Python in
Healthcare

Unsupervised Anomaly Detection: Bayesian Analysis

May 4, 2017

Agenda

- **Who am I? What is Clover?**

- **What is Unsupervised Anomaly Detection?**

- **Healthcare Fraud**

- **What is Bayesian Analysis?**
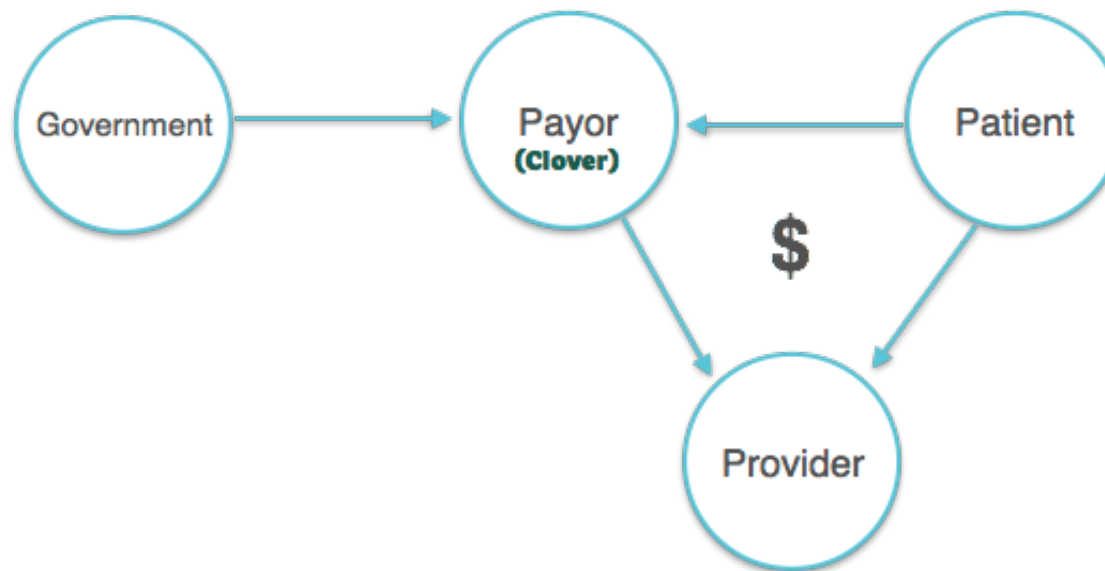
- **Analysis!**

Hi, I'm Vincent!

