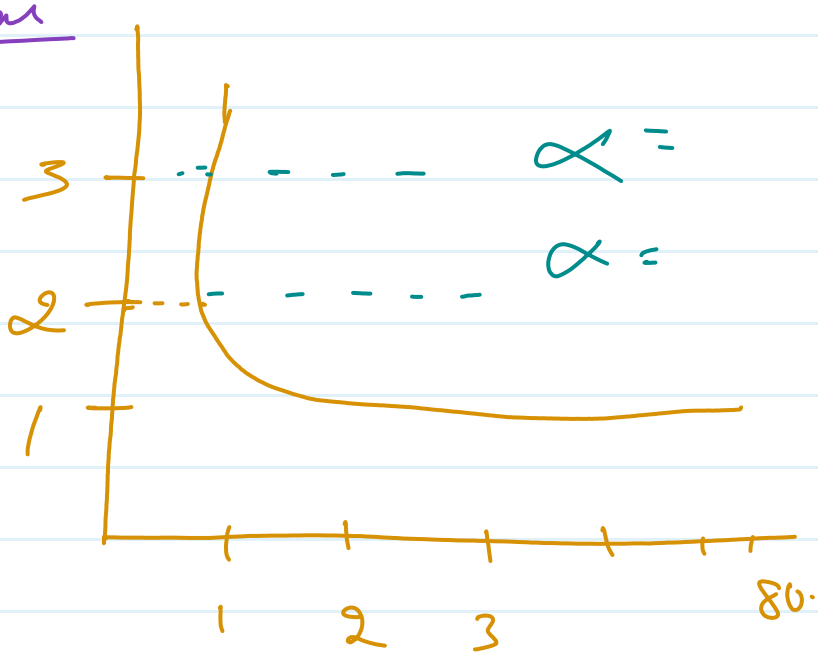
✱ Power law Distribution

80 - 20

Run	Batman
80	20
20	80

W.	E.
80	20
20	80



⇒ 80 - 20

20 - 80

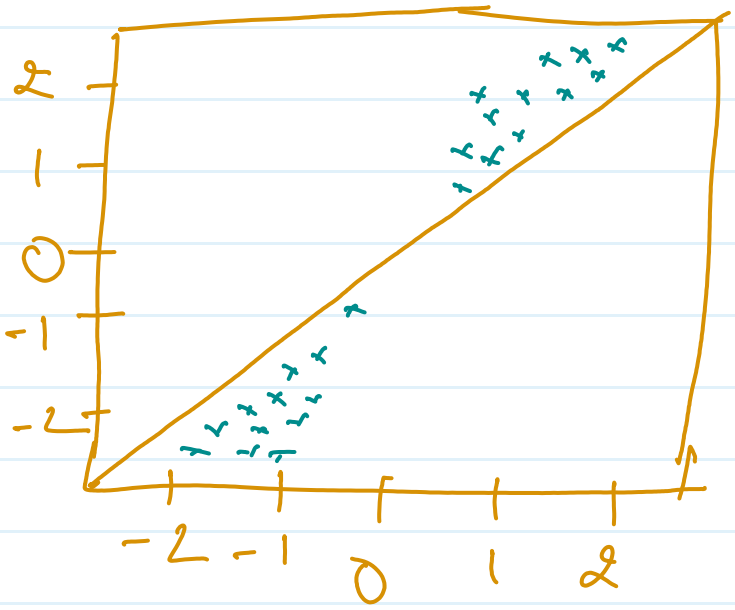
To convert pareto Dist. into gaussian dist. we use Box-Cox transformation.

