Gaussian Distribution

→ $\boxed{\text{CDF}}$

↳ probability that a variable can take a certain value less than or equal to a given value

$$P(H \le 63.5) \Rightarrow 25\%$$

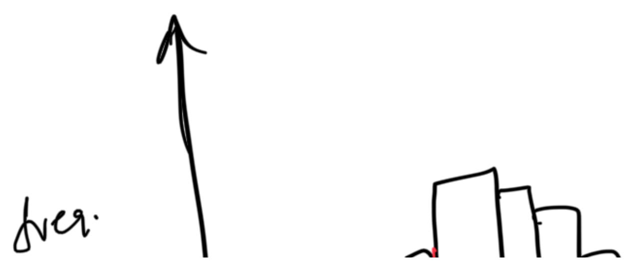
Percentile

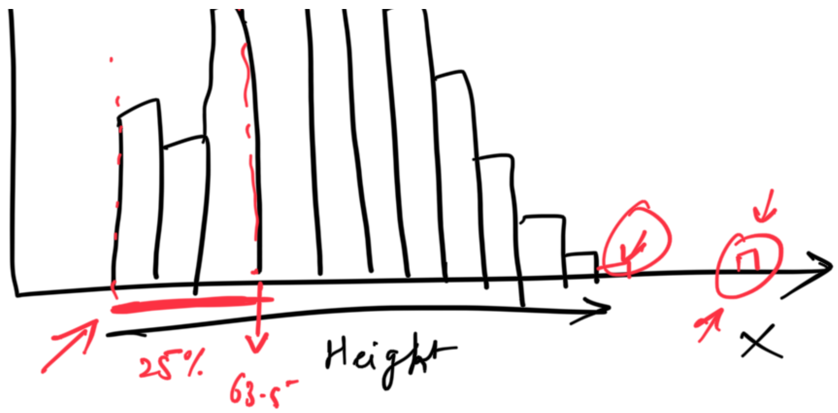
→ 25th percentile value

↳ what is that value below which 25% of the data lies.

n^{th} percentile → n% of data is below a certain value.

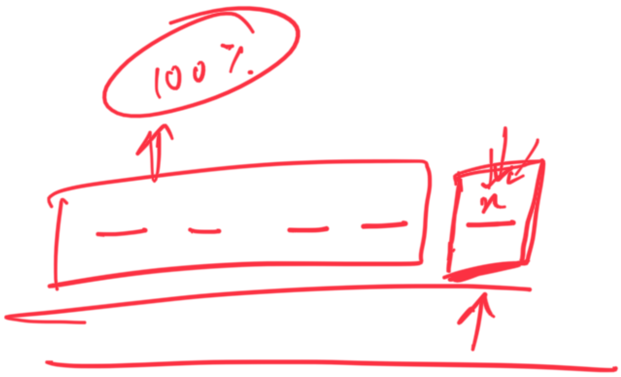
25th percentile (x) → 63.5 (y)





CDF

Probability <= 63.5 => 25%



100%

Below x

Gaussian (Aka Normal Distribution)

Standard Deviation → spread of the data

↳ How far a value is from the mean of the data?

Test Scores: 30, 30, 35, 40, 40

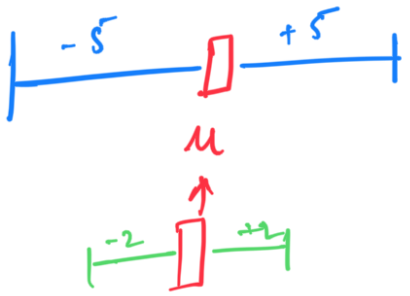
Mean → 35
(μ)

