Limit Theorems

Chia-Ping Chen

Professor
Department of Computer Science and Engineering
National Sun Yat-sen University

Probability

Outline

- Probability inequalities
- Weak law of large numbers
- Convergence of sequence of random variables
- Central limit theorem
- Strong law of large numbers

Probability Inequalities (Bounds)

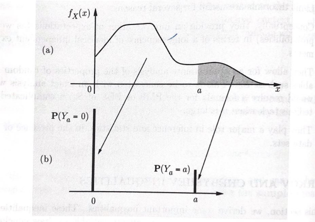
Markov inequality

Let X be a non-negative random variable. For any $a > 0$

$$P(X \geq a) \leq \frac{\mathbf{E}[X]}{a}$$

Define

$$Y_a = \begin{cases} 0, & X < a \\ a, & X \geq a \end{cases}$$



Note $Y_a \leq X$ so $\mathbf{E}[Y_a] \leq \mathbf{E}[X]$. Thus

$$\mathbf{E}[Y_a] = aP(Y_a = a) = aP(X \geq a) \leq \mathbf{E}[X]$$

$$\Rightarrow \; P(X \geq a) \leq \frac{\mathbf{E}[X]}{a}$$

Chebyshev inequality

Let X be a random variable. For any $c > 0$

$$P\left(|X - \mathbf{E}[X]| \geq c\right) \leq \frac{\text{var}(X)}{c^2}$$

Define $Z = (X - \mathbf{E}[X])^2$. Then $Z \geq 0$ and

$$P(|X - \mathbf{E}[X]| \geq c) = P((X - \mathbf{E}[X])^2 \geq c^2) = P(Z \geq c^2)$$

By Markov inequality

$$P(Z \geq c^2) \leq \frac{\mathbf{E}[Z]}{c^2}$$

That is

$$P(|X - \mathbf{E}[Z]| \geq c) \leq \frac{\text{var}(X)}{c^2}$$

Inverse square bound

Let X be a random variable with $\mathbf{E}[X] = \mu$ and $\text{var}(X) = \sigma^2$.

$$P\left(|X - \mu| \geq k\sigma\right) \leq \frac{1}{k^2}$$

By Chebyshev inequality

$$P\left(|X - \mu| \geq c\right) \leq \frac{\sigma^2}{c^2}$$

Let $c = k\sigma$

$$P\left(|X - \mu| \geq k\sigma\right) \leq \frac{\sigma^2}{(k\sigma)^2} = \frac{1}{k^2}$$

Example (Markov and Chebyshev)

Consider $X \sim \textbf{Uni}(0,4)$. We have $\mathbf{E}[X] = 2$ and $\text{var}(X) = \frac{4}{3}$.

- By Markov inequality

$$P(X \geq 2) \leq \frac{\mathbf{E}[X]}{2} = 1$$
$$P(X \geq 3) \leq \frac{\mathbf{E}[X]}{3} = \frac{2}{3}$$

- By Chebyshev inequality

$$P(|X - \mathbf{E}[X]| \geq 1) \leq \frac{\text{var}(X)}{1^2} \;\Rightarrow\; P(|X-2| \geq 1) \leq \frac{4}{3}$$

They are very loose bounds.

Example (5.3)

Let X be a random variable with range $\mathcal{X}$.

- If $\mathcal{X}$ is bounded, **var**(X) is bounded
- For $\mathcal{X} \subseteq [a, b]$

$$\textbf{var}(X) \leq \frac{(b-a)^2}{4}$$

- The probability upper bound in Chebyshev inequality can be further relaxed

$$P(|X - \mathbf{E}[X]| \geq c|) \leq \frac{\textbf{var}(X)}{c^2} \leq \frac{(b-a)^2}{4c^2}$$

