D1EAD – Análise Estatística para Ciência de Dados 2021.1

Data Distributions (Part 2)

Prof. Ricardo Sovat
sovat@ifsp.edu.br
Prof. Samuel Martins (Samuka)
samuel.martins@ifsp.edu.br

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
São Paulo
Campus Campinas

Data Distributions

Motivation

- Some Machine Learning models are designed to work best under some **distribution** **assumptions**.
- Knowing with which **distributions** we are working with can help us to:
    - **identify** which machine-learning models are best to use.
    - Make analysis and inference easier during the **exploratory data analysis**.

- But, let's take a look at some **basic concepts** first.

Notations

https://www.youtube.com/watch?v=mtbJbDwqWLE

Notations

fixed number from population

change from sample to sample

Population Parameter	Sample Statistic	Description
N	n	Number of elements.
μ	$\bar{x}$	Mean
σ	s	Standard deviation
ρ	r	Correlation coefficient.

Basic Concepts

Random experiment

- Process by which we observe something **uncertain**.
- Experiment, trial, or observation that **can be repeated** numerous times under the **same conditions**.
- Ex: toss a coin, roll a die, perc. of calls dropped due to errors over a particular time period, …

Outcome

- A result of a **random experiment**.
- The **outcome** of an individual random experiment **must not be affected by any previous outcome** and **cannot be predicted with certainty**.

Sample space

- The set of all possible outcomes of a random experiment.

Basic Concepts

Random experiment

- Process by which we observe something **uncertain**.
- Experiment, trial, or observation that **can be repeated** numerous times under the **same conditions**.
- Ex: toss a coin, roll a die, perc. of calls dropped due to errors over a particular time period, …

Outcome

- A result of a **random experiment**.
- The **outcome** of an individual random experiment **must not be affected by any previous outcome** and **cannot be predicted with certainty**.

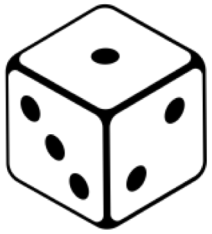
Sample space

- The set of all possible outcomes

Example 1:

Random experiment: roll a die

Sample space: S = {1, 2, 3, 4, 5, 6}.



Example 2:

Random experiment: toss a coin.

Sample space: S = {head, tail}.

Example 3:

Random experiment: number of iPhones sold in Brazil in 2020.

Sample space: S = {0, 1, 2, 3, … }.

Basic Concepts

Event

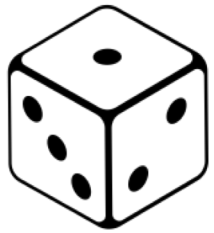
- An **outcome** or a **collection of outcomes** of a random experiment.
- **Any subset** of a sample space.

Ex:

Random experiment: roll a die

Sample space: S = {1, 2, 3, 4, 5, 6}.

Event: Getting an even number -> E = {2, 4, 6}.



Basic Concepts

Random variable

- Variable whose **values** depend on **outcomes** of a random phenomenon (e.g., random experiment).
- Think of it as a rule to decide what number you should record in your dataset after a real-world event happens.
- It can be **discrete** (takes countable number of distinct values) or **continuous** (the values between the range/interval and take infinite numbers).

Basic Concepts

Random variable

- Variable whose **values** depend on **outcomes** of a random phenomenon (e.g., random experiment).
- Think of it as a rule to decide what number you should record in your dataset after a real-world event happens.
- It can be **discrete** (takes countable number of distinct values) or **continuous** (the values between the range/interval and take infinite numbers).

Ex 1:

Random experiment: toss a coin.

Random variable: X = 0 (Head), 1 (Tail)

Basic Concepts

Random variable

- Variable whose **values** depend on **outcomes** of a random phenomenon (e.g., random experiment).
- Think of it as a rule to decide what number you should record in your dataset after a real-world event happens.
- It can be **discrete** (takes countable number of distinct values) or **continuous** (the values between the range/interval and take infinite numbers).

Ex 1:

Random experiment: toss a coin.

Random variable: X = 0 (Head), 1 (Tail)

Ex 2:

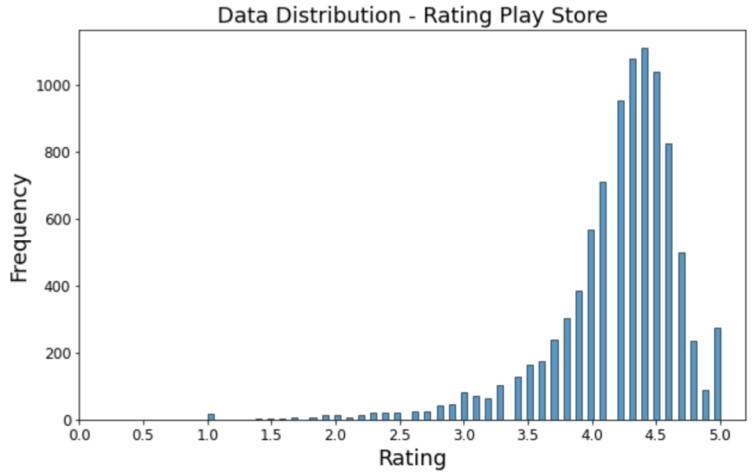
Random experiment: a soccer match.

Basic Concepts

Data Distribution

- Distribution of **individual data points** from a dataset.
- It is a function or a listing which shows **all the possible values (or intervals)** of the data.
- It also tells you how often each value occurs **(frequency)**.
- Often referred to as **probability distributions**.

Ex: Ratings Play Store

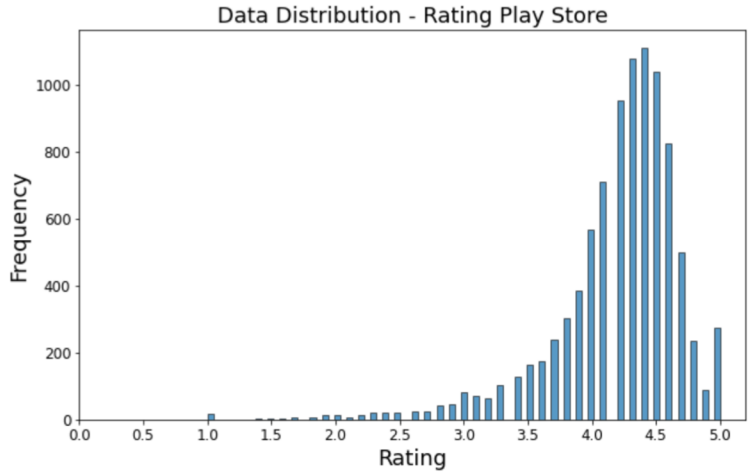


Basic Concepts

Data Distribution

- Distribution of **individual data points** from a dataset.
- It is a function or a listing which shows **all the possible values (or intervals)** of the data.
- It also tells you how often each value occurs **(frequency)**.
- Often referred to as **probability distributions**.

Ex: Ratings Play Store



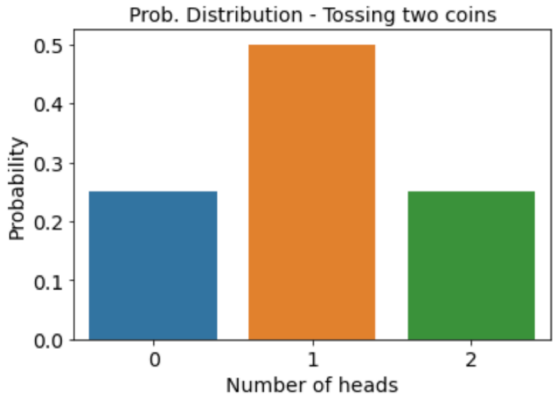
Probability Distribution

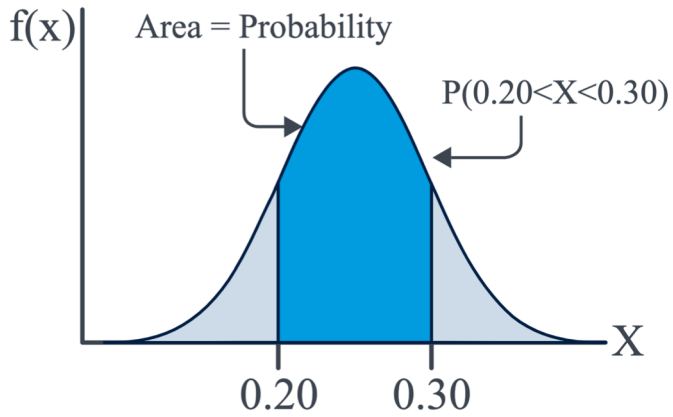
- Mathematical function that gives the **probabilities of occurrence** of different possible **outcomes** for an **experiment**.

Ex: Toss a coin twice

Sample Space: S = {HH, HT, TH, TT}

Event: Prob. of getting heads





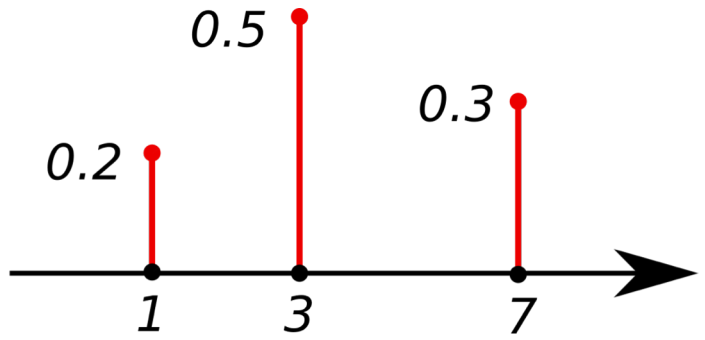
Basic Concepts

Probability Mass Function (PMF)

- The **probability distribution** of a **discrete random variable.**

Properties:

- $\boldsymbol{P(X = x) = f(x)}$
  - Prob. of the random variable X at a **specific x**
- All probabilities are positive: $\boldsymbol{P(x) \geq 0}$
- Any event in the distribution has: $\boldsymbol{0 \leq P(x) \leq 1}$
- The sum of all probabilities is 1. So $\boldsymbol{\sum P(x) = 1}$



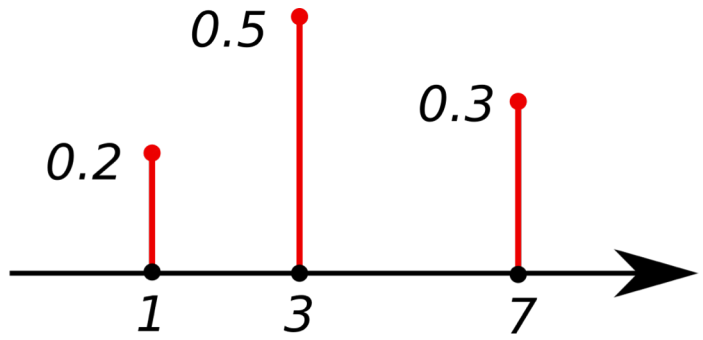
Basic Concepts

Probability Mass Function (PMF)

- The **probability distribution** of a **discrete random variable.**

Properties:

- $\boldsymbol{P(X = x) = f(x)}$
  - Prob. of the random variable X at a **specific x**
- All probabilities are positive: $\boldsymbol{P(x) \geq 0}$
- Any event in the distribution has: $0 \leq \boldsymbol{P(x) \leq 1}$
- The sum of all probabilities is 1. So $\boldsymbol{\sum P(x) = 1}$

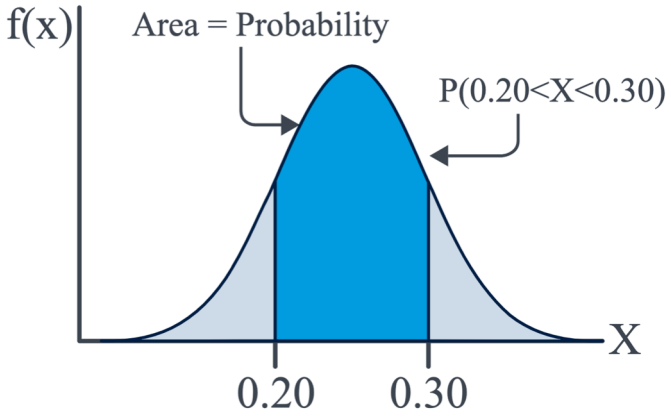


Probability Density Function (PDF)

- The **probability distribution** of a **continuous random variable.**

Properties:

- $\boldsymbol{P(X = x) = 0}$ (it is always zero)
- $\boldsymbol{P(a \leq X \leq b) = \int_a^b f(x)}$
- $\boldsymbol{f(x) \geq 0}$, for all $x \in \mathbb{R}$
- $\boldsymbol{\int_{-\infty}^{+\infty} f(x) = 1}$

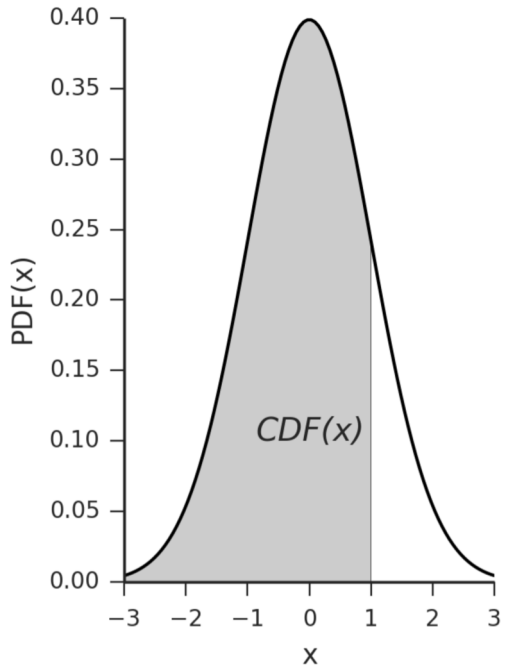


Basic Concepts

Cumulative Distribution Function (CDF)

- Gives the **cumulative value** from $-\infty$ up to a value **x** for a **random variable X (discrete or continuous)**
- It is the **probability function** that X will take a value **less than or equal to x**.

$$\boldsymbol{P(X \leq x)}, \text{ for all } x \in \mathbb{R}$$



$$\boldsymbol{P(X \leq 1)}$$

Basic Concepts

Expected Value

- A practical approach results in a **data/frequency distribution** and a **mean value**
- A theoretical approach results in a **probability distribution** and an **expected value**.

$$\boldsymbol{E(X) = \sum_{x \in S} xP(X = x)}$$

S is the **sample space**

Ex:

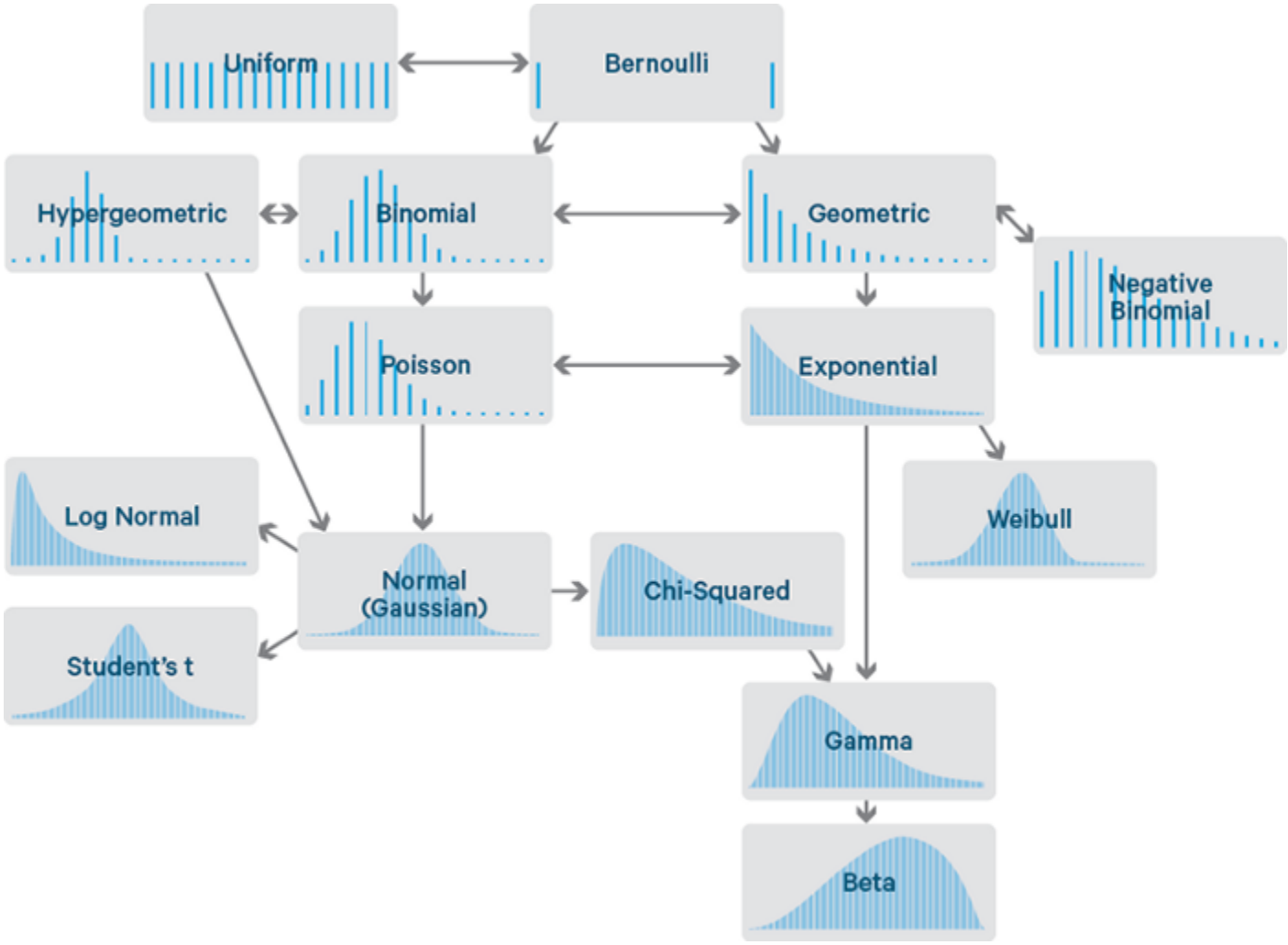
Suppose a **discrete random variable X** with the following sample space and PMF:

$$X = \begin{cases} 1\ with\ probability\ 1/8 \\ 2\ withprobability\ 3/8 \\ 3\ withprobability\ 1/2 \end{cases}$$

The **expected value** for X is:

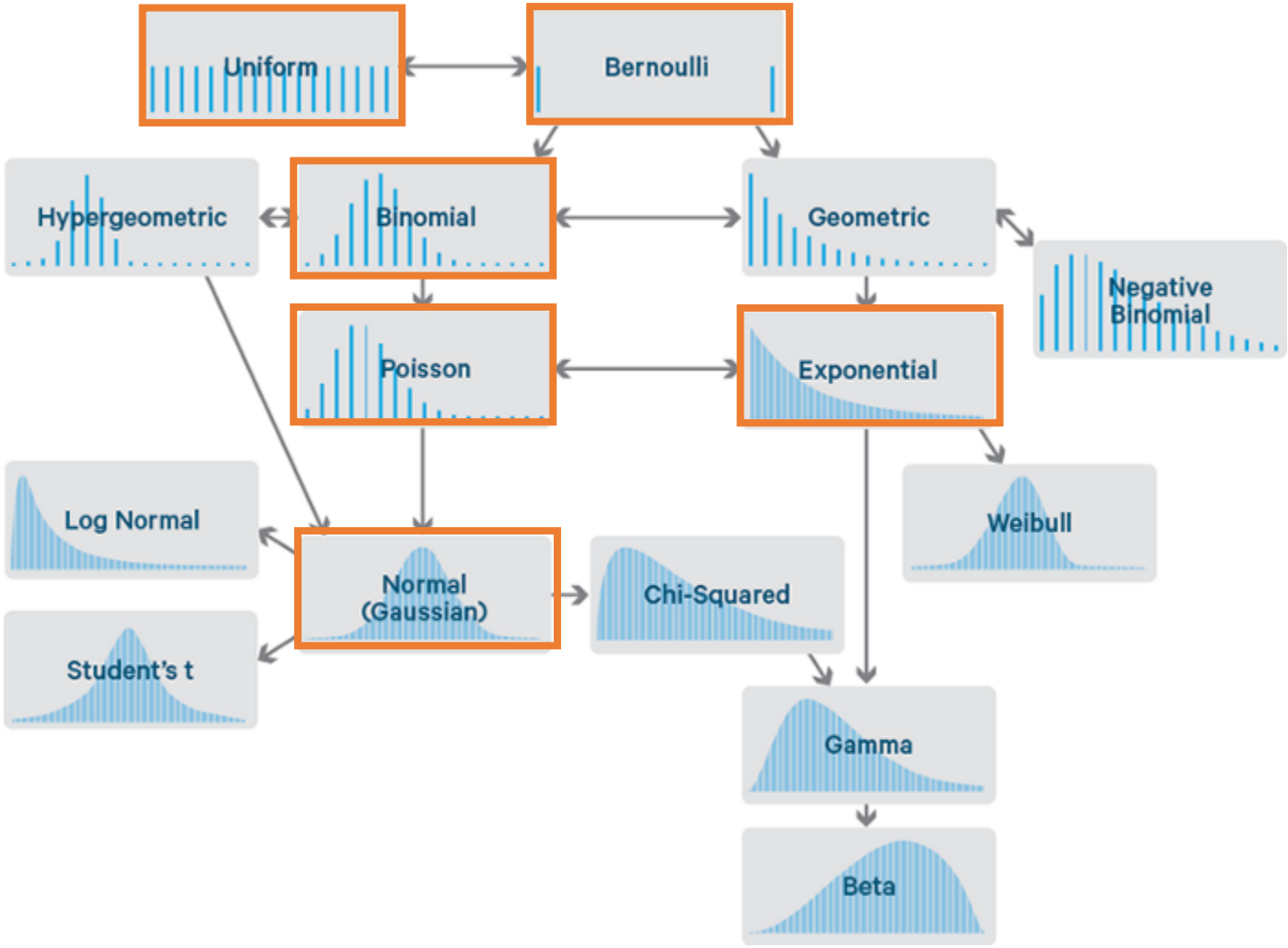
$$\boldsymbol{E(X) = 1 \cdot \left(\frac{1}{8}\right) + 2 \cdot \left(\frac{3}{8}\right) + 3 \cdot \left(\frac{1}{2}\right) = 2.375}$$

Probability Distributions



https://medium.com/diogo-menezes-borges/introduction-to-statistics-for-data-science-7bf596237ac6

Probability Distributions



https://medium.com/diogo-menezes-borges/introduction-to-statistics-for-data-science-7bf596237ac6

1. Bernoulli Distribution

- The simplest distribution.
- Only **two possible outcomes**:
  - 1 (success)
  - 0 (failure)
- A **single trial**.

$$P(X = x) = \begin{cases} 1 - p, & x = 0 \\ p, & x = 1 \end{cases}$$

$$\mu = np$$

$$\sigma = \sqrt{pq}$$

p: probability of success.