Product Analytics at Clover Health

- Dartmouth College – Math and Economics

- Currently doing part-time Masters in CS at Georgia Tech

- Economic Research at the Federal Reserve Board

- Started Python: 2016

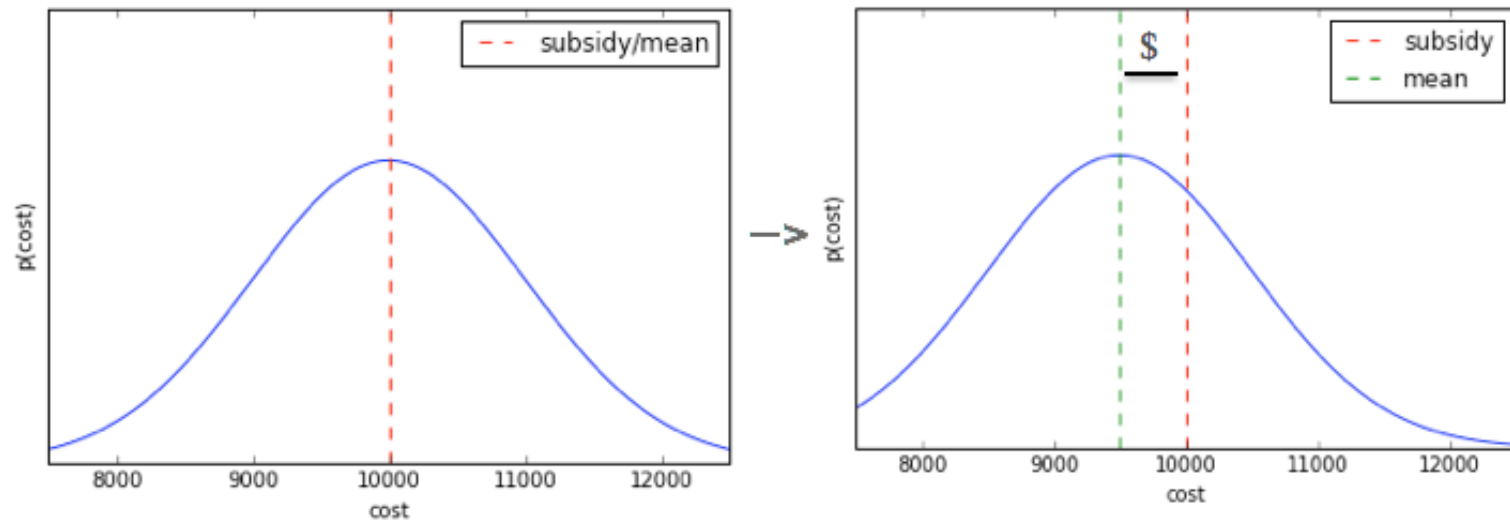# Literally a Health Insurance Company
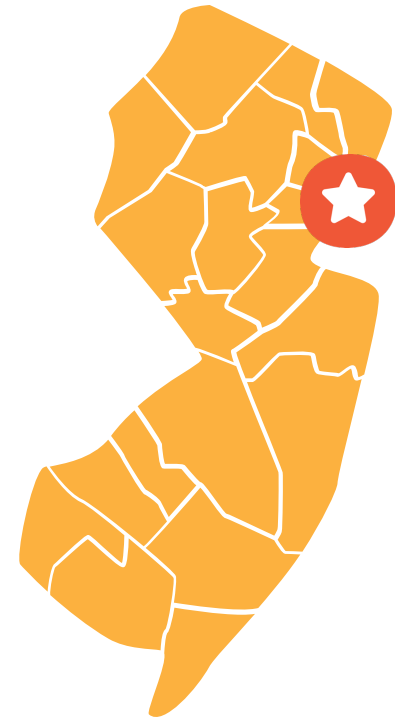
Health Insurance: Medicare Advantage

Medicare beneficiaries are generally 65 and older, have multiple chronic conditions, and engage the healthcare system more than the average population

The Premise

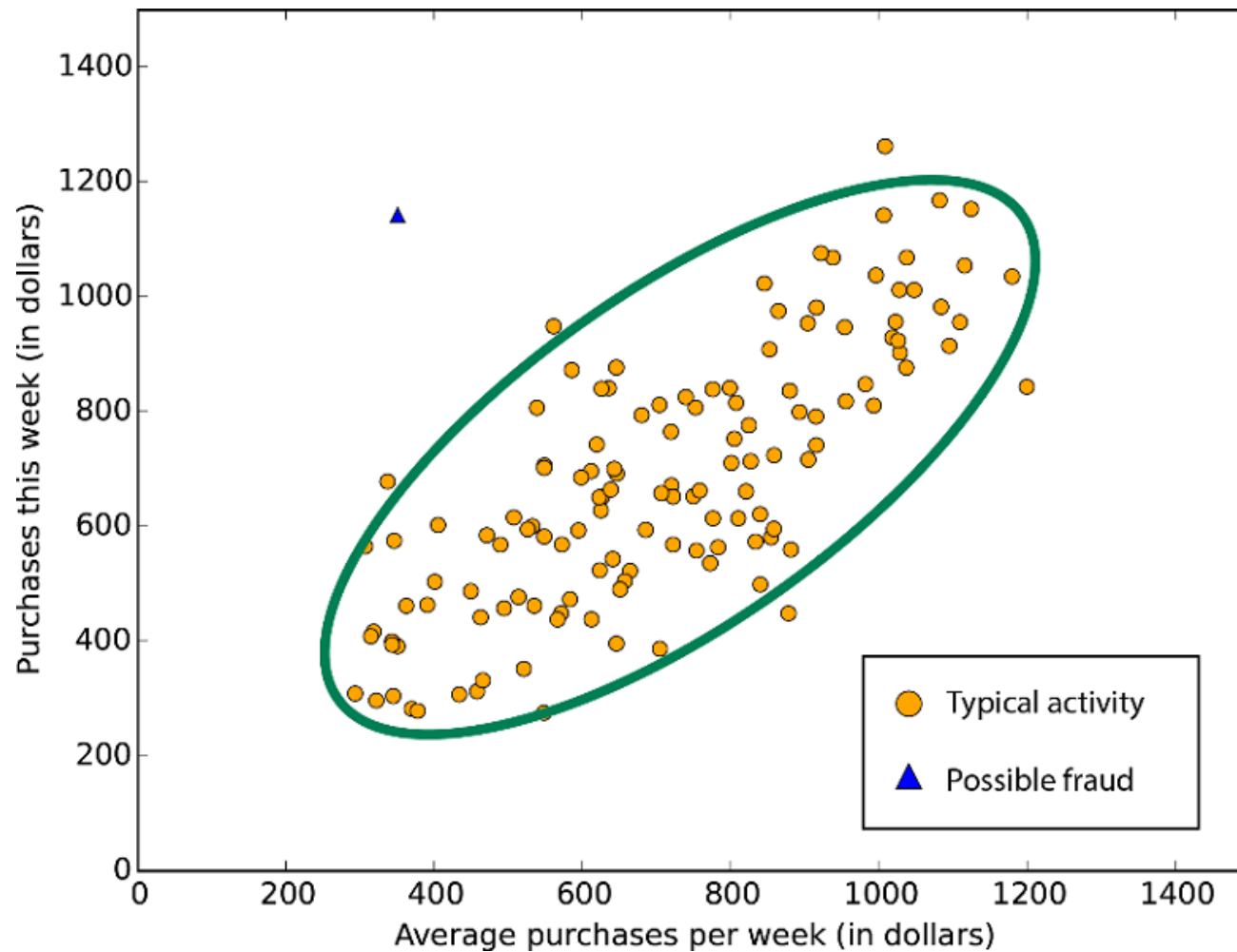Identify care gaps —> Modify risk probabilities —> Member benefits —> Profit

