

## Q1: Making auto parts

The population distribution:

- $\mu = 40.125$
- $\sigma = 0.002$

We are trying to find the mean and standard deviation of the sample mean,  $\bar{X}$  (size  $n = 4$ ).

And we know the properties of the sample mean

- $E(\bar{X}) = \mu = 40.125$
- $Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{0.002^2}{4} = 10^{-6}$

Therefore

- The mean of the numbers recorded is 40.125
- The standard deviation of the numbers recorded is 0.001

## Q2: Dust in coal mines

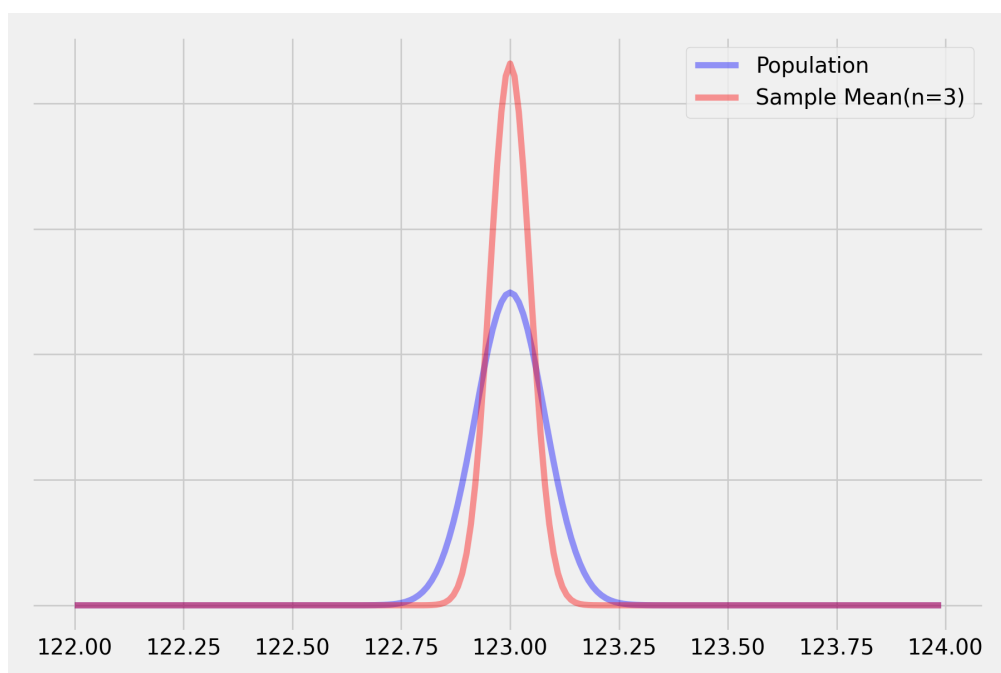
The population distribution is

$$N(\mu = 123, \sigma = 0.08) \quad (1)$$

(a) The laboratory reports the mean of 3 weighings. What is the distribution of this mean?

If individual observations have the  $N(\mu, \sigma)$  distribution, then the sample mean  $\bar{x}$  of an SRS of size  $n$  has the  $N(\mu, \frac{\sigma}{\sqrt{n}})$  distribution.

Therefore, the distribution of this mean is  $N(123, \frac{0.08}{\sqrt{3}})$



(b) What is the probability that the laboratory reports a weight of 124 mg or higher for this filter?

We need to standardize first

$$\bar{x} = 124 \quad (2)$$

$$z = \frac{\bar{x} - \mu}{\sigma} = \frac{124 - 123}{0.08/\sqrt{3}} = 21.65 \quad (3)$$

As you can see,  $z$  value is sky high!! So

$$P(Z \geq 21.65) \approx 0 \quad (4)$$

Therefore, the probability that the laboratory reports a weight of 124 mg or higher for this filter is 0.

### Q3: Pollutants in auto exhausts.

The population distribution with

$$\mu = 0.9 \quad \sigma = 0.15 \quad (5)$$

(a) What is the approximate distribution of the mean NOX emission level  $\bar{x}$  for these cars?

