Ex: Scores obtained by students in MA20205 in the end semester exam are iid random variables with mean 30 marks and standard deviation of 5 marks. Assuming that the class has 200 students, what is the probability that the average of the marks obtained by students lies below 25?

Soln: $X_1, X_2, \ldots, X_{200}$: Marks obtained by students. (iid)

mean $= \mu = 30$; $\sigma =$ s.d. $= 5$.

Required prob $= P\left(\overline{X} = \frac{\sum_{i=1}^{200} X_i}{200} \leq 25\right) = P\left(\sum_{i=1}^{200} X_i \leq 5000\right)$

$\sum_{i=1}^{200} X_i$ has mean $200 \times 30 = 6000$; $\sum_{i=1}^{200} X_i$ has var $200 \times 25 = 5000$ (independence used)

Required prob

$$= P\left(\sum_{i=1}^{200} x_i \leq 5000\right) = P\left(\frac{\sum_{i=1}^{200} x_i - 6000}{\sqrt{5000}} \leq \frac{5000-6000}{\sqrt{5000}}\right)$$

$\because \frac{\sum_{i=1}^{200} x_i - 6000}{\sqrt{5000}}$ approx $\sim N(0,1)$

$$= \Phi\left(\frac{5000-6000}{\sqrt{5000}}\right)$$

where Φ is the CDF of $N(0,1)$.

$=$ ——

Ex: A goods train has 50 wagons. A crane dumps coal in each of these wagons in a coal mine. The amount of coal dumped in each wagon is a random variable with mean 62 tonnes and the standard deviation of 1 tonne. One engine can pull a train with load capacity of 3000 tonnes. What is the probability that one engine is NOT sufficient to pull the train?

Soln: X_i = weight in each wagon, $i = 1, 2, \cdots, 50$.

mean $(X_i) = 62$ tonnes, s.d. $(X_i) = \sigma(X_i) = 1$ tonne.

$\sum_{i=1}^{50} X_i$ = total weight of load in the train

mean of $\sum_{i=1}^{50} X_i = 50 \times 62 = 3100$

variance of $\sum_{i=1}^{50} X_i = 50 \times 1 = 50$ (used independence)

$$P\left(\sum_{i=1}^{50} X_i \geq 3000\right) = P\left(\frac{\sum_{i=1}^{50} X_i - 3100}{\sqrt{50}} \geq \frac{3000-3100}{\sqrt{50}}\right)$$

$$\approx P\left(Z \geq \frac{-100}{\sqrt{50}}\right)$$

where $Z = \frac{\sum_{i=1}^{50} X_i - 3100}{\sqrt{50}}$

$\overset{\text{approx}}{\sim} N(0,1)$

by CLT

Ex: Let $X_1, X_2, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

Define $Y = \dfrac{\sum_{i=1}^{j} X_i}{j}$ and $Z = \dfrac{\sum_{i=j+1}^{n} X_i}{n-j}$

for some $j \in \{1, 2, \cdots, n\}$

i) Find the distribution of Y and Z.

ii) Under what condition on $\alpha, \beta \in \mathbb{R}$

$E(\alpha Y + \beta Z) = \mu$??

In this case, compute $\text{var}(\alpha Y + \beta Z)$.

Ex: Let $X_1, X_2, \cdots, X_{20}$ be a random sample from $N(\mu, \sigma^2)$.

Define $Y = \frac{X_1 + \cdots + X_5}{5}$

i) What is the distribution of

$$\sum_{i=1}^{20} (X_i - \mu)^2$$

ii) What is the distribution of

$$\sum_{i=1}^{17} \left(\frac{X_i - \mu}{\sigma} \right)^2$$

iii) What is the distribution of

$$\frac{\sqrt{3}(Y - \mu)}{\sqrt{\sum_{i=1}^{15} (X_i - \mu)^2}}$$

Ex: In an online question paper with 100 questions, every question is true/false type. A students attempts every question by guessing. The probability of guessing correct answer is 0.7. Assuming that each question is answered independently, compute the prob.

i) at least 25 questions are answered correctly.

ii) exactly 25 questions are answered correctly.

Ex: Normal approximation of Negative-Binomial.

Bernoulli events with probability of success = p.

Number of successes : α.

X = no. of failures before α successes

$$P(X = x) = \binom{x+\alpha-1}{x} p^{\alpha} (1-p)^{x} \qquad x = 0, 1, 2, \cdots$$

$$= 0 \qquad \text{o.w.}$$

Recall that this NB random variable is the sum of α indep. geometric (p) random variables with mean and variance

$$P(X \le x) = \Phi\left(\left[x + \tfrac{1}{2} - \frac{\alpha(1-p)}{p}\right] \Big/ \sqrt{\frac{\alpha(1-p)}{p^2}}\right) - \Phi\left(\frac{x - \frac{1}{2} - \frac{\alpha(1-p)}{p}}{\sqrt{\alpha(1-p)/p^2}}\right)$$

Ex: $X \sim$ Binomial (n, p) (Assume $n \to \infty$, $p \to 0$, & $np =$ const.

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, 2, \ldots, n$$

$$= 0 \qquad \text{o.w.}$$

Assume $\lambda = np \Rightarrow p = \frac{\lambda}{n}$

$$\lim_{n\to\infty} P(X = x) = \lim_{n\to\infty} \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\lambda^x}{x!} \lim_{n\to\infty} \frac{n!}{(n-x)!} \left(\frac{1}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}$$

1st term

$$\lim_{n\to\infty} \frac{n!}{(n-x)!} \left(\frac{1}{n}\right)^x$$

$$= \lim_{n\to\infty} \frac{n(n-1)\cdots(n-x)(n-x-1)\cdots(1)}{(n-x)(n-x-1)\cdots(1)} \left(\frac{1}{n}\right)^x$$

$$= \lim_{n\to\infty} \frac{n}{n}\frac{(n-1)}{n}\cdots\frac{(n-x+1)}{n}$$

$$= \lim_{n\to\infty} 1\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)$$

$$= 1$$

2nd term

$$\lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$$

3rd term

$$\lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{-x} = 1^{-\infty} = 1$$

$$\Rightarrow \lim_{n\to\infty} \quad p(X=x) = \frac{\lambda^x}{x!} e^{-\lambda}$$

Ex: $n=10,\ p=0.3 \qquad (\lambda = np = 3)$

$$p(X=4) = \binom{10}{4} (0.3)^4 (0.7)^6 \overset{??}{\approx} \frac{3^4}{4!} e^{-3}$$

$$p(X=0) = (0.7)^{10} \overset{??}{\approx} \frac{3^0}{0!} e^{-3}$$