Data Analysis

Dr. Coleman Cutchins, PharmD, BCPS

Before we start

- I have no conflicts of interest

- I am a well published Clinical Pharmacist but by no means a statistician

- This does not substitute a formal education in statistics or epidemiology

Objectives

- Understand basic terms of data analysis
- Be able to explain central measures
- Give examples of data collection pertinent to Antimicrobial Stewardship

Statistics

Introduction to Terms

- Population – The complete group of interest
- Sample – Part of the population (subset)
- Variable – a characteristic or property of an item we expect to vary
- Descriptive statics – describes what is going on, what is known
- Inferential statics – predicts, estimates, infers about data that is not completely known. Needs to include a measure of reliability

Why do we need statistics?

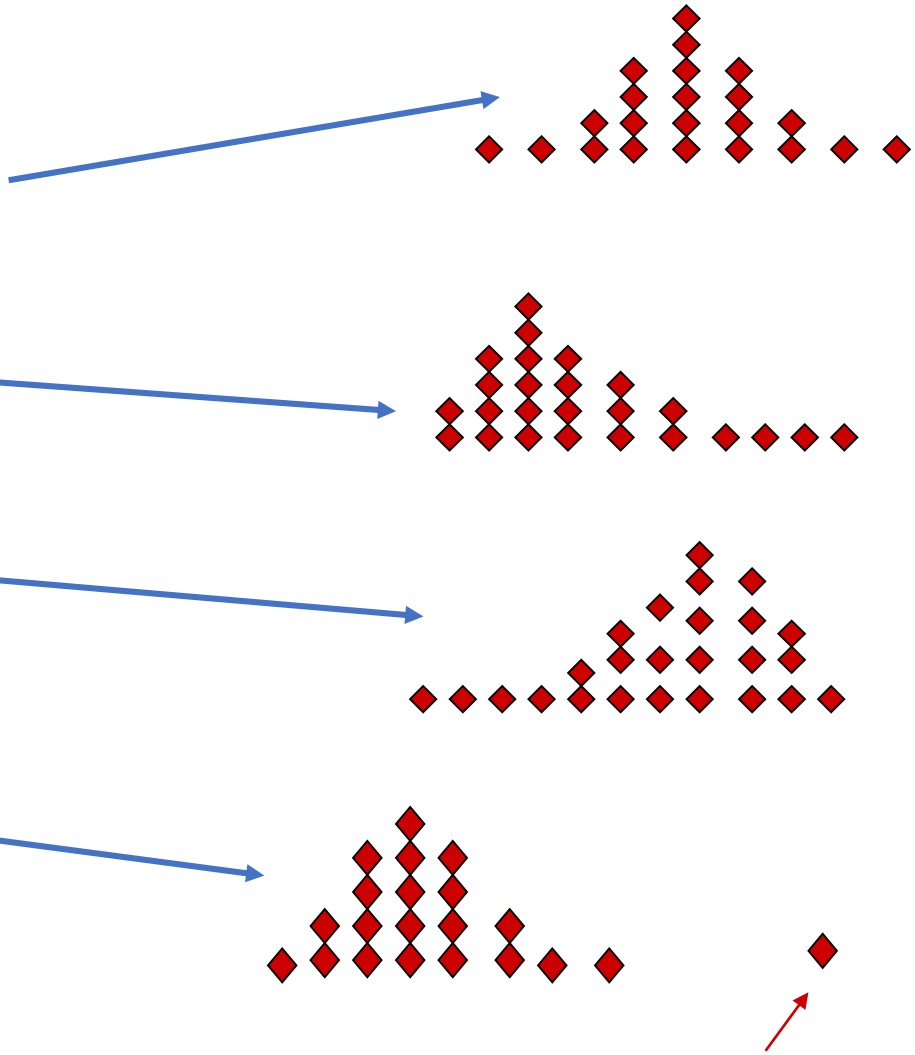
- It’s basically mathematical gambling
- Statistics = science of data
- We study information (sample) that is available and manageable to better understand population principles

Data types

- Qualitative variables - measure a quality characteristic on each experimental unit
- Quantitative variables - measure a numerical quantity on each experimental unit.
  - Discrete - if it can assume only a finite or countable number of values.
  - Continuous - if it can assume the infinitely many values corresponding to the points on a line interval.
- Normalized data – scaled to population

