HW#2 Part 1 Short Solution

ISI Spring 2024

Mahshid Dashtianeh – Reyhane Vahedi

Q1

According to CLT, the distribution of the sample mean of a large enough number of independent and identically distributed (i.i.d) random variables approaches a normal distribution.

We want to find n such that $P(|\bar{x} - \mu| < 1) = 0.95$.

Since we know that the standard deviation of the sample mean ($\bar{x}$) is $\frac{\sigma}{\sqrt{n}}$ we can rewrite the condition as

$$P\left(\left|\frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}\right| < \frac{1}{\frac{\sigma}{\sqrt{n}}}\right) = 0.95$$

Using the properties of the standard normal distribution, we find the corresponding z-scores for the probability 0.95.

$$\frac{1}{\frac{\sigma}{\sqrt{n}}} = 1.96 \rightarrow n = (5 \times 1.96)^2 \approx 96$$

Q2

---

the CDF of $U(n)$ is the probability that the maximum of n variables is less than or equal to u, which is given by the product of the probabilities that each Ui is less than or equal to u:

$F_{U_n}(u) = \left(F(u)\right)^n = u^n$ for 0<u<1

Let $Z_n = n(U_{(n)} - 1)$.

$$F_{Z_n}(z) = P(Z_n \leq z) = P(n(U_n - 1) \leq z) = P\left(U_n \leq \frac{z}{n} + 1\right) = (\frac{z}{n} + 1)^{\ n}$$

$F_{Z_n}(z) \rightarrow e_z \ \ as\, n \rightarrow \infty$ for $-1 \leq z \leq 0$

Q3

---

a)

$All\ products\ made\ in\ n\ days: S_n = X_1 + X_2 + \cdots + X_n$

$E[S_n] = 5n$

$Var[S_n] = 9n$

$P(S_{100} < 440) = \ P(S_{100} < 439.5) = P\left(\frac{S_{100} - 500}{30} \leq \frac{439.5 - 500}{30}\right) \approx \ \emptyset(-2.02) = 1 - \ \emptyset(2.02)$
$= 0.0217$

b)

$P(S_n \geq 200 + 5n) \ \leq 0.05 \ \rightarrow \ P\left(\frac{S_n - 5n}{3\sqrt{n}} \geq \frac{200}{3\sqrt{n}}\right) \leq 0.05$

With normal approximation we have:

$1 - \ \emptyset\left(\frac{200}{3\sqrt{n}}\right) \ \leq 0.05 \ \rightarrow \ \emptyset\left(\frac{200}{3\sqrt{n}}\right) \geq 0.95$

From the table we have: $\emptyset(1.65) \approx 0.95$

$\frac{200}{3\sqrt{n}} \geq 1.65 \ \rightarrow n \leq 1632$

c)

$P(N \geq 220) = \ P(S_{219} \leq 1000) = P\left(\frac{S_{219} - 5(219)}{3\sqrt{219}} \leq \frac{1000 - 5(219)}{3\sqrt{219}}\right) = 1 - \ \emptyset(2.14) = 0.0162$

Q4

Because the die is balanced, $E[X_i] = 7/2$. By the law of large numbers for any $\varepsilon > 0$:

$$P(\left|\frac{S_n}{n} - \frac{7}{2}\right| \geq \varepsilon) \rightarrow 0$$

as n ⟶ ∞ or equivalently:

$$P(\left|\frac{S_n}{n} - \frac{7}{2}\right| < \varepsilon) \rightarrow 1$$

as n ⟶ ∞.

