NORMAL DISTRIBUTION – EMPIRICAL RULE

Calculus

The empirical rule states that for a normal distribution, nearly all of the data will fall within three standard deviations of the mean. The empirical rule can be broken down into three parts:

- 68% of data falls within the first standard deviation from the mean.
- 95% fall within two standard deviations.
- 99.7% fall within three standard deviations.

The rule is also called the **68-95-99 7 Rule** or the **Three Sigma Rule**.

π → radial factor

pi

3.14

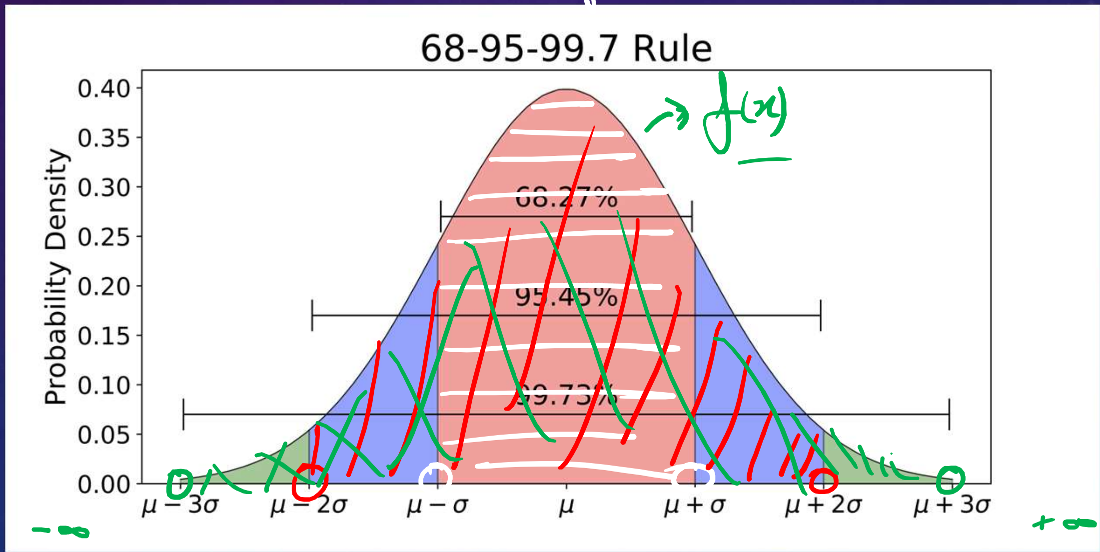
r

e → Natural Number

→ logarithm

2.71

radius of a straight line is ∞



σ → Std. dev

μ → pop. mean

Area Under the Curve → AUC

$\frac{d}{dh}(f(h))$

$\frac{d}{dx}(f(x))$

Calculus

Differentiation

↓

Dividing the curve into small rectangle

Integration

↳ Area under the curve b/w LL & UL

$\int_{LL}^{UL}$

$y = x^2$

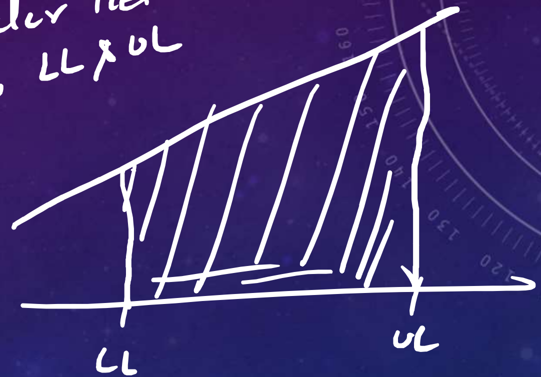
$\frac{dy}{dx} = 2x$

$x = 1$

$\frac{dy}{dx} = 2$

↓

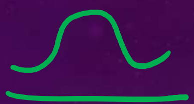
Area of the rectangle at $x = 2$



→ Z distribution

PDF OF A NORMAL DISTRIBUTION AND ORIGIN OF EMPIRICAL FORMULA

N.D. $(\underset{\mu}{35}, \underset{\sigma}{10})$ → Age



ND. $(\overset{\mu}{1000}, \overset{\sigma}{3.71})$

↓

Approx

$\underbrace{\mu = 0, \sigma = 1}$

↓

ND.

↓

Standard Normal Distribution (SND)

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

Normal Distrib.