30-35

(-5) (-5) 0 5 5

S.D. - 5

S.D. → 2



Swiggy

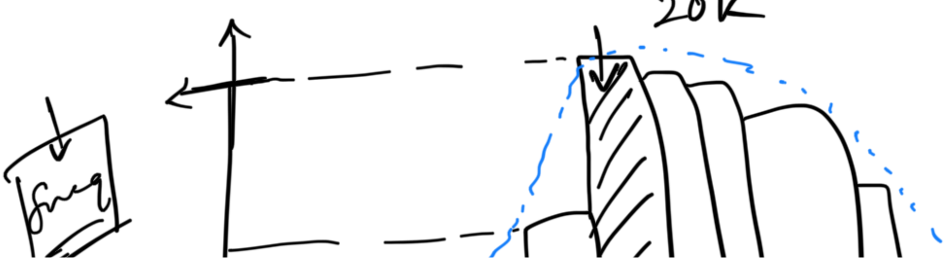
Price

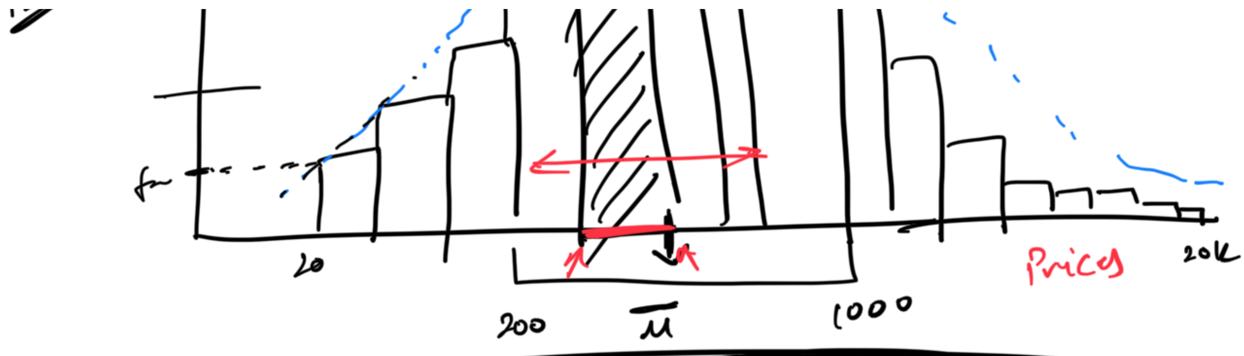
20

20 — 100

100 — 300

300 — 500

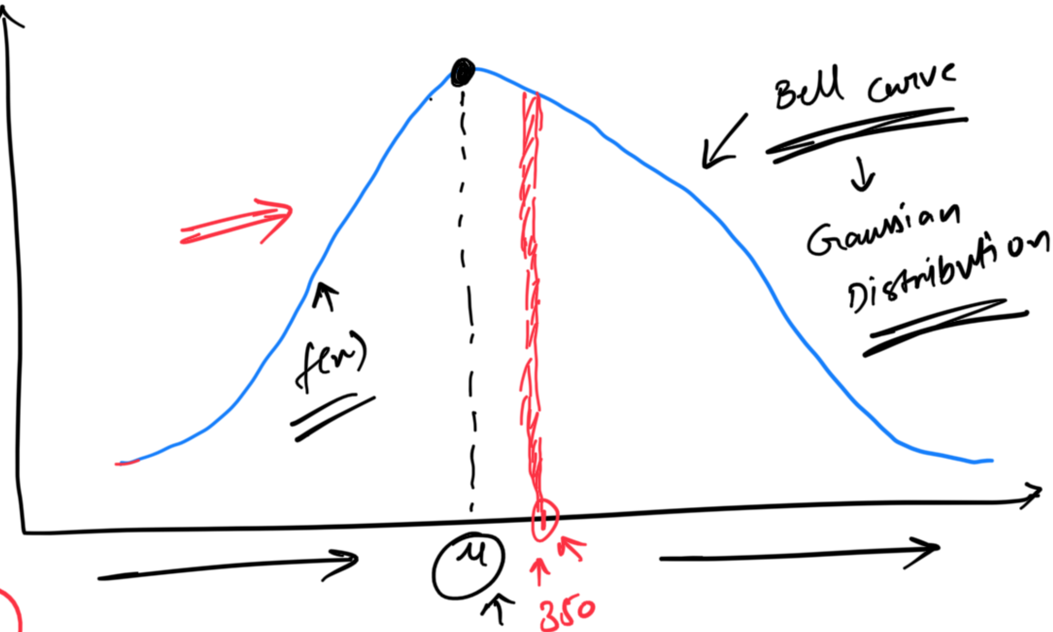




y

Probability Density function

y



Bell curve

Gaussian Distribution

350 - 400

350 - 385

