Statistics

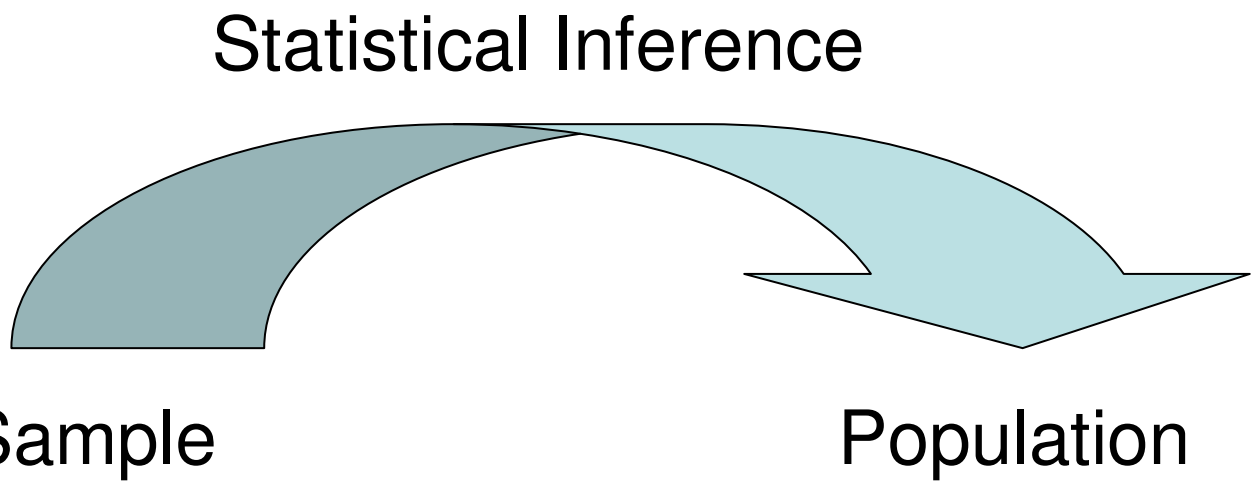
Statistical Inference, test of hypothesis and sample size calculation

January 27, 2010

Jobayer Hossain, Ph.D. & Tim Bunnell, Ph.D.

Nemours Bioinformatics Core Facility

Statistical Inference

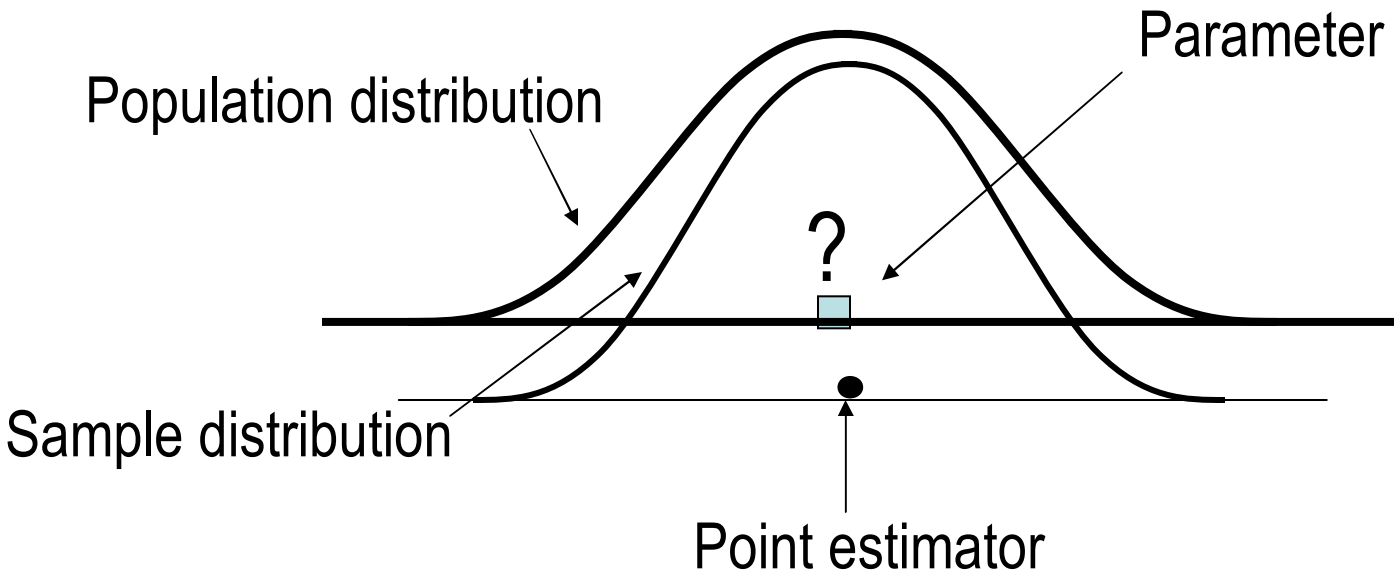


- Statistical inference is the process by which we acquire information about populations from samples.
- Two types of estimates for making inferences:
  - *Point estimation: Estimate the exact value of a specific population parameter.*
  - *Interval estimation: Estimate the range of values (an interval) within which a parameter value will probably lie with a given level of confidence.*

