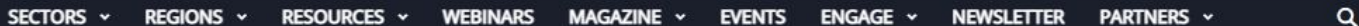


Workshop: How to catch fraudulent consumers in 30 mins using A.I.?

Case study on local power distribution company's data shows great success

Dr Hassan A. Khalid & Dr Wajahat Hussain
NUST, Islamabad





News

MOST READ

[f](#)
[t](#)
[in](#)

The government will pay a \$1 million penalty fee to terminate the \$5 billion loan. The result



May 11, 2021



May 11, 2021



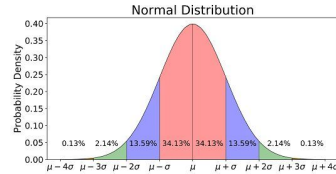
May 12, 2021

Big Data vs Manual کہاتا

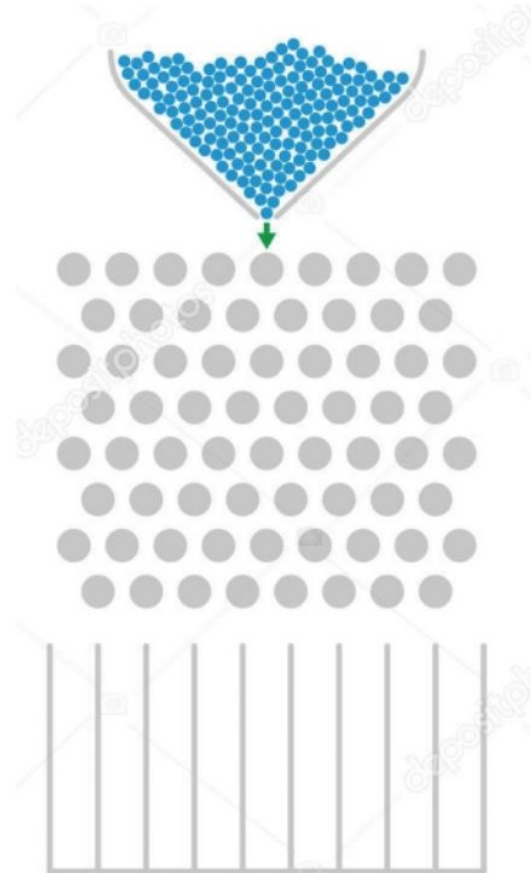


Let's Ask Mr. Gauss What is Normal Behaviour

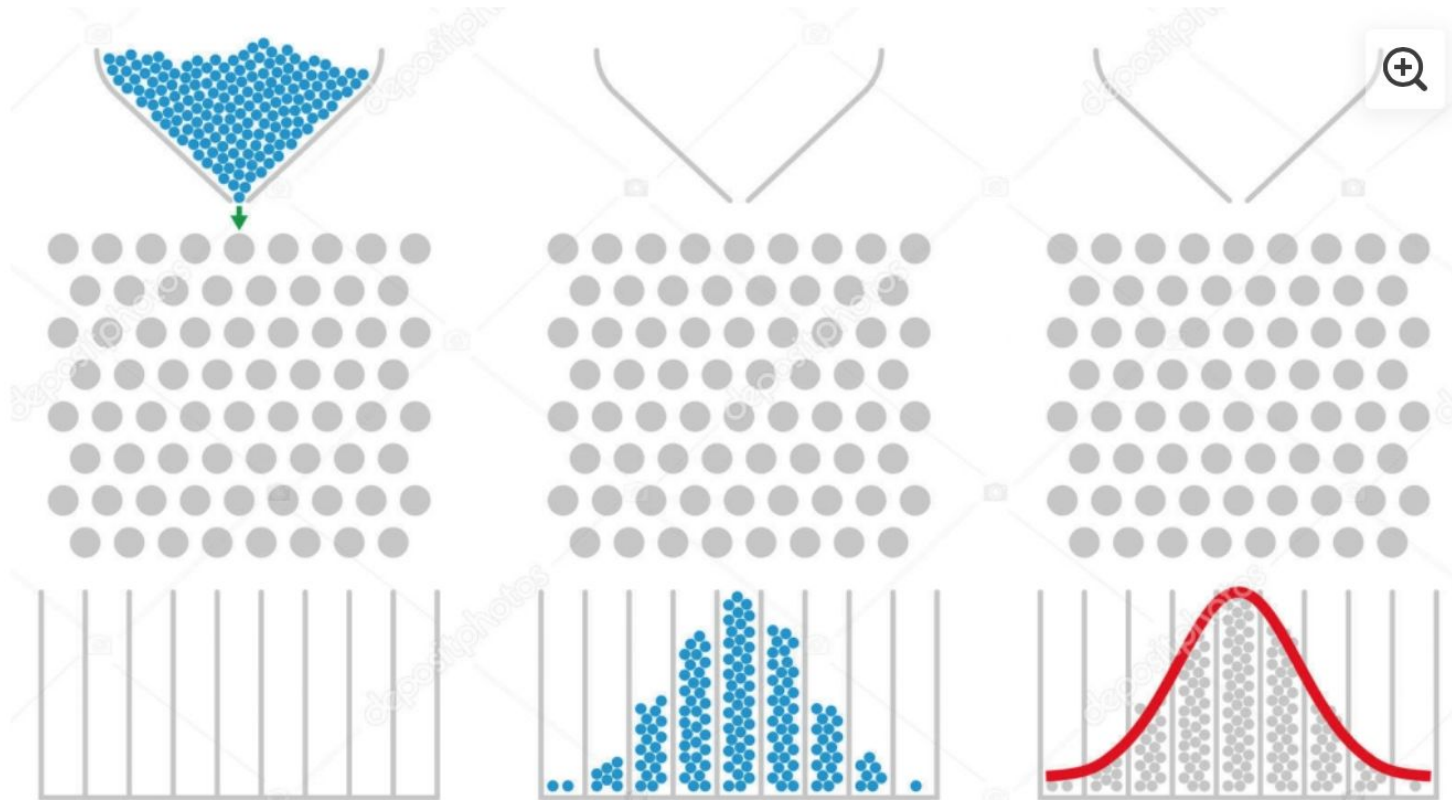
Dr Wajahat Hussain
SEECs, NUST



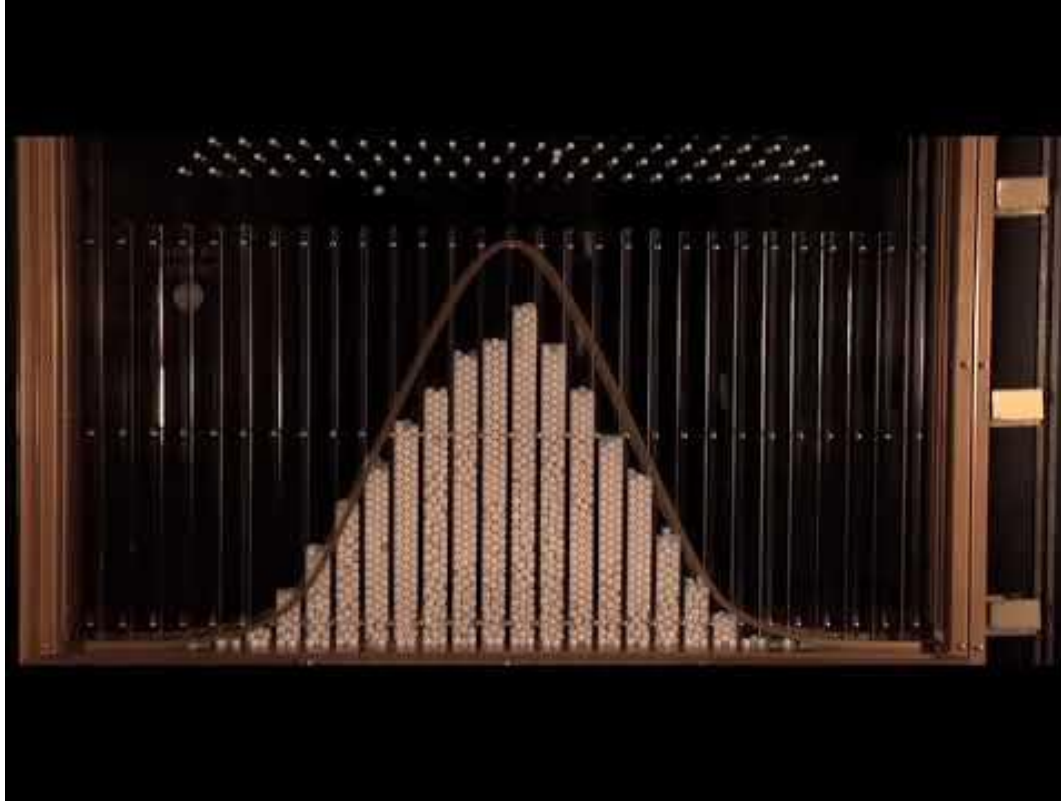
Can you guess in which bin the ball will land?



Can you guess in which bin the ball will land?