Tail vs. Body

Chebyshev (and Markov) inequality bounds "**tail**" probability. We can equivalently bound "**body**" probability. Let X be a random variable. For any $c > 0$

$$P\left(|X - \mathbf{E}[X]| < c\right) \geq 1 - \frac{\mathrm{var}(X)}{c^2}$$

This follows from the Chebyshev inequality

$$\begin{aligned} & P\left(|X - \mathbf{E}[X]| \geq c\right) \leq \frac{\mathrm{var}(X)}{c^2} \\ \Rightarrow\ & 1 - P\left(|X - \mathbf{E}[X]| \geq c\right) \geq 1 - \frac{\mathrm{var}(X)}{c^2} \\ \Rightarrow\ & P\left(|X - \mathbf{E}[X]| < c\right) \geq 1 - \frac{\mathrm{var}(X)}{c^2} \end{aligned}$$

Weak Law of Large Numbers

Definition (sample mean)

Let X be a random variable with finite mean and variance. Let $X_1, X_2, \cdots$ denote independent and repeated measurements (samples) of X. The X_i's are **iid** (independent and identically distributed) random variables with the same probability function as X. Define **sample mean**

$$M_n = \frac{X_1 + \cdots + X_n}{n}$$

The mean and variance of M_n are

$$\mathbf{E}[M_n] = \frac{\mathbf{E}[X_1 + \cdots + X_n]}{n} = \frac{n\mathbf{E}[X]}{n} = \mathbf{E}[X]$$

$$\textbf{var}(M_n) = \frac{1}{n^2}\textbf{var}(X_1 + \cdots + X_n) = \frac{n\text{var}(X)}{n^2} = \frac{\text{var}(X)}{n}$$

Weak law of large numbers

Sample mean converges to expectation. Specifically

$$\lim_{n\to\infty} P(|M_n - \mathbf{E}[X]| < \epsilon) = 1, \ \forall\, \epsilon > 0$$

Apply the Chebyshev corollary to M_n to get

$$P\left(|M_n - \mathbf{E}[M_n]| < \epsilon\right) \geq 1 - \frac{\textbf{var}(M_n)}{\epsilon^2}$$

Substitute $\mathbf{E}[M_n] = \mathbf{E}[X]$ and $\textbf{var}(M_n) = \frac{\text{var}(X)}{n}$

$$1 \geq P\left(|M_n - \mathbf{E}[X]| < \epsilon\right) \geq 1 - \frac{\text{var}(X)}{n\epsilon^2}$$

Thus

$$\lim_{n\to\infty} P(|M_n - \mathbf{E}[X]| < \epsilon) = 1$$

Example (5.4)

Let $(\Omega, \mathcal{F}, P)$ be a probability model based on random experiment $\mathcal{R}$ and $A \in \mathcal{F}$ be an event. The relative frequency of A in a sequence of independent trials of $\mathcal{R}$ converges to $P(A)$.

Define

$$I^A = \begin{cases} 1, & \text{if event } A \text{ occurs in a trial} \\ 0, & \text{otherwise} \end{cases}$$

Let I_i^A indicate the occurrence of A in trial i, then $I_1^A, I_2^A, \cdots$ are **iid** random variables with mean $\mathbf{E}[I^A] = P(A)$. By WLLN, we have $\lim\limits_{n\to\infty} P(|M_n - \mathbf{E}[I^A]| < \epsilon) = 1$ or

$$\lim_{n\to\infty} P\left(\left| \frac{I_1^A + \cdots + I_n^A}{n} - P(A) \right| < \epsilon \right) = 1$$

That is, the relative frequency of A converges to $P(A)$.

Example (5.5)

Let p be the fraction of the population supporting a particular candidate. We poll n persons and record the fraction in the poll supporting the candidate as an estimate of p. How good is this estimation?

Let X indicate "supporting the candidate" of a pollee. Then $X_1, X_2, \ldots$ are approximately **iid** and M_n is the fraction supporting the candidate, with

$$\mathbf{E}[M_n] = \mathbf{E}[X] = P(X = 1) = p, \ \textsf{var}(M_n) = \frac{p(1-p)}{n}$$

By Chebyshev

$$P\left(|M_n - p| < \epsilon\right) \geq 1 - \frac{p(1-p)}{n\epsilon^2} \geq 1 - \frac{1}{4n\epsilon^2}$$

The quality of the estimation of p by M_n depends on n. Take $n = 100$ for example. We have

$$P\left(|M_n - p| < 0.1\right) \geq 1 - \frac{1}{4(100)(0.01)} = 1 - 0.25 = 0.75$$

Definition (margin and confidence)

The quality of estimation is often quantified by margin and confidence.