$$\int_{350}^{355} f(x)\,dx$$

$$y = f(x)$$

mean

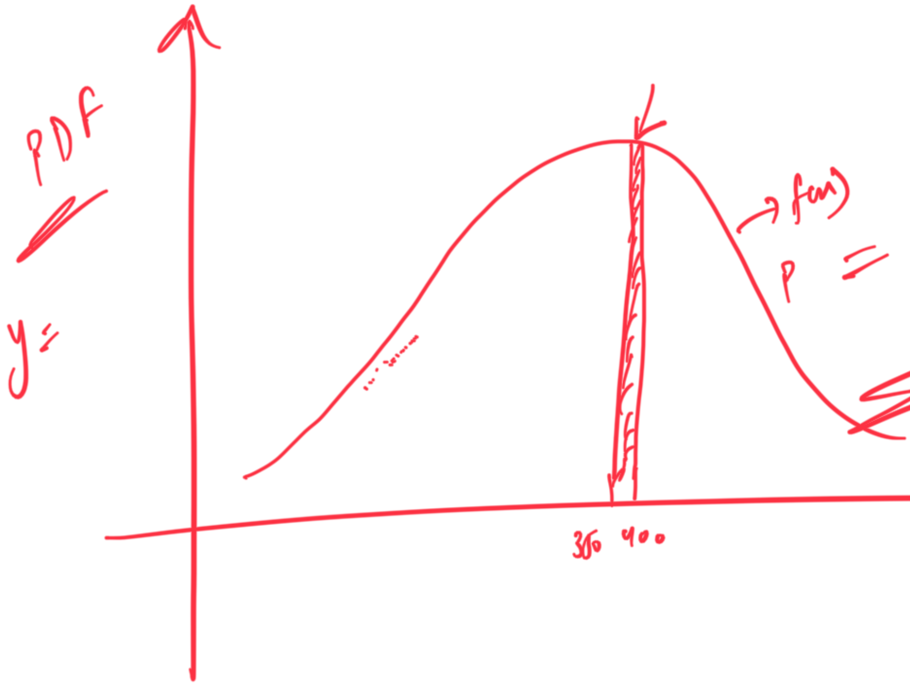
$$\frac{-x^2}{2}$$

Gaussian

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

2.71...

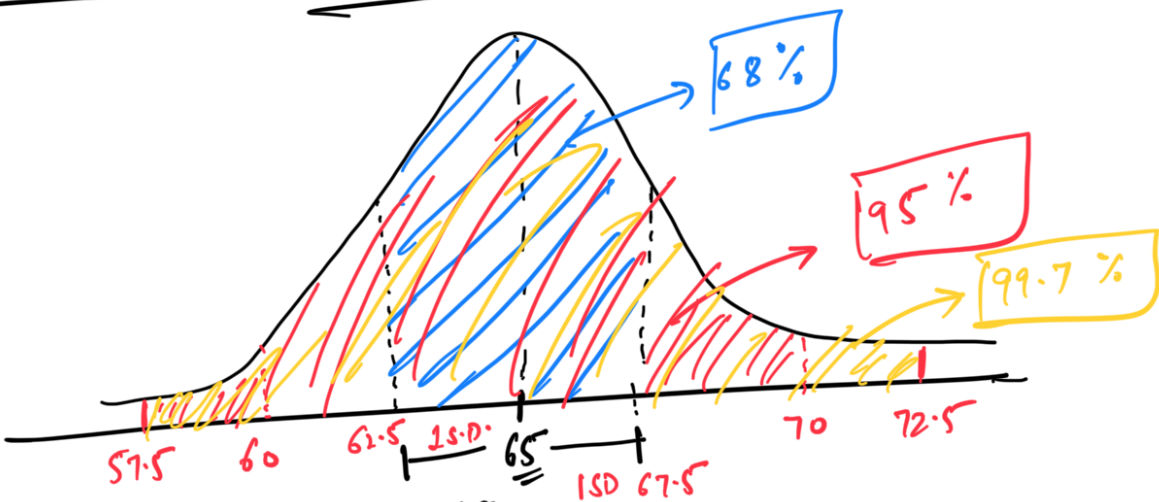
Std. 3.14



Integration

$$P = \int_{350}^{400} f(x)\,dx$$

Empirical Law (68/95/99).