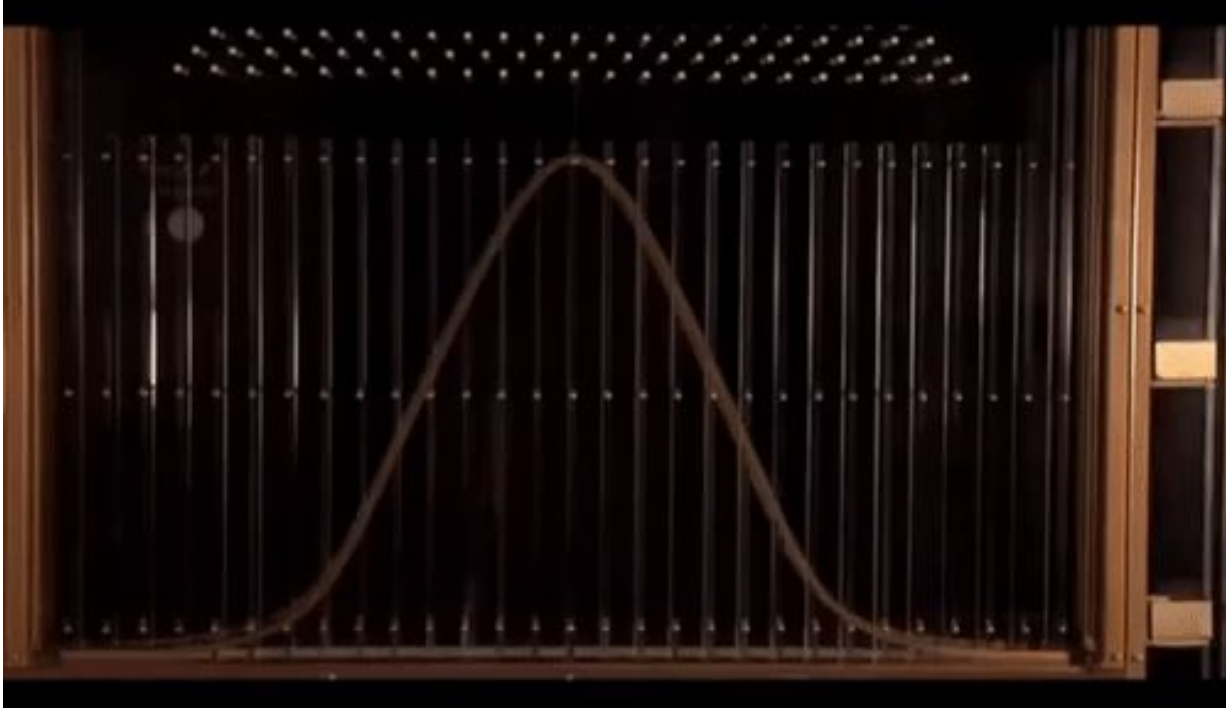


Can you guess in which bin the ball will land? Lets see



<https://www.youtube.com/watch?v=M8t48Vpvl48>

Can you guess in which bin the ball will land? Lets see



https://thumbs.gfycat.com/AggressiveAromaticBuckeyebutterfly-size_restricted.gif

Can you guess in which bin the ball will land? Lets see



<https://www.youtube.com/watch?v=4HpvBZnHOVI>

Can you guess in which bin the ball will land? Lets see