1. Bernoulli Distribution

$$P(X = x) = \begin{cases} 1 - p, & x = 0 \\ p, & x = 1 \end{cases}$$

$$\mu = np$$

$$\sigma = \sqrt{pq}$$

p: probability of success.

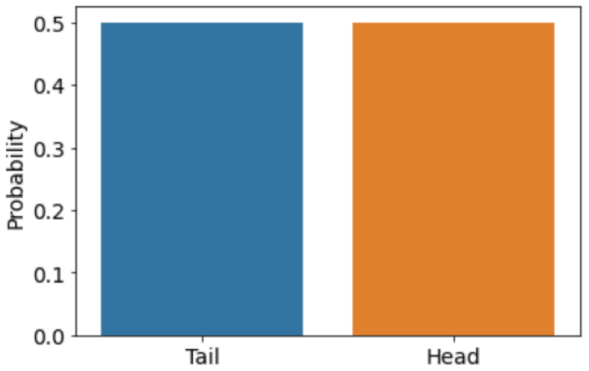
Ex 1: Tossing a coin.

X = {1 (head), 0 (tail)}

$$P(X = x) = \begin{cases} 0.5, & x = 0 \\ 0.5, & x = 1 \end{cases}$$

$$\mu = 0.5$$

$$\sigma = 0.25$$



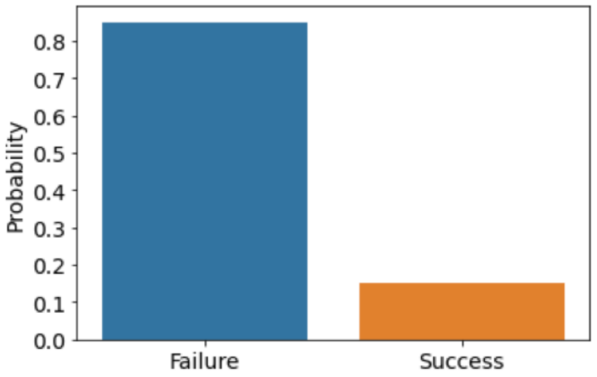
Ex 2: Samuka scoring a goal.

X = {1 (success), 0 (failue)}

$$P(X = x) = \begin{cases} 0.85, & x = 0 \\ 0.15, & x = 1 \end{cases}$$

$$\mu = 0.15$$

$$\sigma = 0.1275$$



2. Binomial Distribution

- It is the **frequency distribution** of the **number of successes (X)** in a given **number of trials (n)** with specified **probability (p) of success** in each trial.
- Ex: getting two heads when tossing three coins, nº of defective PCs in a shipment, nº of girls in a family, etc.

Binomial experiment

1. Fixed number of **identical trials**;
2. Trials are **independent** of each other;
3. Only **two** outcomes are possible (e.g., success and failure, head and tail, true and false, etc);
4. Fixed probability of **success**: **p** (consequently, the probability of **failure** is **q = 1 - p**)

2. Binomial Distribution

$$C_x^n \binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

$$\mu = np$$

$$\sigma = \sqrt{npq}$$

n: number of trials

p: probability of success.

$q = (1 - p)$: probability of failure.

Exercise 1

In an admission test for the Data Science specialization, **10 questions** with **3 possible choices** in each question.

Each question scores equally. Suppose that a candidate have not been prepared for the test. She decided to guess all answers.

Let the test has the **maximum score of 10** and **cut-off score of 5** for being approved for the next stage.

Provide the probability that this candidate will **get 5 questions right**, and the probability that she will **advance to the next stage of the test**.

Exercise 2

In the last World Chess Championship, **the proportion of female participants was 60%.**

The total of teams, with 12 members, in this year's championship is 30.

According to these information, **how many teams should be formed by 8 women?**

3. Poisson Distribution

- Used to describe the **number of occurrences** within a **specific period of time** or **space**.
- Some more examples are:
    - The number of emergency calls recorded at a hospital in a day.
    - The number of thefts reported in an area on a day.
    - The number of customers arriving at a salon in an hour.
    - The number of suicides reported in a particular city.
    - The number of printing errors at each page of the book.

Poisson experiment

A distribution is called **Poisson distribution** when the following assumptions are valid:

1. The **probability of success** is the sample over at the whole interval.
2. Any **successful event should not influence** the outcome of another successful event.
    - The nº of occurrences at a given interval is **independent** from the nº of occurrences at other intervals.
3. The **probability of success** (a given occurrence) is the same at intervals with **equal length**.
4. The **probability of success** in an interval **approaches zero** as the **interval becomes smaller**.

3. Poisson Distribution

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, x = 0, 1, 2, \ldots$$
$$\mu = \lambda$$
$$\sigma = \sqrt{\lambda}$$

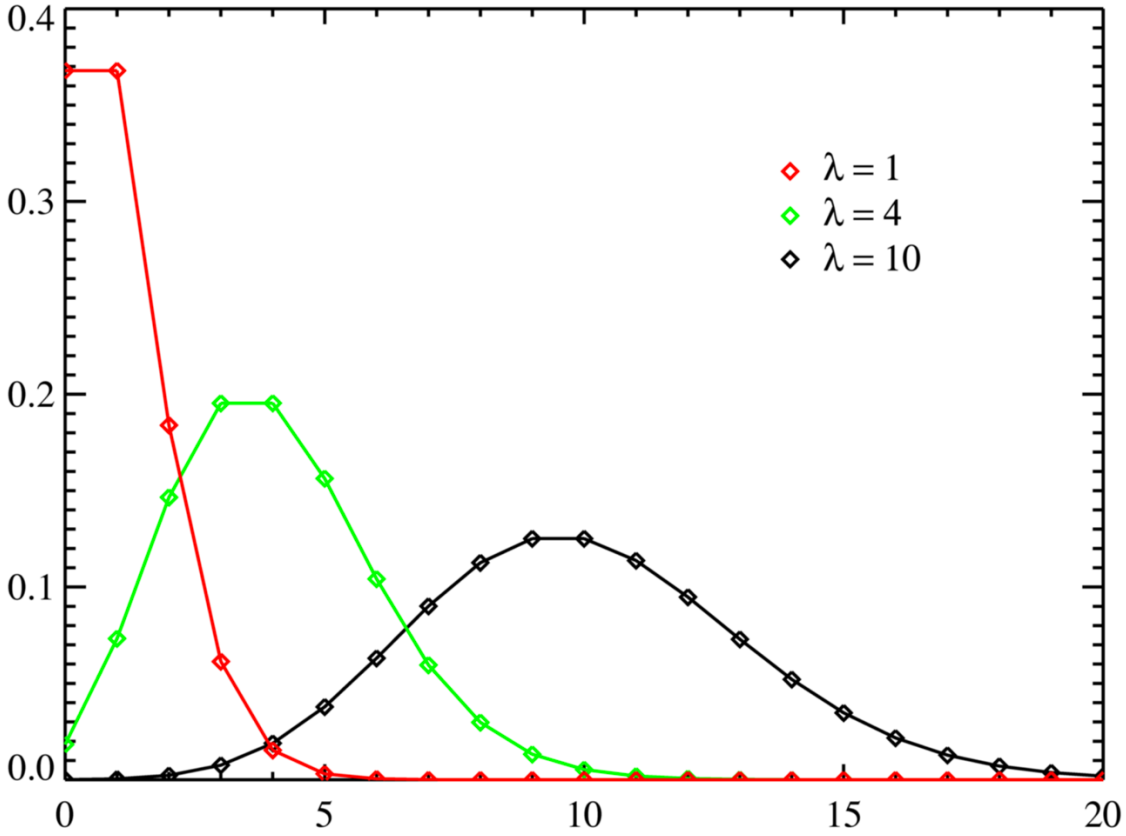
e: 2.71828…

λ: mean nº of occurrences/events (frequency) within a period of time.

x: nº of successes within the period of time.

3. Poisson Distribution

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, x = 0, 1, 2, \ldots$$
$$\mu = \lambda$$
$$\sigma = \sqrt{\lambda}$$



Exercise 1

A restaurant receives **20 orders per hour**. What is the chance that, at a given hour chosen at random, the restaurant will receive **15 orders**?

Exercise 2

Vehicles pass through a junction on a busy road at an average rate of 300 per hour.

Find the probability that none passes in a given minute.

What is the expected number passing in two minutes?

Find the probability that this expected number actually pass through in a given two-minute period.

Exercise 3

Suppose the average number of lions seen on a 1-day safari is 5. What is the probability that tourists will see fewer than four lions on the next 1-day safari?

Uniform Distribution

- Defines **equal probability** over a **given range** for a continuous (or discrete) distribution.

A variable X is said to be **uniformly distributed** if the density function is:

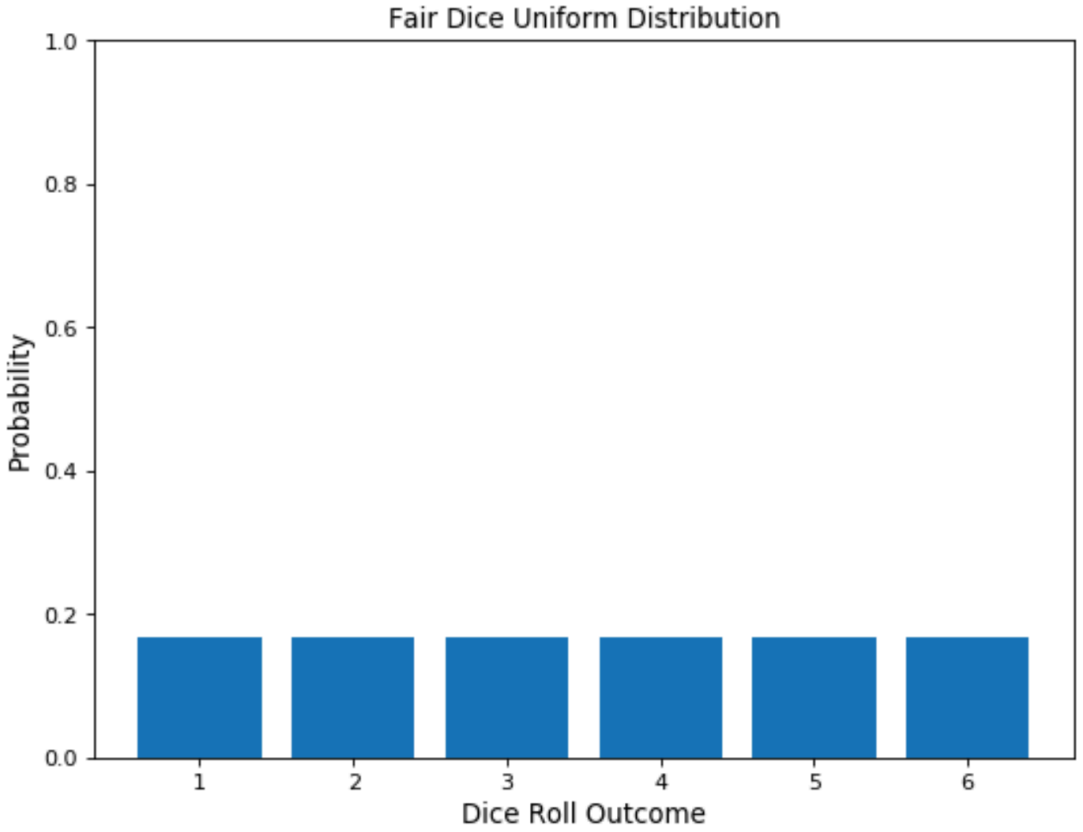
$$f(x) = \frac{1}{b-a}, for\ a \leq x \leq b$$

$$\mu = \frac{(a+b)}{2}$$

$$\sigma = \sqrt{\frac{(b-a)^2}{12}}$$

Exercise 1

Tossing a fair dice.



Exercise 2

The number of bouquets sold daily at a flower shop is **uniformly distributed** with a maximum of 40 and a minimum of 10.

Calculate the probability that the daily sales will fall **between 15 and 30**.

Solution

Exercise 2

The number of bouquets sold daily at a flower shop is **uniformly distributed** with a maximum of 40 and a minimum of 10.

Calculate the probability that the daily sales will fall **between 15 and 30**.

Solution

$$\boldsymbol{P(X = x) = f(x) = \frac{1}{(40 - 10)} = \frac{1}{30} = 0.03333\ldots}$$

