Lesson Plan 3/29

Thursday, March 29, 2018 3:06 PM

Admin
- Reminder: we switched midterm 2 to 4/24

HDSC
- About this Cambridge Analytica Mess:
  - http://www.michalkosinski.com/
    - Psychological targeting as an effective approach to digital mass persuasion
    - Facebook as a Research Tool for the Social Sciences

Project 2
- Reminder: Due 4/6
- Questions, clarifications?

Hypothesis Testing
- The "inverse problem" view of the world
- Null and alternative
- P-value

Multivariate Probability
- Joint Probability
- Conditional Probability
- Bayes’ Rule
Inverse Problem:

\[ \frac{\text{Tweets}}{\text{pop. of}} \]

For a given model as a function of \( p \)

Then set it equal to \( \frac{1}{n} \leq X \)

solve for \( p \)
\[
\Rightarrow \hat{p} = \frac{1}{n} \sum x_i
\]

What is the precision of your estimate?

Central Limit Theorem

\[
\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})
\]

In this case, \( \hat{p} \sim N(\hat{p}, \frac{\sqrt{p(1-p)}}{\sqrt{n}}) \)

by CLT
Hypothesis Testing

How to use "Inverse Problem" procedure to test hypothesis

$p = .5$

1. Estimate $\hat{p}$ (for example $\hat{p} = .65$)

2. What can we say about $\hat{p}$, under our hypothesis $H$?

\[ \hat{p} \sim N(p, \frac{1}{n(p(1-p))}) \]
Round 3

Reject hypothesis or not.

Based on probability of $p$,

\[ 1 - p \left( \frac{\sqrt{n(p - \hat{p})}}{\sqrt{p(1-p)}} \right) \leq 5\% \]

Then reject hypothesis.
\[ P \text{-value:} \]

Null Hypothesis

Alternative

\[ P \text{-value} \]

\[ \rightarrow A/B \quad (a \text{ priori } PA) \]

Had a priori hypothesis

\[ PA = P \quad \text{known} \]
\[ H_0: \ P_B = P_A \]

\[ H_1: \ P_B \neq P_A \]

\[ \rightarrow \ A/B \quad (P_A \text{ is estimated}) \]

\[ H_0: \ P_B = P_A \]

\[ \rightarrow \ \text{Bootstrap} \]

- Use data to also estimate sampling distribution of \( \bar{\theta} \)
Inference: Estimate some parameter by solving "inverse pattern" and make statements about their precision.

Testing: Test a priori hypotheses based on inference.

Probability Distributions:
\[
\rightarrow \text{ Bernoulli:} \\
\quad \text{D: } \{0, 1\} \\
\quad X \sim \text{ Bernoulli}(p) \\
\quad P: P(X = 1) \\
\quad EX: \mu \\
\quad \text{var}X: \mu(1-\mu) \\
\rightarrow \text{ Binomial (multiple trials Bernoulli):} \\
\quad X \sim \text{ Bin}(n, p) \\
\quad D: X = 1, 2, \ldots, n \\
\quad EX: np \\
\quad \text{var}X: n \mu(1-\mu)
\[ X \sim N(\mu, \sigma) \]

\[ P(x \in \mathbb{R}) \]

\[ \text{Ex} = \mu \]

\[ \text{Var} = \sigma^2 \]

-> Joint & Conditional probability

\[ X \in [0, 0.73] : \text{bot-generated} \]

\[ V, s = 13 \] lots of
1) $p(x = x, y = y) \geq 0 \quad \forall (x, y) \in D_x \times D_y$

2) $\sum_{(x, y) \in D_x \times D_y} p(x = x, y = y) = 1$
Conditional Probability

\[ P(Y = y \mid X = 1) \]

Bayes' Rule

\[
\frac{P(Y = y \mid X = 1)}{P(X = 1)} \quad \text{if} \quad P(Y = y \mid X = x) \quad \text{is the same for} \quad x
\]