- Margin bounds estimation error
- Confidence bounds the probability that error is within margin

Let $\hat{p}$ be an estimator of p.

- We aim to establish inequality

$$P\left(|\hat{p} - p| < \epsilon\right) \geq q_0$$

- ϵ is margin and q_0 is confidence
- The probability that $|\hat{p} - p|$ is smaller than ϵ is at least q_0

Example (Margin, confidence, size)

In Example 5.5 we establish

$$P\left(|M_n - p| < \epsilon\right) \geq 1 - \frac{1}{4n\epsilon^2} \geq q_0$$

- The margin ϵ, confidence q_0 and sample size n are related by

$$\frac{1}{4n\epsilon^2} \leq 1 - q_0$$

- For margin ϵ and confidence q_0, the required sample size n is

$$n \geq \frac{1}{4(1 - q_0)\epsilon^2}$$

- For example, for $\epsilon = 0.01$ and $q_0 = 0.95$

$$n \geq \frac{1}{4(0.05)(0.01)^2} = 50000$$

Convergence of Sequence of Random Variables

Convergence

- convergence in probability
- convergence in distribution
- almost-sure convergence

Definition (convergence in probability)

Let $(\Omega, \mathcal{F}, P)$ be a probability model and $Y_1, Y_2, \cdots$ be random variables defined on Ω. The sequence $Y_1, Y_2, \cdots$ converges in probability to Y if

$$\lim_{n\to\infty} P(|Y_n - Y| < \epsilon) = 1, \ \forall \epsilon > 0$$

- Making ϵ arbitrarily small, we can see there is nothing between Y_n and Y with probability 1
- Convergence in probability is denoted by

$$Y_n \xrightarrow{P} Y$$

Example (5.6)

Let X_i be **Uni**$(0,1)$ and $Y_n = \min(X_1, \cdots, X_n)$. Show that

$$Y_n \xrightarrow{P} 0$$

For any $\epsilon > 0$, we have

$$P(|Y_n - 0| < \epsilon) = 1 - P(|Y_n - 0| \geq \epsilon) = 1 - (1-\epsilon)^n$$

So

$$\lim_{n\to\infty} P(|Y_n - 0| < \epsilon) = \lim_{n\to\infty} 1 - (1-\epsilon)^n = 1$$

That is

$$Y_n \xrightarrow{P} 0$$

Example (5.7)

Let Y be **Exp**(1) and $Y_n = \frac{Y}{n}$. Show that

$$Y_n \xrightarrow{P} 0$$

For any $\epsilon > 0$, we have

$$\begin{aligned} P(|Y_n - 0| < \epsilon) &= 1 - P(|Y_n - 0| \geq \epsilon) = 1 - P\left(\frac{Y}{n} \geq \epsilon\right) \\ &= 1 - P\left(Y \geq n\epsilon\right) \\ &= 1 - e^{-n\epsilon} \end{aligned}$$

So

$$\lim_{n\to\infty} P(|Y_n - 0| < \epsilon) = \lim_{n\to\infty} (1 - e^{-n\epsilon}) = \lim_{n\to\infty} 1 - (e^{-\epsilon})^n = 1$$

That is

$$Y_n \xrightarrow{P} 0$$

Definition (convergence in distribution)

Let $(\Omega, \mathcal{F}, P)$ be a probability model and $Y_1, Y_2, \cdots$ be random variables defined on Ω. The sequence $Y_1, Y_2, \cdots$ converges in distribution to Y if the sequence of CDF converges

$$\lim_{n\to\infty} F_{Y_n}(t) = F_Y(t), \ \forall t$$

Convergence in distribution is denoted by

$$Y_n \xrightarrow{D} Y$$

Definition (sample sequence)

Let $(\Omega, \mathcal{F}, P)$ be a probability model and $Y_1, Y_2, \cdots$ be random variables defined on Ω. For $\omega \in \Omega$

$$Y_1(\omega), Y_2(\omega), \cdots$$

is a sample sequence of sequence $Y_1, Y_2, \cdots$.

