SOCI 620: Quantitative methods 2

Jan 17

Multi-parameter posteriors

1. Administrative
2. Modeling income with two parameters
3. MAP estimation

Administrative

Worksheet typos

- Question 2.1 should end with “evenly spaced values between 0.0 and 1.0.”
- Question 2.8 should read “variables named `prior_unopinionated` and `prior_opinionated`”
- Both are corrected at the download link, but you may just want to note the changes

Worksheet deadlines

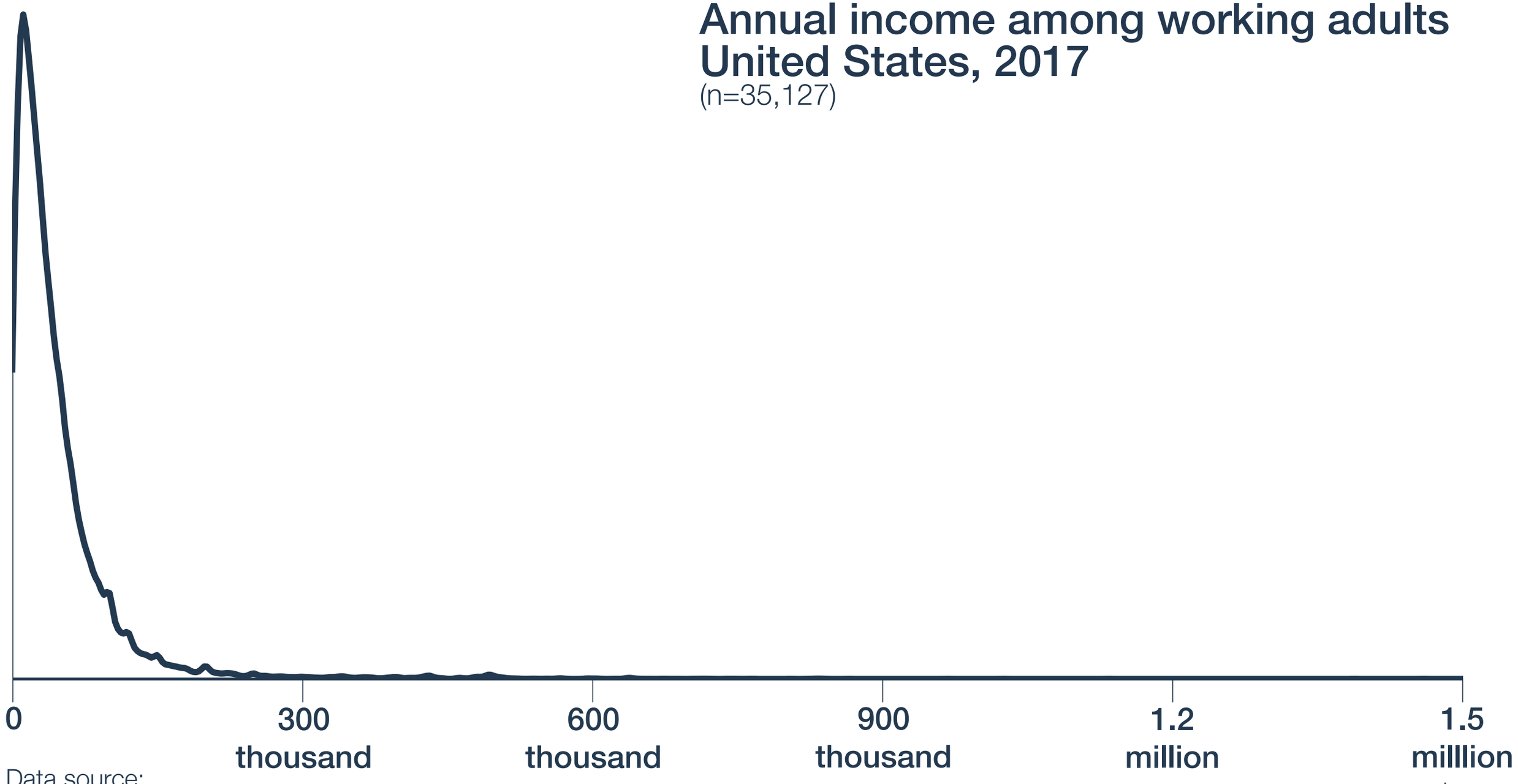
- Should we push the WS1 deadline to Jan 24 and WS2 deadline to Feb 7?