We win when patients win!
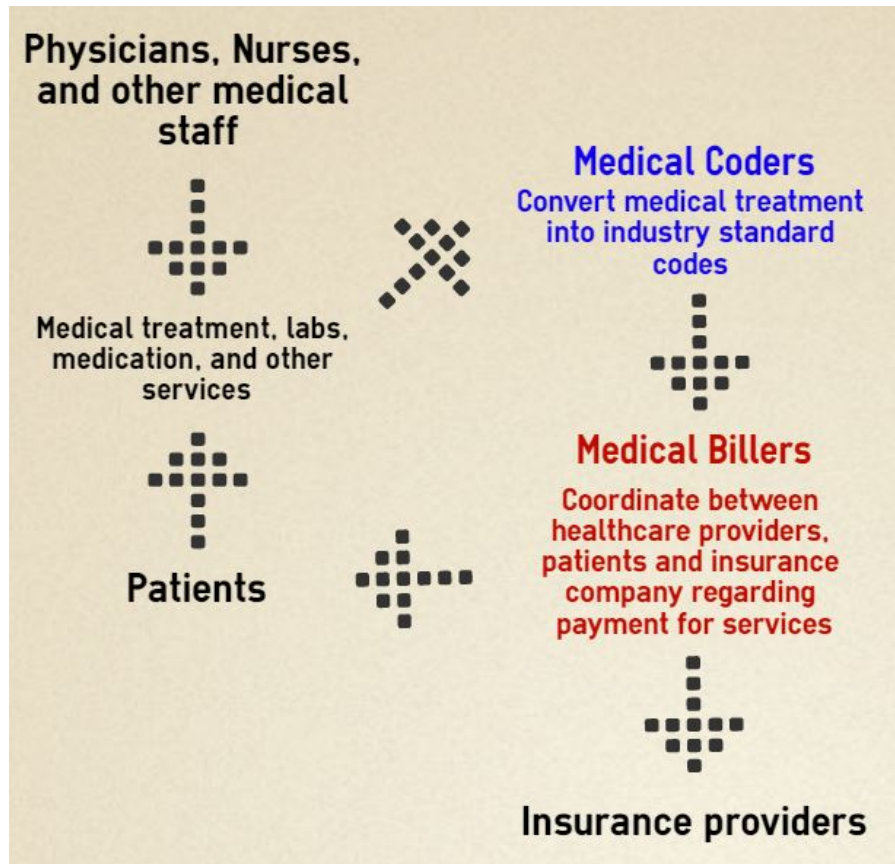
What is Unsupervised Anomaly Detection?

Unsupervised Anomaly Detection

A quick primer on Medicare Billing

Coding and Billing

Physicians, Nurses, and other medical staff

Medical treatment, labs, medication, and other services

Patients

Medical Coders
Convert medical treatment into industry standard codes

Medical Billers
Coordinate between healthcare providers, patients and insurance company regarding payment for services

Insurance providers

It's Complicated!

