Introduction To Sampling

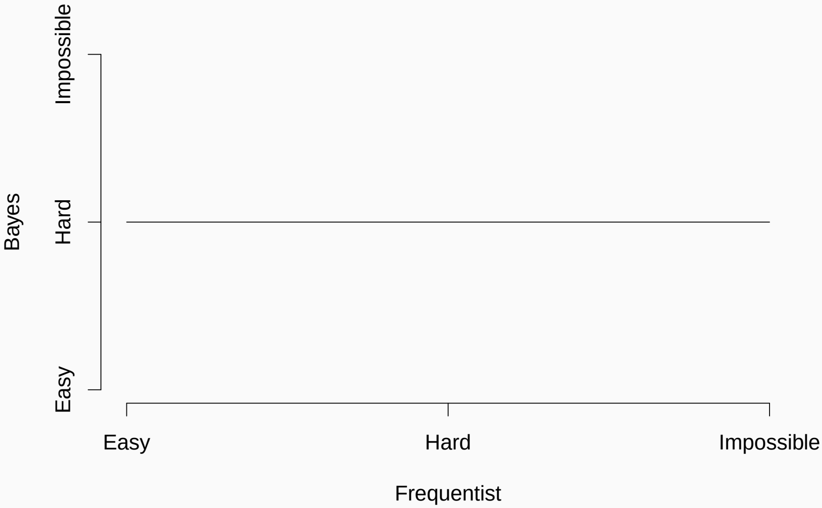
Jeff Rouder

May, 2022

Estimate the parameters in a normal

- I mean, how hard is this
- $\hat{\mu} = \frac{\sum y_i}{n}$
- $\hat{\sigma^2} = \frac{\sum (y_i - \bar{y})^2)}{n-1}$
- $\mathrm{CI} = \pm t(n-1, .975) \times \sqrt{n} \times \frac{\hat{\mu}}{\hat{\sigma}}$

Bayesian. . .



Impossible
Hard
Easy
Bayes
Easy
Hard
Impossible
Frequentist

Sampling as Knowledge

Who Wakes Up First?

Let X denote when Becky wakes up. It is minutes after 7a.

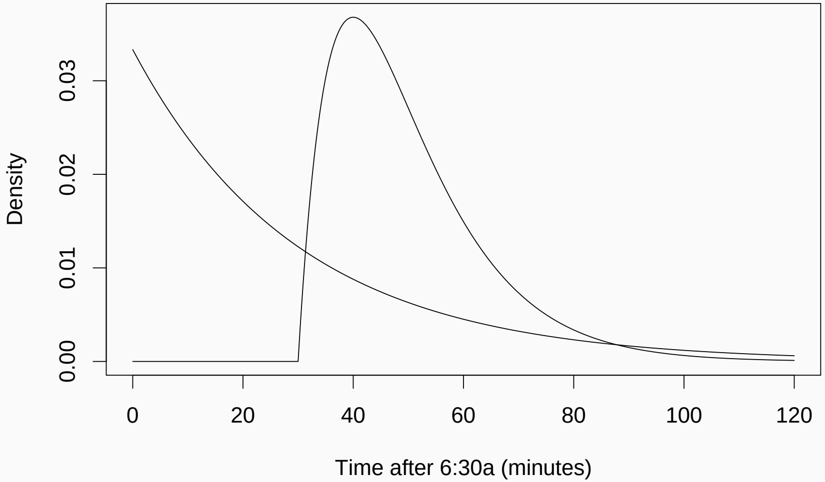
$$X \sim \text{Gamma}(2, 10)$$

Let Y denote when Jeff wakes up. It is minutes after 6:30a.

$$Y \sim \text{Gamma}(1, 30)$$

What is the probability Becky wakes up before Jeff, $Pr(X < Y)$?

