جامعة الملك سعود
King Saud University

College of Medicine, KSU, 2016

Research statistics
Data Analysis: Introduction to the Concept

REEM SALLAM, MBBCH, MSC, PHD

CLINICAL CHEMISTRY UNIT, PATHOLOGY DEPT, KING SAUD UNIVERSITY

Second Research Course
Undergraduate Medical Students
March 2016

The Course Contents

- Basic concepts of statistical analysis
- The basics analysis in SPSS

Data Analysis

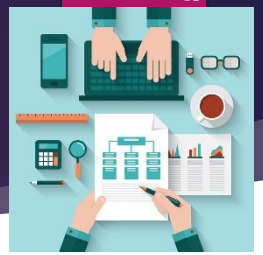
- Data that were gathered in your research are analyzed: this is a first step in ascertaining their meaning
- Before **Getting a feel of your data** hem first looking for:
  - Error
  - Outliers and extremes
  - Data distribution

Data Analysis

- There are 4 basic types of data
- There [illegible] bles that you will be measuring.

Pick your Statistics

There are 4 basic types of data



Categorical (Nominal):

- Data that is in a category or name only, and which values can NOT be placed in any order.

Ordinal:

- The different values of the variable are ordered, but the difference between each value is NOT necessarily the same.

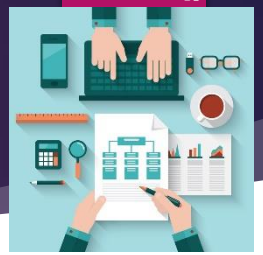
Interval:

- Ordered data in which there is an equal distance between successive levels.

Continuous:

- Numbered data that may be able to go to infinity.

There are 4 basic types of data



Continuous

- Age, weight, gestational age

Categorical (Nominal)

- Gender, race, religion, and yes/no types of answers

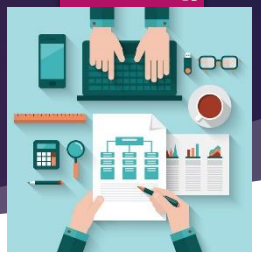
Interval

- Temperature: 37°, 38°, 39°

Ordinal

- One's sense of well-being: good (1), fair (2), or poor (3).

There are 4 basic types of data



Categorical (Nominal):

- Data that is in a category or name only, and which values can NOT be placed in any order.
- E.g. gender, race, religion, and yes/no types of answers

Ordinal:

- The different values of the variable are ordered, but the difference between each value is NOT necessarily the same.
- E.g. One's sense of well-being: good (1), fair (2), or poor (3).

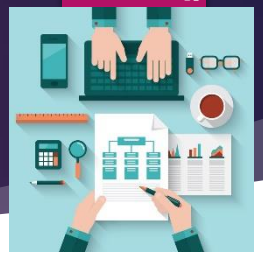
Interval:

- Ordered data in which there is an equal distance between successive levels.
- E.g. temp: the distance between 37° to 38° Is the same as 38°-39°

Continuous:

- Numbered data that may be able to go to infinity.
- E.g. Age, weight, gestational age

There are 4 different types of variables



Demographic:

- Data that describes the characteristics of subjects.

Independent:

- The variables that you control as an investigator.

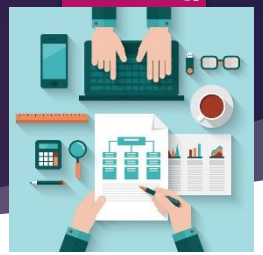
Dependent:

- The outcome variables that you will measure that may be potentially related to or caused by the independent variables.

Confounding:

- Variables that may be associated with the independent variable, and may affect the dependent variable.

There are 4 different types of variables



Independent

- Type of treatment given
- Duration of therapy, or other exposures you may *assign* to subjects.

Demographic

- Gender
- Race
- Previous treatment

Confounding

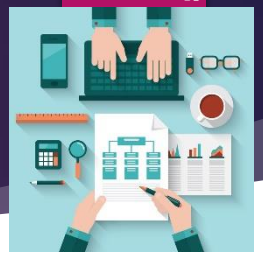
Coffee drinking (independent), death from MI (dependent);

Smoking is a ??? variable

Dependent

- Speed of recovery
- Patient satisfaction

There are 4 different types of variables



Demographic:

- Data that describes the characteristics of subjects.
- E.g. gender, race, previous treatment, etc.

Independent:

- The variables that you control as an investigator.
- E.g. Type of treatment given, duration of therapy, or other exposures you may *assign* to subjects.

Dependent:

- The outcome variables that you will measure that may be potentially related to or caused by the independent variables.
- E.g. speed of recovery, patient satisfaction, etc.

Confounding:

- Variables that may be associated with the independent variable, and may affect the dependent variable.
- E.g. coffee drinking (independent), smoking (potential confounder) and death from MI (dependent).

"Statistics means never having to say you're certain."

Pick your Statistics

- Look at all of the information that you will collect or measure in your study and determine what **type of data** and what **type of variable** each one is
- Based on the type of data, the type of variable, and **what you want to compare**, you can determine what types of statistical tests you may want to use.

There are different types of Statistics

Descriptive:

It describes aspects of your sample.

Confidence intervals:

They indicate that the mean score in the larger population is probably within this interval.

Tests of Significance:

It tells you whether or not a relationship that is found in your sample is likely to exist in the larger population.

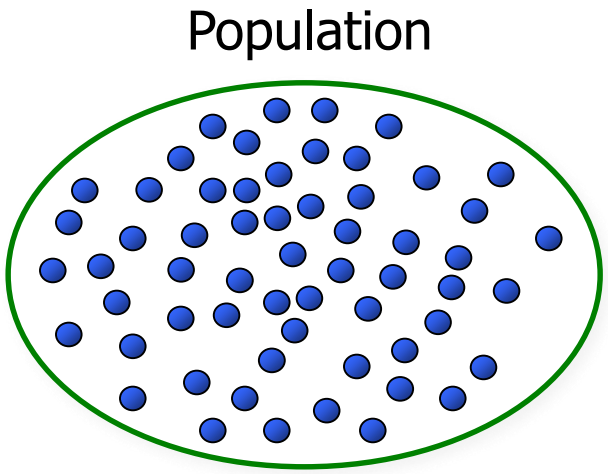
Measures of Association:

It indicates how strong a relationship

Cross-Tab Table:

It shows the inter-relationship between 2 or more variables.