The central limit theorem:

Draw an SRS of size  $n$  from any population with mean  $\mu$  and finite standard deviation  $\sigma$ . When  $n$  is large, the sampling distribution of the sample mean  $\bar{x}$  is approximately Normal:

$$\bar{x} \text{ is approximately } N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \quad (6)$$

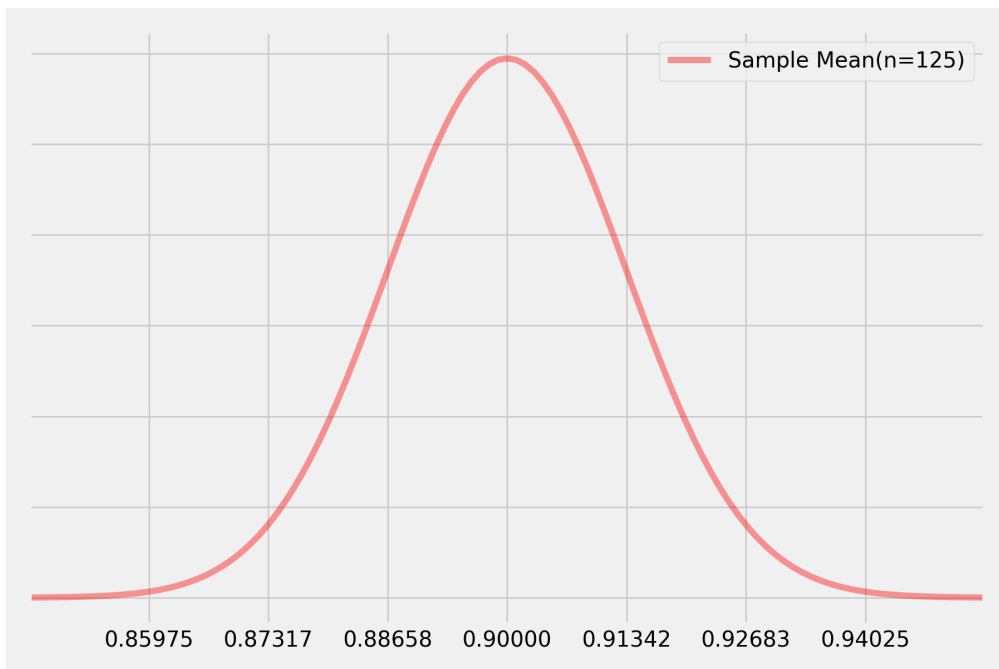
Now

$$n = 125 \quad (7)$$

and it is large enough to apply the central limit theorem

Therefore, the approximate distribution of the mean NOX emission level  $\bar{x}$  for these cars is

$$N\left(\mu = 0.9, \sigma = \frac{0.15}{\sqrt{125}}\right) = N(\mu = 0.9, \sigma = 0.0134) \quad (8)$$



(b) What is the level  $L$  such that the probability that  $\bar{x}$  is greater than  $L$  is only 0.01?

We start from the standard Normal, and trying to find  $z'$

$$P(z \leq z') = 1 - 0.01 = 0.99 \quad (9)$$

From the Z-table, we get

$$P(z \leq 2.33) = 0.9901 \quad (10)$$

So, we can say

$$z' \approx 2.33 \quad (11)$$

Then, we need to unstandardize

$$\frac{L - 0.9}{0.0134} = 2.33 \quad (12)$$

$$L = 0.9312 \quad (13)$$

## Q4: How many people in a car?

The population distribution with

$$\mu = 1.5 \quad \sigma = 0.75 \quad (14)$$

(a) Could the exact distribution of the count be Normal? Why or why not?

The exact distribution could be Normal or "any" distribution. We do not have the information about the population distribution except the mean and standard deviation.

(b) Traffic engineers estimate that the capacity of the interchange is 700 cars per hour. According to the central limit theorem, what is the approximate distribution of the mean number of persons  $\bar{x}$  in 700 randomly selected cars at this interchange?

