PSTAT 120A, Summer 2022: Practice Problems 9

Week 6

Conceptual Review

(a) What is the definition of convergence in distribution? What about convergence in probability?

(b) What is the Law of Large Numbers?

(c) What is the Central Limit Theorem? How does it relate to the DeMoivre-Laplace Theorem?

Problem 1: Poisson Predicament

Provide a probabilistic proof for the identity

Hint: Let $X_i \overset{\text{i.i.d.}}{\sim} \text{Pois}(1)$, and $X := \sum_{i=1}^n X_i$

$$\sum_{k=0}^{c} e^{-n} \cdot \frac{n^k}{k!} \approx \int_{-\infty}^{\frac{c-n}{\sqrt{n}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \, \mathrm{d}z$$

assuming n is quite large.

Solution: Following the hint, we let $X_i \overset{\text{i.i.d.}}{\sim} \text{Pois}(1)$ and so

$$X := \left(\sum_{i=1}^{n} X_i \right) \sim \text{Pois}(n)$$

Therefore, we can see that the sum on the LHS of the equation we wish to prove is

$$\mathbb{P}(X \leq c) = \sum_{k=0}^{c} \mathbb{P}(X = k) = \sum_{k=0}^{c} e^{-n} \cdot \frac{n^k}{k!}$$

Now, the CLT tells us that

$$X \overset{d}{\approx} \mathcal{N}(n, n)$$

meaning

$$\mathbb{P}(X \leq c) = \mathbb{P}\left(\frac{X-n}{\sqrt{n}} \leq \frac{c-n}{\sqrt{n}} \right) \approx \Phi\left(\frac{c-n}{\sqrt{n}} \right) = \int_{-\infty}^{\frac{c-n}{\sqrt{n}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \, \mathrm{d}z$$

Therefore, the string of logic goes:

$$\mathbb{P}(X \leq c) = \sum_{k=0}^{c} \mathbb{P}(X = k) = \sum_{k=0}^{c} e^{-n} \cdot \frac{n^k}{k!} \approx \int_{-\infty}^{\frac{c-n}{\sqrt{n}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \, \mathrm{d}z$$

thereby completing the proof. ■

Problem 2: Money, Money, Money

Every day, Jack goes to the casino. Suppose that on any given day, the amount Jack earns is a random variable with mean \$10 and standard deviation \$2, independent of all other days. Approximate the probability that Jack will earn more than \$310 in a 30-day period.

Solution: Let X_i denote the amount Jack earns on day i, for $i = 1, \cdots, 30$; then $\mathbb{E}[X_i] = 10$ and $\text{Var}(X_i) = 4$. If X denotes Jack's total earnings in a 30-day period, then

$$X := \sum_{i=1}^{30} X_i$$

and, by the Central Limit Theorem,

$$X \overset{\text{d}}{\approx} \mathcal{N}(300, 120)$$

meaning

$$\begin{aligned}\mathbb{P}(X \geq 310) &= \mathbb{P}\left(\frac{X - 300}{\sqrt{120}} \geq \frac{310 - 300}{\sqrt{120}} \right) \\ &= 1 - \mathbb{P}\left(\frac{X - 300}{\sqrt{120}} \leq \frac{10}{\sqrt{120}} \right) = 1 - \Phi\left(\frac{10}{\sqrt{120}} \right) \approx 18\%\end{aligned}$$

Problem 3: Probabilistic Convergence

Prove the following theorem: if X_n is a random variable with mean μ that satisfies the property $\text{Var}(X_n) \to 0$ as $n \to \infty$, then $X_n \xrightarrow{p} \mu$.

Hint: Chebyshev's Inequality

Solution: By Chebyshev's Inequality, for any fixed $\varepsilon > 0$

$$\mathbb{P}(|X_n - \mu| \geq \varepsilon) \leq \frac{\text{Var}(X_n)}{\varepsilon^2}$$

Additionally, since probabilities are definitionally nonnegative, we have

$$0 \leq \mathbb{P}(|X_n - \mu| \geq \varepsilon) \leq \frac{\text{Var}(X_n)}{\varepsilon^2}$$

If $\text{Var}(X_n) \to 0$ as $n \to \infty$, the upper bound above goes to 0 as $n \to \infty$ meaning by the Squeeze Theorem, $\mathbb{P}(|X - \mu| \geq \varepsilon) \to 0$ which is precisely the definition of what it means for X_n to converge in probability to μ. ■

Extra Problems

Problem 4: Random Numbers *(modified from ASV, 4.16)*

We choose 500 numbers uniformly at random from the interval [1.5, 4.8]

(a) Approximate the probability of the event that less than 65 of the numbers start with the digit 1, and justify your choice of approximation.

Solution: Let S denote the number of numbers selected that begin with the digit 1; then $S \sim \text{Bin}(500, p)$ where p denotes the probability that any randomly selected number begins with the digit 1.

We must now compute p. To that end, let U denote the result of selecting a number uniformly from the interval $[1.5, 4.8]$; then $U \sim \text{Unif}[1.5, 4.8]$. Most notably, though, is that

$$p = \mathbb{P}(U < 2) = \int_{1.5}^{2} \frac{1}{4.8 - 1.5} \, \mathrm{d}u = \frac{2 - 1.5}{3.3} = \frac{5}{33}$$

since a number selected from the interval $[1.5, 4.8)$ begins with the digit 1 if and only if it is selected from the interval $[1.5, 2)$.

Thus, $S \sim \text{Bin}(500, 5/33)$. We wish to approximate $\mathbb{P}(S < 65)$; we shall utilize the DeMoivre-Laplace theorem to apply the normal approximation. First, though, we should check that the conditions for the normal approximation to be valid hold:

(1) $n = 500 \geq 100$ ✓

(2) $p = 5/33 \approx 0.15$ is not too extreme

(3) $np(1 - p) = 500(5/33)(28/33) = 64.27916 \geq 10$ ✓

Thus, the normal approximation is valid and we may write

$$\begin{aligned} S &\overset{\mathrm{d}}{\approx} \mathcal{N}\big(np, np(1-p)\big) \\ &\overset{\mathrm{d}}{=} \mathcal{N}(75.7576, 64.27916) \end{aligned}$$

and thus,

$$\mathbb{P}(S < 65) = \mathbb{P}\left(\frac{S - 75.7576}{\sqrt{64.27916}} \leq \frac{65 - 75.7576}{\sqrt{64.27916}} \right) \approx \Phi(-1.34) = 1 - \Phi(1.34) \approx \boxed{0.0901}$$

(b) Approximate the probability of the event that more than 160 of the numbers start with the digit 3, and justify your choice of approximation.

Solution: We proceed similarly as in part (a). The probability that a given uniformly chosen number from [1.5, 4.8] starts with 3 is $q = \frac{1}{3.3} = \frac{10}{33}$. If we denote the number of such numbers among the 500 random numbers by T then $T \sim \text{Bin}(n, q)$ with $n = 500$.

Then

$$\mathbb{P}(T > 160) = \mathbb{P}\left(\frac{T - nq}{\sqrt{nq(1-q)}} > \frac{160 - nq}{\sqrt{nq(1-q)}} \right) \approx \mathbb{P}\left(\frac{T - nq}{\sqrt{nq(1-q)}} > 0.83 \right)$$
$$\approx 1 - \Phi(0.83) \approx 1 - 0.7967 = 0.2033$$