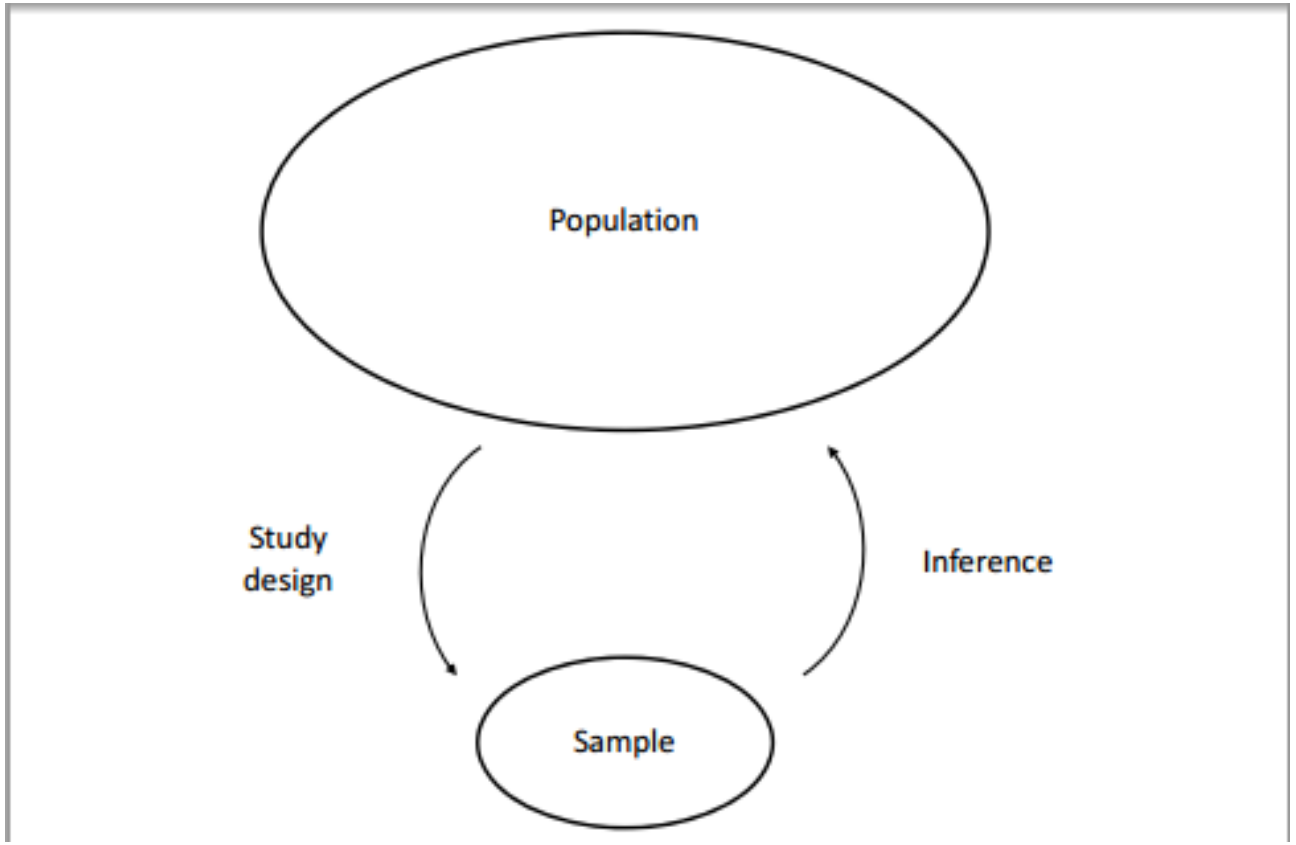
Quiz 1 Review

161:171 Introduction to Biostatistics

The Basis of Statistical Inference:



Study Types and Important Considerations

Observational

-Bias

-Difficult to determine cause-effect relationships due to confounding

Controlled Trials

-Control groups

-Randomized subject assignment

-Double Blinded

Examples of Confounding/Bias:

- Selection Bias: Researchers conduct a telephone poll on a political election (those without telephones are excluded from the poll).
- Non-response Bias: An electronics store provides each of its customers an online survey password to rate their shopping experience. (only customers with extreme opinions will go fill out the survey)
- Confounding: Homicide rates and ice cream sales are associated, however the relationship is explained by weather (both rise during warmer months)

Dealing with Confounding: Example

	Male		Female	
Class	**Total**	**Survived**	**Total**	**Survived**
1st	180	62	145	141
2nd	179	25	106	93
3rd	510	88	196	90
Crew	862	192	23	20

- Suppose we are interested in the survival rates for each gender in the Titanic example.
- What are the overall survival rates?
- Could Class be confounding this relationship?

Dealing with Confounding: Example

	Male		Female	
Class	**Total**	**Survived**	**Total**	**Survived**
1st	180	62	145	141
2nd	179	25	106	93
3rd	510	88	196	90
Crew	862	192	23	20

<u>Steps to adjust for Class:</u>

1) Calculate the overall proportion of observations in each level of class (ignoring gender)
2) Calculate the proportion of survival for each gender in each level of class.
3) Weight the gender survival proportions (step #2) by the proportions in (step #1) and sum them.

Study Interpretation: Confidence Intervals

- “If repeated, the process used to create this interval will capture the true parameter we are interested in 95% of the time”
- Factors influencing the width of a confidence interval:
  - Confidence level (higher = wider, lower = narrower)
  - Sample size (larger = narrower, smaller = wider)
  - Sample variability (more variable = wider, less variable = narrower)

P-values

- “Under the Null Hypothesis, the probability of obtaining a result as extreme or more extreme than the one we observed”
- Small p-values -> reject the Null Hypothesis, our results are unlikely to be due to random chance.
- Question: If a 95% confidence interval for a true population parameter is (1.1, 2.5) and the parameter under the Null Hypothesis is hypothesized to be 0.8, what can you say about the p-value?

Type I and Type II Errors

- Type I Error = Rejecting the Null Hypothesis when its true. (False Discovery)
- Type II Error = Not rejecting the Null Hypothesis when its false. (Missed Discovery)

Type I/Type II Error Example:

- Consider a lab that has conducted 100 studies. Suppose we know that half of the Null Hypotheses in these studies are true. Suppose there is a type I error rate of 10% and a type II error rate of 20%.
- How many false discoveries occurred?
- How many correct rejections occurred?
- How many failed discoveries occurred?

Type I/Type II Error Example:

- How many false discoveries occurred?

  5

- How many correct rejections occurred?

  40

- How many failed discoveries occurred?

  10