STAT 35920

Robert Winter

HW #3

1a. Sample $n>1$ observations $(y_1, \dots, y_n)$ w/ observed mean $\bar{y}$. Assume the likelihood fcn (joint dist.) of $(y_1, \dots, y_n)$ is prop. to

$$\prod_{i=1}^{n} \exp\{-\tfrac{1}{2}\tau(y_i-\mu)^2\}$$

The prior for mean μ is $\mu \sim N(\mu_0, \sigma_0^2)$. Let $\tau_0 = \frac{1}{\sigma_0^2}$. Assume τ, τ_0, μ_0 all known. Derive $p(\bar{y}\mid\mu)$ & $p(\mu\mid\bar{y})$.

- We'll write $\tau = \frac{1}{\sigma^2}$.
- Since $L(y_1, \dots, y_n \mid \mu, \tau) \propto \prod_{i=1}^{n} \exp\{-\frac{1}{2}\tau(y_i-\mu)^2\}$, it follows that for each $i \in \{1, \dots, n\}$, $f(y_i \mid \mu, \tau) \propto \exp\{-\frac{1}{2}\tau(y_i-\mu)^2\}$, so $y_i \sim N(\mu, \frac{1}{\tau})$ $\forall i = 1, \dots, n$.
- Since any linear combination of normal RVs is normal, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ has normal distribution.
- $\mathbb{E}[\bar{y}] = \mathbb{E}[\frac{1}{n}\sum_{i=1}^{n} y_i] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[y_i] = \frac{1}{n}\cdot n\mu = \mu$.
- $\mathrm{Var}(\bar{y}) = \mathrm{Var}(\frac{1}{n}\sum_{i=1}^{n} y_i) = \frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{\tau} = \frac{1}{n^2}\cdot\frac{n}{\tau} = \frac{1}{n\tau}$, where we've used that $y_i \perp\!\!\!\perp y_j$ $\forall i \neq j$. We know this is the case since there are no covariance terms in the likelihood function.
- So, $\bar{y} \mid \mu \sim N(\mu, \frac{1}{n\tau})$, and

$$p(\bar{y}\mid\mu) = \frac{1}{\sqrt{2\pi\cdot\frac{1}{n\tau}}}\exp\left\{-\frac{1}{2\cdot\frac{1}{n\tau}}(\bar{y}-\mu)^2\right\}$$

$$\Rightarrow p(\bar{y}\mid\mu) = \frac{\sqrt{n\tau}}{\sqrt{2\pi}}\exp\left\{-\frac{n\tau}{2}(\bar{y}-\mu)^2\right\}$$

· Then $p(\mu|\bar{y}) \propto p(\bar{y}|\mu)\pi(\mu)$

$$= \frac{\sqrt{\tau n}}{\sqrt{2\pi}} e^{-\frac{\tau n}{2}(\bar{y}-\mu)^2} \cdot \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2}$$

$$\propto \exp\left\{-\frac{1}{2}\tau n(\bar{y}^2 - 2\bar{y}\mu + \mu^2) - \frac{1}{2\sigma_0^2}(\mu^2 - 2\mu_0\mu + \mu_0^2)\right\}$$

$$= \exp\left\{-\frac{1}{2}\tau n\bar{y}^2 + \tau n\bar{y}\mu - \frac{1}{2}\tau n\mu^2 - \frac{1}{2}\tau_0\mu^2 + \tau_0\mu_0\mu - \frac{1}{2}\tau_0\mu_0^2\right\}$$

$$= \exp\left\{\mu^2\left(-\frac{1}{2}\tau n - \frac{1}{2}\tau_0\right) + \mu(\tau n\bar{y} + \tau_0\mu_0)\right\} \cdot \exp\left\{-\frac{1}{2}\tau n\bar{y}^2 - \frac{1}{2}\tau_0\mu_0^2\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left[\mu^2(\tau n + \tau_0) - 2\mu(\tau n\bar{y} + \tau_0\mu_0)\right]\right\}$$

$$= \exp\left\{-\frac{1}{2}(\tau n + \tau_0)\left[\mu^2 - \frac{2(\tau n\bar{y} + \tau_0\mu_0)}{\tau n + \tau_0}\mu\right]\right\}$$

$$= \exp\left\{-\frac{1}{2}(\tau n + \tau_0)\left[\mu^2 - \frac{2(\tau n\bar{y} + \tau_0\mu_0)}{\tau n + \tau_0}\mu + \left(\frac{\tau n\bar{y} + \tau_0\mu_0}{\tau n + \tau_0}\right)^2 - \left(\frac{\tau n\bar{y} + \tau_0\mu_0}{\tau n + \tau_0}\right)^2\right]\right\}$$

$$= \exp\left\{-\frac{1}{2}(\tau n + \tau_0)\left[\mu - \frac{\tau n\bar{y} + \tau_0\mu_0}{\tau n + \tau_0}\right]^2\right\} \cdot \exp\left\{-\frac{1}{2}(\tau n + \tau_0) \cdot -\left(\frac{\tau n\bar{y} + \tau_0\mu_0}{\tau n + \tau_0}\right)^2\right\}$$

$$\propto \exp\left\{-\frac{1}{2}(\tau n + \tau_0)\left(\mu - \frac{\tau n\bar{y} + \tau_0\mu_0}{\tau n + \tau_0}\right)^2\right\}$$

· Thus, $\mu|\bar{y} \sim N\left(\frac{\tau n\bar{y} + \tau_0\mu_0}{\tau n + \tau_0}, \frac{1}{\tau n + \tau_0}\right)$,

and

$$p(\mu|\bar{y}) = \frac{1}{\sqrt{2\pi \cdot \frac{1}{\tau n + \tau_0}}} \exp\left\{\frac{-1}{2\left(\frac{1}{\tau n + \tau_0}\right)}\left(\mu - \frac{\tau n\bar{y} + \tau_0\mu_0}{\tau n + \tau_0}\right)^2\right\}$$

$$p(\mu|\bar{y}) = \frac{\sqrt{\tau n + \tau_0}}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(\tau n + \tau_0)\left(\mu - \frac{\tau n\bar{y} + \tau_0\mu_0}{\tau n + \tau_0}\right)^2\right\}$$

1b. Likelihood $y|\theta \sim Bin(n, \theta)$; prior $\theta \sim Beta(a, b)$. For a new obs y_{new}, what is the posterior predictive distribution $p(y_{new}|y)$? What is posterior predictive mean $\mathbb{E}[y_{new}|y]$?

- $p(y_{new}|y) = \int_{\Theta} p(y_{new}|\theta)p(\theta|y)\,d\theta$

$$= \int_0^1 p(y_{new}|\theta) \cdot p(\theta|y)\,d\theta$$
$$= \int_0^1 \binom{n}{y_{new}} \theta^{y_{new}}(1-\theta)^{n-y_{new}} \cdot p(\theta|y)\,d\theta$$

- $p(\theta|y) \propto \pi(\theta) L(y|\theta)$

$$= \frac{1}{B(a,b)}\theta^{a-1}(1-\theta)^{b-1} \cdot \binom{n}{y}\theta^{y}(1-\theta)^{n-y}$$
$$\propto \theta^{a+y-1}(1-\theta)^{b+n-y-1},$$

so $\theta|y \sim Beta(a+y, b+n-y)$.

- Thus,

$$p(y_{new}|y) = \int_0^1 \binom{n}{y_{new}}\theta^{y_{new}}(1-\theta)^{n-y_{new}} \cdot \frac{1}{B(a+y, b+n-y)}\theta^{a+y-1}(1-\theta)^{b+n-y-1}\,d\theta$$
$$= \binom{n}{y_{new}}\frac{1}{B(a+y,b+n-y)}\int_0^1 \theta^{a+y+y_{new}-1}(1-\theta)^{b+2n-y-y_{new}-1}\,d\theta$$
$$\propto \binom{n}{y_{new}} \cdot B(a+y+y_{new}, b+2n-y-y_{new}),$$

so

$$y_{new}|y \sim BetaBin(n, a+y, b+n-y)$$

Thus,

$$p(y_{new}|y) = \binom{n}{y_{new}}\frac{B(a+y+y_{new}, b+2n-y-y_{new})}{B(a+y, b+n-y)}$$

- So, using the expectation of the Beta-Binomial distribution,

$$\mathbb{E}[y_{new}|y] = \frac{n(a+y)}{(a+y)+(b+n-y)}$$
$$\mathbb{E}[y_{new}|y] = \frac{n(a+y)}{a+b+n}$$

2. Assume $y \sim N(\mu, \sigma^2)$. Sps μ & $\tau = \sigma^{-2}$ both unknown, & joint prior is $\pi(\mu, \tau) = \pi(\mu|\tau)\pi(\tau)$ w/

$$\mu|\tau \sim N(\mu_0, \tau^{-1}\sigma_0^2),$$
$$\tau \sim Gamma(\tfrac{a_0}{2}, \tfrac{b_0}{2}).$$

Derive & recognize $p(\mu|\tau, y)$, $p(\mu|y)$, $p(\tau|y)$.

