Intro to Parametric & Nonparametric Statistics

- Kinds of variables & why we care
- Kinds & definitions of nonparametric statistics
- Where parametric stats come from
- Consequences of parametric assumptions
- Organizing the statistics & models we will cover in this class
- Common arguments for using nonparametric stats
- Common arguments against using nonparametric stats
- Using ranks instead of values to compute statistics

Kinds of variable → The "classics" & some others …

Labels

- aka → identifiers
- values may be alphabetic, numeric or symbolic
- different data values represent unique vs. duplicate cases, trials, or events
- e.g., UNL ID#

Nominal

- aka → categorical, qualitative
- values may be alphabetic, numeric or symbolic
- different data values represent different "kinds"
- e.g., species

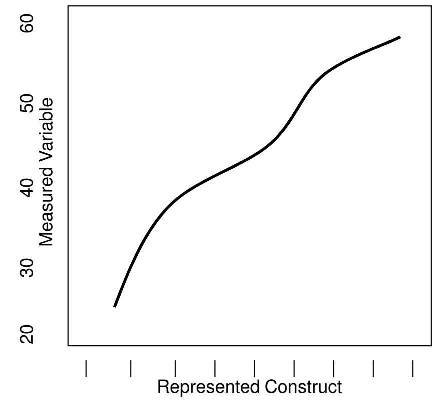
Ordinal

- aka → rank order data, ordered data, seriated data
- values may be alphabetic or numeric
- different data values represent different "amounts"
  - only "trust" the ordinal information in the value
  - don't "trust" the spacing or relative difference information
- has no meaningful "0"
  - don't "trust" ratio or proportional information
- e.g., 10 best cities to live in
  - has ordinal info → 1^{st} is better than 3^{rd}
  - no interval info → 1^{st} & 3^{rd} not "as different" as 5^{th} & 7^{th}
  - no interval info →1st & 5rd not "more different" than 5th & 7th
  - no ratio info → no "0^{th} place"
  - no prop info → 2^{nd} not "twice as good" as 4^{th}
  - no prop dif info→ 1^{st} & 5^{th} not "twice as different" as 1^{st} & 3^{rd}

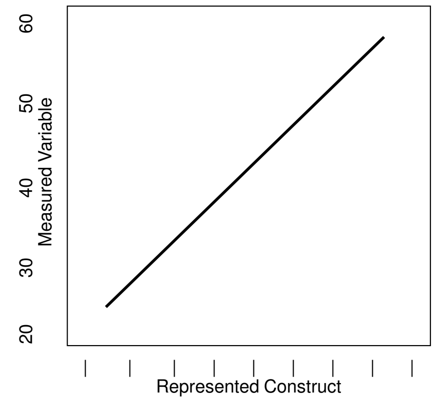
Interval

- aka → numerical, equidistant values
- values are numeric
- different data values represent different “amounts”
  - all intervals of a given extent represent the same difference anywhere along the continuum
  - “trust” the ordinal information in the value
  - “trust” the spacing or relative difference information
- has no meaningful “0” (0 value is arbitrary)
  - don’t “trust” ratio or proportional information
- e.g., # correct on a 10-item spelling test of 20 study words
  - has ordinal info → 8 is better than 6
  - has interval info → 8 & 6 are “as different” as 5 & 3
  - has prop dif info→ 2 & 6 are “twice as different” as 3 & 5
  - no ratio info → 0 not mean “can’t spell any of 20 words”
  - no proportional info → 8 not “twice as good” as 4

Ordinal Measure



Interval Measure



Positive monotonic trace

“more means more but doesn’t tell how much more”

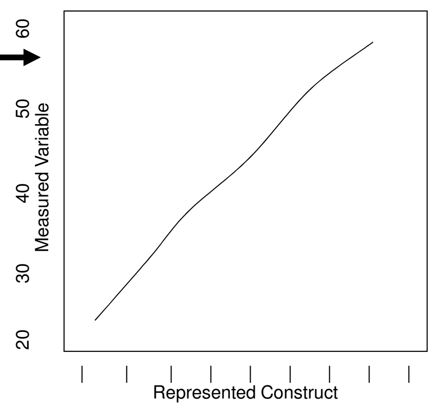
Linear trace

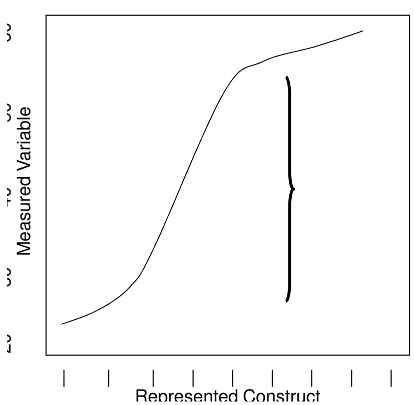
“more tells how much more” but only “difference” not “proportion

y = mx + c

“Nearly” Interval Scale

- “good” summative scales
- how close is “close enough”