$$\boldsymbol{P(15 \leq x \leq 30) = (30\ - 15) * 0.03333 = 0.5}$$

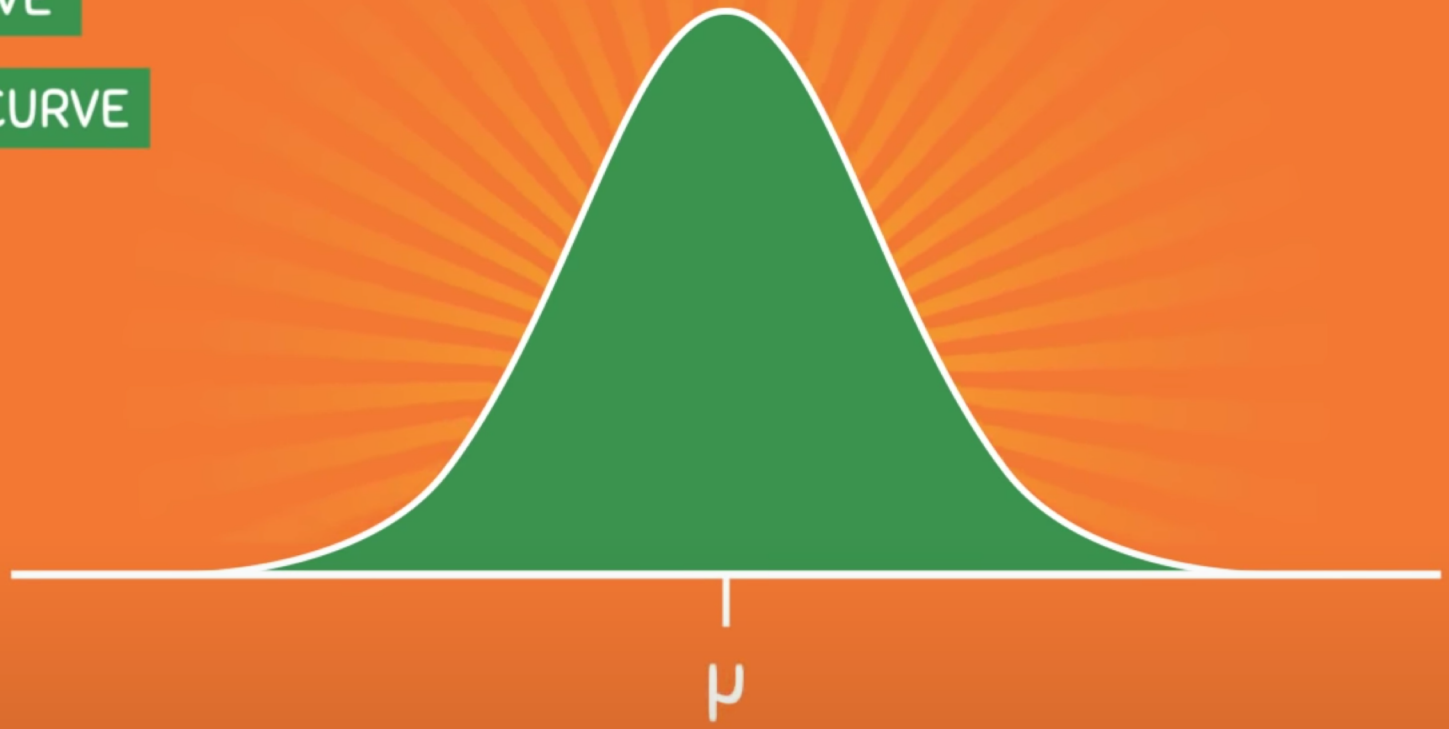
Normal Distribution

Content mainly extracted from the excellent channel **Simple Learning Pro**

https://www.youtube.com/watch?v=mtbJbDwqWLE

BELL CURVE

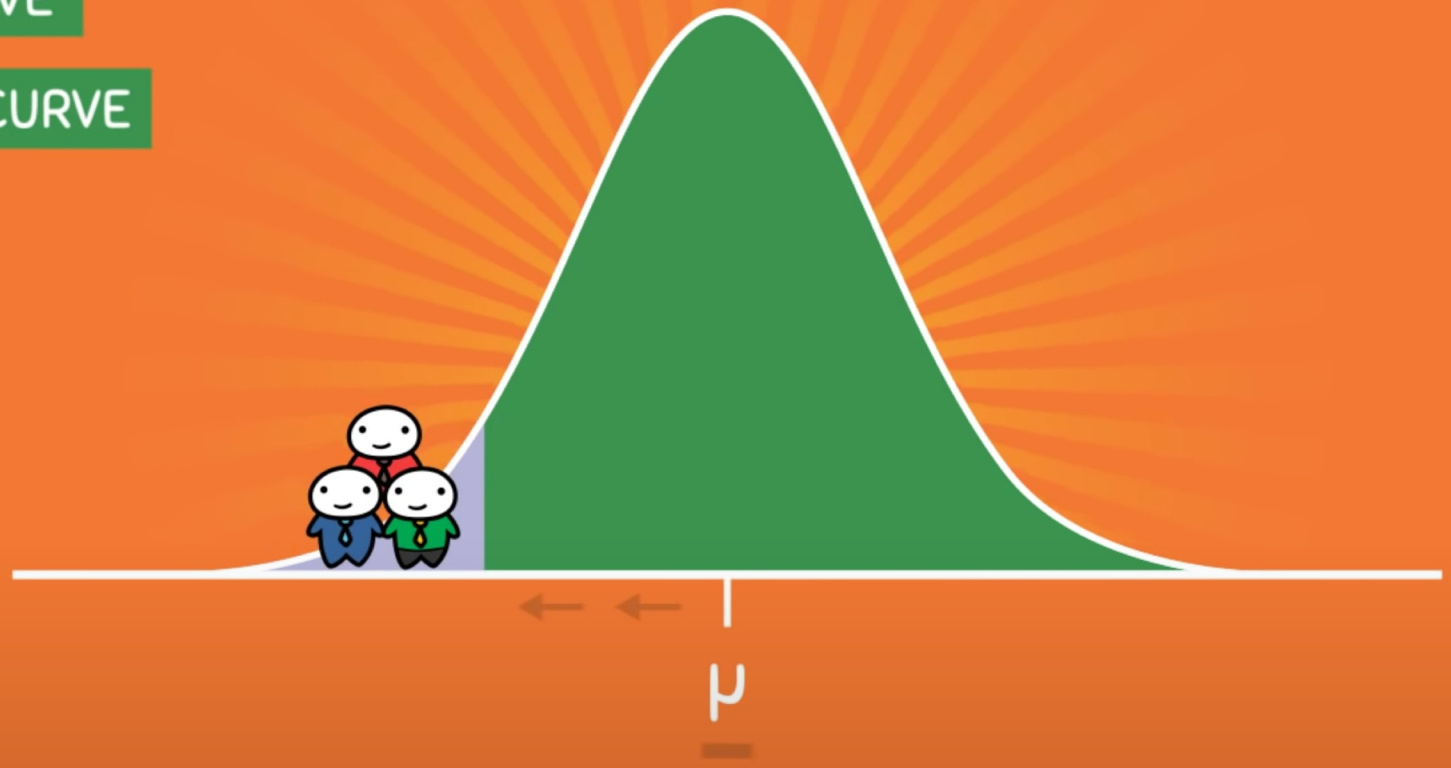
NORMAL CURVE



NORMAL DISTRIBUTION

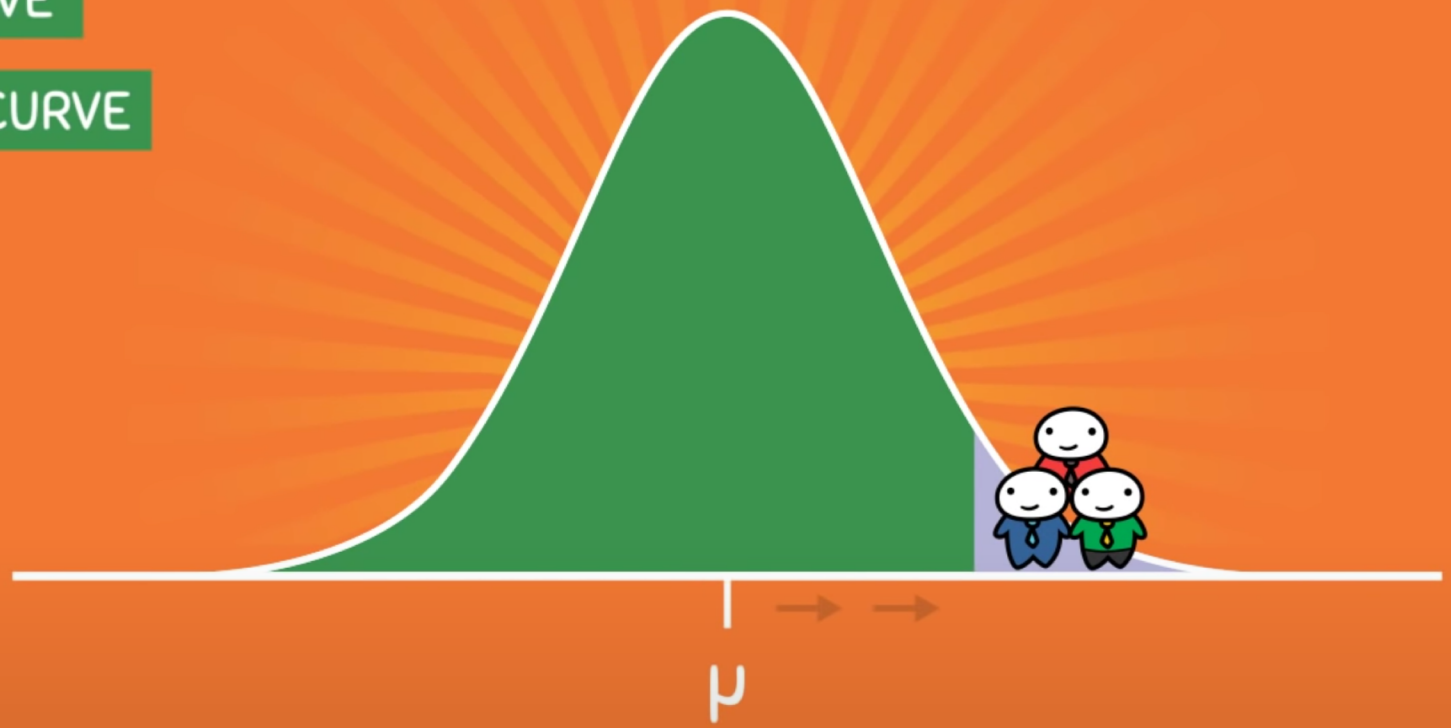
BELL CURVE

NORMAL CURVE



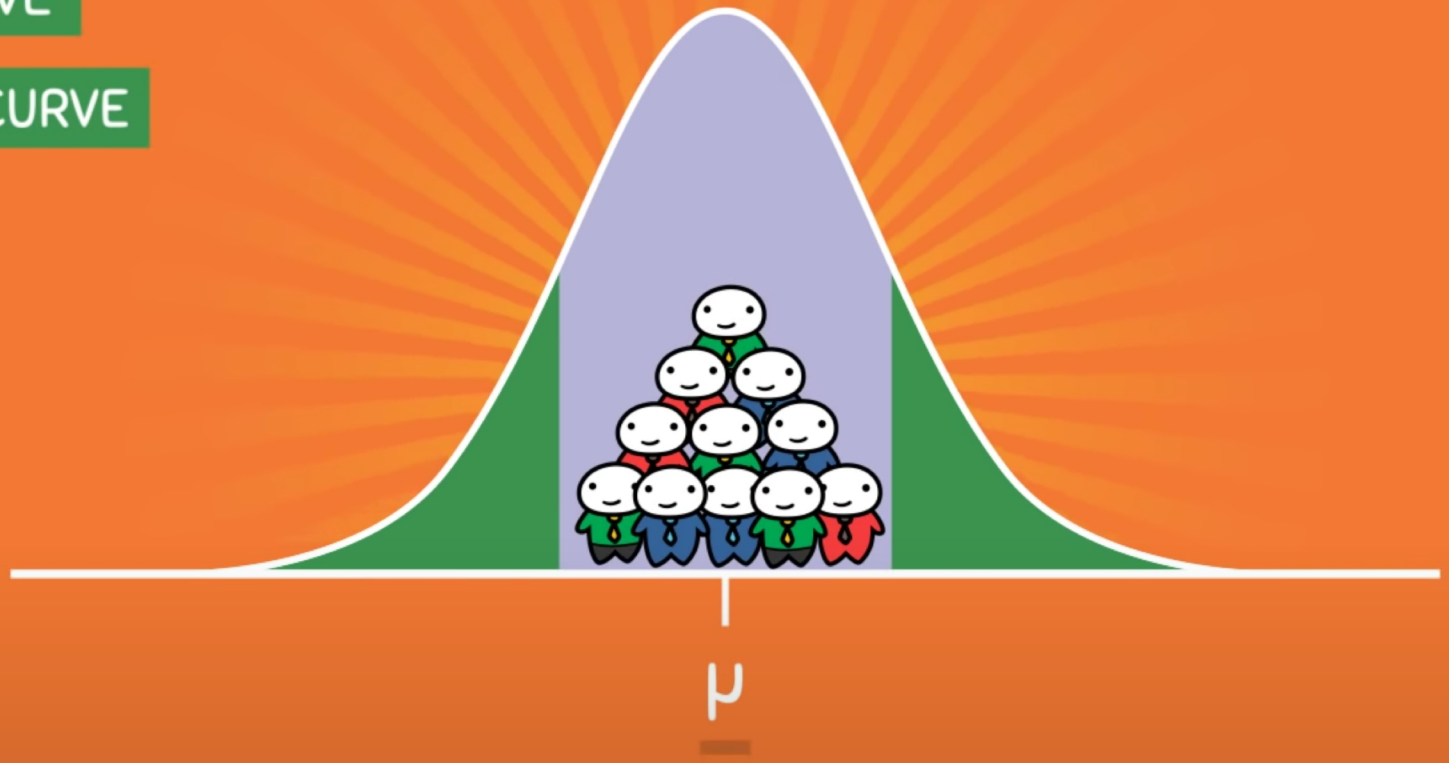
BELL CURVE

NORMAL CURVE



BELL CURVE

NORMAL CURVE



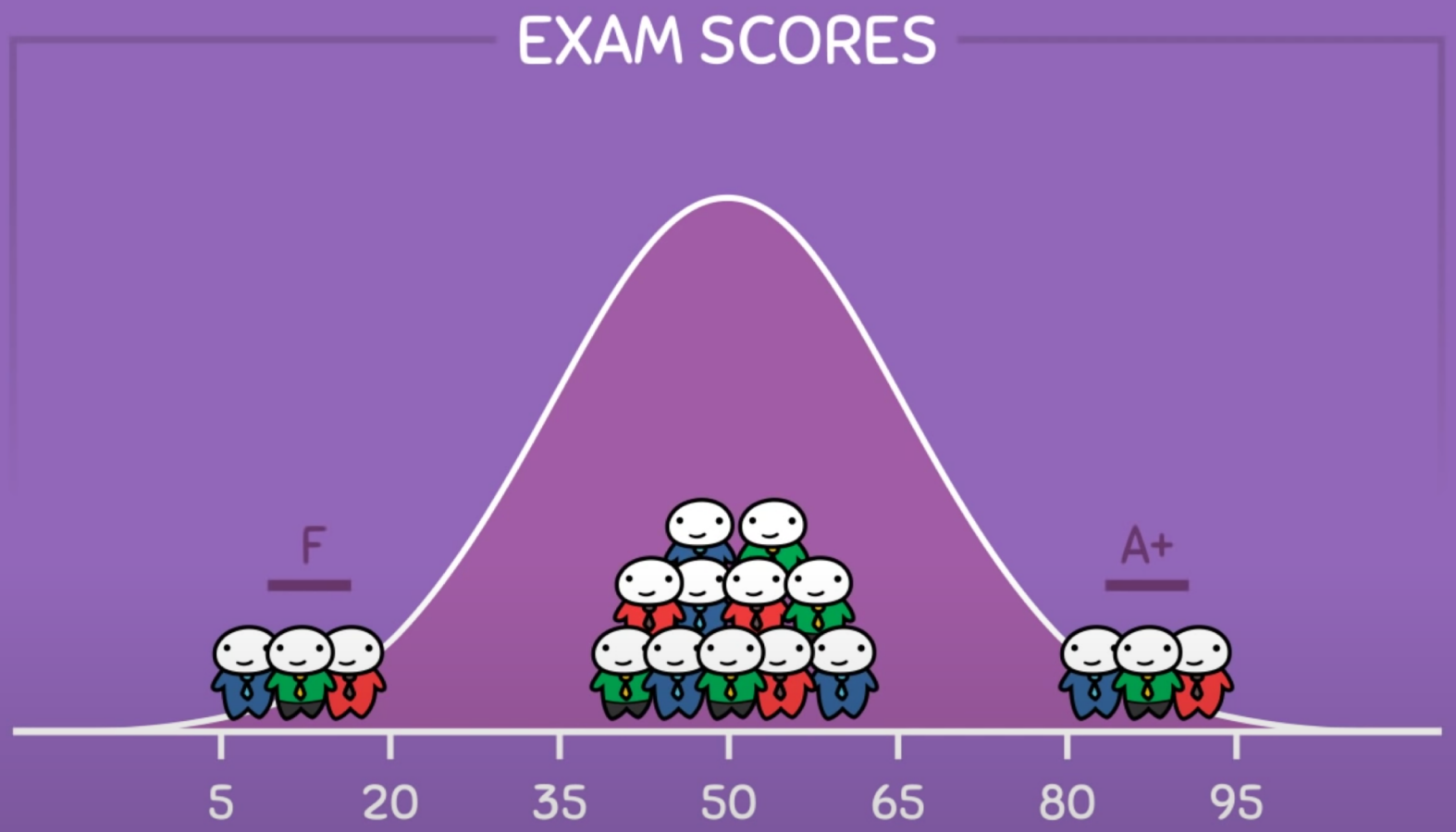
WEIGHT

HEIGHT

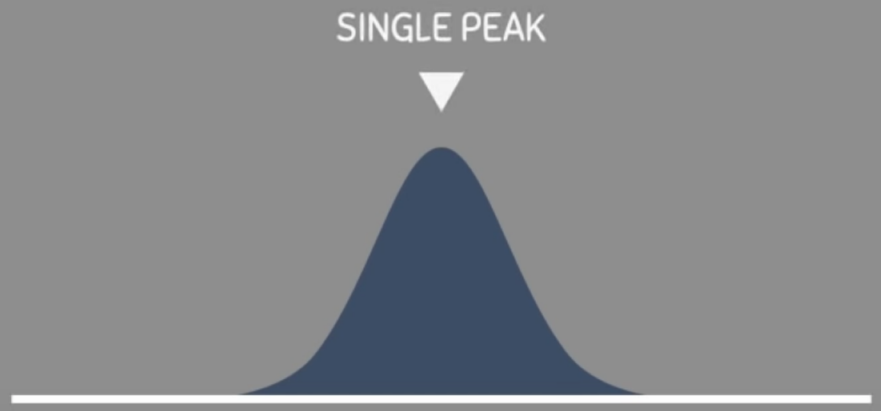
VOLUME

BLOOD PRESSURE

Income

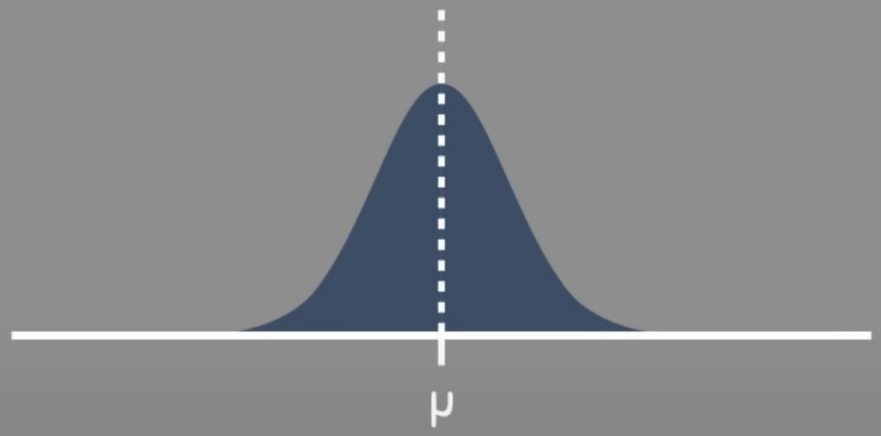

EXAM SCORES
F
A+
5
20
35
50
65
80
95

1 The normal distribution is unimodal



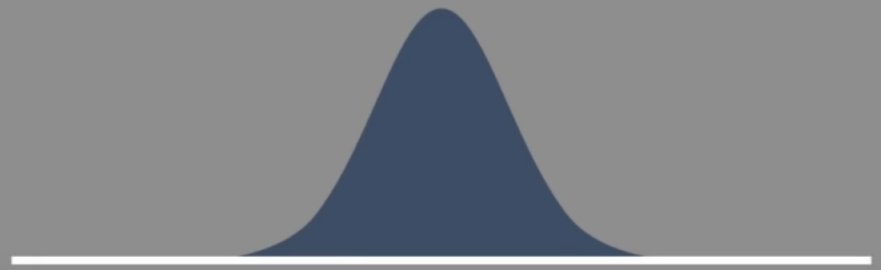
1. The normal distribution is unimodal

2. The normal curve is symmetric about its mean



The mean, median, and mode coincide

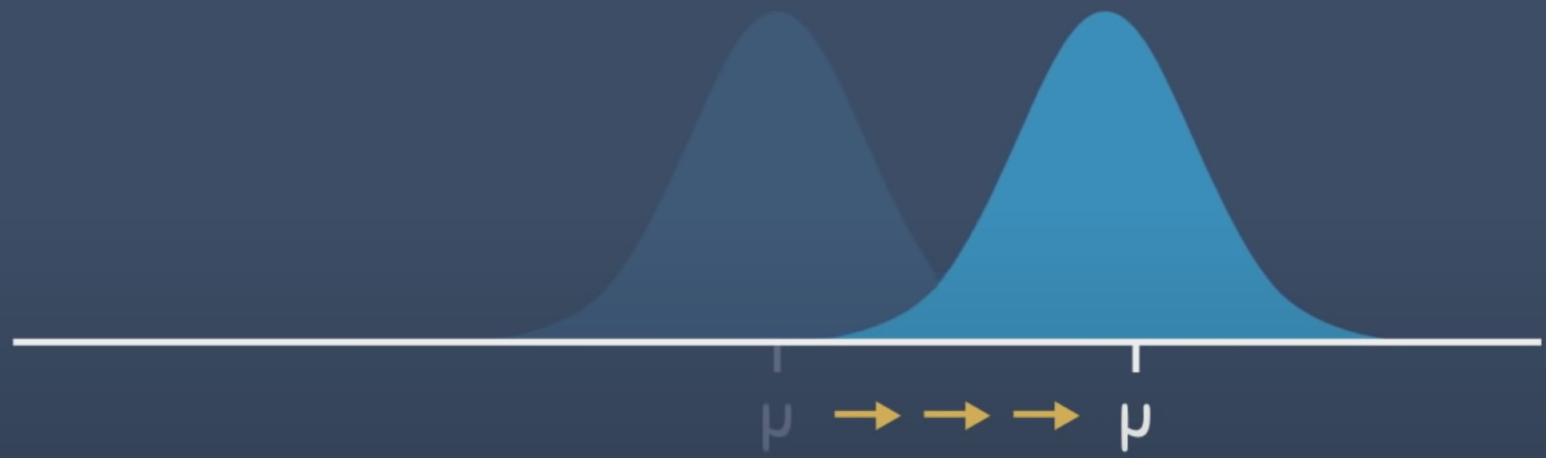
1. The normal distribution is unimodal
2. The normal curve is symmetric about its mean
3. The parameters μ and σ completely characterize the normal distribution



μ

POPULATION MEAN

CHARACTERIZES THE POSITION OF THE NORMAL DISTRIBUTION



σ

POPULATION
STANDARD DEVIATION

CHARACTERIZES THE SPREAD OF THE NORMAL DISTRIBUTION

σ = large

σ

POPULATION
STANDARD DEVIATION

CHARACTERIZES THE SPREAD OF THE NORMAL DISTRIBUTION

lower dispersion

σ = small

σ

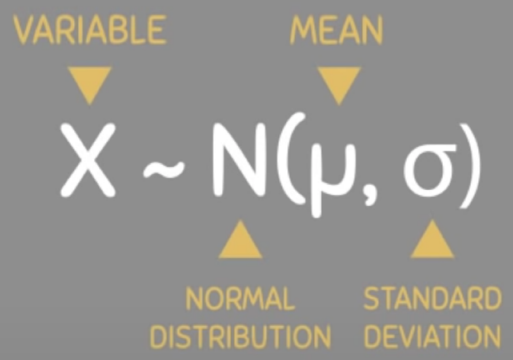
POPULATION
STANDARD DEVIATION

CHARACTERIZES THE SPREAD OF THE NORMAL DISTRIBUTION

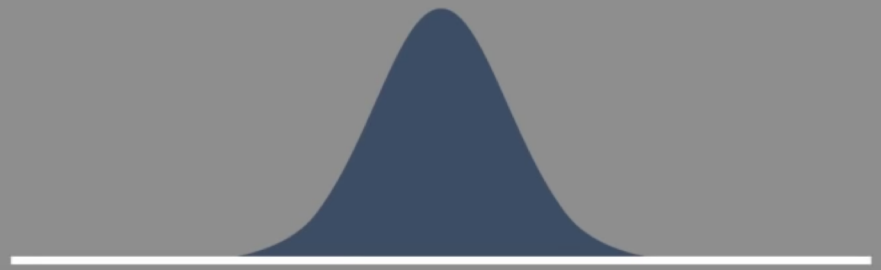
DENSITY CURVE

TOTAL AREA = 100%

1. The normal distribution is unimodal
2. The normal curve is symmetric about its mean
3. The parameters μ and σ completely characterize the normal distribution
4. $X \sim N(\mu, \sigma)$

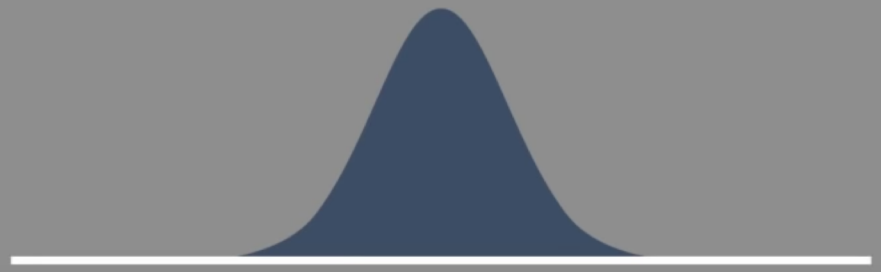


1. The normal distribution is unimodal
2. The normal curve is symmetric about its mean
3. The parameters μ and σ completely characterize the normal distribution
4. $X \sim N(\mu, \sigma)$
5. The ends of the curve tend to the infinite so that, theoretically, they never 'touch' the x-axis;



1. The normal distribution is unimodal
2. The normal curve is symmetric about its mean
3. The parameters μ and σ completely characterize the normal distribution
4. $X \sim N(\mu, \sigma)$
5. The ends of the curve tend to the infinite so that, theoretically, they never 'touch' the x-axis;

Distributions of sample statistics are often normally shaped (central limit theorem)



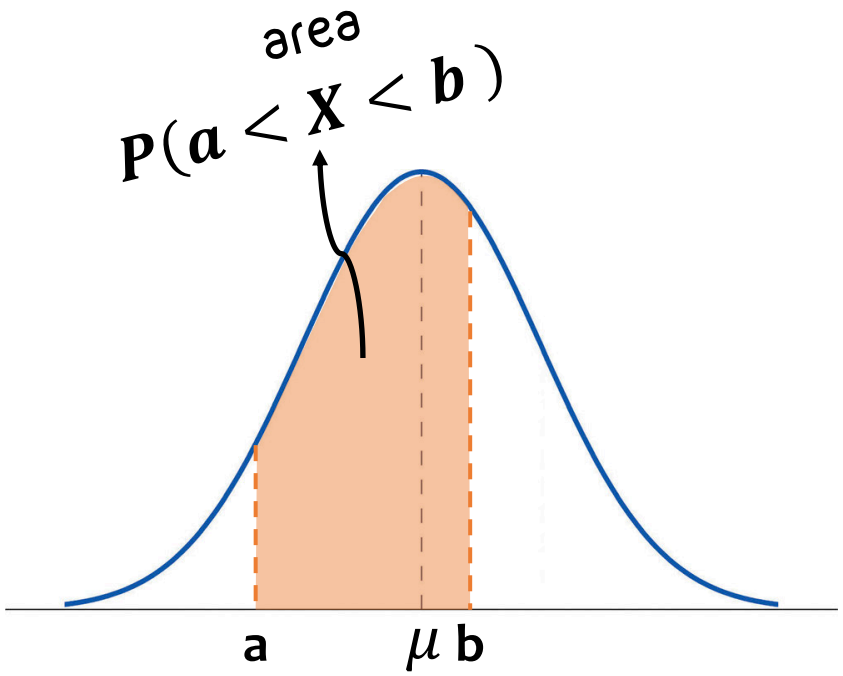
Normal distribution is a powerful tool in the development of mathematical formulas that approximate those distributions.