Population vs. sample



A typically cryptic description

A medical student is conducting a mini research project. The aim is to investigate whether there is a difference in the systolic blood pressure of medical students coming to College by driving as compared to those coming by public transportation.

"The mean systolic blood pressure in group A was 110 mmHg, while in group B it was 104 mmHg …

There was no difference in systolic blood pressure between the groups."

A typically cryptic description

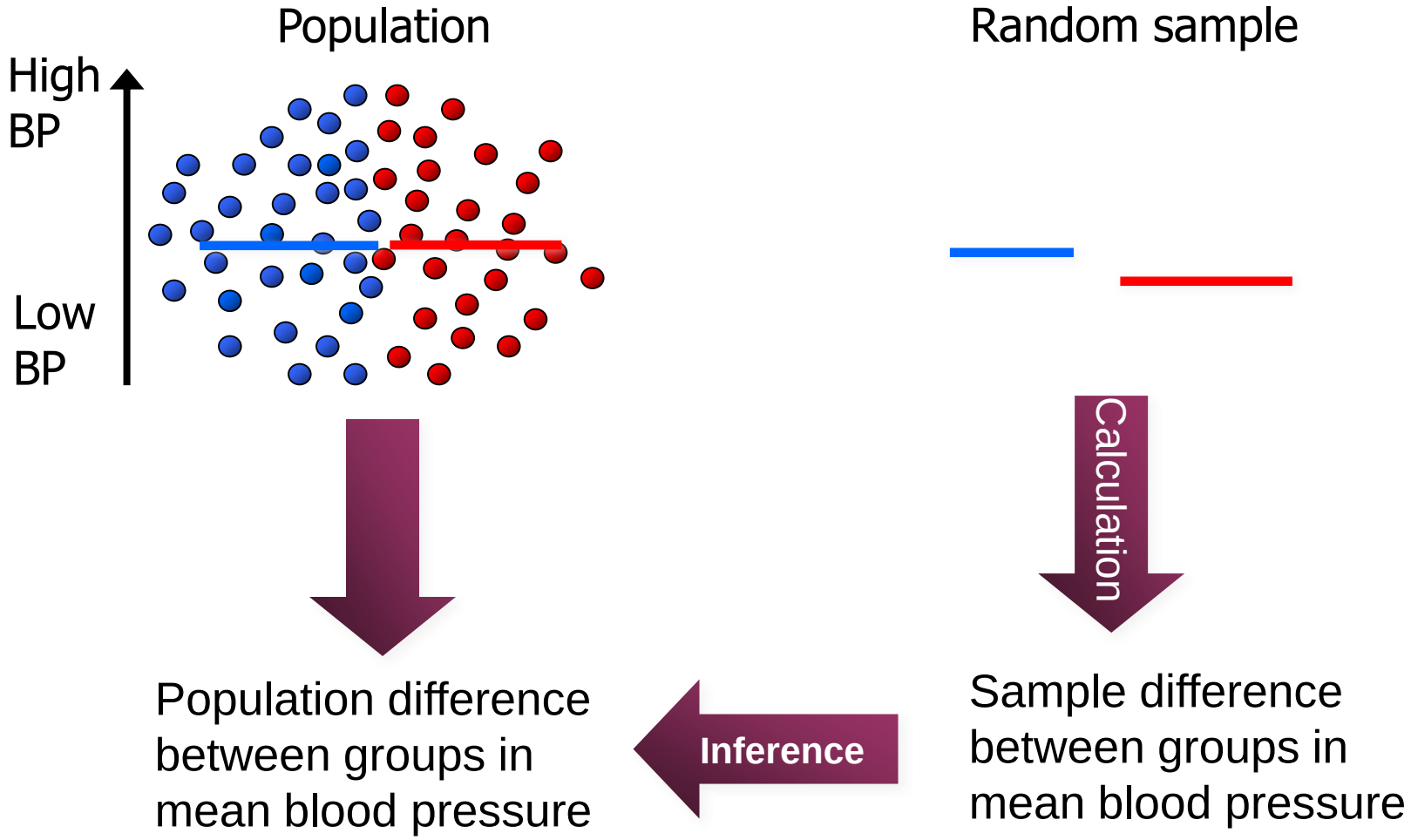
"The mean systolic blood pressure in group A was 110 mmHg, while in group B it was 104 mmHg …

There was no difference in systolic blood pressure between the groups."

Statements like this can be perplexing.

For a start, how can there be no difference when there is clearly a difference?

Population vs. sample



A typically cryptic description

"The mean systolic blood pressure in group A was 110 mmHg, while in group B it was 104 mmHg …

There was no difference in systolic blood pressure between the groups."

A typically cryptic description

Sample

"The mean systolic blood pressure in group A was 110 mmHg, while in group B it was 104 mmHg …

There was no difference in systolic blood pressure between the groups."

Population

A typically cryptic description

"The mean systolic blood pressure in group A was 110 mmHg, while in group B it was 104 mmHg …

There was no difference in systolic blood pressure between the groups."

statistically significant

More on this later …

A typically cryptic description

"The mean systolic blood pressure in group A was 110 mmHg, while in group B it was 104 mmHg …

There was no statistically significant difference in systolic blood pressure between the groups."

mean

Even if the means do not differ significantly *between groups*, systolic blood pressure varies *within each group*.

A typically cryptic description

"The mean systolic blood pressure in group A was 110 mmHg, while in group B it was 104 mmHg …

There was no statistically significant difference in mean systolic blood pressure between the groups."

