M378K: October 17th, 2025.

Problem 11.2. *A new addition of Kafka's "Metamorphosis" has 72 pages. The printing press often malfunctions and introduces typos. The number of typos on each page has a Poisson distribution with mean* $\ln(3)$ *and is independent of the number of typos on other pages (or other books). A book is thrown away if it contains typos on more than 32 pages. Use the normal approximation to estimate the proportion of books that get thrown away.*

$e^{-\lambda} \cdot \frac{\lambda^k}{k!}$

$k = 0, 1, \dots$

$\longrightarrow$: p... probab. that @ least one typo on page

q... a page has no typos

$q = \mathbb{P}[\text{Poisson w/ mean } \ln(3) \text{ is equal to } 0]$

$q = e^{-\ln(3)} = \frac{1}{3} \Rightarrow \boxed{p = \frac{2}{3}}$

Y... # of pages w/ typos

$Y \sim b(72, \frac{2}{3})$

$\mathbb{P}[Y \geq 33] = \mathbb{P}[Y > 32.5]$

$\mathbb{E}[Y] = 72 \cdot \frac{2}{3} = 48;$

$Var[Y] = 72 \cdot \frac{2}{3} \cdot \frac{1}{3} = 16 \Rightarrow SD[Y] = 4$

$\mathbb{P}[Y > 32.5] = \mathbb{P}\left[\underbrace{\frac{Y-48}{4}}_{\approx N(0,1) \sim Z} > \frac{32.5-48}{4}\right]$

$= \mathbb{P}\left[Z > -\frac{15.5}{4}\right] = \mathbb{P}[Z > -3.875]$

≈ 1 □

M378K Introduction to Mathematical Statistics
Problem Set #12
The Central Limit Theorem (CLT).

Let $\{X_n, n = 1, 2, 3, \dots\}$ be a sequence of independent, identically distributed random variables such that $\mu_X = \mathbb{E}[X_1] < \infty$ and $Var[X] = \sigma_X^2 < \infty$. For every $n = 1, 2, \dots$ define

$$S_n = X_1 + X_2 + \cdots + X_n$$

and

$$\bar{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

Problem 12.1. *Find the expected value of S_n and $\bar{X}_n$ for every n.*

$\mathbb{E}[S_n] = n \cdot \mu_X$

$\boxed{\mathbb{E}[\bar{X}_n] = \mu_X}$ $\boxed{\text{accuracy}}$

Problem 12.2. *Find the variance and standard deviation of S_n and $\bar{X}_n$ for every n.*

$Var[S_n] = Var[X_1 + X_2 + \dots + X_n]$ (independence)

$= Var[X_1] + Var[X_2] + \dots + Var[X_n] =$ (i.d.)

$= n \cdot Var[X_1] = n \cdot \sigma_X^2 \Rightarrow SD[S_n] = \sigma_X \sqrt{n}$

$Var[\bar{X}_n] = \frac{\sigma_X^2}{n} \Rightarrow \boxed{SD[\bar{X}_n] = \frac{\sigma}{\sqrt{n}}}$ precision

Theorem 12.1. **The Central Limit Theorem (CLT).** *If the above conditions are satisfied, we have that*

$$\frac{S_n - n\mu_X}{\sigma_X \sqrt{n}} = \frac{\bar{X}_n - \mu_X}{\frac{\sigma_X}{\sqrt{n}}} \xrightarrow{\mathcal{D}} N(0,1) \quad \text{as } n \to \infty.$$

Practically, for "large enough" n, $\bar{X}_n$ is approximately normal with mean μ_X and variance $\frac{\sigma_X^2}{n}$. The rule of thumb is that we use the theorem for $n \geq 30$. If that is the case, we have that for any real $l < r$,

$$\mathbb{P}[l < S_n \leq r] = \mathbb{P}\left[\frac{l - n\mu_X}{\sigma_X\sqrt{n}} < \frac{S_n - n\mu_X}{\sigma_X\sqrt{n}} \leq \frac{r - n\mu_X}{\sigma_X\sqrt{n}}\right] \approx \Phi\left(\frac{r - n\mu_X}{\sigma_X\sqrt{n}}\right) - \Phi\left(\frac{l - n\mu_X}{\sigma_X\sqrt{n}}\right).$$

Similarly, for any real $a < b$,

$$\mathbb{P}[a < \bar{X}_n \le b] = \mathbb{P}\left[\frac{a-\mu_X}{\frac{\sigma_X}{\sqrt{n}}} < \frac{\bar{X}_n-\mu_X}{\frac{\sigma_X}{\sqrt{n}}} \le \frac{b-\mu_X}{\frac{\sigma_X}{\sqrt{n}}}\right] \approx \Phi\left(\frac{b-\mu_X}{\frac{\sigma_X}{\sqrt{n}}}\right) - \Phi\left(\frac{a-\mu_X}{\frac{\sigma_X}{\sqrt{n}}}\right).$$

Problem 12.3. *The* Really Terrible Orchestra[1] *plans a concert at a gazebo in a local park. The orchestra has* 169 *members whose weights are assumed to be independent and identically distributed with mean* 100 *kilos and standard deviation of* 10 *kilos (the weight of the instruments is taken into account here). The gazebo can safely support up to* 17 *tons (each ton is* 1000 *kilos). What is the approximate probability that the gazebo will collapse?*

→: $S = Y_1 + Y_2 + \cdots + Y_n$ w/ Y_i has mean $\mu_Y = 100$ and $\sigma_Y = 10$

$\mathbb{E}[S] = 169 \cdot 100 = 16900$; $SD[S] = 10\sqrt{169} = 130$

$$\mathbb{P}[S > 17000] = \mathbb{P}\left[\underbrace{\frac{S-16900}{130}}_{\approx N(0,1)\sim Z} > \frac{17000-16900}{130}\right] = \mathbb{P}\left[Z > \frac{100}{130}\right]$$

≈ 0.77

$= 1 - \Phi(0.77)$

$= 1 - 0.7794 = 0.2206$ $\square$

Problem 12.4. Source: Sample P exam, Problem #65.

A charity receives 2025 *contributions. Contributions are assumed to be mutually independent and identically distributed with mean* 3125 *and standard deviation* 250. *Calculate the approximate* 90^{th} *percentile for the distribution of the total contributions received.*

→: $S = Y_1 + \cdots + Y_n$ $n = 2025$ w/ Y_i st. $\mu_Y = 3125$ and $\sigma = 250$

$\mathbb{E}[S] = (2025)(3125) = \mu_S$

$SD[S] = 250 \cdot \sqrt{2025} = 250(45) = \sigma_S$

$\pi = ?$ such that $\mathbb{P}[S \le \pi] \approx 0.90$

1st Find the 90th percentile of $Z \sim N(0,1)$

$z^* = \Phi^{-1}(0.9) = \text{qnorm}(0.9) = 1.28$

2nd Apply the linear transform to the above

$\mathbb{P}[Z \le z^*] = 0.90$

$\mathbb{P}[\boxed{\mu_S + \sigma_S \cdot Z} \le \mu_S + \sigma_S \cdot z^*] = 0.9$

$\overset{??}{S}$

$\mathbb{P}[S \le \boxed{(2025)(3125) + 250(45) \cdot 1.28}] \approx 0.90$

π... answer

$\square$

[1] http://thereallyterribleorchestra.com/wordpress/