Probability function density (PDF)

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

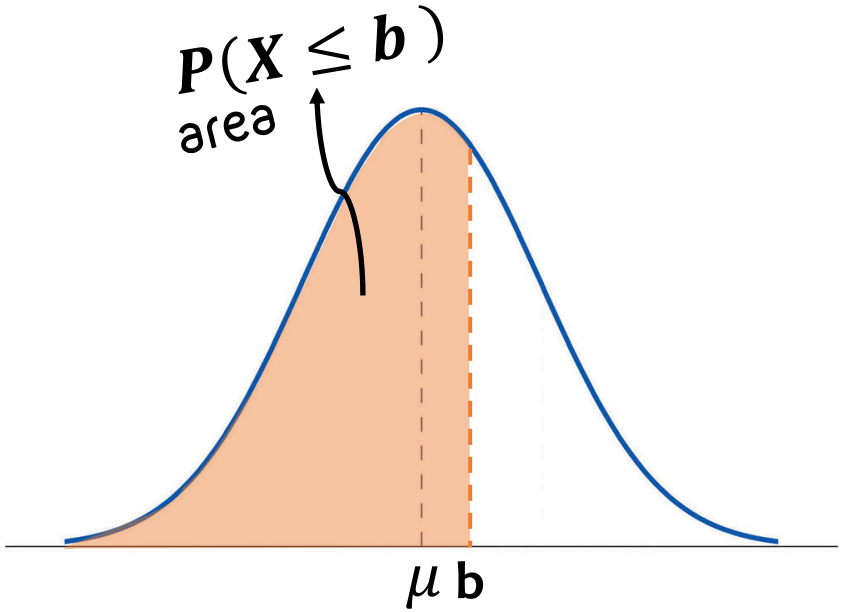
Probability (area under the curve)

$$P(a < X < b\ ) = \int_a^b f(x)dx = \int_a^b \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}\, dx$$

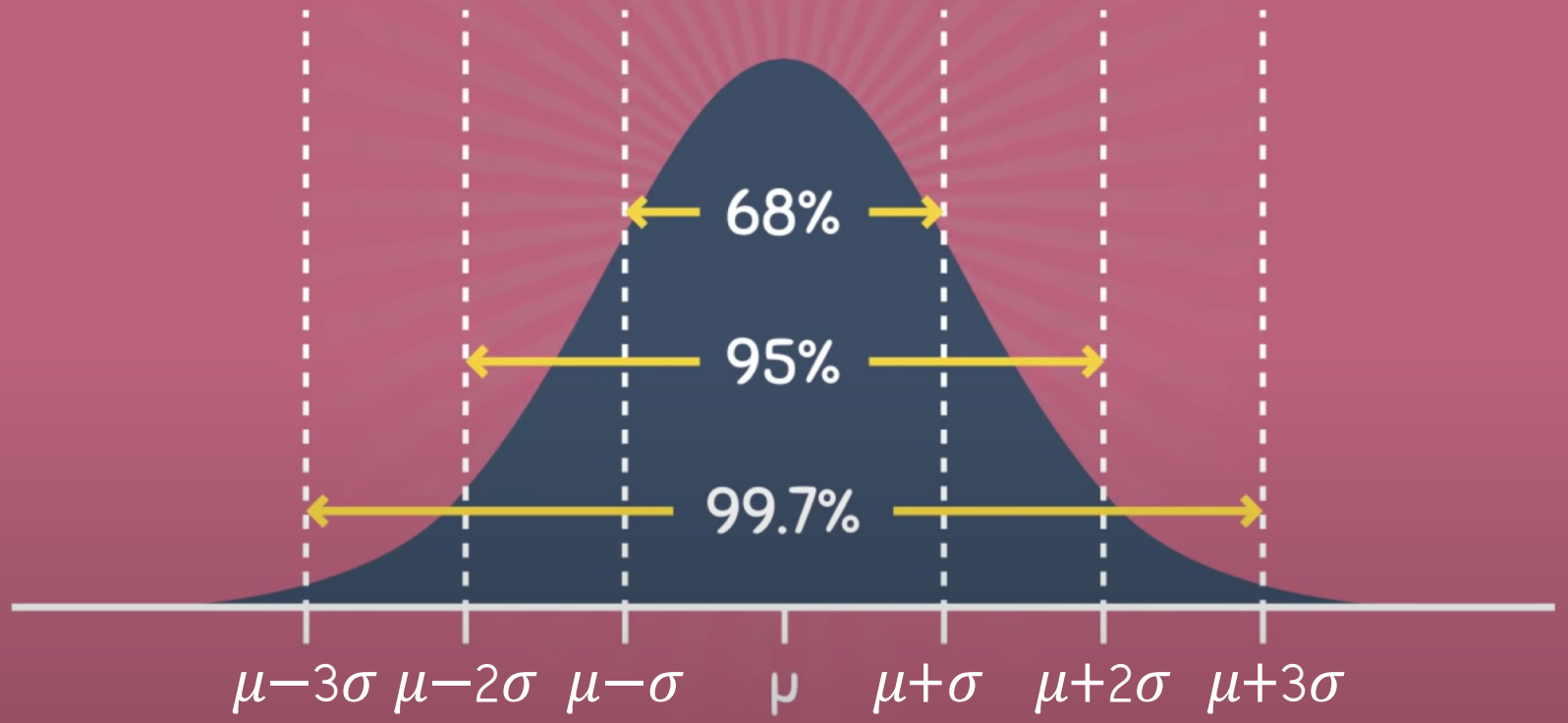


Cumulative Distribution Function (CDF)

$$P(X \leq b\ ) = \int_{-\infty}^b f(x)dx = \int_{-\infty}^b \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}\, dx$$



