Summary---Topic 3: Discrete & Continuous Probability Distributions

Normal Distribution

- A continuous random variable is said to be a normal random variable if its probability density function is given by

$$f(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{1}{2}\right)\left[\frac{X-\mu}{\sigma}\right]^2}$$

➢ A normal distribution depends upon two parameters: "μ" and "σ^2" (or "σ")

➢ Expected value of normal random variable: μ

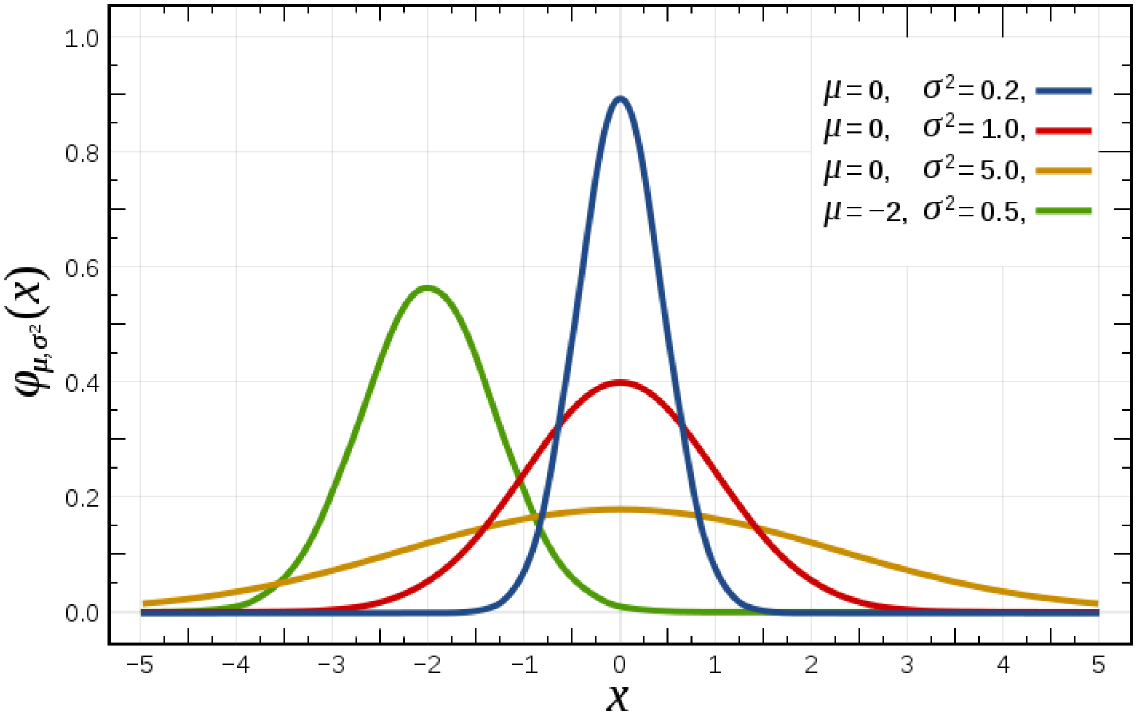
➢ Variance of normal random variable: σ^2

➢ Standard deviation of normal random variable: σ

- $\mathbf{X} \sim \boldsymbol{N}(\boldsymbol{\mu}, \boldsymbol{\sigma}^{\mathbf{2}})$

Normal Distribution

$$f(X) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\left(\frac{1}{2}\right)\left[\frac{X-\mu}{\sigma}\right]^2}$$



- Characteristics of normal distribution

→ "Bell Shaped"

→ symmetric about the mean

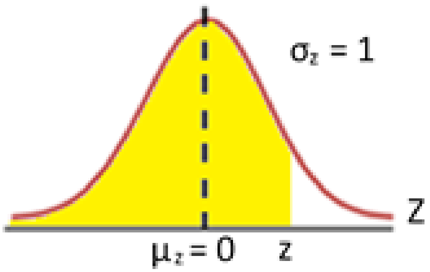
→ mean, median and mode are equal

→ has infinite theoretical range (-∞ to +∞)

Standard Normal Distribution (Z)

- mean **μ = 0** and variance **σ^2 = 1**
- $\mathbf{Z \sim \boldsymbol{N}(0, 1)}$
- Its probability density function is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$



- P(Z ≤ z) = area under the normal curve from -∞ to z
- Total area under the normal curve = 1