Heights $\Rightarrow$ $\mu = 65$

(SD) $\sigma = 2.5$

1SD $\rightarrow$ $65 + 2.5 = 67.5$

2S.D. $\rightarrow$ $65 + 2(2.5) = 70$

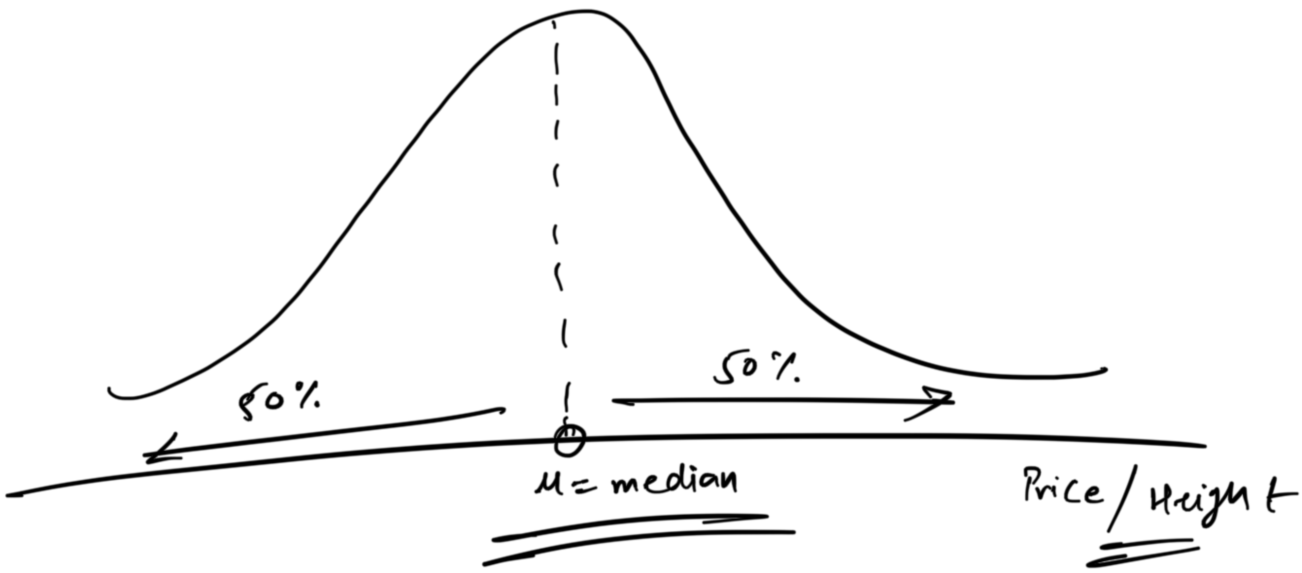
Within 1 S.D. from the mean $\rightarrow$ 68% of the data is captured.

Percentage of people with height :

① within 1 S.D. [62.5 – 67.5] $\rightarrow$ 68%.

" 2 " [60 – 70] $\rightarrow$ 95%

" 3 " [57.5 – 72.5] $\rightarrow$ 99.7%.



$\mu \Rightarrow 65 \longrightarrow 50\%$

$62.5 - 67.5 \rightarrow 68\%$

$65 - 67.5 \rightarrow \frac{68}{2} = 34\%$

50%

+

< 67.5

$50\% + 34\% = 84\%$

Q:

65

2.5

62.5 — 67.5 — 68%

$\boxed{60 - 70}$ → $\boxed{95\%}$

$\boxed{57.5 - 72.5}$ — 99.7%

(60 − 72.5)

(60 − 65) , (65 , 72.5)

$\frac{95}{2} = 47.5 + \frac{99.7}{2} = 49.85$

97.35

66 , 69

$\mu = 65$ → 67.5

$\sigma = 2.5$ ✓

$\rightarrow \boxed{69.1} \rightarrow$ how many S.D. away from the mean?

70 $\rightarrow$ 2 S.D.

67.5 $\rightarrow$ 1 S.D.

Say,

69.1 is 'Z' S.D. away from mean.

$$\rightarrow 67.5 = 65 + 1\,(SD)$$

$$70 = 65 + 2\,(S.D.)$$

Similarly,

$$69.1 = 65 + Z\,(S.D.)$$

$$69.1 = 65 + z\,(2.5)$$

$$z = \frac{69.1 - 65}{2.5}$$

69.1 is $\boxed{Z = 1.64 \text{ S.D.}}$ away from the mean.

$\hookrightarrow$ Z-score

Z - Score

$$\boxed{Z = \frac{x - \mu}{\sigma}}$$

$x \rightarrow$ no. for which we need to cal. Z-score

$\mu \rightarrow$ mea

$\sigma \rightarrow$ S.D.

Z-score $\rightarrow$ no. of S.D. far away from the mean for any data point.