Modeling income

Modeling income

Annual income among working adults United States, 2017

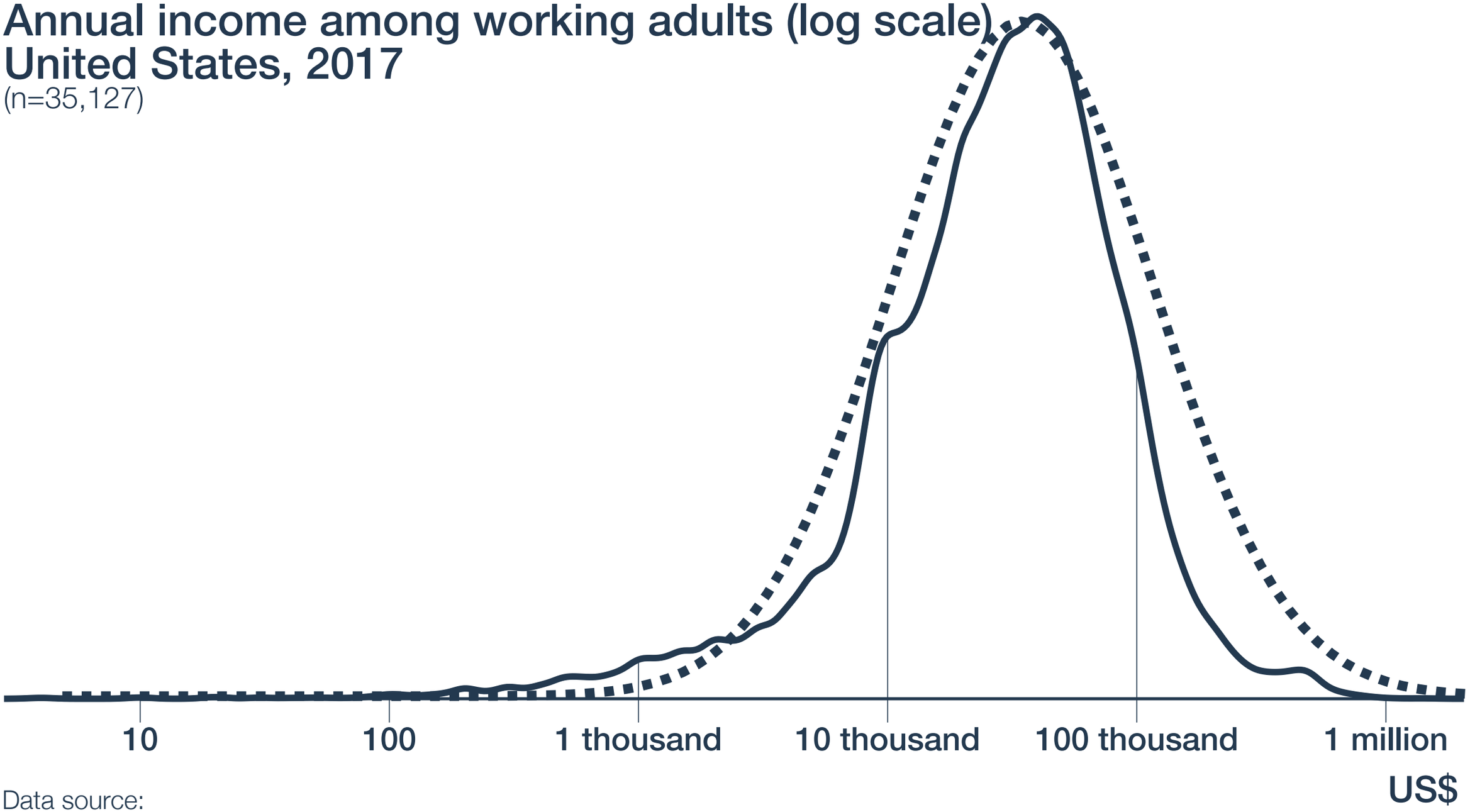
(n=35,127)



Data source:
IPUMS-USA, University of Minnesota, www.ipums.org

Modeling income

Annual income among working adults (log scale)
United States, 2017
(n=35,127)