→ Z-score

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

$(\mu = 0, \sigma = 1)$

Note constants:

π=3.14159

e=2.71828

This is a bell shaped curve with different centers and spreads depending on μ and σ

Empirical Formula of Normal Distribution: The empirical rule, also referred to as the three-sigma rule or 68-95-99.7 rule, is a statistical rule which states that for a normal distribution, almost all data falls within three standard deviations (denoted by σ) of the mean (denoted by μ). Broken down, the empirical rule shows that almost 68% falls within the first standard deviation (μ ± σ), almost 95% within the first two standard deviations (μ ± 2σ), and almost 99.7% within the first three standard deviations (μ ± 3σ).

NORMAL DISTRIBUTION/ORIGIN OF Z-SCORE

Standardizing $\leftarrow Z = \frac{x-\mu}{\sigma} \rightarrow$ going to convert ANY (N.D) to (SND)

Z Score gives how many standard deviation away from mean a value is. However, to understand the probability associated with it, we need to refer to Z-Table.

Test → Maths → 100

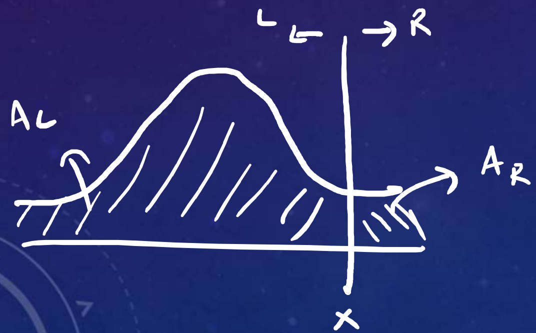
36 → 40

Dad !

Toffee

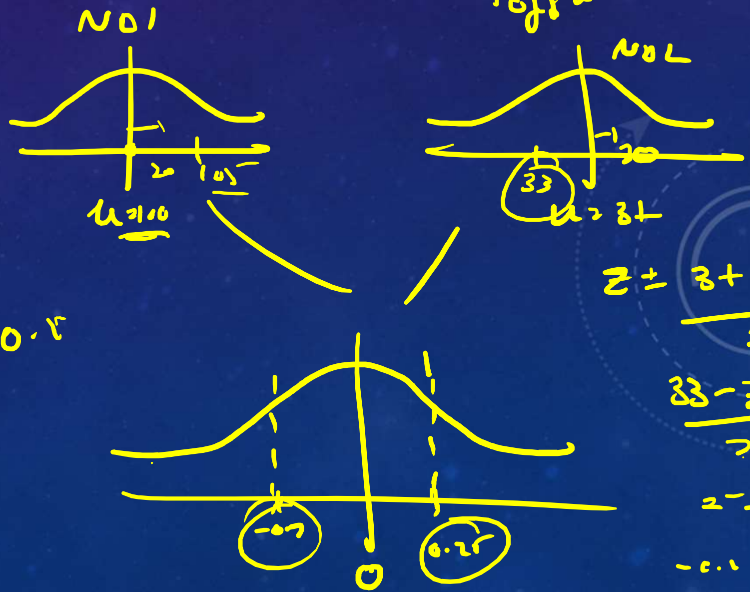
Type 1: Comparison of 2 different Normally Distributed values (Z-Score is enough)

Type 2: Finding the probability or percentage of values. (Need Z-table)



$Z = \frac{100-100}{20} \rightarrow 0$

$\frac{105-100}{20} = \frac{5}{20} = 0.25$



$Z = \frac{37-37}{20}$

$\frac{33-37}{20}$

$= \frac{-4}{20}$

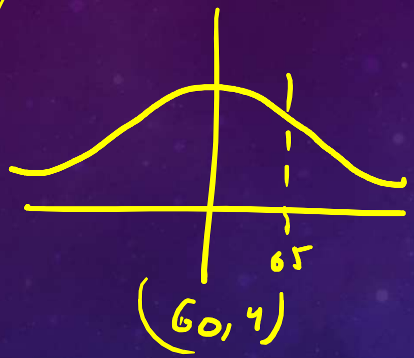
$= -0.2$

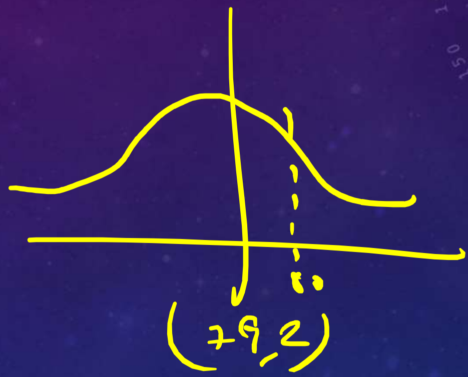
Type 1: Happy and Ekta are two students. Happy Scored 65 marks in Math Exam while Ekta scored 80 in English Exam. Given that both Math and English marks follows an approx. Normal Distribution, who performed better?