(SD = 4.2 mmHg)

(SD = 5.0 mmHg)

SD is the standard deviation.

Estimates of variability are essential.

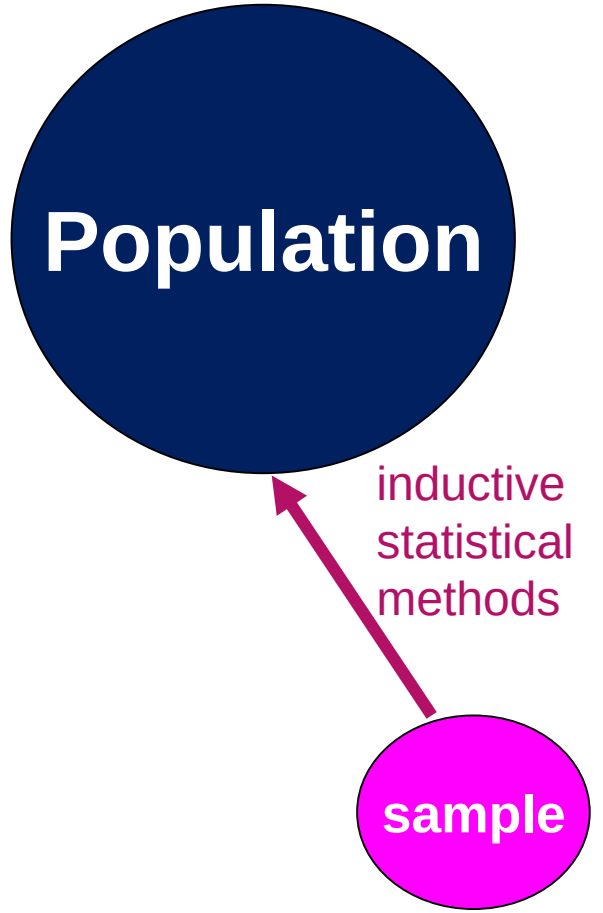
The concept of variability

Variability is always present. Failing to report estimates of variability can be misleading and can also make it impossible for the reader to verify results.

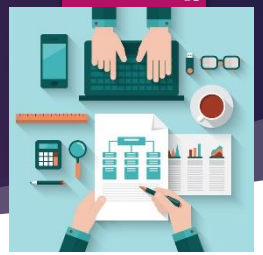
There are different types of Statistics

Descriptive & Inferential:

- Descriptive statistics describe aspects of your sample.
- E.g. averages (mean, median, mode), and dispersion (range, SD).
- If you use your sample to guess the mean or SD of a larger population, they become part of inferential statistics



There are different types of Statistics



Confidence intervals (CIs):

- CIs are put around the sample's mean scores.
- They indicate that the mean score in the larger population is probably within this interval.
- You know that other samples will produce different mean scores. You have only your own sample to work from. You calculate the CI to help represent this fact.

Confidence intervals: Example1

A national poll reports that **42%** of a sample of KSU students support increasing lectures' duration to 120 minutes. It says these results are **accurate "± 3%".**

This means that the % of students in the whole University who support increasing the lectures' duration is **probably between 39% and 45%.**

Hypothesis testing by measuring mean, SD, & Confidence intervals: Example

Pediatrics. 2004 Dec;114(6):e667-71. Epub 2004 Nov 15.

Bedside limited echocardiography by the emergency physician is accurate during evaluation of the critically ill patient.

Pershad J[1], Myers S, Plouman C, Rosson C, Elam K, Wan J, Chin T.

Author information

Abstract

OBJECTIVE: Echocardiography can be a rapid, noninvasive, objective tool in the assessment of ventricular function and preload during resuscitation of a critically ill or injured child. We sought to determine the accuracy of bedside limited echocardiography by the emergency physician (BLEEP) in estimation of (1) left ventricular function (LVF) and (2) inferior vena cava (IVC) volume, as an indirect measure of preload.

METHODS: We conducted a prospective observational study of a convenience sample of patients who were admitted to our intensive care unit. All patients underwent BLEEP followed by an independent formal echocardiogram by an experienced pediatric echocardiography provider (PEP). IVC volume was assessed by measurement of the maximal diameter of the IVC. LVF was determined by calculating shortening fraction (SF) using M-mode measurements on the parasternal short-axis view at the level of the papillary muscle. An independent blinded pediatric cardiologist reviewed all images for accuracy and quality. Estimates of SF obtained on the BLEEP examination were compared with those obtained by the PEP.

RESULTS: Thirty-one patients were enrolled. The mean age was 5.1 years (range: 23 days-16 years); 48.4% (15 of 31) were girls; 58.1% (18 of 31) were on mechanical ventilatory support at the time of their study. There was good agreement between the emergency physician (EP) and the PEP for estimation of SF (r = 0.78). The mean difference in the estimate of SF between the providers was 4.4% (95% confidence interval: 1.6%-7.2%). This difference in estimate of SF was statistically significant. Similarly, there was good agreement between the EP and the PEP for estimation of IVC volume (r = 0.8). The mean difference in the estimate of IVC diameter by the PEP and the EP was 0.068 mm (95% confidence interval: -0.16 to 0.025 mm). This difference was not statistically significant.

CONCLUSIONS: Our study suggests that PEP sonographers are capable of obtaining images that permit accurate assessment of LVF and IVC volume. BLEEP can be performed with focused training and oversight by a pediatric cardiologist.

Hypothesis testing

Goal: to compare ECG measurements made by emergency physicians and experienced pediatric echocardiography providers.