- $p(\mu|\tau, y) \propto \pi(\mu|\tau) L(y|\mu,\tau)$

$$= \frac{1}{\sqrt{2\pi\tau^{-1}\sigma_0^2}} e^{-\frac{1}{2\tau^{-1}\sigma_0^2}(\mu-\mu_0)^2} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

$$\propto \exp\left\{-\frac{1}{2\tau^{-1}\sigma_0^2}(\mu-\mu_0)^2 - \frac{1}{2\sigma^2}(y-\mu)^2\right\}$$

$$= \exp\left\{-\frac{1}{2\tau^{-1}\sigma_0^2}(\mu^2 - 2\mu_0\mu + \mu_0^2) - \frac{1}{2\sigma^2}(y^2 - 2y\mu + \mu^2)\right\}$$

$$= \exp\left\{-\frac{\mu^2}{2\tau^{-1}\sigma_0^2} + \frac{\mu_0\mu}{\tau^{-1}\sigma_0^2} - \frac{\mu_0^2}{2\tau^{-1}\sigma_0^2} - \frac{y^2}{2\sigma^2} + \frac{y\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right\}$$

$$= \exp\left\{\mu^2\left(\frac{-1}{2\tau^{-1}\sigma_0^2} - \frac{1}{2\sigma^2}\right) + \mu\left(\frac{\mu_0}{\tau^{-1}\sigma_0^2} + \frac{y}{\sigma^2}\right)\right\}$$

$$\cdot \exp\left\{\frac{-\mu_0^2}{2\tau^{-1}\sigma_0^2} - \frac{y^2}{2\sigma^2}\right\}$$

$$\propto \exp\left\{\mu^2\left(\frac{-1}{2\tau^{-1}\sigma_0^2} - \frac{\sigma_0^2}{2\tau^{-1}\sigma_0^2}\right) + \mu\left(\frac{\mu_0}{\tau^{-1}\sigma_0^2} + \frac{y\sigma_0^2}{\tau^{-1}\sigma_0^2}\right)\right\}$$

$$= \exp\left\{\left(\frac{-1-\sigma_0^2}{2\tau^{-1}\sigma_0^2}\right)\left[\mu^2 + \mu\left(\frac{\mu_0 + y\sigma_0^2}{\tau^{-1}\sigma_0^2}\right) \Big/ \left(\frac{-1-\sigma_0^2}{2\tau^{-1}\sigma_0^2}\right)\right]\right\}$$

$$= \exp\left\{-\frac{(1+\sigma_0^2)}{2\tau^{-1}\sigma_0^2}\left[\mu^2 + \mu \cdot \frac{\mu_0 + y\sigma_0^2}{\cancel{\tau^{-1}\sigma_0^2}} \cdot \frac{2\cancel{\tau^{-1}\sigma_0^2}}{-(1+\sigma_0^2)}\right]\right\}$$

$$= \exp\left\{-\frac{(1+\sigma_0^2)}{2\tau^{-1}\sigma_0^2}\left[\mu^2 - 2\mu\left(\frac{\mu_0 + y\sigma_0^2}{1+\sigma_0^2}\right)\right]\right\}$$

$$= \exp\left\{-\frac{(1+\sigma_0^2)}{2\tau^{-1}\sigma_0^2}\left[\mu^2 - 2\mu\left(\frac{\mu_0 + y\sigma_0^2}{1+\sigma_0^2}\right) + \left(\frac{\mu_0 + y\sigma_0^2}{1+\sigma_0^2}\right)^2 - \left(\frac{\mu_0 + y\sigma_0^2}{1+\sigma_0^2}\right)^2\right]\right\}$$

$$= \exp\left\{-\frac{(1+\sigma_0^2)}{2\tau^{-1}\sigma_0^2}\left[\mu - \left(\frac{\mu_0 + y\sigma_0^2}{1+\sigma_0^2}\right)\right]^2\right\} \exp\left\{-\frac{(1+\sigma_0^2)}{2\tau^{-1}\sigma_0^2} \cdot -\left(\frac{\mu_0 + y\sigma_0^2}{1+\sigma_0^2}\right)^2\right\}$$

$$\propto \exp\left\{\frac{-1}{2\left(\frac{\tau^{-1}\sigma_0^2}{1+\sigma_0^2}\right)}\left[\mu - \left(\frac{\mu_0 + y\sigma_0^2}{1+\sigma_0^2}\right)\right]^2\right\}$$

$\longrightarrow$

• That is,

$$\mu \mid \tau, y \sim N\left(\frac{\mu_0 + y\sigma_0^2}{1+\sigma_0^2}, \frac{\tau^{-1}\sigma_0^2}{1+\sigma_0^2}\right)$$

So,

$$p(\mu \mid \tau, y) = \frac{1}{\sqrt{2\pi \cdot \frac{\tau^{-1}\sigma_0^2}{1+\sigma_0^2}}} \exp\left\{\frac{-1}{2\left(\frac{\tau^{-1}\sigma_0^2}{1+\sigma_0^2}\right)}\left(\mu - \frac{\mu_0 + y\sigma_0^2}{1+\sigma_0^2}\right)^2\right\}$$

- To find $p(\tau|y)$, we begin by finding $p(\mu,\tau|y)$.
- $p(\mu,\tau|y) \propto L(y|\mu,\tau)\,\pi(\mu|\tau)\,\pi(\tau)$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y-\mu)^2} \cdot \frac{1}{\sqrt{2\pi\tau^{-1}\sigma_0^2}} e^{\frac{-1}{2\tau^{-1}\sigma_0^2}(\mu-\mu_0)^2}$$

$$\cdot \frac{(b_0/2)^{a_0/2}}{\Gamma(a_0/2)} \tau^{\frac{a_0}{2}-1} e^{-b_0\tau/2}$$

$$= \frac{\tau^{1/2}}{\sqrt{2\pi}} e^{-\frac{1}{2}\tau(y-\mu)^2} \cdot \frac{\tau^{1/2}}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{1}{2\sigma_0^2}\tau(\mu-\mu_0)^2}$$

$$\cdot \frac{(b_0/2)^{a_0/2}}{\Gamma(a_0/2)} \tau^{\frac{a_0}{2}-1} e^{-b_0\tau/2}$$

$$\propto \tau^{\frac{a_0}{2}+\frac{1}{2}-1} \tau^{\frac{1}{2}} e^{-b_0\tau/2} \exp\left\{-\frac{\tau}{2}(y-\mu)^2\right\} \exp\left\{-\frac{\tau}{2\sigma_0^2}(\mu-\mu_0)^2\right\}$$

$$= \tau^{\frac{a_0}{2}+\frac{1}{2}-1} \tau^{\frac{1}{2}} e^{\frac{-b_0\tau}{2}} \exp\left\{-\frac{\tau}{2}(y-\mu)^2 - \frac{\tau}{2\sigma_0^2}(\mu-\mu_0)^2\right\}$$

- Then

$$p(\tau|y) \propto \int_{-\infty}^{\infty} \tau^{\frac{a_0}{2}-\frac{1}{2}} \tau^{\frac{1}{2}} e^{-\frac{b_0\tau}{2}} \exp\left\{-\frac{\tau}{2}(y-\mu)^2 - \frac{\tau}{2\sigma_0^2}(\mu-\mu_0)^2\right\} d\mu$$

$$= \tau^{\frac{a_0}{2}-\frac{1}{2}} \tau^{\frac{1}{2}} e^{-\frac{b_0\tau}{2}} \int_{-\infty}^{\infty} \exp\left\{-\frac{\tau}{2}(y-\mu)^2 - \frac{\tau}{2\sigma_0^2}(\mu-\mu_0)^2\right\} d\mu$$