Data source:
IPUMS-USA, University of Minnesota, www.ipums.org

Plan: model *log* income as a normally distributed random variable

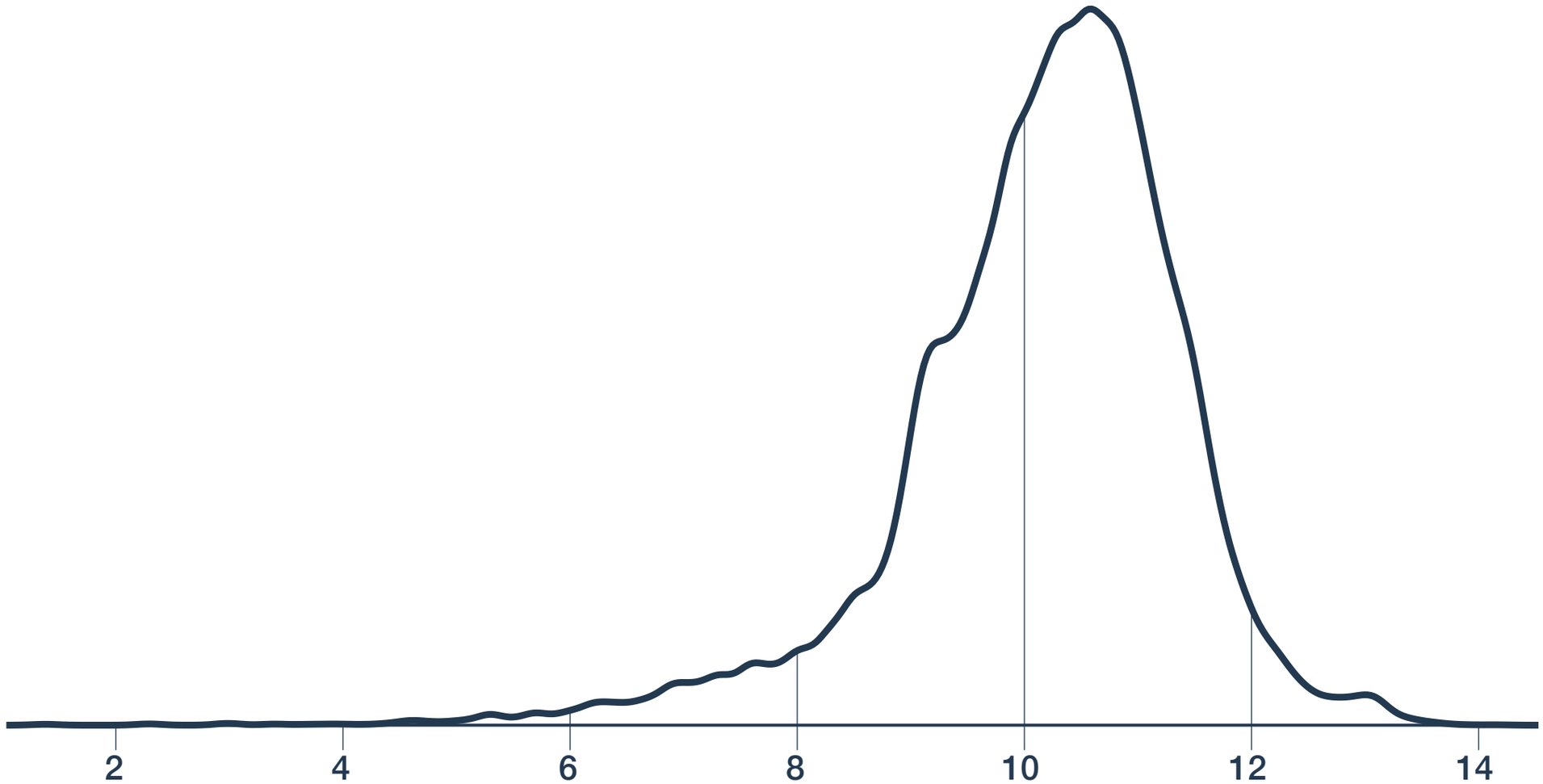
A two-parameter model

Model of normally distributed log income:

$$y_i \sim \text{Norm}(\mu, \sigma)$$

Two parameters:

μ determines location
σ determines width



A two-parameter model

Model of normally distributed log income:

$$y_i \sim \text{Norm}(\mu, \sigma)$$

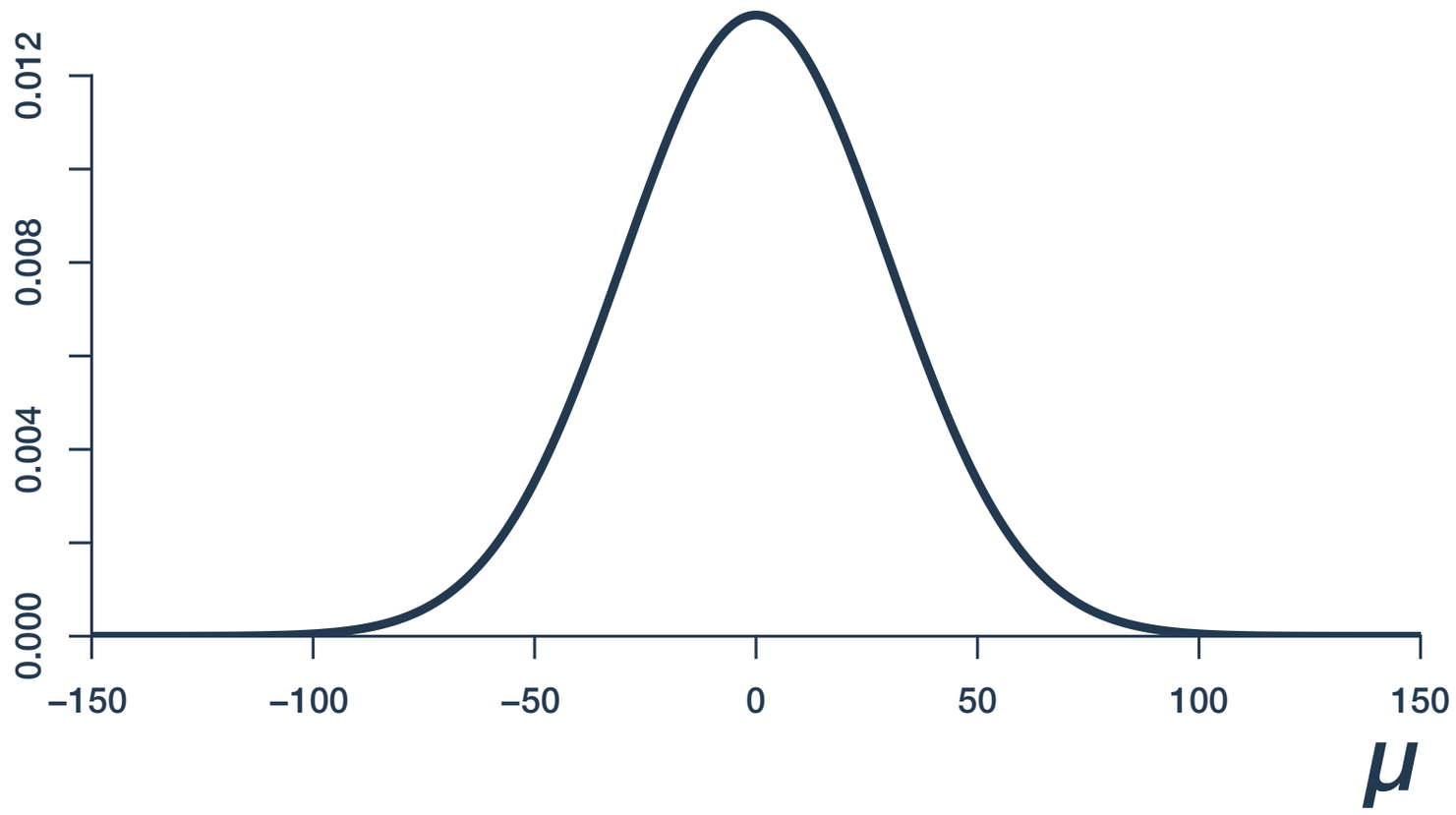
$$\mu \sim \textbf{?}$$

A two-parameter model

Model of normally distributed log income:

$$y_i \sim \text{Norm}(\mu, \sigma)$$

$$\mu \sim \text{Norm}(0, 30)$$



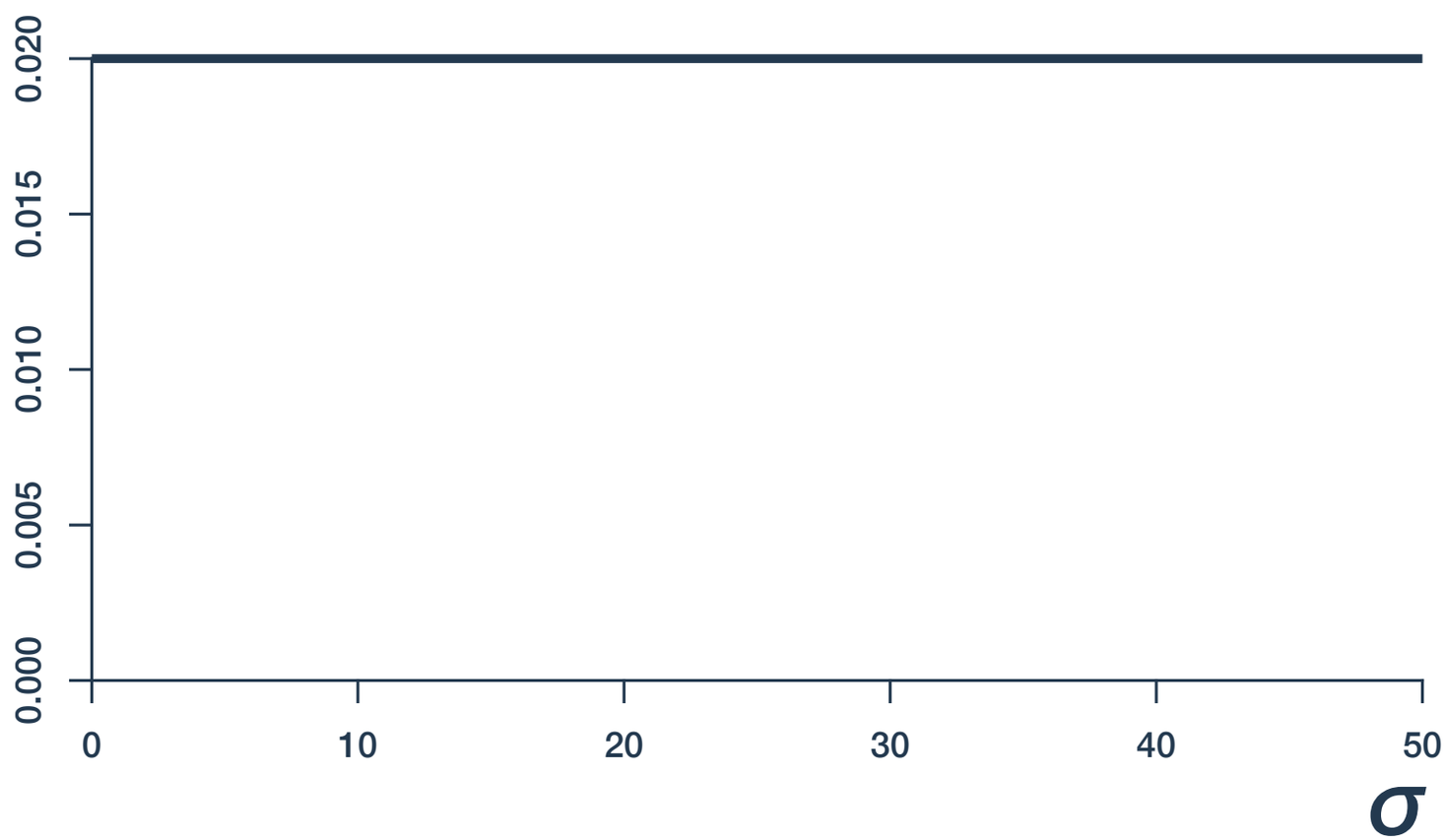
A two-parameter model

Model of normally distributed log income:

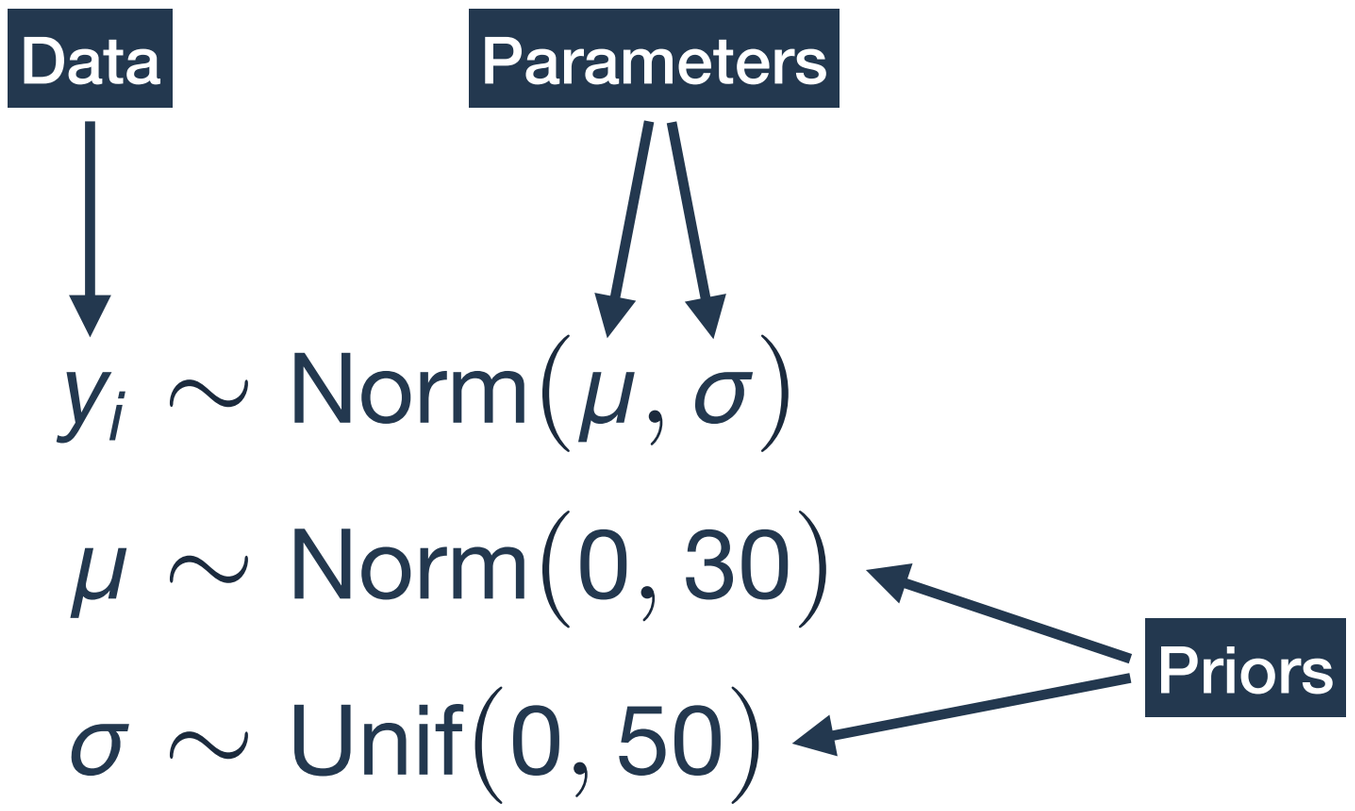
$$y_i \sim \text{Norm}(\mu, \sigma)$$

$$\mu \sim \text{Norm}(0, 30)$$

$$\sigma \sim \text{Unif}(0, 50)$$



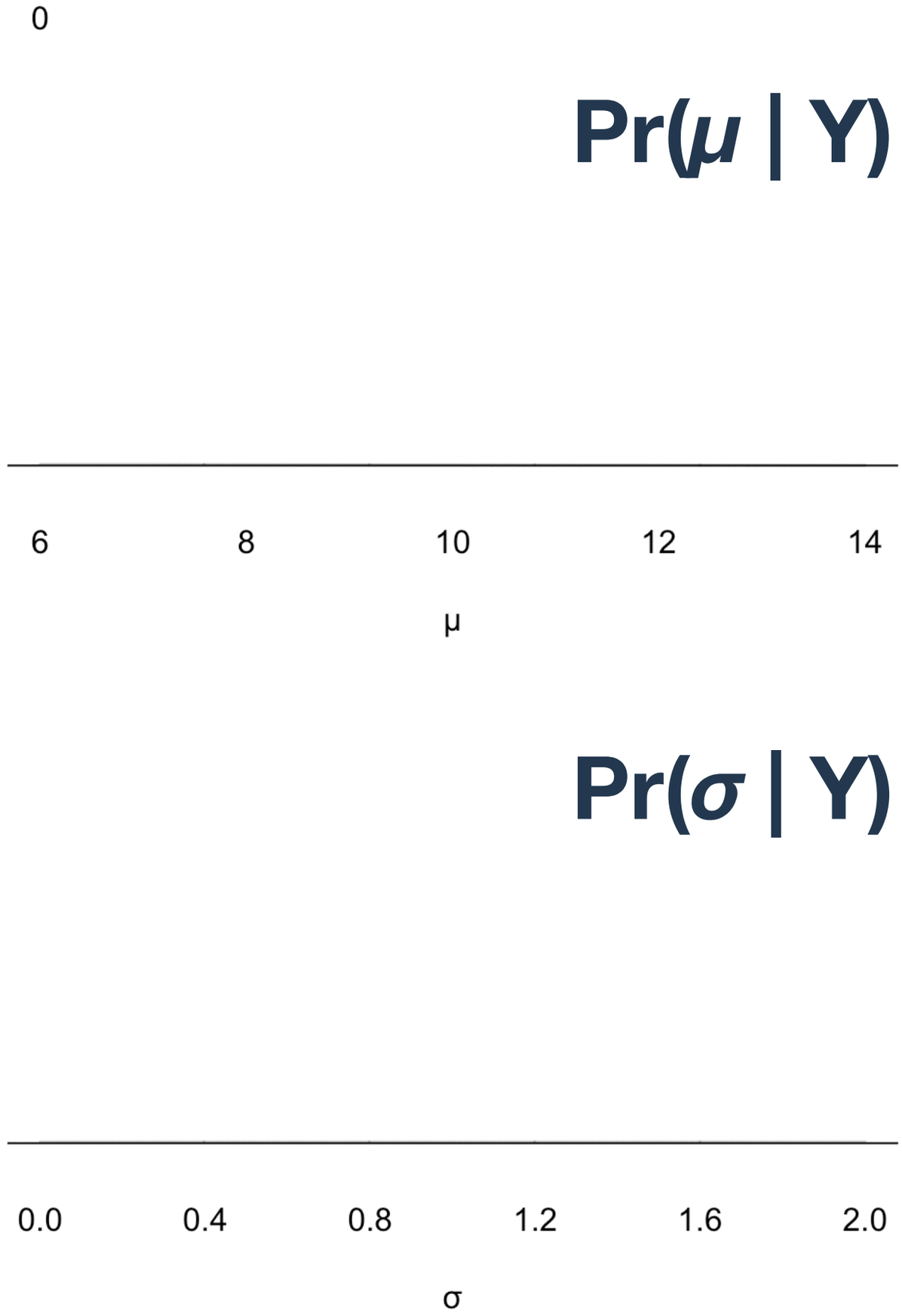
A two-parameter model



Multi-parameter posteriors

Updating marginal posterior distributions for μ and σ

First 100 samples