Shortening fraction
SF (%)

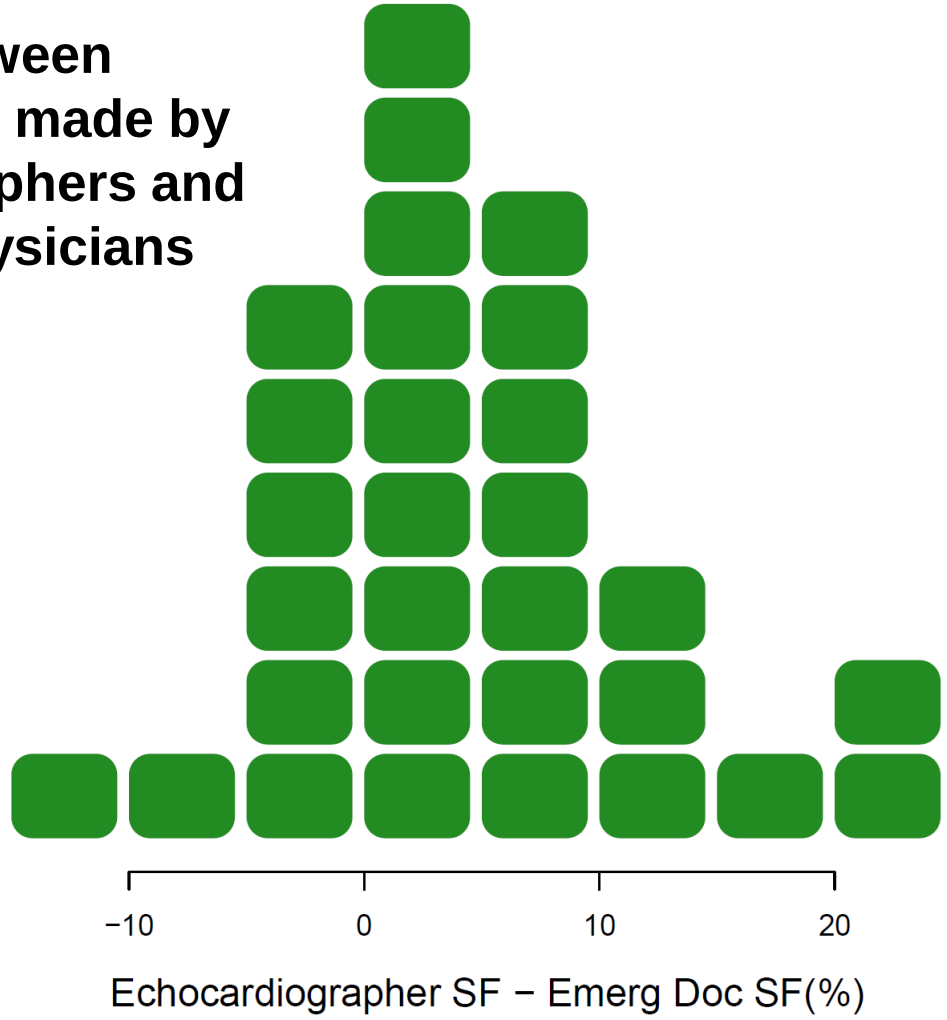
SF is the is a slightly different way of measuring left ventricle performance. It measures and ratios the change in the diameter of the left ventricle between the contracted and relaxed states

$$\text{LV } \frac{\text{end-diastolic diameter - LV end-systolic diameter}}{\text{LV end-diastolic diameter-}} \times 100$$

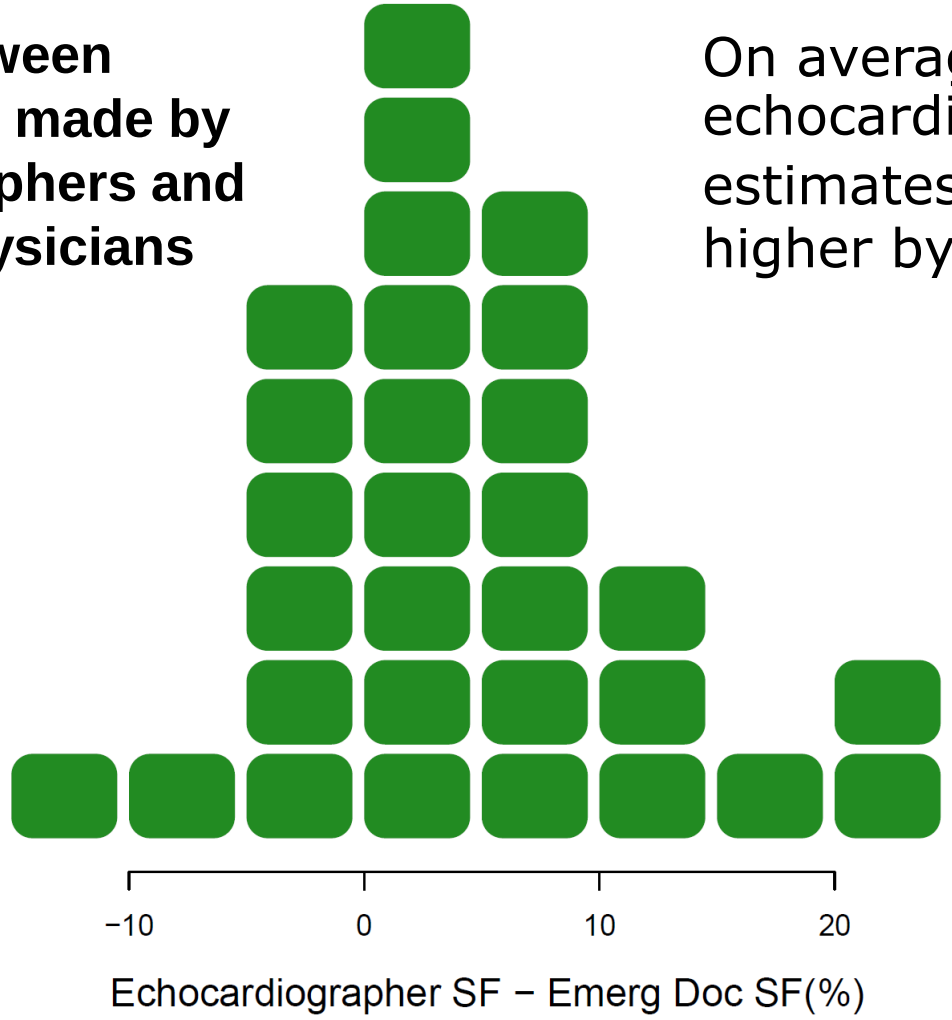
The normal range is 0.18-0.42, or 18-42%
(reference: *Measurements in Cardiology*).