$$y_i = \frac{x_i^{\alpha} - 1}{\alpha}$$

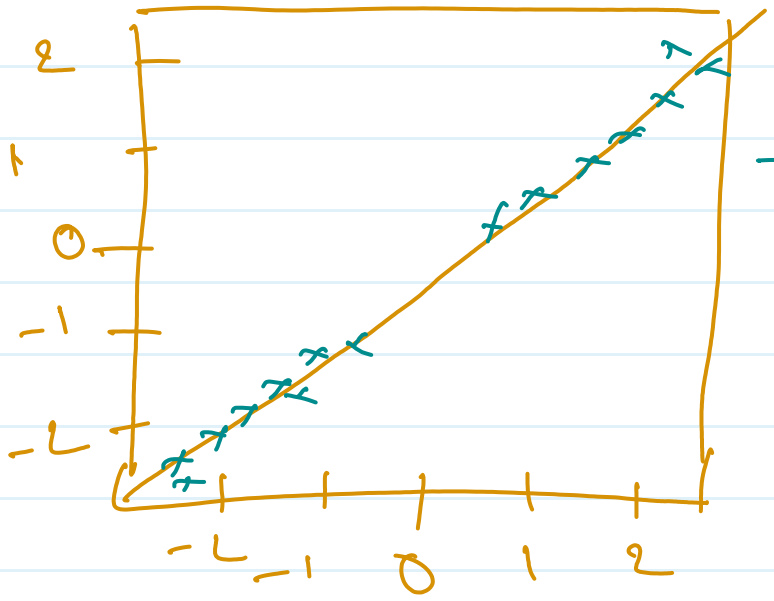
∵ α is height in distribution.

Q-Q plot

(Quantile - Quantile plot)



— non-normal dist.



— normal distribution

Reliability and validity

Reliability

① Test-Retest Reliability :-

use pearson correlation coefficie.

② Internal Consisteney :-

We use cronbanch's alpha (α)

③ Inter-Rater Reliability

We use cohen's kappa (κ)

validity

① content validity :-

It often involves expert judgement.

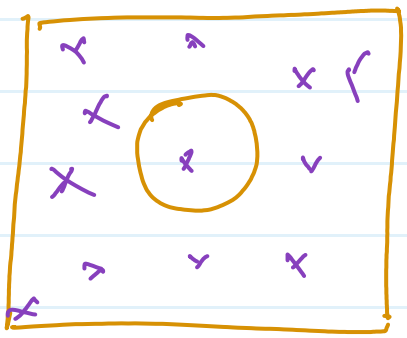
② construct validity :-

We use Average variance Extracted

③ Criterian - Related validity

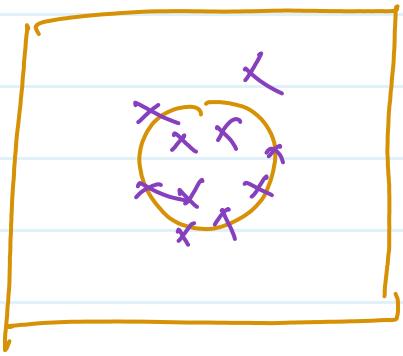
pearson correlation coefficient.

Reliable ✗ validity



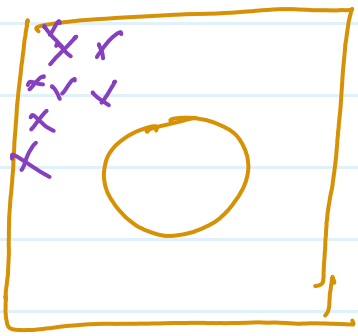
gun

Reliable validity



gun

Reliable ✓ validite ✗



gun

Sample size

How to determine the sample size?

It determine in two steps.

① Calculate sample size for infinite population

② Adjust the sample size for required population.

infinite $$s = z^2 \times p \times (1-p)/m^2$$

where s = Sample size for ∞ popu.

Z = Zscore

p = population proportion [50%]

m = margin of error

Z-score is determined based on confidence level.

Confidence level – The probability that the value of parameta falls withing specified range of value.

conf. leve.	Z-value
90%	1.645
95%	1.960
99%	2.576

if consider 95%, Z = 1.960

margin of error = 0.05

p = 50% = 0.5

$$S = (Z\text{score})^2 \times p \times (1-p) / m^2$$

$$= (1.960)^2 \times 0.5 \times (1-0.5) / (0.05)^2$$

$$\boxed{S = 384.16}$$

Now we must adjust sample size to the required population.