“Limited” Interval Scale

- provides interval data only over part of the possible range of the scale values / construct
- summative/aggregated scales

Binary Items

Nominal

- for some constructs different values mean different kinds
- e.g., loves dogs = 1   loves cats = 2

Ordinal

- for some constructs can rank-order the categories
- e.g., fail = 0   pass = 1

Interval

- only one interval, so "all intervals of a given extent represent the same difference anywhere along the continuum"

So, you will see binary variables treated as categorical or numeric, depending on the research question and statistical model.

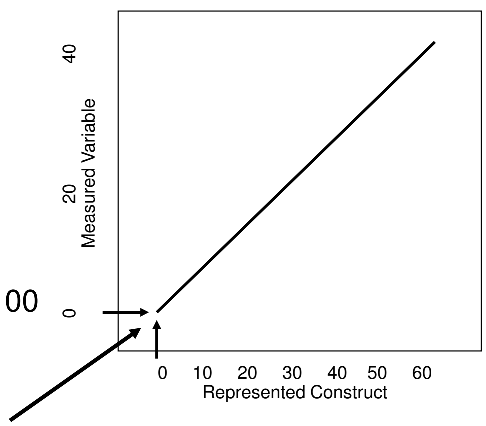
Ratio

- aka → numerical, "real numbers"
- values are numeric
- different data values represent different "amounts"
  - "trust" the ordinal information in the value
  - "trust" the spacing or relative difference information
- has a meaningful "0"
  - "trust" ratio or proportional information
- e.g., number of treatment visits
  - has ordinal info → 9 is better than 7
  - has interval info → 9 & 6 are "as different" as 5 & 2
  - has prop dif info→ 2 & 6 are "twice as different" as 3 & 5
  - has ratio info → 0 does mean "didn't visit"
  - has proportional info → 8 is "twice as many" as 4

Pretty uncommon in Psyc & social sciences

- tend to use arbitrary scales
- usually without a zero
  - 20 5-point items → 20-100

Ratio Measure



Linear scale & "0 means none"

Linear trace w/ 0

"more how much more"

y = mx + c

Kinds of variables → Why we care …

Reasonable mathematical operations

Nominal → $\neq$ =

Ordinal → $\neq$ < = >

Interval → $\neq$ < = > + - (see note below about * /)

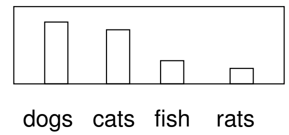
Ratio → $\neq$ < = > + - * /

Note: For interval data we cannot * or / numbers, but can do so with differences. E.g., while 4 can not be said to be twice 2, 8 & 4 are twice as different as are 5 & 3.

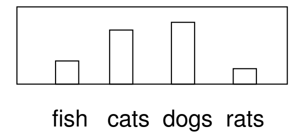
Data Distributions

We often want to know the "shape" of a data distribution.

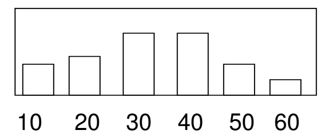
Nominal → can't do → no prescribed value order

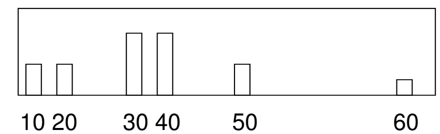


vs.

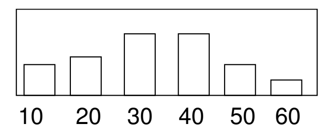


Ordinal → can't do well → prescribed order but not spacing





Interval & Ratio → prescribed order and spacing



There are two kinds of statistics commonly referred to as "nonparametric"...

Statistics for quantitative variables w/out making "assumptions about the form of the underlying data distribution"
- univariate stats -- median & IQR
- univariate stat tests -- 1-sample test of median
- bivariate -- analogs of the correlations, t-tests & ANOVAs

Statistics for qualitative variables
- univariate -- mode & #categories
- univaraite stat tests -- goodness-of-fit X^2
- bivariate -- Pearson's Contingency Table X^2

Have to be careful!! → X^2 tests are actually parametric (they assume an underlying normal distribution – more later)

Defining nonparametric statistics ...