Patient	Emerg Doc	Cardiographer
1	40.60	38.00
2	45.09	51.00
3	31.64	53.40
4	52.68	53.60
5	40.78	46.00
6	34.27	30.50
7	23.33	38.80
8	37.97	45.30
9	33.54	46.40
10	31.19	33.00
11	28.25	40.00
12	23.70	23.70
13	31.44	NA
14	30.57	42.70
15	31.71	36.30
16	29.28	50.00
17	38.83	41.00
18	37.53	42.40
19	33.55	36.00
20	51.38	51.10
21	21.91	30.30
22	23.50	28.70
23	55.00	46.00
24	49.12	44.40
25	48.81	49.40
26	40.91	43.40
27	15.56	24.20
28	40.30	46.20
29	28.40	31.80
30	47.83	47.10
31	74.34	63.20

Difference between measurements made by echocardiographers and emergency physicians

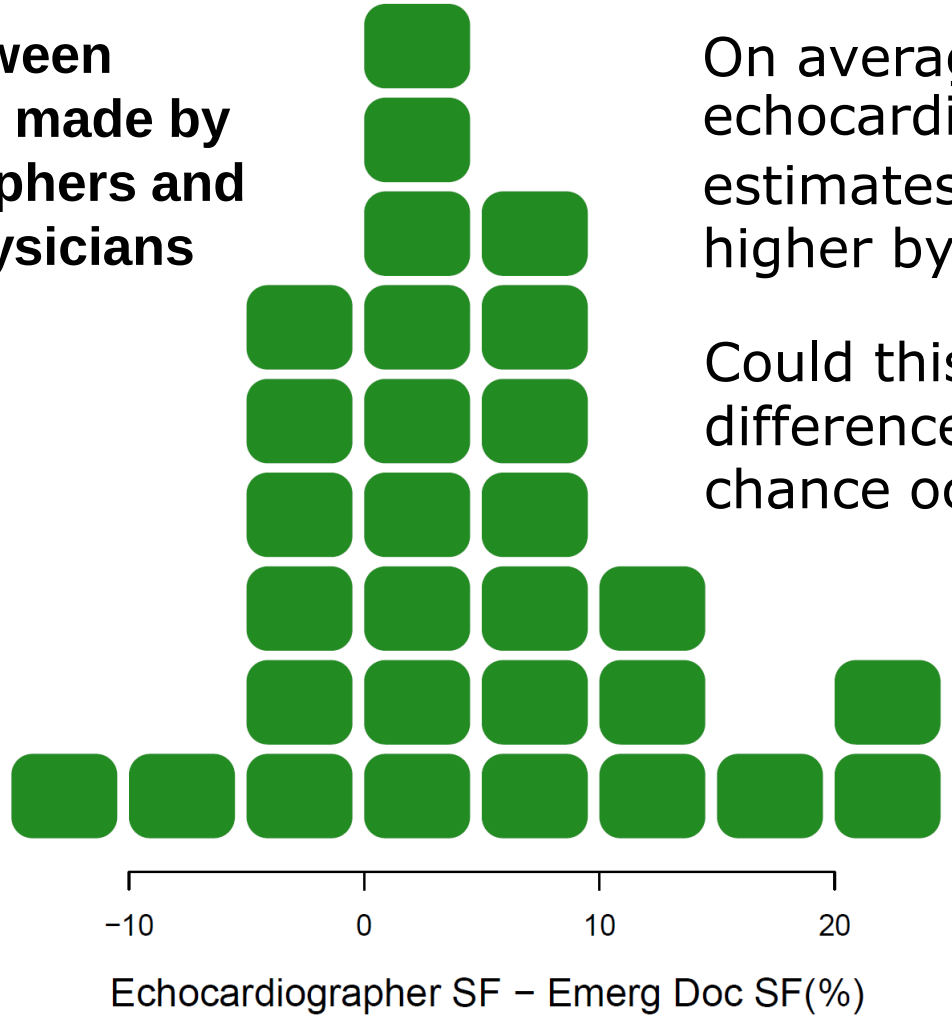


Difference between measurements made by echocardiographers and emergency physicians



On average echocardiographers' estimates were higher by 4.4%.

Difference between measurements made by echocardiographers and emergency physicians

