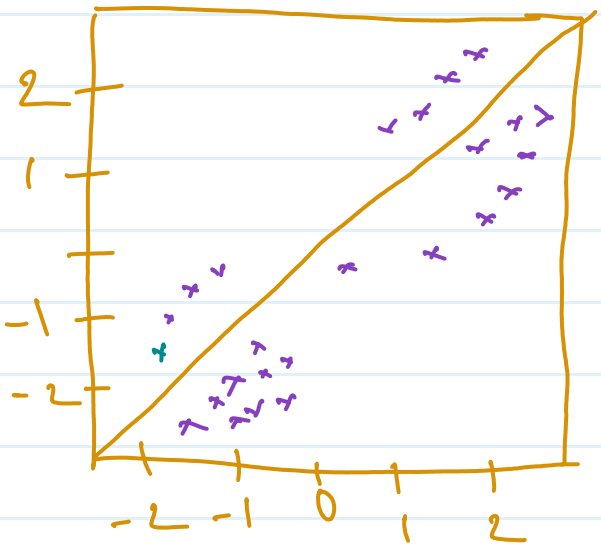
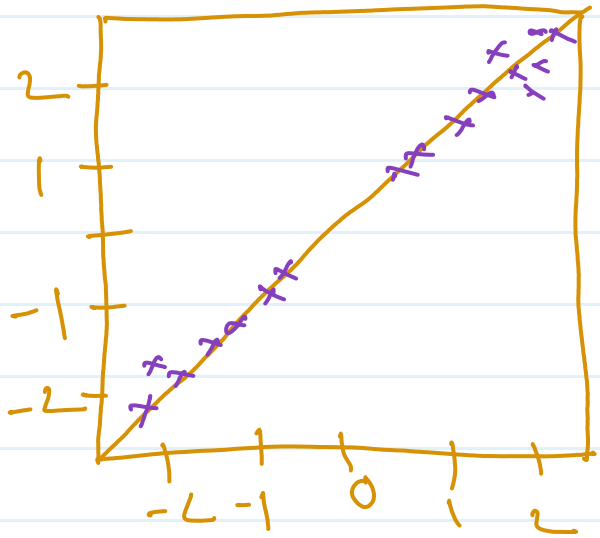
✱ Q-Q plot

Quartile and Quantile plot

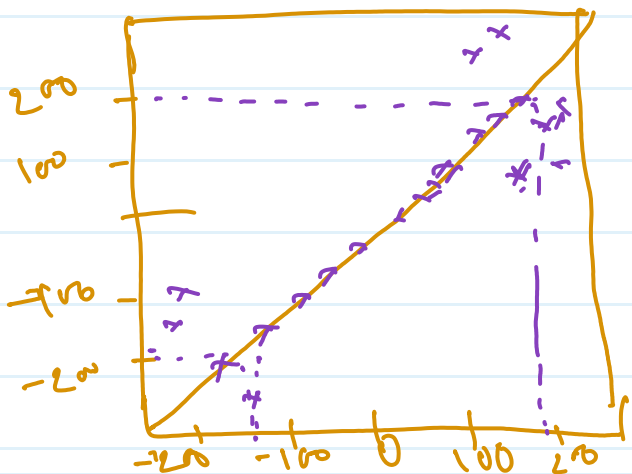


— It is non-Gaussian dist.



— Q-Q.

It is normal dist.



Spearman Rank correlation

$$\gamma_s = \frac{COV(R_x R_y)}{\sigma_{R_x} * \sigma_{R_y}}$$

Eg.

	Eco growth X	Money Y
4	2.1	8
3	2.5	12
2	3.6	10
1	4.0	14

Rank X	Rank Y
4	4
3	2
2	3
1	1

mean x = 2.5

mean y = 2.5

$(X-\bar{X})$			$(Y-\bar{Y})$	
$4-2.5$	$=$	1.5	$(4-2.5)$	1.5
$3-2.5$	$=$	0.5	$(2-2.5)$	-0.5
$2-2.5$	$=$	-0.5	$(3-2.5)$	0.5
$1-2.5$	$=$	-1.5	$(1-2.5)$	-1.5