68–95–99.7 RULE



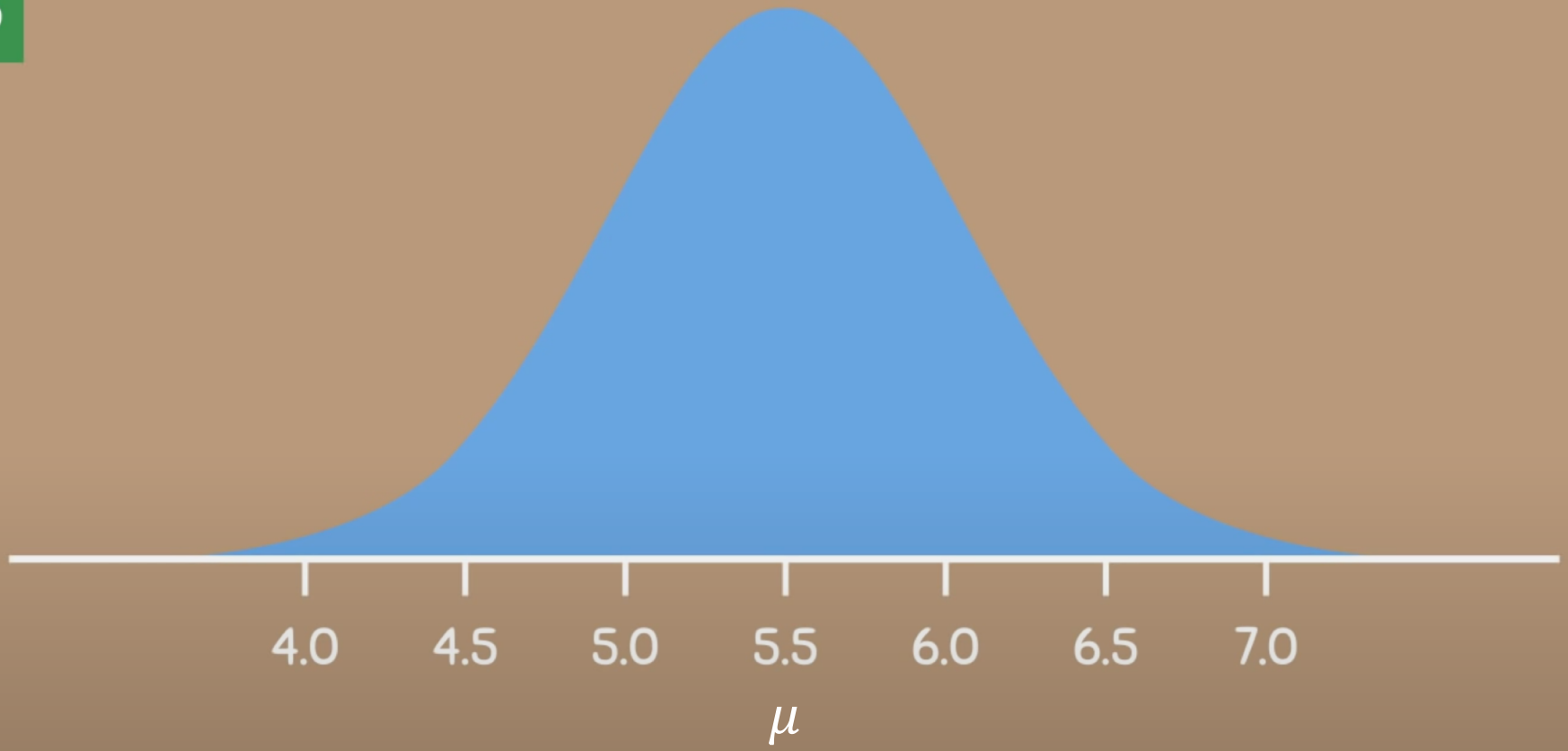
Example

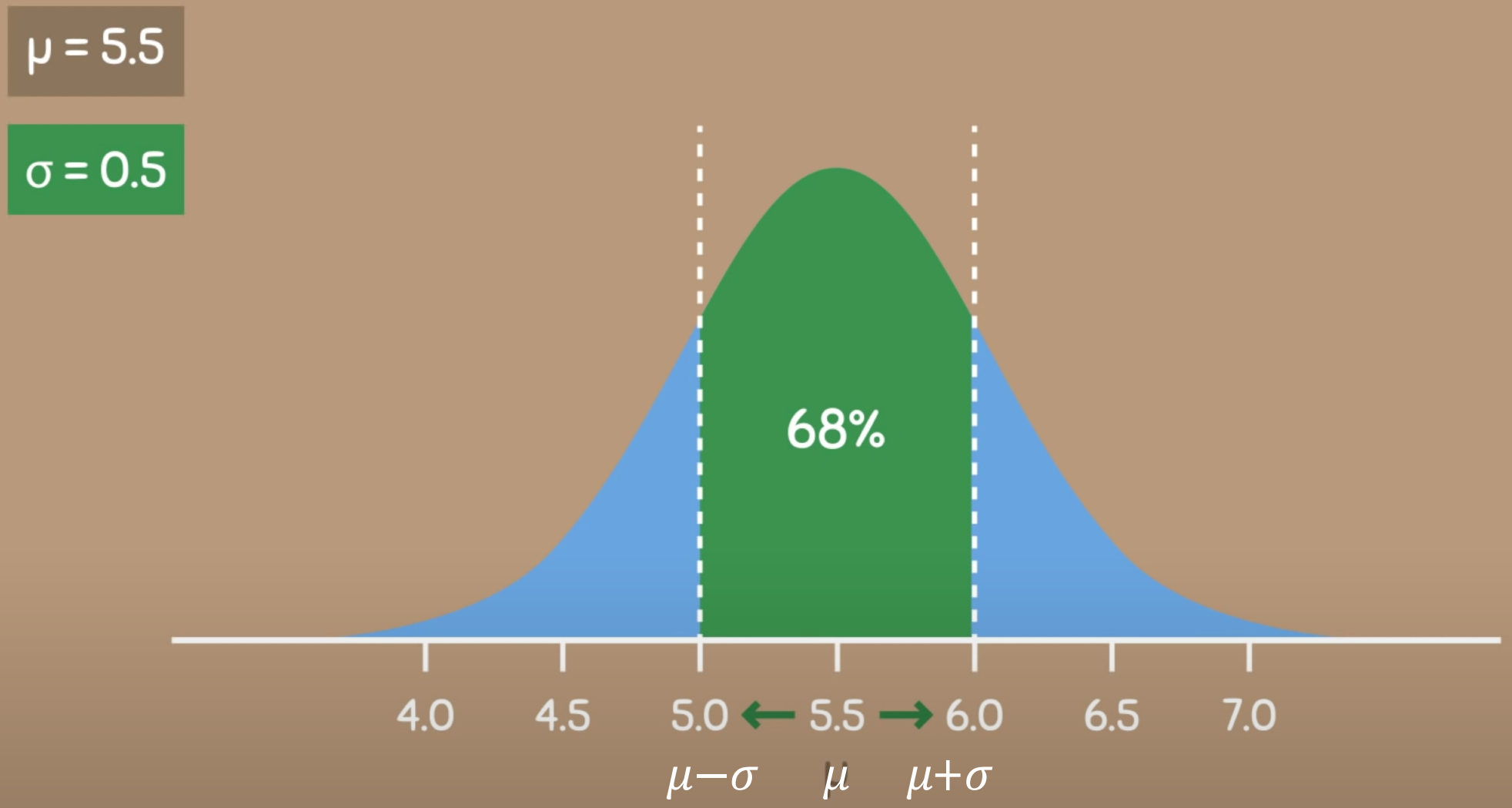
$X \sim N(\mu, \sigma)$

- X = HEIGHT (ft)
- $\mu = 5.5$
- $\sigma = 0.5$

$\mu = 5.5$

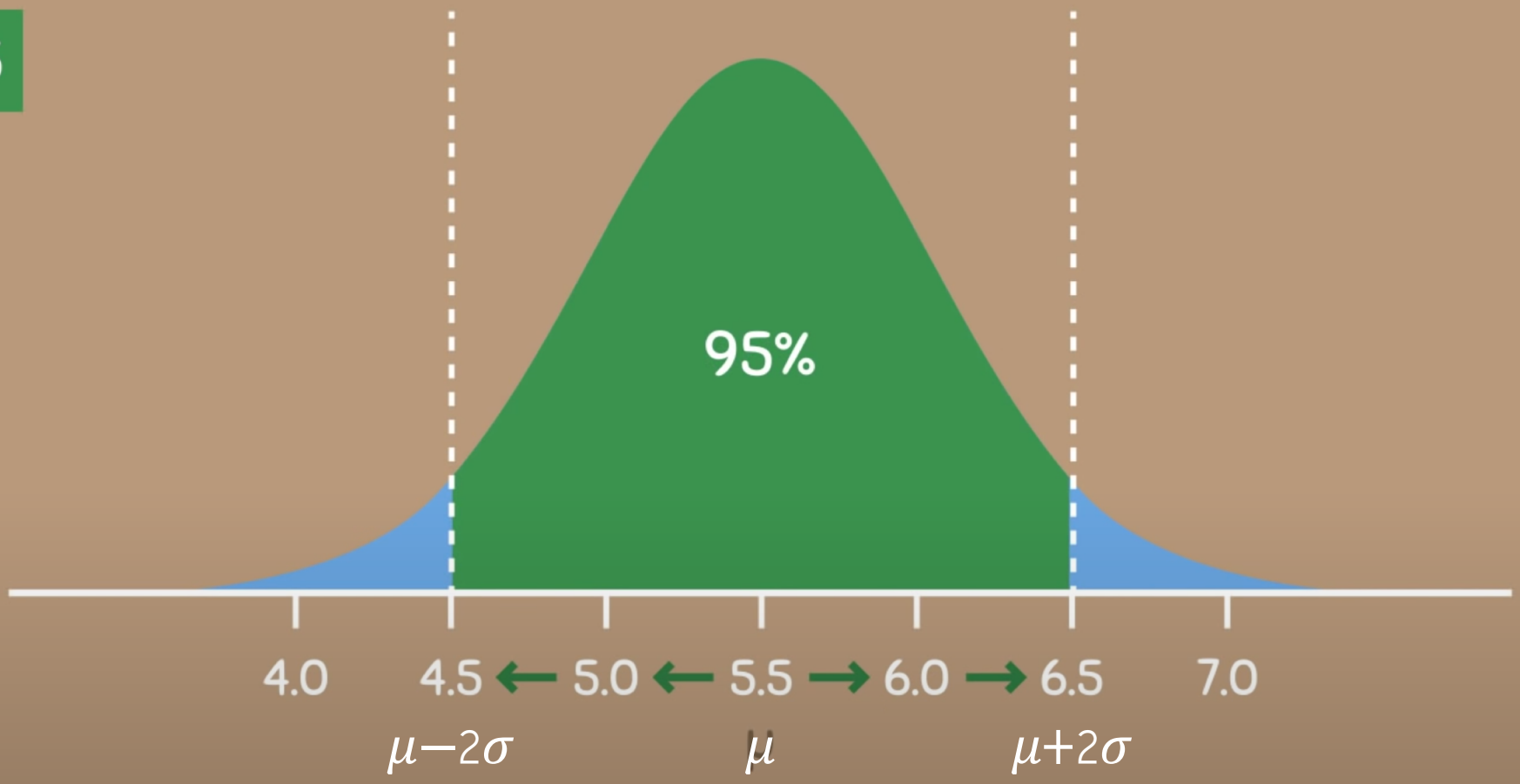
$\sigma = 0.5$

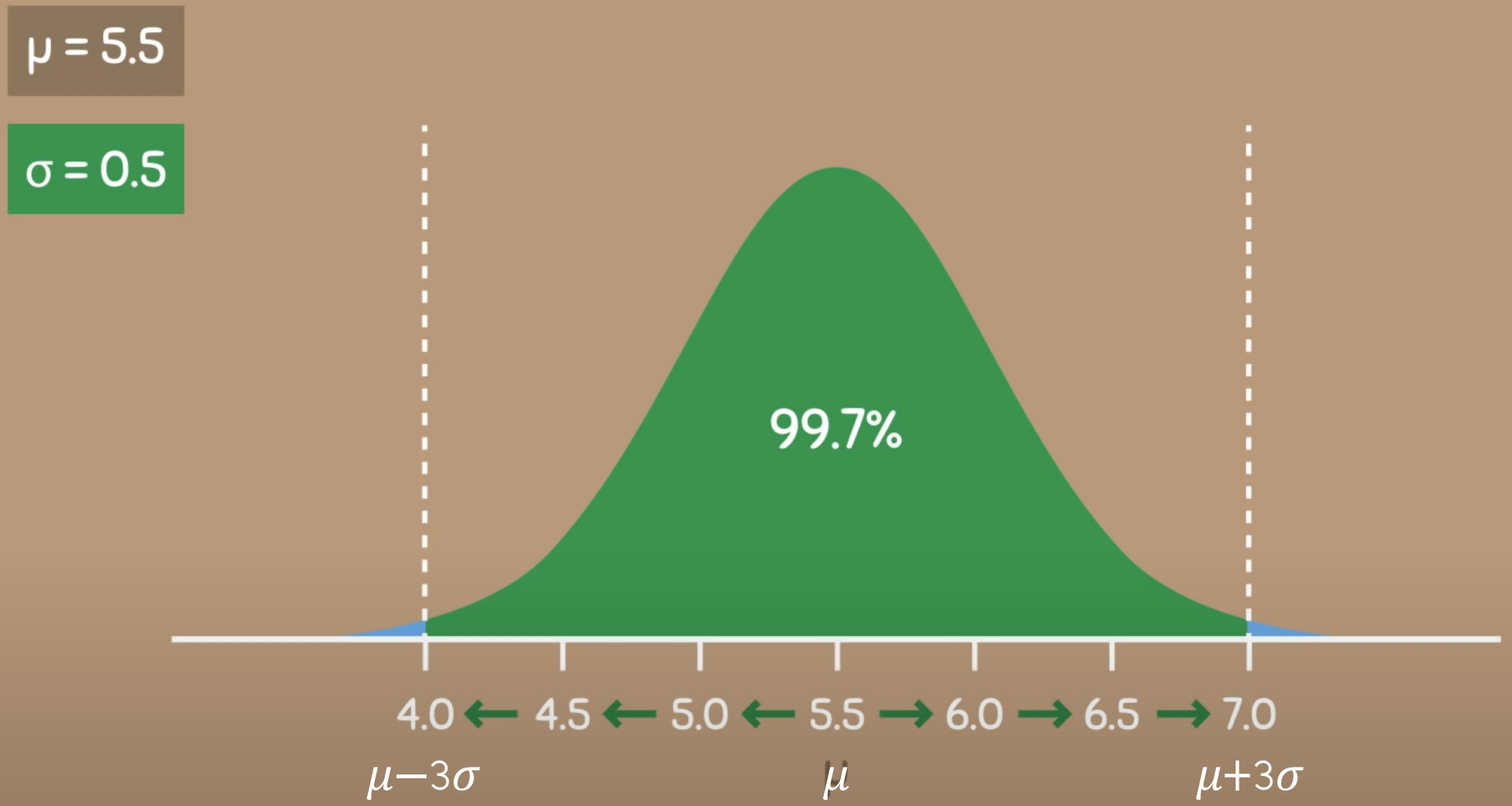



μ = 5.5
σ = 0.5
68%
4.0
4.5
5.0
5.5
6.0
6.5
7.0
μ−σ
μ
μ+σ

$\mu = 5.5$

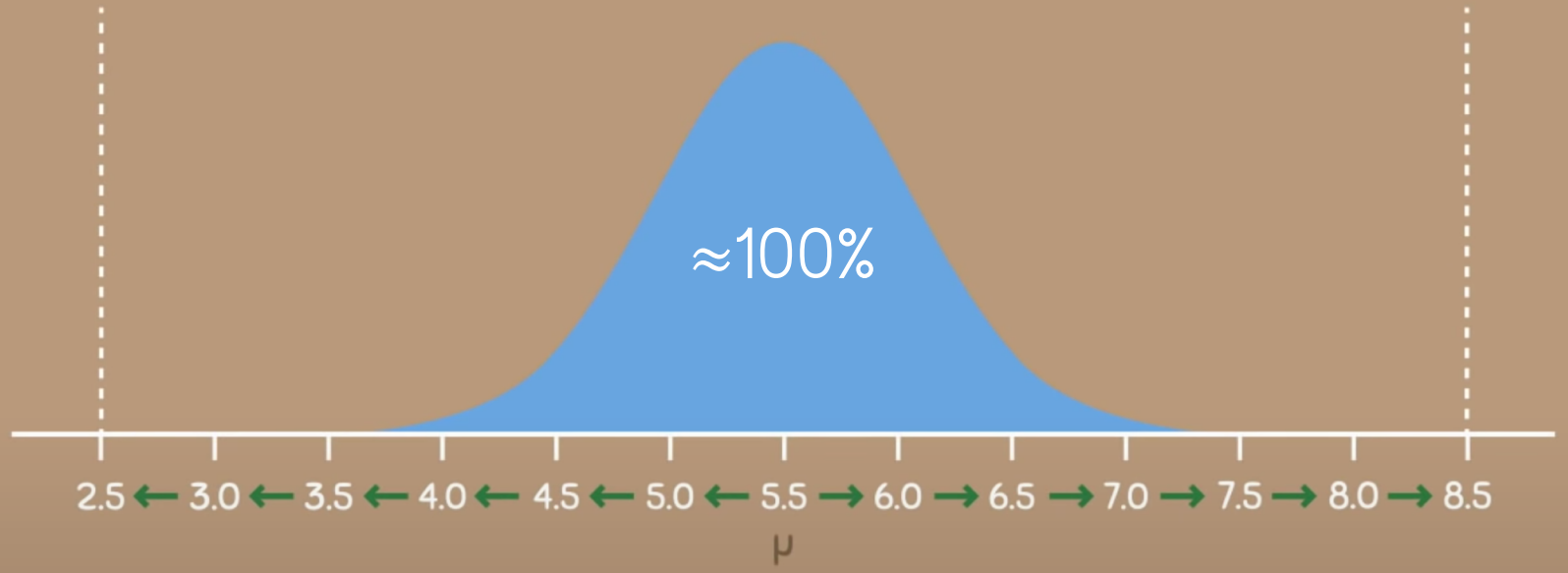
$\sigma = 0.5$




μ = 5.5
σ = 0.5
99.7%
4.0
4.5
5.0
5.5
6.0
6.5
7.0
μ−3σ
μ
μ+3σ

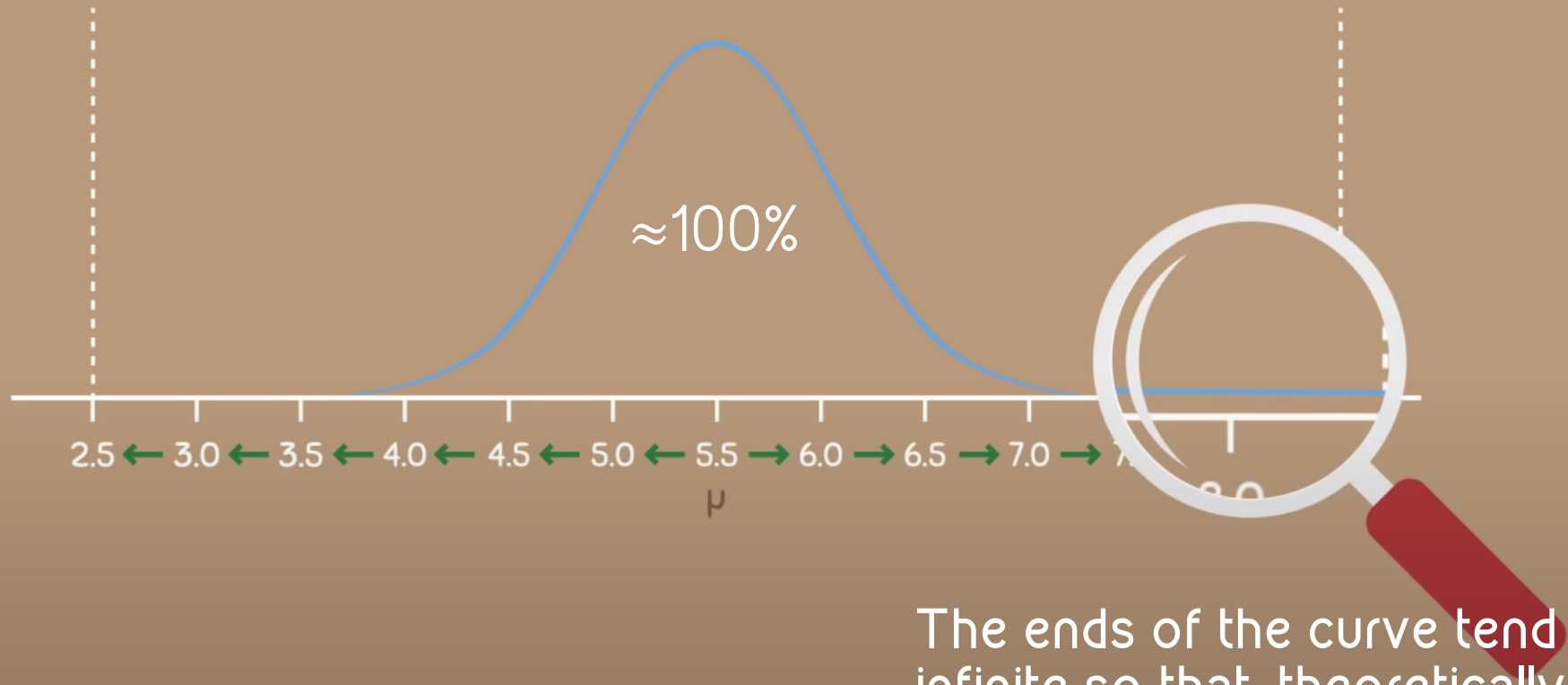
$\mu = 5.5$

$\sigma = 0.5$



$\mu = 5.5$

$\sigma = 0.5$



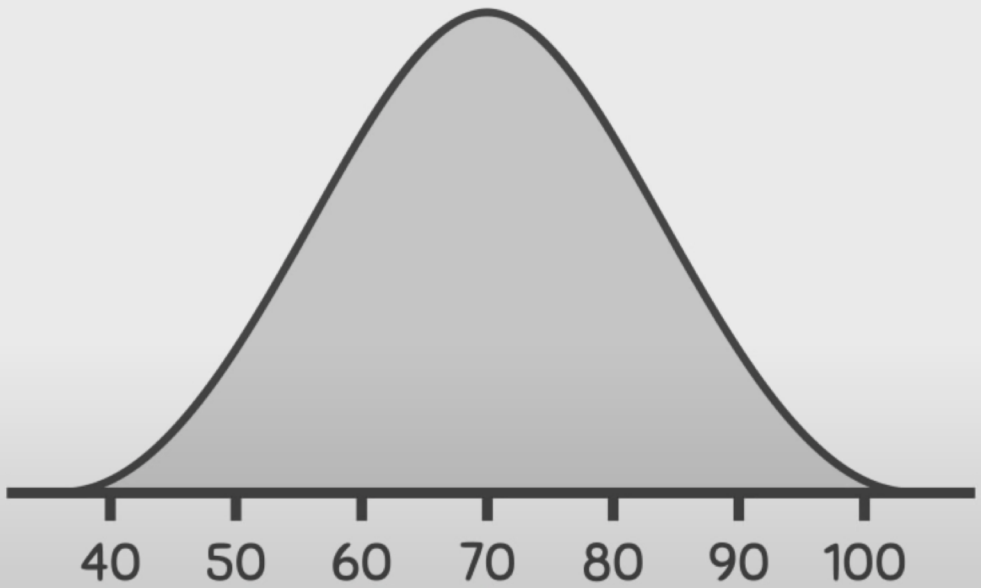
The ends of the curve tend to the infinite so that, theoretically, they never 'touch' the x-axis;

PRACTICE QUESTIONS

❶ The normal distribution below has a standard deviation of 10. Approximately what area is contained between 70 and 90?

$\mu = 70$

$\sigma = 10$

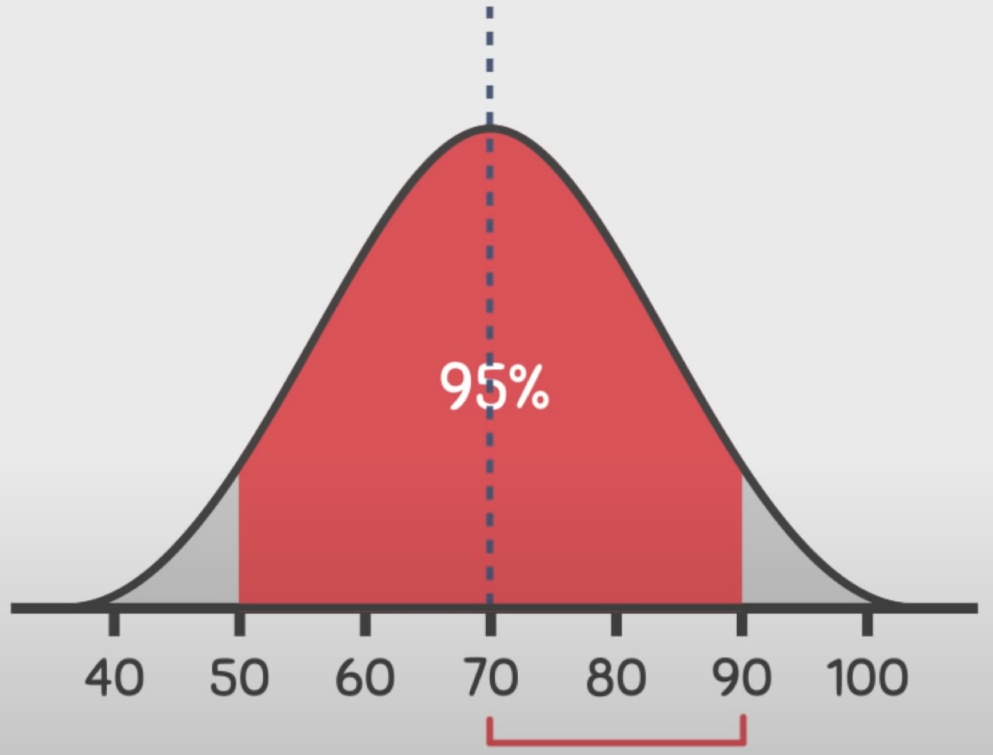


PRACTICE QUESTIONS

❶ The normal distribution below has a standard deviation of 10. Approximately what area is contained between 70 and 90?

$\mu = 70$

$\sigma = 10$

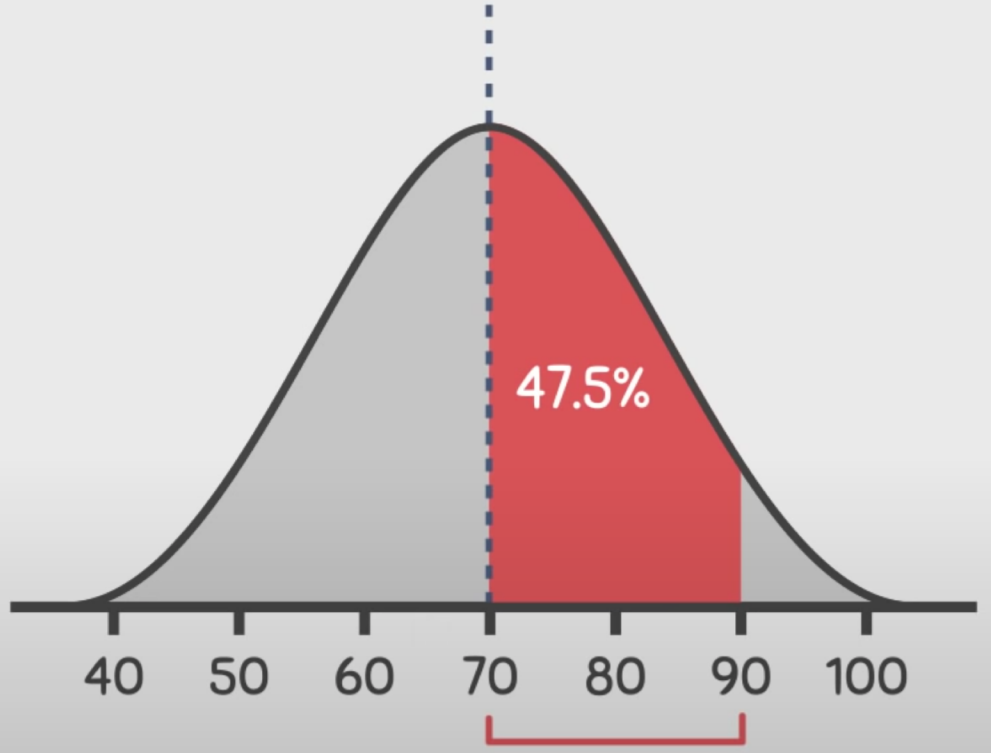


PRACTICE QUESTIONS

❶ The normal distribution below has a standard deviation of 10. Approximately what area is contained between 70 and 90?

$\mu = 70$

$\sigma = 10$

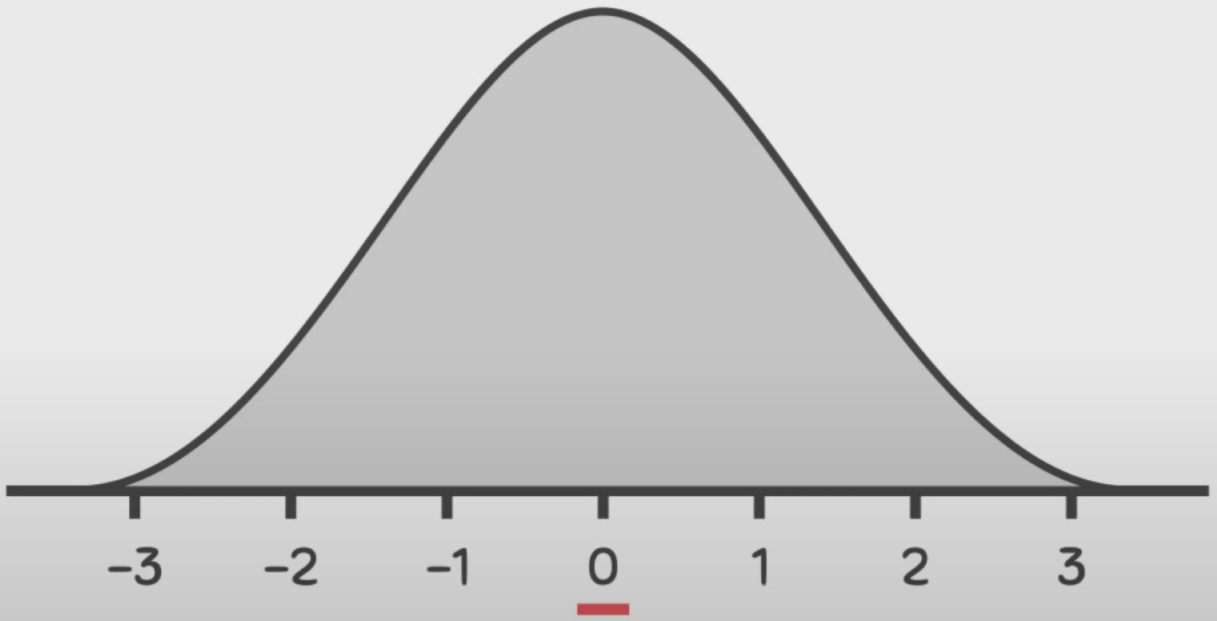


PRACTICE QUESTIONS

❷ For the normal distribution below, approximately what area is contained between −2 and 1?

$\mu = 0$

$\sigma = 1$

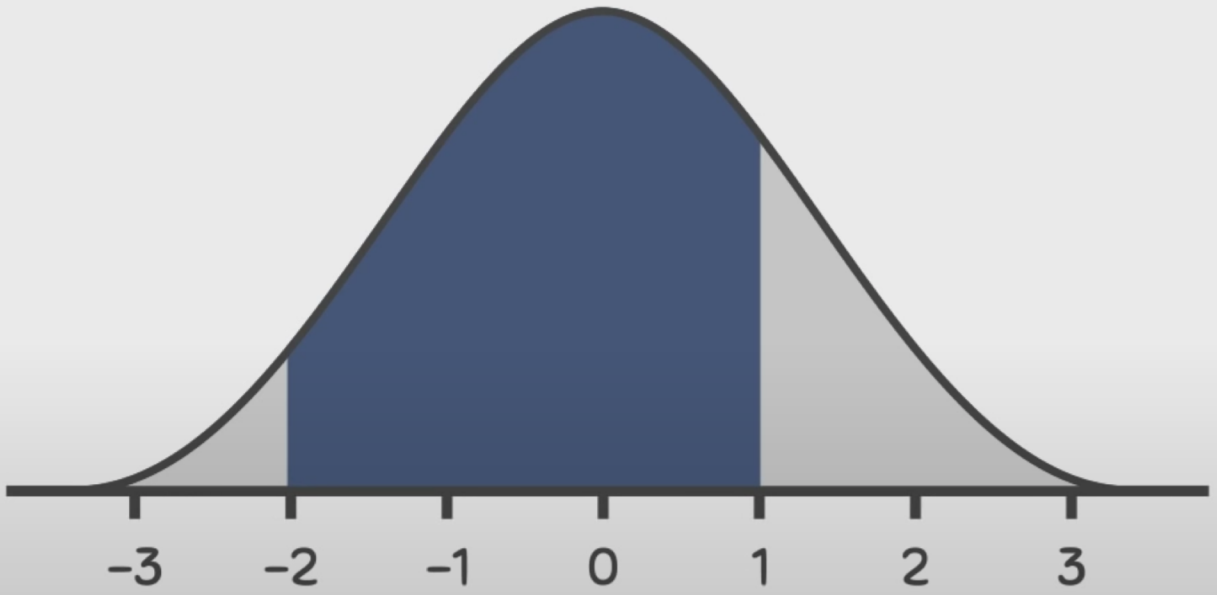


PRACTICE QUESTIONS

❷ For the normal distribution below, approximately what area is contained between −2 and 1?

$\mu = 0$

$\sigma = 1$

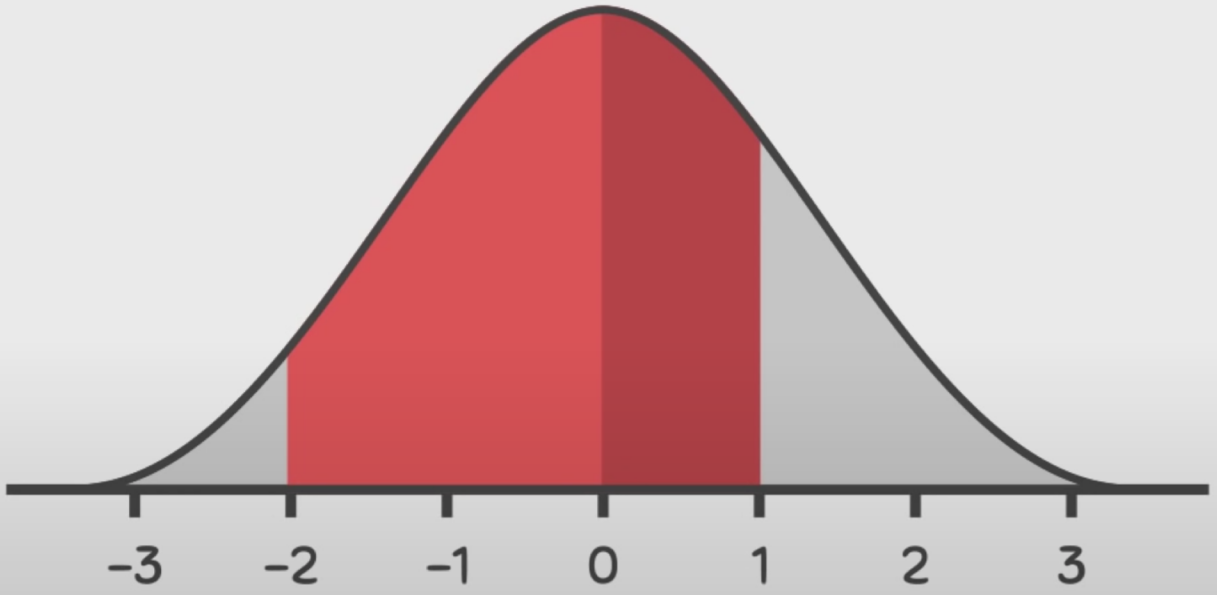


PRACTICE QUESTIONS

❷ For the normal distribution below, approximately what area is contained between −2 and 1?

