STA2001 Probability and Statistics (I)

Lecture 18

Tianshi Chen

The Chinese University of Hong Kong, Shenzhen

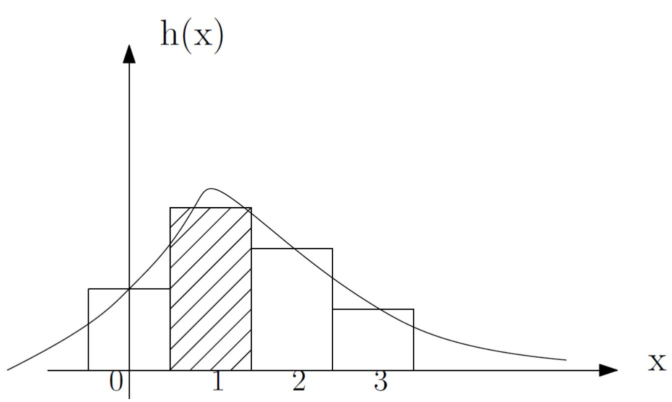
Review of Last Lecture

Histogram of and Approximation for discrete distribution

Consider a discrete RV Y with pmf $f(y) : \overline{S} \to (0, 1]$ with $\overline{S} = \{0, 1, \cdots, n\}$.Then the histogram for Y is

$$h(y) = f(k), y \in (k - \frac{1}{2}, k + \frac{1}{2}), k = 0, 1, \cdots, n$$

The area below the histogram corresponds to probability, which make the histogram has similar property as the pdf of continuous distribution.



If it is possible to find a continuous distribution with pdf "close" to the histogram of the discrete distribution, then we can compute the probability of discrete distribution approximately by using the continuous distribution.

Review of Last Lecture

Half-unit correction for continuity

Now, if $Y = \sum_{i=1}^{n} X_i$, where $X_1, \cdots, X_n$ are i.i.d. random sample drawn from discrete distributions with mean μ and variance σ^2.

$P(Y = k)$	$\approx$	$P(k - \frac{1}{2} < Y < k + \frac{1}{2})$
discrete RV		approximate by continuous RV
pmf $f(y)$		by CLT for large n, Y can be approximated by $N(n\mu, n\sigma^2)$ in the sense that the pdf of the normal distribution is close to the histogram of Y
hard to calculate		easy to calculate

Review of Last Lecture

[Chebyshev's inequality]

If the RV X has a finite mean μ and finite nonzero variance σ^2, then for every $k \geq 1$,

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

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

If $\varepsilon = k\sigma$, then

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

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

Convergence in Probability

$Z_n \xrightarrow{p} Z$

Definition

A sequence of RVs $Z_1, Z_2, \cdots$, is said to converge in probability to a RV Z, often denoted by, $Z_n \xrightarrow{p} Z$, if for any $\varepsilon > 0$,

$$\lim_{n\to\infty} P(|Z_n - Z| \geq \varepsilon) = 0.$$

Example 1

Assume that Z_n has an exponential distribution with $\theta = 1/n$, $n = 1, 2, \cdots$. Then show that $Z_n \xrightarrow{p} 0$, i.e., the sequence of RVs $Z_1, Z_2, \cdots$, converges in probability to $Z = 0$.

$\theta = \frac{1}{n}$

Example 1

Assume that Z_n has an exponential distribution with $\theta = 1/n$, $n = 1, 2, \cdots$. Then show that $Z_n \xrightarrow{p} 0$, i.e., the sequence of RVs $Z_1, Z_2, \cdots$, converges in probability to $Z = 0$.

For any $\epsilon > 0$,

$$
\begin{aligned}
\lim_{n\to\infty} P\big(|Z_n - 0| \geq \epsilon\big) &= \lim_{n\to\infty} P\big(Z_n \geq \epsilon\big) \\
&= \lim_{n\to\infty} e^{-n\epsilon} \\
&= 0,
\end{aligned}
$$

where the first equation is true because $Z_n \geq 0$, and the second equation is obtained by using the fact that Z_n has an exponential distribution with $\theta = 1/n$.