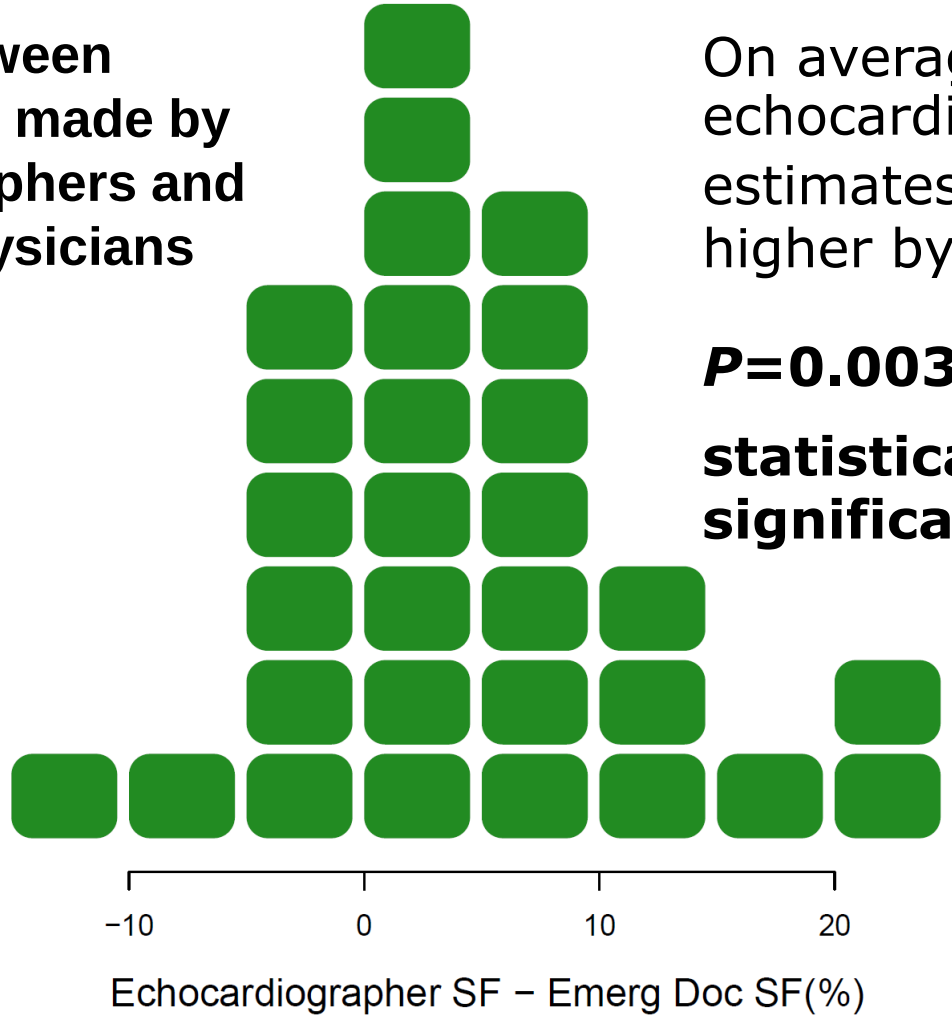


On average echocardiographers' estimates were higher by 4.4%.

Could this difference be a chance occurrence?

They tested the hypothesis that *in the* ***population*** there is no difference.

Difference between measurements made by echocardiographers and emergency physicians



On average echocardiographers' estimates were higher by 4.4%.

***P*=0.003 (t-test)**

statistically significant

Statistical vs. Clinical significance

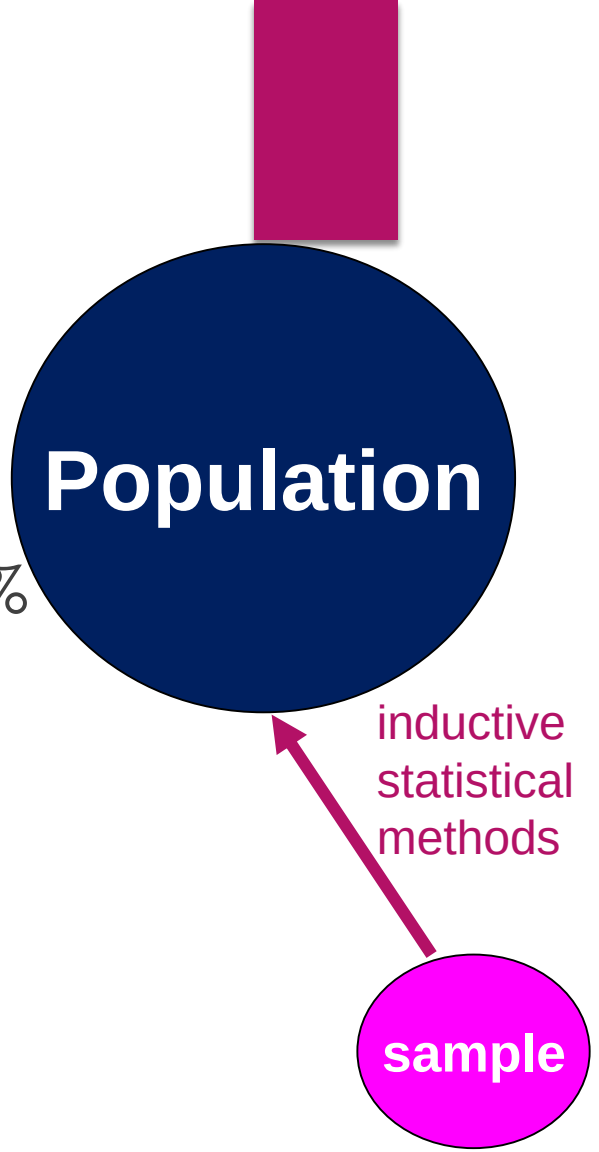
- But the authors note, *"Although **statistically significant**, the difference of 4.4% in the estimation of SF **may not be clinically relevant**."*
- A statistically significant finding is not always clinically significant.
- Subject area judgment is always needed.

Treating a statistically significant finding as important without considering whether it is clinically relevant

*Statistical significance simply **rules out chance** as an explanation for the results. It does not necessarily mean that the results are clinically significant.*