Definition (sure convergence)

Let $(\Omega, \mathcal{F}, P)$ be a probability model and $Y_1, Y_2, \cdots$ be random variables defined on Ω.

- The sequence $Y_1, Y_2, \cdots$ converges surely if

$$\left\{\omega \in \Omega \;\middle|\; \lim_{n\to\infty} Y_n(\omega) \text{ exists}\right\} = \Omega$$

- Sure convergence means every sample sequence converges

Definition (almost-sure convergence)

Let $(\Omega, \mathcal{F}, P)$ be a probability model and $Y_1, Y_2, \cdots$ be random variables defined on Ω.

- The sequence $Y_1, Y_2, \cdots$ converges almost surely if

$$P\left(S = \left\{\omega \in \Omega \,\middle|\, \lim_{n\to\infty} Y_n(\omega) \text{ exists}\right\}\right) = 1$$

- Almost-sure convergence means $Y_1, Y_2. \cdots$ converges with probability 1
- The event S^c that $Y_1(\omega), Y_2(\omega) \cdots$ does not converge has probability 0

Almost-sure convergence to a random variable

Let $(\Omega, \mathcal{F}, P)$ be a probability model and $Y_1, Y_2, \cdots$ be random variables defined on Ω. The sequence $Y_1, Y_2, \cdots$ converges almost surely to Y if

$$P\left(S = \left\{\omega \in \Omega \,\middle|\, \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\right\}\right) = 1$$

- An element in S satisfies 2 conditions: $Y_n(\omega)$ converges, and it converges to $Y(\omega)$
- Almost-sure convergence is denoted by

$$P\left(\lim_{n\to\infty} Y_n = Y\right) = 1 \text{ or } Y_n \xrightarrow{\text{a.s.}} Y$$

- Suppose $Y_1, Y_2, \cdots$ converges almost surely. Then $Y_n \xrightarrow{\text{a.s.}} Y$ where $Y(\omega) = \lim_{n\to\infty} Y_n(\omega)$ for $\omega \in S$.

Example (5.14)

Let $X_1, X_2, \cdots$ be **iid Uni**$(0,1)$ and $Y_n = \min(X_1, \cdots, X_n)$. Show that the sequence $Y_1, Y_2, \cdots$ converges to 0 almost surely.

Any sample sequence $Y_1(\omega), Y_2(\omega), \cdots$ converges because it is non-increasing and bounded below by 0. Thus the sequence $Y_1, Y_2, \cdots$ converges surely. For any $\epsilon > 0$, we have

$$P\left(\lim_{n\to\infty} Y_n \geq \epsilon\right) = P\left(\bigcap_{i=1}^{\infty}(X_i \geq \epsilon)\right) = \lim_{n\to\infty}(1-\epsilon)^n = 0$$

So

$$P\left(\lim_{n\to\infty} Y_n = 0\right) = 1 - P\left(\lim_{n\to\infty} Y_n > 0\right) = 1$$

Thus, the sequence $Y_1, Y_2, \cdots$ converges to 0 almost surely.

Strength of convergence*

Almost-sure convergence implies convergence in probability.

Let $S = \left(\lim_{n\to\infty} Y_n = Y\right)$ and $S_n(\epsilon) = (|Y_n - Y| < \epsilon)$.

$$\begin{aligned}\omega \in S &\Rightarrow \lim_{n\to\infty} Y_n(\omega) = Y(\omega) \\ &\Rightarrow \lim_{n\to\infty} |Y_n(\omega) - Y(\omega)| = 0 \\ &\Rightarrow \exists n_0 \text{ s.t. } |Y_n(\omega) - Y(\omega)| < \epsilon \text{ for all } n > n_0 \\ &\Rightarrow \exists n_0 \text{ s.t. } \omega \in S_n(\epsilon) \text{ for all } n > n_0\end{aligned}$$