Statistical Inference

- Point estimate:

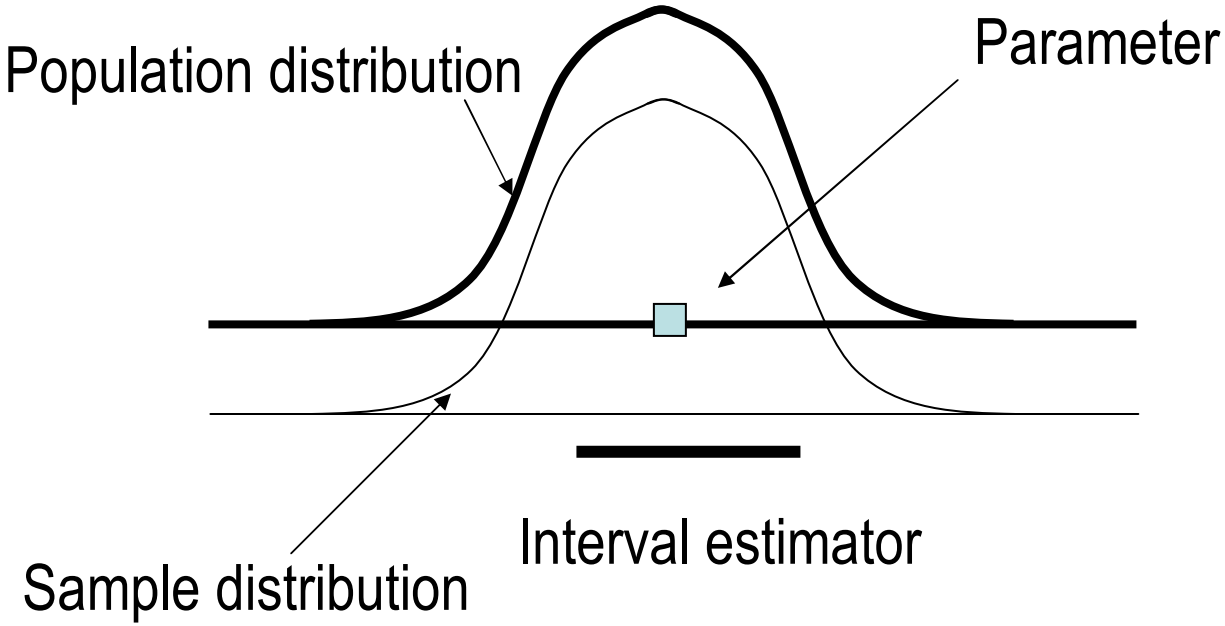
  Draws inference about a population parameter by estimating its specific value. E.g. sample mean estimates the value of the population mean.



Statistical Inference

- Interval estimate:

  Draws inferences about a population by estimating the value of an unknown parameter using an interval.



Statistical Inference

- Precision of an estimate:
  - If we repeat sampling of the same size from a population, the statistic varies as sample varies. This variability is known sampling variability.
  - Standard error (SE) of an estimate measures the sampling variability or the precision of that estimate.
    - It indicates how precisely one can estimate a population value from a given sample.
    - For a large sample, approximately 68% of times sample estimate will be within one SE of the population value.

Test of hypothesis

- Hypothesis
  - Statement or belief about some characteristic of a population or population distribution such as it's mean or standard deviation
- Types of Hypotheses
  - Null Hypothesis (H_0): The hypothesis that is assumed to be true unless contradicted by data observed in a sample. It's often a statement of statistical neutral stance and it assumes that any observed difference in data is due to chance or sampling variability. E.g. there is no effect of test drug
  - Alternative Hypothesis(H_1): The hypothesis that assumes that the observed difference in data is not just due to chance and it is the hypothesis that one must assumed if the H_0 is rejected.
- H_0 & H_1 are complementary
- Normally want to reject H_0 in favor of H_1

Test of hypothesis

		Real Situation	
Decision		**Ho is true**	**Ho is false**
	Reject Ho	**Type I error**	**Correct Decision**
	Accept Ho	**Correct Decision**	**Type II Error**

$$\alpha = P(Type\ I\ Error) \quad \beta = P(Type\ II\ Error)$$

- Goal: Keep α, β reasonably small

Test of hypothesis

- **Level of Significance**: The probability of type I error is known as the level of significance.
- **Power (1-β):** The probability of rejecting H_0 when alternative hypothesis is true at a fixed level of significance (α). That is, the probability of rejecting null hypothesis for a specified value of an alternative hypothesis.
  - Power of a test is a function of sample size and the parameter of interest.
  - We calculate power for a particular value of the parameter in alternative hypothesis.
  - An increased sample size increases power of a test.