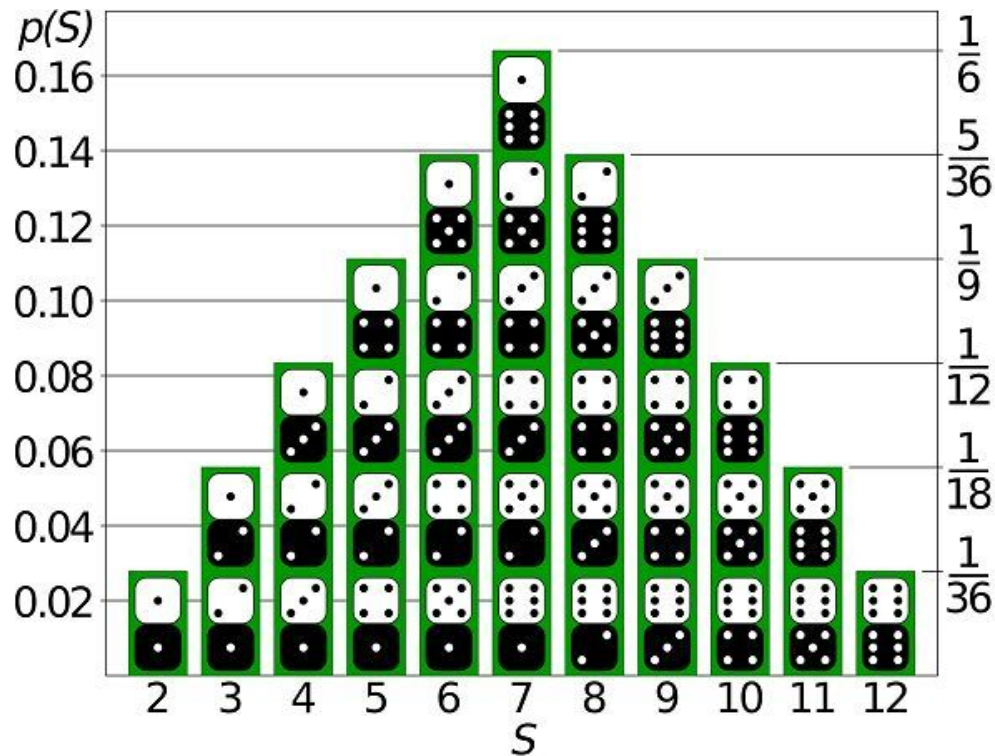


<https://www.youtube.com/watch?v=Bampgm0HKDU>

Rolling A Dice



Rolling A Dice

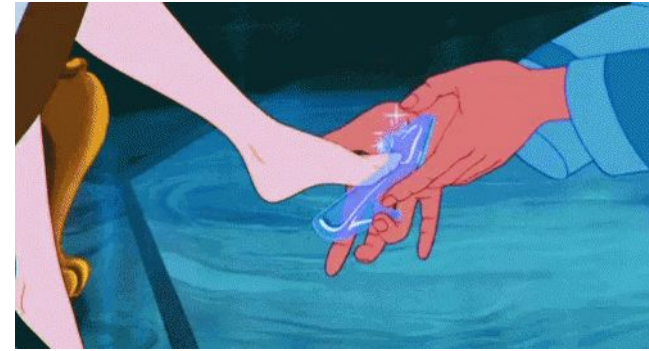
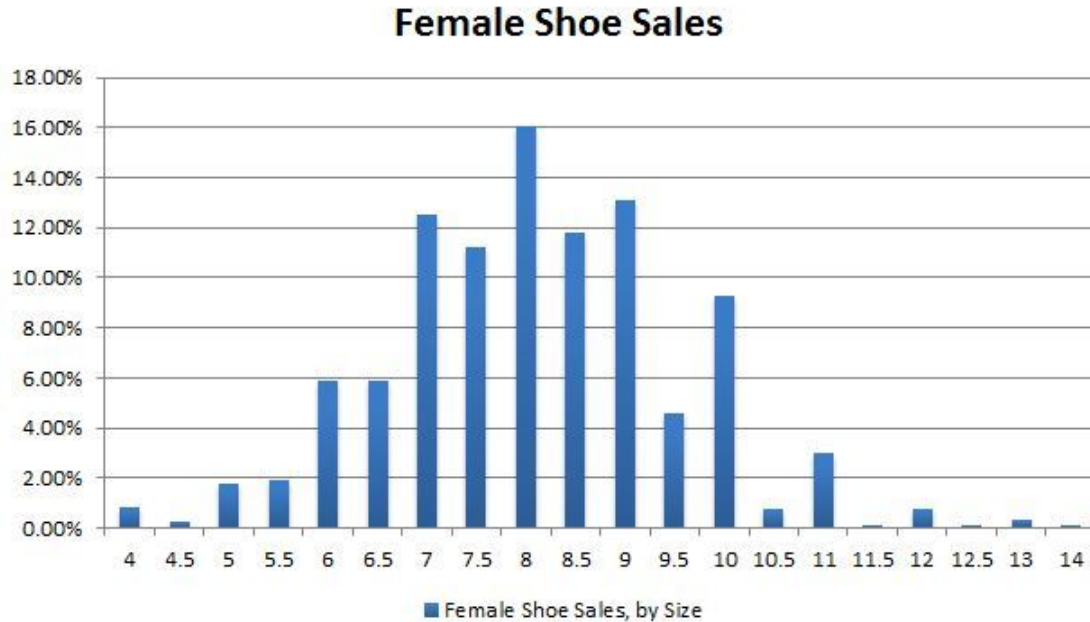


What is the most common shoe size?

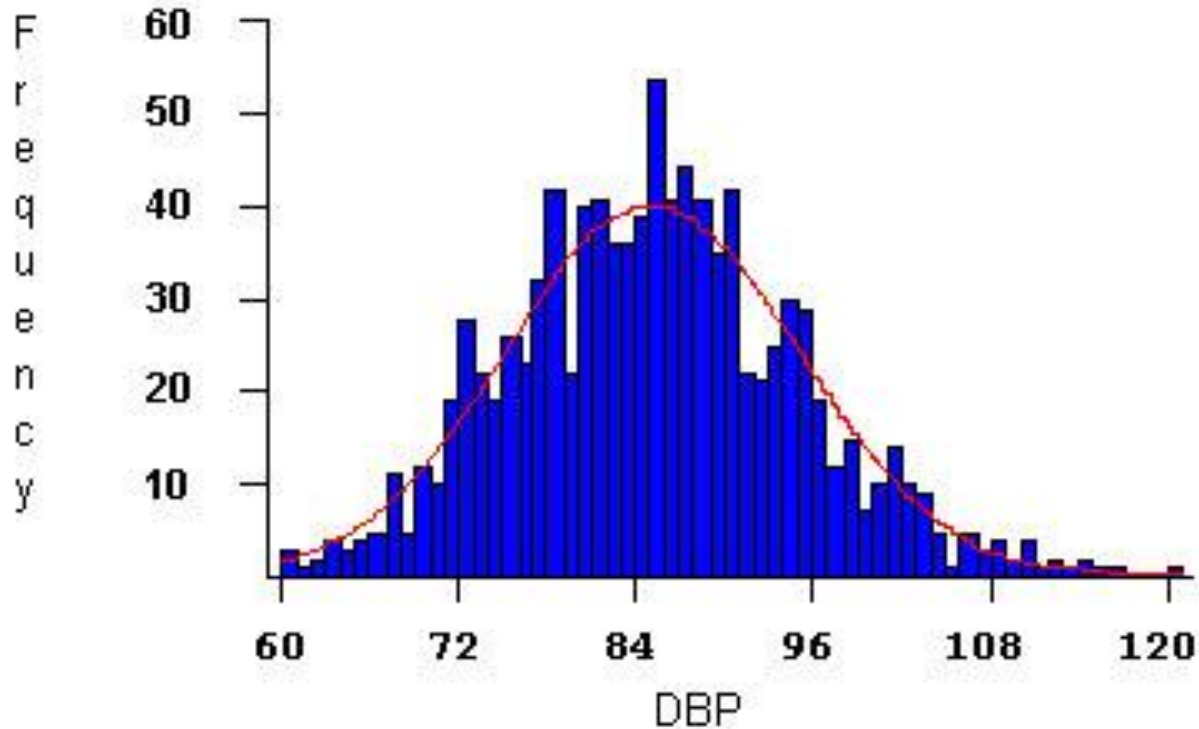


<https://studiousguy.com/real-life-examples-normal-distribution/>

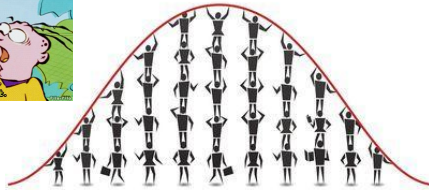
What is the most common shoe size?