Beyond the p-value

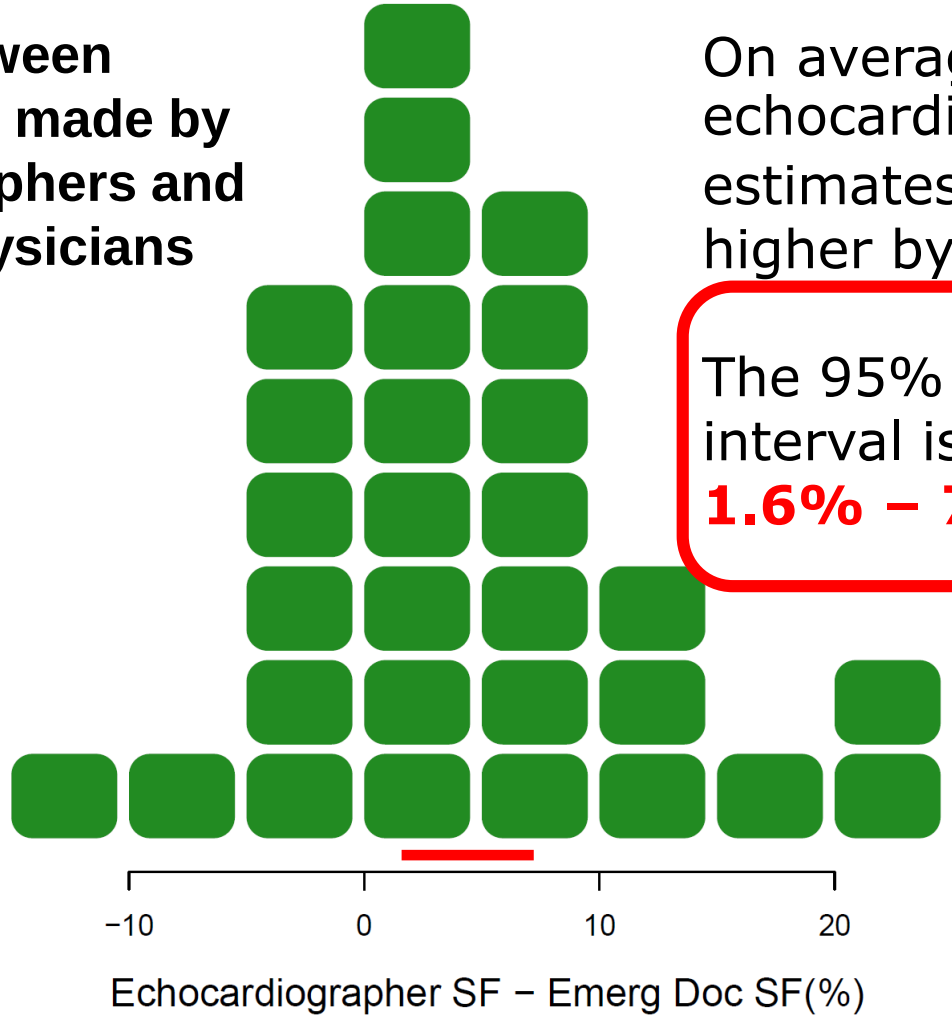
- In the previous example, the mean sample difference is 4.4%
- The mean population difference could be larger or smaller.
- We know (with 95% confidence) that the population difference is greater than zero.
- Can we know anything more? **Yes!** By calculating the CI



Confidence intervals

- A much more useful result than a p-value is a **confidence interval**.
- A confidence interval tells us **what population values (of the difference in means) are consistent with our data.**
- Values outside the confidence interval are ruled out.
- A key issue is the **clinical relevance of the values contained in a confidence interval.**

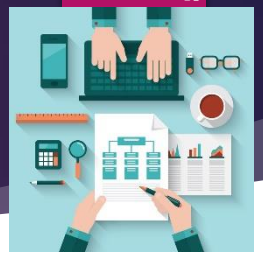
Difference between measurements made by echocardiographers and emergency physicians



On average echocardiographers' estimates were higher by **4.4%**.

The 95% confidence interval is **1.6% – 7.2%**.

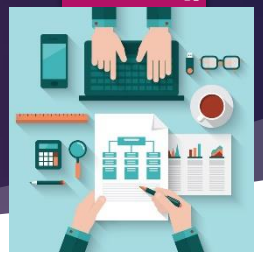
There are different types of Statistics



Tests of Significance:

- It tells you whether or not a relationship that is found in your sample is likely to exist in the larger population.
- E.g. the t-test, F-test, and Chi square.
- It answers: "Does a relationship really exist?"
- A p-value is assigned to this statistics. If the test statisctics is significant: the relationship probably exists in the larger population as well.
- A relationship is significant only because it is likely to really exist; this is not an indicator of importance.

Tests of Significance: Example



A researcher looked at the relation[illegible]en gender and the number of hours of free time people hav[illegible] reported that **men have an average of 1.5 hr more free** [illegible]have. Using t-test, this statistic had a **p-value of 0.31**.

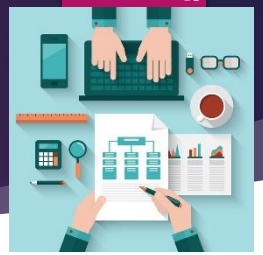
Larger than 0.05

We cannot have faith that this difference really exists in the larger population.

? Lot of variation in people's answers which makes it likely that another sample might not show this same relationship.

This researcher has no statistically significant evidence that **men and women have different amounts of free time in a week.**

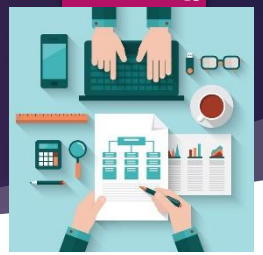
There are different types of Statistics



Measures of Association:

- It indicates how strong a relationship is. E.g. a correlation coefficient tells you how closely related 2 variables are to one another.
- It answers: " How meaningful is this relationship?"
- They typically range from 0.0 (no relationship) to 1.0 (extremely strong relationship).
- Values 0.3-0.5: indicates a relationship exists; while values 0.7-0.8: indicates a stronger relationship.

Measure of Association: Example



A researcher compares scores of men and women [illegible] edge about leisure options in the community. Women sc[illegible]en and the statistic (a t-test) reports a **p-value of 0.01.**

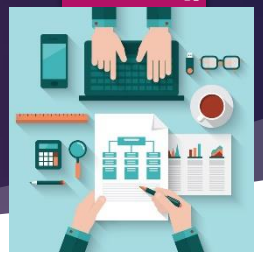
Less than 0.05

She concludes that this finding probably represents a real difference b[illegible] men and women in the larger population. .

Less than 0.3

However, another measure of association (gamma) shows a value of **0.18**. what does this tell her?