Nonparametric statistics (also called “distribution free statistics”) are those that can describe some attribute of a population, test hypotheses about that attribute, its relationship with some other attribute, or differences on that attribute across populations , across time or across related constructs, that require no assumptions about the form of the population data distribution(s) nor require interval level measurement.

Now, about that last part…

… that require no assumptions about the form of the population data distribution(s) nor require interval level measurement.

This is where things get a little dicey. Today we get just a taste , but we will examine this more carefully as we discuss the relevant models …

Most of the statistics you know have a fairly simple “computational formula”.

As examples...

Here are formulas for two familiar parametric statistics:

The mean ... $$M = \Sigma X \; / \; N$$

The standard deviation ... $$S = \sqrt{\frac{\Sigma (X - M)^2}{N}}$$

But where to these formulas “come from” ???

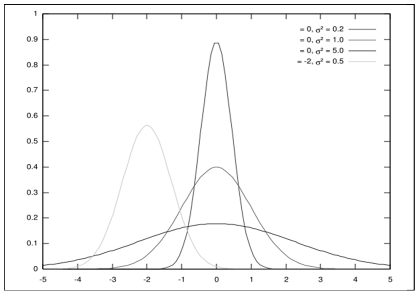
As you’ve heard many times, “computing the mean and standard deviation assumes the data are drawn from a population that is normally distributed.”

What does this really mean ???

formula for the normal distribution:

$$f(x) = \frac{e^{-(x-\mu)^2 / 2\sigma^2}}{\sigma \sqrt{2\pi}}$$

For a given mean (μ) and standard deviation (σ), plug in any value of x to receive the proportional frequency of that value in that particular normal distribution.



The computational formula for the mean and std are derived from this formula.

First …

Since the computational formulas for the mean and the std are derived based upon the assumption that the normal distribution formula describes the data distribution…

if the data are not normally distributed …

then the formulas for the mean and the std don't provide a description of the center & spread of the population distribution.

Same goes for all the formulae that you know !!

Pearson's corr, Z-tests, t-tests, F-tests, X^2 tests, etc…..

Second …

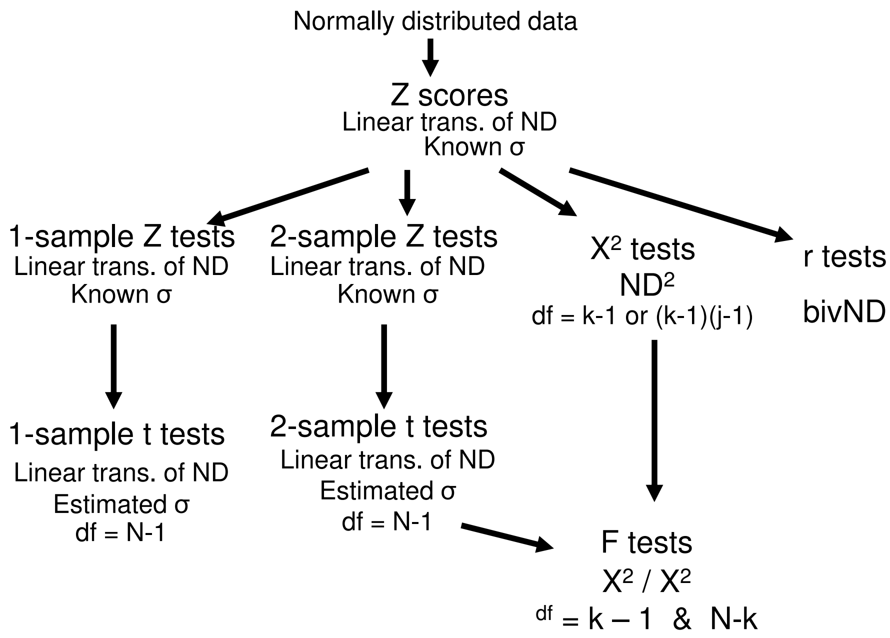
Since the computational formulas for the mean and the std use +, -, * and /, they assume the data are measured on an interval scale (such that equal differences anywhere along the measured continuum represent the same difference in construct values, e.g., scores of 2 & 6 are equally different than scores of 32 & 36)

if the data are not measured on an interval scale …

then the formulas for the mean and the std don't provide a description of the center & spread of the population distribution.

Same goes for all the formulae that you know !!

Pearson's corr, Z-tests, t-tests, F-tests, X^2 tests, etc…..



Organizing nonparametric statistics ...

