Estimating the Probability of Compliance with Physical Activity Guidelines from a Bayesian Perspective

Bryan Stanfill, Dave Osthus, Alicia Carriquiry and Sarah Nusser
Department of Statistics, Iowa State University
ICDAM 8, May 14 - 17, 2012

Introduction

Currently, the Center for Disease Control (CDC) in the United States issues guidelines with recommended levels of physical activity for children, adults, and older adults. In light of the current obesity levels in America, it is hypothesized that compliance with these guidelines is low. The objective of this work is to estimate the proportion of individuals in a population with usual (or long-run average) physical activity levels that meet the CDC guidelines. We use data from the Physical Activity Measurement Survey (PAMS) project, and adopt Bayesian methods to model physical activity. Our approach integrates the complex design of the PAMS and classical Bayesian methods to extend the results beyond the sample.

USA's Physical Activity Guidelines

The current U.S. Physical Activity guidelines can be met in three ways:

1. 150 minutes of moderate-intensity aerobic activity (3-6 METs) per week.
2. 75 minutes of vigorous-intensity aerobic activity (over 6 METs) per week.
3. An equivalent mix of moderate- and vigorous-intensity aerobic activity per week.

Key: These minutes of physical activity of specified intensity levels must come in at least 10 continuous minute intervals, known as bouts. For the purposes of this project we consider an individual compliant with the CDC guidelines if on 5 or more days a week that individual engages in 90 MET-minutes of physical activity at an intensity of at least 3 METs observed during bouts of at least 10 minutes.

Data Collection and Description

In the PAMS study, physical activity was collected from a representative sample of over 1400 Iowa adults ages 18-70. Four counties were selected then split into high and low minority tracts. Within each tract a random sample of houses was chosen. Within each household an eligible adult was chosen at random. Each participant followed the following procedure.

- Selected individuals wore SenseWear Monitor for 24 hours on two non-consecutive days.
- Minute by minute energy expenditure information is recorded.
- Data collection was evenly distributed over two years, partitioned into eight quarters.

Covariates were also gathered for each individual. Age and BMI distributions for the data set we consider are summarized below.

Variable	Statistic	Females n=269	Males n=201
Age	Q1	45.00	37.00
	Median	54.00	48.00
	Q3	62.00	60.00
BMI	Q1	25.76	25.91
	Median	29.62	28.89
	Q3	36.02	32.95

Modeling Goals and Considerations

Modeling goals:
Our goal is to develop a statistical model that can plausibly "produce" the observed data while at the same time, aides us in answer three primary questions. They are:

1. What proporting of Iowans engage in some PA?
2. What proportion of Iowans comply with the CDC PA guidelines?
3. What covariates are associated with compliance?