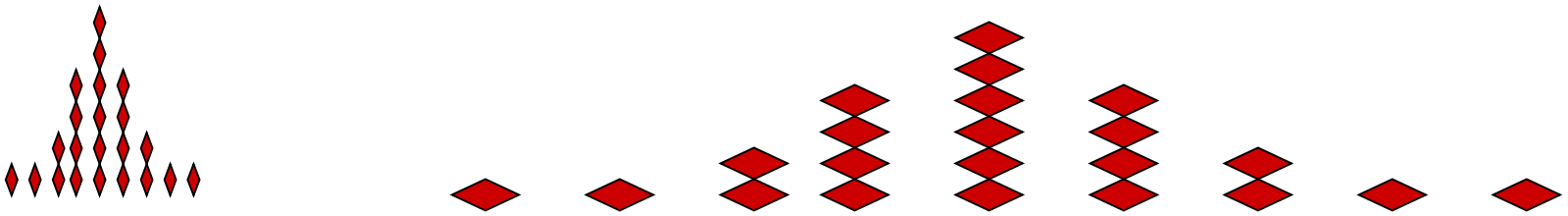
Central tendency

- Mean – Average
- Median – Middle
- Normal distribution
  - Mean = Median
- Skewed right
  - Mean > Median
- Skewed left
  - Mean < Median
- Outliers
  - Extreme values



Variability

- Range = (Max value) – (min value)
- Standard Deviation (SD) – Measure of spread (scatter pattern) ≥ 0
  - SD = 0 -> Indicates no variability, data is constant
  - As SD increases the data is scattered more



Measures of Relative Standing

- Percentiles
  - How one value compares with entire set
    - 75^{th} percentile = 75% of the values were lower and 25% higher
- Standardized scores
  - z or t scores
    - Value of 0 = the mean, positive is above negative is below
      - The number vales tell distance from the mean measured in SD

Bell Curve (Normal)

Objective of Statistics

Use a sample to make inference about a population with some level of trust/confidence.

Null Hypothesis

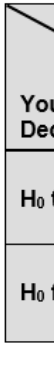
- H_0, null hypothesis: conventional belief, status quo, prevailing viewpoint.
- H_a, alternative hypothesis: competing belief, the change we're looking for.
  - For the “sake of argument,” we assume the Ho theory is true.

Now how do we judge accuracy?

- α is set at the beginning of the study as the risk willing to take that if you say there is a difference you are wrong
  - Does it have to be 0.05?
  - p value is what is reported in results as the chance that the results are wrong
- β is the chance that if you say there is no difference and the really is
  - Power is 1 – β and related to sample size, larger sample size has a better chance of showing difference if one exists
    - **If difference is shown Power for the most part doesn't matter**

Are we correct?

- What is the probability that our test will reject a false Ha
- <u>Type I Error:</u>
  - Our sample misleads us to Reject a true Ho This probability = reported p-value
- <u>Type II Error:</u>
  - Our sample leads us to not reject a false Ho. This probability = 1 – the reported power



When we Infer

- Estimation:
  - Do with some desired level of confidence or assurance
  - Confidence interval (CI) and Margin of Error (ME)
    - I'm 95% sure Average ± ME
    - I'm 98% sure percentage ± ME
  - Odds Ratio (OR)
    - Odds of an event
      - If OR crosses 1 there is no difference in groups
      - Example: OR 3.4 (CI: 0.97-5.2)

Other things to consider

- **Features of data collection that affect our results:**
- Characteristics of the sample and its generalizability
  - Inclusion Criteria & Exclusion Criteria
    - What was included or excluded
  - Statistical significance vs. Clinical significance

Application

Out Patient Antimicrobial Stewardship

- This is a new quality project for centers of Medicare Services (CMS)
- Now What?

Where to start?

- Fluroquinolone (FQ) –Not a “Bad” drug class
  - Commonly used for CAP and UTI
  - Great for drug resistant GNRs
  - Risk factor for FQ resistance is FQ exposure
- Clindamycin – More than just a risk of c. diff
  - Commonly used for SSTI
  - More broad spectrum than generally needed (anaerobic coverage)
- Azithromycin (Z-PAK) – approaching never for monotherapy coverage for anything
  - Turning into the “Placebo antibiotic”

Initial Measures

- Using the Medicare Data we can match the Part A and B encounter to the Part D Prescription fill for patients seen In the ER and not admitted we will calculate: % of antibiotic (ABX) prescriptions that are fluroquinolone (FQ) – FQ/total_ABX
  - %ABX for UTI that are FQ – FQ/total_ABX_UTI
  - %ABX for community acquired pneumonia that are FQ – FQ/total_ABX_CAP
  - Any ABX prescribed for URI_acute broncitis – ABX/URI_Dx
  - % antibacterial (so not including Tamiflu) prescribed for influenza (without pneumonia dx) – ABX/Influenza
  - %ABX for SSTI that is Clindamycin – Clinda/ total_ABX_ SSTI
  - %Clindamycin for Pharyngitis – Clinda/ total_ABX_ Pharyngitis
  - %ABX prescriptions that are Azithomycin – Azithro/Total_ABX

Summary and further questions

- Very complex problem
- It effects all walks of life
- Why does the USA have the 8th highest drug addiction rate in the world?

Objectives

- Understand basic terms of data analysis
- Be able to explain central measures
- Give examples of data collection pertinent to Antimicrobial Stewardship

THANK YOU!
Questions and Comments

References