Distribution of Blood Pressure

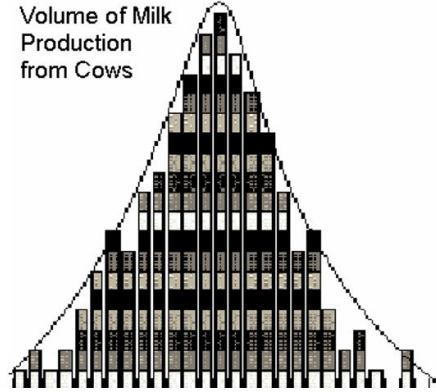


Normal distribution examples

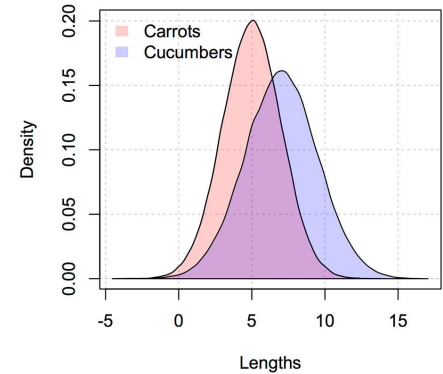


Height Distribution

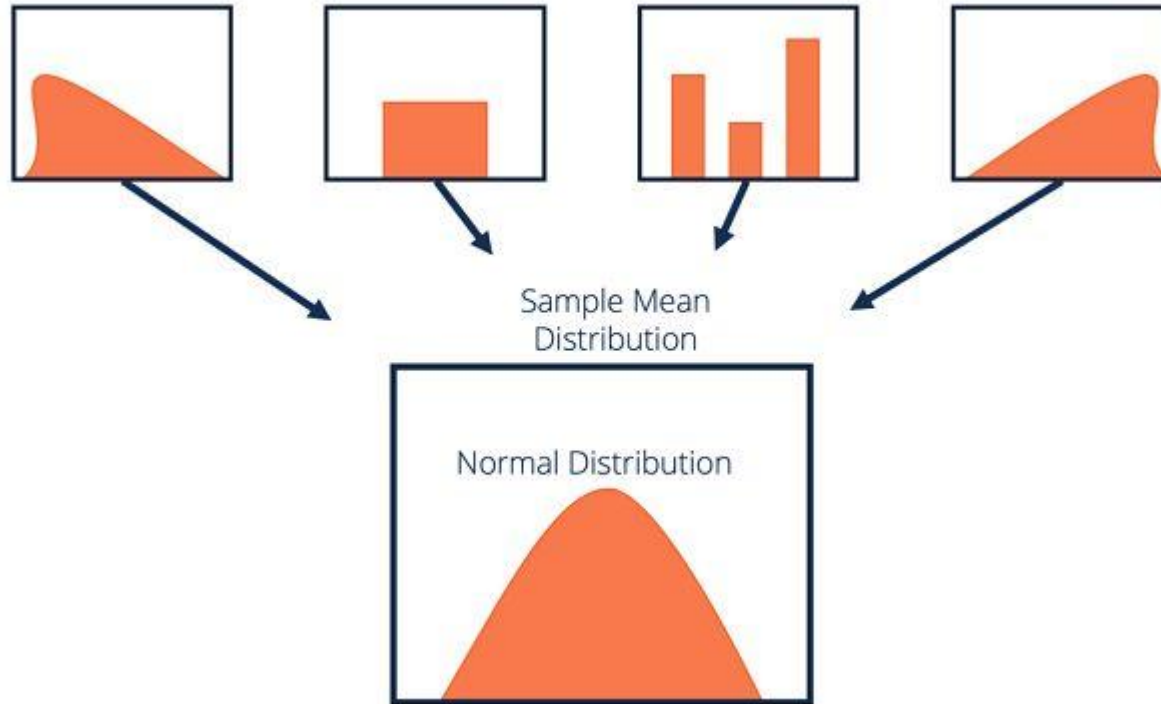
Volume of Milk
Production
from Cows



Distribution of carrots and cucumbers

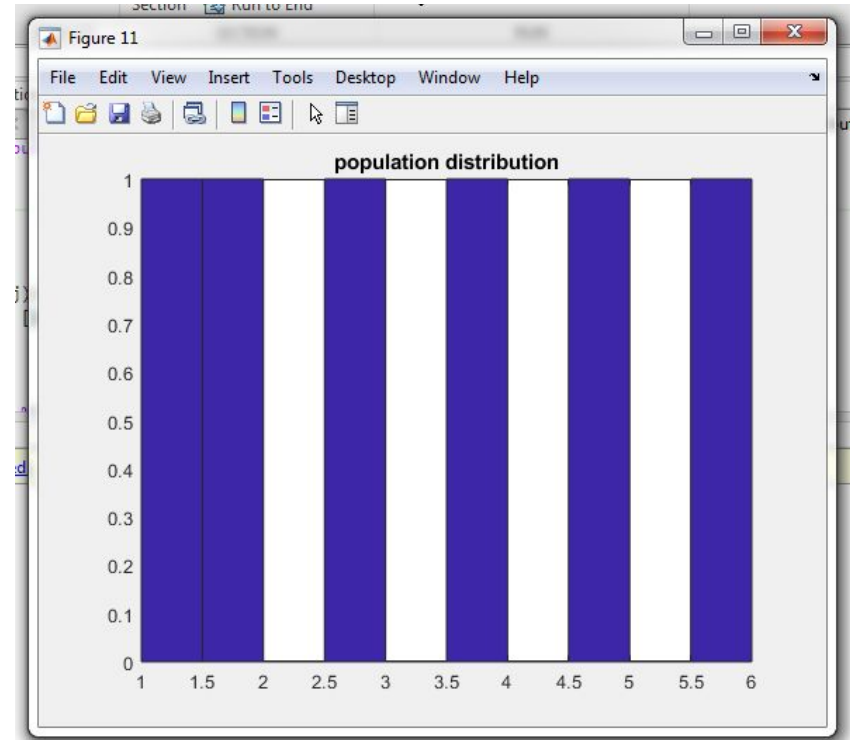


Why normal distribution is normal?



