Which police officers have a high propensity to escalate force?

Greg Ridgeway

Rebecca W. Bushnell Professor of Criminology
Professor of Statistics and Data Science
University of Pennsylvania

September 20, 2024

Some officers seem inclined to escalate

- Laquan McDonald shooting, October 20, 2014
- CPD Officer Van Dyke fired 16 rounds
- Officer Walsh fired no rounds, holstering his firearm

Type of force depends on officer and environment

- $Y = y$ indicates type of force, $y \in \{0, 1, 2, 3\}$
- Each officer has λ, latent propensity to escalate force
- Environment $\mathbf{z}$ (e.g., time, place, lighting, suspect, policies and laws)

$$P(Y_i = y|\mathbf{z}) = \frac{\exp(\theta_y + s_y(h(\mathbf{z}) + \lambda_i))}{\sum_{k=0}^{3} \exp(\theta_k + s_k(h(\mathbf{z}) + \lambda_i))}$$

- Derived from a flexible multinomial distribution
- With an order constraint on $\mathbf{s}$, identical to Anderson (1984) ordinal stereotype model

Type of force depends on officer and environment

- $Y = y$ indicates type of force, $y \in \{0, 1, 2, 3\}$
- Each officer has λ, latent propensity to escalate force
- Environment $\mathbf{z}$ (e.g., time, place, lighting, suspect, policies and laws)

$$P(Y_i = y|\mathbf{z}) \quad = \quad \frac{\exp(\theta_y + s_y(h(\mathbf{z}) + \lambda_i))}{\sum_{k=0}^{3} \exp(\theta_k + s_k(h(\mathbf{z}) + \lambda_i))}$$

- Derived from a flexible multinomial distribution
- With an order constraint on $\mathbf{s}$, identical to Anderson (1984) ordinal stereotype model

Type of force depends on officer and environment

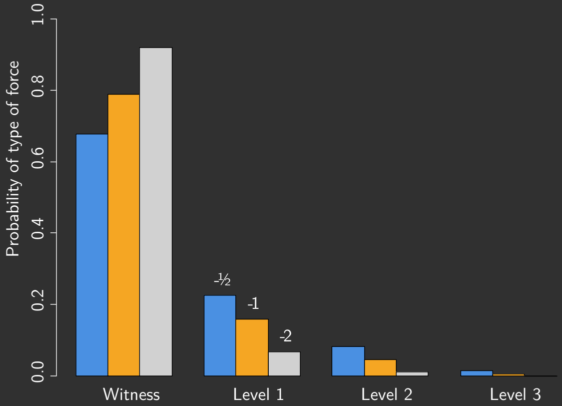
- $Y = y$ indicates type of force, $y \in \{0, 1, 2, 3\}$
- Each officer has λ, latent propensity to escalate force
- Environment $\mathbf{z}$ (e.g., time, place, lighting, suspect, policies and laws)

$$P(Y_i = y|\mathbf{z}) \quad = \quad \frac{\exp(\theta_y + s_y(h(\mathbf{z}) + \lambda_i))}{\sum_{k=0}^{3} \exp(\theta_k + s_k(h(\mathbf{z}) + \lambda_i))}$$

- Derived from a flexible multinomial distribution
- With an order constraint on s, identical to Anderson (1984) ordinal stereotype model

Ordinal stereotype implies force type distribution

- θ set to match Seattle's force rate
- $h(\mathbf{z}) = 0$
- $\mathbf{s} = \{0, 1, \frac{3}{2}, 2\}$
- $\lambda_1 = -\frac{1}{2}$, $\lambda_2 = -1$, $\lambda_3 = -2$



Estimating each officer's λ

Consider a moment with features $\mathbf{z}$ with Officers 1, 2, and 3

- one officer does nothing
- one officer physically restrains
- one officer strikes baton to the head

What's the probability $Y_1 = 0$, $Y_2 = 1$, and $Y_3 = 3$?

$$P(Y_1 = 0, Y_2 = 1, Y_3 = 3 | \mathbf{k} = \{1, 1, 0, 1\}, \mathbf{s}, \lambda, \theta, h(\mathbf{z})) = \frac{e^{s_0\lambda_1 + s_1\lambda_2 + s_3\lambda_3}}{e^{s_0\lambda_1 + s_1\lambda_2 + s_3\lambda_3} + \ldots + e^{s_3\lambda_1 + s_1\lambda_2 + s_0\lambda_3}}$$

Estimating each officer's λ

Consider a moment with features $\mathbf{z}$ with Officers 1, 2, and 3

- one officer does nothing
- one officer physically restrains
- one officer strikes baton to the head

What's the probability $Y_1 = 0$, $Y_2 = 1$, and $Y_3 = 3$?

$$P(Y_1 = 0, Y_2 = 1, Y_3 = 3 | \mathbf{k} = \{1, 1, 0, 1\}, \mathbf{s}, \lambda, \theta, h(\mathbf{z})) = \frac{e^{s_0\lambda_1 + s_1\lambda_2 + s_3\lambda_3}}{e^{s_0\lambda_1 + s_1\lambda_2 + s_3\lambda_3} + \ldots + e^{s_3\lambda_1 + s_1\lambda_2 + s_0\lambda_3}}$$

Advantages of conditional likelihood

For a general incident with m officers

$$P(\mathbf{Y} = \mathbf{y}|\mathbf{s}, \lambda, \mathbf{k}) = \frac{\exp\left(\sum_{i=1}^{m} s_{y_i}\lambda_i\right)}{\sum_{\mathbf{y}^* \in \mathcal{K}} \exp\left(\sum_{i=1}^{m} s_{y_i^*}\lambda_i\right)}$$

