Fall 2018 STSCI 5080 Discussion 8 (10/19)

Normal approximation to Poisson

Theorem 1 (Continuity theorem for mgfs)**.** *Let X_n and X have mgfs ψ_n and ψ, respectively. If $\psi_n(\theta) \to \psi(\theta)$ for any θ in an open interval containing the origin, then $X_n \xrightarrow{d} X$.*

Theorem 2. *Let $X_n \sim Po(\lambda_n)$ and $\lambda_n \to \infty$ as $n \to \infty$. Then*

$$\frac{X_n - \lambda_n}{\sqrt{\lambda_n}} \xrightarrow{d} N(0,1).$$

Proof. Recall that $E(X_n) = \lambda_n$ and $\mathrm{Var}(X_n) = \lambda_n$. The mgf of X_n is

$$\psi_{X_n}(\theta) = e^{\lambda_n(e^\theta - 1)}.$$

Hence the mgf of $Y_n = (X_n - \lambda_n)/\sqrt{\lambda_n}$ is

$$\begin{aligned}\psi_{Y_n}(\theta) = E(e^{\theta Y_n}) = E(e^{\theta(X_n - \lambda_n)/\sqrt{\lambda_n}}) &= e^{-\theta\sqrt{\lambda_n}}E(e^{\theta X_n/\sqrt{\lambda_n}}) \\ &= e^{-\theta\sqrt{\lambda_n}}\psi_{X_n}(\theta/\sqrt{\lambda_n}) \\ &= e^{-\theta/\sqrt{\lambda_n}}e^{\lambda_n(e^{\theta/\sqrt{\lambda_n}}-1)}.\end{aligned}$$

By Taylor's theorem, we can expand e^x as

$$e^x = 1 + x + \frac{x^2}{2} + x^2R(x),$$

where $\lim_{x\to 0} R(x) = 0$. Hence

$$\lambda_n(e^{\theta/\sqrt{\lambda_n}} - 1) = \theta\sqrt{\lambda_n} + \frac{\theta^2}{2} + \theta^2R(\theta/\sqrt{\lambda_n}),$$

so that

$$e^{-\theta/\sqrt{\lambda_n}}e^{\lambda_n(e^{\theta/\sqrt{\lambda_n}}-1)} = e^{-\theta^2/2} + \theta^2R(\theta/\sqrt{\lambda_n}).$$

Since $\lim_{n\to\infty} R(\theta/\sqrt{\lambda_n}) = 0$, we conclude that

$$\lim_{n\to\infty}\psi_{Y_n}(\theta) = e^{-\theta^2/2},$$

which is the mgf of $N(0,1)$. By the continuity theorem, we have $Y_n \xrightarrow{d} N(0,1)$. □

Problems

1. Show that if $E(|X_n|) \to 0$, then $X_n \xrightarrow{P} 0$.
2. Let X_n be such that

   $$P(X_n = n) = \frac{1}{n} \quad \text{and} \quad P(X_n = 0) = 1 - \frac{1}{n}.$$

   Show that $E(X_n) = 1$ but $X_n \xrightarrow{P} 0$. So convergence in probability does not imply moment convergence.

3. (**Rice 5.4.1**) Let $X_1, X_2, \ldots$ be a sequence of independent random variables with $E(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma_i^2$. Show that if $n^{-2}\sum_{i=1}^{n}\sigma_i^2 \to 0$, then $\overline{X}_n \xrightarrow{P} \mu$.

4. (**Rice 5.4.3**) Suppose that the number of insurance claims, N, filed in a year is Poisson distributed with $E(N) = 10{,}000$. Use the normal approximation to the Poisson to approximate $P(N > 10{,}200)$.

5. (**Rice 5.4.5**) Using mgfs, show that as $n \to \infty, p \to 0$, and $np \to \lambda$, the binomial distribution with parameters n and p tends to the Poisson distribution.

6. Let $X_1, \ldots, X_n \sim U[0,1]$ i.i.d. Show that $nX_{(1)} \xrightarrow{d} Ex(1)$ by directly evaluating the cdf of $X_{(1)}$.

7. (**Rice 5.4.10**) A six-sided die is rolled 100 times. Using the normal approximation, find the probability that the face showing a six turns up between 15 and 20 times. Find the probability that the sum of the face values of the 100 trials is less than 300.

8. (**Rice 5.4.16**) Suppose that $X_1, \ldots, X_{20}$ are independent random variables with density function
$$f(x) = 2 \quad \text{if } 0 \le x \le 1.$$
Let $S = X_1 + \cdots + X_{20}$. Use the central limit theorem to approximate $P(S \le 10)$.