Dice Example

population mean is 3.500000



Dice Example: Two throws

population mean is 3.500000

```
population mean is 3.500000
```

```
sampleMeanDist =
```

```
Columns 1 through 9
```

1.00	1.50	2.00	2.50	3.00	3.50	1.50	2.00	2.50
------	------	------	------	------	------	------	------	------

```
Columns 10 through 18
```

3.00	3.50	4.00	2.00	2.50	3.00	3.50	4.00	4.50
------	------	------	------	------	------	------	------	------

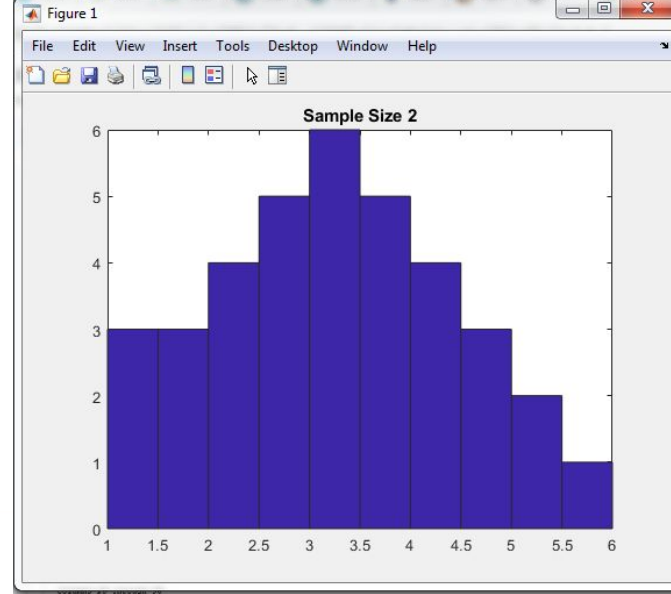
```
Columns 19 through 27
```

2.50	3.00	3.50	4.00	4.50	5.00	3.00	3.50	4.00
------	------	------	------	------	------	------	------	------

```
Columns 28 through 36
```

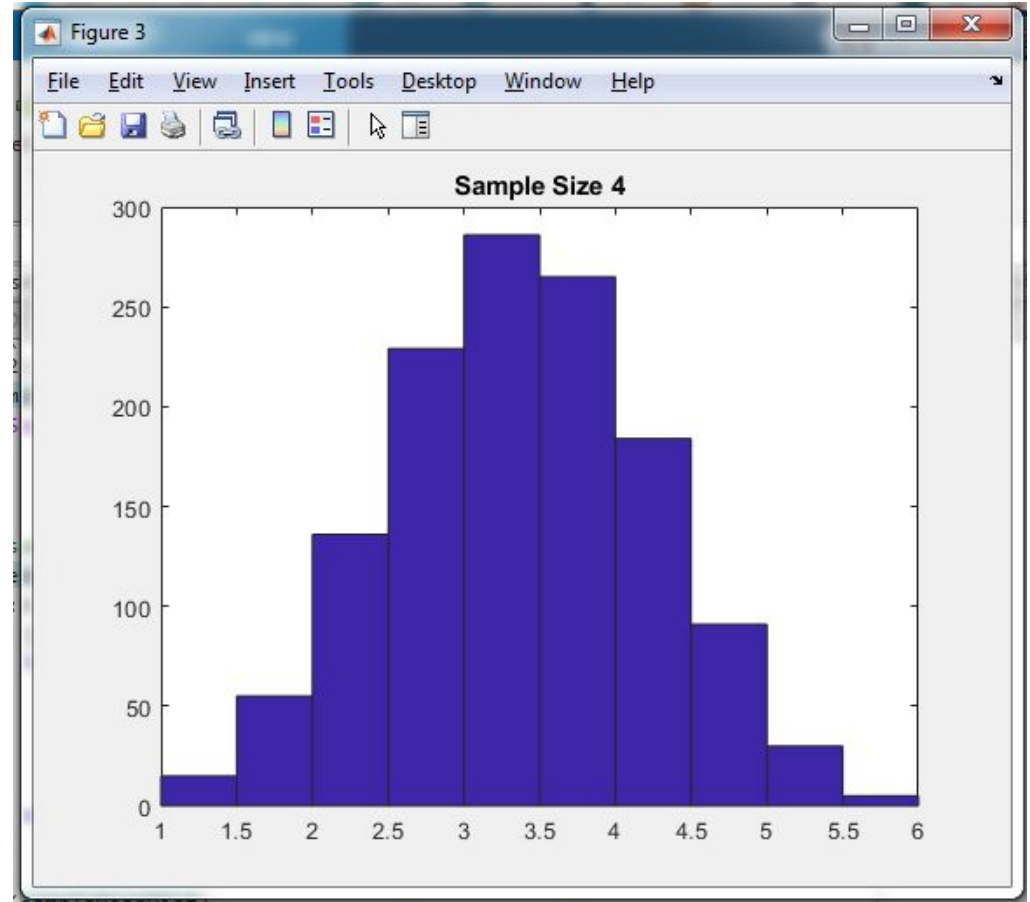
4.50	5.00	5.50	3.50	4.00	4.50	5.00	5.50	6.00
------	------	------	------	------	------	------	------	------

```
mean of mean is 3.500000
```