- No need for environmental/situational measures
- Moments when all officers have $y = 0$ have no information
- Use-of-force incidents involving a single officer have no information

The only times and places with information for the conditional likelihood are those with multiple officers on the scene of a use-of-force incident

- Denominator is a combinatorial challenge
- λs are not identifiable, only some $\lambda_i - \lambda_j$ are identifiable

Advantages of conditional likelihood

For a general incident with m officers

$$P(\mathbf{Y} = \mathbf{y}|\mathbf{s}, \lambda, \mathbf{k}) = \frac{\exp\left(\sum_{i=1}^{m} s_{y_i}\lambda_i\right)}{\sum_{\mathbf{y}^*\in\mathcal{K}} \exp\left(\sum_{i=1}^{m} s_{y_i^*}\lambda_i\right)}$$

- No need for environmental/situational measures
- Moments when all officers have $y = 0$ have no information
- Use-of-force incidents involving a single officer have no information

The only times and places with information for the conditional likelihood are those with multiple officers on the scene of a use-of-force incident

- Denominator is a combinatorial challenge
- λs are not identifiable, only some $\lambda_i - \lambda_j$ are identifiable

Advantages of conditional likelihood

For a general incident with m officers

$$P(\mathbf{Y}=\mathbf{y}|\mathbf{s},\lambda,\mathbf{k}) = \frac{\exp\left(\sum_{i=1}^m s_{y_i}\lambda_i\right)}{\sum_{\mathbf{y}^*\in\mathcal{K}}\exp\left(\sum_{i=1}^m s_{y_i^*}\lambda_i\right)}$$

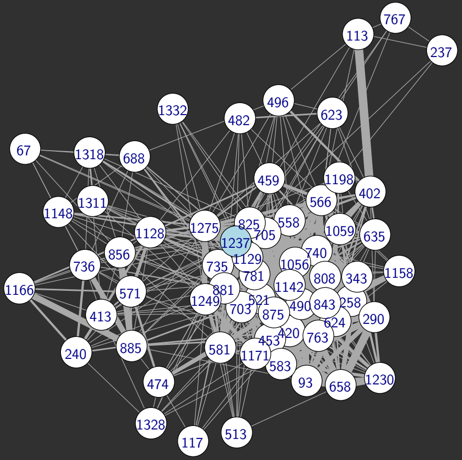
- No need for environmental/situational measures
- Moments when all officers have $y=0$ have no information
- Use-of-force incidents involving a single officer have no information

> The only times and places with information for the conditional likelihood are those with multiple officers on the scene of a use-of-force incident

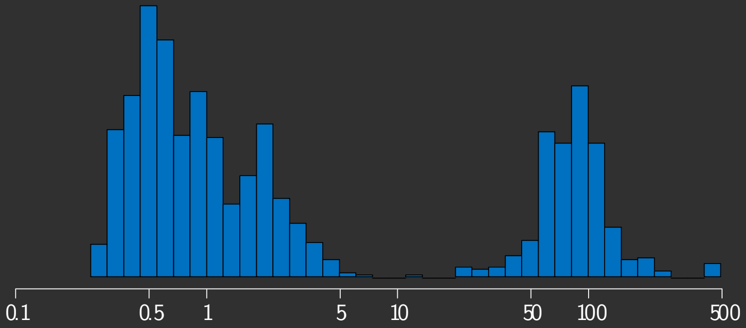
- Denominator is a combinatorial challenge
- λs are not identifiable, only some $\lambda_i-\lambda_j$ are identifiable

Precise comparisons possible only between well-connected officers

- Use-of-force subgraph from Seattle PD



$\mathrm{Var}(\lambda_i|\lambda_{1237})$ reveals well-connected officers



- Histogram shows $\mathrm{Var}(\lambda_i|\lambda_{1237})$
- $\lambda_{1237} - \lambda_i$ is strongly identifiable only if Officer 1237 and Officer i share enough information
- Officers in disconnected subgraphs or with few shared incidents will have a large conditional variance

Flag officers with high force escalation

Officer ID	Count of force type used				Peers	Prob. rank
	Witness	Level 1	Level 2	Level 3		top 5%
1237	5	11	7	0	626	0.94

Flag officers with high force escalation

Officer ID	Count of force type used				Peers	Prob. rank
	Witness	Level 1	Level 2	Level 3		top 5%
412	0	10	7	0	515	1.00
18	6	22	1	0	638	1.00
434	0	6	2	0	514	1.00
911	0	7	0	0	515	1.00
251	0	7	1	1	514	0.99
479	0	10	1	0	514	0.99
478	0	4	2	0	514	0.97
746	2	0	5	0	555	0.97
1237	5	11	7	0	626	0.94

Summary

- Conditional likelihood solves the long-standing problem of confounding by assignment
- Currently working with Seattle PD to incorporate in their Early Warning System
- Interesting combination of policing, statistics, mathematics, and computer science
  - Anderson's ordinal stereotype model, conditional likelihood, Poisson-Multinomial, Heap's algorithm, discrete Fourier transform, Schur complement, Markov Chain Monte Carlo, parallel computing

Which police officers have a high propensity to escalate force?

Greg Ridgeway

Rebecca W. Bushnell Professor of Criminology
Professor of Statistics and Data Science
University of Pennsylvania

September 20, 2024