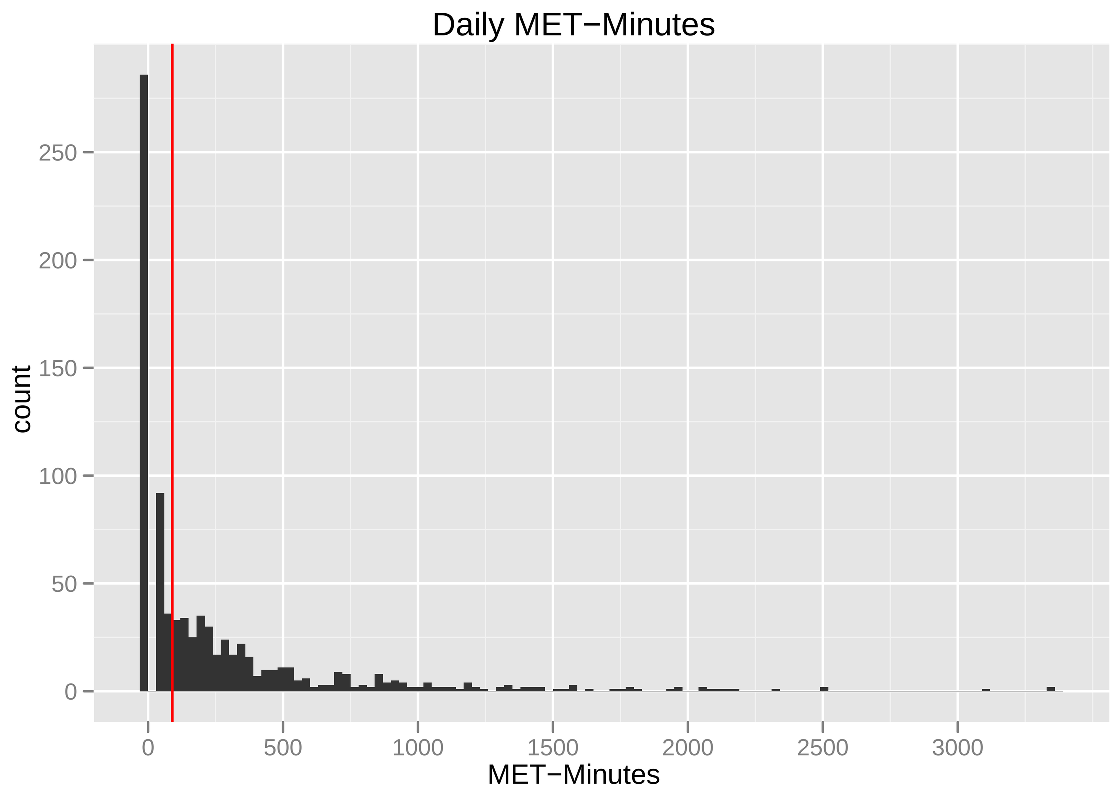

Figure 1-A histogram of our dataset.

Modeling considerations:

1. MET-minutes are constrained to a portion of the positive real-line
2. Spike at 0 MET-minutes
3. Long tail / right-skewed

We plan to accomplish these goals through the following mixture model. Let Y_{kij} be a random variable associated with the MET-minutes of individual i on day j in stratum k. We will define Y_{kij} as follows:

$$Y_{kij} = (1 - V_{kij}) + V_{kij}W_{kij}$$

where

$$V_{kij} \sim \text{Bernoulli}(p_{ki}) \quad \text{and} \quad W_{kij} \sim \text{Shifted Gamma}(\mu_{ki}, \phi_k).$$

Further,

$$\text{logit}(p_{ki}) = \mathbf{x}_{1ki}^T \boldsymbol{\beta}_{1k} + \gamma_{1ki} \quad \text{and} \quad log(\mu_{ki}) = \mathbf{x}_{2ki}^T \boldsymbol{\beta}_{2k} + \gamma_{2ki};$$

$\boldsymbol{x}_{1ki}$ and $\boldsymbol{x}_{2ki}$ are matrices of potentially different covariates, γ_{1ki} and γ_{2ki} are random person effects and $\boldsymbol{\beta}_{1k}$ and $\boldsymbol{\beta}_{2k}$ are vectors of unknown parameters to be estimated. Here we use age and BMI as covariates in both the bernoulli and gamma distributions.

Results

Figure 2 is a plot of the estimated mean probability of a bout as a function of gender and BMI. Figure 3 displays the estimated probability of complying with CDC guidelines if survey design is ignored (in red) or used (in blue).

