NO TRUMP!: A statistical exercise in priming

Jonathan Falk and Andrew Gelman

NERA Economic Consulting and Columbia University

ICML workshop on data-efficient machine learning
24 June 2016


Results: The following table gives the summary results from these three tournaments:[6]

	Vanderbilt 2015	Vanderbilt 1999	Dutch 2015
Hands Played	781	205	127
Percentage Played in No Trump E(NT)	28.81%	25.98%	25.98%
Percentage of "Made" No Trump Hands E(NTxM)	19.97%	12.60%	18.63%
Conditional Success Rate E(M|NT)	69.32%	48.50%	71.71%

[6] We plan to release the dataset after we have published this research in five different journals (B. Frey, passim). In the meantime, we are not concerned with any data transcription errors since they would simply lead to attenuation bias which would make our results even stronger (Reinhart and Rogoff, 2010).

"As is often the case, priming effects are subtle, unexpected, and newsworthy, while at the same time being perfectly coherent with theory."

The Fluctuating Female Vote: Politics, Religion, and the Ovulatory Cycle

Psychological Science
XX(X) 1–10

Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797612466416
pss.sagepub.com
SAGE

Kristina M. Durante[1], Ashley Rae[1], and Vladas Griskevicius[2]
[1]College of Business, University of Texas, San Antonio, and [2]Carlson School of Management, University of Minnesota

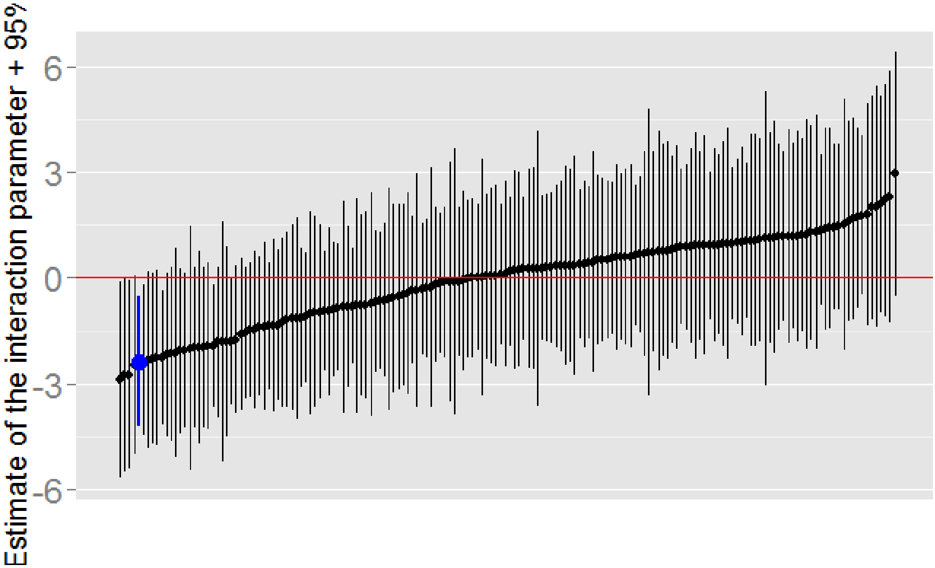
Abstract
Each month, many women experience an ovulatory cycle that regulates fertility. Although research has found that this cycle influences women's mating preferences, we proposed that it might also change women's political and religious views. Building on theory suggesting that political and religious orientation are linked to reproductive goals, we tested how fertility influenced women's politics, religiosity, and voting in the 2012 U.S. presidential election. In two studies with large and diverse samples, ovulation had drastically different effects on single women and women in committed relationships. Ovulation led single women to become more liberal, less religious, and more likely to vote for Barack Obama. In contrast, ovulation led women in committed relationships to become more conservative, more religious, and more likely to vote for Mitt Romney. In addition, ovulation-induced changes in political orientation mediated women's voting behavior. Overall, the ovulatory cycle not only influences women's politics but also appears to do so differently for single women than for women in relationships.

Choices!

1. Exclusion criteria based on cycle length (3 options)
2. Exclusion criteria based on "How sure are you?" response (2)
3. Cycle day assessment (3)
4. Fertility assessment (4)
5. Relationship status assessment (3)

168 possibilities (after excluding some contradictory combinations)

Living in the multiverse



Policy!

