KEY WORDS

Data: Any Observations that have been collected
Stastics: Collect, Analyze, Summarize, Interpret, Draw conclusion
Population: The compelete set of elements being studied.
Samples: Some subset of a population.
Census: Collecting from every member of a population.
* If you collect a sample, it must be collected random.

Types of Data:

P-P Parameter: A characteristic of a population
S-S Stastic: A characteristic of a sample

Qualitative: Non-numeric
(Catagorical) & mathematical are meaningless
Quantitative: Numeric
↳ Discrete, Continuous

4 Levels of Measurement

Nominal: Categories not ordered (Religion)
Ordinal: Can be ordered (Differences are meaningless)
Interval: Ordered, Differences are meaningful,
No "Nature Zero". Like temperature
Ratio: Just like Interval, But with a Nature Zero
"Amount of money"

Design of Experiements / Observations

OBS vs EXP

Measures specific traits, but does not modify subjects.

Apply a treatment and then measure the effect on the subjects.

Random

Each member of a population has an equal chance of being selected in the sample.

4 common Sampling Techniques.

1. Convenience Sample: Use the result that are easy to get (Not Random)
2. Systematic Sampling: Put a Pop in some order and select every "k^{TH}" Member.
3. Stratafied Sample: Break pop into subgroups based on some characteristic. then sample each group.
4. Cluster Sample: Divide pop into "Clusters" (Regardless of characteristic), Randomly select A certain # of clusters, and then collect data from the entire cluster.

Discribing Data

5 characteristics:

1. center — mean, median, mode (In order)
2. variation — how the data is spread
3. distribution
4. outliers
5. changes over time

n = # of items in a sample $\quad \bar{X} = \frac{\Sigma x}{n}$ Sample mean

N = # of items in a population $\quad \mu = \frac{\Sigma x}{N}$ Population mean

Variation:

Ways to measure.

1. Range (does not consider all values)
2. Standard deviation (s^2, σ^2)

Greatly affected by outliers

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

or $$s = \sqrt{\frac{n\Sigma(x^2) - (\Sigma x)^2}{n(n-1)}}$$

Nomal Distribution. Empirical Rule

- 68% w
- 95% will fall w/in 2 std dev "usual"
- 99.7% . . . 3 . .

Measures of relative standing

Z-score: The # of std dev. A data value (x) is away from the mean.

Sample $z = \frac{x-\bar{x}}{s}$ $\quad$ Pop $z = \frac{x-\mu}{\sigma}$

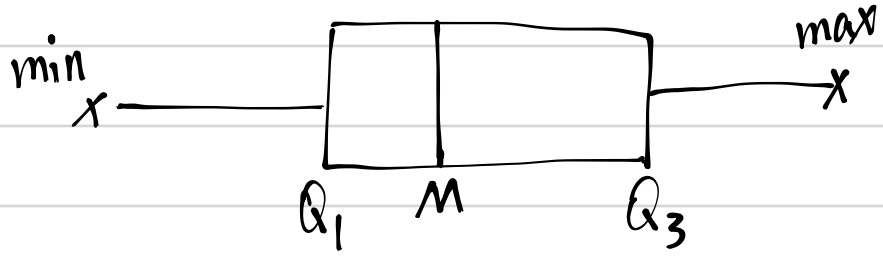
Quartiles Percentiles → 100 Parts

Box Plot

minimum maximum

Q_1 Q_3

median



"At least one" means one or more

$$P(\text{"At least one"}) = 1 - P(\text{"None"})$$

Rules for Permutations

1. n : different items
2. r : the # of items to be arranged out of n
3. Arrangement and order matters.

$$_nP_r = \frac{n!}{(n-r)!}$$

For non-distinct items

$$\frac{n!}{n_1!\,n_2!\,n_3!\cdots}$$

$n_1, n_2 \cdots$ are non-distinct items

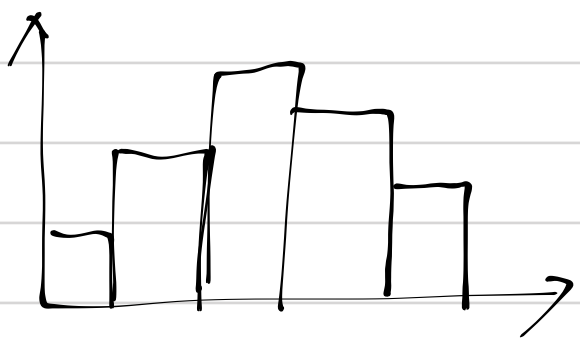
Sometimes order doesn't matter.

Combinations:

1. n. distinct rides
2. select r items
3. Arrangement doesn't matter.

$$_nC_r = \frac{n!}{(n-r)!\,r!}$$

Histogram form probility distribution



(Expected Value)

Mean: $\mu = \frac{\Sigma(xf)}{N} = \Sigma[x \cdot P(x)]$

Variance:

$$\sigma^2 = \Sigma[x^2 \cdot P(x)] - \mu^2$$