Setup

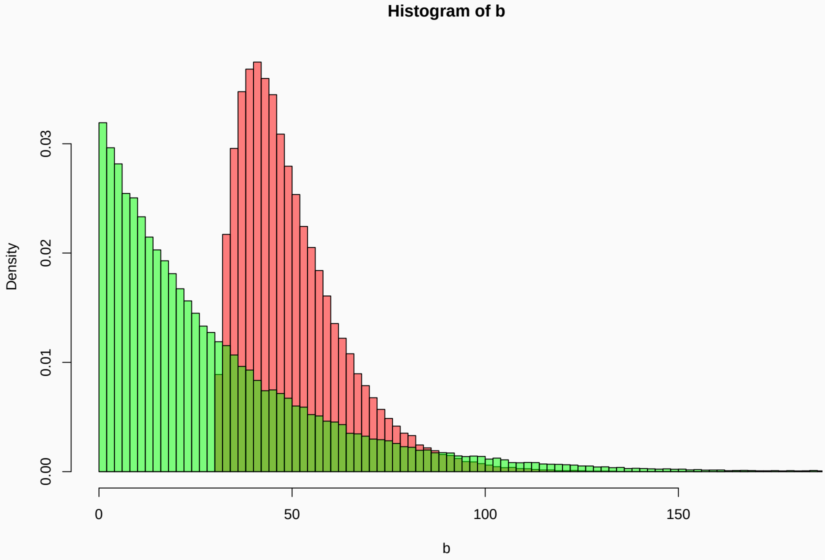


Density
0.00
0.01
0.02
0.03
0
20
40
60
80
100
120
Time after 6:30a (minutes)

Simulation Code

```
M=100000
j=rgamma(M,shape=1,scale=30)
b=rgamma(M,shape=2,scale=10)+30
hist(b,breaks=seq(0,500,2),col=rgb(1,0,0,.5),
     xlim=c(0,180),prob=T)
hist(j,add=T,col=rgb(0,1,0,.5),breaks=seq(0,500,2),prob=T)
print(mean(b<j))
```

Simulations




```
## [1] 0.2064
```

MCMC techniques

- Samples from posterior
- e.g. $(\mu, \sigma^2 | \boldsymbol{Y})$
- Samples are enough to learn

How To?

- Program Your Own Sampler To Learn About MCMC
  - fast to execute
  - you feel smart
  - insightful
- JAGS
  - fast but autocorrelation
  - great manual
  - stable
- stan
  - best samplers,
  - changing, info scattered
  - very popular (cool kids)

Model + Data

$$
\begin{aligned}
Y_i \sim & \mathrm{N}(\mu, \sigma^2) \\
\mu \sim & \mathrm{N}(a, b) \\
\sigma^2 \sim & \text{Inverse Gamma}(q, r)
\end{aligned}
$$

$Y =$ 92, 115, 100, 98

Demo, Roll My Own

$$\mu \mid \sigma^2, \boldsymbol{Y} \sim \text{Normal}(vc, v), \quad v = \left(\frac{N}{\sigma^2} + \frac{1}{b}\right)^{-1}, \quad c = \frac{N\bar{Y}}{\sigma^2} + \frac{a}{b}$$

$$\sigma^2 \mid \mu, \boldsymbol{Y} \sim \text{IG}\left(q + \frac{N}{2}, r + \frac{SSE}{2}\right)$$

]

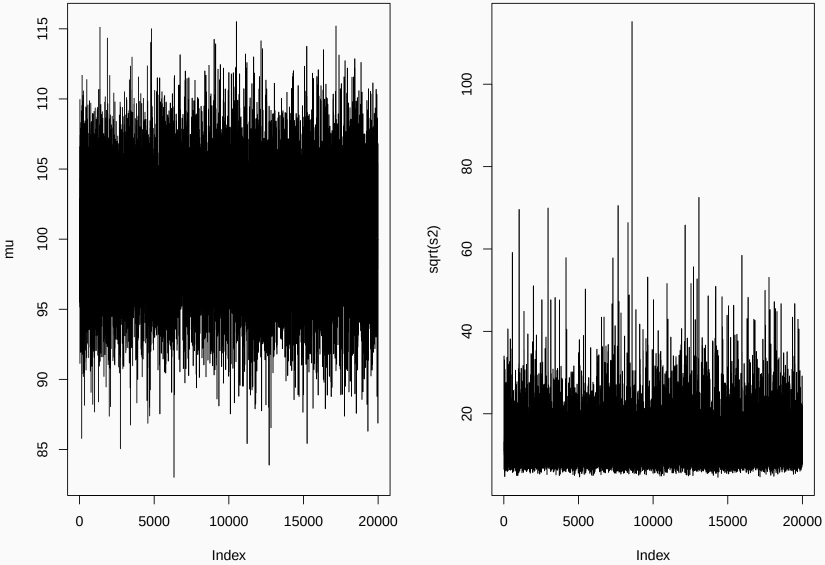
Demo, Roll My Own

```
Y=c(92,115,100,98)
ybar=mean(Y)
N=length(Y)
a=100
b=25
q=.5
r=.5*15^2
M=20000 # number of samples
mu=1:M
s2=1:M
mu[1]=100 #initial value
s2[1]=100 #initial value
```

