

STAT 231 — LECTURE 7

Bartosz Antczak

Instructor: Michael Wallace

September 22, 2017

7.1 Statistical Inference

So far we've focused on **descriptive statistics**, which is the portrayal of the data, or parts of it, in numerical or graphical ways to highlight features of interest. When data is used to create general conclusions, that is called **statistical inference**. Two ways of arriving at these conclusions exist:

- **Inductive reasoning:** extrapolating reasoning from the specific to the general (e.g., 25% of undergraduate students said that they like poutine, which might mean that 25% of students in general do as well)
- **Deductive reasoning:** using axioms to prove mathematical statements

The methods of statistical inference will be used to examine three main types of problems:

- **Estimation Problems:** estimates one or more attributes of a process or population (e.g., estimate the number of STAT 231 students who like poutine)
- **Hypothesis Testing:** use our data to assess the truth of some question or hypothesis (e.g., is it true that a higher proportion of statistic students than cs students like poutine?)
- **Prediction problems:** use the data to predict the future value of a particular variate

Note: you can never be too precise when discussing statistical terminology. On exams and tests, your terminology CANNOT be ambiguous.

7.2 Chapter 2: Statistical Models and Maximum Likelihood Estimation

A **statistical model** is a mathematical model that incorporates probability (or uncertainty). In our models, we define a **random variable** as a representation of a particular variate of a randomly selected unit from the population or process. The probability models we'll be focusing on in this course include:

- Binomial(θ)
- Poisson(θ)
- Exponential(θ)
- Gaussian(θ), where $\theta = (\mu, \sigma)$