- $\int_{-\infty}^{\infty} \exp\left\{-\frac{\tau}{2}(y-\mu)^2 - \frac{\tau}{2\sigma_0^2}(\mu-\mu_0)^2\right\} d\mu$

$$= \int_{-\infty}^{\infty} \exp\left\{-\frac{\tau}{2}(y^2-2\mu y+\mu^2) - \frac{\tau}{2\sigma_0^2}(\mu^2-2\mu\mu_0+\mu_0^2)\right\} d\mu$$

$$= \int_{-\infty}^{\infty} \exp\left\{\mu^2\left(-\frac{\tau}{2}-\frac{\tau}{2\sigma_0^2}\right) + \mu\left(\frac{-\tau}{2}\cdot -2y - \frac{\tau}{2\sigma_0^2}\cdot -2\mu_0\right) + \left(-\frac{\tau}{2}y^2 - \frac{\tau}{2\sigma_0^2}\mu_0^2\right)\right\} d\mu$$

$$= \int_{-\infty}^{\infty} \exp\left\{-\mu^2\left(\frac{\tau\sigma_0^2+\tau}{2\sigma_0^2}\right) + \mu\left(\tau y + \frac{\tau\mu_0}{\sigma_0^2}\right) - \left(\frac{\tau}{2}y^2 + \frac{\tau\mu_0^2}{2\sigma_0^2}\right)\right\} d\mu$$

$$= \int_{-\infty}^{\infty} \exp\left\{-\left[\mu^2 \underbrace{\left(\frac{\tau\sigma_0^2+\tau}{2\sigma_0^2}\right)}_{:=\alpha} - \mu \underbrace{\left(\frac{\tau y\sigma_0^2+\tau\mu_0}{\sigma_0^2}\right)}_{:=\beta} + \underbrace{\left(\frac{\tau\sigma_0^2 y^2+\tau\mu_0^2}{2\sigma_0^2}\right)}_{:=\gamma}\right]\right\} d\mu$$

$$= \sqrt{\frac{\pi}{\alpha}} \exp\left\{\frac{\beta^2}{4\alpha} - \gamma\right\}$$

$$= \sqrt{\frac{\pi}{\left(\frac{\tau\sigma_0^2+\tau}{2\sigma_0^2}\right)}} \exp\left\{\left(\frac{1}{4}\right)\left(\frac{\tau y\sigma_0^2+\tau\mu_0}{\sigma_0^2}\right)^2 \cdot \frac{2\sigma_0^2}{\tau\sigma_0^2+\tau} - \frac{\tau\sigma_0^2 y^2+\tau\mu_0^2}{2\sigma_0^2}\right\}$$

$$= \sqrt{\frac{2\pi\sigma_0^2}{\tau(\sigma_0^2+1)}} \exp\left\{\left(\frac{1}{2}\right) \frac{\tau^2(y\sigma_0^2+\mu_0)^2}{\sigma_0^4} \cdot \frac{\sigma_0^2}{\tau(\sigma_0^2+1)} - \frac{\tau\sigma_0^2 y^2+\tau\mu_0^2}{2\sigma_0^2}\right\}$$

$$= \tau^{-1/2}\sqrt{\frac{2\pi\sigma_0^2}{\sigma_0^2+1}}\exp\left\{\frac{\tau(y\sigma_0^2+\mu_0)^2}{2\sigma_0^2(\sigma_0^2+1)} - \frac{\tau(\sigma_0^2y^2+\mu_0^2)}{2\sigma_0^2}\right\}$$

$$\propto \tau^{-1/2}\exp\left\{\frac{\tau}{2\sigma_0^2}\left(\frac{(y\sigma_0^2+\mu_0)^2}{\sigma_0^2+1} - \frac{(\sigma_0^2+1)(\sigma_0^2y^2+\mu_0^2)}{\sigma_0^2+1}\right)\right\}$$

$$= \tau^{-1/2}\exp\left\{\frac{\tau}{2\sigma_0^2}\left(\frac{y^2\sigma_0^4+2\mu_0y\sigma_0^2+\mu_0^2-(\sigma_0^4y^2+\mu_0^2\sigma_0^2+\sigma_0^2y^2+\mu_0^2)}{\sigma_0^2+1}\right)\right\}$$

$$= \tau^{-1/2}\exp\left\{\frac{\tau}{2(\sigma_0^2+1)}\left(\frac{2\mu_0y\sigma_0^2-\mu_0^2\sigma_0^2-\sigma_0^2y^2}{\sigma_0^2}\right)\right\}$$

$$= \tau^{-1/2}\exp\left\{\frac{\tau}{2(\sigma_0^2+1)}\left(\frac{-\sigma_0^2(y^2-2\mu_0y+\mu_0^2)}{\cancel{\sigma_0^2}}\right)\right\}$$

$$= \tau^{-1/2}\exp\left\{\frac{-\tau}{2(\sigma_0^2+1)}(y-\mu_0)^2\right\}$$

- So,

$$p(\tau|y) \propto \tau^{\frac{a_0}{2}-\frac{1}{2}}\tau^{\frac{1}{2}}e^{\frac{-b_0\tau}{2}}\int_{-\infty}^{\infty}\exp\left\{-\frac{\tau}{2}(y-\mu)^2-\frac{\tau}{2\sigma_0^2}(\mu-\mu_0)^2\right\}d\mu$$

$$\propto \tau^{\frac{a_0}{2}-\frac{1}{2}}\tau^{\frac{1}{2}}e^{\frac{-b_0\tau}{2}}\tau^{-\frac{1}{2}}\exp\left\{\frac{-\tau}{2(\sigma_0^2+1)}(y-\mu_0)^2\right\}$$

$$= \tau^{\frac{a_0}{2}-\frac{1}{2}}\exp\left\{-\tau\left[\frac{b_0}{2}+\frac{1}{2(\sigma_0^2+1)}(y-\mu_0)^2\right]\right\},$$

which is the kernel of a Gamma distribution. So,

$$\tau|y \sim \text{Gamma}\left(\frac{a_0}{2}+\frac{1}{2}, \frac{b_0}{2}+\frac{(y-\mu_0)^2}{2(\sigma_0^2+1)}\right)$$

Thus,

$$p(\tau|y) = \frac{\left(\frac{b_0}{2}+\frac{(y-\mu_0)^2}{2(\sigma_0^2+1)}\right)^{\frac{a_0}{2}+\frac{1}{2}}}{\Gamma\left(\frac{a_0}{2}+\frac{1}{2}\right)}\tau^{\frac{a_0}{2}-\frac{1}{2}}\exp\left\{-\tau\left[\frac{b_0}{2}+\frac{(y-\mu_0)^2}{2(\sigma_0^2+1)}\right]\right\}$$

with support $(0,\infty)$

· So,

$$p(\mu|y) = \int_0^\infty p(\mu|\tau,y)\,p(\tau|y)\,d\tau$$

$$= \int_0^\infty \frac{1}{\sqrt{2\pi \cdot \frac{\tau^{-1}\sigma_0^2}{1+\sigma_0^2}}} \exp\left\{\frac{-1}{2\left(\frac{\tau^{-1}\sigma_0^2}{1+\sigma_0^2}\right)}\left(\mu - \frac{\mu_0 + y\sigma_0^2}{1+\sigma_0^2}\right)^2\right\}$$

$$\cdot \frac{\left(\frac{b_0}{2} + \frac{(y-\mu_0)^2}{2(\sigma_0^2+1)}\right)^{\frac{a_0}{2}+\frac{1}{2}}}{\Gamma\left(\frac{a_0}{2}+\frac{1}{2}\right)} \tau^{\frac{a_0}{2}-\frac{1}{2}} \exp\left\{-\tau\left[\frac{b_0}{2} + \frac{(y-\mu_0)^2}{2(\sigma_0^2+1)}\right]\right\} d\tau$$

$$\propto \int_0^\infty \sqrt{\frac{1+\sigma_0^2}{2\pi\tau^{-1}\sigma_0^2}} \cdot \tau^{\frac{a_0}{2}-\frac{1}{2}} \cdot \exp\left\{-\frac{1+\sigma_0^2}{2\tau^{-1}\sigma_0^2}\left(\mu - \frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2}\right)^2 - \tau\left[\frac{b_0}{2} + \frac{(y-\mu_0)^2}{2(\sigma_0^2+1)}\right]\right\} d\tau$$

$$\propto \int_0^\infty \tau^{\frac{1}{2}} \cdot \tau^{\frac{a_0}{2}-\frac{1}{2}} \exp\left\{-\tau\left[\left(\frac{1+\sigma_0^2}{2\sigma_0^2}\right)\left(\mu - \frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2}\right)^2 + \left(\frac{b_0}{2} + \frac{(y-\mu_0)^2}{2(\sigma_0^2+1)}\right)\right]\right\} d\tau$$

$$= \int_0^\infty \underbrace{\tau^{\frac{a_0}{2}}}_{:=\tau^\beta} \exp\left\{-\tau\underbrace{\left[\left(\frac{1+\sigma_0^2}{2\sigma_0^2}\right)\left(\mu - \frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2}\right)^2 + \left(\frac{b_0}{2} + \frac{(y-\mu_0)^2}{2(\sigma_0^2+1)}\right)\right]}_{:=\alpha}\right\} d\tau$$

$$= \frac{\Gamma(\beta+1)}{\alpha^{\beta+1}}$$

$$= \frac{\Gamma\left(\frac{a_0}{2}+1\right)}{\left[\left(\frac{1+\sigma_0^2}{2\sigma_0^2}\right)\left(\mu - \frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2}\right)^2 + \left(\frac{b_0}{2} + \frac{(y-\mu_0)^2}{2(\sigma_0^2+1)}\right)\right]^{\frac{a_0}{2}+1}}$$

$$\propto \left[\left(\frac{1+\sigma_0^2}{2\sigma_0^2}\right)\left(\mu - \frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2}\right)^2 + \left(\frac{b_0}{2} + \frac{(y-\mu_0)^2}{2(\sigma_0^2+1)}\right)\right]^{-\left(\frac{a_0+2}{2}\right)}$$