$\mu = 0$

$\sigma = 1$

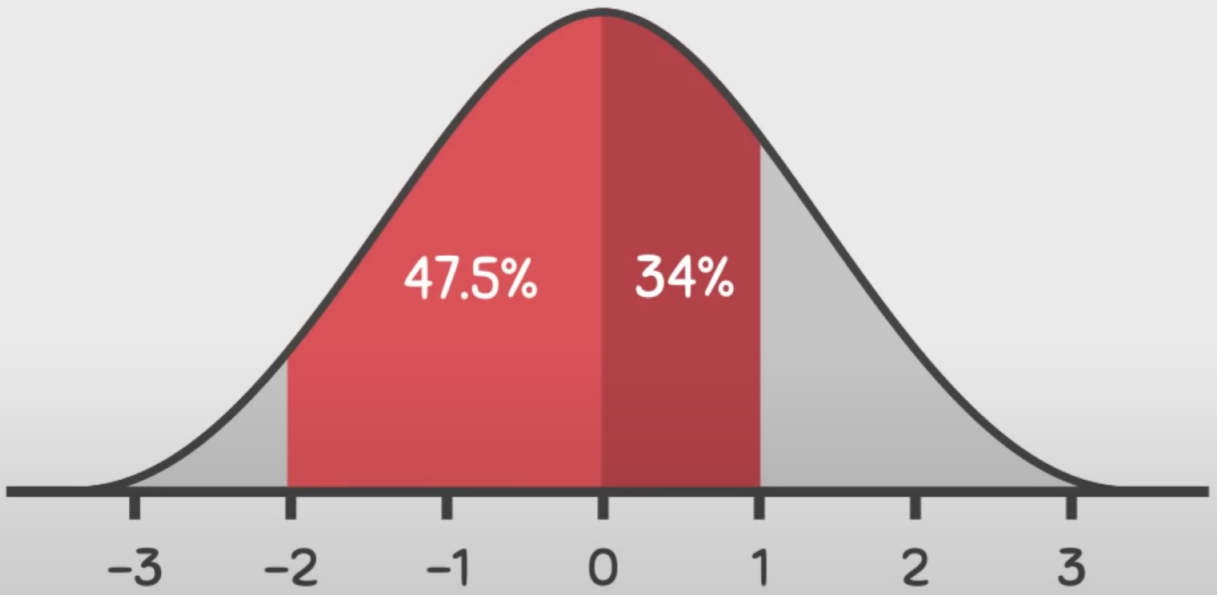


PRACTICE QUESTIONS

❷ For the normal distribution below, approximately what area is contained between −2 and 1?

$\mu = 0$

$\sigma = 1$

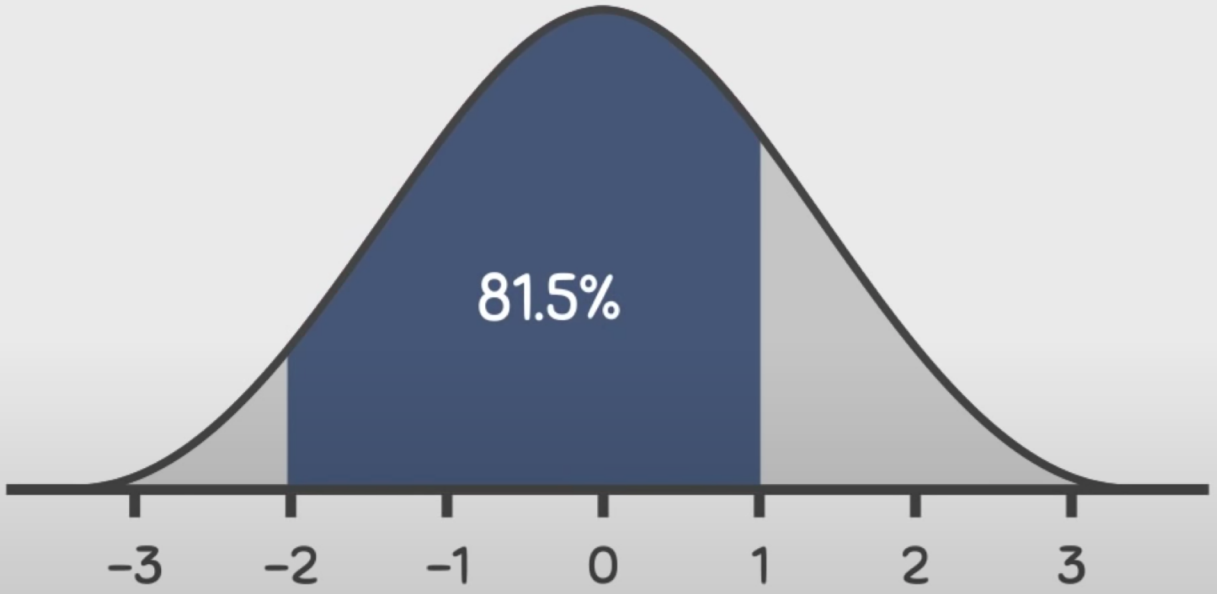


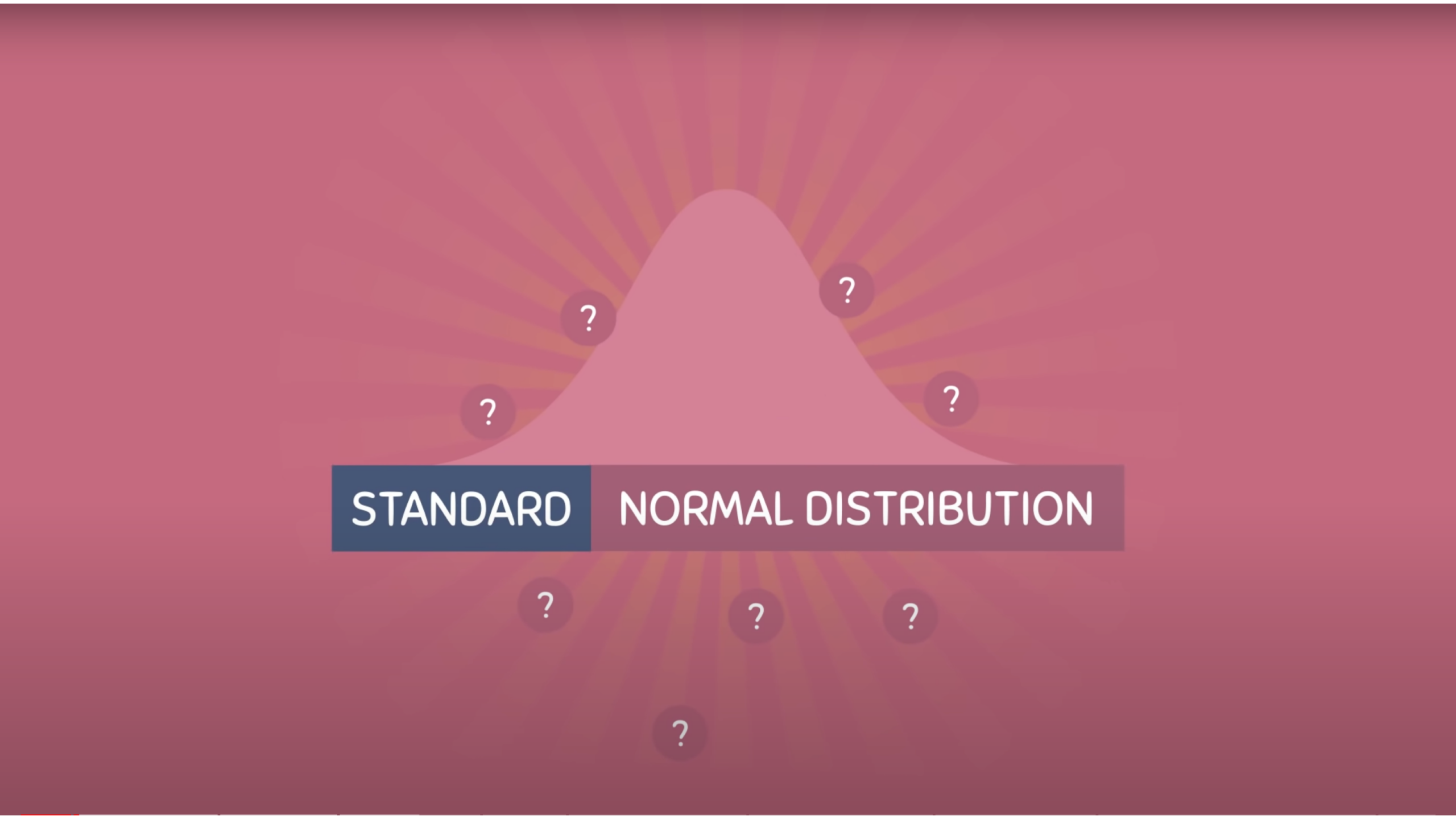
PRACTICE QUESTIONS

❷ For the normal distribution below, approximately what area is contained between −2 and 1?

$\mu = 0$

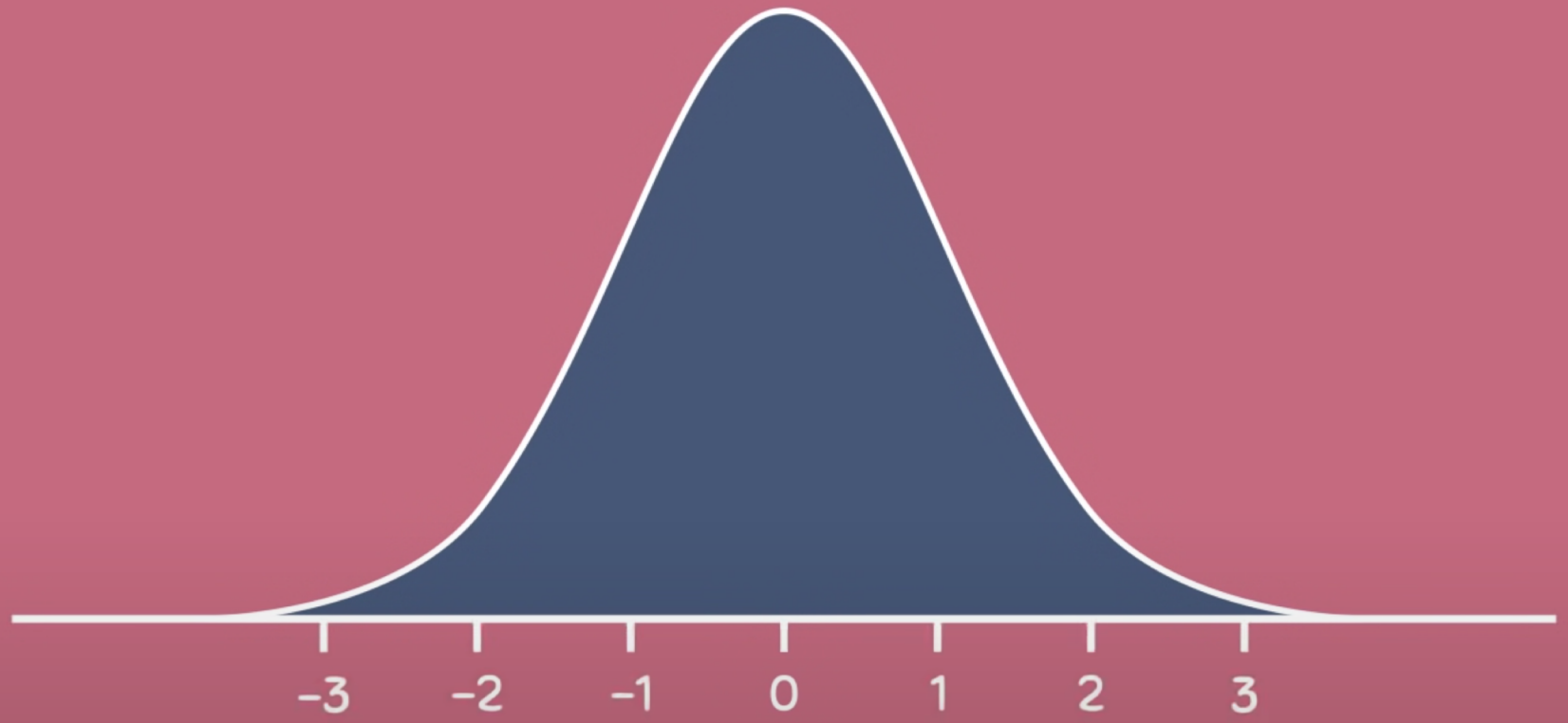
$\sigma = 1$




STANDARD NORMAL DISTRIBUTION

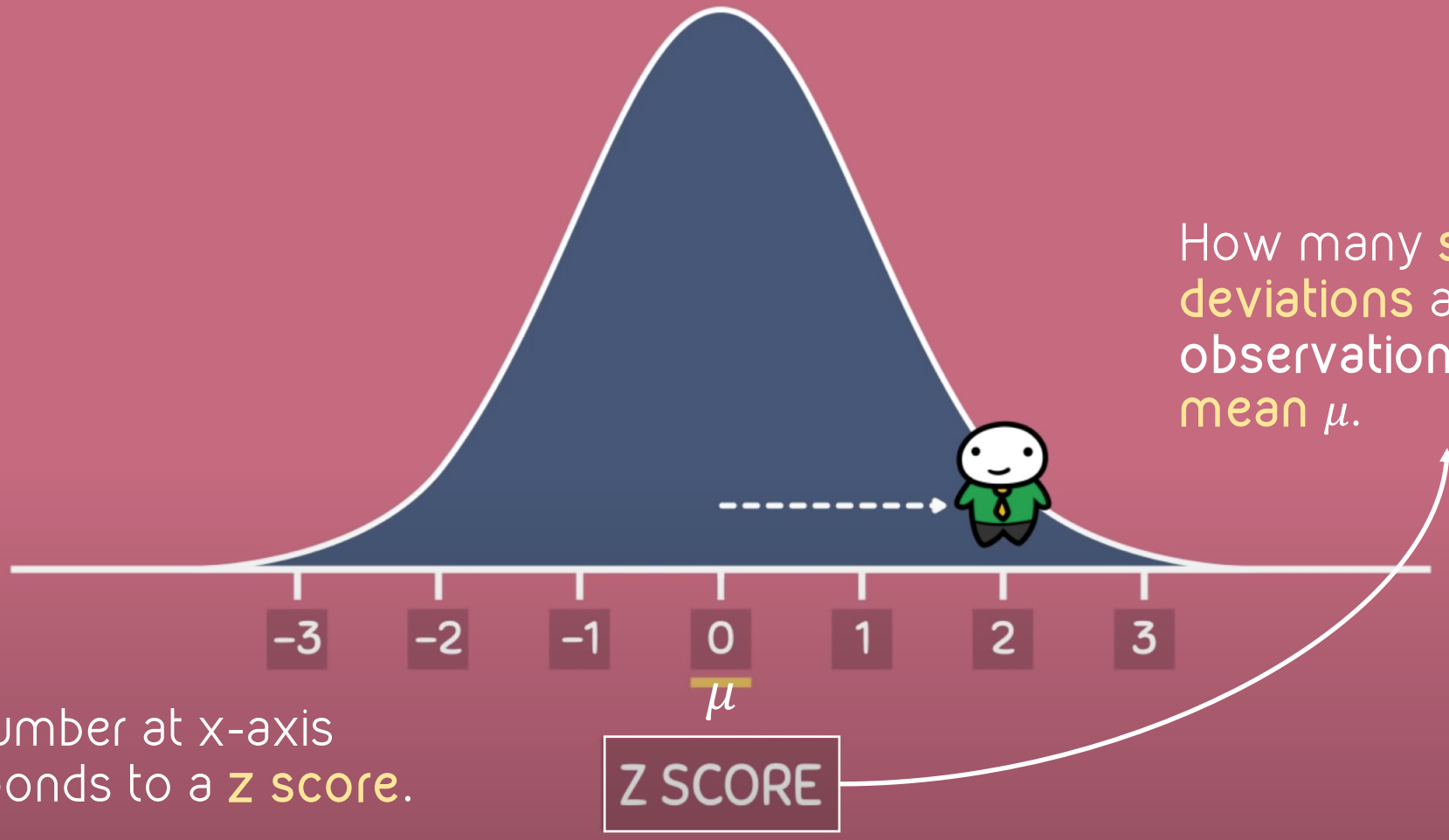
STANDARD NORMAL DISTRIBUTION

- $\mu = 0$
- $\sigma = 1$



STANDARD NORMAL DISTRIBUTION

- $\mu = 0$
- $\sigma = 1$

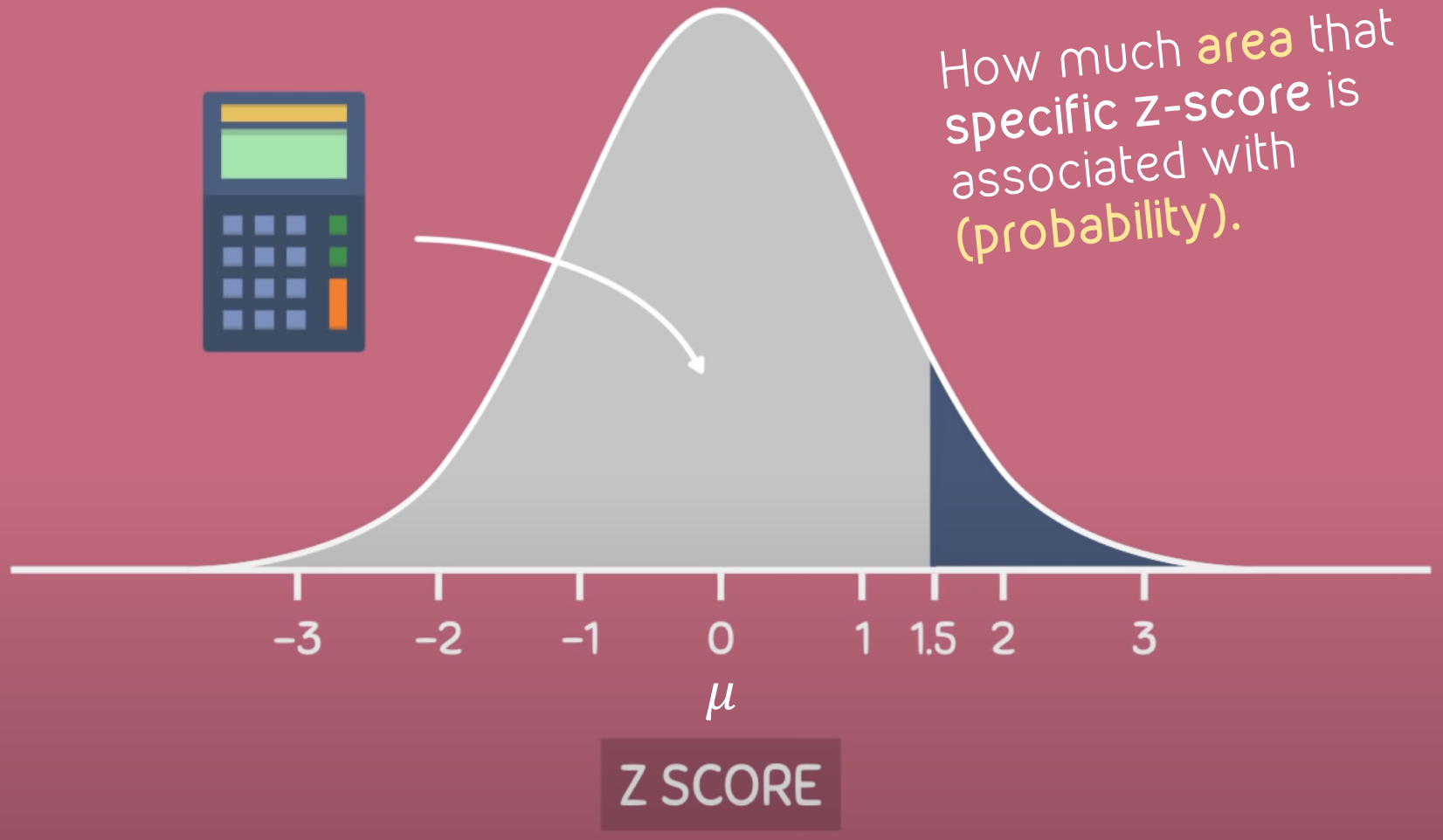


How many **standard deviations** a given **observation** is from the **mean** μ.

Each number at x-axis corresponds to a **z score**.