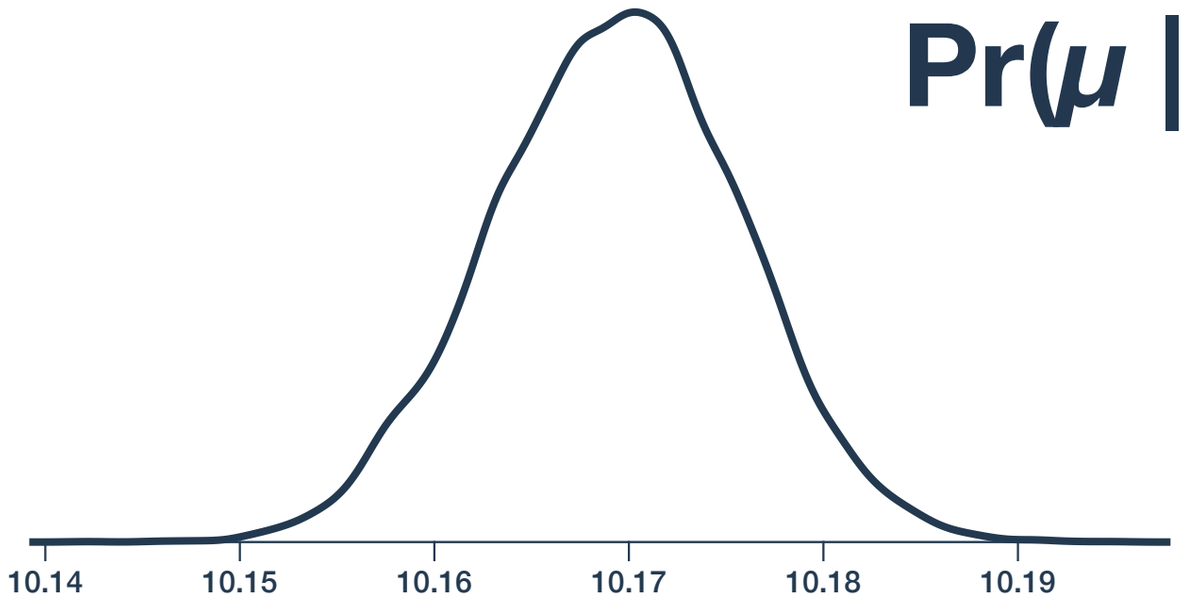
https://soci620.netlify.com/media/updates_mu_sigma.gif

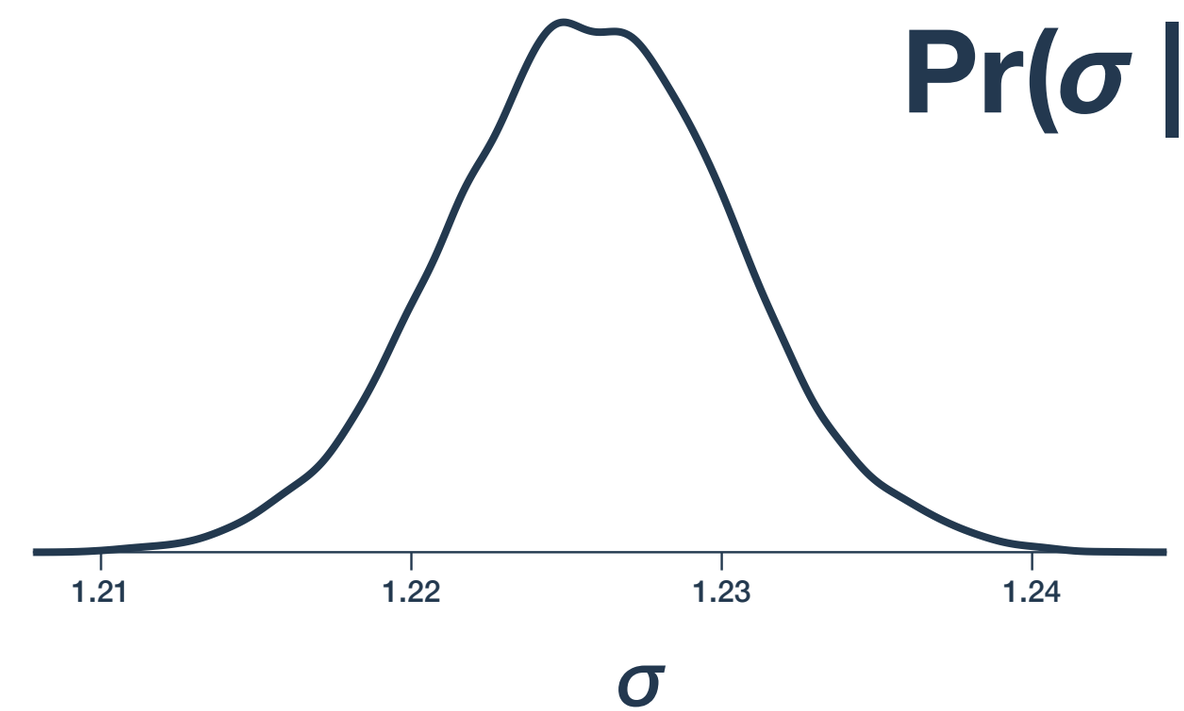
Multi-parameter posteriors

Updated marginal posterior distributions for μ and σ

Full sample (n=35,127)

- Nearly normal in shape
- Very narrow bounds
- Lumpy because built from posterior sample