  → P(Z ≤ 0) = 0.5 and P(Z ≥ 0) = 0.5

Standard Normal Table : $\mathrm{P}(Z < k) = ?$

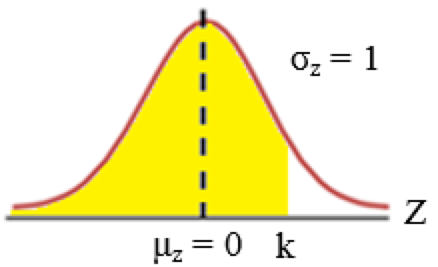
Formatting:

- The label for rows contains the integer part and the first decimal place of Z.
- The label for columns contains the second decimal place of Z.
- The value within the table gives the probability from Z=$-\infty$ up to the desired Z value

Z	**0.00**	**0.01**	**0.02**
0.0	0.5000	0.5040	0.5080
0.1	0.5398	0.5438	0.5478
0.2	0.5793	0.5832	0.5871
0.3	0.6179	0.6217	0.6255

$P(Z < 0.21) = 0.5832$

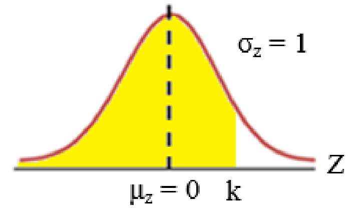
Value within the table is the cumulative probability from -∞ to a particular k value



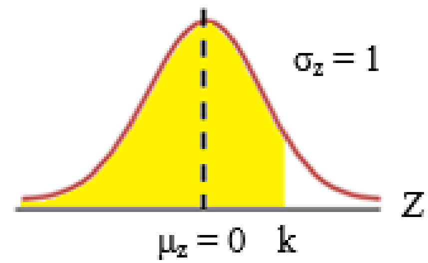
Standard Normal Table

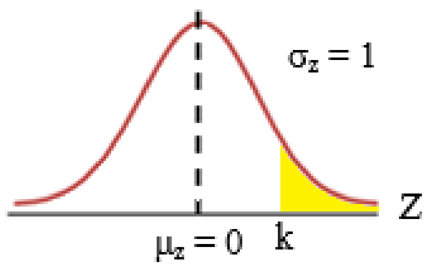
Z	0.00	0.01	0.02
0.0	0.5000	0.5040	0.5080
0.1	0.5398	0.5438	0.5478
0.2	0.5793	0.5832	0.5871
0.3	0.6179	0.6217	0.6255

Value within the table is the cumulative probability from $-\infty$ to a particular k value

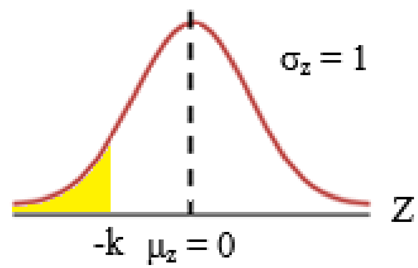


With the knowledge of $P(Z \leq k)$, we can easily know the value of $P(Z \geq k)$, $P(Z \leq -k)$, and $P(-k \leq Z \leq k)$ where k can be any real number.

