Pattern vs Process

Lecture #17 | GEOG 510
GIS & Spatial Analysis in Public Health

Varun Goel

Outline

- Geographic Pattern and Process
- Deterministic and stochastic
- Complete Spatial Randomness
- Processes
  - 1st and 2nd order

Pattern and Process

- Simple difference
  - A spatial pattern is generally the result of some spatial process
- Similar to the difference between "analysis" types, descriptive and explanatory
  - We map patterns (observe, describe)
  - We model processes (understand, explain)

Process Types

- Deterministic spatial process
  - No random element
    - Always produces the same result
  - Completely predictable
    - Perfectly modeled as a mathematical formula
  - Generally, <u>not</u> observed in the complex phenomena we study
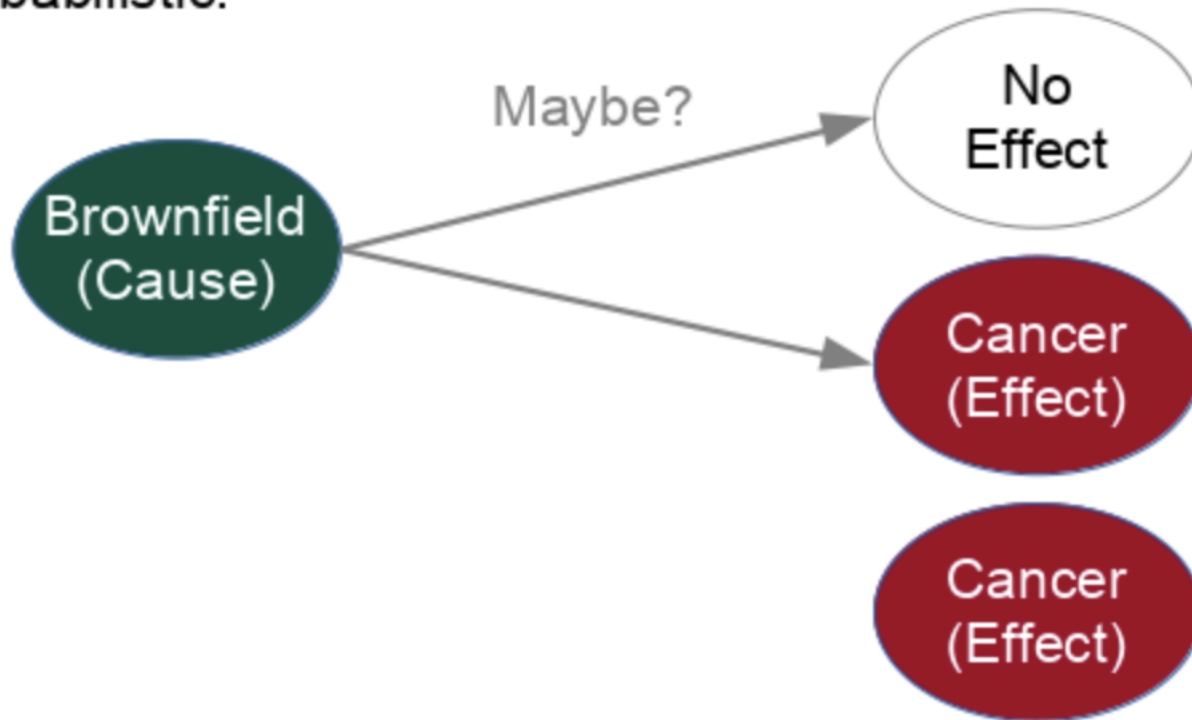   - Humans!

Process Types

- Stochastic spatial process
  - Contains random element
  - Not perfectly predictable
    - Often, we can describe these processes using mathematical formulas... but not perfectly
  - Generally, this is what we will observe in the complex phenomena we study
    - ...and, the reason we use statistics!
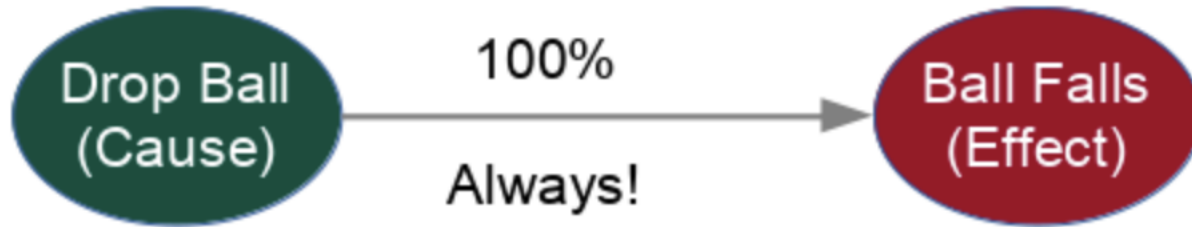
Deterministic vs. Probabilistic Causation