Math ~ N(60,4)

English ~ N(79,2)

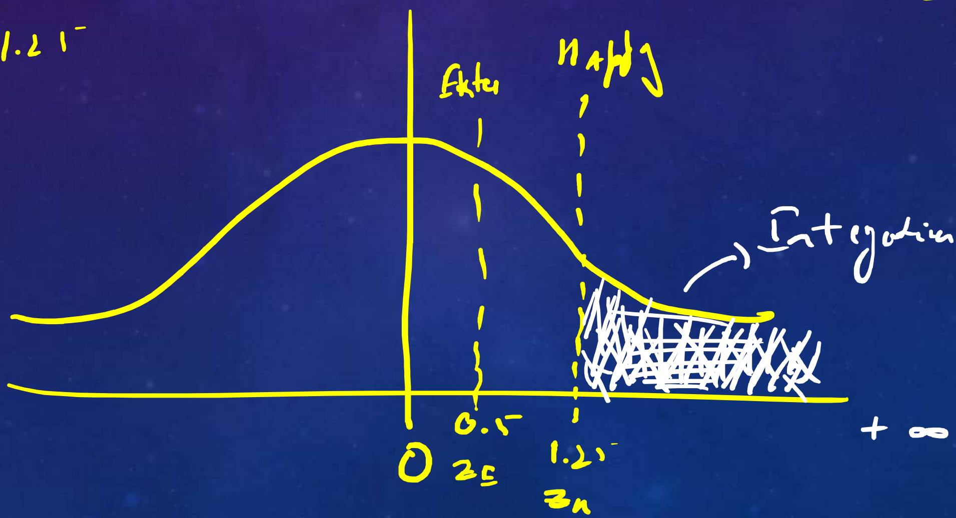
$$z = \frac{x-\mu}{\sigma}$$





$$z_{Happy} = \frac{65-60}{4} = 1.25$$

$$z_{Ekta} = \frac{80-79}{2} = 0.5$$



Integration (AUC)

LL = 1.25

UL = + ∞

= 0.1056

= 10.56%

Type 2: According to the Canter for Disease Control, heights for U.S. adult females and males are approximately normal.

