Central limit theorem

Central limit theorem: When sample size is large (≥30), the average of a set of independent identically distributed random variables ($\bar{X}$) is always approximately normally distributed with mean **population mean** and variance **population variance divided by sample size**. i.e

When n is large, $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

if $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ it indicates that

$E(\bar{X}) = \boldsymbol{\mu}$

$\Rightarrow E(\frac{X1 + X2 + \ldots + Xn}{n}) = \boldsymbol{\mu}$

$\Rightarrow \frac{1}{n} E(X_1 + X_2 + \ldots + Xn) = \boldsymbol{\mu}$ [Formula, E(cx)= cE(x)]

$\Rightarrow E(X_1 + X_2 + \ldots + Xn) = n\boldsymbol{\mu}$

$\Rightarrow E(\sum_{i=1}^{n} X_i) = n\boldsymbol{\mu}$

$V(\bar{X}) = \frac{\sigma^2}{n}$

$\Rightarrow V(\frac{X1 + X2 + \ldots + Xn}{n}) = \frac{\sigma^2}{n}$

$\Rightarrow \frac{1}{n^2} V(X_1 + X_2 + \ldots + Xn) = \frac{\sigma^2}{n}$ [Formula $V(cX) = c^2V(X)$]

$\Rightarrow V(X_1 + X_2 + \ldots + Xn) = n^2. \frac{\sigma^2}{n}$

$\Rightarrow V(X_1 + X_2 + \ldots + Xn) = n\sigma^2$

$\Rightarrow V(\sum_{i=1}^{n} X_i) = n\boldsymbol{\sigma}^2$

i.e $\sum_{i=1}^{n} X_i \sim N(n\boldsymbol{\mu}, n\boldsymbol{\sigma}^2)$

Example: The number of flaws in a glass sheet has a Poisson distribution with a parameter λ = 0.5. a) What is the distribution of the **average** number of flaws per sheet in 100 sheets of glass? b) Calculate the probability that this average is between 0.45 and 0.55. c) What is the distribution of the **total** number of flaws in 100 sheets of glass? d) Calculate the probability that there are fewer than 40 flaws in 100 sheets of glass?

Solution: The probability mass function of Poisson distribution is

$$P(X=x) = \frac{e^{-0.5}0.5^x}{x!} \qquad x=0,1,2,3\ldots$$

Expectation E(X) = 0.5

Variance V(X) = 0.5

a) The distribution of the average number of flaws per sheet in 100 sheets of glass follows normal distribution with mean = 0.5

variance = $\frac{0.5}{100}$

i.e $\bar{X}$ ~ N(0.5, $\frac{0.5}{100}$)

b) P(0.45 < $\bar{X}$ <0.55)

$= P(\frac{0.45-0.5}{\sqrt{\frac{0.5}{100}}} < \frac{\bar{X}-0.5}{\sqrt{\frac{0.5}{100}}} < \frac{0.55-0.5}{\sqrt{\frac{0.5}{100}}})$

= P(-0.707<Z<0.707)

= F(.707) – F(-.707)

= 0.7611- 0.2389

= 0.5222

c) The distribution of the total number of flaws in 100 sheets of glass

$X_1 + X_2 + \ldots + X_{100}$ ~ N(100*0.5, 100*0.5)

$=\sum_{i=1}^{100} X_i$ ~ N(50,50)

$d)\ P(\sum_{i=1}^{100} X_i < 40) = P(\frac{\sum_{i=1}^{100} X_i - 50}{\sqrt{50}} < \frac{40-50}{\sqrt{50}})$ = P(Z<-1.41) = F(-1.41)

= 0.0793