

One way which you can describe a variable is by detailing its values and the number of times those values occur — the distribution.

Histogram — represents a distribution, by describing the frequency of a variable.

Mode — is the most common value.

Normal/Gaussian Distrib. — bell shape distrib.

↳ Tail prejudice on either side of the distrib.

Uniform Distribution — all values occur with the same frequency.

Outliers

Data which egregiously falls outside an acceptable context.

Histograms

- Make the most frequent values obvious.
- Not optimal for comparing distributions
 - ↳ does not account for differences in the sample size.
- a complete description of the distributions sample
 - ↳ can reconstruct the values in the sample (no-order)
- details of a distrib. can be found in a hist.
 - ↳ however it is oft better to use descr. stat.,

- central tendency - vals. cluster around a point?
- mode - is there more than one cluster?
- spread - how much spread is there in the values?
- tails - how quickly do probabilities drop off as we move from the modes?
- outliers - are there extreme values far from the modes?

These type of statistics are called summary statistics.

mean - describe central tendency of distrib.

mean calculation:

$$\bar{x} = \frac{1}{n} \sum_i x_i$$

average - differs in one of several summary statistics that can be chosen to describe a central tendency

Variance

In absence of a central tendency, can describe instead using

- mean
- variance - describing the variability & spread of a distrib. The dispersion from the mean.

Squaring the Deviation has these advantages,

1 ensures the deviations

can be used constructively. The squares remove negativity

$$S^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$$

known as the deviation from the

and prevent cancellation from competing terms.

Variance is the mean squared deviation.

② The exponent is appropriate for non-linear relationship of a variable's value; where a deviation's scale is proportionate to its significance.

Standard Deviation is $\sqrt{s^2}$

③ Reconciling the squaring is manageable \rightarrow more so, than other styles of weighting and enforced constructionism.

Variance is not a good summary stat., but it is useful in some cases.

Effect Size

- an evaluation of how much a variable is influenced between two groups.

Often the differences between groups doesn't have much context; there isn't a popular scale which to root the delta, like University's letter system.

It is a summary statistic, calculating the size of effect.

A common choice is to calculate the difference between the means of two segments for a particular variable.

An alternative method is to use Cohen's Effect Size (Cohen's d).

Instead by leveraging the amount of variation in both segments range of scores the diff. can be contextualized.

$d = 0.2 \rightarrow$ small
 $d = 0.5 \rightarrow$ medium
 $d = 0.8 \rightarrow$ large
 $d = 1 \rightarrow$ (2) standard dev. in diff.

The idea is to use σ as the unit of change for change.

To compare the difference in effect relative to the typical dispersion.

+ semblance of a scale can be found.

Formula.

$$d = \frac{M_E - M_C}{\text{Sample SD Pooled}} \times \left(\frac{N-3}{N-2.25} \right) \times \sqrt{\frac{N-2}{N}}$$

Sample SD Pooled =

$$\sqrt{\frac{(s_1^2 + s_2^2)}{2}}$$

VERSION 2

$$\frac{M_1 - M_2}{\sqrt{\frac{s_1^2 + s_2^2}{2}}}$$

Correction factor for samples < 50

$n_1 = \text{len}(\text{sample}_1)$
 $n_2 = \text{len}(\text{sample}_2)$
 $M_1 = \text{mean}(\text{sample}_1)$
 $M_2 = \text{mean}(\text{sample}_2)$

If the scores of one segment are entirely different from the other, this

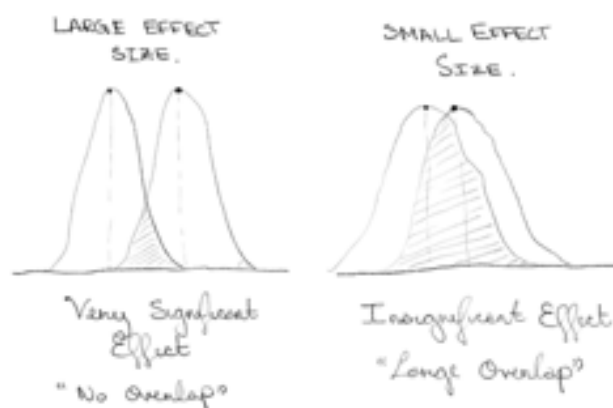
* Can also use Hedges's g , which employs a less potent correction factor.
 * Sample SD. Pooled can be replaced with the Standard Deviation of the population.

would suggest a high degree of difference.

$$Z\text{-score} = \frac{x_i - \bar{x}}{\sigma} \left\{ \begin{array}{l} \text{diff from the} \\ \text{mean in units of} \\ (\sigma) \end{array} \right.$$

Interpretation of Effect Size - Cohen's d

- Measures the difference in means in terms of the standard deviation.
- Effect size is equivalent to the z-score of a standard normal distribution.
- The effect size can be used to express the z-score, or the experiment's mean as a percentile relative to the control group. That percentile can be converted to rank-order describing the placement of the experiment's mean amongst the control group.
- Can also convert to a probability of successfully guessing a datapoint's membership to the control or experiment group.



Effect can be gauged by assessing the variation in common—"the overlap", and using the commonality to contextualize the change in effect.

- ④ r-based Binomial Effect Size Display (BESD)
 - "understandability", and somewhat interpretable.
- ⑤ Common Language Effect Size (CLES)
 - ↳ the probability of a score sampled from one distribution will be greater than the score sampled from another.

Raw Cohen's "d" Value

- 0.2 → small Δ height between 15 & 16 yr. old girls
- 0.5 → med. Δ height 14 & 18 yr. old girls
- 0.8 → large Δ height 13 & 18 yr. old girls.

Relationship Between E.S. & Significance:

- E.S. lends significance to the amount of difference.
- statistical significance - the likelihood that difference is ^{not} coincidental.
- effect size does not account for this.

Margin for Error in estimating Effect Size:

- more data/larger sample → more accuracy / stat. sig.
- can quantify margin of err. w- confidence interval.

$$-5.17 \quad \overline{11.11} \quad .$$