Measure of Association: Example



Gamma can answer the questions:
1. Is there an association?
2. How strong is the association?
3. What direction (because level is ordinal) is it?

She concludes that this relationship, which probably really does exist, of a Gamma =0.18, **is so small, it really isn't very important.**

Value	Strength
0.0 - 0.30	Weak
0.30 - 0.60	Moderate
> 0.60	Strong

The slippery slope of causation

EXAMPLE:

"The new surgical technique produced quicker recovery."

Compare with:

"Patients who received the new surgical technique recovered quicker."

Perhaps the patients who received the new surgical technique had less serious conditions.

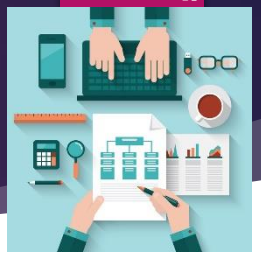
When there is an association, be very cautious **NOT TO** ascribe causation

A well-executed RCT provides the most solid grounds for causal inferences

"Correlation does not imply causation."

An observational study can detect an association, but not (by itself) causation.

There are different types of Statistics



Cross-Tab Table:

- It shows the inter-relationship between 2 or more variables. The variables must be categorical or ordinal. Interval data can be collapsed into ordinal and then analyzed in a cross-tab table.

- It shows the frequencies as the simple number in each cell. You can best judge if there is (or there is no) relationship by comparing percentages, not actual cell frequencies.

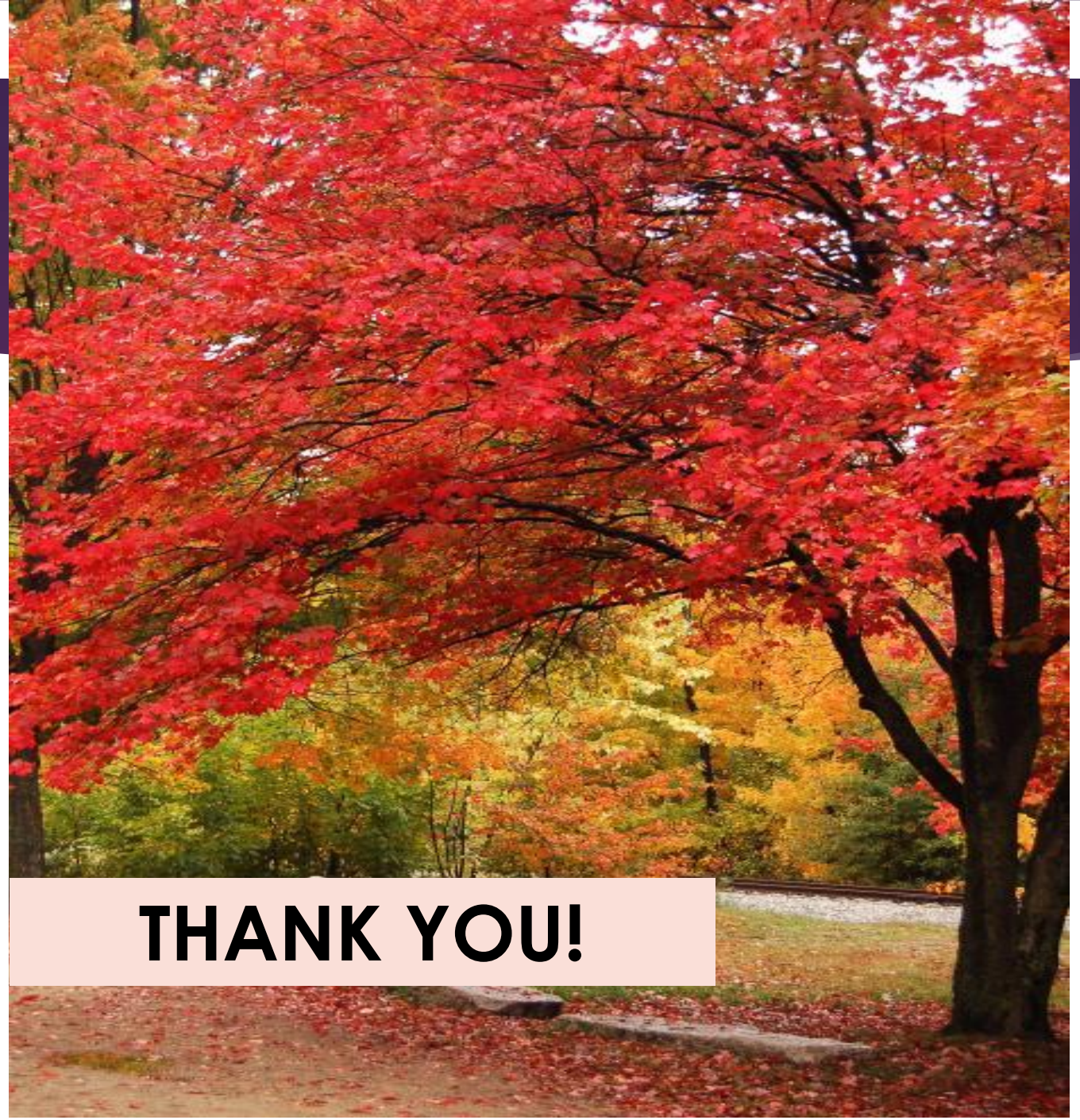
Cross-Tab Table : Example1

You want to study the relationship between gender and exercising during the workday. You recruited 50 male, and 50 female and distributed a simple questionnaire regarding. 19 male and 12 female answered "Yes" to the question: Are you exercising during the workday in at least 3 days of the week for at least 30 minute/day?

	Males	Females	Total
Excercise	19 (38%)	12 (24%)	31 (31%)
No Exercise	31 (62%)	38 (76%)	69 (69%)
Total	50 (100%)	50 (100%)	100 (100%)

The Frequencies are the simple numbers in ach cell

The percentages are better used for comparison


THANK YOU!