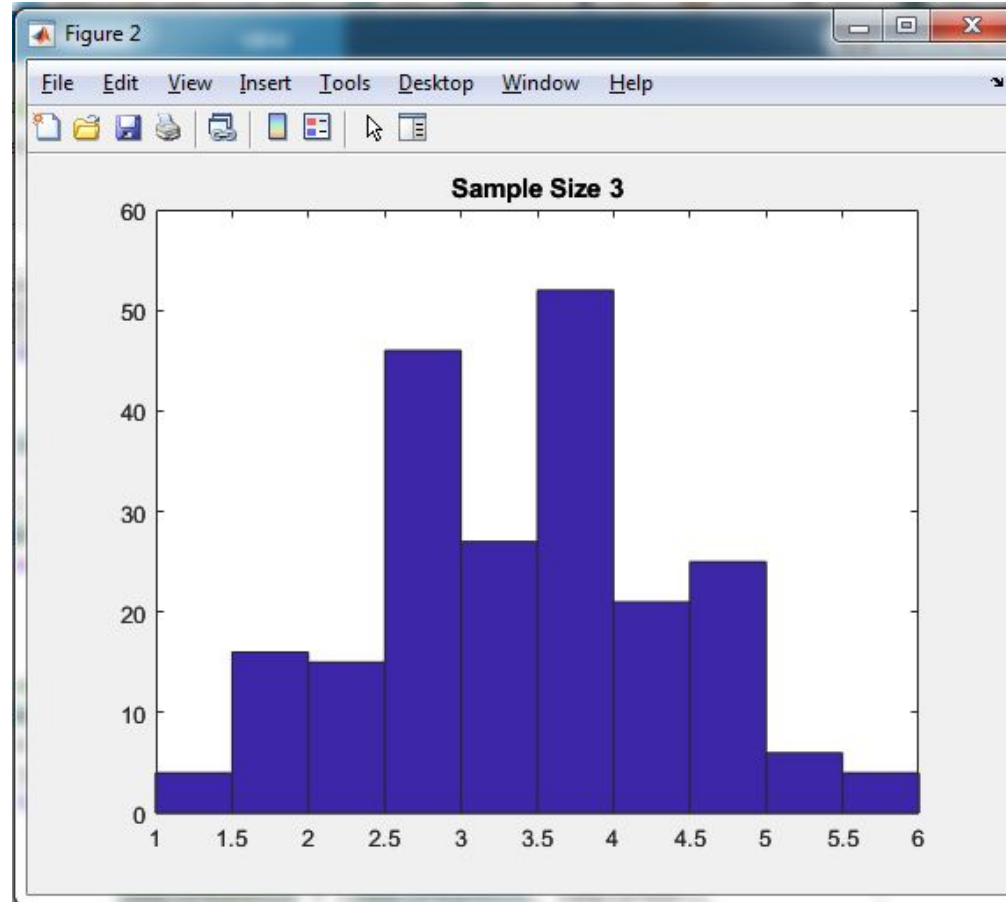
Dice Example: Four throws

population mean is 3.500000



Dice Example: Four throws

population mean is 3.500000



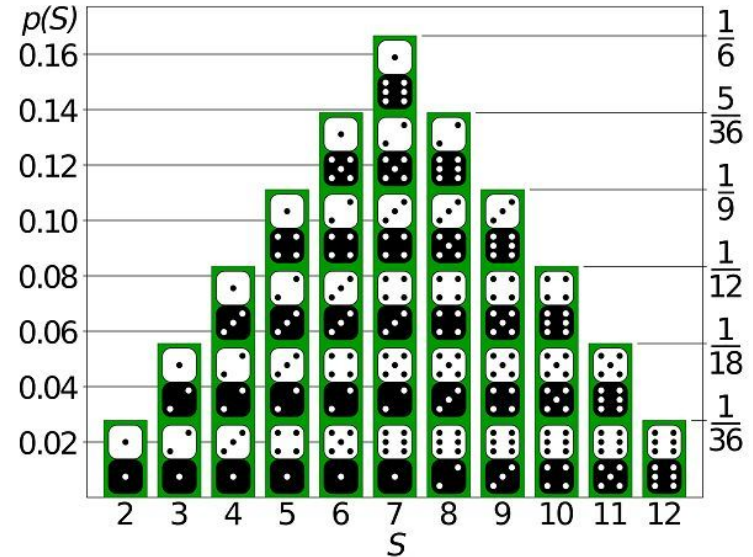
Normal Distribution

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

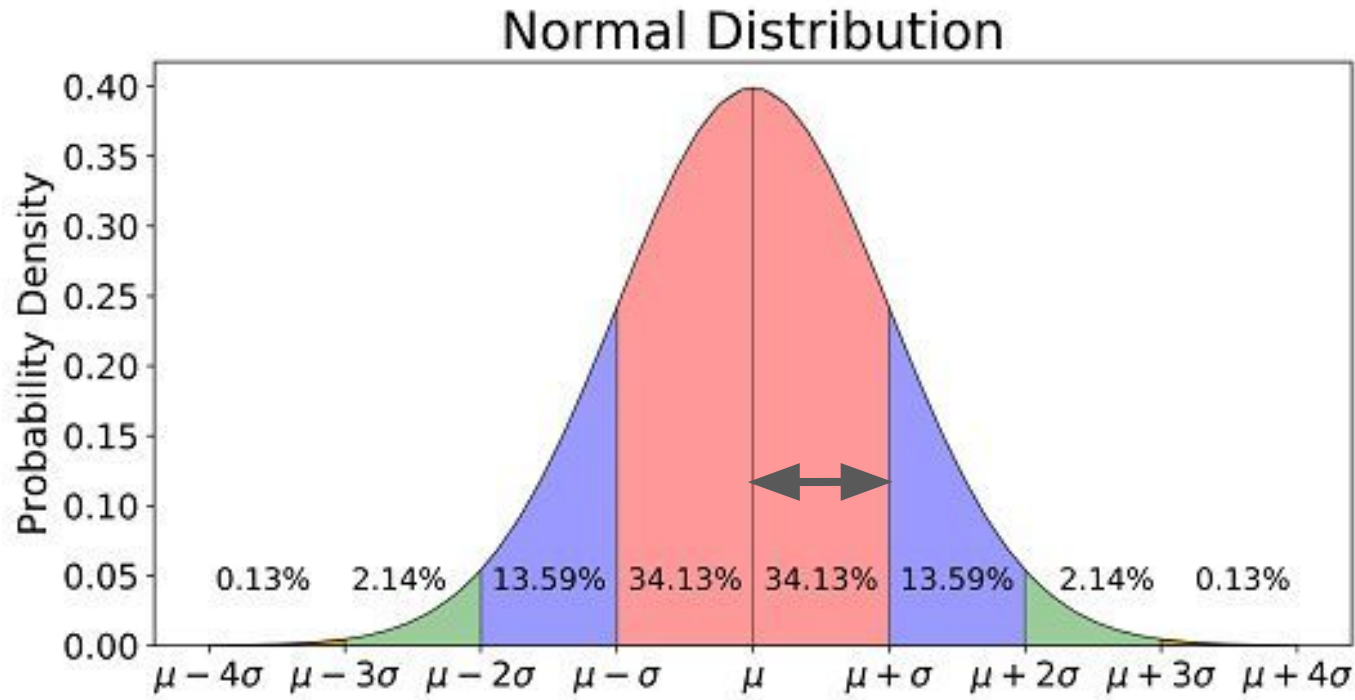
$f(x)$ = probability density function

σ = standard deviation

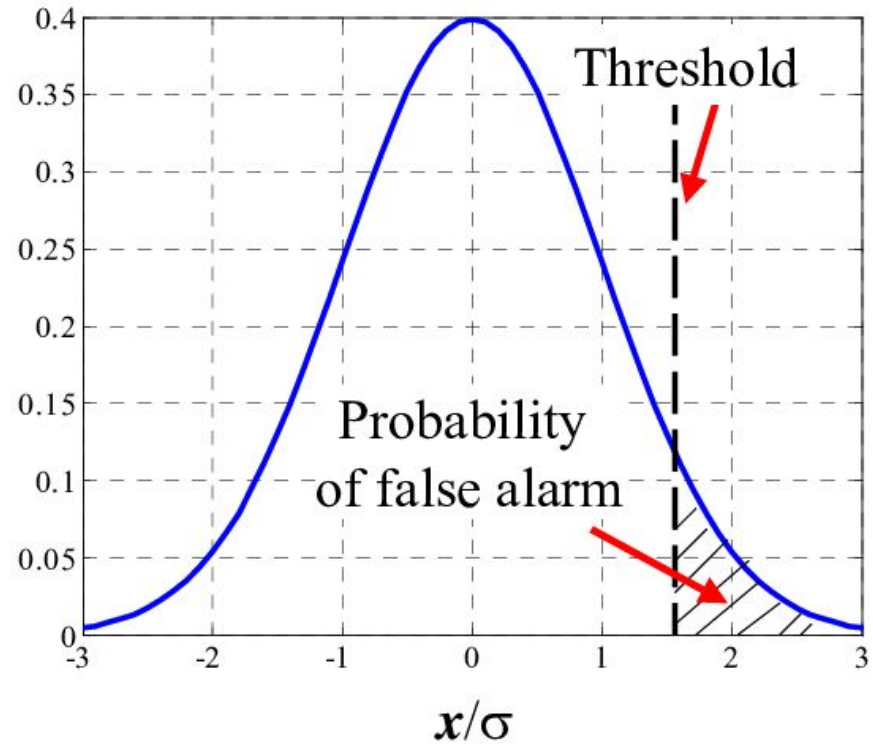
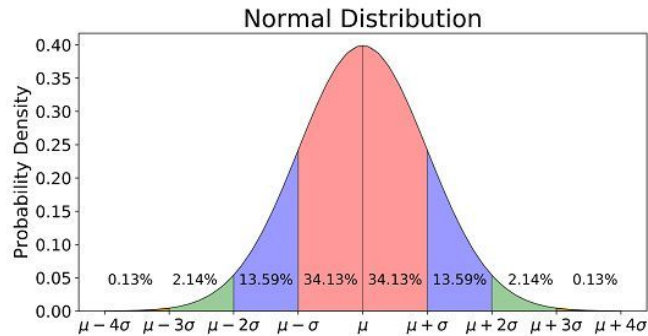
μ = mean



