Stat651 Mini Project 1

Arthur Lui

19 September 2014

1 Binomial Data

1.1 Conjugate Prior Distribution

The conjugate prior distribution is the **Beta** distribution.

Proof:

$$
\begin{aligned}
y_i|\theta &\sim Binomial(a,b) \\
\theta &\sim Beta(a,b) \\
\pi(\theta) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}
\end{aligned}
$$

$$
g(\lambda|\boldsymbol{y}) = \frac{\pi(\lambda)\prod_{i=1}^{n}f(y_i|\lambda)}{\int_0^\infty \pi(\lambda)\prod_{i=1}^{n}f(y_i|\lambda)d\lambda}
$$

$$
= \frac{\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}\prod_{i=1}^{n}\binom{n_i}{y_i}\theta^{y_i}(1-\theta)^{n_i-y_i}}{\int_0^1 \dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}\prod_{i=1}^{n}\binom{n_i}{y_i}\theta^{y_i}(1-\theta)^{n_i-y_i}d\theta}
$$

$$
= \frac{\theta^{\sum y_i}(1-\theta)^{\sum n_i-y_i}}{\int_0^\infty \theta^{\sum y_i}(1-\theta)^{\sum n_i-y_i}d\theta}
$$

$$
= \frac{\dfrac{\Gamma(a+\sum n_i+b)}{\Gamma(\sum y_i+a)\Gamma(\sum n_i-y_i+b)}\theta^{a+\sum y_i-1}(1-\theta)^{b+\sum n_i-y_i-1}}{\int_0^\infty \dfrac{\Gamma(a+\sum n_i+b)}{\Gamma(\sum y_i+a)\Gamma(\sum n_i-y_i+b)}\theta^{a+\sum y_i-1}(1-\theta)^{b+\sum n_i-y_i-1}d\theta}
$$

$$
\implies \lambda|\boldsymbol{y} \sim Beta(a+\sum_{i=1}^{n}y_i, b+\sum_{i=1}^{n}n_i-y_i)
$$

1.2 Chosen Values for Prior distribution

$\theta \sim Beta(a=1, b=9)$

I have chosen $a = 1$, and $b = 9$ since I do not know enough about baseball to put a lot of confidence in my belief. $a + b = 10$, which is relatively small compared to the sample size of 153. I do think that

homeruns should be relatively uncommon. I would expect no more than 10% of all hits to be homeruns. So, I also chose a and b such that the expected value of θ is 10%.

1.3 Derivation of Maximum Likelihood Estimator

$$
\begin{aligned}
\mathcal{L}(\theta|\boldsymbol{y}) &= \prod_{i=1}^{n} \binom{n_i}{y_i} \theta^{y_i} (1-\theta)^{\sum(n_i - y_i)} \\
l(\theta|\boldsymbol{y}) &= log\{\prod_{i=1}^{n} \binom{n_i}{y_i} \theta^{y_i} (1-\theta)^{\sum(n_i - y_i)}\} \\
&= \sum_{i=1}^{n} log\binom{n_i}{y_i} + log(\theta)\sum_{i=1}^{n} y_i + log(1-\theta)\sum_{i=1}^{n}(n_i - y_i) \\
l'(\theta|\boldsymbol{y}) &= \frac{\sum y_i}{\theta} - \frac{\sum n_i - y_i}{1-\theta} \\
l''(\theta|\boldsymbol{y}) &= \frac{-\sum y_i}{\theta^2} - \frac{\sum n_i - y_i}{(1-\theta)^2} \\
&< 0 \quad (\implies LocalMax)
\end{aligned}
$$

$$
\begin{aligned}
l'(\theta|\boldsymbol{y}) &:= 0 \\
\frac{\sum y_i}{\hat{\theta}} - \frac{\sum n_i - y_i}{1-\hat{\theta}} &= 0 \\
\sum y_i - \hat{\theta}\sum y_i - \hat{\theta}(\sum n_i - \sum y_i) &= 0 \\
\hat{\theta} &= \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} n_i}
\end{aligned}
$$

## 1.4 Plot of MLE, and Prior and Posterior Distributions
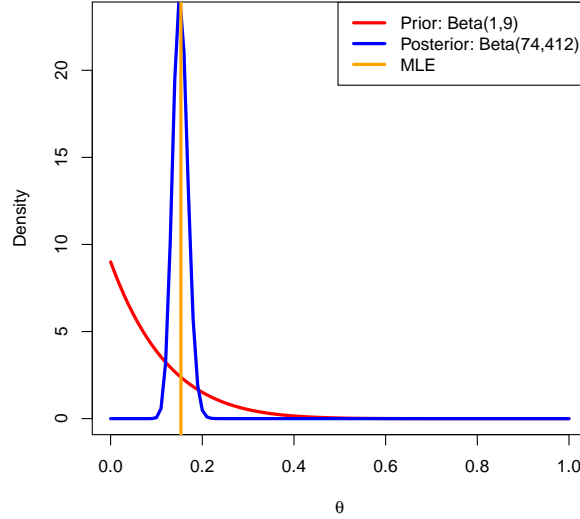
1.5 Posterior Mean, Mode, Median, Variance, Standard Deviation, and 95% Central Credible Interval

Since the posterior distribution is $Beta(74, 412)$,
the posterior mean is $74/(74 + 412) = 0.1523$.
the posterior mode is $(74 - 1)/(74 + 412 - 2) = 0.1508$.
the posterior median is 0.1518.
the posterior variance is $3e - 04$.
the posterior standard deviation is 0.0163.
the 95% credible interval is $(0.141, 0.1855)$.