So $S \subset S_n(\epsilon)$ for all $n > n_0$. Suppose $Y_n \xrightarrow{\text{a.s.}} Y$. We have

$$P(S) = 1 \Rightarrow P(S_{n>n_0}(\epsilon)) = 1 \Rightarrow \lim_{n\to\infty} P(S_n(\epsilon)) = 1 \Rightarrow Y_n \xrightarrow{P} Y$$

Example (5.15)

In an arrival process, the time slots are partitioned into consecutive intervals

$$\{2^k, \cdots, 2^{k+1}-1\},\ k = 0, 1, \cdots$$

In each interval, there is exactly one arrival, and all slots within the interval are equally likely. Define $Y_n = 1$ for an arrival at slot n, and $Y_n = 0$ otherwise. Show that $Y_n \xrightarrow{P} 0$, but **not** $Y_n \xrightarrow{\text{a.s.}} 0$.

$$\begin{aligned}
\lim_{n\to\infty} P(|Y_n - 0| < \epsilon) &= 1 - \lim_{n\to\infty} P(|Y_n - 0| \geq \epsilon) \\
&= 1 - \lim_{n\to\infty} P(Y_n = 1) = 1 - \lim_{n\to\infty} \frac{1}{2^{\lfloor \log_2 n \rfloor}} \\
&= 1
\end{aligned}$$

Since any sample sequence of $Y_1\, Y_2 \cdots$ is non-convergent, we have

$$P\left(\lim_{n\to\infty} Y_n = 0\right) = 0 \neq 1$$

Central Limit Theorem

Definition (sample sum and standardized sample mean)

Let X be a random variable with $\mathbf{E}[X] = \mu$ and $\mathrm{var}(X) = \sigma^2$. Let $X_1, X_2, \cdots$ denote independent samples of X. Define sample sum

$$S_n = X_1 + \cdots + X_n$$

and standardized sample mean

$$Z_n = \frac{S_n/n - \mu}{\sigma/\sqrt{n}} = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

$$\mathbf{E}[S_n] = \mathbf{E}[X_1 + \cdots + X_n] = n\,\mathbf{E}[X_i] = n\mu$$

$$\mathbf{var}(S_n) = \mathbf{var}(X_1 + \cdots + X_n) = n\,\mathbf{var}(X_i) = n\sigma^2$$

$$\mathbf{E}[Z_n] = \mathbf{E}\left[\frac{S_n - n\mu}{\sigma\sqrt{n}}\right] = 0,\ \ \mathbf{var}(Z_n) = \frac{\mathbf{var}(S_n)}{n\sigma^2} = 1$$

Central limit theorem

Let $X_1, X_2, \cdots$ be **iid** samples of a random variable with finite mean and variance. Then the sequence of standardized sample means $Z_1, Z_2, \cdots$ converges in distribution to the standard normal.

That is

$$Z_n \xrightarrow{D} Y \sim \mathcal{N}(0,1)$$

In other words

$$F_{Z_n}(t) \xrightarrow{n \to \infty} \Phi(t), \quad \forall t$$

Normal approximation

Let $X_1, X_2, \cdots$ be **iid** samples of random variable X with mean μ and variance σ^2. For a large n

- standardized sample mean is approximately normal

$$Z_n \dot{\sim} \mathcal{N}(0,1)$$

- sample sum is approximately normal

$$S_n \dot{\sim} \mathcal{N}\left(n\mu, n\sigma^2\right)$$

- sample mean is approximately normal

$$M_n \dot{\sim} \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

Example (5.9)

The weight of package is $X \sim \textbf{Uni}(5, 50)$. We load 100 such packages on a plane. What is the probability that the total weight exceeds 3000 pounds?