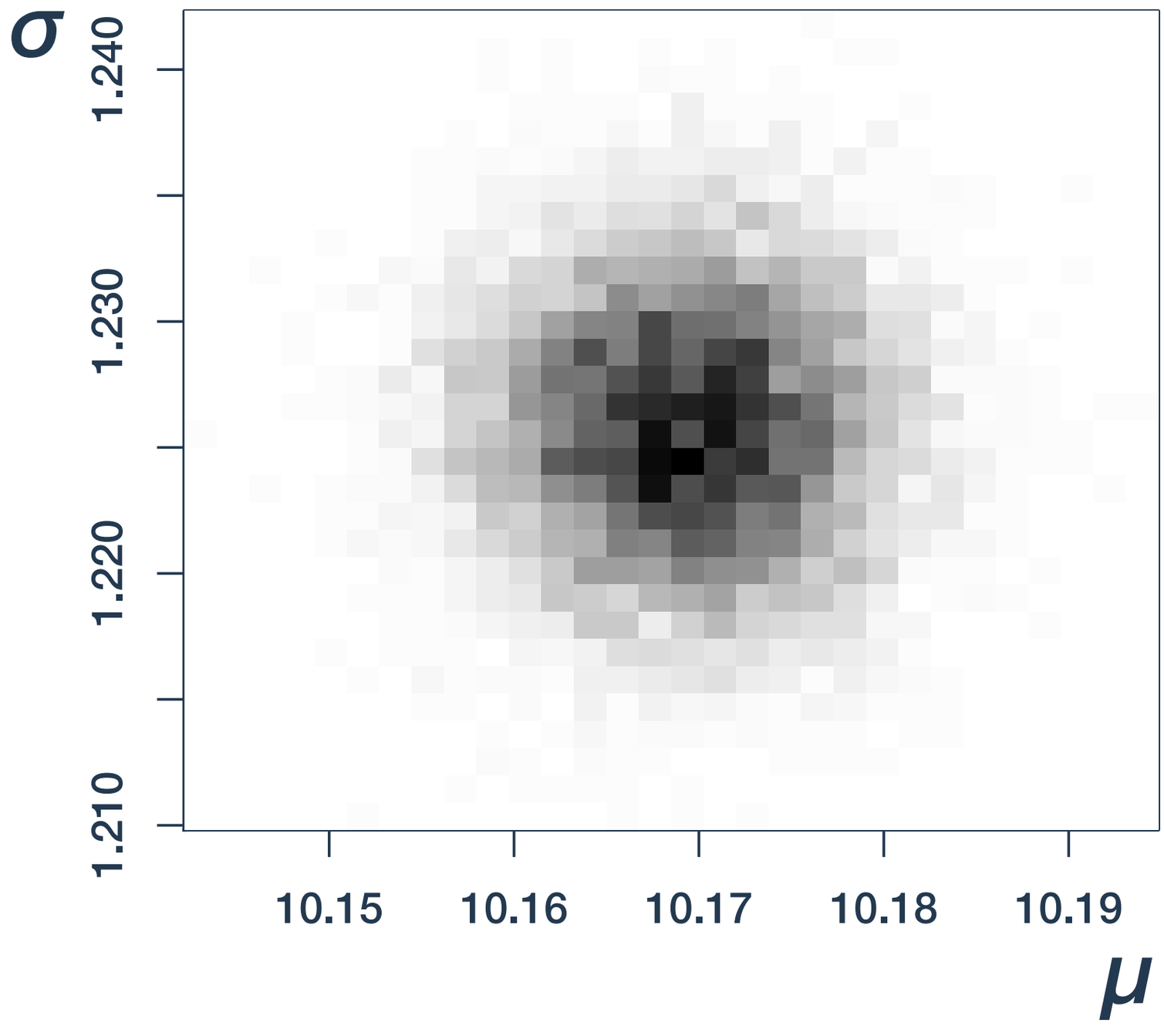




Multi-parameter posteriors

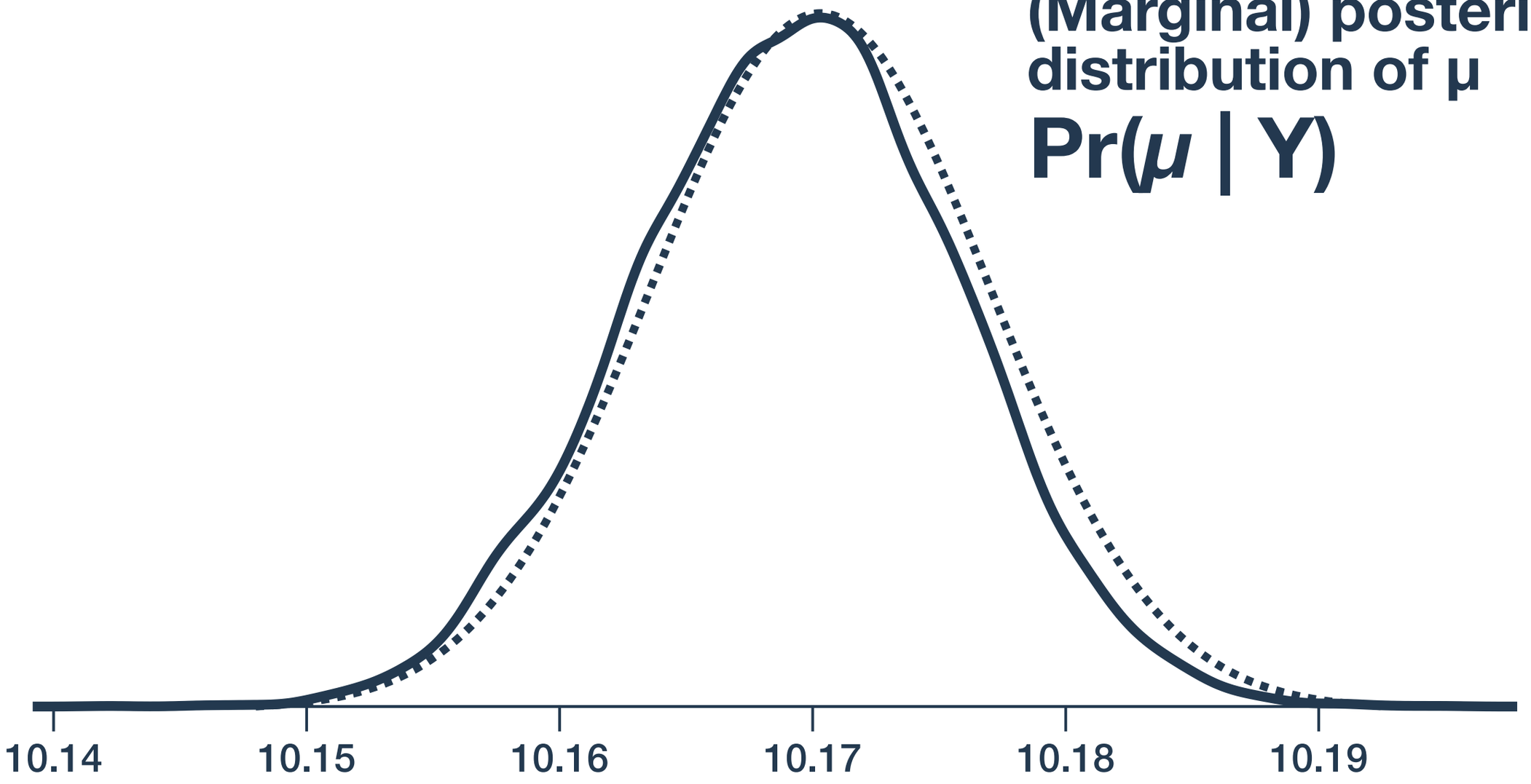
Updated joint posterior distribution for μ and σ