1.6 Three Other Prior Distributions & Plots of Posterior Distributions

I chose Beta(90,10) as one of my priors to illustrate how the posterior would change if I were very certain that the probability of home runs were high. I chose Beta (100,90) to illustrate how the posterior would change if I were very certain that the probability of home runs were low. I chose Beta(5,5) as my last prior to illustrate how the posterior would change if I were not certain what the probability of home runs should be.
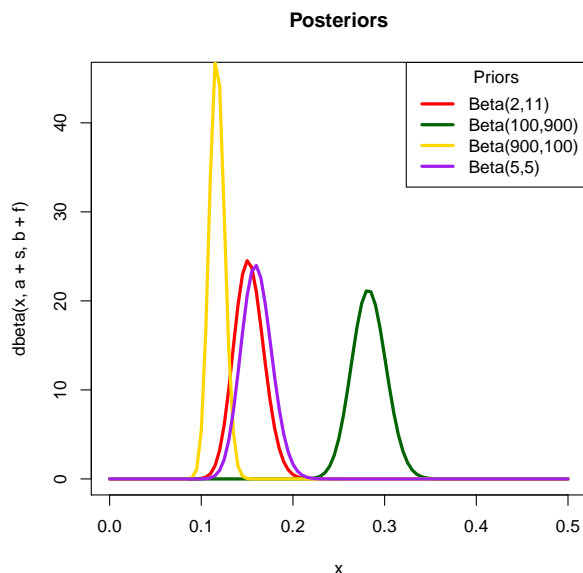
2 Poisson Data

2.1 Conjugate Prior Distribution

The conjugate prior distribution is the **Gamma** distribution.
Proof:

$$
\begin{aligned}
y_i|\lambda &\sim POI(\lambda) \\
\lambda &\sim Gamma(a,b)
\end{aligned}
$$

$$
f(y_i|\lambda) = \frac{e^{-\lambda}\lambda^y)i}{y_i!}
$$

$$
g(\lambda|\boldsymbol{y}) = \frac{\pi(\lambda)\prod_{i=1}^{n} f(y_i|\lambda)}{\int_0^\infty \pi(\lambda)\prod_{i=1}^{n} f(y_i|\lambda)d\lambda}
$$

$$
= \frac{\frac{1}{\Gamma(a)b^a}\lambda^{a-1}e^{-\lambda/b}\prod_{i=1}^{n}\frac{e^{\lambda}\lambda^{y_i}}{y_i!}}{\int_0^\infty \frac{1}{\Gamma(a)b^a}\lambda^{a-1}e^{-\lambda/b}\prod_{i=1}^{n}\frac{e^{\lambda}\lambda^{y_i}}{y_i!}d\lambda}
$$

$$
= \frac{\frac{1}{\Gamma(a)b^a}\lambda^{a-1}e^{-\lambda/b}\frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n}y_i}}{\prod_{i=1}^{n}y_i!}}{\int_0^\infty \frac{1}{\Gamma(a)b^a}\lambda^{a-1}e^{-\lambda/b}\frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n}y_i}}{\prod_{i=1}^{n}y_i!}d\lambda}
$$

$$
= \frac{e^{-(n+1/b)\lambda}\lambda^{\sum_{i=1}^{n}y_i+a-1}}{\int_0^\infty e^{-(n+1/b)\lambda}\lambda^{\sum_{i=1}^{n}y_i+a-1}d\lambda}
$$

$$
= \frac{e^{-(n+1/b)\lambda}\lambda^{\sum_{i=1}^{n}y_i+a-1}\dfrac{1}{\Gamma(\sum_{i=1}^{n}y_i+a)(\frac{1}{n+1/b})^{\sum_{i=1}^{n}y_i+a}}}{\int_0^\infty e^{-(n+1/b)\lambda}\lambda^{\sum_{i=1}^{n}y_i+a-1}\dfrac{1}{\Gamma(\sum_{i=1}^{n}y_i+a)(\frac{1}{n+1/b})^{\sum_{i=1}^{n}y_i+a}}d\lambda}
$$

$$
= \frac{1}{\Gamma(\sum_{i=1}^{n}y_i+a)(\frac{1}{n+1/b})^{\sum_{i=1}^{n}y_i+a}}e^{-(n+1/b)\lambda}\lambda^{\sum_{i=1}^{n}y_i+a-1}
$$