The central limit theorem:

Draw an SRS of size  $n$  from any population with mean  $\mu$  and finite standard deviation  $\sigma$ . When  $n$  is large, the sampling distribution of the sample mean  $\bar{x}$  is approximately Normal:

$$\bar{x} \text{ is approximately } N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \quad (15)$$

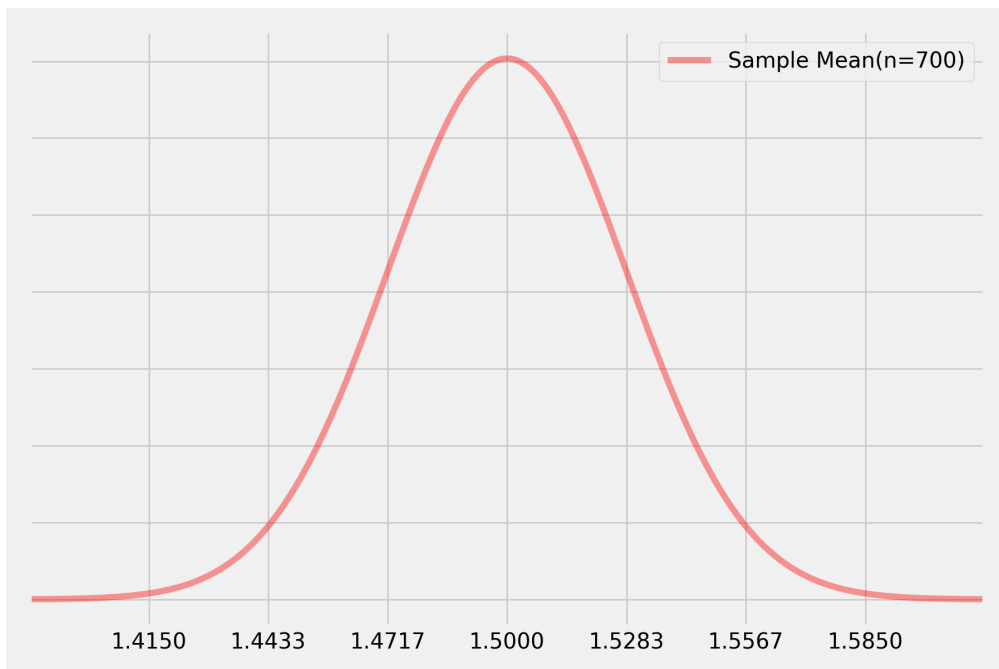
Now

$$n = 700 \quad (16)$$

and it is large enough to apply the central limit theorem

Therefore, the approximate distribution of the mean number of persons  $\bar{x}$  in 700 randomly selected cars at this interchange is

$$N\left(\mu = 1.5, \sigma = \frac{0.75}{\sqrt{700}}\right) = N(\mu = 1.5, \sigma = 0.0283) \quad (17)$$



(c) What is the probability that 700 cars will carry more than 1075 people?

700 cars carry more than 1075 people, it is equivalent to

$$\bar{x} = \frac{1075}{700} = 1.536 \quad (18)$$

Therefore, this question is to ask

$$P(\bar{x} > 1.536) \quad (19)$$

In order to use the Z-table, we need to standardize

$$z' = \frac{1.536 - 1.5}{0.0283} = 1.27 \quad (20)$$

From the Z-table, we get

$$P(z > 1.27) = 1 - P(z \leq 1.27) = 1 - 0.8980 = 0.102 \quad (21)$$

Therefore, the probability that 700 cars will carry more than 1075 people is 0.102

## Q5: Random-digit dialing

$$X \sim \text{Binomial}(n = 15, p = 0.2) \quad (22)$$

(a) What is the probability that exactly 3 calls reach a person?

This question is trying to find

$$\begin{aligned} P(X = 3) &= \binom{15}{3} 0.2^3 (1 - 0.2)^{15-3} \\ &= \frac{15!}{3!12!} (8 \times 10^{-3}) (0.0687) \\ &= 455 \times 0.5496 \times 10^{-3} \\ &= 0.25 \end{aligned}$$

Therefore, the probability that exactly 3 calls reach a person is 0.25

(b) What is the probability that 3 or fewer calls reach a person?

This question is trying to find

$$\begin{aligned} P(X \leq 3) &= P(0) + P(1) + P(2) + P(3) \\ &= \binom{15}{0} 0.2^0 (1 - 0.2)^{15} + \binom{15}{1} 0.2^1 (1 - 0.2)^{15-1} \\ &\quad + \binom{15}{2} 0.2^2 (1 - 0.2)^{15-2} + 0.25 \\ &= 0.035 + 0.132 + 0.23 + 0.25 = 0.647 \end{aligned}$$

Therefore, the probability that 3 or fewer calls reach a person is 0.647

## Q6: Multiple-choice tests

(a) Use the Normal approximation to find the probability that Jodi scores 70% or lower on a 100-question test.