$$\text{S.D. } \sigma_X = \sqrt{\frac{(1.5)^2+(0.5)^2+(-1.5)^2+(-0.5)^2}{4-1}}$$

$$= \sqrt{1.66} = 1.288$$

$$\text{S.D. } \sigma_Y = \sqrt{\frac{(1.5)^2+(-0.5)^2+(0.5)^2+(-1.5)^2}{4-1}}$$

$$= \sqrt{1.66} = 1.288$$

$$\rho(x,y) = \frac{(x_i-\bar{x})(y_i-\bar{y})}{n-1}$$

$$= \frac{(1.5)\times(1.5)+(0.5)\times(-0.5)+(-0.5)(0.5)+(-1.5)\times(-1.5)}{4-1}$$

$$= \frac{2.25-0.25-0.25+2.25}{3}$$

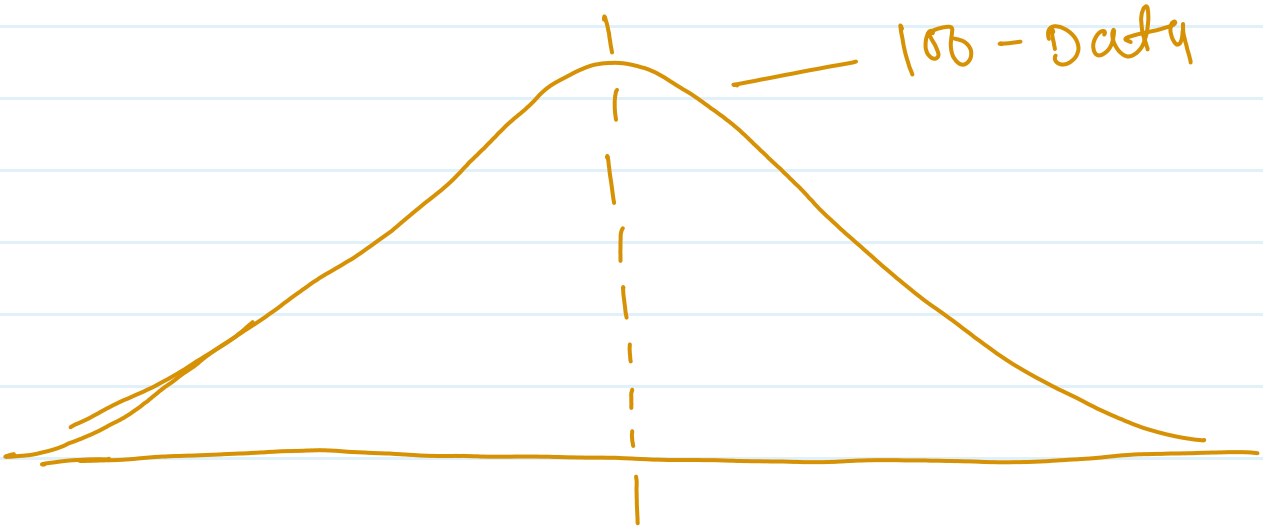
$$\Rightarrow \quad \frac{4}{3} \doteq 1.33$$