$$
\implies \lambda|\boldsymbol{y} \sim Gamma(\sum_{i=1}^{n}y_i+a, (n+\frac{1}{b})^{-1})
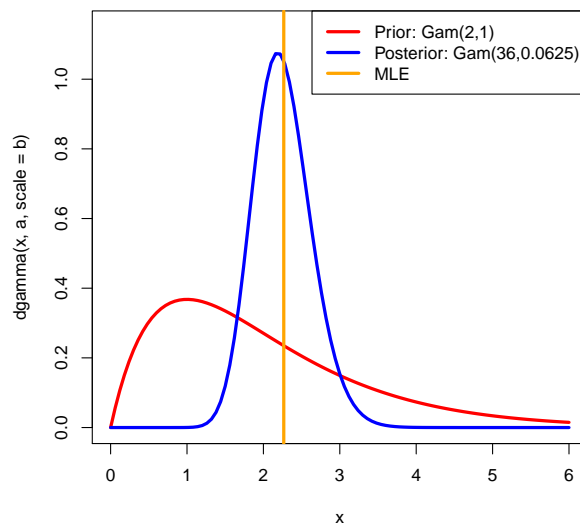$$

2.2   Chosen Values for Prior Distribution

I chose a Gamma(2,1) prior because it was unlikely (probability $< .0005$) for random draws from prior to exceed 10. And the variance is not large (4).

2.3 Derivation of Maximum Likelihood Estimator

$$\mathcal{L}(\lambda|\boldsymbol{y}) \quad = \quad \prod_{i=1}^{n} f(y_i|\lambda)$$

$$= \quad \prod_{i=1}^{n} \frac{e^{\lambda}\lambda^{y_i}}{y_i!}$$

$$= \quad \frac{e^{-n\lambda}\lambda^{\sum\limits_{i=1}^{n} y_i}}{\prod\limits_{i=1}^{n} y_i!}$$

$$l(\lambda|\boldsymbol{y}) \quad = \quad log(\mathcal{L}(\lambda|\boldsymbol{y}))$$

$$= \quad -n\lambda + \sum_{i=1}^{n} y_i log(\lambda) - \sum_{i=1}^{n} y_i!$$

$$= \quad -n + \frac{\sum\limits_{i=1}^{n} y_i}{\lambda}$$

$$l' \quad := \quad 0$$

$$-n + \frac{\sum\limits_{i=1}^{n} y_i}{\hat{\lambda}} \quad = \quad 0$$

$$\hat{\lambda} \quad = \frac{\sum\limits_{i=1}^{n} y_i}{\lambda} = \bar{\boldsymbol{y}}$$

## 2.4 Plot of MLE, and Prior and Posterior Distributions

2.5 Posterior Mean, Mode, Median, Variance, Standard Deviation, and 95% Credible Interval

$$
\begin{aligned}
\lambda|\boldsymbol{y} &\sim Gamma(a,b) \\
\implies \quad E[\lambda|\boldsymbol{y}] &= ab &&= 2 \\
Var[\lambda|\boldsymbol{y}] &= a^2 b &&= 4 \\
SD[\lambda|\boldsymbol{y}] &= \sqrt{a^2 b} &&= 2 \\
Mode[\lambda|\boldsymbol{y}] &= (a-1)b &&= 1 \\
Median[\lambda|\boldsymbol{y}] &&&= 1.6783 \\
95\% CI[\lambda|\boldsymbol{y}] &&&= (0.2422, 5.5716)
\end{aligned}
$$

2.6 Three Other Prior Distributions & Plots

I chose the priors Gamma(2,2), Gamma(4,1), and Gamma(1,4) because they don't generate draws greater than 10 with high probability. They each have an expected value of 4, but their variances are 8, 16, and 4, respectively. Their posteriors are very similar to the original posterior.
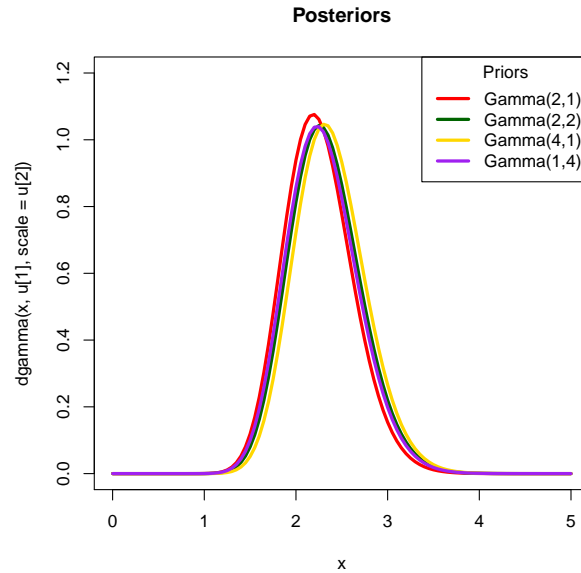
3 Exponential Data

3.1 Conjugate Prior Distribution

The conjugate prior distribution is the **Inverse Gamma** distribution.