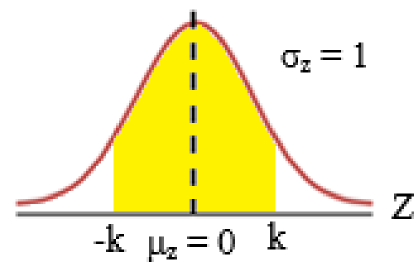




$P(Z \geq k) = 1 - P(Z \leq k)$



$P(Z \leq -k) = P(Z \geq k)$

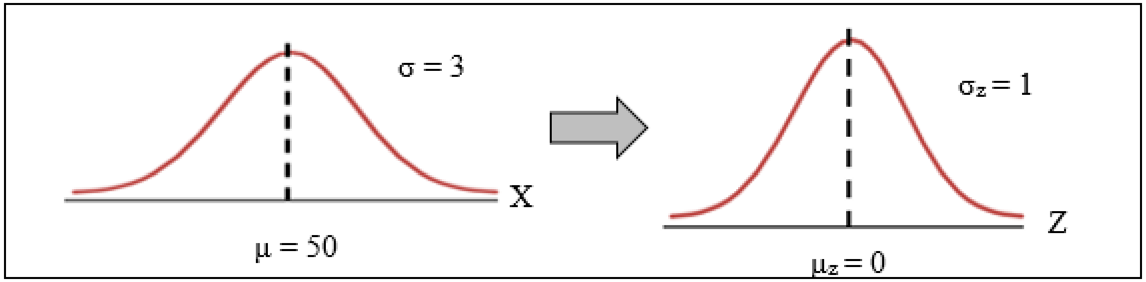


$P(-k \leq Z \leq k) =$
$P(Z \leq k) - P(Z \leq -k)$

Standardization of Normal Distribution

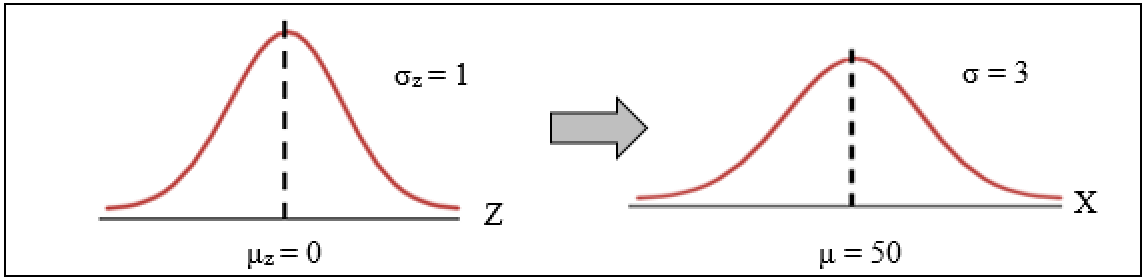
- Standardization: $\mathbf{X} \sim \boldsymbol{N}\left(\boldsymbol{\mu}, \boldsymbol{\sigma}^{2}\right) \rightarrow \boldsymbol{Z}=\frac{\boldsymbol{X}-\boldsymbol{\mu}}{\boldsymbol{\sigma}} \sim \boldsymbol{N}(\mathbf{0}, \mathbf{1})$

From non-standard normal distribution to standard normal distribution: $\boldsymbol{Z}=\frac{\boldsymbol{X}-\boldsymbol{\mu}}{\boldsymbol{\sigma}}$



If X = 50, Z = 0; if X = 80, Z = 10; if X = 20, Z = -10

From standard normal distribution to non-standard normal distribution: $\mathrm{X}=\boldsymbol{Z}\boldsymbol{\sigma}+\boldsymbol{\mu}$



If Z = 0, X = 50; if Z = 10, X = 80; if Z = -10, X = 20

Exercises and Solutions

Q11. Given a normal distribution with $\mu = 100$ and $\sigma = 10$, what is the probability that

a) X > 85?

b) X < 80?

c) X < 80 or X > 110?

d) 80% of the values are between what two X values (symmetrically distributed around the mean)?

Solution:

According to the question, we know that $X \sim N(\mu = 100, \sigma^2 = 10^2)$.

To compute non-standard normal probabilities:

1. Do standardization: from the non-standard normal distribution to standard normal distribution: $\boldsymbol{Z} = \frac{\boldsymbol{X-\mu}}{\boldsymbol{\sigma}}$.
2. Check the Standard Normal Table.

a) $\mathrm{P}(\mathrm{X} > 85) = \mathrm{P}\left(\frac{X-100}{10} > \frac{85-100}{10}\right) = \mathrm{P}(Z > -1.5) = 1 - \mathrm{P}(Z \leq -1.5) = 1 - 0.0668 = 0.9332$

Q11. Given a normal distribution with $\mu = 100$ and $\sigma = 10$, what is the probability that

a) X > 85?

b) X < 80?

c) X < 80 or X > 110?

d) 80% of the values are between what two X values (symmetrically distributed around the mean)?

To compute non-standard normal probabilities:

1. Do standardization: from the non-standard normal distribution to standard normal distribution: $\boldsymbol{Z} = \frac{\boldsymbol{X}-\boldsymbol{\mu}}{\boldsymbol{\sigma}}$.
2. Check the Standard Normal Table.

b) $\mathrm{P}(\mathrm{X} < 80) = \mathrm{P}\left(\frac{X-100}{10} < \frac{80-100}{10}\right) = \mathrm{P}(Z < -2) = 0.0228$

c) $\mathrm{P}(\mathrm{X} < 80 \text{ or } \mathrm{X} > 110) = \mathrm{P}(\mathrm{X} < 80) + \mathrm{P}(\mathrm{X} > 110)$

$= 0.0228 + \mathrm{P}\left(\frac{X-100}{10} > \frac{110-100}{10}\right)$

$= 0.0228 + \mathrm{P}(Z > 1) = 0.0228 + 1 - \mathrm{P}(Z \leq 1)$

$= 0.0228 + 1 - 0.8413 = 0.1815$

Q11. Given a normal distribution with $\mu = 100$ and $\sigma = 10$, what is the probability that

d) 80% of the values are between what two X values (symmetrically distributed around the mean)?