Let $S_n = X_1 + \cdots + X_n$ where $n = 100$. With $\mu = \mathbf{E}[X] = 27.5$ and $\sigma^2 = \textbf{var}(X) = 168.75$, we have

$$S_n \dot\sim \mathcal{N}(n\mu, n\sigma^2) \text{ i.e. } S_{100} \dot\sim \mathcal{N}(2750, 16875)$$

Thus

$$\begin{aligned} P(S_{100} > 3000) &= 1 - P(S_{100} \leq 3000) \\ &= 1 - P\left(\frac{S_{100} - 2750}{\sqrt{16875}} \leq \frac{3000 - 2750}{\sqrt{16875}}\right) \\ &\approx 1 - P(Y \leq 1.92) \\ &= 1 - \Phi(1.92) \\ &= 0.0274 \end{aligned}$$

Example (5.10)

The processing time of a part is $T \sim \textbf{Uni}(1,5)$. Estimate the probability that the number of parts processed within 320 time units, denoted by N_{320}, is at least 100.

Let S_{100} be the time to process 100 parts. Note $(N_{320} \geq 100) = (S_{100} \leq 320)$. With $\mu = \mathbf{E}[T] = 3$ and $\sigma^2 = \textbf{var}(T) = \frac{4}{3}$, we have

$$S_{100} \dot{\sim} \mathcal{N}\left(100\,\mu, 100\,\sigma^2\right) = \mathcal{N}\left(300, \frac{400}{3}\right)$$

Thus

$$\begin{aligned} P(S_{100} \leq 320) &= P\left(\frac{S_{100} - 300}{\sqrt{\frac{400}{3}}} \leq \frac{320 - 300}{\sqrt{\frac{400}{3}}}\right) \\ &\approx P(Y \leq 1.73) = \Phi(1.73) \\ &= 0.9582 \end{aligned}$$

Bernoulli sample mean

Let $X_1, X_2, \ldots$ be samples of $X \sim \textbf{Ber}(p)$. Then

$$M_n \dot{\sim} \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$$

The mean and variance of X are

$$\mu = \mathbf{E}[X] = p$$

$$\sigma^2 = \textbf{var}(X) = p(1-p)$$

It follows that

$$M_n \dot{\sim} \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) = \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$$

Example (5.11)

Consider Example 5.5 that estimates p by M_n of a sample of size n. Using the normal approximation of M_n, we have

$$
\begin{aligned}
P(|M_n - p| \geq \epsilon) &\approx 2\,P(M_n - p \geq \epsilon) \\
&= 2\,P\left(\frac{M_n - p}{\sqrt{\frac{p(1-p)}{n}}} \geq \frac{\epsilon}{\sqrt{\frac{p(1-p)}{n}}}\right) \\
&\approx 2\left[1 - \Phi\left(\sqrt{\frac{1}{p(1-p)}}\sqrt{n}\epsilon\right)\right] \\
&\leq 2\left[1 - \Phi(2\sqrt{n}\epsilon)\right]
\end{aligned}
$$

For example, with $n = 100$ and $\epsilon = 0.1$, we have

$$
\begin{aligned}
P(|M_{100} - p| \geq 0.1) &\leq 2\left[1 - \Phi(2 \cdot \sqrt{100} \cdot 0.1)\right] \\
&= 2\,[1 - \Phi(2)] \\
&= 0.0456
\end{aligned}
$$

Sample size

Consider the estimation of p by M_n with margin ϵ and confidence q_0. We want a sample size n to guarantee

$$P(|M_n - p| < \epsilon) \geq q_0$$

- By Bernoulli Normal approximation of M_n

$$P(|M_n - p| < \epsilon) \geq 1 - 2\left[1 - \Phi(2\sqrt{n}\epsilon)\right] \geq q_0$$

$$\Phi(2\sqrt{n}\epsilon) \geq 1 - \frac{1 - q_0}{2}$$

- By Chebyshev inequality

$$P\left(|M_n - p| < \epsilon\right) \geq 1 - \frac{1}{4n\epsilon^2} \geq q_0$$

$$n \geq \frac{1}{4(1 - q_0)\epsilon^2}$$

Example (Sample size)

Consider the case of $\epsilon = 0.01$ and $q_0 = 0.95$.

- By Normal approximation of M_n

$$\Phi(2\sqrt{n}\epsilon) \geq 0.975$$

$$2 \cdot \sqrt{n} \cdot 0.01 \geq \Phi^{-1}(0.975) = 1.96$$

$$n \geq 9604$$

- By Chebyshev inequality

$$n \geq \frac{1}{4(1-q_0)\epsilon^2} = 50000$$

Binomial Normal approximation

Let S be **Bin**(n, p).

