QS101: Introduction to Quantitative Methods in Social Science

Week 4: Reliability, Validity, and Exploration of Single Variables

Florian Reiche
Teaching Fellow in Quantitative Methods
Course Director BA Politics and Sociology
Deputy Director of Student Experience and Progression

23.10.2014

Reliability versus Validity
- Reliability
- Validity

Exploration of Single Variables
- Describing the Centre of the Data
- Describing Variability of the Data

Reliability versus Validity

Reliability

- Stability
- Internal Reliability
- Inter-Observer Consistency

See also the discussion in Adcock, R.N. and David Collier. 2001. Measurement Validity: A Shared Standard for Qualitative and Quantitative Research. American Political Science Review, vol. 95, no. 3, 529-46

Stability

- Is a measure stable over time?
- Remember last week's discussion of measuring economic development
- Tests of stability can be conducted by administering a test of a measure at two points of time t_1 and t_2, with $t_1 \neq t_2$

Internal Reliability

- Refers to the coherence of different attributes used to measure a concept
- In order to test internal reliability, the *split-half* method can be used

Inter-Observer Consistency

- Becomes a problem when a lot of subjective judgement is required in coding
- E.g. content analysis
- Are decisions consistent across observers?
- Example: coding of Polity IV

Validity

- Face Validity
- Concurrent Validity
- Predictive Validity
- Construct Validity
- Convergent Validity

Face Validity

- Does the measure reflect the content of the concept in question
- Established by asking knowledgeable people
- Intuitive Process

Concurrent Validity

- Employ a criterion on which cases are known to differ (e.g. students) and that is relevant to the concept

Concurrent Validity

- Employ a criterion on which cases are known to differ (e.g. students) and that is relevant to the concept
- Assume we measure Satisfaction with QS101 lectures, and employ absenteeism

Concurrent Validity

- Employ a criterion on which cases are known to differ (e.g. students) and that is relevant to the concept
- Assume we measure Satisfaction with QS101 lectures, and employ absenteeism
- Are those dissatisfied with the lecture more likely to be absent than those who are not?

Concurrent Validity

- Employ a criterion on which cases are known to differ (e.g. students) and that is relevant to the concept
- Assume we measure Satisfaction with QS101 lectures, and employ absenteeism
- Are those dissatisfied with the lecture more likely to be absent than those who are not?
- If so, then absenteeism has concurrent validity

Predictive Validity

- Researcher uses a future criterion, rather than a contemporary one
- E.g. future levels of absenteeism

Construct Validity

- Form hypotheses from a *theory* that is relevant to the concept

Convergent Validity

- Compare a measure to measures of the same concept developed through other methods
- E.g. not just ask students how much time they spend on preparing seminars, but let them do a diary, logging preparation time

Exploration of Single Variables

Mean

- The sum of the observations divided by the number of observations.

i	age
1	19
2	20
3	33
4	22
5	21

Mean

- The sum of the observations divided by the number of observations.
- $19 + 20 + 33 + 22 + 21 = 115$

i	age
1	19
2	20
3	33
4	22
5	21

Mean

- The sum of the observations divided by the number of observations.
- $19 + 20 + 33 + 22 + 21 = 115$
- $\frac{19+20+33+22+21}{5} = 23$

i	age
1	19
2	20
3	33
4	22
5	21

Mean

- The sum of the observations divided by the number of observations.
- $19 + 20 + 33 + 22 + 21 = 115$
- $\frac{19+20+33+22+21}{5} = 23$
- $\bar{y} = \frac{\Sigma y_i}{n}$

i	age
1	19
2	20
3	33
4	22
5	21

Some Conventions

- Sample size: n

Some Conventions

- Sample size: n
- Observations of a variable y are denoted as $y_1, y_2, y_3, \ldots, y_n$

Some Conventions

- Sample size: n
- Observations of a variable y are denoted as $y_1, y_2, y_3, \ldots, y_n$
- The sample mean is denoted as $\bar{y}$

Weighted Average

- Denote the sample means for two sets of data with sample sizes n_1 and n_2
- The overall sample mean for the combined set of $(n_1 + n_2)$ can then be written as:

$$\bar{y} = \frac{n_1\bar{y}_1 + n_2\bar{y}_2}{n_1 + n_2}$$

Median

- The median splits the sample into two parts with equal numbers of observations, when they are ordered from lowest to highest

Median

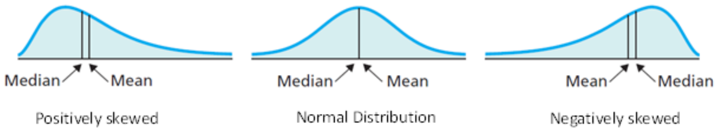
- The median splits the sample into two parts with equal numbers of observations, when they are ordered from lowest to highest
- If the sample size n is odd, a single observation occurs in the middle.

Median

- The median splits the sample into two parts with equal numbers of observations, when they are ordered from lowest to highest
- If the sample size n is odd, a single observation occurs in the middle.
- If the sample size n is even, two middle observations occur, and the median is the midpoint (average) between the two.