To compute non-standard normal probabilities:

1. Do standardization: from the non-standard normal distribution to standard normal distribution: $\boldsymbol{Z} = \frac{\boldsymbol{X}-\boldsymbol{\mu}}{\boldsymbol{\sigma}}$.
2. Check the Standard Normal Table.

d) 80% of the values are between two values → $\mathrm{P}(\mathrm{a} < \mathrm{X} < \mathrm{b}) = 0.8$

symmetrically distributed around the mean→

$$\begin{cases} \mathrm{P}(X < a) = 0.1 \\ \mathrm{P}(X < b) = 0.9 \end{cases}$$

Do the standardization, we have

$$\begin{cases} \mathrm{P}\left(\frac{X-100}{10} < \frac{a-100}{10}\right) = 0.1 \\ \mathrm{P}\left(\frac{X-100}{10} < \frac{b-100}{10}\right) = 0.9 \end{cases} \Rightarrow \begin{cases} \mathrm{P}\left(Z < \frac{a-100}{10}\right) = 0.1 \\ \mathrm{P}\left(Z < \frac{b-100}{10}\right) = 0.9 \end{cases}$$

$$\Rightarrow \begin{cases} \frac{a-100}{10} = -1.28 \\ \frac{b-100}{10} = 1.28 \end{cases} \Rightarrow \begin{cases} a = 87.2 \\ b = 112.8 \end{cases}$$

Q12*. The breaking strength of plastic bags used for packaging produce is normally distributed, with a mean of 5 pounds per square inch and a standard deviation of 1.5 pounds per square inch. What proportion of the bags have a breaking strength of

a) Less than 3.11 pounds per square inch?

b) At least 3.8 pounds per square inch?

c) Between 5 and 5.5 pounds per square inch?

d) 95% of the breaking strength will be contained between what two values symmetrically distributed around the mean?

Let X be the breaking strength of plastic bags, then $X \sim N(\mu = 5, \sigma^2 = 1.5^2)$

a) $\mathrm{P(X < 3.11)} =?$
b) $\mathrm{P(X \geq 3.8)} =?$
c) $\mathrm{P(5 < X < 5.5)} =?$
d) If $\mathrm{P(a < X < b)} = 0.95$, and “a” and “b” are symmetrically distributed around the mean, then what are the values of “a” and “b”?

Q12*. $X \sim N(\mu = 5, \sigma^2 = 1.5^2)$ $\boldsymbol{\Rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)}$

a) $\mathrm{P}(\mathrm{X} < 3.11) = ?$

b) $\mathrm{P}(\mathrm{X} \geq 3.8) = ?$

c) $\mathrm{P}(5 < \mathrm{X} < 5.5) = ?$

d) If $\mathrm{P}(\mathrm{a} < \mathrm{X} < \mathrm{b}) = 0.95$, and "a" and "b" are symmetrically distributed around the mean, then what are the values of "a" and "b"?

To compute non-standard normal probabilities:

1. Do standardization: from the non-standard normal distribution to standard normal distribution
2. Check the Standard Normal Table.

Solution:

a) $\mathrm{P}(\mathrm{X} < 3.11) = \mathrm{P}\left(\frac{X-5}{1.5} < \frac{3.11-5}{1.5}\right) = \mathrm{P}(Z < -1.26) = 0.1038$

b) $\mathrm{P}(\mathrm{X} \geq 3.8) = \mathrm{P}\left(\frac{X-5}{1.5} \geq \frac{3.8-5}{1.5}\right) = \mathrm{P}(Z \geq -0.8) = 1 - \mathrm{P}(Z < -0.8) = 0.7881$

c) $\mathrm{P}(5 < \mathrm{X} < 5.5) = \mathrm{P}\left(\frac{5-5}{1.5} < \frac{X-5}{1.5} < \frac{5.5-5}{1.5}\right) = \mathrm{P}(0 < Z < 0.33)$