First, this question should be expressed as

$$\begin{aligned} X &\sim \text{Binomial}(n = 100, p = 0.75) \\ \mu &= np = 75 \\ \sigma &= \sqrt{np(1 - p)} = 4.33 \end{aligned}$$

As the number of observations  $n$  gets larger, the binomial distribution gets close to a Normal distribution.

$$Y \sim N(\mu = 75, \sigma = 4.33)$$

This question is trying to find

$$P(Y \leq 70)$$

Then, we standardize

$$\frac{70 - 75}{4.33} = -1.155 \approx -1.16$$

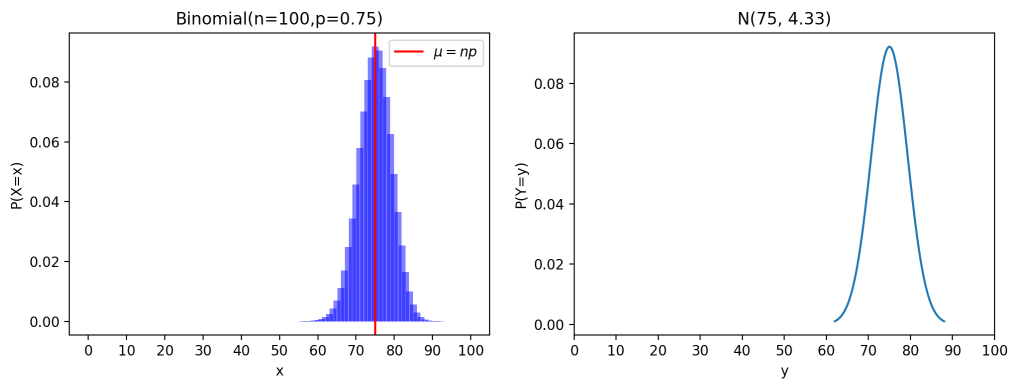
From the Z-table, we get

$$P(Z \leq 1.16) = 0.877$$

And from the property of symmetry, we get

$$P(Z \leq -1.16) = 1 - 0.877 = 0.123$$

Therefore, the probability that Jodi scores 70% or lower on a 100-question test is 0.123



(b) If the test contains 250 questions, what is the probability that Jodi will score 70% or lower?

First, this question should be expressed as

$$\begin{aligned} X &\sim \text{Binomial}(n = 250, p = 0.75) \\ \mu &= np = 187.5 \\ \sigma &= \sqrt{np(1-p)} = 6.85 \end{aligned}$$

And we also use the Normal distribution to approximate

$$Y \sim N(\mu = 187.5, \sigma = 6.85)$$

For 250 questions, score 70% is

$$250 \times 0.7 = 175$$

Therefore, this question becomes to find

$$P(Y \leq 175)$$

Then, we standardize

$$\frac{175 - 187.5}{6.85} = -1.82$$

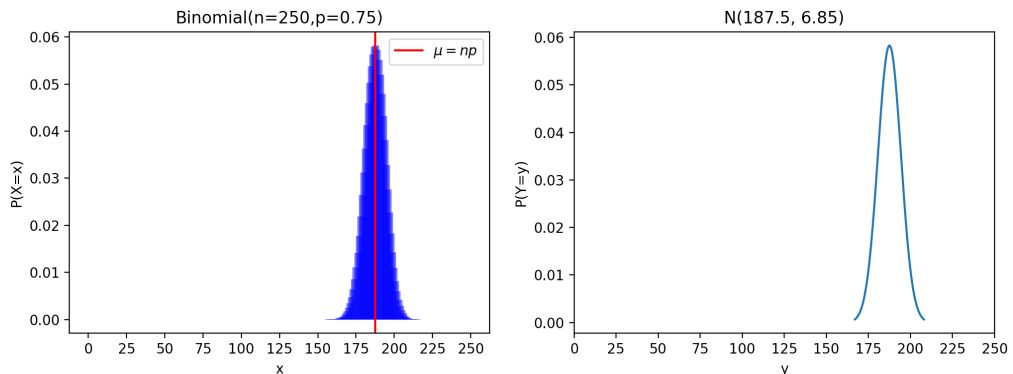
From the Z-table, we get

$$P(Z \leq 1.82) = 0.9656$$

And from the property of symmetry, we get

$$P(Z \leq -1.82) = 1 - 0.9656 = 0.0344$$

Therefore, the probability that Jodi scores 70% or lower on a 250-question test is 0.0344



## Q7: Reaching dropouts

(a) If the mailing list can be considered a random sample of the population, what is the mean number of high school dropouts who will receive the flyer?