Theorem [Law of Large Number]

cvg in distribution $\neq$ cvg in probability

Theorem (Law of Large Numbers)

Let $X_1, X_2, \cdots, X_n$ be a random sample of size n drawn from a distribution with finite mean μ and finite nonzero variance, and let $\overline{X}$ be the sample mean. Then $\overline{X}$ converges in probability to μ, i.e., for any $\varepsilon > 0$,

大数定理

$$\lim_{n\to\infty} P\left(\left|\overline{X} - \mu\right| \geq \varepsilon\right) = 0$$

$$\lim_{n\to\infty} P(|\overline{X}-\mu| \geq \varepsilon) = 0$$

Proof of Law of Large Number

Note that

$$E(\overline{X}) = \mu, \quad Var(\overline{X}) = \frac{1}{n}\sigma^2$$

By the Chebyshev's inequality, i.e., Corollary 5.8-1, for every $\varepsilon > 0$.

$$P(|\overline{X} - \mu| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}\frac{\sigma^2}{n}$$

Proof of Law of Large Number

Taking limits on both sides yield

$$0 \leq \lim_{n\to\infty} P(|\overline{X} - \mu| \geq \varepsilon) \leq \lim_{n\to\infty} \frac{\sigma^2}{\varepsilon^2 n} = 0$$

$$\Rightarrow \lim_{n\to\infty} P(|\overline{X} - \mu| \geq \varepsilon) = 0$$

or equivalently $\lim_{n\to\infty} P(|\overline{X} - \mu| < \varepsilon) = 1$

Section 5.9 Limiting Moment Generating Functions

Motivation

Binomial distribution $b(n, p)$ can be approximated by the Poisson distribution with $\lambda = np$ when n is large and p is fairly small:

- the approximation is good if $n \geq 20$ and $p \leq 0.05$
- the approximation is very good if $n \geq 100$ and $p \leq 0.1$
- the approximation becomes better with larger n and smaller p

Example 1, page 227

Question

Let $Y \sim b(50, 1/25)$. Q: $P(Y \leq 1)$?

Example 1, page 227

Question

Let $Y \sim b(50, 1/25)$. Q: $P(Y \leq 1)$?

1. By definition,

$$P(Y \leq 1) = P(Y = 0) + P(Y = 1)$$
$$= \left(\frac{24}{25}\right)^{50} + 50\left(\frac{1}{25}\right)\left(\frac{24}{25}\right)^{49} = 0.4$$

2. By approximation with Poisson distribution $\lambda = np = 2$

$$P(Y \leq 1) \approx \frac{2^0 e^{-2}}{0!} + \frac{2^1 e^{-2}}{1!} = 3e^{-2} = 0.406$$

Why? and an Interesting Observation

First, recall the mgf of $b(n, p)$ is $M(t) = (1 - p + pe^t)^n$.

Then, we will consider the limit of $M(t) = (1 - p + pe^t)^n$ as $n \to \infty$ such that $np = \lambda$ is a constant.

$$M(t) = (1 - \frac{\lambda}{n} + \frac{\lambda}{n}e^t)^n = [1 + \frac{\lambda(e^t - 1)}{n}]^n$$

$e^{\lim_{n\to\infty} [f(x)-1]\, g(x)}$

Since $\lim_{n\to\infty}(1 + \frac{b}{n})^n = e^b$

$= e^b$

$$\lim_{n\to\infty} M(t) = e^{\lambda(e^t-1)} \to \text{mgf for Poisson distribution}$$

Theorem 5.9-1, page 226

Limiting mgf technique

Let $\{M_n(t)\}_{n=1}^{\infty}$ be a sequence of mgfs for t in an open interval around $t = 0$. If $\lim_{n\to\infty} M_n(t) = M(t)$, for t in the open interval around $t = 0$. Then the sequence of RVs

$$Z_n \xrightarrow{d} Z,$$

cvg in distribution

where $M_n(t)$ and $M(t)$ are mgfs of Z_n and Z, respectively.

$$\boxed{\begin{array}{c} \lim_{n\to\infty} M_n(t) = M(t) \\ \updownarrow \qquad\qquad \updownarrow \\ \text{distribution} \\ Z_n \xrightarrow{d} Z \end{array}}$$

Convergence of $b(n, p)$

$b(n, p)$ converges in distribution to

- a Poisson distribution according to the limiting mgf technique

Convergence of $b(n, p)$

$b(n, p)$ converges in distribution to

- a Poisson distribution according to the limiting mgf technique
- a standard normal distribution in some sense according to CLT

How to understand?