Q5

---

a) Because $U_i's$ are uniform in (-0.5,0.5): $\sum U_i = 0 \, , Var(U_i) = 1/12$

Let $S_n = \sum_{i=1}^{n} X_i$ and $K_n = \sum_{i=1}^{n} J_i$

Then:

$P(|S_n - K_n| \le a) = P\left(-a \le \sum (X_i - J_i) \le a\right) = P\left(-a \le \sum U_i \le a\right)$

By the CLT $\frac{\sum_{i=1}^{n} U_i - 0}{\sqrt{n/12}} \sim N(0,1)$ as n → ∞.

b) For n = 300; a=5. Using the normal table:

$$P\left\{\frac{-5}{\sqrt{300/12}} \le \frac{\sum_{i=1}^{n} U_i}{\sqrt{300/12}} \le \frac{5}{\sqrt{300/12}}\right\} = 0.68$$

Q6

Let X_i denote the success of *i*-th customer. Then each X_i follows *Bernoulli* distribution with probability 0.03, and $E(X_i) = 0.03$ and $Var(X_i) = 0.0291$. Let $S_{2500} = \sum_{i=1}^{2500} X_i$. From the CLT,

$\frac{S_{2500} - 2500 \cdot 0.03}{\sqrt{0.0291 \cdot 2500}}$ follows approximately $N(0,1)$. Hence, we have

$$P(S_{2500} \ge 80) = P\left(\frac{S_{2500} - 2500 \cdot 0.03}{\sqrt{0.0291 \cdot 2500}} \ge \frac{80 - 2500 \cdot 0.03}{\sqrt{0.0291 \cdot 2500}}\right) \approx P(Z \ge 0.586) = 0.279.$$

Q7

Let $X_i = 1$ if the *i*th person in the sample is color blind and 0 otherwise. Then each X_i follows *Bernoulli* distribution with estimated probability 0.02, and $E(X_i) = 0.02$ and $Var(X_i) = 0.0196$. Let $S_n = \sum_{i=1}^{n} X_i$. We want $P(S_n \ge 1) = 0.99$. By the CLT, $\frac{S_n / n - 0.02}{\sqrt{0.0196 / n}}$ follows approximately $N(0,1)$. Then:

$$0.99 = P(S_n \ge 1) \approx P\left(Z \ge \frac{1/n - 0.02}{\sqrt{0.0196/n}}\right).$$

Using the normal table, $\frac{1/n - 0.02}{\sqrt{0.0196/n}} = -2.33$. Solving this equation, we have $n = 359.05$. Thus, the sample size must be at least 360.

Q8

a)

$\lambda_1.\lambda_2.\dots.\lambda_n \qquad \lambda_n \to \infty$

$\{X_n\} \sim Poisson \qquad E[X_n] = Var[X_n] = \lambda_n$

If we want to approximate the Poisson distribution with Normal distribution, their means and variances should be the same. Standardization:

$$Z_n = \frac{X_n - E[X_n]}{\sqrt{Var[X_n]}} = \frac{X_n - \lambda_n}{\sqrt{\lambda_n}} \sim N(0.1)$$

$$M_{X_n}(t) = e^{\lambda_n(1-e^{-t})}$$

$$M_{Z_n}(t) = e^{-t\sqrt{\lambda_n}} M_{X_n}\left(\frac{t}{\sqrt{\lambda_n}}\right) = e^{-t\sqrt{\lambda_n}} e^{\lambda_n(e^{\frac{t}{\sqrt{\lambda_n}}}-1)}$$

$$logM_{Z_n}(t) = -t\sqrt{\lambda_n} + \lambda_n(e^{\frac{t}{\sqrt{\lambda_n}}} - 1)$$

With Taylor expansion of e^x:

$$\lim_{n\to\infty} logM_{Z_n}(t) = \frac{t^2}{2} \to \lim_{n\to\infty} M_{Z_n}(t) = e^{\frac{t^2}{2}}$$

As expected, the moment generating function in the standard normal distribution is also as follows.

b)

$$P(X > 950) = P\left(\frac{X-900}{\sqrt{900}} > \frac{950-900}{\sqrt{900}}\right) \approx 1 - \emptyset\left(\frac{5}{3}\right) = 0.04779$$