Females: mean of 64 inches and SD of 2 inches

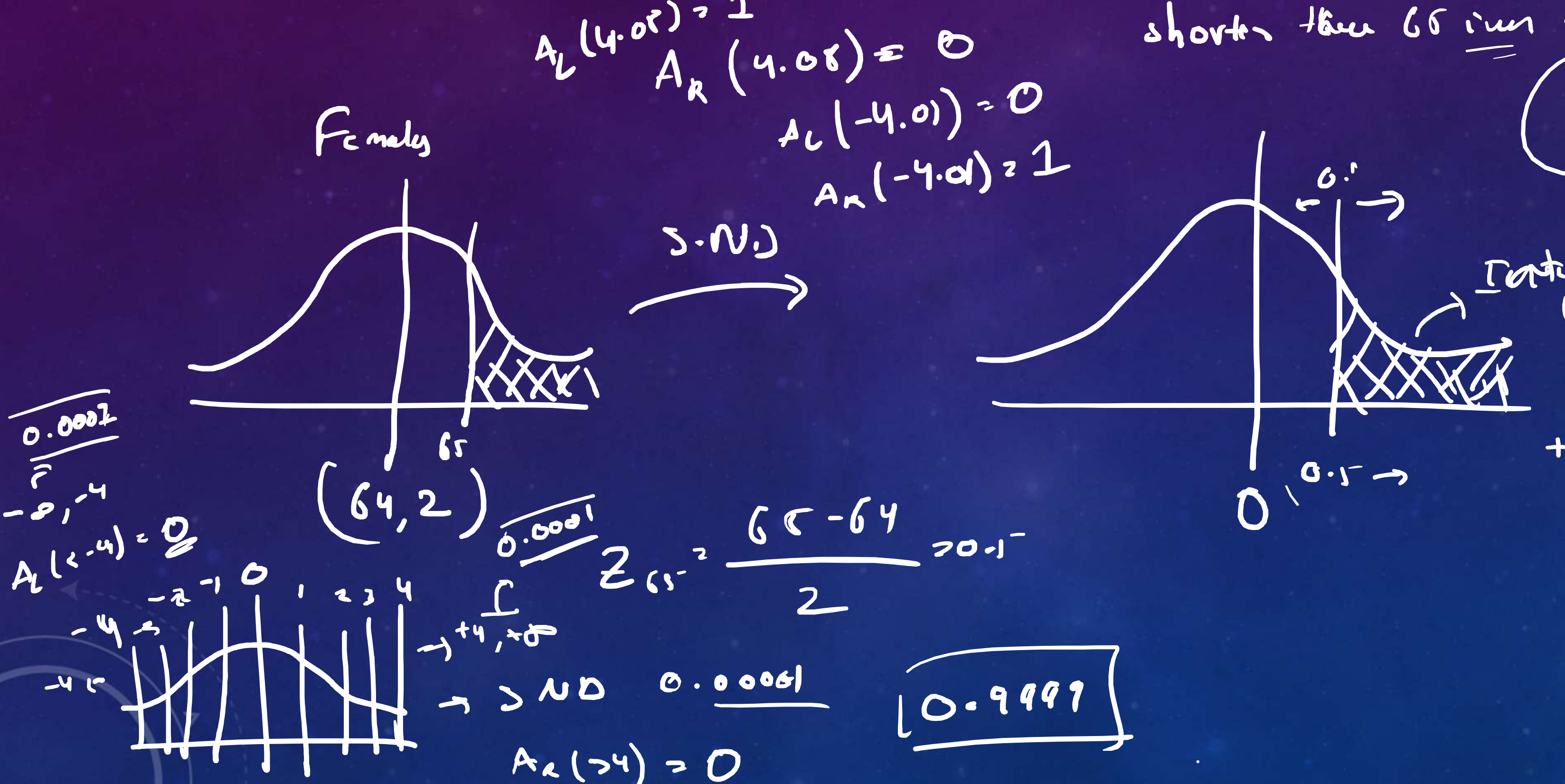
Males: mean of 69 inches and SD of 3 inches

Find the probability of a randomly selected U.S. adult female being taller than 65 inches. → ~ 31%

shorter than 65 inch ~ 69%

$A_R + A_L = 1$

$A_R = 1 - A_L$



$$Z_{65} = \frac{65-64}{2} = 0.5$$

$\boxed{0.9999}$

LL = 0.5

UL = +∞

$p = 0.30853$

% = 30.85%

→ Integration to find the AUC

↳ Problem → Z - Table → A_L $(-4, +4)$

-ve z Table $(-4, 0)$

+ve z Table $(0, +4)$

↓

$[-\infty, (-4, +4)]$

$(-\infty, -4)$ $\quad -\infty, -3.99$

-3.98

-3.97

$\vdots$

$+3.98$

$+3.99$

$+4.00$

$(-\infty, -1.15)$

↓

A_L

-4

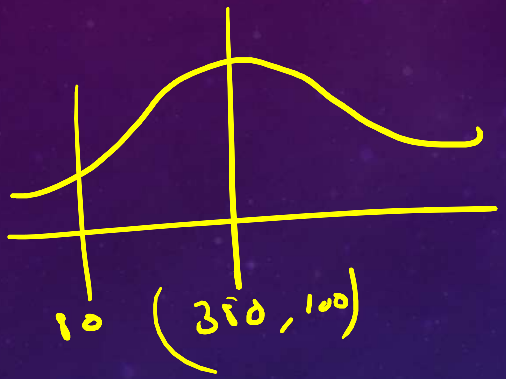
$-\infty$ $\quad 4$

A_L

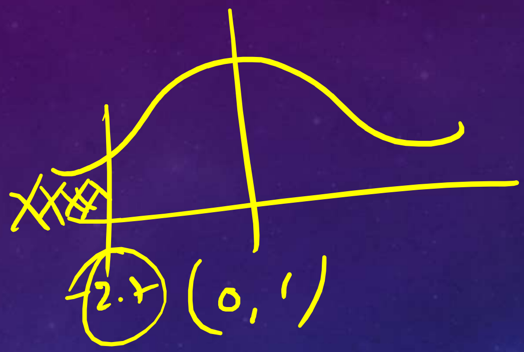
$A_R = 1 - A_L$

Books in the library are found to have average length of 350 pages with standard deviation of 100 pages. What is the z-score corresponding to a book of length 80 pages?