Proof:

$$
\begin{aligned}
y_i|\lambda &\sim EXP(\lambda) \\
\lambda &\sim \Gamma^{-1}(a,b) \\
\pi(\lambda) &= \frac{1}{\Gamma(a)b^a}\lambda^{-(a+1)}e^{\frac{-1}{b\lambda}}
\end{aligned}
$$

$$
g(\lambda|\boldsymbol{y}) = \frac{\pi(\lambda)\prod\limits_{i=1}^{n}f(y_i|\lambda)}{\int_0^\infty \pi(\lambda)\prod\limits_{i=1}^{n}f(y_i|\lambda)d\lambda}
$$

$$
= \frac{\frac{1}{\Gamma(a)b^a}\lambda^{-(a+1)}e^{\frac{-1}{b\lambda}}\prod\limits_{i=1}^{n}\frac{e^{-y_i/\lambda}}{\lambda}}{\int_0^\infty \frac{1}{\Gamma(a)b^a}\lambda^{-(a+1)}e^{\frac{-1}{b\lambda}}\prod\limits_{i=1}^{n}\frac{e^{-y_i/\lambda}}{\lambda}d\lambda}
$$

$$
= \frac{\dfrac{1}{\Gamma(a+n)\left(\dfrac{1}{\sum\limits_{i=1}^{n}y_i+\frac{1}{b}}\right)^{a+n}}(\lambda)^{-(a+n+1)}e^{\frac{-1}{\left(\frac{1}{\sum\limits_{i=1}^{n}y_i+\frac{1}{b}}\right)}\lambda}}{\displaystyle\int_0^\infty \dfrac{1}{\Gamma(a+n)\left(\dfrac{1}{\sum\limits_{i=1}^{n}y_i+\frac{1}{b}}\right)^{a+n}}(\lambda)^{-(a+n+1)}e^{\frac{-1}{\left(\frac{1}{\sum\limits_{i=1}^{n}y_i+\frac{1}{b}}\right)}\lambda}d\lambda}
$$

$$
= \frac{1}{\Gamma(a+n)\left(\dfrac{1}{\sum\limits_{i=1}^{n}y_i+\frac{1}{b}}\right)^{a+n}}(\lambda)^{-(a+n+1)}e^{\frac{-1}{\left(\frac{1}{\sum\limits_{i=1}^{n}y_i+\frac{1}{b}}\right)}\lambda}
$$

$$
\implies \lambda|\boldsymbol{y} \sim \Gamma^{-1}\left(a+n,\left(\sum_{i=1}^{n}y_i+\frac{1}{b}\right)^{-1}\right)
$$

3.2 Chosen Values for Prior distribution

I chose a = 3 and b = 1/2000 because the mean of the inverse gamma distribution is $\frac{1}{b(a-1)} = 1000$. And the variance is $\frac{1}{b^2(a-1)^2(a-2)} = 10^6$, for a > 2. I suspect the length of rivers to be highly variable, and 1000 miles seems to be reasonable guess.

3.3 Derivation of Maximum Likelihood Estimator

$$
\begin{aligned}
\mathcal{L}(\lambda|\boldsymbol{y}) \ &= \ \prod_{i=1}^{n} f(y_i|\lambda) \\
&= \ \prod_{i=1}^{n} \frac{exp(-y_i/\lambda)}{\lambda} \\
&= \ \frac{exp(-\sum_{i=1}^{n} y_i/\lambda)}{\lambda^n} \\
l \ &= \ log(\mathcal{L}(\lambda|\boldsymbol{y})) \\
&= \ -\sum_{i=1}^{n} y_i/\lambda - n \ log(\lambda) \\
l' \ &= \ \frac{\sum_{i=1}^{n} y_i}{\lambda^2} - \frac{n}{\lambda} \\
l' \ &:= \ 0 \\
\frac{\sum_{i=1}^{n} y_i}{\hat{\lambda}^2} - \frac{n}{\hat{\lambda}} \ &= \ 0 \\
\sum_{i=1}^{n} y_i - n\hat{\lambda} \ &= \ 0 \\
\hat{\lambda} \ &= \ \frac{\sum_{i=1}^{n} y_i}{n} \\
\hat{\lambda} \ &= \ \bar{\boldsymbol{y}}
\end{aligned}
$$

## 3.4 Plot of MLE, and Prior and Posterior Distributions
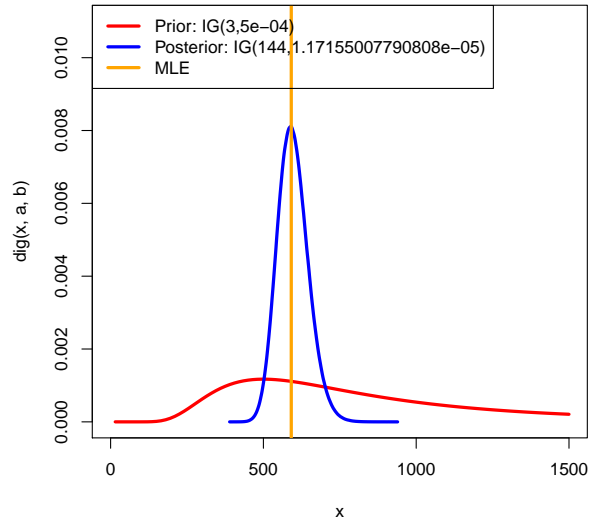
3.5 Posterior Mean, Mode, Median, Variance, Standard Deviation, and 95% Credible Interval