$$= \left\{\left(\frac{b_0}{2} + \frac{(y-\mu_0)^2}{2(\sigma_0^2+1)}\right)\left[1 + \frac{\left(\frac{1+\sigma_0^2}{2\sigma_0^2}\right)\left(\mu - \frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2}\right)^2}{\left(\frac{b_0}{2} + \frac{(y-\mu_0)^2}{2(\sigma_0^2+1)}\right)}\right]\right\}^{-\left(\frac{a_0+1+1}{2}\right)}$$

$$\propto \left[1+\left(\frac{1+\sigma_0^2}{2\sigma_0^2}\right)\left(\mu-\frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2}\right)^2\left(\frac{2(\sigma_0^2+1)}{b_0(\sigma_0^2+1)+(y-\mu_0)^2}\right)\right]^{-\left(\frac{(a_0+1)+1}{2}\right)}$$

$$=\left[1+\frac{(\sigma_0^2+1)^2}{b_0\sigma_0^2(\sigma_0^2+1)+\sigma_0^2(y-\mu_0)^2}\left(\mu-\frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2}\right)^2\right]^{-\left(\frac{(a_0+1)+1}{2}\right)}$$

$$=\left[1+\frac{1}{a_0+1}(a_0+1)\left(\frac{(\sigma_0^2+1)^2}{b_0\sigma_0^2(\sigma_0^2+1)+\sigma_0^2(y-\mu_0)^2}\right)\left(\mu-\frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2}\right)^2\right]^{-\left(\frac{(a_0+1)+1}{2}\right)}$$

$$=\left[1+\frac{1}{a_0+1}\left(\frac{(a_0+1)(\sigma_0^2+1)^2}{b_0\sigma_0^2(\sigma_0^2+1)+\sigma_0^2(y-\mu_0)^2}\right)\left(\mu-\frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2}\right)^2\right]^{-\left(\frac{(a_0+1)+1}{2}\right)}$$

$$=\left[1+\frac{1}{a_0+1}\cdot\frac{\left(\mu-\frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2}\right)^2}{\frac{b_0\sigma_0^2(\sigma_0^2+1)+\sigma_0^2(y-\mu_0)^2}{(a_0+1)(\sigma_0^2+1)^2}}\right]^{-\left(\frac{(a_0+1)+1}{2}\right)}$$

which is the kernel of a location-scale t distribution. So,

$$\mu|y \sim LSt\left(\frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2},\ \frac{b_0\sigma_0^2(\sigma_0^2+1)+\sigma_0^2(y-\mu_0)^2}{(a_0+1)(\sigma_0^2+1)^2},\ a_0+1\right)$$

Thus,

$$p(\tau|y)=\frac{\Gamma\left(\frac{a_0}{2}+1\right)\left[1+\frac{1}{a_0+1}\left(\frac{(a_0+1)(\sigma_0^2+1)^2}{b_0\sigma_0^2(\sigma_0^2+1)+\sigma_0^2(y-\mu_0)^2}\right)\left(\mu-\frac{\mu_0+y\sigma_0^2}{1+\sigma_0^2}\right)^2\right]^{-\left(\frac{a_0}{2}+1\right)}}{\Gamma\left(\frac{a_0+1}{2}\right)\sqrt{\pi(a_0+1)\cdot\frac{b_0\sigma_0^2(\sigma_0^2+1)+\sigma_0^2(y-\mu_0)^2}{(a_0+1)(\sigma_0^2+1)^2}}}$$

w/ support $(-\infty,\infty)$

3a. $X \sim Bin(50, p)$; observe $x=30$. Derive the posterior distribution for the prior

$$p \sim \tfrac{1}{2}\{Beta(10,20) + Beta(20,10)\}$$

- The prior PDF is

$$\pi(p) = \tfrac{1}{2}\left(\frac{1}{B(10,20)} p^{10-1}(1-p)^{20-1}\right) + \tfrac{1}{2}\left(\frac{1}{B(20,10)} p^{20-1}(1-p)^{10-1}\right), \ p \in [0,1]$$

$$= \frac{1}{2} \cdot \frac{1}{B(10,20)} p^{9}(1-p)^{19} + \frac{1}{2} \cdot \frac{\Gamma(20+10)}{\Gamma(20)\Gamma(10)} p^{19}(1-p)^{9}$$

$$= \frac{1}{2} \cdot \frac{1}{B(10,20)} p^{9}(1-p)^{19} + \frac{1}{2} \cdot \frac{1}{B(10,20)} p^{19}(1-p)^{9}$$

$$= \frac{1}{2B(10,20)}\left[p^{9}(1-p)^{19} + p^{19}(1-p)^{9}\right]$$

- The likelihood is

$$f(x|p) = \binom{50}{x} p^{x}(1-p)^{50-x}$$

- So, the posterior PDF is

$$f(p|x) \propto \pi(p) f(x|p)$$

$$= \frac{1}{2B(10,20)}\left[p^{9}(1-p)^{19} + p^{19}(1-p)^{9}\right] \cdot \binom{50}{x} p^{x}(1-p)^{50-x}$$

$$\propto p^{x}(1-p)^{50-x}\left[\tfrac{1}{2}p^{9}(1-p)^{19} + \tfrac{1}{2}p^{19}(1-p)^{9}\right]$$

$$= \tfrac{1}{2}p^{9+x}(1-p)^{50-x+19} + \tfrac{1}{2}p^{19+x}(1-p)^{50-x+9}$$

$$= \tfrac{1}{2}p^{9+x}(1-p)^{69-x} + \tfrac{1}{2}p^{19+x}(1-p)^{59-x},$$

which is the average of the kernels of two Beta distributions, so

$$p|x \sim \tfrac{1}{2} Beta(9+x+1, 69-x+1) + \tfrac{1}{2} Beta(19+x+1, 59-x+1)$$

$$= \tfrac{1}{2} Beta(10+x, 70-x) + \tfrac{1}{2} Beta(20+x, 60-x)$$

- Since we observe $x=30$,

$$p|x=30 \sim \tfrac{1}{2} Beta(10+30, 70-30) + \tfrac{1}{2} Beta(20+30, 60-30)$$

$$\Rightarrow p|x=30 \sim \tfrac{1}{2} Beta(40,40) + \tfrac{1}{2} Beta(50,30)$$

So, the posterior PDF of p is

$$f(p|x=30) = \frac{1}{2B(40,40)} p^{39}(1-p)^{39} + \frac{1}{2B(50,30)} p^{49}(1-p)^{29}$$

with support $p \in [0,1]$.

$\rightarrow$

3b. In general, for data (x, n) w/ n fixed & $x \sim Bin(n, p)$, if the prior is $p \sim \sum_{i=1}^{K} w_i \, Beta(a_i, b_i)$ w/ $\sum_{i=1}^{K} w_i = 1$, $w_i > 0 \ \forall i$, derive the posterior distribution in closed form.

· Now the prior PDF is

$$\pi(p) = \sum_{i=1}^{K} w_i \cdot \frac{1}{B(a_i, b_i)} p^{a_i - 1}(1-p)^{b_i - 1}$$

· The likelihood function is

$$f(x|p) = \binom{n}{x} p^x (1-p)^{n-x}$$

· So, the posterior PDF is

$$f(p|x) \propto \pi(p) f(x|p)$$
$$= \sum_{i=1}^{K} w_i \cdot \frac{1}{B(a_i, b_i)} p^{a_i - 1}(1-p)^{b_i - 1} \cdot \binom{n}{x} p^x (1-p)^{n-x}$$
$$\propto p^x (1-p)^{n-x} \sum_{i=1}^{K} w_i \frac{1}{B(a_i, b_i)} p^{a_i - 1}(1-p)^{b_i - 1}$$
$$= \sum_{i=1}^{K} w_i \frac{1}{B(a_i, b_i)} p^{a_i + x - 1}(1-p)^{b_i + n - x - 1},$$

which is a weighted average of the PDFs of Beta distributions, so

$$p|x \sim \sum_{i=1}^{K} w_i \, Beta(a_i + x, b_i + n - x)$$

· So, the posterior PDF of p is

$$f(p|x) = \sum_{i=1}^{K} w_i \cdot \frac{1}{B(a_i + x, b_i + n - x)} p^{a_i + x - 1}(1-p)^{b_i + n - x - 1}$$

with support $p \in [0, 1]$

4. Bayesian linear regression. $Y = (Y_1, \ldots, Y_n)$,
$\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$, $\beta = (\beta_0, \beta_1, \ldots, \beta_{k-1})$,

$$X = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1,k-1} \\ 1 & X_{21} & \cdots & X_{2,k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \cdots & X_{n,k-1} \end{bmatrix}.$$

$$Y_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_{k-1} X_{i,k-1} + \varepsilon_i, \quad \varepsilon_i \sim \text{iid } N(0, \sigma^2)$$

OLS: $\hat{b} = (X^T X)^{-1} X^T Y$, $\hat{\sigma}^2 = \frac{1}{n-k}(y - X\hat{b})^T (y - X\hat{b})$.