Suppose we have 1,00,000

Formula

$$s = \frac{(s)/1}{+ [(s-1)/\text{population}]}$$

$$s = 384.$$

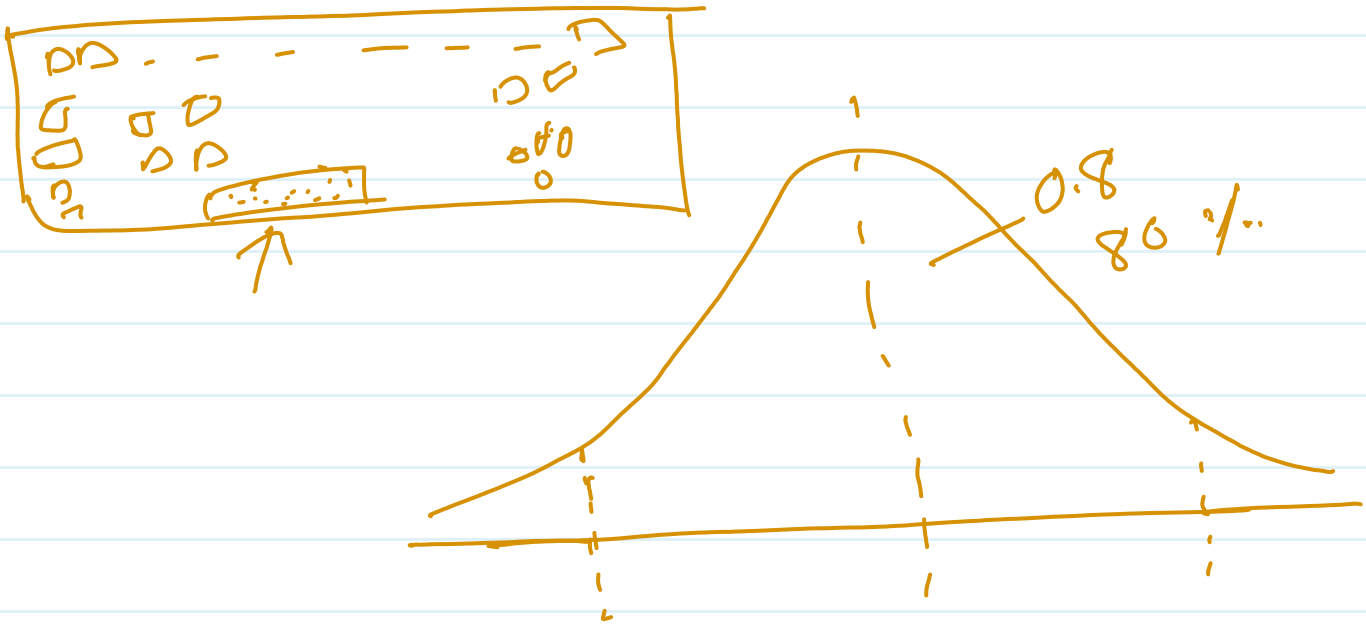
$$s = \frac{384.16}{1 + \left[\frac{(384.16-1)}{100000}\right]}$$

$$= \frac{384.16}{1.0038316}$$

$$= 383.69$$

$$\boxed{s = 384}$$

P - value.



H_0 (Null hypo) – Treat everything same or equal.

For null hypo. testing p value is

$p = 0.01$ we consider.

If we repeate this expe. 100 times then 1

Eg:- coin – four coin

Head – 50 time —

Tail – 50 time. —

We experiment for - 100 times

① Get head 60 times => we reject null hypo.

② head 30 times => reject null hypo

③ head ≤ 1 time => we fail to reject null hypothesis.

Note:- p-value - It is the probability for the null hypo. to be true.