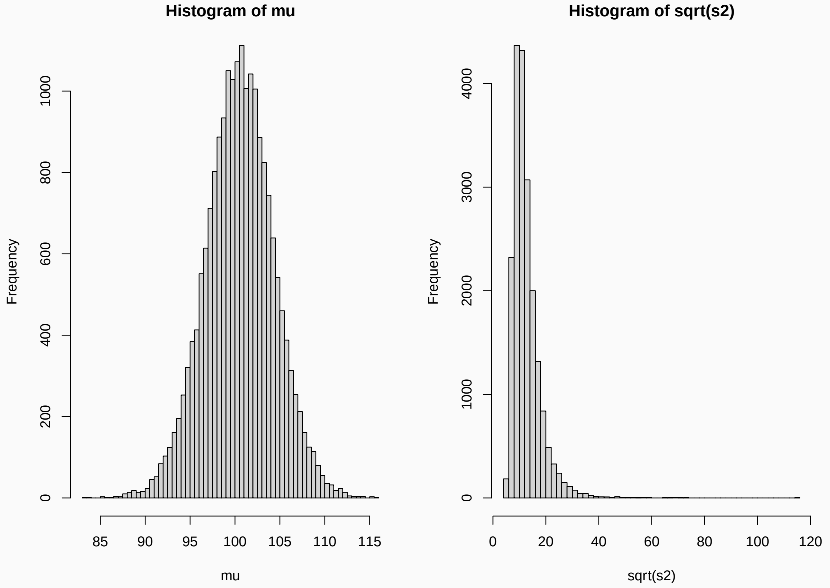
Demo, Roll My Own

```r
for (m in 2:M){
  v=1/(N/s2[m-1]+1/b)
  c=N*ybar/s2[m-1]+a/b
  mu[m]=rnorm(1,c*v,sqrt(v))
  SSE=sum((Y-mu[m])^2)
  s2[m]=rinvgamma(1,N/2+q,SSE/2+r)
  }
```

Demo Roll My Own



Demo Roll My Own



JAGS & stan

- Examples in their directories
- It is too hard to typeset here

Two-sample analysis

- Half the kids get Smarties (treatment)
- Half don't (control)
- All get an IQ test an hour later
- What was the effect of Smarties?
- You make up the data to explore the following questions
- Let Y_{ij} denote the IQ for the jth kiddo, $j = 1, \ldots, J$ in the ith group, $i = 1, 2$. ($i = 1$ for control, $i = 2$ for treatment)
- Y_1= 101, 82, 136, 77, 102, 99
- Y_2= 104, 101, 115, 88, 83, 122

Two Models, Which Is Better?

Cell Means Parameterization:

$$Y_{1j} \sim \mathrm{N}(\mu_1, \sigma^2)$$
$$Y_{2j} \sim \mathrm{N}(\mu_2, \sigma^2)$$

Same, but more compact:

$$Y_{ij} \sim \mathrm{N}(\mu_i, \sigma^2)$$

Effects Parameterization:

$$Y_{1j} \sim \mathrm{N}(\alpha - \theta/2, \sigma^2)$$
$$Y_{2j} \sim \mathrm{N}(\alpha + \theta/2, \sigma^2)$$

Same but more compact:

$$Y_{ij} \sim \mathrm{N}(\alpha + x_i\theta, \sigma^2), \quad x_i = c(-1/2, 1/2)$$

Two Models, Which Is Better?

Cell Means:

$$
\begin{aligned}
Y_{ij} &\sim \mathrm{N}(\mu_i, \sigma^2) \\
\mu_i &\sim \mathrm{N}(a, b) \\
\sigma^2 &\sim \mathrm{IG}(1/2, r/2)
\end{aligned}
$$

Effects:

$$
\begin{aligned}
Y_{ij} &\sim \mathrm{N}(\alpha + x_i\theta, \sigma^2) \\
\alpha &\sim \mathrm{N}(a_1, b_1) \\
\theta &\sim \mathrm{N}(a_2, b_2) \\
\sigma^2 &\sim \mathrm{IG}(1/2, r/2)
\end{aligned}
$$

- $a = a_1 = 100$, $b = b_1 = 5^2$, \$r=15^2