Improper priors: $p(\beta) \propto 1$, $\beta \in \mathbb{R}^k$
$p(\sigma^2) \propto \frac{1}{\sigma}$

Derive $p(\beta \mid X, Y)$ & $p(\sigma^2 \mid X, Y)$ up to normalizing constant & name them.

- Recall from lecture notes that we may write the likelihood function as

$$L(\beta, \sigma^2 \mid Y, X) \propto \sigma^{-n} \exp\left\{-\frac{1}{2\sigma^2}\left[\hat{\sigma}^2(n-k) + (\beta - \hat{b})^T X^T X (\beta - \hat{b})\right]\right\}$$
$$= (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\hat{\sigma}^2(n-k)\right\} \exp\left\{-\frac{1}{2\sigma^2}(\beta - \hat{b})^T X^T X (\beta - \hat{b})\right\}$$

- The joint prior for (β, σ^2) is

$$p(\beta, \sigma^2) = p(\beta \mid \sigma^2)\, p(\sigma^2)$$
$$= p(\beta)\, p(\sigma^2) \quad \text{since the marginal priors indicate } \beta \perp\!\!\!\perp \sigma^2$$
$$\propto 1 \cdot \frac{1}{\sigma}$$
$$= \frac{1}{\sigma}$$

- Then

$$p(\beta, \sigma^2 \mid Y, X) \propto p(\beta, \sigma^2)\, L(\beta, \sigma^2 \mid Y, X)$$
$$= \frac{1}{\sigma} \cdot (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\hat{\sigma}^2(n-k)\right\} \exp\left\{-\frac{1}{2\sigma^2}(\beta - \hat{b})^T X^T X (\beta - \hat{b})\right\}$$
$$= (\sigma^2)^{-(n+1)/2} \exp\left\{-\frac{1}{2\sigma^2}\hat{\sigma}^2(n-k)\right\} \exp\left\{-\frac{1}{2\sigma^2}(\beta - \hat{b})^T X^T X (\beta - \hat{b})\right\}$$

- So,

$$p(\sigma^2 \mid Y, X) \propto \int_{\mathbb{R}^k} p(\beta, \sigma^2 \mid Y, X)\, d\beta$$
$$= \int_{\mathbb{R}^k} (\sigma^2)^{-(n+1)/2} \exp\left\{-\frac{1}{2\sigma^2}\hat{\sigma}^2(n-k)\right\} \exp\left\{-\frac{1}{2\sigma^2}(\beta - \hat{b})^T X^T X (\beta - \hat{b})\right\} d\beta$$
$$= (\sigma^2)^{-(n+1)/2} \exp\left\{-\frac{1}{2\sigma^2}\hat{\sigma}^2(n-k)\right\} \int_{\mathbb{R}^k} \exp\left\{-\frac{1}{2\sigma^2}(\beta - \hat{b})^T X^T X (\beta - \hat{b})\right\} d\beta$$

$$=(\sigma^2)^{\frac{-(n+1)}{2}}\sqrt{\frac{\pi^k}{\det(\frac{1}{2\sigma^2}X^TX)}}\exp\left\{-\frac{1}{2\sigma^2}\hat{\sigma}^2(n-k)\right\}$$

$$=(\sigma^2)^{\frac{-(n+1)}{2}}\sqrt{\frac{\pi^k}{(\frac{1}{2\sigma^2})^k\det(X^TX)}}\exp\left\{-\frac{1}{2\sigma^2}\hat{\sigma}^2(n-k)\right\}$$

$$=(\sigma^2)^{\frac{-(n+1)}{2}}\sqrt{\frac{(2\pi)^k(\sigma^2)^k}{\det(X^TX)}}\exp\left\{-\frac{1}{2\sigma^2}\hat{\sigma}^2(n-k)\right\}$$

$$\propto(\sigma^2)^{-(n+1)/2}\cdot(\sigma^2)^{k/2}\exp\left\{-\frac{1}{2\sigma^2}\hat{\sigma}^2(n-k)\right\}$$

$$=(\sigma^2)^{\frac{1}{2}(-n-1+k)}\exp\left\{-\frac{1}{2\sigma^2}\hat{\sigma}^2(n-k)\right\}$$

$$=(\sigma^2)^{-(n-k-1)/2+1}\exp\left\{-\frac{1}{2\sigma^2}\hat{\sigma}^2(n-k)\right\}$$

$$=(\sigma^2)^{-(1+\frac{n-k-1}{2})}\exp\left\{-\frac{(n-k-1)(n-k)\hat{\sigma}^2}{2\sigma^2(n-k-1)}\right\},$$

which is the PDF of an inverse-scaled-χ^2 distribution, so

$$\sigma^2|y,X\sim IS\chi^2\left(n-k-1,\frac{(n-k)\hat{\sigma}^2}{n-k-1}\right)$$

- Also, letting $\tau := \frac{1}{\sigma^2}$,

$$p(\beta|y,X) \propto \int_0^\infty (\sigma^2)^{-(n+1)/2} \exp\{-\tfrac{1}{2\sigma^2}\hat{\sigma}^2(n-k)\}\exp\{-\tfrac{1}{2\sigma^2}(\beta-\hat{b})^T X^T X(\beta-\hat{b})\} d\sigma^2$$

$$\left(\begin{array}{ll} \tau := \frac{1}{\sigma^2} = (\sigma^2)^{-1} & \sigma^2 \to 0 \Rightarrow \tau \to \infty \\ d\tau = \frac{-1}{\sigma^4} d\sigma^2 \Rightarrow d\sigma^2 = -\sigma^4 d\tau = -\tau^{-2} d\tau & \sigma^2 \to \infty \Rightarrow \tau \to 0 \end{array}\right.$$

$$= \int_\infty^0 (\tau^{-1})^{-(n+1)/2} \exp\{-\tfrac{\tau}{2}[\hat{\sigma}^2(n-k) + (\beta-\hat{b})^T X^T X(\beta-\hat{b})]\} \cdot -\tau^{-2} d\tau$$

$$= \int_0^\infty \tau^{(n+1)/2-2} \exp\{-\tfrac{\tau}{2}[\hat{\sigma}^2(n-k) + (\beta-\hat{b})^T X^T X(\beta-\hat{b})]\} d\tau$$

$$= \int_0^\infty \underbrace{\tau^{\frac{1}{2}(n-3)}}_{:=\tau^\beta} \exp\{-\tau \cdot \underbrace{\tfrac{1}{2}[\hat{\sigma}^2(n-k) + (\beta-\hat{b})^T X^T X(\beta-\hat{b})]}_{:=\alpha}\} d\tau$$

$$= \frac{\Gamma(\frac{1}{2}(n-3)+1)}{[\frac{1}{2}\hat{\sigma}^2(n-k) + \frac{1}{2}(\beta-\hat{b})^T X^T X(\beta-\hat{b})]^{\frac{1}{2}(n-3)+1}}$$

$$\propto \left[\frac{\hat{\sigma}^2(n-k)}{2} + \frac{(\beta-\hat{b})^T X^T X(\beta-\hat{b})}{2}\right]^{-(\frac{n}{2}-\frac{3}{2}+1)}$$

$$= \left[\frac{\hat{\sigma}^2(n-k)}{2}\left(1 + \frac{(\beta-\hat{b})^T X^T X(\beta-\hat{b})}{\hat{\sigma}^2(n-k)}\right)\right]^{-(\frac{n}{2}-\frac{1}{2})}$$

$$\propto \left[1 + \frac{(\beta-\hat{b})^T X^T X(\beta-\hat{b})}{\hat{\sigma}^2(n-k)}\right]^{-(n-1)/2}$$

$$= \left[1 + \frac{(n-k-1)}{(n-k-1)}\frac{(\beta-\hat{b})^T X^T X(\beta-\hat{b})}{\hat{\sigma}^2(n-k)}\right]^{-\frac{1}{2}[(n-k-1)+k]}$$

$$= \left[1 + \frac{1}{n-k-1}\frac{(\beta-\hat{b})^T X^T X(\beta-\hat{b})}{\hat{\sigma}^2(n-k)/(n-k-1)}\right]^{-\frac{1}{2}[(n-k-1)+k]},$$

which is the kernel of a multivariate t distribution, so

$$\beta|y,X \sim MVT\left(\hat{b}, \frac{\hat{\sigma}^2(X^TX)^{-1}(n-k)}{n-k-1}, n-k-1\right)$$

5. n identical light bulbs. After L hours of lighting, k have gone out w/ lighting times $y_1, y_2, \dots, y_k$, all $< L$. Remaining $(n-k)$ bulbs still on, but expt ends. $y_i \sim$ iid $Exp(\theta)$. $z_i := \mathbb{I}(y_i > L)$. Write likelihood fcn, as a fcn of θ, for observed data $(y_1, z_1), \dots, (y_n, z_n)$. Find a conjugate prior for θ & its posterior.