$= \mathrm{P}(Z < 0.33) - \mathrm{P}(Z < 0) = 0.6293 - 0.5 = 0.1293$

d) $\begin{cases} \mathrm{P}(X < a) = 0.025 \\ \mathrm{P}(X < b) = 0.975 \end{cases} \Rightarrow \begin{cases} \mathrm{P}\left(\frac{X-5}{1.5} < \frac{a-5}{1.5}\right) = 0.025 \\ \mathrm{P}\left(\frac{X-5}{1.5} < \frac{b-5}{1.5}\right) = 0.975 \end{cases} \Rightarrow \begin{cases} \frac{a-5}{1.5} = -1.96 \\ \frac{b-5}{1.5} = 1.96 \end{cases} \Rightarrow \begin{cases} a = 2.06 \\ b = 7.94 \end{cases}$

Q13*A statistical analysis of 1,000 long-distance telephone calls made from the headquarters of the Bricks and Clicks Computer Corporation indicates that the length of these calls is normally distributed with μ = 220 seconds and σ = 30 seconds.

a) What is the probability that a call lasted less than 175 seconds?

b) What is the probability that a call lasted between 175 and 265 seconds?

c) What is the probability that a calls lasted between 115 and 175 seconds?

d) What is the length of a call if only 1% of all calls are shorter?

Solution:

Let X be the length of long-distance telephone call. $X \sim N(\mu = 220, \sigma^2 = 30^2)$

a) $\mathrm{P}(\mathrm{X} < 175) = \mathrm{P}\left(\frac{X-220}{30} < \frac{175-220}{30}\right) = \mathrm{P}(Z < -1.5) = 0.0668$

b) $\mathrm{P}(175 < \mathrm{X} < 265) = \mathrm{P}\left(\frac{175-220}{30} < \frac{X-220}{30} < \frac{265-220}{30}\right) = \mathrm{P}(-1.5 < Z < 1.5)$

$= \mathrm{P}(Z < 1.5) - \mathrm{P}(Z < -1.5) = 0.9332 - 0.0668 = 0.8664$

c) $\mathrm{P}(115 < \mathrm{X} < 175) = \mathrm{P}\left(\frac{115-220}{30} < \frac{X-220}{30} < \frac{175-220}{30}\right) = \mathrm{P}(-3.5 < Z < -1.5)$

$= \mathrm{P}(Z < -1.5) - \mathrm{P}(Z < -3.5) = 0.0668 - 0.00023 = 0.06657$

d) We need find the value of "a" such that $\mathrm{P}(\mathrm{X} < a) = 0.01$

$\Rightarrow \mathrm{P}\left(\frac{X-220}{30} < \frac{a-220}{30}\right) = 0.01 \Rightarrow \mathrm{P}\left(Z < \frac{a-220}{30}\right) = 0.01 \Rightarrow \frac{a-220}{30} = -2.33 \Rightarrow a = 150.1$

Q14. The exam marks of a large class of students follow a normal distribution with mean μ and standard deviation σ. 1% of the students got 90 or above. 10% of the students got 40 or below. The passing mark is 50.

a) Find the values of μ and σ.

b) Find the chance that a randomly selected student passes the exam.

Solution:

a) Let X be the exam marks of the student, $X \sim N(\mu, \sigma^2)$

$$\begin{cases} \mathrm{P}(X \geq 90) = 0.01 \\ \mathrm{P}(X \leq 40) = 0.1 \end{cases} \rightarrow \begin{cases} \mathrm{P}(X < 90) = 0.99 \\ \mathrm{P}(X \leq 40) = 0.1 \end{cases} \rightarrow \begin{cases} \mathrm{P}\left(\frac{X-\mu}{\sigma} < \frac{90-\mu}{\sigma}\right) = 0.99 \\ \mathrm{P}\left(\frac{X-\mu}{\sigma} < \frac{40-\mu}{\sigma}\right) = 0.1 \end{cases}$$

$$\rightarrow \begin{cases} \mathrm{P}\left(Z < \frac{90-\mu}{\sigma}\right) = 0.99 \\ \mathrm{P}\left(Z < \frac{40-\mu}{\sigma}\right) = 0.1 \end{cases} \rightarrow \begin{cases} \frac{90-\mu}{\sigma} = 2.33 \\ \frac{40-\mu}{\sigma} = -1.28 \end{cases} \rightarrow \begin{cases} \mu = 57.73 \\ \sigma = 13.85 \end{cases}$$

b) $X \sim N(57.73, 13.85^2)$

$$\mathrm{P}(\mathrm{X} \geq 50) = \mathrm{P}\left(\frac{X-57.73}{13.85} \geq \frac{50-57.73}{13.85}\right) = \mathrm{P}(Z \geq -0.56) = 1 - \mathrm{P}(Z < -0.56)$$
$$= 1 - 0.2877 = 0.7123$$