What is abnormal pattern?



What is abnormal pattern?



What is abnormal pattern?

Normal Distribution

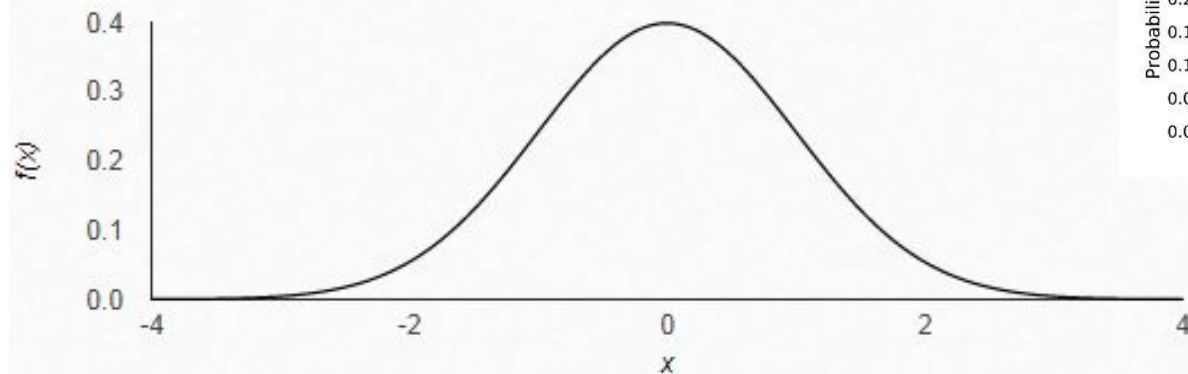
$$X \sim N(\mu, \sigma)$$

$$\mu = 0$$

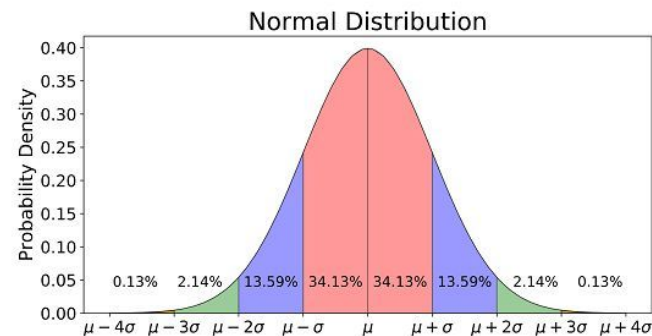
$$\sigma = 1$$

$$x =$$

$$P(X > x) =$$



$$\mu = E(X) = 0 \quad \sigma = SD(X) = 1 \quad \sigma^2 = Var(X) = 1$$



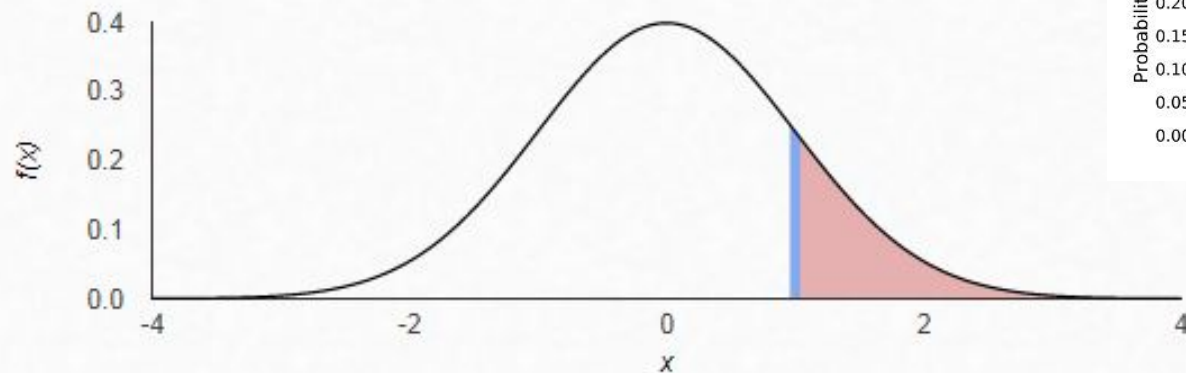
What is abnormal pattern?

Normal Distribution

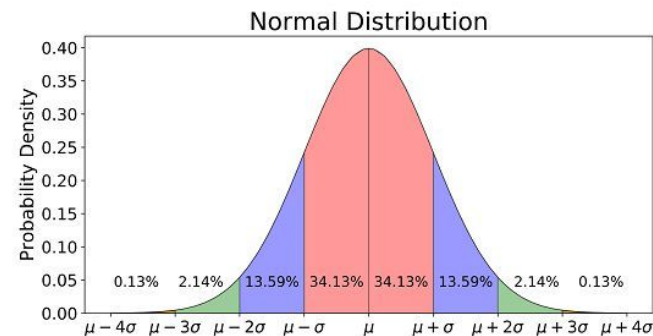
$$X \sim N(\mu, \sigma)$$

$\mu =$ $\sigma =$

$x =$ $P(X > x) =$



$$\mu = E(X) = 0 \quad \sigma = SD(X) = 1 \quad \sigma^2 = Var(X) = 1$$



What is abnormal pattern?