Labor Market Returns to Early Childhood Stimulation: a 20-year Followup to an Experimental Intervention in Jamaica

Paul Gertler, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M. Chang, Sally Grantham-McGregor

We find large effects on the earnings of participants from a randomized intervention that gave psychosocial stimulation to stunted Jamaican toddlers living in poverty. The intervention consisted of one-hour weekly visits from community Jamaican health workers over a 2-year period that taught parenting skills and encouraged mothers to interact and play with their children in ways that would develop their children's cognitive and personality skills. We re-interviewed the study participants 20 years after the intervention. Stimulation increased the average earnings of participants by 42 percent. Treatment group earnings caught up to the earnings of a matched non-stunted comparison group. These findings show that psychosocial stimulation early in childhood in disadvantaged settings can have substantial effects on labor market outcomes and reduce later life inequality.

Political science!

- Monthly cycle and voting
- Fat arms and political attitudes
- Subliminal smiley faces
- College football
- Shark attacks

- What if it were all true??

Toward Routine Use of Informative Priors

Andrew Gelman

Department of Statistics and Department of Political Science
Columbia University, New York

ICML workshop on data-efficient machine learning
24 June 2016

Bad Bayes

- Model:
  - $y|\theta \sim N(\theta, 1)$
  - $p(\theta) \propto 1$
- Data:
- Inference:
- Wanna bet??

Bad Bayes

- Model:
  - $y|\theta \sim \mathrm{N}(\theta, 1)$
  - $p(\theta) \propto 1$
- Data:
- Inference:
- Wanna bet??

Bad Bayes

- Model:
  - $y|\theta \sim \mathrm{N}(\theta, 1)$
  - $p(\theta) \propto 1$
- Data:
  - $y = 1$
- Inference:
- Wanna bet??

Bad Bayes

- Model:
  - $y|\theta \sim \mathrm{N}(\theta, 1)$
  - $p(\theta) \propto 1$
- Data:
  - $y = 1$
- Inference:
- Wanna bet??

Bad Bayes

- Model:
  - $y|\theta \sim \mathrm{N}(\theta, 1)$
  - $p(\theta) \propto 1$
- Data:
  - $y = 1$
- Inference:
  - $\theta|y \sim \mathrm{N}(y, 1)$
  - $\Pr(\theta > 0|y) = .84$
- Wanna bet??

Bad Bayes

- Model:
  - $y|\theta \sim \mathrm{N}(\theta, 1)$
  - $p(\theta) \propto 1$
- Data:
  - $y = 1$
- Inference:
  - $\theta|y \sim \mathrm{N}(y, 1)$
  - $\Pr(\theta > 0|y) = .84$
- Wanna bet??

Bad Bayes

- Model:
  - $y|\theta \sim \mathrm{N}(\theta, 1)$
  - $p(\theta) \propto 1$
- Data:
  - $y = 1$
- Inference:
  - $\theta|y \sim \mathrm{N}(y, 1)$
  - $\Pr(\theta > 0|y) = .84$
- Wanna bet??

Bad Bayes

- Model:
  - $y|\theta \sim \mathrm{N}(\theta, 1)$
  - $p(\theta) \propto 1$
- Data:
  - $y = 1$
- Inference:
  - $\theta|y \sim \mathrm{N}(y, 1)$
  - $\Pr(\theta > 0|y) = .84$
- Wanna bet??

Bad Bayes

- Model:
  - $y|\theta \sim \mathrm{N}(\theta, 1)$
  - $p(\theta) \propto 1$
- Data:
  - $y = 1$
- Inference:
  - $\theta|y \sim \mathrm{N}(y, 1)$
  - $\Pr(\theta > 0|y) = .84$
- Wanna bet??

Bad Bayes

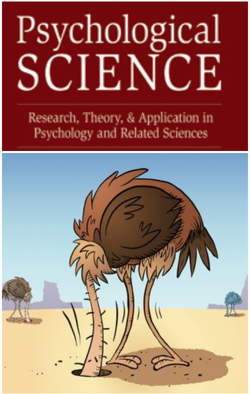
- Model:
  - $y|\theta \sim \mathrm{N}(\theta, 1)$
  - $p(\theta) \propto 1$
- Data:
  - $y = 1$
- Inference:
  - $\theta|y \sim \mathrm{N}(y, 1)$
  - $\Pr(\theta > 0|y) = .84$
- Wanna bet??