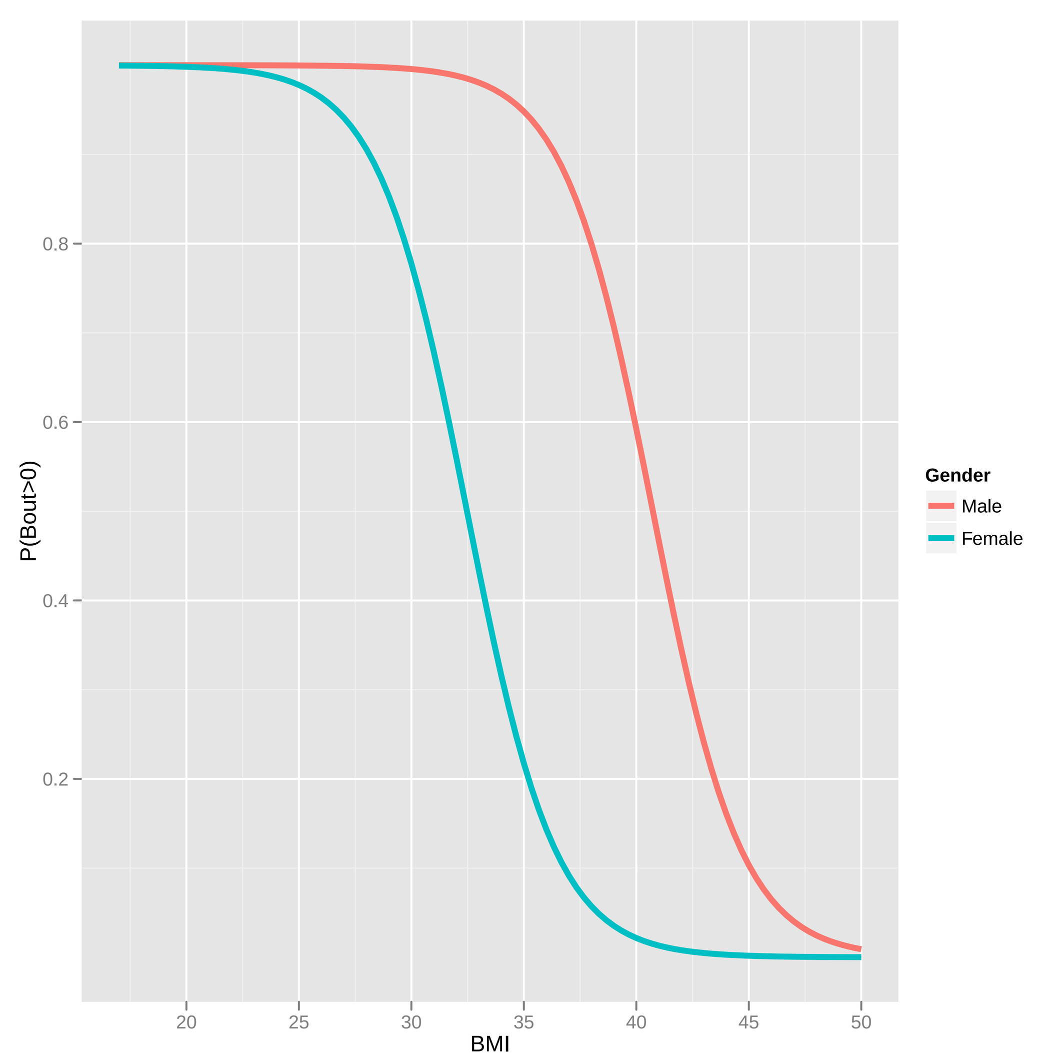

Figure 2-P(Bout) as a function of BMI and Gender

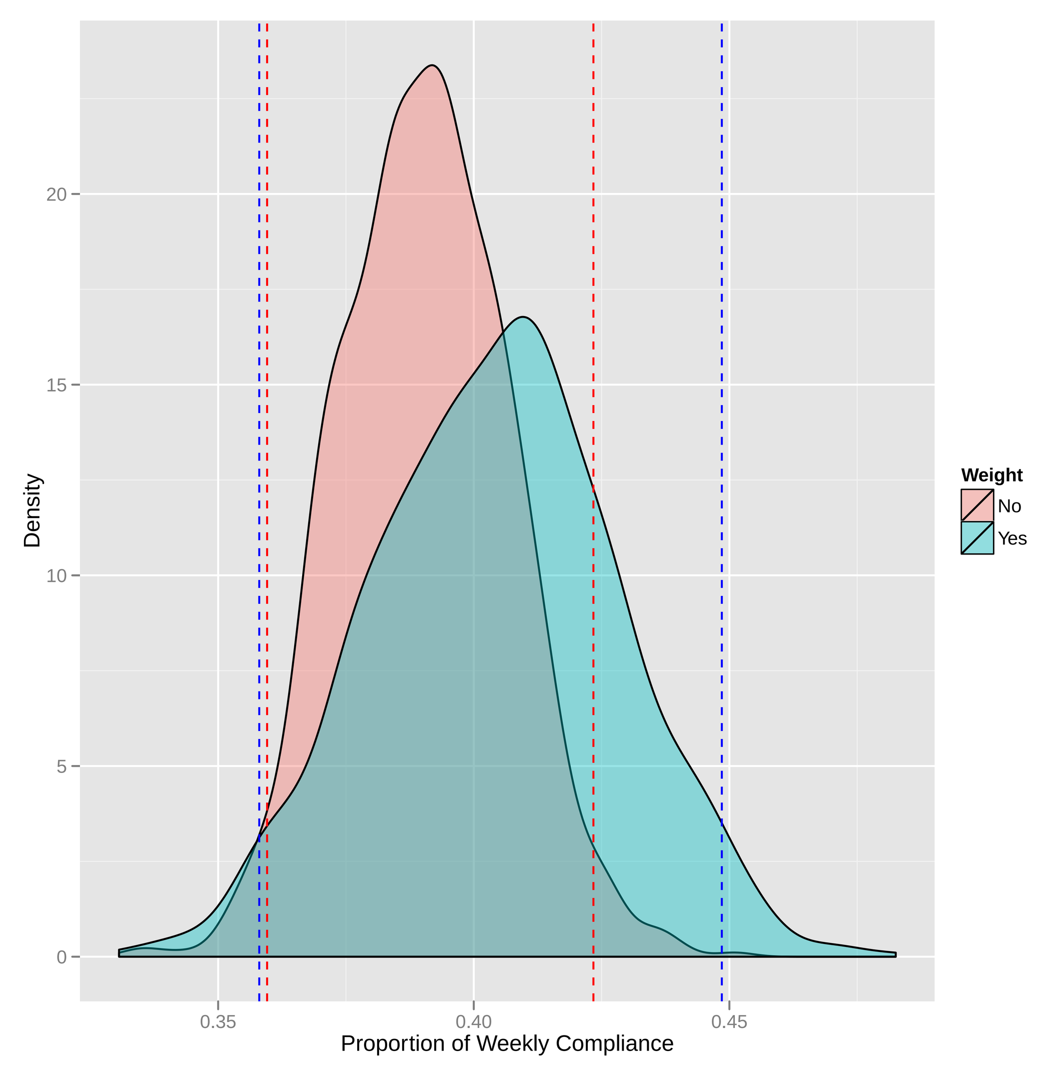

Figure 3-P(Compliance) for Iowans

- Parameter estimates for logistic regression: $\hat{\beta}_{BMI} = -0.281\ (0.085)$, $\hat{\beta}_{Male} = 2.34\ (1.07)$
- For a given gender, the odds an individual engages in a bout decrease by 75% with every unit increase in BMI
- At a given BMI, a male has odds 10.34 times greater than that of a female of engaging in a bout
- Results for shifted gamma: $\hat{\beta}_{BMI} = -0.089\ (0.066)$, $\hat{\beta}_{Male} = 0.86\ (0.44)$

Funding and Thanks

The PAMS study was funded by a NIH grant awarded to Dr. Greg Welk, Dr. Sarah Nusser and Dr. Alicia Carriquiry (R01 HL91024-01A1) titled "*A Measurement Error Approach to Estimating Usual Daily Physical Activity Distributions*". A special thanks to Dr. Miguel Calabro for providing the raw data files needed for the analysis.