Test of hypothesis

- We aim to make inferences controlling both type I and type II errors.
- The reduction in one results in an increase in the other
- The consequence of type I error seems to be more severe than that of type II error.
- That's why we choose a test that minimizes the probability of type II error (maximize the power of the test) keeping the probability of type I error at a fixed low level (say 0.05).

Test of hypothesis

- **Reasons for estimating power prior to conducting a research study:**
  - To determine the **sample size** required for a sufficient power to correctly detect a significant difference.

Test of hypothesis

- **Increasing Power**: Following steps can increase power of a test:
  - Increase the sample size
  - Increase the significance level (α)
  - Reduce the variability
  - Enlarge the effect size

Test of hypothesis

- Test Statistic
  - A numerical summary of a set of data that reduces the data to a single value that can be used to perform a hypothesis test
  - t, F, Chi-square, Sign, Mann-Whitney U, Wilcoxon rank sum are some commonly used test statistics.

Test of hypothesis

- P-value:
  - The probability, assuming H_0 is true, that the test statistic would take a value as extreme or more extreme than that actually observed.
  - A statement of the likelihood that the the null hypothesis is true i.e. any observed difference in data is due to chance.
  - The smaller the P-value, the stronger the evidence against H_0 provided by the data.
  - A p-value of 0.05 implies that there is a 1 in 20 chance of a Type I error, i.e., rejecting H_0 when it is actually true.

Test of hypothesis

- Confidence Interval (CI):
  - An interval within which the value of the parameter lies with a specified probability
  - A large sampling variability leads a wide interval reflecting the uncertainty of the estimate
  - It measures the precision of an estimate
  - A 95% CI of a parameter implies that if one repeats a study 100 times, the true measure of population (parameter) will lie inside the CI in 95 out of 100 measures.
  - If a parameter does not lie within 95% CI, indicates the significance at 5% level of significance

Test of hypothesis

CI = point estimate $\pm$ (measure of how confident we want to be) $\times$ (standard error)

- What effect does a larger sample size have on the confidence interval?
- It reduces standard error and makes CI narrower indicating more precision of the estimate

Test of hypothesis

- Hypothesis testing steps:

  1. Specify null (H_0) and alternative (H_1) hypotheses

  2. Select significance level (alpha) - say 0.05 or 0.01

  3. Calculate test statistic – e.g. t, F, Chi-square

  4. Calculate probability value (p-value) or confidence Interval (CI)?

  5. Describe the result and statistic in an understandable way.

Sample size calculation

- Select primary variables of interest and formulate hypotheses
- Determine/estimate standard deviation
- Decide a tolerance level of significance (α)
- Determine test statistic to use
- Determine desired power or confidence level
- Determine the effect size -- a scientifically or clinically meaningful difference

Sample size calculation

- Factors affecting sample size:
  - The larger the standard deviation, the larger the sample size needs to be
  - The smaller the likelihood of a Type I error, the larger the sample size must be
  - The greater the amount of power desired, the larger the sample size must be
  - The smaller the expected effect, the larger the sample size must be

Sample size calculation

- An Example:
  - Aim of the study: To test the hypothesis that kids with a parental history of obesity are at a higher risk of being obese themselves.
  - Suppose prevalence of obesity is 2% among US kids, whereas it is 5% among US kids with a parental history (?).
  - How many kids should we recruit to detect this difference, (.05 - .02 = .03) with a 90% power using a two-sided test with α =.05?
  - Calculation shows that 340.2 $\approx$ 341 kids are needed for this study.
- Question: What are the possible H_0, H_1, and the test statistic we need to use?

Sample size calculation

- SPSS has a separate software ‘SamplePower’ for power and sample size calculation. It is not included in the basic SPSS editor for data analysis.
- Below are some useful links for free power and sample size calculation:
  - http://stat.ubc.ca/~rollin/stats/ssize/index.html
  - http://hedwig.mgh.harvard.edu/sample_size/size.html
  - http://www.stat.uiowa.edu/~rlenth/Power/index.html
  - http://cct.jhsph.edu/javamarc/index.htm
  - http://statpages.org/#Power