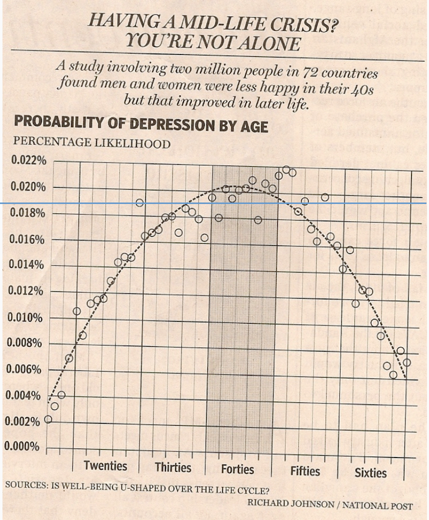
Mathematical Statistics II

STA2212H S LEC9101

Week 1

January 13 2021



Today

Start Recording

1. Course Overview
2. Review of Likelihood AoS Ch 9
3. HW3 from MS I

- January 13 3.15 – 4.15 Ruoqi Yu Zoom Link
  “Matching Methods for Observational Studies Derived from Large Administrative Databases”
- January 18 5.15 – 6.15 “Statistical Learning with Electronic Health Records Data”
  Jesse Gronsbell

Course Overview

Link

STA 2212S: Mathematical Statistics II
Wednesday, 10-12 am; Friday 10-11 am Eastern January 13 – April 12 2021

From the calendar:
This course is a continuation of STA2112H. It is designed for graduate students in statistics and biostatistics. Topics include: Likelihood inference, Bayesian methods, Significance testing, Linear and generalized linear models, Goodness-of-fit, Computational methods
Prerequisite: STA2112H

I will definitely cover the first 3 topics, and the 5th, and we'll see how time goes for the others. "Computational methods" was probably meant to be shorthand for "bootstrap" and "MCMC", and will be touched on in the other topics.

Course Delivery:

S STA 2212S: Mathematical Statistics II Link
Syllabus Spring 2021

Week	Date	Methods	References
1	Jan 13/15	Review of parametric inference	AoS Ch 9
2	Jan 20/22	Significance testing	AoS Ch 10.1,2,6,7; SM Ch 7.3
3	Jan 27/29	Significance testing	
4	Feb 3/5	Goodness of fit testing	AoS Ch 10.3,4,5,8
5	Feb 10/12	Multiple testing and FDR	AoS Ch 10.7, EH Ch 15.1,2
6	Feb 17/19	Break	
7	Feb 24/26	Bayesian Inference	AoS Ch 11.1-4; SM Ch 11.1,2; EH

Link

STA2212: Inference and Likelihood

A. Notation

One random variable: Given a model for X which assumes X has a density $f(x;\theta), \quad \theta \in \Theta \subset \mathbb{R}^k$, we have the following definitions:

likelihood function	$L(\theta; x) = c(x) f(x;\theta)$	
log-likelihood function	$\ell(\theta; x) = \log L(\theta; x) = \log f(x;\theta) + a(x)$	
score function	$u(\theta) = \partial \ell(\theta; x)/\partial\theta$	$s(x;\theta)$
observed information function	$j(\theta) = -\partial^2 \ell(\theta; x)/\partial\theta\partial\theta^T$	
expected information (in one observation)	$i(\theta) = \mathrm{E}_\theta\{U(\theta)U(\theta)^T\}^1$	$I(\theta)$ (9.11)

Independent observations: When we have X_i independent, identically distributed from $f(x_i;\theta)$, then, denoting the observed sample $\boldsymbol{x} = (x_1, \ldots, x_n)$ we have:

likelihood function $\quad L(\theta; \boldsymbol{x}) = \prod_{i=1}^n f(x_i;\theta) \quad \mathcal{L}_n(\theta)$

Notes A Notation

Notes B Asymptotics, scalar

Notes B Asymptotics, vector

Your friend the delta-method

maximum likelihood estimators are asymptotically normally distributed

maximum likelihood estimators are “equivariant”

More about likelihood

- maximum likelihood estimators are consistent Th. 9.13
- maximum likelihood estimators are asymptotically normally distributed Th. 9.18
- among all consistent estimators, maximum likelihood estimators have the smallest asymptotic variance Th. 9.23
- i.e., maximum likelihood estimators are asymptotically efficient

- likelihood functions depend on the data only through the sufficient statistic[1]
- sufficient statistics have all the information about the parameters in the model
- algebraically

$$f(\boldsymbol{x};\theta) \propto f(t;\theta)f(a \mid t),$$

where $(t,a) \leftrightarrow \boldsymbol{x}$ is a one-to-one transformation of $\boldsymbol{x} = (x_1,\dots,x_n)$
- examples Poisson, normal, gamma, logistic regression

[1] In fact AoS defines sufficiency this way

Exponential families have sufficient statistics

Exponential families are ‘smooth’

Not all families are smooth

Calculating maximum likelihood estimators

Newton-Raphson

EM-algorithm

STAT 2112: Homework 3

The homework is worth 60 pts. Problem 2c and Problem 4 are optional and worth up to 15 points of extra credit.

Problem 1

Suppose that the number of accidents per week at an intersection follows a Poisson distribution with parameter μ. Over the past 52 weeks, there were 0 accidents in 30 weeks and one or more accidents in 22 weeks. Assume that the 52 weeks are independent.

(a) 10 pts

Given this data, find the method of moments estimator of μ. Keep in mind that we only know if there were 0 or > 0 accidents for any week.

Profile likelihood

$$f(x_i; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x_i^{\alpha-1} e^{-x/\beta}$$