- Since $y_i \sim$ iid $Exp(\theta)$, each y_i has PDF

$$f(y) = \begin{cases} \theta e^{-\theta y}, & y \in [0, \infty) \\ 0, & o/w \end{cases}$$

- We decompose f into survival & hazard functions:

$$f(y) = \theta \cdot e^{-\theta y}$$
$$= \lambda(y) \cdot S(y),$$

where $\lambda(y) = \theta$, $S(y) = e^{-\theta y}$.

- Then the likelihood function is

$$\mathcal{L}(\theta \mid y_i\text{'s}, z_i\text{'s}) = \prod_{i=1}^{n} \lambda(y_i)^{1-z_i} S(y_i)$$
$$= \prod_{i=1}^{n} \theta^{1-z_i} e^{-\theta y_i}$$
$$= \theta^{\sum_{i=1}^{n}(1-z_i)} e^{\sum_{i=1}^{n} -\theta y_i}$$
$$= \theta^{n - \sum_{i=1}^{n} z_i} e^{-\theta \sum_{i=1}^{n} y_i}$$
$$= \theta^{n-(n-k)} e^{-\theta \sum_{i=1}^{n} y_i}$$

$$\Rightarrow \mathcal{L}(\theta \mid y_i\text{'s}, z_i\text{'s}) = \theta^{k} \exp(-\theta \textstyle\sum_{i=1}^{n} y_i)$$

- If we treat $y_i = L$ for those bulbs that were censored, we can further simplify:

$$\mathcal{L}(\theta \mid y_i\text{'s}, z_i\text{'s}) = \theta^{k} e^{-\theta [\sum_{i=1}^{k} y_i + \sum_{i=k+1}^{n} L]}$$
$$= \theta^{k} e^{-\theta \sum_{i=1}^{k} y_i} e^{-\theta L \sum_{i=k+1}^{n}}$$

$$\mathcal{L}(\theta \mid y_i\text{'s}, z_i\text{'s}) = \theta^{k} \exp(-\theta \textstyle\sum_{i=1}^{k} y_i) \exp(-\theta L (n-k))$$

- By inspection, suppose $\Theta \sim \text{Gamma}(\alpha, \beta)$, so

$$\pi(\theta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} \Theta^{\alpha-1} e^{-\beta\Theta}, & \Theta \in (0, \infty) \\ 0, & \text{o/w} \end{cases}$$

- Then

$$p(\Theta | y_i\text{'s}, z_i\text{'s}) \propto \pi(\theta) \mathcal{L}(\Theta | y_i\text{'s}, z_i\text{'s})$$
$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \Theta^{\alpha-1} e^{-\beta\Theta} \cdot \Theta^k e^{-\Theta \sum_{i=1}^n y_i}$$
$$\propto \Theta^{\alpha+k-1} e^{-(\beta + \sum_{i=1}^n y_i)\Theta},$$

so

$$\Theta | y_i\text{'s}, z_i\text{'s} \sim \text{Gamma}(a+k, \beta + \textstyle\sum_{i=1}^n y_i).$$

- Since the posterior distribution of Θ is also Gamma, a conjugate prior for Θ is $\Theta \sim \text{Gamma}(a, \beta)$, w/ posterior distribution
$$\Theta | y_i\text{'s}, z_i\text{'s} \sim \text{Gamma}(a+k, \beta + \textstyle\sum_{i=1}^n y_i)$$

6. N race cars. $N \sim Geom(\quad)$ w/ $E[N]=100$:

$$h(N) = \frac{1}{100}\left(\frac{99}{100}\right)^{N-1},$$

$N \in \mathbb{N}$. Observe $x = 203$.

$$f(x|N) = \begin{cases} \frac{1}{N}, & x \in \{1, \dots, N\} \\ 0, & x > N \end{cases}$$

6a. Find $h(N|x)$, $\mathbb{E}[N|x]$, $sd(N|x)$.

- $h(N|x) \propto f(x|n)h(N)$

$$= \begin{cases} \frac{1}{N} \cdot \frac{1}{100}\left(\frac{99}{100}\right)^{N-1}, & x \leq N \\ 0 \cdot \frac{1}{100}\left(\frac{99}{100}\right)^{N-1}, & x > N \end{cases}$$

$$\propto \begin{cases} \frac{1}{N}\left(\frac{99}{100}\right)^{N-1}, & N \geq x \\ 0, & N < x \end{cases}$$

$$= \begin{cases} \frac{c}{N}\left(\frac{99}{100}\right)^{N-1}, & N \geq x \\ 0, & N < x \end{cases} \quad \text{for some } c \in \mathbb{R}.$$

- In particular,

$$h(N|x=203) = \begin{cases} \frac{c}{N}\left(\frac{99}{100}\right)^{N-1}, & N \geq 203 \\ 0, & N < 203 \end{cases}.$$

- c must solve $\sum_{n=203}^{\infty} \frac{c}{n}\left(\frac{99}{100}\right)^{n-1} = 1$.
- $\sum_{n=203}^{\infty} \frac{c}{n}\left(\frac{99}{100}\right)^{n-1} = c\sum_{n=203}^{\infty} \frac{1}{n}\left(\frac{99}{100}\right)^{n}\left(\frac{99}{100}\right)^{-1}$

$$= \frac{100}{99} c \sum_{n=203}^{\infty} \frac{1}{n}\left(\frac{99}{100}\right)^{n}$$

- Note that $\int_0^{0.99} \sum_{n=203}^{\infty} t^{n-1}\,dt = \sum_{n=203}^{\infty}\int_0^{0.99} t^{n-1}\,dt$ (by Fubini's/Tonelli's Thms)

$$= \sum_{n=203}^{\infty} \frac{1}{n} t^n \Big|_0^{0.99}$$

$$= \sum_{n=203}^{\infty}\left(\frac{1}{n}\left(\frac{99}{100}\right)^n - \frac{1}{n}(0)\right)$$

$$= \sum_{n=203}^{\infty}\frac{1}{n}\left(\frac{99}{100}\right)^n.$$

- But for $t \in [0, \frac{99}{100}]$,

$$\sum_{n=203}^{\infty} t^{n-1} = \sum_{m=0}^{\infty} t^{m+203-1}$$

$$= \sum_{m=0}^{\infty} t^m t^{202}$$

$$= \frac{t^{202}}{1-t}$$

$\longrightarrow$

- So, $\sum_{n=203}^{\infty} \frac{1}{n}\left(\frac{99}{100}\right)^n = \int_0^{0.99} \sum_{n=203}^{\infty} t^{n-1}\,dt$

$$= \int_0^{0.99} \frac{t^{202}}{1-t}\,dt$$

$$\approx 0.047 \text{ (via WolframAlpha).}$$

- Thus, $\sum_{n=203}^{\infty} \frac{c}{n}\left(\frac{99}{100}\right)^{n-1} \approx \frac{100}{99}c(0.047) \approx 1$

$$\Rightarrow c \approx \frac{99}{100}/0.047$$

$$\Rightarrow c \approx 21.253$$

- So,

$$h(N|x=203) \approx \begin{cases} \frac{21.253}{N}\left(\frac{99}{100}\right)^{N-1}, & N \geq 203 \\ 0, & o/w \end{cases}$$

- $\mathbb{E}[N|x=203] = \sum_{n=203}^{\infty} n\,h(n|x=203)$

$$\approx \sum_{n=203}^{\infty} n \cdot \frac{21.253}{n}\left(\frac{99}{100}\right)^{n-1}$$

$$= 21.253 \sum_{n=203}^{\infty} \left(\frac{99}{100}\right)^{n-1}$$

$$= 21.253 \sum_{m=0}^{\infty} \left(\frac{99}{100}\right)^{203+m-1}$$

$$= 21.253 \sum_{m=0}^{\infty} \left(\frac{99}{100}\right)^{m}\left(\frac{99}{100}\right)^{202}$$

$$= 21.253 \left(\frac{99}{100}\right)^{202} \sum_{m=0}^{\infty} \left(\frac{99}{100}\right)^{m}$$

$$= 21.253 \left(\frac{99}{100}\right)^{202} \cdot \frac{1}{1-\frac{99}{100}}$$

$$\Rightarrow \mathbb{E}[N|x=203] \approx 279.089$$

- $Var(N|x=203) = \mathbb{E}[N^2|x=203] - \mathbb{E}[N|x=203]^2$
- $\mathbb{E}[N^2|x=203] = \sum_{n=203}^{\infty} n^2 h(n|x=203)$

$$\approx \sum_{n=203}^{\infty} n^2 \cdot \frac{1}{n}(21.253)\left(\frac{99}{100}\right)^{n-1}$$

$$= 21.253 \sum_{n=203}^{\infty} n\left(\frac{99}{100}\right)^{n-1}$$

- Note that $\sum_{n=203}^{\infty} n\left(\frac{99}{100}\right)^{n-1} = \sum_{m=0}^{\infty} (m+203)\left(\frac{99}{100}\right)^{203+m-1}$

$$= \sum_{m=0}^{\infty} (m+203)\left(\frac{99}{100}\right)^{m-1}\left(\frac{99}{100}\right)^{203}$$

$$= \left(\frac{99}{100}\right)^{203}\left[\sum_{m=0}^{\infty} m\left(\frac{99}{100}\right)^{m-1} + 203\sum_{m=0}^{\infty}\left(\frac{99}{100}\right)^{m}\right]$$

$$= \left(\frac{99}{100}\right)^{203}\left[\sum_{m=0}^{\infty} m\left(\frac{99}{100}\right)^{m-1} + 203 \cdot \frac{1}{1-0.99}\right]$$