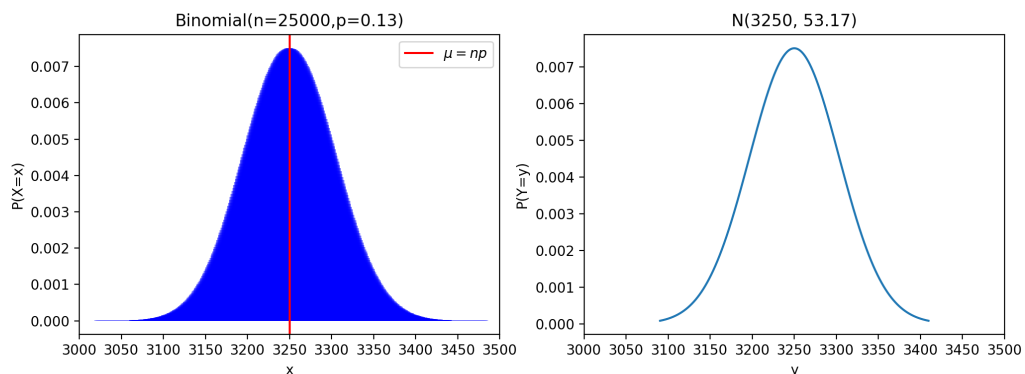
First, this question should be expressed as

$$X \sim \text{Binomial}(n = 25000, p = 0.13)$$

$$\mu = np = 3250$$

$$\sigma = \sqrt{np(1-p)} = 53.17$$

Therefore, the mean number of high school dropouts who will receive the flyer is 3250



(b) What is the probability that at least 3500 dropouts will receive the flyer?

Here, we use the Normal distribution to approximate

$$Y \sim N(\mu = 3250, \sigma = 53.17)$$

This question becomes to find

$$P(Y \geq 3500)$$

Then, we standardize

$$z = \frac{3500 - 3250}{53.17} = 4.7$$

So, we transform the original question to find

$$P(Z \geq 4.7) = 1 - P(Z < 4.7)$$

where

$$Z \sim N(0, 1)$$

From the largest number of Z-table,

$$P(Z \leq 3.49) = 0.9998$$

So, I conclude that

$$P(Z \geq 4.7) = 1 - P(Z < 4.7) \approx 0$$

Therefore, the probability that at least 3500 dropouts will receive the flyer is 0

## Q8-1: Chebyshev's inequality

Chebyshev Inequality:

Let  $X$  be a random variable with  $E(X) = \mu$  and  $\text{Var}(X) = \sigma^2$

For  $k > 0$

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

(a) Let  $X$  denote the amount of rainfall. received per week in a region. Assume that  $\mu = 1.00$  inch and  $\sigma = 0.25$  inch. Would it be unusual for this region to receive more than 2 inches of rain in a given week?

First, the probability that this region receives more than 2 inches of rain in a given week is

$$P(X \geq 2)$$

And we try to change the original equation

$$P(X \geq 2) = P(X - 1 \geq 1)$$

Now, I let  $k = 1$ , and use Chebyshev's inequality

$$P(|X - 1| \geq 1) \leq \frac{\sigma^2}{k^2} = \frac{0.25^2}{1^2} = 0.0625$$

It further becomes

$$\begin{aligned} P(X - 1 \leq -1) + P(X - 1 \geq 1) &\leq 0.0625 \\ P(X \leq 0) + P(X \geq 2) &\leq 0.0625 \end{aligned}$$

Then, I denote a positive number  $\alpha = P(X \leq 0) > 0$  (the probability that there is no rainfall), and get



$$P(X \geq 2) \leq 0.0625 - \alpha$$

As you can see, the probability that this region receive more than 2 inches of rain in a given week must be smaller than 0.0625, so I can say it is unusual for that.