Q15. The fill amount of bottles of soft drink has been found to be normally distributed with a mean amount of 2.0 liters and a standard deviation of 0.05 liter. Bottles that contain less than 95% of the listed net content (**1.90 liters in this case**) can make the manufacturer **subject to penalty by the Consumer Council**, whereas bottles that have a net content **above 2.12** liters may cause **opening excess spillage upon**

a) **What proportion of the bottles is subject to penalty by the Consumer Council?**

b) **What proportion of the bottles is risking to excess spillage upon opening?**

c) In an effort to reduce the possible penalty due to insufficient net content in the bottles, the manufacturer has set out the following quality control requirement: 99% of bottles should comply with the Consumer Council's standard. To achieve this, the bottler decides to set the filling machine to a new mean amount. Determine the mean amount to be set for the bottle filling machine such that the above requirement can be met.

Solution:

Let X be the fill amount of bottles of soft drink, then $X \sim N(\mu = 2, \sigma^2 = 0.05^2)$

a) $\mathrm{P}(\mathrm{X} < 1.9) = \mathrm{P}\left(\frac{X-2}{0.05} < \frac{1.9-2}{0.05}\right) = \mathrm{P}(Z < -2) = 0.0228$

b) $\mathrm{P}(\mathrm{X} > 2.12) = \mathrm{P}\left(\frac{X-2}{0.05} > \frac{2.12-2}{0.05}\right) = \mathrm{P}(Z > 2.4) = 1 - \mathrm{P}(Z \leq 2.4) = 0.0082$

Q15. The fill amount of bottles of soft drink has been found to be normally distributed with a mean amount of 2.0 liters and a standard deviation of 0.05 liter. Bottles that contain less than 95% of the listed net content (1.90 liters in this case) can make the manufacturer subject to penalty by the Consumer Council, whereas bottles that have a net content above 2.12 liters may cause excess spillage upon opening.

c) In an effort to reduce the possible penalty due to insufficient net content in the bottles, the manufacturer has set out the following quality control requirement: 99% of bottles should comply with the Consumer Council's standard. To achieve this, the bottler decides to set the filling machine to a new mean amount. Determine the mean amount to be set for the bottle filling machine such that the above requirement can be met.

Solution:

c) $X \sim N(\mu, \sigma^2 = 0.05^2)$

→ $\mathrm{P}(\mathrm{X} < 1.9) = 1 - 0.99 = 0.01$

→ $\mathrm{P}\left(\frac{X-2}{0.05} < \frac{1.9-\mu}{0.05}\right) = \mathrm{P}\left(Z < \frac{1.9-\mu}{0.05}\right) = 0.01$

→ $\frac{1.9-\mu}{0.05} = -2.33$ → $\mu = 2.0165$

Q16. At the CityU Computer Service Centre, the loading time for e-Portal page on Internet Explorer is normally distributed with mean 3 seconds.

a) Without doing the calculations, for a randomly selected student, which of the following intervals of loading time (in second) is the most likely to be: 2.9-3.1, 3.1-3.3, 3.3-3.5, 3.5-3.7? Which interval of loading time is the least likely to be? Explain.

b) What is the chance that the loading time is exactly 2 seconds?

Solution:

a) The loading time is normally distributed with mean of 3 seconds

- Most likely: 2.9-3.1, since it lies in the central part of the normal distribution model, which has the largest area, thus the largest probability to occur.
- Less likely: 3.5-3.7, since it is the farthest interval from the mean, thus has the least probability to occur under the normal distribution model.

b) $P(X = 2) = 0$, since it is a line, not an area, this probability = 0.

Q17*. The volume of a randomly selected bottle of a new type of mineral water is known to have a normal distribution with **a mean of 995ml** and a **standard deviation of 5ml**.

What is the volume that should be stamped on the bottle so that only 3% of bottles are underweight?

Solution:

Let X be the volume that should be stamped on the bottle,

then $X \sim N(995, 5^2)$

We need find the value of "a", such that $\mathrm{P}(\mathrm{X} < a) = 0.03$

$$\rightarrow \mathrm{P}\left(\frac{X-995}{5} < \frac{a-995}{5}\right) = \mathrm{P}\left(Z < \frac{a-995}{5}\right) = 0.03$$

$$\rightarrow \frac{a-995}{5} = -1.88 \quad \rightarrow a = 985.6$$