$$
\begin{aligned}
E[\lambda|\boldsymbol{y}] &= 596.8979 \\
Var[\lambda|\boldsymbol{y}] &= 2510.3609 \\
SD[\lambda|\boldsymbol{y}] &= 50.1035 \\
Mode[\lambda|\boldsymbol{y}] &= 590.639 \\
Median[\lambda|\boldsymbol{y}] &= 594.1357 \\
95\%CI[\lambda|\boldsymbol{y}] &= (506.6403, 702.9484)
\end{aligned}
$$

3.6 Three Other Prior Distributions & Plots

I chose the Priors IG(3,1/200), IG(3,1/20), and IG(3,1/2) to illustrate how changing b affects the posterior. The variance of the prior is changed drastically as b chages by a factor of 10. But the posteriors are similar.
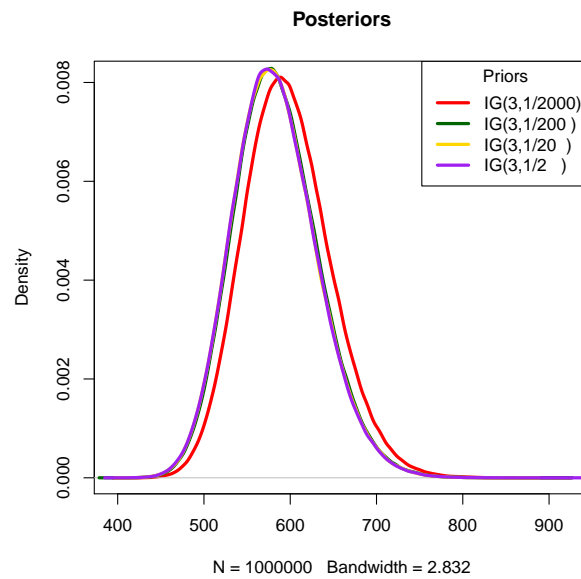
4 Normal Data - σ known

4.1 Conjugate Prior Distribution

The conjugate prior distribution is the **Normal** distribution.
Proof:

9

$$
\begin{aligned}
y_i|\mu &\sim Normal(\mu, \sigma = 9) \\
mu &\sim Normal(m, s) \\
\pi(\mu) &= (2\pi s^2)^{-.5} exp\{-\tfrac{(y_i-m)^2}{2s^2}\}
\end{aligned}
$$

$$
g(\mu|\boldsymbol{y}) = \frac{(2\pi s^2)^{-.5} exp\{-\dfrac{(\mu-m)^2}{2s^2}\}\displaystyle\prod_{i=1}^{n}(2\pi\sigma^2)^{-.5}exp\{-\dfrac{(y_i-\mu)^2}{2\sigma^2}\}}{\displaystyle\int_{-\infty}^{\infty}(2\pi s^2)^{-.5}exp\{-\dfrac{(\mu-m)^2}{2s^2}\}\displaystyle\prod_{i=1}^{n}(2\pi\sigma^2)^{-.5}exp\{-\dfrac{(y_i-\mu)^2}{2\sigma^2}\}d\mu}
$$

$$
= \frac{exp\{\displaystyle\sum_{i=1}^{n}\dfrac{-y_i^2+2\mu y_i-\mu^2}{2\sigma^2}+\dfrac{-\mu^2+2\mu m-m^2}{2s^2}\}}{\displaystyle\int_{-\infty}^{\infty}exp\{\displaystyle\sum_{i=1}^{n}\dfrac{-y_i^2+2\mu y_i-\mu^2}{2\sigma^2}+\dfrac{-\mu^2+2\mu m-m^2}{2s^2}\}d\mu}
$$

$$
= \frac{exp\{\dfrac{n\mu\bar{y}}{\sigma^2}-\dfrac{n\mu^2}{2\sigma^2}-\dfrac{\mu^2}{2s^2}+\dfrac{\mu m}{s^2}\}}{\displaystyle\int_{-\infty}^{\infty}exp\{\dfrac{n\mu\bar{y}}{\sigma^2}-\dfrac{n\mu^2}{2\sigma^2}-\dfrac{\mu^2}{2s^2}+\dfrac{\mu m}{s^2}\}d\mu}
$$

$$
= \frac{exp\{\dfrac{-\mu^2}{2}(\dfrac{n}{\sigma^2}+\dfrac{1}{s^2})+\dfrac{2\mu}{2}(\dfrac{n\bar{y}}{\sigma^2}+\dfrac{m}{s^2})\}}{\displaystyle\int_{-\infty}^{\infty}exp\{\dfrac{-\mu^2}{2}(\dfrac{n}{\sigma^2}+\dfrac{1}{s^2})+\dfrac{2\mu}{2}(\dfrac{n\bar{y}}{\sigma^2}+\dfrac{m}{s^2})\}d\mu}
$$