Convergence of $b(n, p)$

$b(n, p)$ converges in distribution to

- a Poisson distribution according to the limiting mgf technique
- a standard normal distribution in some sense according to CLT

How to understand?

Convergence of $b(n, p)$

1. as $n \to \infty$ with $\lambda = np$ being a constant, and let $Z_n \sim b(n, p)$ and $Z \sim \text{Poisson}(\lambda)$, then $Z_n \xrightarrow{d} Z$.

   MGF

2. as $n \to \infty$ with p being a constant, and let $Z_n \sim b(n, p)$ and $Z \sim N(0, 1)$, then

$$\frac{Z_n/n - p}{\sqrt{p(1-p)/n}} \xrightarrow{d} Z$$

CLT

Proof of CLT

CLT

Let $\overline{X}$ be the sample mean of the random sample of size n, $X_1, X_2, \cdots, X_n$ from a distribution with a finite mean μ and a finite nonzero variance σ^2, then as $n \to \infty$, the random variable $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ converge in distribution to $N(0,1)$.

The idea of the proof:

1. Let
$$W_n = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}},$$
and then show the mgf of W_n, say $M_n(t)$, converges to the mgf of $N(0,1)$ for t in an open interval around $t = 0$
2. By Theorem 5.9-1, $W_n \xrightarrow{d} N(0,1)$

Proof of CLT, page 208

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

[independence]

Proof of CLT, page 208

Let

$$Z_i = \frac{X_i - \mu}{\sigma}, \quad i = 1, \cdots, n$$

Then $Z_1, \cdots, Z_n$ are i.i.d..

Let

$$M(t) = E\left\{\exp\left(tZ_i\right)\right\}, \quad |t| < h$$

be the common mgf for $Z_i, i = 1, \cdots, n$.

Then

$$E[e^{tW_n}] = \left[M\left(\frac{t}{\sqrt{n}}\right)\right]^n, \quad \left|\frac{t}{\sqrt{n}}\right| < h.$$

Proof of CLT, page 208

Now consider $M(t)$. Actually, $Z_1, \ldots, Z_n$ are i.i.d. with mean 0 and variance 1. Then

$$M(0) = 1, M'(0) = 0, M''(0) = 1.$$

By using Taylor's expansion, there exists $|t_1| \leq |t|$ such that

$$\begin{aligned} M(t) &= M(0) + M'(0)t + \frac{1}{2}M''(t_1)t^2 \\ &= 1 + \frac{1}{2}M''(t_1)t^2 = 1 + \frac{1}{2}t^2 + \frac{1}{2}t^2[M''(t_1) - 1]. \end{aligned}$$

Proof of CLT, page 208

Then

$$E(e^{tW_n}) = \left[M(\frac{t}{\sqrt{n}})\right]^n = \left[1 + \frac{1}{2}\frac{t^2}{n} + \frac{1}{2}\frac{t^2}{n}[M''(t_1) - 1]\right]^n,$$

$$\left|\frac{t}{\sqrt{n}}\right| < h, \quad |t_1| \leq \frac{|t|}{\sqrt{n}}.$$

Since $M''(t)$ is continuous at $t = 0$ and as $n \to \infty$, $t_1 \to 0$,

$$\lim_{n\to\infty} M''(t_1) - 1 = 1 - 1 = 0.$$

Note that

$$\lim_{n\to\infty} \left(1 + \frac{b}{n}\right)^n = e^b$$

Proof of CLT, page 208

we have

$$\lim_{n\to\infty} E(e^{tW_n}) = \lim_{n\to\infty}\left[1+\frac{1}{2}\frac{t^2}{n}+\frac{1}{2}\frac{t^2}{n}[M''(t_1)-1]\right]^n$$

$$= e^{\frac{t^2}{2}} \to \text{mgf of } N(0,1)$$

By Theorem 5.9-7,

$$W_n = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1)$$

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