What is the problem we are trying to solve?

- Routine use of informative priors

Why is it important?

- Existing unregularized estimates are too noisy
- Problems with big data
- Replication crisis in science



How is it solved today, and what are the limitations of current solutions?

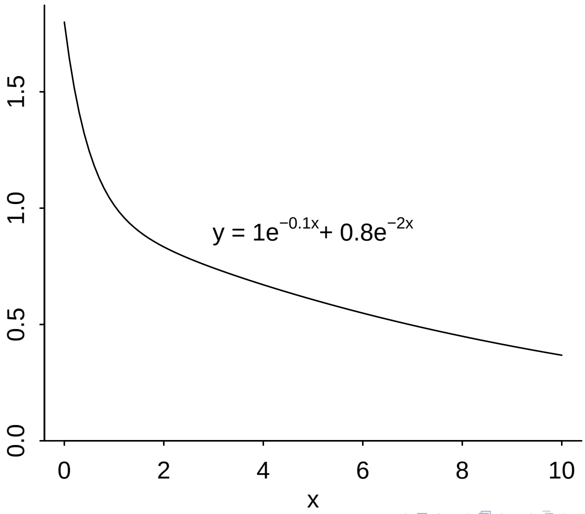
- Ignore the problem
  - Limitation: we're buried in noise
- Elicit priors from experts
  - Limitation: experts can't assess uncertainty, and they can be biased
  - "Judge, jury, and executioner" problem
- Hierarchical modeling, lasso, etc.
  - Limitation: what do you do when you have a single parameter?

How will we solve the problem?

- Embedding in a hierarchical model
- Getting away from the idea that a no-prior estimate generally exists!
- Weakly informative priors
- A small amount of information can do a lot of regularization

A simple but hard inference problem

- Sum of declining exponentials: $y = a_1 e^{-b_1 x} + a_2 e^{-b_2 x}$
- Statistical version: $y_i = \left(a_1 e^{-b_1 x_i} + a_2 e^{-b_2 x_i}\right) \cdot \epsilon_i$



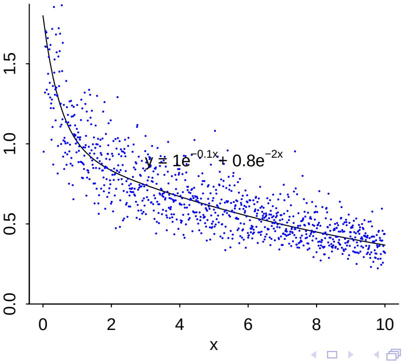
Stan code

```
data {
  int N;
  vector[N] x;
  vector[N] y;}
parameters {
  vector[2] log_a;
  ordered[2] log_b;
  real<lower=0> sigma;}
transformed parameters {
  vector<lower=0>[2] a;
  vector<lower=0>[2] b;
  a <- exp(log_a);
  b <- exp(log_b);}
model {
  vector[N] ypred;
  ypred <- a[1]*exp(-b[1]*x) + a[2]*exp(-b[2]*x);
  y ~ lognormal(log(ypred), sigma);
}
```

Simulate fake data in R

```
a <- c(1, 0.8)
b <- c(0.1, 2)
sigma <- 0.2

x <- (1:1000)/100
N <- length(x)
ypred <- a[1]*exp(-b[1]*x) + a[2]*exp(-b[2]*x)
y <- ypred*exp(rnorm(N, 0, sigma))
```



Fit the model in Stan

```
Inference for Stan model: exponentials.
4 chains, each with iter=1000; warmup=500; thin=1;
post-warmup draws per chain=500, total post-warmup draws=2000.

      mean se_mean   sd  25%  50%  75% n_eff Rhat
a[1]  1.00    0.00 0.03 0.99 1.00 1.02   494    1
a[2]  0.70    0.00 0.08 0.65 0.69 0.75   620    1
b[1]  0.10    0.00 0.00 0.10 0.10 0.10   532    1
b[2]  1.71    0.02 0.34 1.48 1.67 1.90   498    1
sigma 0.19    0.00 0.00 0.19 0.19 0.20   952    1
```

