Statistics in the News

- What is statistics?
- What do statisticians do?
- Why do I have to take that statistics course?

Statistics

- What is statistics?
- What do statisticians do?
- Why do I have to take that statistics course?
- An information science
- A guide to randomness
- A guide to variation
- A framework for understanding surprises
- Work on research teams
- Study methods for collection and analysis of data
- Study the science of information
- Teach statistics
Statistics

• What is statistics?
• What do statisticians do?
  • Why do I have to take that statistics course?

“Lies, Damned Lies and Statistics”
• Use current news items for topics
• Learn the statistics behind the headlines
• Show the breadth of statistical application

Inspiration

“A course called Chance”
J. Laurie Snell
Dartmouth College
Morrie DeGroot
CMU

This week in the news

Outline

• Can chocolate save your life?
  – Types of studies, “statistically significant”
• Where are all the girls?
  – Experiments and small effects, bell curves
• The Netflix Grand Prize
  – Supercrunching and data mining, ensemble methods, regularization
• How is your recovery money spent?
  – Graphical displays of information
• 2005 study: 10 males and 10 females with high blood pressure
• Two week period of “treatment” with dark chocolate
• One week washout period
• Two week period of “treatment” with white chocolate
• Blood pressure decreased during the dark chocolate period: **-11.0 SBP**  **-6.2 DBP**
  
```
  +/- 6.3  +/- 4.2
```

• A randomized, controlled, cross-over study
• 100g chocolate daily for two weeks
• 20 people? two weeks? Convincing?
• Why was the study carried out?
• Cocoa has flavonols, which are thought to have various health benefits
• How else can we study this?
**The EPIC Study**

- **European Prospective Investigation into Cancer**
- 19,357 people, aged between 35 and 65
- Followed for at least ten years
- Asked about diet at beginning of study
- Other factors measured at the beginning of study
- Information collected on heart attacks and strokes and blood pressure

**Results of the EPIC Study**

- Decrease in blood pressure: -1.0 SBP; -0.9 DBP
  
  +/- 0.4        +/- 0.3

- Relative risk of heart attack and stroke: 0.61 (0.44–0.87)

- 39% reduction in relative risk
- From top quartile of consumption to bottom
- “If people in the group eating the least amount of chocolate increased their chocolate intake by six grams a day, 85 fewer heart attacks and strokes per 10,000 people could be expected to occur over a period of about ten years.”

**Confounding factors**

- Do people who eat more chocolate eat more everything? ✓
- How do people report how much chocolate they eat? ✓
- How do we know who had heart attacks and strokes? ✓
- Do people who eat more chocolate already have lower blood pressure? ✓
Meta-analysis: Desch et al. (2010)

**SBP down 4.5**

• EPIC Study drafted a press release
Outline

- Can chocolate save your life?
- Where are all the girls?
- The Netflix Grand Prize
- How is your recovery money spent?

Can girls really do math?

Why don’t girls do math?

- Lack of abilities, e.g. spatial skills
- Lack of beliefs about abilities
- “Stereotype threat”
- Lack of supportive environment for college students
- Lack of supportive environment for college faculty
- Implicit bias  [https://implicit.harvard.edu/implicit/]
- Workplace bias
Result: The “leaky pipeline”

Advanced placement calculus

<table>
<thead>
<tr>
<th>Boys</th>
<th>Girls</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>114,285</td>
<td>108,249</td>
<td>5.2%</td>
</tr>
</tbody>
</table>

Grade

<table>
<thead>
<tr>
<th>Boys</th>
<th>Girls</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>2.8</td>
<td>9.7%</td>
</tr>
</tbody>
</table>

Mathematics Degrees

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bachelor’s</td>
<td>56%</td>
<td>34%</td>
</tr>
<tr>
<td>Doctoral</td>
<td>70%</td>
<td>30%</td>
</tr>
</tbody>
</table>

College Faculty

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untenured</td>
<td>77.2%</td>
<td>22.8%</td>
</tr>
<tr>
<td>Tenured</td>
<td>79.4%</td>
<td>20.6%</td>
</tr>
</tbody>
</table>
Why don’t girls do math?