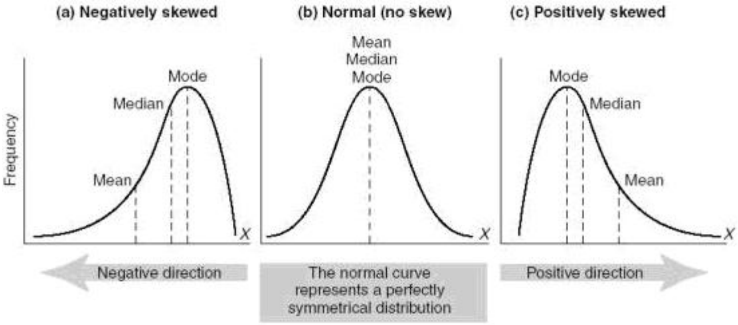
Mean and Median

- For symmetric distributions, mean and median are identical
- For skewed distributions, the mean lies toward the direction of the slew (the longer tail), relative to the median
- The median is unaffected by outliers (!)

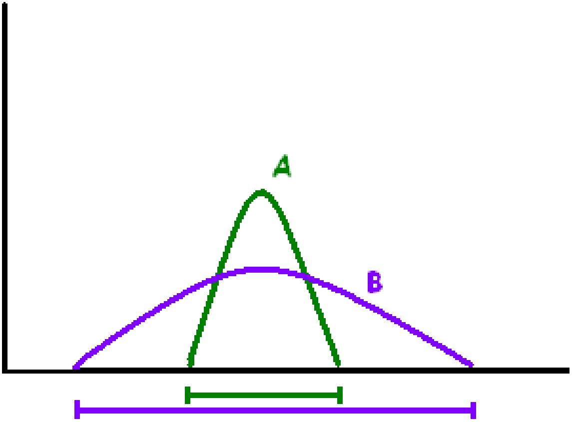


Mode

- The mode is the value that occurs most frequently.



Why the measure of centre doesn't do it



Range

- Is simply the difference between the largest and the smallest observation.

Deviation

- The deviation d of an observation y_i from the sample mean $\bar{y}$ is the difference between them:

$$d = y_i - \bar{y}$$

- The deviation is positive when the observation falls above the mean.
- The deviation is negative when the observation falls below the mean.
- Unfortunately, $\Sigma(y_i - \bar{y}) = 0$

Standard Deviation

- The standard deviation s of n observations is

$$s = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{n-1}} = \sqrt{\frac{\text{sum of squared deviations}}{\text{sample size}-1}}$$

- This is the positive square root of the variance s^2, which is

$$s^2 = \frac{\Sigma(y_i - \bar{y})^2}{n-1} = \frac{(y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \ldots + (y_n - \bar{y})^2}{n-1}$$

- The variance is approximately an average of the squared deviations
- As the units of measurement are squared, this makes the variance difficult to interpret

Properties of the Standard Deviation

- $s \geq 0$

Properties of the Standard Deviation

- $s \geq 0$
- $s = 0$ only when all observations have the same value

Properties of the Standard Deviation

- $s \geq 0$
- $s = 0$ only when all observations have the same value
- The greater the variability around the mean, the larger is the value of s.

Properties of the Standard Deviation

- $s \geq 0$
- $s = 0$ only when all observations have the same value
- The greater the variability around the mean, the larger is the value of s.
- We will return to the interpretation of the magnitude of S next week, when we look at distributions.

Quartiles and Percentiles

- The p^{th} percentile is the point such that $p\%$ of the observations fall below or at that point and $(100 - p)$ % fall above it.

Quartiles and Percentiles

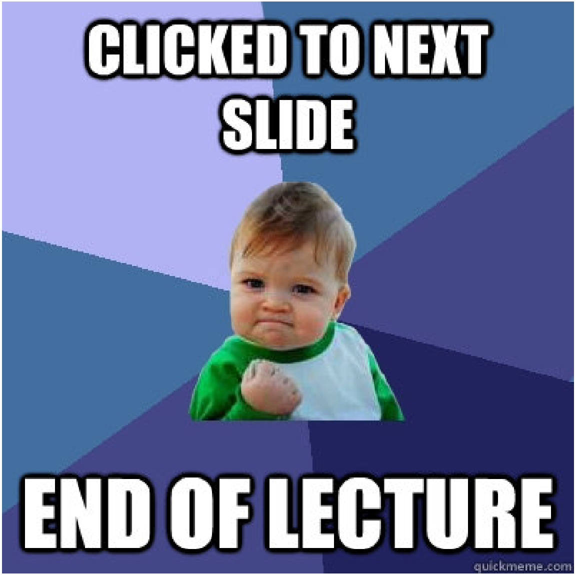
- The p^{th} percentile is the point such that $p\%$ of the observations fall below or at that point and $(100 - p)$ % fall above it.
- The 25^{th} percentile is called the lower quartile. The 75^{th} percentile is called the upper quartile.

Quartiles and Percentiles

- The p^{th} percentile is the point such that p% of the observations fall below or at that point and $(100 - p)$ % fall above it.
- The 25^{th} percentile is called the lower quartile. The 75^{th} percentile is called the upper quartile.
- One quarter of the data fall below the lower quartile, and one quarter of the data fall above the upper quartile.

Quartiles and Percentiles

- The p^{th} percentile is the point such that $p\%$ of the observations fall below or at that point and $(100 - p)$ % fall above it.
- The 25^{th} percentile is called the lower quartile. The 75^{th} percentile is called the upper quartile.
- One quarter of the data fall below the lower quartile, and one quarter of the data fall above the upper quartile.
- The difference between the two quartiles is called the interquartile range, denoted by IQR.


CLICKED TO NEXT SLIDE
END OF LECTURE
quickmeme.com