STANDARD NORMAL DISTRIBUTION

- $\mu = 0$
- $\sigma = 1$



Z-SCORE TABLE

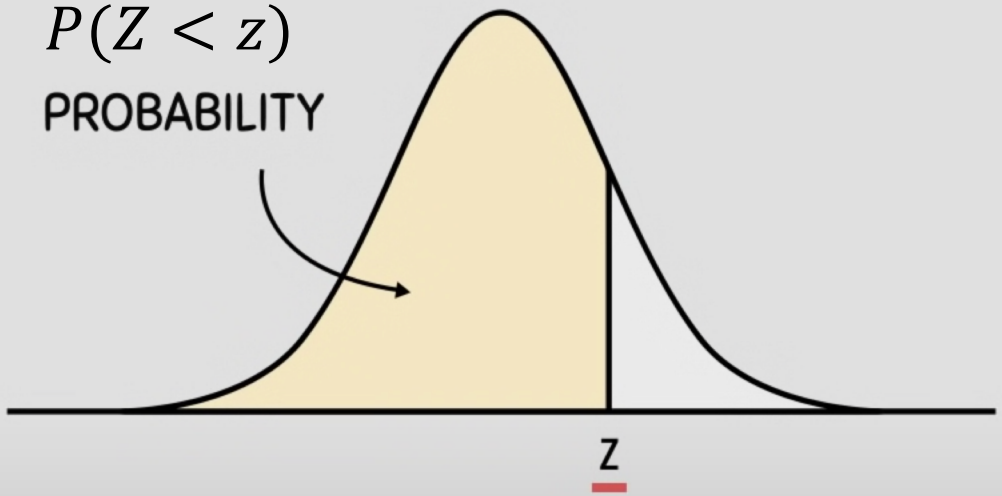
STANDARD NORMAL TABLE

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0224	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6631	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

z values

areas
(probabilities)

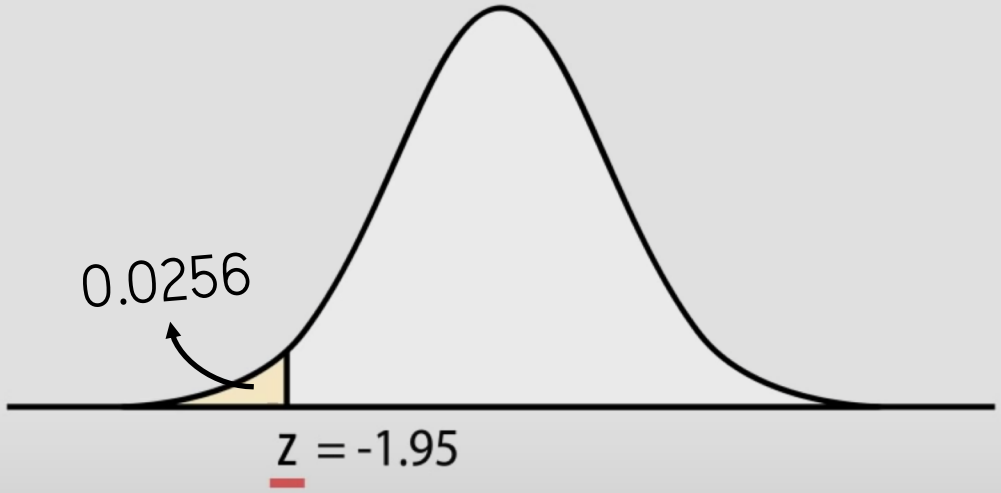


Tells the **total amount of area** (probability) contained to the **left side** of value of z.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0224	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

area

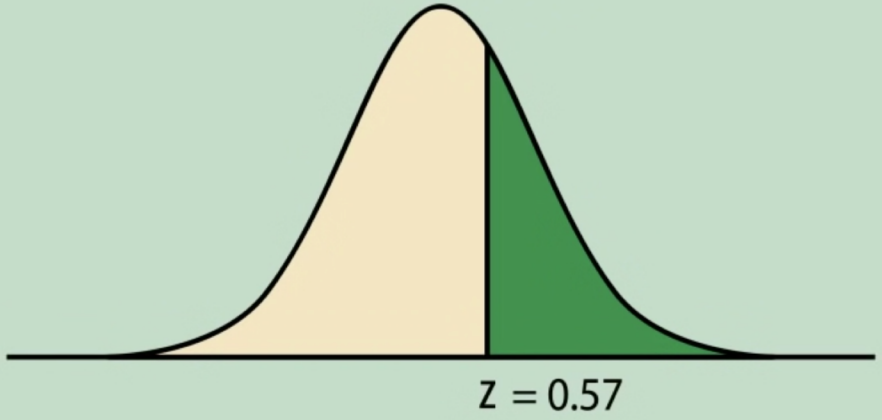
$P(Z < -1.95) = 0.0256$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0224
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660

$$1 - \text{AREA}_{\text{LEFT}} = \text{AREA}_{\text{RIGHT}}$$

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



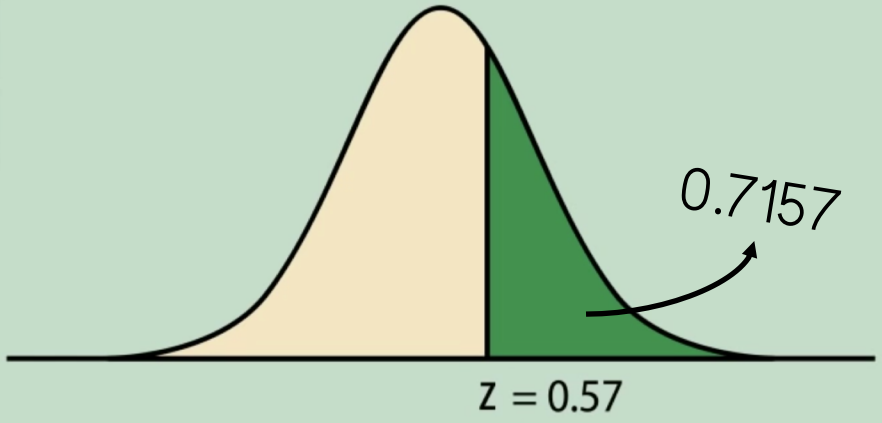
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6631	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389

$$1 - 0.7157 = 0.2843$$

$$P(Z \geq 0.57) = 1 - P(Z < 0.57)$$

DENSITY CURVE

TOTAL AREA = 1

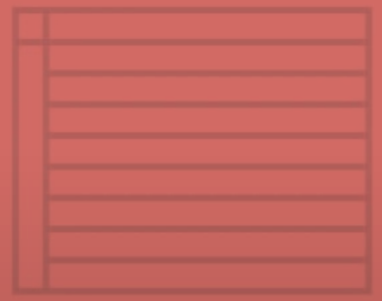


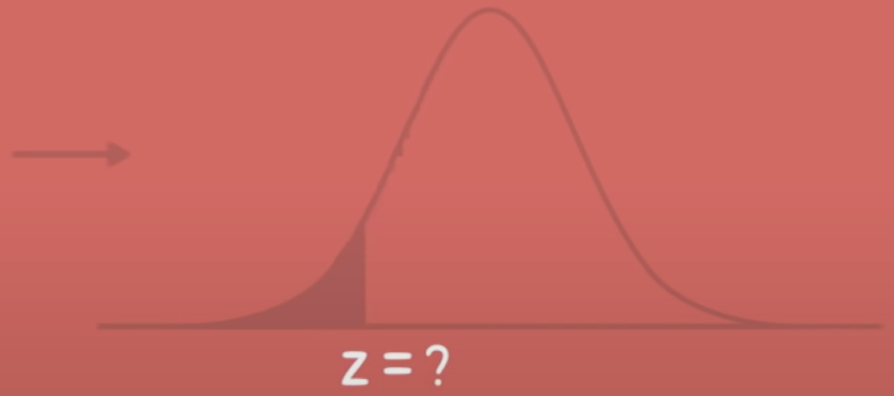
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6631	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389

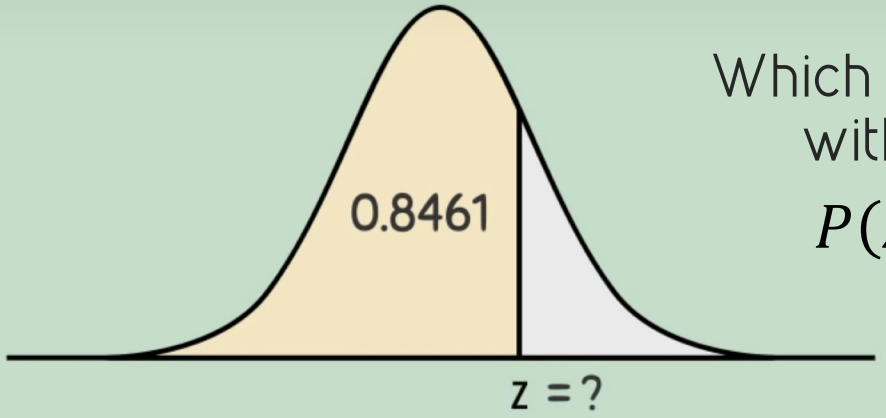
Which z-score is associated with
a specific area (probability)?

REVERSE LOOK-UP

standard normal table



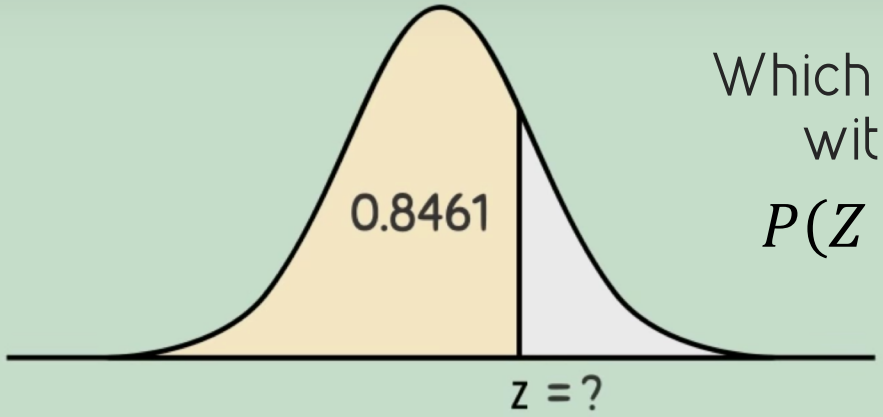




Which z-score is associated with the area 0.8461?

$$P(Z < z) = 0.8461$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6631	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319



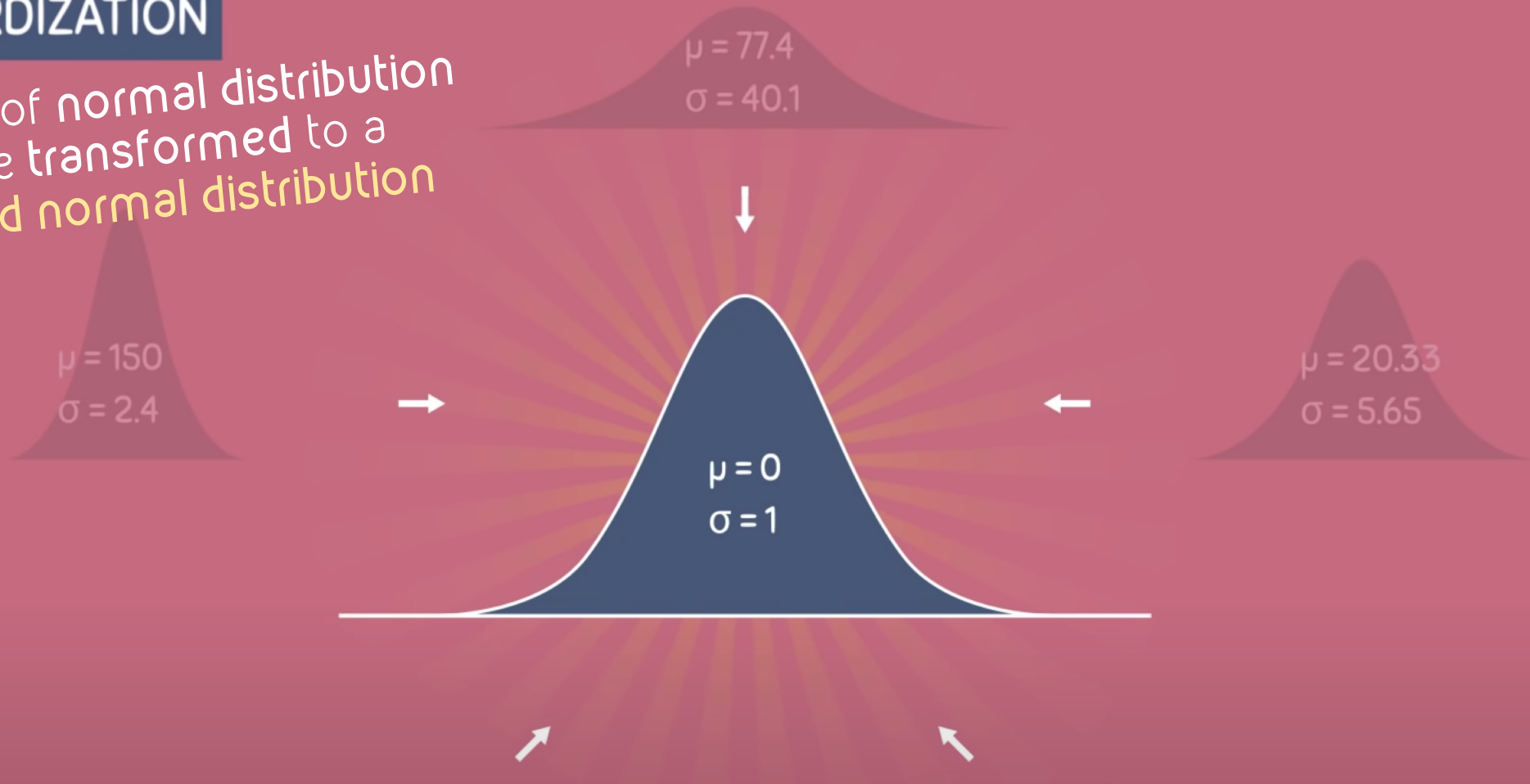
Which z-score is associated with the area 0.8461?

$$P(Z < \mathbf{1.02}) = 0.8461$$

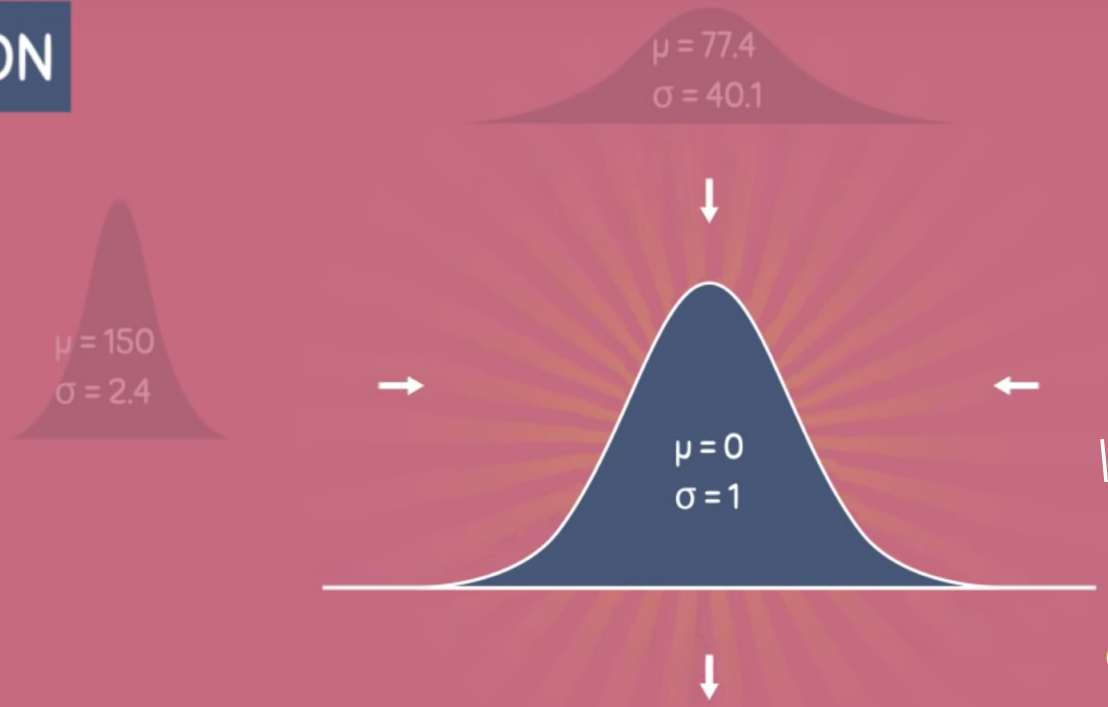
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6631	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319

STANDARDIZATION

Any type of **normal distribution** can be **transformed** to a **standard normal distribution**

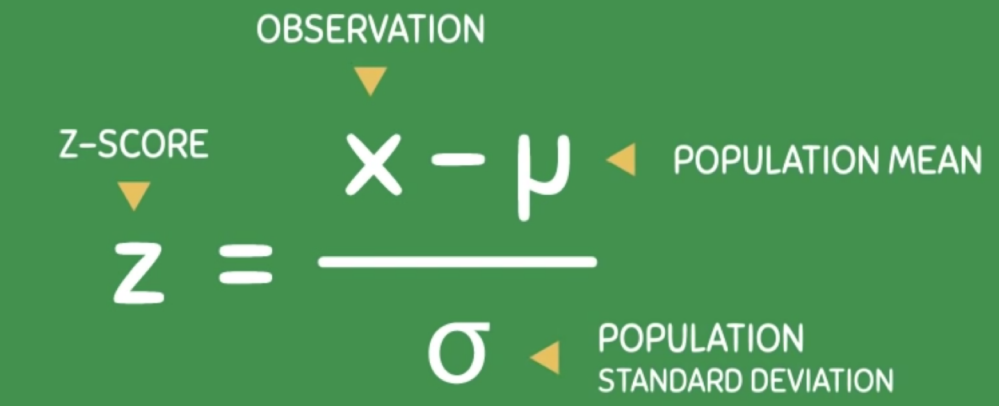


STANDARDIZATION



It allows us to use the **z-score** table to calculate **exact areas** from **any** given **normally distributed population** with **any** value of μ and σ

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6631	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319

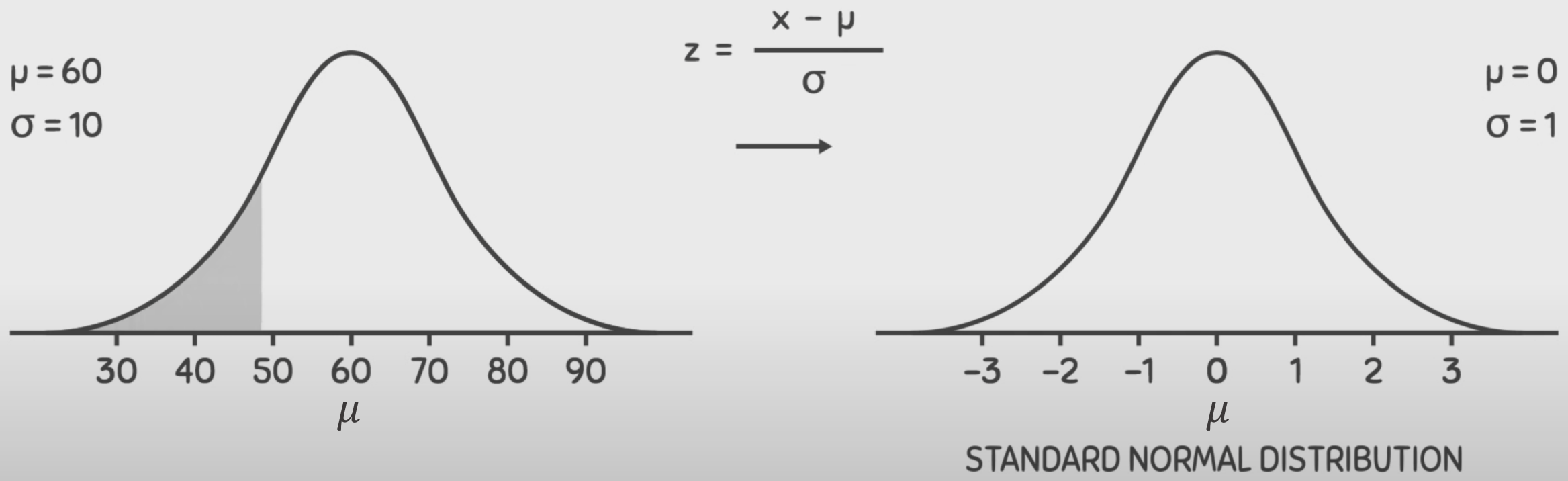


STANDARDIZATION
FORMULA

EXAMPLE

Suppose that we gathered data from last year's final chemistry exam and found that it followed a normal distribution with a mean of 60 and a standard deviation of 10. What proportion of students scored less than 49 on the exam?