CSR (Complete Spatial Randomness)

- Important concept!
  - No observable pattern
  - No process (other than randomness) governing the locations or spatial distribution
    - e.g., for points, locations are simply distributed randomly throughout a region
    - e.g., for polygons or cells (with attributes), values of the attributes are randomly distributed throughout a region

CSR (Complete Spatial Randomness)

- Important concept!
  - We use CSR as a comparison for tests of our observed spatial patterns
    - Basically, is the observed pattern different than a random pattern?
  - From geography, we know that many geographic patterns are not completely random
    - e.g., spatial autocorrelation

Spatial Processes

- First order effects
  - Observed spatial variation is due to an extrinsic (external) factor (or factors)
    - For example
      - Disease cases and population counts
      - Health outcomes and poverty
    - Often, we use correlation/regression based methods to capture first order effects
      - Spatial regression

Spatial Processes

- Second order effects
  - Spatial variation is due to interaction or intrinsic factors
    - For example
      - Directly transmittable disease (cases spawn more cases)
      - Some retail behavior (competition), Dispersion
    - Diffusion models

Pattern and Process

- The "map" as hinderance to the advancement of GIS
  - Maps are great and we map patterns all the time (describe)
    - Sometimes, this is highly valuable in itself
  - We use the underlying data values to evaluate the pattern
    - Spatial autocorrelation and clustering

Pattern and Process

- The "map" as hinderance to the advancement of GIS
  - We use models to understand processes (explain, understand)
    - Can be somewhat limited because nature of geographic health information
  - For your projects, you should be thinking about "process" rather than "pattern"
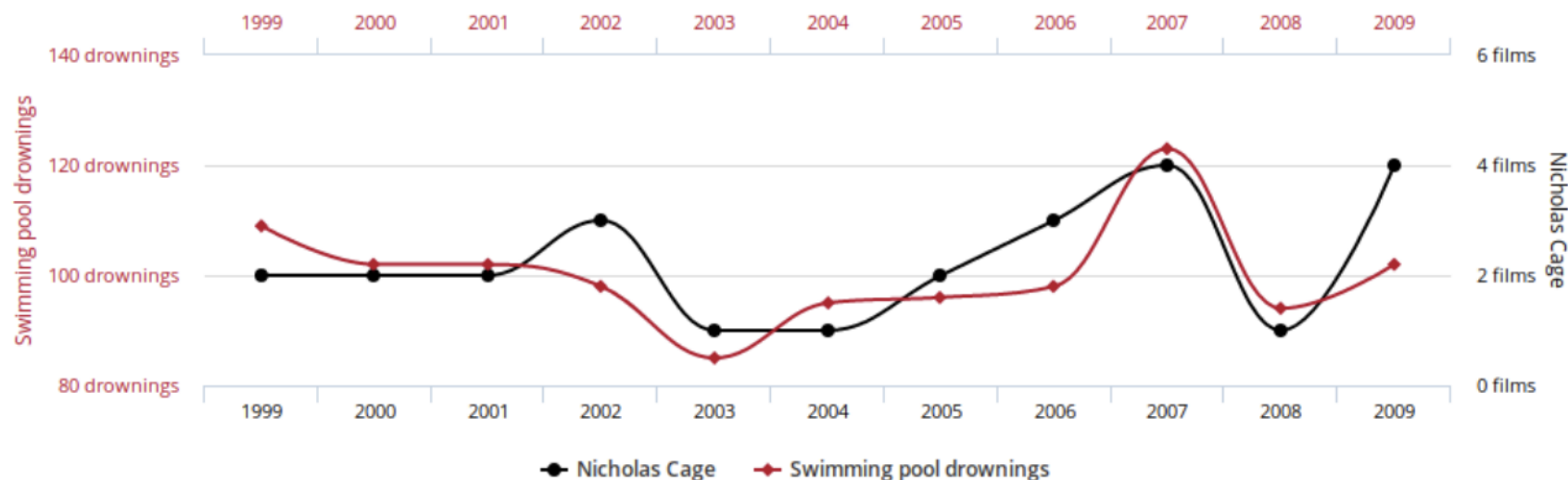
Pattern and Process

- Main Question
  - What "Process(es)" may be driving the spatial "Pattern(s)" that we observe
    - In most statistics (frequentist): What is probability of observed data given a realized 'process'.
    - Bayesian statistics: What is the probability of a realized 'process' given the data (and prior evidence)?