$$
= \frac{exp\{\dfrac{-\mu^2+2\mu(\frac{n\bar{y}}{\sigma^2}+\frac{m}{s^2})(\frac{s^2\sigma^2}{s^2+\sigma^2})}{2(\frac{n}{\sigma^2}+\frac{1}{s^2})^{-1}}\}}{\displaystyle\int_{-\infty}^{\infty}exp\{\dfrac{-\mu^2+2\mu(\frac{n\bar{y}}{\sigma^2}+\frac{m}{s^2})(\frac{s^2\sigma^2}{s^2+\sigma^2})}{2(\frac{n}{\sigma^2}+\frac{1}{s^2})^{-1}}\}d\mu}
$$

$$
= \frac{exp\{\dfrac{-\mu^2+2\mu(\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})-(\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})^2+(\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})^2}{2(\frac{s^2\sigma^2}{ns^2+\sigma^2})}\}}{\displaystyle\int_{-\infty}^{\infty}exp\{\dfrac{-\mu^2+2\mu(\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})-(\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})^2+(\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})^2}{2(\frac{s^2\sigma^2}{ns^2+\sigma^2})}\}d\mu}
$$

$$
= \frac{(2\pi\sqrt{\dfrac{s^2\sigma^2}{ns^2+\sigma^2}}^{2})^{-.5}exp\{\dfrac{-\mu^2+2\mu(\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})-(\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})^2+(\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})^2}{2(\sqrt{\frac{s^2\sigma^2}{ns^2+\sigma^2}}^{2})}\}}{\displaystyle\int_{-\infty}^{\infty}(2\pi\sqrt{\dfrac{s^2\sigma^2}{ns^2+\sigma^2}}^{2})^{-.5}exp\{\dfrac{-\mu^2+2\mu(\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})-(\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})^2+(\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})^2}{2(\sqrt{\frac{s^2\sigma^2}{ns^2+\sigma^2}}^{2})}\}d\mu}
$$

$$
= (2\pi\sqrt{\tfrac{s^2\sigma^2}{ns^2+\sigma^2}}^{2})^{-.5}exp\{-\dfrac{(\mu-\frac{n\bar{y}s^2+m\sigma^2}{ns^2+\sigma^2})^2}{2\sqrt{\frac{s^2\sigma^2}{ns^2+\sigma^2}}^{2}}\}
$$

$$\implies \mu|\boldsymbol{y} \sim Normal(\frac{n\bar{y}s^2 + m\sigma^2}{ns^2 + \sigma^2}, \sqrt{\frac{s^2\sigma^2}{ns^2 + \sigma^2}}), \ \ where \ \ \sigma = 9.$$
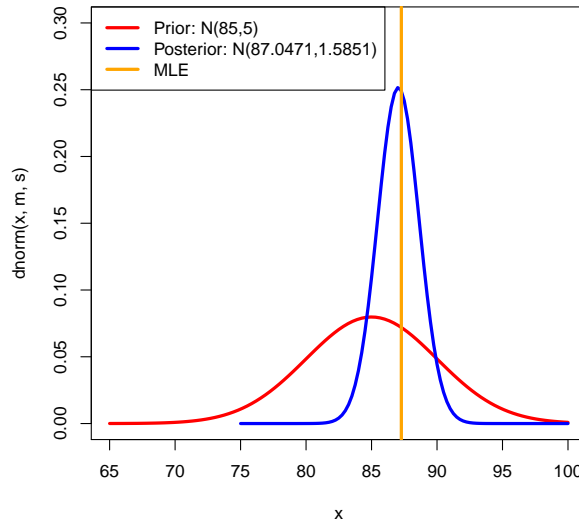
4.2 Chosen Values for Prior distribution

I chose the $m = 85$ and $s = 5$ because Stat 221 apparently has easy exams.

4.3 Derivation of Maximum Likelihood Estimator

$$
\begin{aligned}
\mathcal{L}(\lambda|\boldsymbol{y}) &= \prod_{i=1}^{n} f(y_i|\lambda) \\
&= \prod_{i=1}^{n} (2\pi\sigma^2)^{-.5} exp\{-\frac{(y_i-\mu)^2}{2\sigma^2}\} \\
l(\lambda|\boldsymbol{y}) &= -\frac{n}{2}log(2\pi\sigma^2) - \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\sigma^2} \\
l' &= \frac{2\sum_{i=1}^{n}(y_i-\mu)}{2\sigma^2} \\
&= \frac{\sum_{i=1}^{n}y_i - n\mu}{\sigma^2} \\
l' &:= 0 \\
\frac{\sum_{i=1}^{n}y_i - n\hat{\mu}}{\sigma^2} &= 0 \\
\hat{\mu} &= \frac{\sum_{i=1}^{n}y_i}{n} \\
\hat{\mu} &= \bar{\boldsymbol{y}}
\end{aligned}
$$