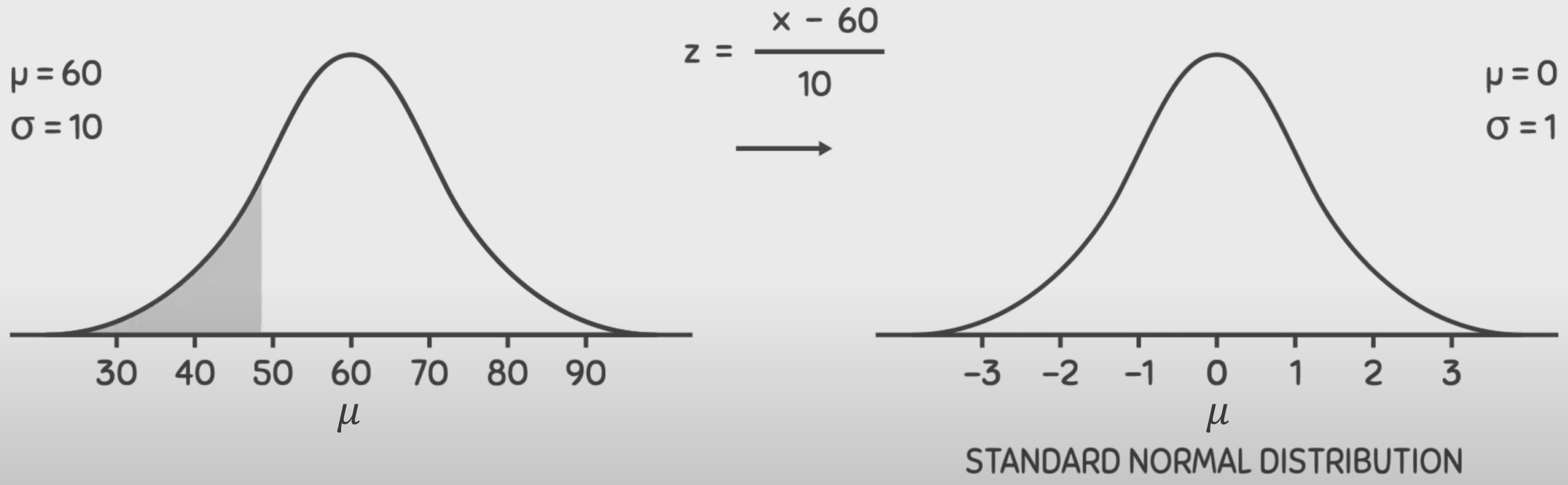
$$P(X < 49) = ?$$



EXAMPLE

Suppose that we gathered data from last year's final chemistry exam and found that it followed a normal distribution with a mean of 60 and a standard deviation of 10. What proportion of students scored less than 49 on the exam?

$$P(X < 49) = ?$$

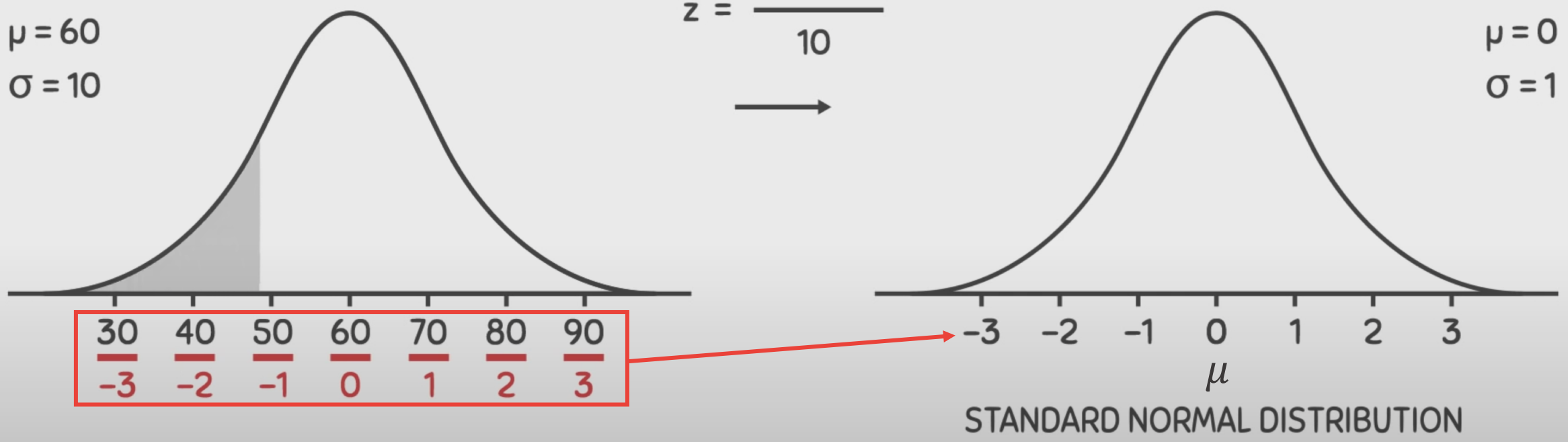


EXAMPLE

Suppose that we gathered data from last year's final chemistry exam and found that it followed a normal distribution with a mean of 60 and a standard deviation of 10. What proportion of students scored less than 49 on the exam?

$$P(X < 49) = ?$$

$$z = \frac{x - 60}{10}$$



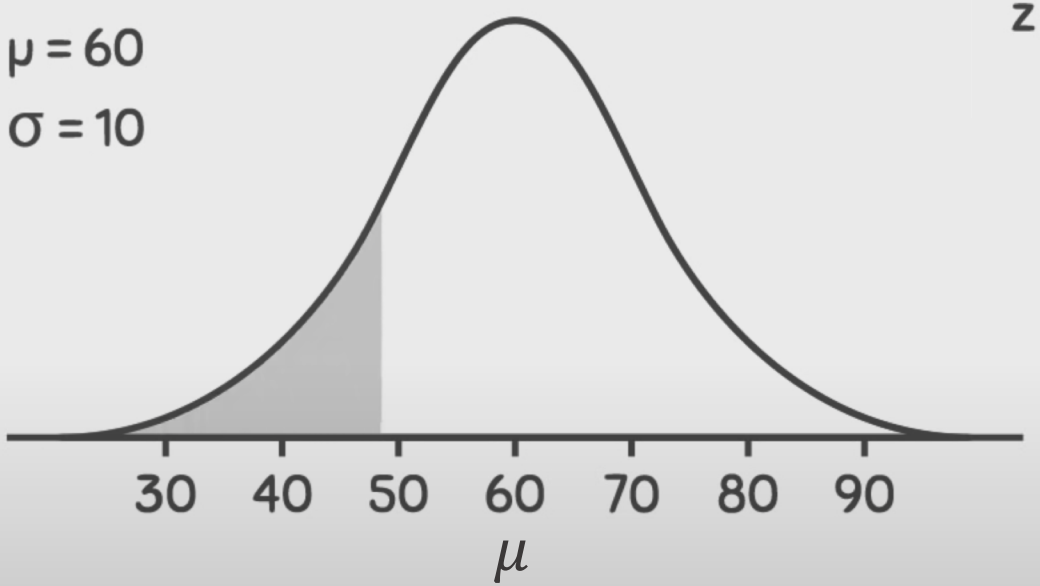
EXAMPLE

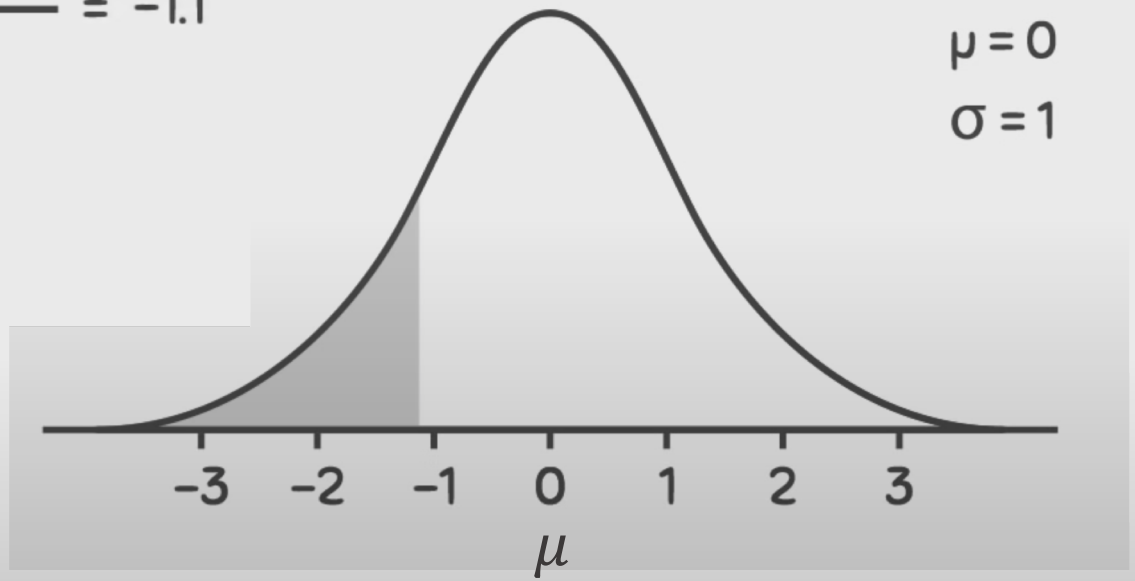
Suppose that we gathered data from last year's final chemistry exam and found that it followed a normal distribution with a mean of 60 and a standard deviation of 10. What proportion of students scored less than 49 on the exam?

$P(X < 49) = ?$

$$z = \frac{49 - 60}{10} = -1.1$$

$P(Z < -1.1) = ?$





STANDARD NORMAL DISTRIBUTION

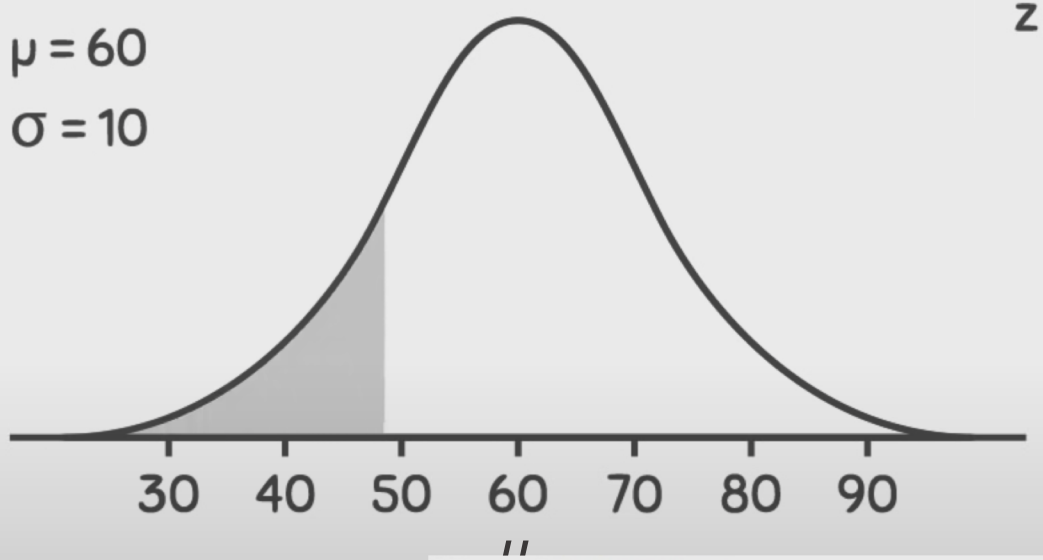
EXAMPLE

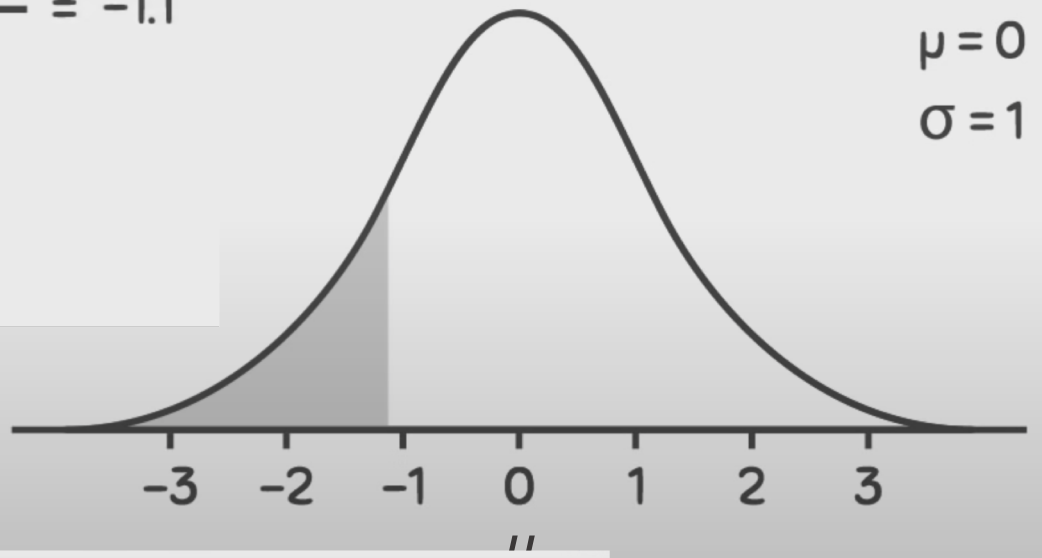
Suppose that we gathered data from last year's final chemistry exam and found that it followed a normal distribution with a mean of 60 and a standard deviation of 10. What proportion of students scored less than 49 on the exam?

$P(X < 49) = 0.1357$

$P(Z < -1.1) = 0.1357$

$$z = \frac{49 - 60}{10} = -1.1$$





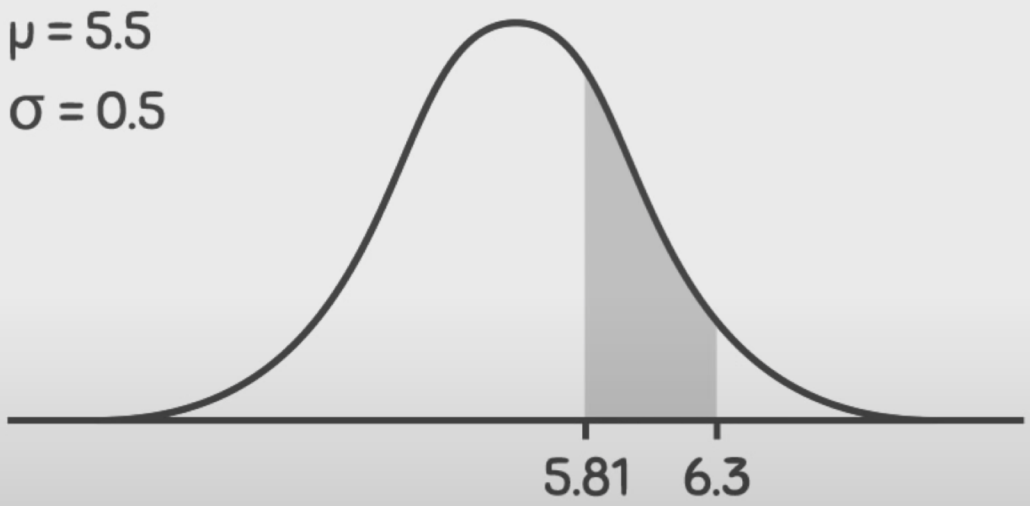
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379

ISTRIBUTION

EXAMPLE

When measuring the heights of all students at a local university, it was found that it was normally distributed with a mean height of 5.5 feet, and a standard deviation of 0.5 feet. What proportion of students are between 5.81 feet, and 6.3 feet tall?

$$P(5.81 < X < 6.3) = ?$$



EXAMPLE

When measuring the heights of all students at a local university, it was found that it was normally distributed with a mean height of 5.5 feet, and a standard deviation of 0.5 feet. What proportion of students are between 5.81 feet, and 6.3 feet tall?

$P(5.81 < X < 6.3) = ?$

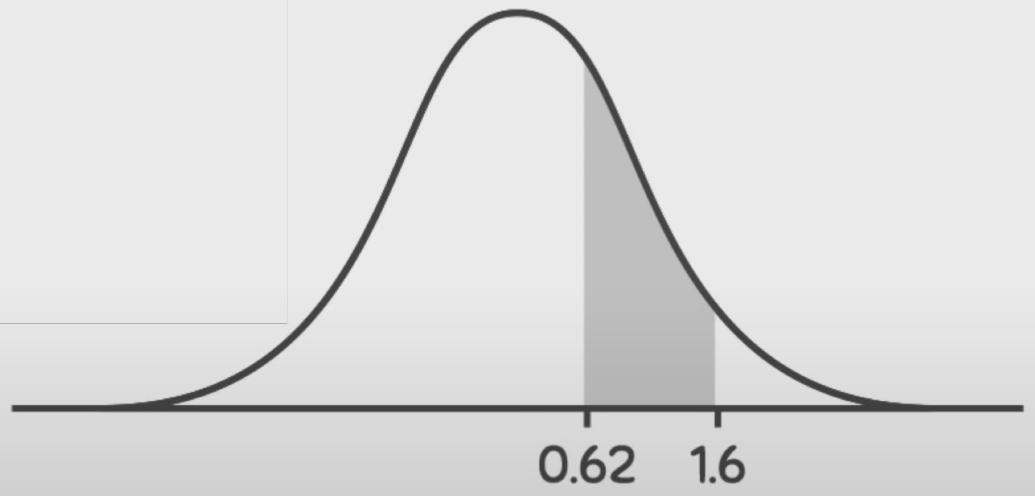
$\mu = 5.5$

$\sigma = 0.5$

$$z = \frac{x - \mu}{\sigma}$$

$P(0.62 < Z < 1.6) = ?$

$P(Z < 1.6) - P(Z < -0.62)$



EXAMPLE

When measuring the heights of all students at a local university, it was found that it was normally distributed with a mean height of 5.5 feet, and a standard deviation of 0.5 feet. What proportion of students are between 5.81 feet, and 6.3 feet tall?

$P(5.81 < X < 6.3) = ?$

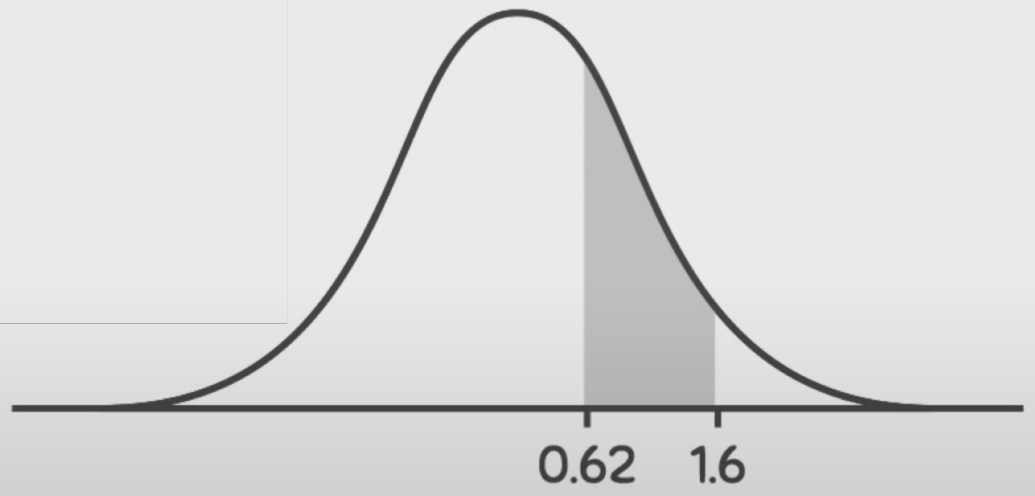
$\mu = 5.5$

$\sigma = 0.5$

$$z = \frac{x - \mu}{\sigma}$$

$P(0.62 < Z < 1.6) = ?$

$P(Z < 1.6) - P(Z < -0.62)$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
↓										
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
↓										
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633

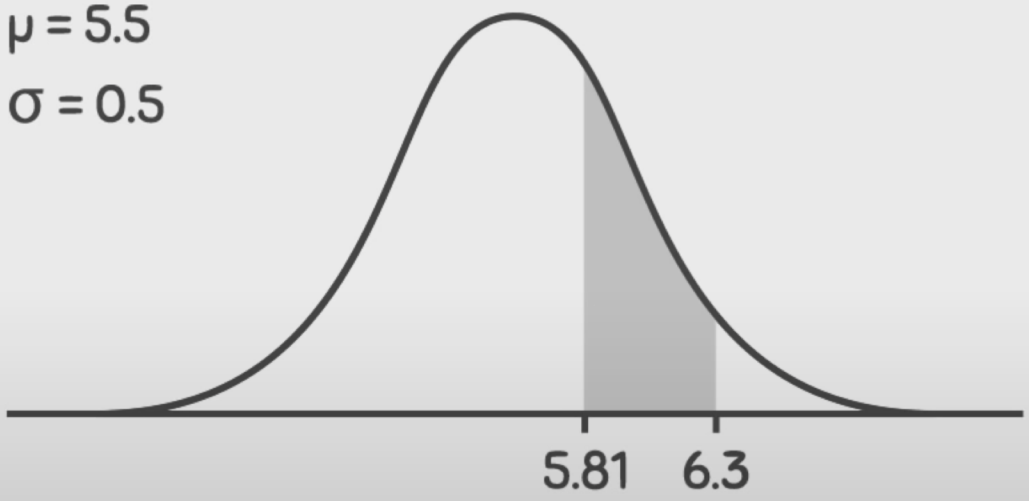
EXAMPLE

When measuring the heights of all students at a local university, it was found that it was normally distributed with a mean height of 5.5 feet, and a standard deviation of 0.5 feet. What proportion of students are between 5.81 feet, and 6.3 feet tall?

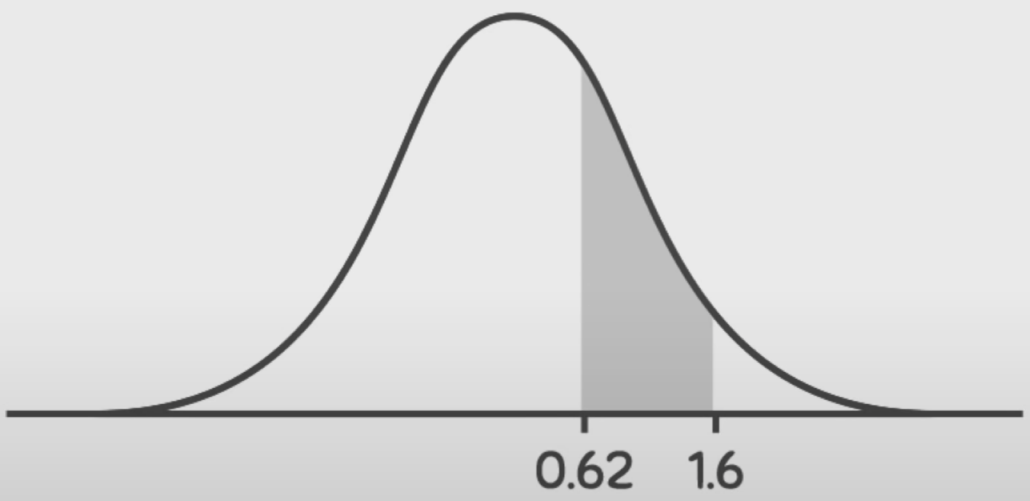
$P(5.81 < X < 6.3) = 0.2128$

$\mu = 5.5$

$\sigma = 0.5$



$P(0.62 < Z < 1.6) = 0.2128$



$P(Z < 0.62) = 0.7324$

$P(Z < 1.6) = 0.9452$

Exercise

When studying the height of the inhabitants of Pompeia, it was found that its **distribution is approximately normal**, with **mean** of 1.70 m and **standard deviation** of 0.1. Calculate the:

(1) Probability of a person, selected by chance, is less than 1.8m tall?

(2) Probability of a person, selected by chance, is between 1.6m and 1.8m tall?

(3) Probability of a person, selected by chance, is over 1.9m tall?

D1EAD – Análise Estatística para Ciência de Dados 2021.1

Data Distributions (Part 2)

Prof. Ricardo Sovat
sovat@ifsp.edu.br
Prof. Samuel Martins (Samuka)
samuel.martins@ifsp.edu.br

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
São Paulo
Campus Campinas