Normal Distribution

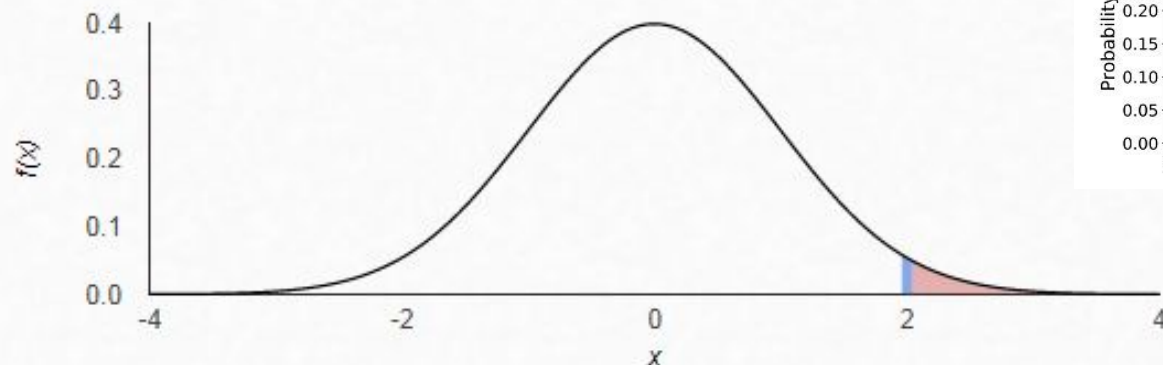
$$X \sim N(\mu, \sigma)$$

$\mu = 0$

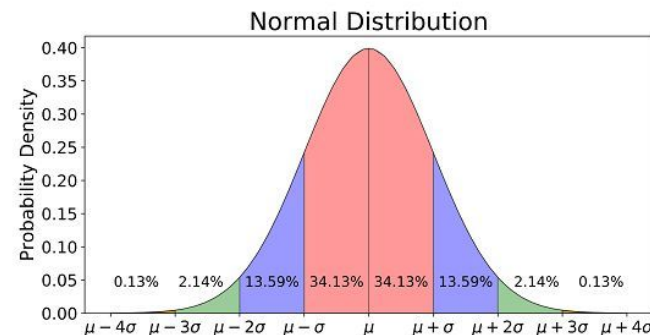
$\sigma = 1$

$x = 2$

$P(X > x) = 0.02275$



$$\mu = E(X) = 0 \quad \sigma = SD(X) = 1 \quad \sigma^2 = Var(X) = 1$$



What is abnormal pattern?

Normal Distribution

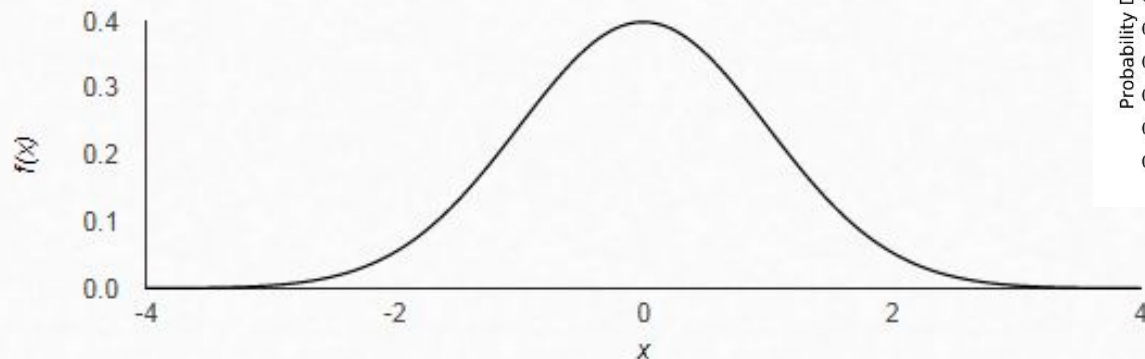
$$X \sim N(\mu, \sigma)$$

$\mu =$

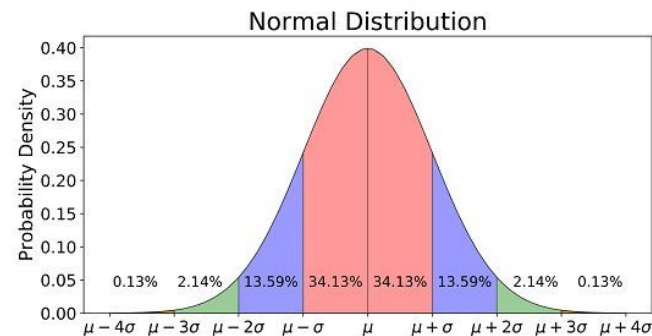
$\sigma =$

$x =$

$P(X > x) =$



$$\mu = E(X) = 0 \quad \sigma = SD(X) = 1 \quad \sigma^2 = Var(X) = 1$$



Central Limit Theorem Helps Us Guess About Population

Before we introduce the formal statistical terms and procedures, we'll work our chapter-opening sports-complex problem all the way through. Recall that the aluminum roofing sheets have a claimed average thickness of 0.04 inch and that they will be unsatisfactory if they are too thick or too thin. The contractor takes a sample of 100 sheets and determines that the sample mean thickness is 0.0408 inch. On the basis of past experience, he knows that the population standard deviation is 0.004 inch. Does this sample evidence indicate that

$$\begin{aligned}\sigma_{\bar{x}} &= \frac{\sigma}{\sqrt{n}} \\ &= \frac{0.004 \text{ in.}}{\sqrt{100}} \\ &= \frac{0.004 \text{ in.}}{10} \\ &= 0.0004 \text{ in.}\end{aligned}$$

$$\begin{aligned}z &= \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \\ &= \frac{0.0408 - 0.04}{0.0004} \\ &= 2 \leftarrow \text{Standard errors of the mean}\end{aligned}$$

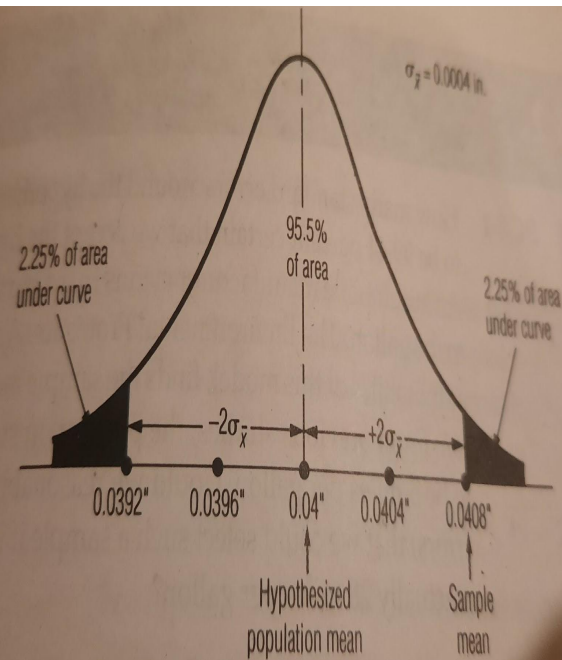


FIGURE 8-1 Probability that \bar{x} will differ from hypothesized μ by 2

Big Data vs Manual کہاتا