## 4.4 Plot of MLE, and Prior and Posterior Distributions

11

4.5 Posterior Mean, Mode, Median, Variance, Standard Deviation, and 95% Credible Interval

$$
\begin{aligned}
E[\lambda|\boldsymbol{y}] &= 87.0471 \\
Var[\lambda|\boldsymbol{y}] &= 1.5851 \\
SD[\lambda|\boldsymbol{y}] &= 1.259 \\
Mode[\lambda|\boldsymbol{y}] &= 87.0471 \\
Median[\lambda|\boldsymbol{y}] &= 87.0471 \\
95\%CI[\lambda|\boldsymbol{y}] &= (83.9405, 90.1538)
\end{aligned}
$$

4.6 Three Other Prior Distributions & Plots

I chose the Priors N(90,5) and N(80,5) to show how changing the mean of the parameter changes the posterior. I chose N(85,10) to show how changing the standard deviation of the parameter changes the posterior. All the posteriors look similar. So, the posterior of the mean parameter is not affected much by the priors.
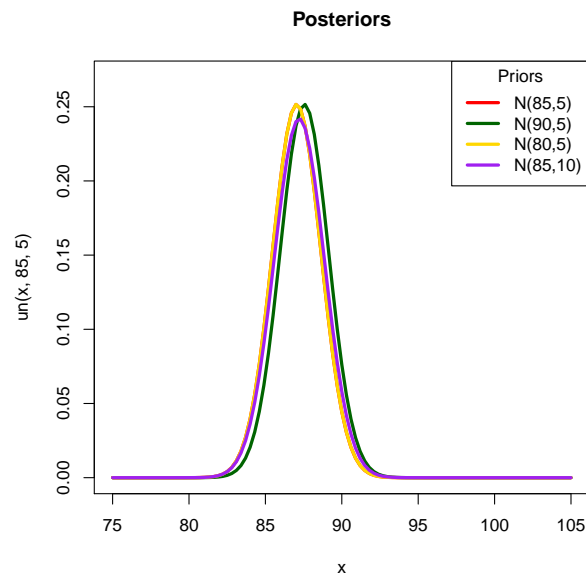
5 Normal Data - μ known

5.1 Conjugate Prior Distribution

The conjugate prior distribution is the **this** distribution.

Proof:

$$
\begin{aligned}
y_i|\sigma^2 &\sim Normal(\mu = 87, \sigma^2)\\
\sigma^2 &\sim \Gamma^{-1}(a,b)\\
\pi(\sigma^2) &= \frac{1}{\Gamma(a)b^a}(\sigma^2)^{-(a+1)}e^{\frac{-1}{b\sigma^2}}
\end{aligned}
$$

$$
\begin{aligned}
g(\sigma^2|\boldsymbol{y}) &= \frac{\pi(\sigma^2)\prod\limits_{i=1}^{n}f(y_i|\sigma^2)}{\int_0^{\infty}\pi(\sigma^2)\prod\limits_{i=1}^{n}f(y_i|\sigma^2)d\sigma^2}\\[2em]
&= \frac{\frac{1}{\Gamma(a)b^a}(\sigma^2)^{-(a+1)}e^{\frac{-1}{b\sigma^2}}\prod\limits_{i=1}^{n}(2\pi\sigma^2)^{-.5}exp\{-\frac{(y_i-\mu)^2}{2\sigma^2}\}}{\int_0^{\infty}\frac{1}{\Gamma(a)b^a}(\sigma^2)^{-(a+1)}e^{\frac{-1}{b\sigma^2}}\prod\limits_{i=1}^{n}(2\pi\sigma^2)^{-.5}exp\{-\frac{(y_i-\mu)^2}{2\sigma^2}\}d\sigma^2}\\[2em]
&= \frac{(2\pi\sigma^2)^{-n/2}exp\{\frac{-\sum\limits_{i=1}^{n}(y_i-\mu)^2}{2\sigma^2}\}\frac{1}{\Gamma(a)b^a}(\sigma^2)^{-(a+1)}e^{\frac{-1}{b\sigma^2}}}{\int_0^{\infty}(2\pi\sigma^2)^{-n/2}exp\{\frac{-\sum\limits_{i=1}^{n}(y_i-\mu)^2}{2\sigma^2}\}\frac{1}{\Gamma(a)b^a}(\sigma^2)^{-(a+1)}e^{\frac{-1}{b\sigma^2}}d\sigma^2}\\[2em]
&= \frac{(\sigma^2)^{-(a+\frac{n}{2}+1)}exp\left\{-\left(\frac{1}{2/\sum\limits_{i=1}^{n}(y_i-\mu)^2}+\frac{1}{b}\right)\frac{1}{\sigma^2}\right\}}{\int_0^{\infty}(\sigma^2)^{-(a+\frac{n}{2}+1)}exp\left\{-\left(\frac{1}{2/\sum\limits_{i=1}^{n}(y_i-\mu)^2}+\frac{1}{b}\right)\frac{1}{\sigma^2}\right\}d\sigma^2}\\[2em]
&= \frac{\left\{\frac{1}{\Gamma(a+\frac{n}{2})\left(\frac{\sum\limits_{i=1}^{n}(y_i-\mu)^2}{2}+\frac{1}{b}\right)^{-1}}\right\}(\sigma^2)^{-(a+\frac{n}{2}+1)}exp\left\{-\left(\frac{1}{2/\sum\limits_{i=1}^{n}(y_i-\mu)^2}+\frac{1}{b}\right)\frac{1}{\sigma^2}\right\}}{\int_0^{\infty}\left\{\frac{1}{\Gamma(a+\frac{n}{2})\left(\frac{\sum\limits_{i=1}^{n}(y_i-\mu)^2}{2}+\frac{1}{b}\right)^{-1}}\right\}(\sigma^2)^{-(a+\frac{n}{2}+1)}exp\left\{-\left(\frac{1}{2/\sum\limits_{i=1}^{n}(y_i-\mu)^2}+\frac{1}{b}\right)\frac{1}{\sigma^2}\right\}d\sigma^2}\\[2em]
&= \left\{\frac{1}{\Gamma(a+\frac{n}{2})\left(\frac{\sum\limits_{i=1}^{n}(y_i-\mu)^2}{2}+\frac{1}{b}\right)^{-1}}\right\}(\sigma^2)^{-(a+\frac{n}{2}+1)}exp\left\{-\left(\frac{1}{2/\sum\limits_{i=1}^{n}(y_i-\mu)^2}+\frac{1}{b}\right)\frac{1}{\sigma^2}\right\}
\end{aligned}
$$