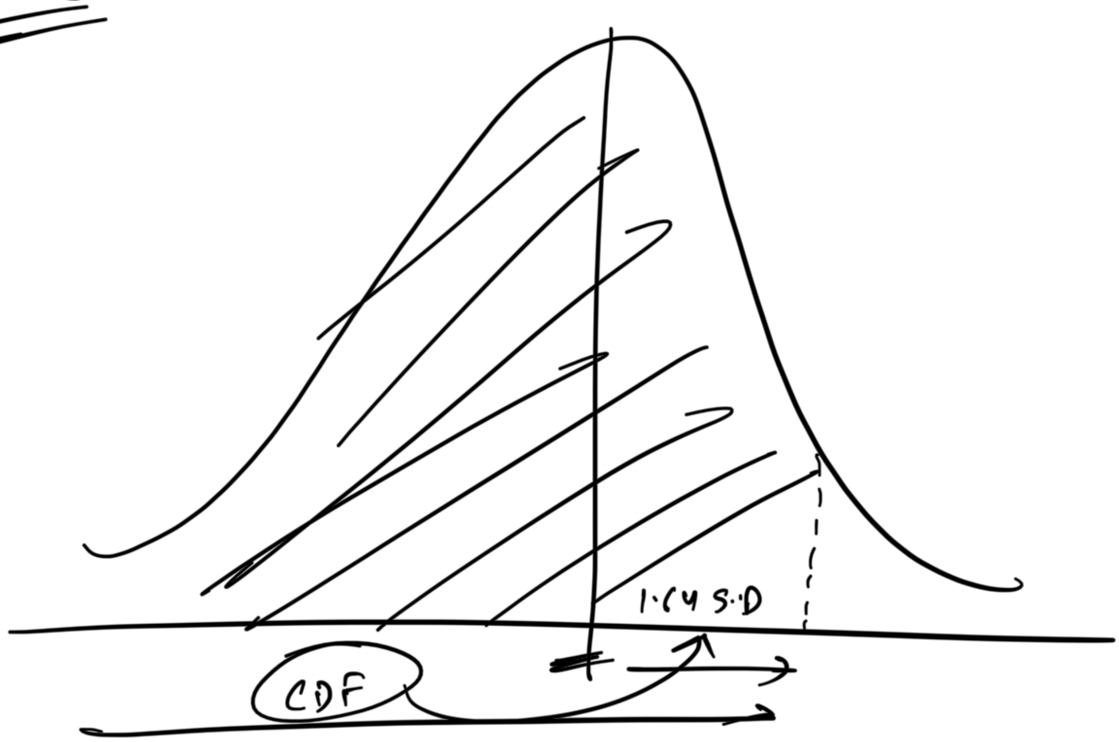
Next $\rightarrow$ What fraction / % of people are shorter than 69.1 inches?

x S.D

$P(\leq \boxed{1.64}$ $\leftarrow$ CDF

$z = 1.64$

Scipy



$$z\text{ - score} = 1.5$$

$$\mu = 50$$

$$\sigma = 2$$

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

$$(z * \sigma) + \mu = x$$

$$x = (1.5 \times 2) + 50$$

$$\boxed{= 53}$$

Q. 96 % people are shorter than me.

Q. 96 % people

What is my height?

PPF → Percent Point Function

norm.ppf( )

$z =$ norm.ppf (0.96)

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

Microsoft

Q: skaters take a mean of 7.42 secs & S.D. 0.34 secs.

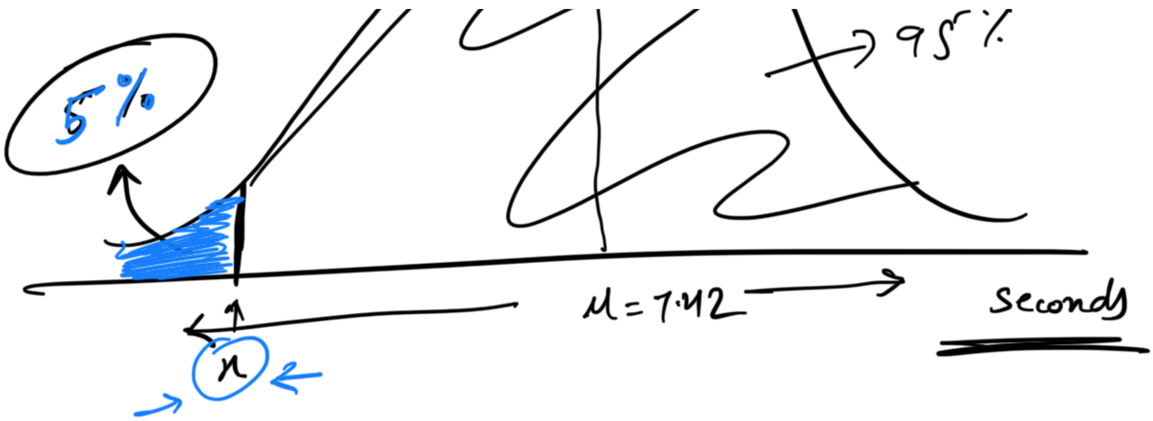
for 500 meters

i ?

What should be the $\boxed{\text{speed}}$ such that the skater is faster than 95% of the competitors?

$\mu = 7.42$

$\sigma = 0.34$



$x \longrightarrow$ 5th percentile $\rightarrow x$

Speed of the skater

5% of the times are $< x$

5th perc $= x$

SK $\longrightarrow$ time x

95%

$x \longrightarrow$ 5% of people lesser time than X

S'