- Lack of abilities, e.g. spatial skills
- Lack of beliefs about abilities
- Stereotype threat
- Lack of supportive environment for college students
- Lack of supportive environment for college faculty
- Implicit bias  https://implicit.harvard.edu/implicit/
- Workplace bias

Stereotype threat

N  There are no gender differences in mathematical abilities, Researchers Say

S  Stereotype: Women’s body in art; Women’s unique experience

E  Expectations are responsible for gender differences in mathematical abilities, Researchers Say

G  Genes are involved in mathematical abilities, Researchers Say

---

G  Environment

E  No differences

N  standard

Dar-Nimrod and Heine, 2006: Study 1

---

G  Environment

E  No differences

N  standard

Dar-Nimrod and Heine, 2006: Study 2
The bell curve strikes

- Assume scores are drawn from normal distribution
- Compute probability of an observed difference if the scores came from the same normal distribution
- If this probability is small, result is “statistically significant”
- Effect of stereotype threat in Dar-Nimrod and Heine’s experiment is statistically significant

\[
\int \frac{\partial^2 f(x, y)}{\partial h(x) u(y)} dx dy
\]

Advanced Calculus

- Stereotype threat
  - “math test aimed at measuring your mathematical abilities”
- Non-threat condition
  - “this mathematics test shows no gender differences”

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threat</td>
<td>3.13</td>
<td>3.08</td>
</tr>
<tr>
<td>Non-threat</td>
<td>3.60</td>
<td>2.60</td>
</tr>
</tbody>
</table>

Course grades similar between men and women
Netflix Prize

The Contest

• Cinematch error 0.9525
• Goal 10% reduction
• Target error 0.8572
• Winners, July 26, 2009 0.8567

Leaderboard

The Data

• 18,000 movies; 500,000 users; 7 years
• 100 million ratings
• Build a prediction model for ratings
• Netflix tests this on a hold-out set of 4 million
• Computes a measure of prediction error

The Statistics

1. The Amazon strategy: “people who liked this movie also liked...”
• Nearest neighbour models

\[ \hat{r}_{ui} = \frac{\sum_{j \in N(i;u)} s_{ij} \hat{r}_{ij}}{\sum_{j \in N(i;u)} s_{ij}} \]

• \( N(i; u) \) Set of movies \( i \) rated by user \( u \)
• \( s_{ij} \) Similarity between movies \( i \) and \( j \)
2. **Matrix factorization**
- Characterize movies by latent factors
- Characterize users by latent factors
- Amount of violence, drama/comedy, chick flick
- User preferences for these

$$\hat{r}_{ui} = q_i^T p_u$$

- Vectors of movie/user factors $q_i$, $p_u$

3. **Baseline predictors**
- Some movies are more popular than others
- Some users are easier to please than others

$$\hat{r}_{ui} = \mu + b_i + b_u + q_i^T p_u$$

- These things change over time $b_i(t), b_u(t)$
- And so do latent factors $q_i(t)^T p_u(t)$

- Estimating the unknowns
  - Least squares

$$\min_{b, p, q} \sum_{(u,i) \in T} (r_{ui} - \mu - b_i - b_u - q_i^T p_u)^2 +$$

$$\lambda_1 (\sum_u b_u^2 + \sum_i b_i^2) + \lambda_2 \sum_i ||q_i||^2 + \lambda_3 \sum_u ||p_u||^2$$

- Regularization/shrinkage
**The Surprises**

- Averaging predictions from different methods was quick and effective
- “At the end of the first year of the competition, our submission was a linear combination of 107 prediction sets”
- 8.43% improvement
- Different teams on the leaderboard also merged
- Big Chaos/ BellKor/ Pragmatic Theory

**The Lessons**

- Generated interest in recommender systems
- Increased collaboration among computer scientists, statisticians, engineers
- Benefits of ensemble methods remarkable
- Combination of first and second place teams reduced error to **0.8555**  
  **0.8572**
  **0.8567**
- None of these things are easy! (contest 2)
Visualization
http://www.flickr.com/photos/chef_ely/3791293142/sizes/l/in/
set-72157621825510293/

Outline
- Can chocolate save your life?
- Where are all the girls?
- The Netflix Grand Prize
- How is your recovery money spent?
Statistical Graphics

Bad Graphs!

USA TODAY Snapshot

And good
... references

Netflix
- Koren, Y. et al. (2009). The BellKor Solution to the NGP.
- Tüscher, A. et al. (2009). The BigChaos Solution to the NGP.
- Pirotte, M. et al. (2009). The Pragmatic Theory solution to the NGP.
- http://www.netflixprize.com/index
- The Ensemble Visualization:
  http://www.the-ensemble.com/content/netflix-prize-movie-similarity-visualization

Graphics
- http://www.recovery.gov/
- http://www.gapminder.org/

References

Chocolate

Girls