- S is the sample sum of **iid** samples of **Ber**(p)
- For large n

$$S \dot{\sim} \mathcal{N}(np, np(1-p))$$

Let l be an integer.

$$P(S \leq l) = P\left(\frac{S - np}{\sqrt{np(1-p)}} \leq \frac{l - np}{\sqrt{np(1-p)}}\right) \approx \Phi\left(\frac{l - np}{\sqrt{np(1-p)}}\right)$$

Since S only takes integer values, a better approximation is

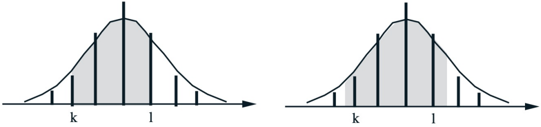
$$P(S \leq l) \approx \Phi\left(\frac{l + \frac{1}{2} - np}{\sqrt{np(1-p)}}\right)$$

De Moivre-Laplace approximation

Let S be **Bin**(n,p) and $k \le l$ be integers.

$$P(k \le S \le l) \approx \Phi\left(\frac{l+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k-\frac{1}{2}-np}{\sqrt{np(1-p)}}\right)$$

$$\begin{aligned} P(k \le S \le l) &= P(S \le l) - P(S < k) \\ &= P(S \le l) - P(S \le k-1) \\ &\approx \Phi\left(\frac{l+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k-\frac{1}{2}-np}{\sqrt{np(1-p)}}\right) \end{aligned}$$



(a) (b)

Example (5.12)

Let S be **Bin**$(36, 0.5)$.

- Applying the De Moivre-Laplace approximation, we get

$$P(S \leq 21) \approx \Phi\left(\frac{21.5 - 18}{3}\right) = 0.879$$

$$\begin{aligned} P(S = 19) &= P(S \leq 19) - P(S \leq 18) \\ &\approx \Phi\left(\frac{19.5 - 18}{3}\right) - \Phi\left(\frac{18.5 - 18}{3}\right) = 0.124 \end{aligned}$$

- The real probabilities are

$$P(S \leq 21) = \sum_{k=0}^{21} \binom{36}{k} (0.5)^{36} = 0.8785$$

$$P(S = 19) = \binom{36}{19} (0.5)^{36} = 0.1251$$

Strong Law of Large Numbers

Strong law of large numbers

Let X be a random variable with $\mathbf{E}[X] = \mu$ and $\text{var}(X) = \sigma^2$. Let $X_1, X_2, \cdots$ denote independent samples of X. Then the sequence of sample means converges almost surely to μ

$$M_n \xrightarrow{\text{a.s.}} \mu$$

That is

$$P\left(\lim_{n\to\infty} M_n = \mu\right) = 1$$

Summary 1

Markov inequality

$$P(X \geq r) \leq \frac{\mathbf{E}[X]}{r}$$

Chebyshev inequality

$$P\left(|X - \mathbf{E}[X]| \geq c\right) \leq \frac{\textsf{var}(X)}{c^2}$$

Summary 2

Convergence in probability $\left(Y_n \xrightarrow{P} Y\right)$

$$\lim_{n\to\infty} P(|Y_n - Y| < \epsilon) = 1, \ \forall \epsilon > 0$$

Convergence in distribution $\left(Y_n \xrightarrow{D} Y\right)$

$$\lim_{n\to\infty} F_{Y_n}(t) = F_Y(t) \text{ for } t \text{ where } F_Y(t) \text{ is continuous}$$

Almost sure convergence $\left(Y_n \xrightarrow{\text{a.s.}} Y\right)$

$$P\left(\lim_{n\to\infty} Y_n = Y\right) = 1$$

Summary 3

The weak law of large numbers

$$\left(M_n \xrightarrow{P} \mu\right)$$

The strong law of large numbers

$$\left(M_n \xrightarrow{\text{a.s.}} \mu\right)$$

The central limit theorem

$$\lim_{n\to\infty} F_{Z_n}(t) = \Phi(t)$$