$$\implies \sigma^2 \sim \Gamma^{-1}\left(a + \frac{n}{2}, \left(\frac{\sum_{i=1}^{n}(y_i - \mu)^2}{2} + \frac{1}{b}\right)^{-1}\right), \quad where \ \mu = 87.$$

5.2　Chosen Values for Prior distribution

I chose a=3 and b = 1/3 because the mean and variance for σ^2 would be 1.5 and 2.25 respectively, which seems to be a good guess for standard deviation in an "easy" statistics class.

5.3　Derivation of Maximum Likelihood Estimator

$$
\begin{aligned}
\mathcal{L}(\sigma^2|\boldsymbol{y}) &= \prod_{i=1}^{n} f(y_i|\sigma^2) \\
&= \prod_{i=1}^{n}(2\pi\sigma^2)^{-.5} exp\{-\frac{(y_i-\mu)^2}{2\sigma^2}\} \\
l &= \frac{-n}{2}log(2\pi\sigma^2) - \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\sigma^2} \\
l' &= \frac{-n}{2\sigma^2} + \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2(\sigma^2)^2} \\
\\
l' &:= 0 \\
\frac{-n}{2\hat{\sigma^2}} + \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2(\hat{\sigma^2})^2} &= 0 \\
-n\hat{\sigma^2} + \sum(y_i-\mu)^2 &= 0 \\
\hat{\sigma^2} &= \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{n}, \quad where \ \mu = 87.
\end{aligned}
$$

5.4　Plot of MLE, and Prior and Posterior Distributions
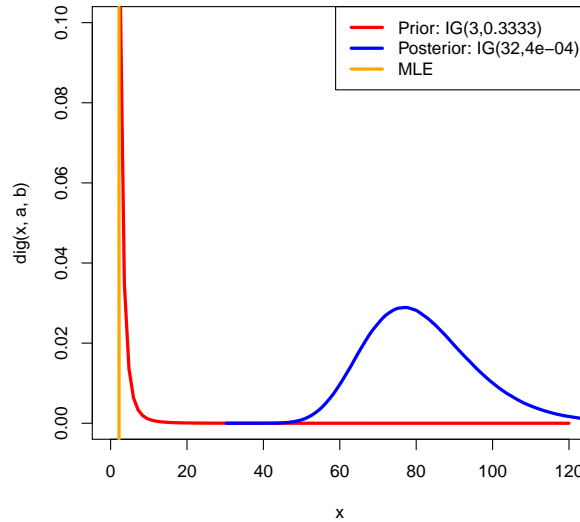
5.5 Posterior Mean, Mode, Median, Variance, Standard Deviation, and 95% Credible Interval

$$
\begin{aligned}
E[\lambda|\boldsymbol{y}] &= 81.7434 \\
Var[\lambda|\boldsymbol{y}] &= 222.5596 \\
SD[\lambda|\boldsymbol{y}] &= 14.9184 \\
Mode[\lambda|\boldsymbol{y}] &= 77.3846 \\
Median[\lambda|\boldsymbol{y}] &= 80.0219 \\
95\%CI[\lambda|\boldsymbol{y}] &= (57.5852, 115.7609)
\end{aligned}
$$

5.6 Three Other Prior Distributions & Plots

I chose the Priors IG(5,1/3), IG(7,1/3), and IG(9,1/3) to illustrate how changing b affects the posterior. The posteriors are noticeably affected by small changes in a.

Posteriors

N = 1000000   Bandwidth = 0.02842

15