(b) Let  $X$  denote the amount of cases of rabies reported in a given state per week. Assume that  $\mu = 0.5$  and  $\sigma^2 = 0.04$ . Would it be unusual to observe two cases in a given week?

First, the probability that observe at least two cases in a given week

$$P(X \geq 2)$$

And we try to change the original equation

$$P(X \geq 2) = P(X - 0.5 \geq 1.5)$$

Now, I let  $k = 1.5$ , and use Chebyshev's inequality

$$P(|X - 0.5| \geq 1.5) \leq \frac{0.04}{1.5^2} = 0.018$$

It further becomes

$$\begin{aligned} P(X - 0.5 \leq -1.5) + P(X - 0.5 \geq 1.5) &\leq 0.018 \\ P(X \leq -1) + P(X \geq 2) &\leq 0.018 \end{aligned}$$

And because the amount of cases can not be negative number, so

$$P(X \leq -1) = 0$$

Finally, we get

$$P(X \geq 2) \leq 0.018$$

As you can see, the probability that observe at least two cases in a given week must be smaller than 0.018, so I can say it is unusual for that.

## Q8-2

$$E(x) = \mu = 60, \quad \sigma = 7$$

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

### 1. 求在46分與74分之間的人數有多少人

First, let us check:

$$\begin{aligned} 60 - 46 &= 14 \\ 60 - 74 &= -14 \end{aligned}$$

Therefore, we can express  $46 \leq X \leq 74$  as

$$|X - 60| \leq 14$$

and this question is to ask

$$P(46 \leq X \leq 74) = P(|X - 60| \leq 14)$$

Now, I let  $k = 14 = 2\sigma$ , and use Chebyshev's inequality

$$P(|X - 60| \geq 2\sigma) \leq \frac{\sigma^2}{4\sigma^2} = \frac{1}{4}$$

Finally, we get

$$P(|X - 60| \leq 14) = 1 - P(|X - 60| \geq 14) = 1 - \frac{1}{4} = \frac{3}{4}$$
$$4000 \times \frac{3}{4} = 3000$$

Therefore, there are 3000 people in the range from 46 to 74.

## 2. 求最多佔全部1/9的分數範圍

This question is to find  $a$  and  $b$  in the following equation:

$$P(a \leq X \leq b) \leq \frac{1}{9}$$

We can not find the solution from Chebyshev's inequality.

However, we can find the solution if the question becomes "求最少佔全部1/9的分數範圍".

I will illustrate it in the last part of this homework.

## 2. 求最少佔全部1/9的分數範圍 (Revised Version)

This question is to find  $a$  and  $b$  in the following equation:

$$P(a \leq X \leq b) \geq \frac{1}{9}$$

We can further change the equation

$$P(a - 60 \leq X - 60 \leq b - 60) \geq \frac{1}{9}$$

If we let  $a - 60 = -k$  and  $b - 60 = k$ , we get

$$P(|X - 60| \leq k) \geq \frac{1}{9}$$
$$-P(|X - 60| \leq k) \leq -\frac{1}{9}$$
$$1 - P(|X - 60| \leq k) \leq \frac{8}{9}$$

Because  $\sum_i P(X_i) = 1$ , we know

$$P(|X - 60| \geq k) = 1 - P(|X - 60| \leq k)$$

therefore, we transform the original question to find  $k$

$$P(|X - 60| \geq k) \leq \frac{8}{9}$$

and combine with Chebyshev's inequality

$$P(|X - 60| \geq k) \leq \frac{\sigma^2}{k^2}$$

Doing simple algebra

$$\frac{\sigma^2}{k^2} = \frac{8}{9}$$

$$k^2 = \frac{9}{8}\sigma^2$$

$$k = \frac{3}{2\sqrt{2}}\sigma = 7.425$$

Substitute this result into

$$P(|X - 60| \leq k) \geq \frac{1}{9}$$

We get

$$\begin{aligned} |X - 60| &\leq 7.425 \\ -7.425 &\leq X - 60 \leq 7.425 \\ 52.575 &\leq X \leq 67.425 \end{aligned}$$

Therefore,

$$P(52.575 \leq X \leq 67.425) \geq \frac{1}{9}$$

最少佔全部1/9的分數範圍 在52.575分與67.425分之間