- Compare to true values:

```
a <- c(1, 0.8)
b <- c(0.1, 2)
sigma <- .2
```

Let's make the problem harder

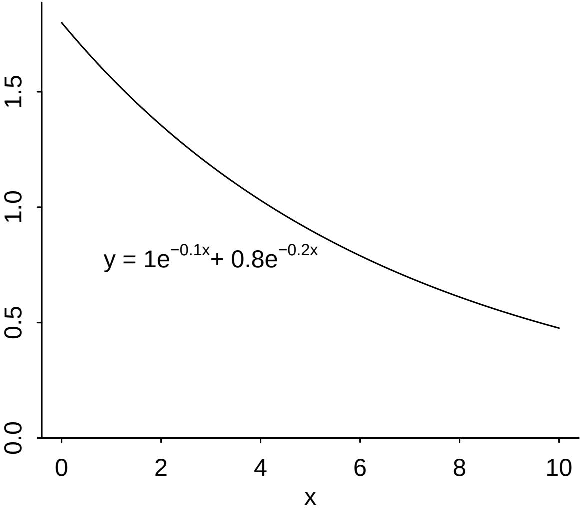
- Simulate new data using these new parameter values:

```
a <- c(1, 0.8)
b <- c(0.1, 0.2)
```

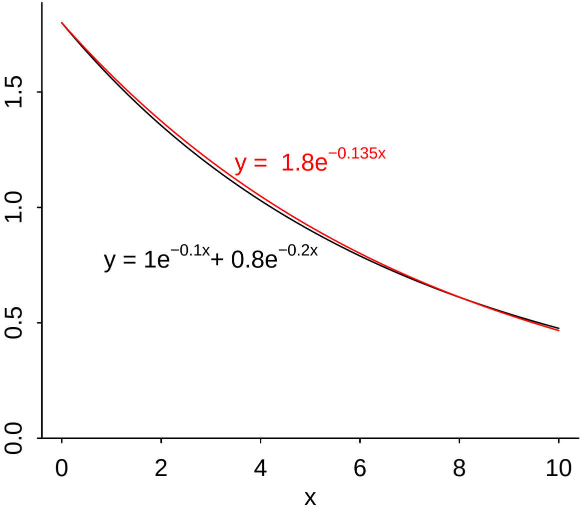
- Then fit the model:

```
           mean se_mean   sd  25%       50%       75% n_eff Rhat
a[1]   1.33e+00    0.54 0.77 1.28  1.77e+00  1.79e+00     2 44.2
a[2]  2.46e+294     Inf  Inf 0.00  0.00e+00  1.77e+00  2000  NaN
b[1]   1.00e-01    0.04 0.06 0.10  1.30e-01  1.30e-01     2 33.6
b[2]  3.09e+305     Inf  Inf 0.50 1.15e+109 4.77e+212  2000  NaN
sigma  2.00e-01    0.00 0.00 0.19  2.00e-01  2.00e-01    65  1.0
```

What went wrong?



What went wrong?



Informative prior distribution

```
log_a ~ normal(0, 1);
log_b ~ normal(0, 1);
```

Happy ending

	mean	se_mean	sd	25%	50%	75%	n_eff	Rhat
a[1]	1.56	0.09	0.32	1.52	1.72	1.75	13	1.25
a[2]	0.32	0.08	0.28	0.14	0.22	0.37	13	1.20
b[1]	0.13	0.00	0.01	0.12	0.13	0.13	22	1.14
b[2]	1.94	0.20	2.29	0.22	1.26	3.00	127	1.05
sigma	0.20	0.00	0.00	0.19	0.20	0.20	656	1.00

- Compare to true values:

```
a <- c(1, 0.8)
b <- c(0.1, 0.2)
sigma <- .2
```

Areas of application

- Ill-posed problems ($y = a_1 e^{-b_1 x} + a_2 e^{-b_2 x}$, differential equation models in pharmacometrics)
- Weakly-informative priors for logistic regression
- Fitting models with less "hand-holding"
- A new way to think about junk science

Summary

- Get rid of the idea that classical methods (including noninformative Bayes) are "safe" or "conservative"
  - Examples: 1 ± 1, separation in logistic regression, junk science, ... good science
  - The statistical significance filter is real.
- Prior information is all around
- Resolve the GIGO problem by embedding in a hierarchical model
- Avoid the no-pooling or complete-pooling choice
- Fit in Stan!