FROM SAMPLE TO POPULATION: MAKING INFERENCE

Gbenga Ogunfowokan
Lead, Nigerian Regional Faculty
The Global Health Network
19th May 2017
Objectives

- Know what is Inference
- Know what is parameter estimation
- Understand hypothesis testing & the “types of errors” in decision making.
- Know what the $\alpha$-level means.
- Learn how to use test statistics to examine hypothesis about population mean, proportion
What is inference
<table>
<thead>
<tr>
<th>Statistic</th>
<th>Parameter</th>
<th>Mean:</th>
<th>Standard deviation:</th>
<th>Proportion:</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{X} )</td>
<td>( \mu )</td>
<td></td>
<td>( s )</td>
<td>( \sigma )</td>
</tr>
<tr>
<td>( p )</td>
<td>( \pi )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*estimates from sample* | *estimates from entire population*
Inference

• **Two ways to make inference**
  – Estimation of parameters
    * Point Estimation ($\bar{X}$ or $p$)
    * Intervals Estimation
  – Hypothesis Testing
Basic statistics workshop organised by The Global Health Network Nigerian Regional Faculty
Nigerian women 
N = pop size 
μ

\[ \mu = \frac{x_1 + x_2 + \cdots + x_N}{N} \]

\[ \sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N}} \]

LA: \( x_{AL,1}, x_{AL,2}, \ldots, x_{AL,1000} \) \( \rightarrow \bar{x}_{AL} \)

KA: \( x_{NC,1}, x_{NC,2}, \ldots, x_{NC,1000} \) \( \rightarrow \bar{x}_{NC} \)

PLATEAU: \( x_{WY,1}, x_{WY,2}, \ldots, x_{WY,1000} \) \( \rightarrow \bar{x}_{WY} \)

Sampling distribution
\[ \text{mean}(\bar{x}) \approx \mu \]
\[ SD(\bar{x}) < \sigma \]

Basic statistics workshop organised by The Global Health Network Nigerian Regional Faculty
Central Limit Theorem (CLT): The distribution of sample statistics is nearly normal, centered at the population mean, and with a standard deviation equal to the population standard deviation divided by square root of the sample size.

\[ \bar{x} \sim N \left( \text{mean} = \mu, SE = \frac{s}{\sqrt{n}} \right) \]

Conditions for the CLT:

1. **Independence:** Sampled observations must be independent
   - random sample/assignment
   - if sampling without replacement, \( n < 10\% \) of population

2. **Sample size/skew:** Either the population distribution is normal, or if the population distribution is skewed, the sample size is large (rule of thumb: \( n > 30 \)).
A plausible range of values for the population parameter is called a

**confidence interval.**

- If we report a **point estimate**, we probably won’t hit the exact population parameter.
- If we report a range of plausible values we have a good shot at capturing the parameter.
Central Limit Theorem (CLT):
\[ \bar{x} \sim N \left( \text{mean} = \mu, SE = \frac{\sigma}{\sqrt{n}} \right) \]

approximate 95% CI: \( \bar{x} \pm 2SE \)

margin of error (ME)
Confidence interval for a population mean: Computed as the sample mean plus/minus a margin of error (critical value corresponding to the middle XX% of the normal distribution times the standard error of the sampling distribution).

\[ \bar{x} \pm z^* \frac{s}{\sqrt{n}} \]

Conditions for this confidence interval:
1. **Independence:** Sampled observations must be independent.
   ‣ random sample/assignment
   ‣ if sampling without replacement, \( n < 10\% \) of population
2. **Sample size/skew:** \( n \geq 30 \), larger if the population distribution is very skewed.
Confidence Interval

$$\alpha/2$$

$$1 - \alpha$$

$$\bar{X} - 1.96 \text{SE}$$

$$\bar{X} + 1.96 \text{SE}$$

95% Samples

Basic statistics workshop organised by The Global Health Network Nigerian Regional Faculty
finding the critical value 95% confidence

\[
\frac{(1 - 0.95)}{2} = 0.025
\]

\[ Z^* = \frac{0.025}{0.95} \]

\[
qnorm(0.025) = -1.96
\]
Confidence Interval

• CI for Mean: given by
  – $x \pm 1.96*\text{SE}(x)$ for 95% CI where $\text{SE}(x)=\frac{\sigma}{\sqrt{n}}$
  – $X\pm 1.65*\text{SE}(x)$ for 90% CI