Nonparametric statistics (also called "distribution free statistics") are those that can describe some attribute of a population,, test hypotheses about that attribute, its relationship with some other attribute, or differences on that attribute across populations, across time, or across related constructs, that require no assumptions about the form of the population data distribution(s) nor require interval level measurement.

describe some attribute of a population	univariate stats
test hypotheses about that attribute	univariate statistical tests
its relationship with some other attribute	tests of association
differences on that attribute across populations	between groups comparisons
across time, or across related constructs	within-groups comparisons

Statistics We Will Consider

DV ➡	Categorical	Parametric Interval/ND	Nonparametric Ordinal/~ND
univariate stats	mode, #cats	mean, std	median, IQR
univariate tests	gof X^2	1-grp t-test	1-grp Mdn test
association	X^2	Pearson's r	Spearman's r Kendall's Tau
2 bg	X^2	t- / F-test	M-W K-W Mdn
k bg	X^2	F-test	K-W Mdn
2wg	McNem Crn's	t- / F-test	Wil's Fried's
kwg	Crn's	F-test	Friedman's

M-W -- Mann-Whitney U-Test Wil's -- Wilcoxin's Test Fried's -- Friedman's F-test
K-W -- Kruskal-Wallis Test
Mdn -- Median Test McNem -- McNemar's X^2 Crn's – Cochran's Test

Things to notice…

Statistics We Will Consider

DV	Categorical	Parametric Interval/ND	Nonparametric Ordinal/~ND
univariate stats	mode, #cats	mean, std	median, IQR
univariate tests	gof X^2	1-grp t-test	1-grp Mdn test
association	X^2	Pearson's r	Spearman's r
2 bg	X^2	t- / F-test	M-W K-W Mdn
k bg	X^2	F-test	K-W Mdn
2wg	McNem Wil's	t- / F-test	Wil's Fried's
kwg	Cochran's	F-test	Fried's

M-W -- Mann-Whitney U-Test
K-W -- Kruskal-Wallis Test
Mdn -- Median Test
Wil's -- Wilcoxin's Test
Fried's -- Friedman's F-test
McNem -- McNemar's X^2

X^2 is used for tests of association between categorical variables & for between groups comparisons with a categorical DV

These WG-comparisons can only be used with binary DVs

k-condition tests can also be used for 2-condition situations

Common reasons/situations ***FOR*** using Nonparametric stats
• & a caveat to consider

Data are not normally distributed

• r, Z, t, F and related statistics are rather "robust" to many violations of these assumptions

Data are not measured on an interval scale.

• Most psychological data are measured "somewhere between" ordinal and interval levels of measurement. The good news is that the "regular stats" are pretty robust to this influence, since the rank order information is the most influential (especially for correlation-type analyses).

Sample size is too small for "regular stats"

• Do we really want to make important decisions based on a sample that is so small that we change the statistical models we use? Remember the importance of sample size to stability.

Common reasons/situations ***AGAINST*** using Nonparametric stats
• & a caveat to consider

Robustness of parametric statistics to most violated assumptions

• Difficult to know if the violations or a particular data set are "enough" to produce bias in the parametric statistics. One approach is to show convergence between parametric and nonparametric analyses of the data.

Poorer power/sensitivity of nonpar statistics (make Type II errors)

• Parametric stats are only more powerful when the assumpt-ions upon which they are based are well-met. If assumptions are violated then nonpar statistics are more powerful.

Mostly limited to uni- and bivariate analyses

• Most research questions are bivariate. If the bivariate results of parametric and nonparametric analyses converge, then there may be increased confidence in the parametric multivariate results.

continued…

Not an integrated family of models, like GLM

- There are only 2 families -- tests based on summed ranks and tests using X^2 (including tests of medians), most of which converge to Z-tests in their "large sample" versions.

H0:s not parallel with those of parametric tests

- This argument applies best to comparisons of "groups" using quantitative DVs. For these types of data, although the null is that the distributions are equivalent (rather than that the centers are similarly positioned → H0: for t-test and ANOVA), if the spread and symmetry of the distributions are similar (as is often the case & the assumption of t-test and ANOVA), then the centers (medians instead of means) are what is being compared by the significance tests.
- In other words, the H0:s are similar when the two sets of analyses make the same assumptions.

Working with "Ranks" instead of "Values"

All of the nonparametric statistics for use with quantitative variables work with the ranks of the variable values, rather than the values themselves.

S#	score	rank
1	12	3.5
2	20	6
3	12	3.5
4	10	2
5	17	5
6	8	1

Converting values to ranks…

- smallest value gets the smallest rank
- highest rank = number of cases
- tied values get the mean of the involved ranks
- cases 1 & 3 are tied for 3rd & 4th ranks, so both get a rank of 3.5

Why convert values to ranks?

Because distributions of ranks are "better behaved" than are distributions of values (unless there are many ties).