$$\gamma_s = \frac{1.33}{1.28 \times 1.28}$$

$$= 0.81 \quad = 81\%$$

✯ Hypothesis Testing

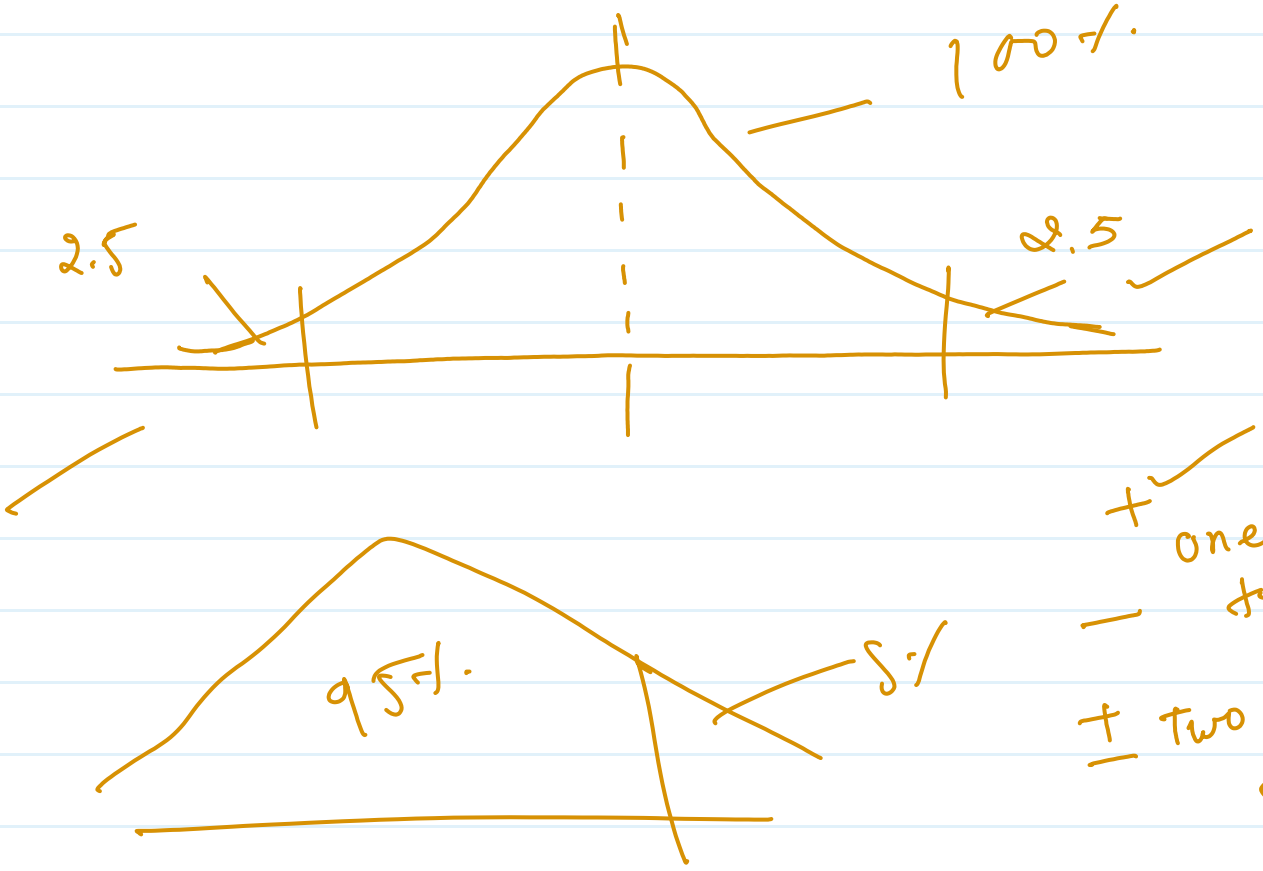
✯ α = cricial value / significant value



Domain expert :- tolerance ± 0.5 mm

sales monthly target — +110
100
— 90

$\Rightarrow \quad \alpha = 5\%$



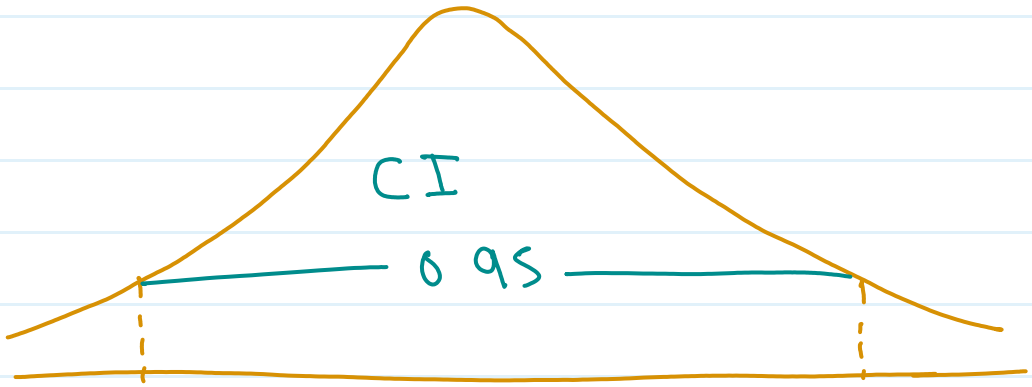
+ one tail test

± two tail test

if we don't have domain expert so we take α value as 5%

100 — 1

5% — 0.05 $\quad 1 - 0.05 = 0.95$



Confidence interval

① Sample size ≥ 30
population SD. σ = given.

Z-test :

② Sample size < 30
Sample SD s is given

t-test

Numeric data

③ Categorical data

chi-square test