• E.g mean=24.2, Standard deviation 5.9, sample size 100.
  – SE $= \frac{5.9}{\sqrt{100}} = 0.6$
  – 95% CI; 24.2±1.96*0.6; 1.96*0.6=1.2
  – 95% CI; 24.2-1.2 to 24.2+1.2 = 23.0 to 25.4

• Interpretation: we are 95% confident that the mean in the population lies within this interval.
In a survey of 140 asthmatics, 35% had allergy to house dust. Construct the 95% CI for the population proportion.

\[
\pi = p \pm Z \sqrt{\frac{P(1-p)}{n}}
\]

\[
SE = \sqrt{\frac{0.35(1-0.35)}{140}} = 0.04
\]

\[
0.35 - 1.96 \times 0.04 \leq \pi \leq 0.35 + 1.96 \times 0.04
\]

\[
0.27 \leq \pi \leq 0.43
\]

\[
27% \leq \pi \leq 43%
\]
Estimation of parameters

Population

Mean, $\mu$, is unknown

Sample

Point estimate

Mean $\bar{X} = 50$

Interval estimate

I am 95% confident that $\mu$ is between 40 & 60

Basic statistics workshop organised by The Global Health Network Nigerian Regional Faculty
Interpretation of CI

Probabilistic

In repeated sampling $100(1-\frac{\alpha}{2})\%$ of all intervals around sample means will in the long run include \[ ? \]

Practical

We are $100(1-\frac{\alpha}{2})\%$ confident that the single computed CI contains \[ ? \]
Standard Error

**Quantitative Variable**

\[ SE \text{ (Mean)} = \frac{s}{\sqrt{n}} \]

**Qualitative Variable**

\[ SE \text{ (p)} = \sqrt{\frac{p(1-p)}{n}} \]
Confidence Interval

• CI for proportions
  – \( \hat{p} \pm 1.96 \times \text{SE}(\hat{p}) \), for 95% CI where \( \text{SE}(\hat{p}) = \sqrt{\hat{p}(1-\hat{p})/n} \)

• For small samples, the t distribution is used to estimate CI:
  – Multiplier will be the value of t corresponding to t two-sided \( p=0.05 \) with \( df=n-1 \)

• CI also calculated for RR & OR: estimates that the true association lies within the interval:
  – OR \( e^{\pm 1.96 \sqrt{(1/a + 1/b + 1/c + 1/d)}} \)
Practical

• PRACTICAL DEMONSTRATION WITH EPI INFO SOFTWARE
• 1) Quantitative Variable
• Point Estimate
• Measure of dispersion
• Inference using 95% CI
Practical

- PRACTICAL DEMONSTRATION USING EPI INFO SOFTWARE
- Qualitative variable
- Point estimate using Proportion, Percentage
- Measure of dispersion
- Inference using 95% CI
HYPOTHESIS TESTING
Hypothesis testing

A statistical method that uses sample data to evaluate a hypothesis about a population parameter. It is intended to help researchers differentiate between real and random patterns in the data.
An assumption about the population parameter.

I assume the mean SBP of participants is 120 mmHg.
agreement of CI and HT

- A two sided hypothesis with threshold of $\alpha$ is equivalent to a confidence interval with $CL = 1 - \alpha$.
- A one sided hypothesis with threshold of $\alpha$ is equivalent to a confidence interval with $CL = 1 - (2 \times \alpha)$.

If $H_0$ is rejected, a confidence interval that agrees with the result of the hypothesis test should not include the null value.

If $H_0$ is failed to be rejected, a confidence interval that agrees with the result of the hypothesis test should include the null value.
hypothesis test as a trial

If we again think of a hypothesis test as a criminal trial then it makes sense to frame the verdict in terms of the null and alternative hypotheses:

\[ H_0 : \text{Defendant is innocent} \]
\[ H_A : \text{Defendant is guilty} \]

Which type of error is being committed in the following circumstances?

- Declaring the defendant innocent when they are actually guilty
  - Type 2 error
- Declaring the defendant guilty when they are actually innocent
  - Type 1 error
Which error is the worst error to make?

- **Type 2**: Declaring the defendant innocent when they are actually guilty
- **Type 1**: Declaring the defendant guilty when they are actually innocent

“better that ten guilty persons escape than that one innocent suffer”
type I error rate

- We reject $H_0$ when the p-value is less than 0.05 ($\alpha = 0.05$).
- This means that, for those cases where $H_0$ is actually true, we do not want to incorrectly reject it more than 5% of those times.
- In other words, when using a 5% significance level there is about 5% chance of making a Type I error if the null hypothesis is true.

$$P(\text{Type I error} \mid H_0 \text{ true}) = \alpha$$

- This is why we prefer small values of $\alpha$ – increasing $\alpha$ increases the Type I error rate.