- Real interactions need to get translated into codes.
- We're going to focus on Evaluation and Management (E/M) Codes

Evaluation and Management (E/M) Codes

Very Common

- Medicare pays for over 200 million office visits a year
- $32.3 billion in payments in 2010

Types of Services

- Consultation
- Review of illness history
- Physical Exam

Five Different Levels

- Providers bill for different levels of intensity
- Higher intensity = Higher reimbursement
- **Possible area of abuse!**

Upcoding: Billing for services at a higher intensity than rendered

"

Evaluation and management services are visits performed by physicians to assess and manage a beneficiary's health. Medicare paid $32.3 billion for E/M services in 2010. In total, Medicare inappropriately paid $6.7 billion for claims for E/M services in 2010 that were a result of incorrectly coded and/or lacking documentation, representing 21 percent of Medicare payments for E/M services that year.

Department of Human and Health Services:
Office of Inspector General

Improper Payments For Evaluation and Management Services Cost Medicare Billions in 2010

Bayesian Framework

How to think like a Bayesian

Likelihood
Probability of collecting this data when our hypothesis is true

Bill Howe, UW

Prior
The probability of the hypothesis being true before collecting data

$$P(H|D) = \frac{P(D|H)\ P(H)}{P(D)}$$

Posterior
The probability of our hypothesis being true given the data collected

Marginal
What is the probability of collecting this data under all possible hypotheses?

Example:

- Posterior: Probability of having cancer given a positive screening

- Likelihood: Probability of a positive screening given cancer

- Prior: Probability of having cancer

- Marginal: Probability of a positive screening

Beta Binomial Distribution
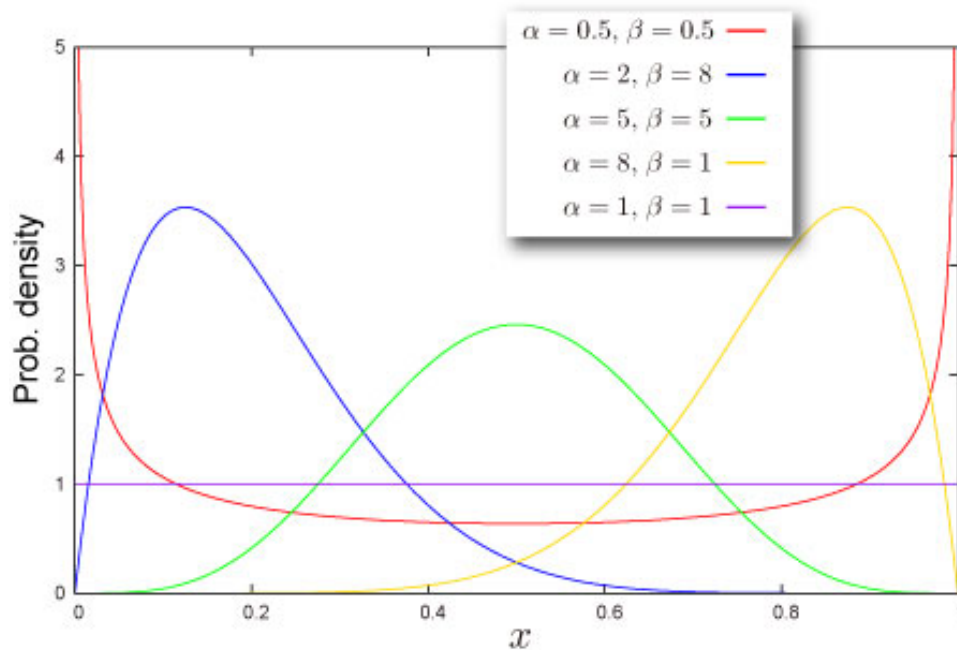
Great for percentage-type measurements!

- Binomial distribution but p varies

- Shape parameters $\alpha > 0$ and $\beta > 0$.

- Detect anomalies of prevalance of level 5 billed E/M codes

The Beta distribution has nice **mathematical properties**. Let's talk after if you're interested.

Analysis!

Limitations

Not Perfect

- Inherent member selection bias!

- What constitutes as an anomaly?

- Results are sensitive to priors

Thank you & contact information

Vincent La

Data Scientist, Clover Health

vincent.la@cloverhealth.com

LinkedIn
@vincentla