that $n$ is sufficiently large if $np \geqslant 5$ and $n(1-p) \geqslant 5$.



# Approximation for discrete distributions

A random variable $Y$ having a Poisson distribution with mean $\lambda$ can be thought of as the sum of $\lambda$ Poisson distributed random variables with mean 1. Thus,

$$W = \frac{Y - \lambda}{\lambda}$$

has a distribution that is approximately $N(0, 1)$, and the distribution of $Y$ is approximately $N(\lambda, \lambda)$.

# Approximation for discrete distributions

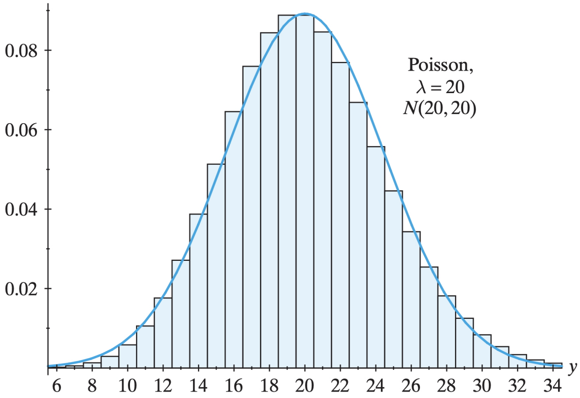
For a discrete distribution, $P(Y = k)$ can be represented by the are of the rectangle with a height of $P(Y = k)$ and a base of length 1 centered at $k$. When approximating the probability using a normal distribution, we use the area under the PDF of a normal distribution between $k - \frac{1}{2}$ and $k + \frac{1}{2}$. This is often referred to as the **half-unit correction for continuity**.

$$P(Y \leqslant k) \approx \Phi(\frac{k + 1/2 - \mu}{\sigma})$$

$$P(Y < k) \approx \Phi(\frac{k - 1/2 - \mu}{\sigma})$$

# Approximation for discrete distributions

The normal approximation for a Poisson distribution is “good” when $\lambda \geqslant 5$.



# Approximation for discrete distributions

**Example**: Let $Y$ have a binomial distribution with $n = 10$ and $p = 0.5$. Using normal approximation to find $P(3 \leqslant Y < 6)$.

The mean and variance of $Y$ is $10 \times 0.5 = 5$ and $10 \times 0.5 \times (1 - 0.5) = 2.5$.

$$\begin{aligned} P(3 \leqslant Y < 6) &= P(2.5 \leqslant Y \leqslant 5.5) \\ &= P(\frac{2.5 - 5}{\sqrt{2.5}} \leqslant \frac{Y - 5}{\sqrt{2.5}} \leqslant \frac{5.5 - 5}{\sqrt{2.5}}) \\ &= \Phi(0.316) - \Phi(-1.581) \\ &= 0.5672 \end{aligned}$$

We can also calculate the probability based on binomial distribution:

$$\begin{aligned} P(3 \leqslant Y < 6) &= P(Y = 3) + P(Y = 4) + P(Y = 5) \\ &= 0.1172 + 0.2051 + 0.2461 \\ &= 0.5683 \end{aligned}$$