Basic statistics workshop organised by The Global Health Network Nigerian Regional Faculty
The goal is to keep $\alpha$ and $\beta$ low.

### Table

<table>
<thead>
<tr>
<th>Truth</th>
<th>Decision</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$ true</td>
<td>fail to reject</td>
<td>$1 - \alpha$</td>
<td>Type I error, $\alpha$</td>
</tr>
<tr>
<td>$H_A$ true</td>
<td>reject $H_0$</td>
<td></td>
<td>Type 2 error, $\beta$</td>
</tr>
</tbody>
</table>

- **Type 1 error** is rejecting $H_0$ when you shouldn’t have, and the probability of doing so is $\alpha$ (significance level).
- **Type 2 error** is failing to reject $H_0$ when you should have, and the probability of doing so is $\beta$.
- **Power** of a test is the probability of correctly rejecting $H_0$, and the probability of doing so is $1 - \beta$.
Factors Increasing Type II Error

- True Value of Population Parameter
  - Increases When Difference Between Hypothesized Parameter & True Value Decreases
- Significance Level $\alpha$
  - Increases When $\alpha$ Decreases
- Population Standard Deviation $\sigma$
  - Increases When $\sigma$ Increases
- Sample Size $n$
  - Increases When $n$ Decreases
$p$ Value Test

• Probability of Obtaining a Test Statistic More Extreme ($\leq$ or $\geq$) than Actual Sample Value Given $H_0$ Is True

• Called Observed Level of Significance

• Used to Make Rejection Decision
  – If $p$ value $\geq \alpha$, Do Not Reject $H_0$
  – If $p$ value $< \alpha$, Reject $H_0$
Hypothesis Testing: Steps

Test the Assumption that the true mean SBP of participants is 120 mmHg.

State $H_0$: $H_0: \mu \neq 120$

State $H_1$: $H_1: \mu = 120$

Choose $\alpha$: $\alpha = 0.05$

Choose $n$: $n = 100$

Choose Test: $Z, t, X^2$ Test (or $p$ Value)
Hypothesis Testing: Steps

Compute Test Statistic *(or compute P value)*

Search for Critical Value

Make Statistical Decision rule

Express Decision
One sample-mean Test

- Assumptions
  - Population is normally distributed

- t test statistic

\[
t = \frac{\text{sample mean} - \text{null value}}{\text{standard error}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}
\]
What is **normal body temperature**? Is it actually 37.6°C (on average)?

State the null and alternative hypotheses

\[ H_0: \mu = 37.6^\circ C \]

\[ H_a: \mu \neq 37.6^\circ C \]
Example Normal Body Temp (cont)

Data: random sample of $n = 18$ normal body temps

<table>
<thead>
<tr>
<th>Temperature</th>
<th>n</th>
<th>Mean (°C)</th>
<th>SD</th>
<th>SE</th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18</td>
<td>37.22</td>
<td>0.68</td>
<td>0.161</td>
<td>2.38</td>
<td>0.029</td>
</tr>
</tbody>
</table>

Summarize data with a test statistic

$$t = \frac{\text{sample mean} - \text{null value}}{\text{standard error}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$
<table>
<thead>
<tr>
<th>Degrees of freedom</th>
<th>0.10</th>
<th>0.05</th>
<th>0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.314</td>
<td>12.706</td>
<td>63.657</td>
</tr>
<tr>
<td>5</td>
<td>2.015</td>
<td>2.571</td>
<td>4.032</td>
</tr>
<tr>
<td>10</td>
<td>1.813</td>
<td>2.228</td>
<td>3.169</td>
</tr>
<tr>
<td>17</td>
<td>1.740</td>
<td>2.110</td>
<td>2.898</td>
</tr>
<tr>
<td>20</td>
<td>1.725</td>
<td>2.086</td>
<td>2.845</td>
</tr>
<tr>
<td>24</td>
<td>1.711</td>
<td>2.064</td>
<td>2.797</td>
</tr>
<tr>
<td>25</td>
<td>1.708</td>
<td>2.060</td>
<td>2.787</td>
</tr>
<tr>
<td>∞</td>
<td>1.645</td>
<td>1.960</td>
<td>2.576</td>
</tr>
</tbody>
</table>
Example Normal Body Temp (cont)

Find the $p$-value

Df = n – 1 = 18 – 1 = 17

From SPSS: $p$-value = 0.029

From t Table: $p$-value is between 0.05 and 0.01.