- Now note that for $x \in (0,1)$,

$$\sum_{m=0}^{\infty} m x^{m-1} = \sum_{m=0}^{\infty} \frac{d}{dx} x^m$$

$$= \frac{d}{dx}\sum_{m=0}^{\infty} x^m \quad \text{since the series converges uniformly}$$

$$= \frac{d}{dx}\left(\frac{1}{1-x}\right)$$

$$= \frac{1}{(1-x)^2},$$

So $\sum_{m=0}^{\infty} m\left(\frac{99}{100}\right)^{m-1} = \frac{1}{(1-0.99)^2} = 10{,}000,$

so $\sum_{n=203}^{\infty} n\left(\frac{99}{100}\right)^{n-1} = \left(\frac{99}{100}\right)^{203}\left[10{,}000 + 20{,}300\right]$

$$\approx 3965.667,$$

so $\mathbb{E}[N^2 | x = 203] \approx (21.253)(3965.667)$

$$\approx 84{,}284.730.$$

· Thus,

$$Var(N|x=203) \approx 84{,}284.730 - (279.089)^2$$

$$\Rightarrow Var(N|x=203) \approx 6394.333$$

$$\Rightarrow sd(N|x=203) \approx \sqrt{6394.333}$$

$$\Rightarrow sd(N|x=203) \approx 79.965$$

Homework 3, Exercise 6(b)

Robert Winter

6(b) Find a 95% HPD credible interval for N (you do not need to match 95% exactly. Get as close as possible).

First, we construct the posterior probability mass function $h(N|x = 203)$, and generate some data $(203, h(203|x = 203)), \ldots (1000, h(1000|x = 203))$.

```
# Posterior PMF
post = function(n){
  c = 21.253
  if(n>=203){
    h = (c/n)*(0.99)^(n-1)
  }
  else{
    h = 0
  }
  return(h)
}
```

```
# Data for plot
ns = c(203:1000)
hs = c()
for(i in 1:length(ns)){
  hs[i] = post(ns[i])
}
data = cbind(ns, hs) %>% as.data.frame()
```

Since the posterior PMF is monotonically decreasing, it is unimodal, with its mode at the leftmost point of its support, $n = 203$. Since the HPD credible interval for a unimodal distribution contains the mode, it follows that the HPD interval begins at $n = 203$. After some trial and error with the rightmost endpoint of the HPD interval, I found that the interval $(203, 438)$ contains approximately 95.02% of the mass of the distribution, making it the (best possible approximate) 95% HPD credible interval.

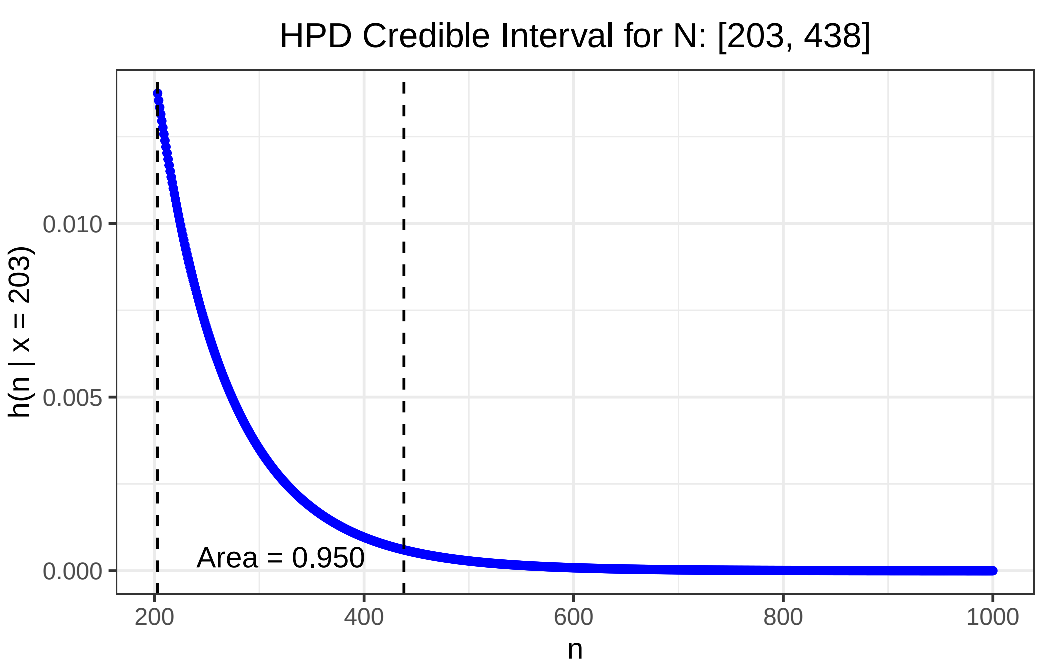
```
# HPD interval
L = 203
R = 438

area = 0
for(n in L:R){
  area = area + post(n)
}
area # 0.9502
```

```
[1] 0.9502448
```

We visualize the 95% HPD credible interval below.

```
ggplot(data, aes(x=ns, y=hs)) +
  theme_bw() +
  geom_point(size=1, col = "blue") +
  xlab("n") +
  ylab("h(n | x = 203)") +
  geom_vline(xintercept = c(L,R), linetype = "dashed") +
  ggtitle("HPD Credible Interval for N: [203, 438]") +
  theme(plot.title = element_text(hjust = 0.5)) +
  annotate("text", x = 0.5*(L+R), y = 0.0004, label = "Area = 0.950")
```



7. Poisson sampling model $f(n|\lambda) = \text{Pois}(\lambda)$. Find Jeffrey's prior for λ.

- $f(n|\lambda) = \dfrac{\lambda^n e^{-\lambda}}{n!}$

- $\Rightarrow \log f(n|\lambda) = \log\left(\dfrac{\lambda^n e^{-\lambda}}{n!}\right)$

$$= \log(\lambda^n) + \log(e^{-\lambda}) - \log(n!)$$
$$= n\log\lambda - \lambda - \log(n!)$$

- $\Rightarrow \dfrac{\partial \log f(n|\lambda)}{\partial \lambda} = \dfrac{n}{\lambda} - 1$

- $\Rightarrow \dfrac{\partial^2 \log f(n|\lambda)}{\partial \lambda^2} = \dfrac{-n}{\lambda^2}$

- $\Rightarrow I(\lambda) = -\mathbb{E}\left[\dfrac{\partial^2 \log f(n|\lambda)}{\partial \lambda^2}\right]$

$$= -\mathbb{E}\left[\frac{-n}{\lambda^2}\right]$$
$$= \frac{1}{\lambda^2}\mathbb{E}[n]$$
$$= \frac{1}{\lambda^2}\cdot\lambda$$
$$= \frac{1}{\lambda}$$

- $\Rightarrow \pi_J(\lambda) \propto I(\lambda)^{1/2}$

$$= (1/\lambda)^{1/2}$$

- Thus, Jeffrey's prior is

$$\pi_j(\lambda) \propto \frac{1}{\sqrt{\lambda}}$$

Note that this is an improper prior.

8. Coin. θ is prob of heads. $H_1: \theta < \frac{1}{2}$, $H_2: \theta = \frac{1}{2}$, $H_3: \theta > \frac{1}{2}$. $p(H_2) = \frac{1}{2}$, $p(H_1) = p(H_3) = \frac{1}{4}$. $\theta | H_1 \sim U(0, \frac{1}{2})$, $\theta | H_3 \sim U(\frac{1}{2}, 1)$. Observe $n=1$ toss. $x_1 \in \{0, 1\}$ indicator of heads.

8a. Find $p(x_1=1 \mid H_1)$, $p(x_1=1 \mid H_2)$, $p(x_1=1 \mid H_3)$.

