STT 465 (Fall, 2016): Bayesian Statistical Methods

- ⇒ Instructor: Gustavo de los Campos (Asoc. Prof., EPI-Biostat & Statistics)
  Contact: gustavoc@msu.edu
  Primary office: IQ building (http://iq.research.msu.edu/),
  office 1311
  Websites: http://quantgen.github.io

- ⇒ Office Hours: MW 9:00-10:00 AM (Wells, office TBA)

- ⇒ Course website: https://github.com/gdlc/stt465

- ⇒ Syllabus: see course website for info about the course, pre-requisites, grading, etc.

- ⇒ Pre-requisites, rules, exams & grading policy (see syllabus)

- ⇒ Tentative schedule (see website)

Introduction

⇒ **Statistical learning**: we use data collected on finite number of subjects (sample) to learn about features of the entire population.

⇒ **Example**: Survey N individuals about likely vote in a coming presidential election; our goal is to make inferences about preferences about the entire population of likely voters.

⇒ The process of statistical learning

Data => Estimates and Measures of Uncertainty

Statistical Inference

$\Rightarrow$ Review of basic concepts:

- Population $Y=\{y_1,\ldots\}$ (may be finite or infinitely large).
- Sample We collect a sample of size n from the population $Y_s=\{y_{1,\ldots,}y_n)$

- Estimator: $\theta(Y_s)$

$\Rightarrow$ Inference: we make statements about population parameters based on the sampled data.

$\Rightarrow$ Two approaches:

- Classical (frequentist) inference: statements are based on the sampling distribution of the estimator over conceptual repeated sampling

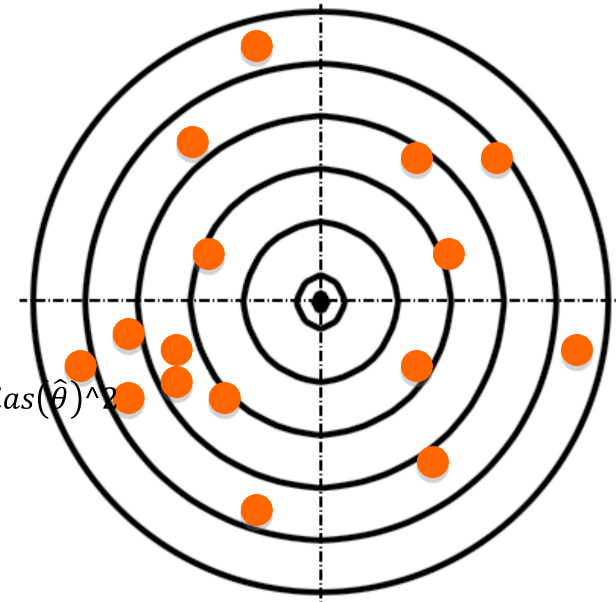
- Bayesian: statements are conditional on the observed data (the only sample we have drawn from the population).

Frequentist approach

$$Sampling \;\; Model \quad \mathrm{p}\left(Y_S \middle| \theta\right)$$

$$Sample \; Y_s = \left\{ y_i \right\} \quad Estimator \;\; \hat{\theta}\left(Y_s\right)$$

Sampling Distribution of Estimates



$$Bias\left(\hat{\theta}\right) = E\left(\hat{\theta}\right) - \theta$$

$MSE(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right] = Var(\hat{\theta}) + Bias(\hat{\theta})\hat{}2$

$$Var\left(\hat{\theta}\right) = E\left(\hat{\theta} - E\left(\hat{\theta}\right)\right)^2$$

$$MSE(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right] = Var(\hat{\theta}) + Bias(\hat{\theta})\hat{}2$$

Squared-Bias

Variance

Bayesian approach

Elements of a Bayesian Model

(1) Sampling model:

- Describes the probability of the data given a set of parameters $\mathrm{p}(Y|\theta)$

(2) Prior distribution: describe our beliefs about possible values of the parameters before we observe data. $\mathrm{p}(\theta)$

(3) Posterior distribution: the probability of the parameters given the sample $\mathrm{p}(\theta|Y)$

Describe our beliefs after we observe data.

(4) We arrive at the posterior distribution using Bayes' Theorem

$$\mathrm{p}(\theta|Y) = \frac{\mathrm{p}(Y|\theta)\mathrm{p}(\theta)}{\mathrm{p(Y)}}$$