Area to left of $t = -2.11$ equals area to right of $t = +2.11$.

The value $t = 2.38$ is between column headings 2.110 & 2.898 in table, and for df = 17, the $p$-values are 0.05 and 0.01.
Example Normal Body Temp (cont)

Decide whether or not the result is statistically significant based on the $p$-value

Using $\alpha = 0.05$ as the level of significance criterion, the results are statistically significant because 0.029 is less than 0.05. In other words, we can reject the null hypothesis.

Report the Conclusion

We can conclude, based on these data, that the mean temperature in the human population does not equal 37.6.
One-sample test for proportion

- Involves categorical variables
- Fraction or % of population in a category
- Sample proportion (p)
  - Test is called Z test
    - Z is computed value
    - \( \pi \) is proportion in population (null hypothesis value)

\[
p = \frac{X}{n} = \frac{\text{number of successes}}{\text{sample size}}
\]

\[
Z = \frac{p - \pi}{\sqrt{\pi (1 - \pi)} \frac{1}{n}}
\]

Critical Values: 1.96 at \( \alpha=0.05 \)
2.58 at \( \alpha=0.01 \)
Example

• In a survey of diabetics in a large city, it was found that 100 out of 400 have diabetic foot. Can we conclude that 20 percent of diabetics in the sampled population have diabetic foot.

• Test at the $\alpha = 0.05$ significance level.
Solution

H₀: \( \pi = 0.20 \)
H₁: \( \pi \neq 0.20 \)

\[
Z = \sqrt{\frac{0.25 - 0.20}{0.20(1-0.20) + 1.96}} = 2.50
\]

Critical Value: 1.96

Decision:

We have sufficient evidence to reject the Ho value of 20%
We conclude that in the population of diabetic the proportion who have diabetic foot does not equal 0.20
Flow chart of commonly used statistical tests

- **Exposure variable**
  - Normal
    - 1 group: One-sample t test
    - 2 groups: Two-sample t test
    - Paired: Paired t test
    - >2 groups: One-way ANOVA test
    - Continuous: Pearson Corr / Linear Reg
  - Skew
    - Sign test / Signed rank test
    - Mann-Whitney U test
    - Wilcoxon signed rank test
    - Kruskal Wallis test
    - Spearman Corr / Linear Reg

- **Outcome variable**
  - Categorical
    - 1 group: Chi-square test / Exact test
    - 2 groups: Chi-square test / Fisher’s exact test / Logistic regression
    - Paired: McNemar’s test / Kappa statistic
    - >2 groups: Chi-square test / Fisher’s exact test / Logistic regression
    - Continuous: Logistic regression / Sensitivity & specificity / ROC

- Survival
  - 2 groups: KM plot with Log-rank test
  - >2 groups: KM plot with Log-rank test
  - Continuous: Cox regression
## Which Statistical Test?

Use the table to obtain information on how to carry out the test in SPSS and how to report and present the results.

Move the cursor over the boxes that classify the tests for further details. Click on the statistical tests for more details.

<table>
<thead>
<tr>
<th>Number of groups / Exposure (independent) variable</th>
<th>Outcome (dependent) variable</th>
<th>Survival Time to event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous and Normally distributed (Parametric)</td>
<td>Continuous and skewed / Ordinal (Non-parametric)</td>
<td>Binary (2 categories)</td>
</tr>
<tr>
<td>1 group</td>
<td>One-sample t test</td>
<td>Sign test / Signed rank test</td>
</tr>
<tr>
<td>2 independent groups</td>
<td>Two-sample t test</td>
<td>Mann-Whitney U test</td>
</tr>
<tr>
<td></td>
<td>Linear regression</td>
<td></td>
</tr>
<tr>
<td>Paired (related) sample (2 time points)</td>
<td>Paired t test</td>
<td>Wilcoxon signed rank test</td>
</tr>
<tr>
<td></td>
<td>Bland-Altman method</td>
<td></td>
</tr>
<tr>
<td>&gt;2 independent groups</td>
<td>One-way ANOVA test</td>
<td>Kruskal-Wallis test</td>
</tr>
<tr>
<td></td>
<td>Linear regression</td>
<td></td>
</tr>
<tr>
<td>&gt;2 related samples (&gt;2 time points)</td>
<td>Repeated measures ANOVA</td>
<td>Friedman’s Test</td>
</tr>
<tr>
<td>Continuous</td>
<td>Pearson’s correlation</td>
<td>Spearman’s rank correlation</td>
</tr>
<tr>
<td></td>
<td>Linear Regression</td>
<td>Linear regression</td>
</tr>
<tr>
<td>Epidemiological data</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Comparing a binary outcome between two groups – data presented as a 2x2 table