- $h(\theta|H_1) = \begin{cases} \frac{1}{\frac{1}{2}-0}, & \theta \in (0, \frac{1}{2}) \\ 0, & o/w \end{cases} = \begin{cases} 2, & \theta \in (0, \frac{1}{2}) \\ 0, & o/w \end{cases}$

- $h(\theta|H_2) = \begin{cases} 1, & \theta = \frac{1}{2} \\ 0, & o/w \end{cases}$

- $h(\theta|H_3) = \begin{cases} \frac{1}{1-\frac{1}{2}}, & \theta \in (\frac{1}{2}, 1) \\ 0, & o/w \end{cases} = \begin{cases} 2, & \theta \in (\frac{1}{2}, 1) \\ 0, & o/w \end{cases}$

- $$\begin{aligned} p(x_1=1 \mid H_1) &= \int_{\theta \in H_1} p(x_1=1 \mid \theta) h(\theta|H_1) d\theta \\ &= \int_0^{1/2} \theta \cdot 2 d\theta \\ &= 2 \cdot \tfrac{1}{2} \theta^2 \Big|_0^{1/2} \\ &= (\tfrac{1}{2})^2 - 0^2 \\ &= \tfrac{1}{4} \end{aligned}$$

- $$\begin{aligned} p(x_1=0 \mid H_1) &= \int_{\theta \in H_1} p(x_1=0 \mid \theta) h(\theta|H_1) d\theta \\ &= \int_0^{1/2} (1 - p(x_1=1 \mid \theta)) 2 d\theta \\ &= 2 \int_0^{1/2} (1-\theta) d\theta \\ &= 2(\theta - \tfrac{1}{2}\theta^2) \Big|_0^{1/2} \\ &= 2[(\tfrac{1}{2} - \tfrac{1}{2}(\tfrac{1}{2})^2) - (0 - \tfrac{1}{2}0^2)] \\ &= 2[\tfrac{1}{2} - \tfrac{1}{8}] \\ &= \tfrac{3}{4} \end{aligned}$$

- So, $p(x_1|H_1) = \begin{cases} \frac{1}{4}, & x_1 = 1 \\ \frac{3}{4}, & x_1 = 0 \end{cases}$

and in particular, $p(x_1=1 \mid H_1) = \frac{1}{4}$

- $p(x_1=1 \mid H_2) = p(x_1=1 \mid \theta=\frac{1}{2})h(\theta=\frac{1}{2} \mid H_2)$
$= (\frac{1}{2})(1)$
$= \frac{1}{2}$
- $p(x_1=0 \mid H_2) = p(x_1=0 \mid \theta=\frac{1}{2})h(\theta=\frac{1}{2} \mid H_2)$
$= (1-p(x_1=1 \mid \theta=\frac{1}{2}))(1)$
$= (1-\frac{1}{2})$
$= \frac{1}{2}$
- So, $p(x_1 \mid H_2) = \begin{cases} \frac{1}{2}, & x_1=1 \\ \frac{1}{2}, & x_1=0 \end{cases}$

and in particular,
$p(x_1=1 \mid H_2) = \frac{1}{2}$

- $p(x_1=1 \mid H_3) = \int_{\theta \in H_3} p(x_1=1 \mid \theta)h(\theta \mid H_3)d\theta$
$= \int_{1/2}^{1} \theta \cdot 2 d\theta$
$= 2 \cdot \frac{1}{2}\theta^2 \Big|_{1/2}^{1}$
$= (1^2 - (\frac{1}{2})^2)$
$= \frac{3}{4}$
- $p(x_1=0 \mid H_3) = \int_{\theta \in H_3} p(x_1=0 \mid \theta)h(\theta \mid H_3)d\theta$
$= \int_{1/2}^{1} (1-p(x_1=1 \mid \theta))2d\theta$
$= 2\int_{1/2}^{1} (1-\theta)d\theta$
$= 2(\theta - \frac{1}{2}\theta^2)\Big|_{1/2}^{1}$
$= 2[(1-\frac{1}{2}(1)^2) - (\frac{1}{2} - \frac{1}{2}(\frac{1}{2})^2)]$
$= 2[\frac{1}{2} - \frac{1}{2} + \frac{1}{8}]$
$= \frac{1}{4}$
- So, $p(x_1 \mid H_3) = \begin{cases} \frac{3}{4}, & x_1=1 \\ \frac{1}{4}, & x_1=0 \end{cases}$

and in particular,
$p(x_1=1 \mid H_3) = \frac{3}{4}$

8b. Find $p(x_1=1)$.

- By the Law of Total Probability,

$$p(x_1=1) = p(x_1=1 \mid H_1)p(H_1) + p(x_1=1 \mid H_2)p(H_2) + p(x_1=1 \mid H_3)p(H_3)$$
$$= \frac{1}{4}\left(\frac{1}{4}\right) + \frac{1}{2}\left(\frac{1}{2}\right) + \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)$$
$$= \frac{1}{16} + \frac{1}{4} + \frac{3}{16}$$
$$= \frac{1}{2}$$

- Similarly,

$$p(x_1=0) = p(x_1=0 \mid H_1)p(H_1) + p(x_1=0 \mid H_2)p(H_2) + p(x_1=0 \mid H_3)p(H_3)$$
$$= \frac{3}{4}\left(\frac{1}{4}\right) + \frac{1}{2}\left(\frac{1}{2}\right) + \frac{1}{4}\left(\frac{1}{4}\right)$$
$$= \frac{3}{16} + \frac{1}{4} + \frac{1}{16}$$
$$= \frac{1}{2}$$

- So, $p(x_1) = \begin{cases} \frac{1}{2}, & x_1 = 1 \\ \frac{1}{2}, & x_1 = 0 \end{cases}$

and in particular, $p(x_1=1) = \frac{1}{2}$

9. Rayleigh dist. w/ PDF $f(x|\delta) = \delta x e^{-\delta x^2/2} \mathbb{I}_{(0,\infty)}(x)$.
$x = \{x_1, \ldots, x_n\}$ is a realization of a random smpl. from this model. Prior dist. of δ is Gamma(a, b).

9a. Find posterior $h(\delta|x)$. Name? $\mathbb{E}[\delta|x]$? $Var(\delta|x)$?

- $\pi(\delta) = \frac{b^a}{\Gamma(a)} \delta^{a-1} e^{-b\delta}$
- $h(\delta|x) \propto \prod_{i=1}^n \delta x_i e^{-\delta x_i^2/2} \cdot \frac{b^a}{\Gamma(a)} \delta^{a-1} e^{-b\delta}$

$$= \delta^n \cdot \prod_{i=1}^n x_i \cdot e^{-\frac{\delta}{2}\sum_{i=1}^n x_i^2} \cdot \frac{b^a}{\Gamma(a)} \delta^{a-1} e^{-b\delta}$$

$$\propto \delta^{n+a-1} e^{-b\delta - \frac{1}{2}\delta \sum_{i=1}^n x_i^2}$$

$$= \delta^{n+a-1} e^{-\delta(b + \frac{1}{2}\sum_{i=1}^n x_i^2)}$$

- Thus, $\delta|x \sim \text{Gamma}(a+n, b + \frac{1}{2}\sum_{i=1}^n x_i^2)$

and $h(\delta|x) = \frac{(b + \frac{1}{2}\sum_{i=1}^n x_i^2)^{a+n}}{\Gamma(a+n)} \delta^{a+n-1} e^{-\delta(b+\frac{1}{2}\sum_{i=1}^n x_i^2)}$

- Using the moments of the Gamma distribution,
  - $\mathbb{E}[\delta|x] = \frac{a+n}{b + \frac{1}{2}\sum_{i=1}^n x_i^2}$

  and

  - $Var(\delta|x) = \frac{a+n}{(b + \frac{1}{2}\sum_{i=1}^n x_i^2)^2}$

9b. Find Jeffrey's prior & corresponding posterior dist conditional on one obs. x from Rayleigh model.

- $f(x|\delta) = \delta x e^{-\delta x^2/2}$

$\Rightarrow \log f(x|\delta) = \log(\delta x e^{-\delta x^2/2})$

$= \log\delta + \log x - \frac{1}{2}\delta x^2$

$\Rightarrow \frac{\partial \log f(x|\delta)}{\partial \delta} = \frac{1}{\delta} - \frac{1}{2}x^2$

$\Rightarrow \frac{\partial^2 \log f(x|\delta)}{\partial \delta^2} = \frac{-1}{\delta^2}$

$\Rightarrow I(\delta) = -\mathbb{E}\left[\frac{\partial^2 \log f(x|\delta)}{\partial \delta^2}\right]$

$= -\mathbb{E}\left[\frac{-1}{\delta^2}\right]$

$= \frac{1}{\delta^2}$

$\Rightarrow h_J(\delta) = (1/\delta^2)^{1/2} = \frac{1}{\delta}$

- That is, Jeffrey's prior is

$h_J(\delta) = \frac{1}{\delta}$

Note that this is an improper prior.

- $f(\delta|x) \propto h_J(\delta) \cdot f(x|\delta)$

$= \frac{1}{\delta} \cdot \delta x e^{-\delta x^2/2}$

$= x e^{-\delta x^2/2}$

$\propto e^{-\frac{1}{2}x^2\delta}$

- Thus, $\delta|x \sim Exp(\frac{1}{2}x^2)$, and

$f(\delta|x) = \frac{1}{2}x^2 e^{-\frac{1}{2}x^2\delta}$