Pattern and Process

- How do we infer process
  - Theory-driven approaches
  - Data-driven approaches
    - Hypothesis testing
      - Do we have enough evidence to reject the "null" hypothesis? (think p-values)
    - Correlation vs Causation
      - Correlation/Association – How is a process correlated/associated with the pattern?
      - Causation – How does a process cause the pattern

Highly recommend: https://michaelminn.net/tutorials/correlation/
https://www.tylervigen.com/spurious-correlations
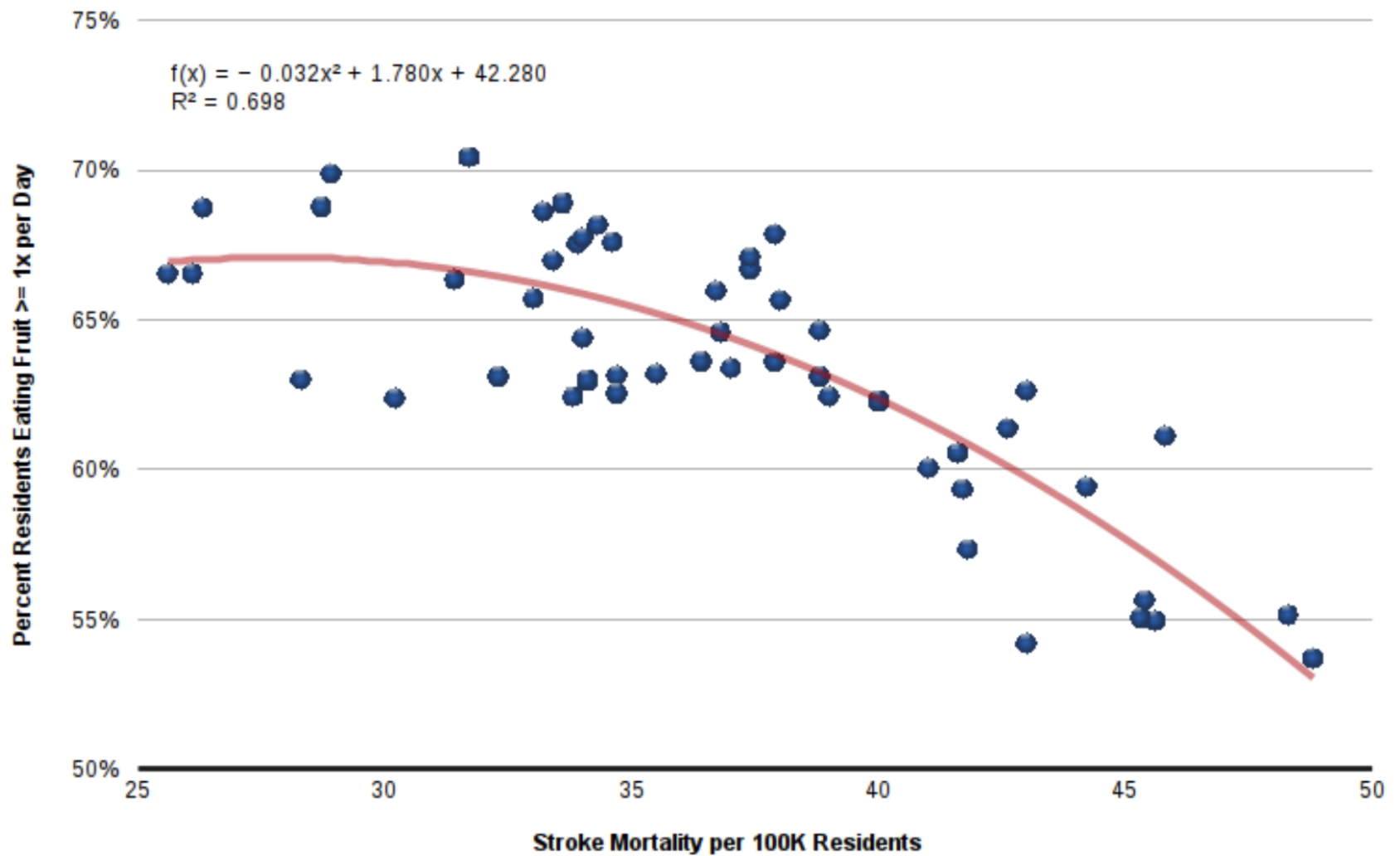
2017 Stroke Mortality vs Fruit Consumption by State (CDC)
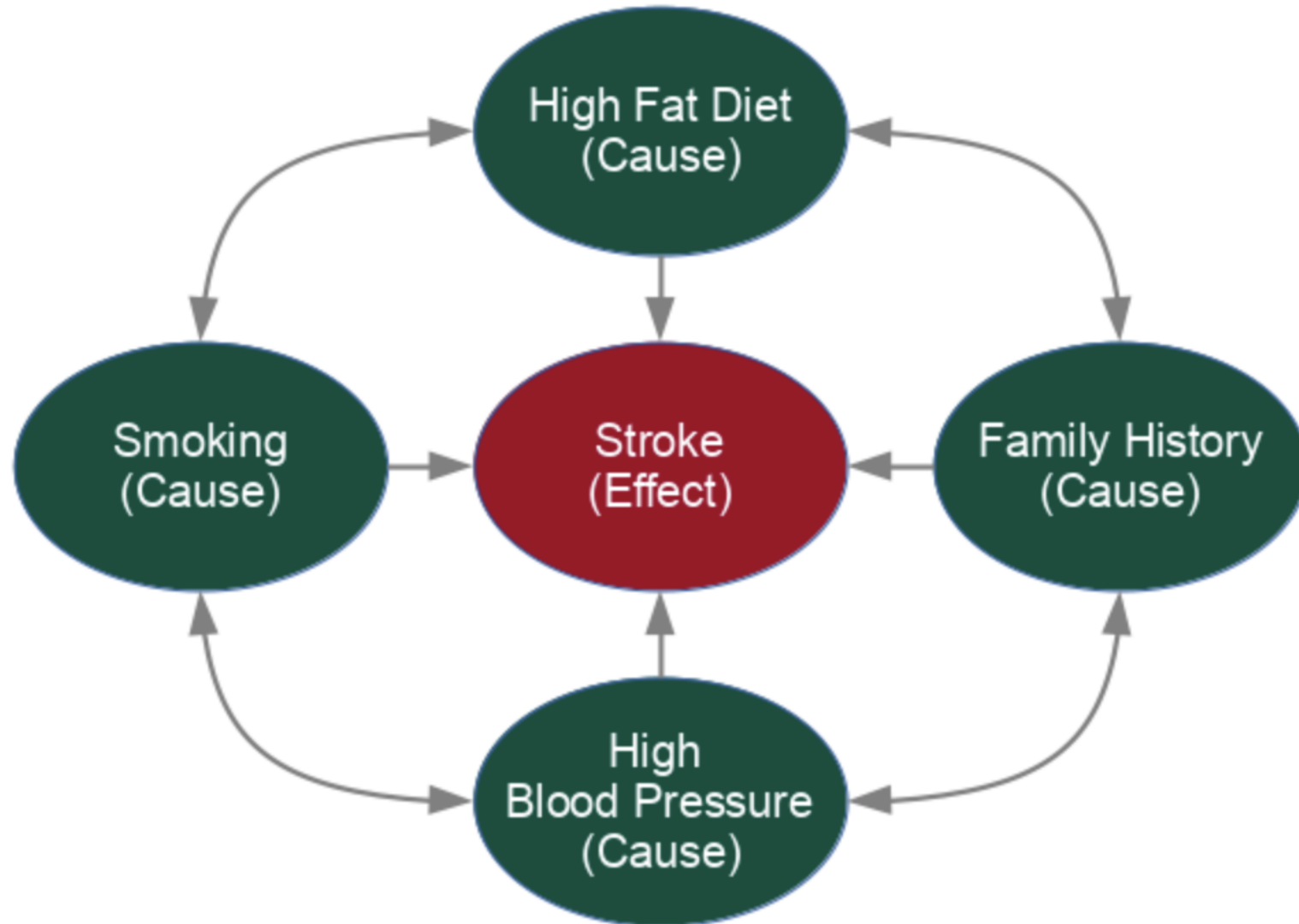
Pattern and Process

- Correlation VS Causation
  - For this class, you will mainly focus on correlations/associations
    - Regression modeling
    - Simple statistical tests
      - Are 2 variables correlated?
  - Causation is complex
    - Requires solid understanding of theory
    - Study design, counterfactual modeling
    - Natural Experiments, Randomized Control Trials

Pattern and Process

- Correlation VS Causation
  - For this class, you will mainly focus on correlations/associations
    - Regression modeling
    - Simple statistical tests
      - Are 2 variables correlated?
  - Causation is complex
    - Requires solid understanding of theory
    - Study design, counterfactual modeling
    - Natural Experiments, Randomized Control Trials

Keywords

- Geographic Pattern and Process
- Deterministic and stochastic
- Complete Spatial Randomness
- Processes
  - 1st and 2nd order
- Causation
- Correlation