Solutions

1. By Markov's inequality,
$$P(|X_n| > \varepsilon) \le \frac{E(|X_n|)}{\varepsilon} \to 0$$
as $n \to \infty$. This implies that $X_n \xrightarrow{P} 0$.

2. We have
$$E(X_n) = 0 \cdot \left(1 - \frac{1}{n}\right) + n \cdot \frac{1}{n} = 1.$$
In addition, for any $\varepsilon > 0$,
$$P(|X_n| > \varepsilon) = P(X_n = n) = \frac{1}{n}$$
for sufficiently large n. Hence
$$\lim_{n\to\infty} P(|X_n| > \varepsilon) = 0,$$
which implies that $X_n \xrightarrow{P} 0$.

3. (**Rice 5.4.1**) By Chebyshev's inequality, we have
$$P(|\overline{X}_n - \mu| > \varepsilon) \le \frac{\mathrm{Var}(\overline{X}_n)}{\varepsilon^2} = \frac{n^{-2}\sum_{i=1}^n \sigma_i^2}{\varepsilon^2}.$$
Since $n^{-2}\sum_{i=1}^n \sigma_i^2 \to 0$ by assumption, we have $\overline{X}_n \xrightarrow{P} \mu$.

4. (**Rice 5.4.3**) $N \sim Po(\lambda)$ with $\lambda = 10000$. So
$$P(N > 10200) = P\left(\frac{N - 10000}{100} > \frac{10200 - 10000}{100}\right) \approx 1 - \Phi(2) \approx 0.023.$$

5. (**Rice 5.4.5**) The mgf of $X_n \sim Bin(n,p)$ is
$$\psi_n(\theta) = \{1 + p(e^\theta - 1)\}^n.$$
Let $\lambda_n = np$ so that $p = \lambda_n/n$. Then
$$\psi_n(\theta) = \left\{1 + \frac{\lambda_n}{n}(e^\theta - 1)\right\}^n.$$
Since $\lambda_n \to \lambda$, the RHS $\to e^{\lambda(e^\theta - 1)}$, which is the mgf of $Po(\lambda)$. By the continuity theorem, we have $X_n \xrightarrow{d} Po(\lambda)$.

6. The cdf of $X_{(1)}$ is
$$P(X_{(1)} \le x) = 1 - (1-x)^n$$
for $0 \le x \le 1$. Hence, for $x \ge 0$,
$$P(nX_{(1)} \le x) = P(X_{(1)} \le x/n) = 1 - (1 - x/n)^n \to 1 - e^{-x}.$$
If $x < 0$, then $P(nX_{(1)} \le x) = 0$, so that we have $nX_{(1)} \xrightarrow{d} Ex(1)$.

7. (**Rice 5.4.10**) Let X be the number that the face shows a six; then $X \sim Bin(100, 1/6)$. We want to find $P(15 < X < 20)$. We have $E(X) = 100/6 = 50/3$ and $\mathrm{Var}(X) = 100 \cdot 5/36 = 125/9$. Hence,

$$P(15 < X < 20) = P\left(\frac{15 - 50/3}{125/9} < \frac{X - 50/3}{125/9} < \frac{20 - 50/3}{125/9}\right)$$
$$\approx \Phi(0.24) - \Phi(-0.12) = \Phi(0.24) + \Phi(0.12) - 1 \approx 0.182.$$

Next, let Y_i denote the face value of the i-th trial, and we want to find $P(S \leq 300)$ where $S = \sum_{i=1}^{100} Y_i$. The mean and variance of Y_i are

$$E(Y_i) = \sum_{j=1}^{6} \frac{j}{6} = \frac{7}{2} \quad \text{and} \quad \mathrm{Var}(Y_i) = E(Y_i^2) - \{E(Y_i)\}^2 = \sum_{j=1}^{6} \frac{j^2}{6} - \frac{49}{4} = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}.$$

Hence,

$$P(S \leq 300) = P\left(\frac{S - 350}{\sqrt{875/3}} \leq \frac{300 - 350}{\sqrt{875/3}}\right) \approx \Phi(-2.93) \approx 0.002.$$

8. (**Rice 5.4.16**) The mean and variance of X_i are

$$E(X_i) = \int_0^1 2x^2 dx = \frac{2}{3} \quad \text{and} \quad \mathrm{Var}(X_i) = E(X_i^2) - \{E(X_i)\}^2 = \int_0^1 2x^3 dx - \frac{4}{9} = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}.$$

So,

$$P(S \leq 10) = P\left(\frac{S - 40/3}{\sqrt{10/9}} \leq \frac{10 - 40/3}{\sqrt{10/9}}\right) \approx \Phi(-3.16) \approx 0.001.$$