% of Book that are having pages < 80.



SND →



$$Z_{80} = \frac{80-350}{100}$$

$$= -2.7$$

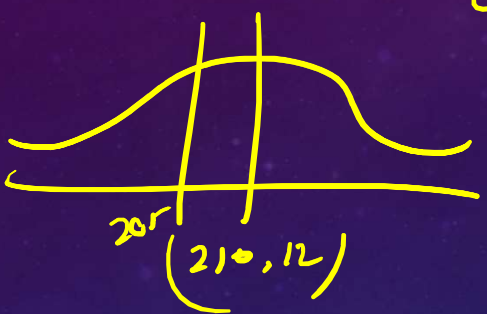
$$A_L = 0.00347$$

0.347%

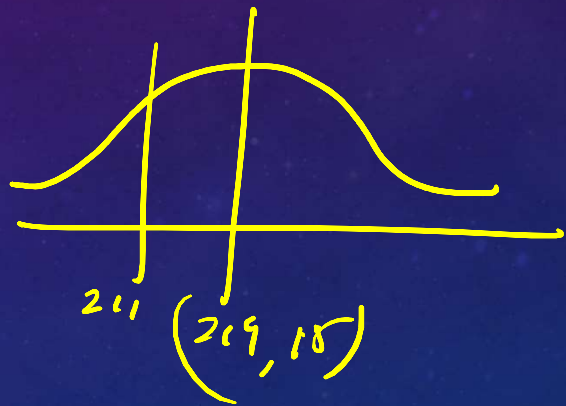
0.35%

In the first round of a tournament, Keri bowled a 205. The first round's mean score was 210, with a standard deviation of 12. In the second round, she bowled a 211. The second round's mean was 219, with a standard deviation of 15. Find the Z-scores for Keri's two scores, and use them to determine which score was better relative to the other bowlers' scores in that round.

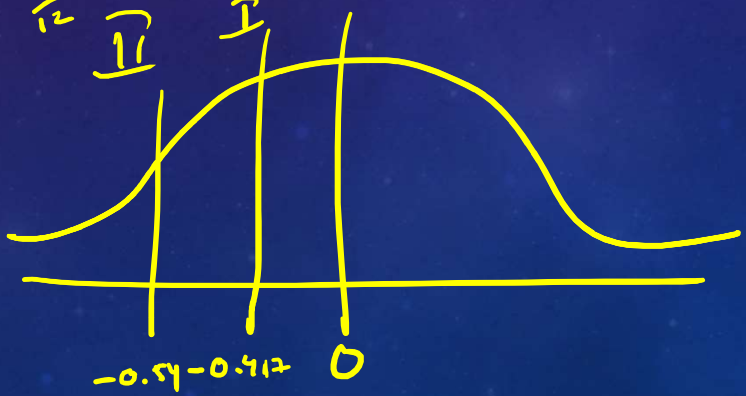
I round



II round



SND



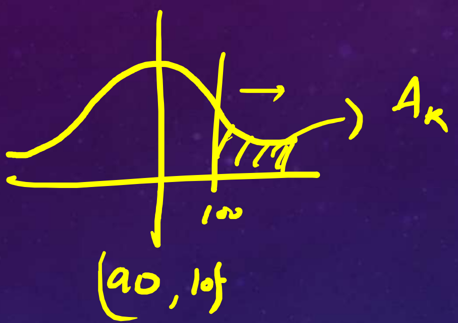
$$z_I = \frac{205-210}{12} = \frac{-5}{12}$$

$$z_{II} = \frac{211-219}{15} = \frac{-8}{15}$$

-0.54

Speed camera

The speeds of cars are measured using a radar unit, on a motorway. The speeds are normally distributed with a mean of 90 km/hr and a standard deviation of 10 km/hr. What is the probability that a car selected at chance is moving at more than 100 km/hr?



$$Z_{100} = \frac{100-90}{10} = +1 \rightarrow A_L \rightarrow Z\text{-table}$$

$$A_R = 1 - A_L$$

$$A_R = 1 - 0.84134$$

$$A_R = 0.15866$$

$$\Rightarrow 15.87\%$$