Full sample (n=35,127)



Maximum a posteriori estimation (MAP/QUAP)

MAP estimation



(Marginal) posterior distribution of μ

Pr(*μ* | Y)

Maximum a posteriori (quadratic approximation):

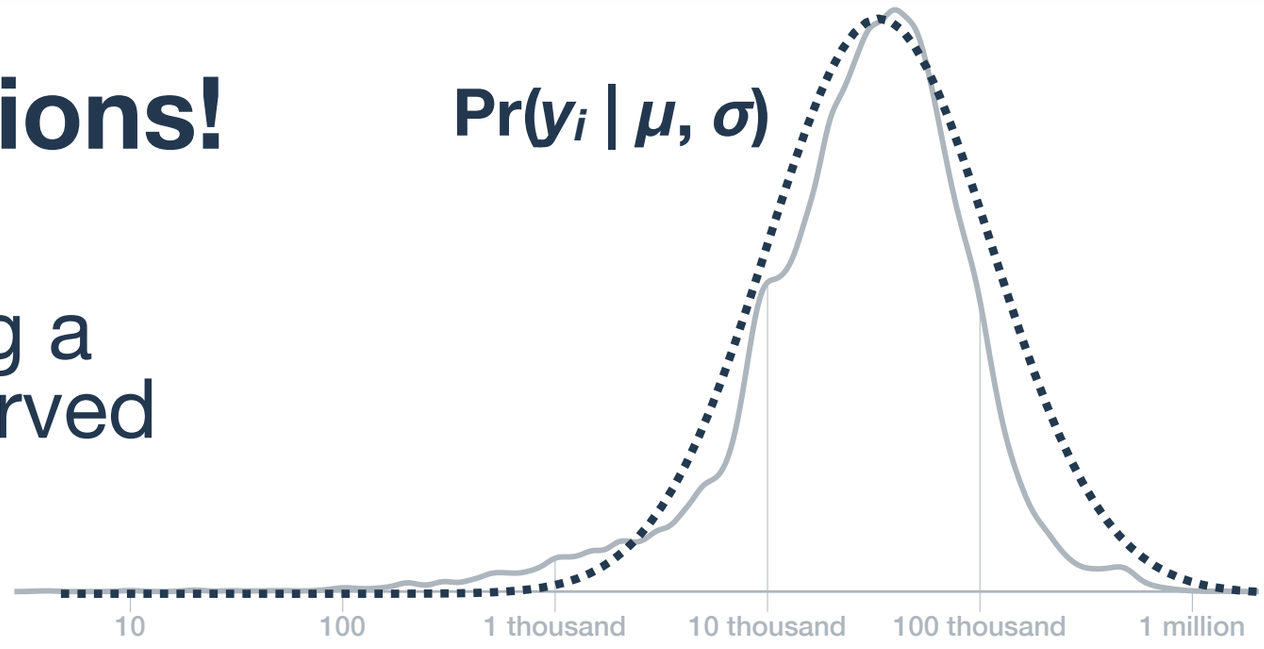
1. Find the ‘peak’ of the posterior and call it m_μ.
2. Estimate the ‘spread’ of the posterior and call it s_μ.
3. Approximate the posterior with Norm(m_μ, s_μ)

MAP estimation

So many normal distributions! This can get confusing!

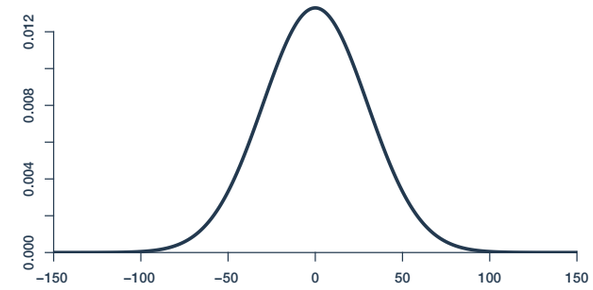
- ***First***, we modeled ***income*** using a normal distribution with unobserved parameters μ and σ:

$$y_i \sim \text{Norm}(\mu, \sigma)$$



- ***Then***, we gave μ a normal ***prior***:

$$\mu \sim \text{Norm}(0, 30)$$



- And ***then*** we used a normal distribution to approximate the ***posterior*** distributions of μ and σ:

$$\mu|Y \sim \text{Norm}(m_\mu, s_\mu)$$

$$\sigma|Y \sim \text{Norm}(m_\sigma, s_\sigma)$$