<table>
<thead>
<tr>
<th></th>
<th>Unfit</th>
<th>Fit</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard</strong></td>
<td>80</td>
<td>140</td>
<td>220</td>
</tr>
<tr>
<td></td>
<td>(a)</td>
<td>(b)</td>
<td>(a+b)</td>
</tr>
<tr>
<td><strong>Enhanced</strong></td>
<td>20</td>
<td>220</td>
<td>240</td>
</tr>
<tr>
<td></td>
<td>(c)</td>
<td>(d)</td>
<td>(c+d)</td>
</tr>
</tbody>
</table>

- Table shows results from our trial (number of patients)
- Difference in proportion of Fit between groups (absolute difference): 
  \[
  \frac{d}{c+d} - \frac{b}{a+b}
  \]
- An alternative parameter is the relative risk (multiplicative difference): 
  \[
  \frac{d}{c+d} \times \frac{b}{a+b}
  \]
- Another alternative is the odds ratio: 
  \[
  \frac{d}{c} \div \frac{b}{a} = \frac{a \times d}{b \times c}
  \]

Chi-square test and Fisher’s exact test show if there is any association between the two independent variables, but it doesn’t provide the effect size between the groups regarding the outcome of interest, e.g. Fit.
### Percentage of Fit in standard group: 140/220 (63.6%)
### Percentage of Fit in enhanced group: 220/240 (91.7%)

<table>
<thead>
<tr>
<th>Parameter (95% CI)</th>
<th>Parameter (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute difference in proportions ( \frac{d}{c+d} - \frac{b}{a+b} )</td>
<td>28.1% (21%, 35%)*</td>
</tr>
<tr>
<td>Relative risk ( \frac{d}{c+d} ) ( \frac{c}{a+b} )</td>
<td>1.44 (1.29, 1.60)‡</td>
</tr>
<tr>
<td>Odds ratio ( \frac{a \times d}{b \times c} )</td>
<td>6.29 (3.69, 10.72)‡</td>
</tr>
</tbody>
</table>

* Asymptotic 95% confidence intervals (calculated in CIA)
‡ 95% confidence intervals calculated in SPSS

**Reminder:** Report confidence intervals for **ALL** key parameter estimates
- If 95% confidence interval for a difference excludes 0 \( \Rightarrow \) statistically significant e.g. Absolute difference
- If 95% confidence interval for a ratio excludes 1 \( \Rightarrow \) statistically significant e.g. Relative risk and Odds ratio
## Advantages and disadvantages of absolute and relative changes, and odds ratios

| **Absolute difference** | simplest to calculate and to interpret  
|                        | when applied to number of subjects in a group gives number of subjects expected to benefit  
|                        | 1/(absolute difference) gives NNT – ‘number needed to treat’ to see one additional positive response |
| **Relative risk**      | intuitively appealing  
|                        | a multiplicative effect – proportion (risk) of failure in the treatment group examined relative to (or compare to) that in the reference group  
|                        | different result depending on whether risks of ‘Fit’ or ‘Unfit’ are examined and whether ‘Standard exercise’ group is selected as the reference level  
|                        | natural parameter for cohort studies |
| **Odds ratio**         | difficult to understand – unless you’re a betting person!  
|                        | ratio of ‘number of successes expected per number of failures’ between the treatment group of interest and the reference group  
|                        | invariant to whether rate of ‘Fit’, ‘Unfit’, or rate of taking ‘Enhanced exercise’ are examined  
|                        | logistic regression in terms of odds ratios  
|                        | natural parameter for case-control studies |
• PRACTICAL DEMONSTRATION USING EPI INFO AND EXCEL
Practicals

- RESEARCH TITLE
- RESEARCH OBJECTIVES
- DESCRIPTIVE STATISTICS
- INFERENTIAL STATISTICS
THANK YOU FOR YOUR ATTENTION

Basic statistics workshop organised by The Global Health Network Nigerian Regional Faculty

CORBIS/Brian Leng (05065)