BIOS 779 HW1

Mingwei Fei

January 17, 2023

1 Problem 1- Carlin and Louis

Diagnostic test, represent $P(D = d|T = t)$ in terms of test sensitivity, $P(T = 1|D = 1)$. Specificity, $P(T = 0|D = 0)$. and disease prevalence, $P(D = 1)$, and relate to Bayes' theorem.

$$\begin{aligned}
P(D = d|T = t) &= \frac{P(D = d, T = t)}{P(T = t)} \\
&= \frac{P(D = d, T = t)}{\sum_{D=d} P(T = t, D = d)P(D = d)} \\
&= \frac{P(T = t|D = d)P(D = t)}{P(T = t, D = 0)P(D = 0) + P(T = t, D = 1)P(D = 1)} \\
&= \frac{P(T = t|D = d)P(D = d)}{P(T = t|D = 0)P(D = 0)^2 + P(T = t|D = 1)P(D = 1)^2}
\end{aligned}$$

So we have

$$
\begin{aligned}
P(D=1|T=1) &= \frac{P(T=1|D=1)P(D=1)}{P(T=1|D=0)(1-P(D=1))^2+P(T=1|D=1)P(D=1)^2} \\
&= \frac{P(T=1|D=1)P(D=1)}{(1-P(T=0|D=0))(1-P(D=1))^2+P(T=1|D=1)P(D=1)^2} \\
P(D=0|T=1) &= \frac{P(T=1|D=0)(1-P(D=1))}{P(T=1|D=0)(1-P(D=1))^2+P(T=1|D=1)P(D=1)^2} \\
&= \frac{(1-P(T=0|D=0))(1-P(D=1))}{(1-P(T=0|D=0))(1-P(D=1))^2+P(T=1|D=1)P(D=1)^2} \\
P(D=0|T=0) &= \frac{P(T=0|D=0)(1-P(D=1))}{P(T=0|D=0)(1-P(D=1))^2+P(T=0|D=1)P(D=1)^2} \\
&= \frac{P(T=0|D=0)(1-P(D=1))}{P(T=0|D=0)(1-P(D=1))^2+(1-P(T=1|D=1))P(D=1)^2} \\
P(D=1|T=0) &= \frac{P(T=0|D=1)P(D=1)}{P(T=0|D=0)(1-P(D=1))^2+P(T=0|D=1)P(D=1)^2} \\
&= \frac{(1-P(T=1|D=1))P(D=1)}{P(T=0|D=0)(1-P(D=1))^2+(1-P(T=1|D=1))P(D=1)^2}
\end{aligned}
$$

2 Problem 2

Suppose $X_1, .., X_n$ is a random sample from X, where X has density

$$
p(x|r,\theta) = \begin{cases} \binom{r+x-1}{x}\theta^r(1-\theta)^x & x = 0,1,2... \\ 0 & \text{otherwise} \end{cases} \tag{1}
$$

where $0 < \theta < 1, r$ is known, and θ is unknown.

(a) Derive Jeffreys's prior for θ.

$$
\begin{aligned}
p(\theta) &\propto \sqrt{detI(\theta)} \\
logp(x|r,\theta) &= log\begin{pmatrix} r+x-1 \\ x \end{pmatrix} + rlog\theta + xlog(1-\theta) \\
\frac{\partial logp(x|r,\theta)}{\partial\theta} &= \frac{r}{\theta} - \frac{x}{1-\theta} \\
\frac{\partial^2 logp(x|r,\theta)}{\partial\theta\,\partial\theta} &= -\frac{r}{\theta^2} - \frac{x}{(1-\theta)^2} \\
I(\theta) &= -E[\frac{\partial^2 logp(x|r,\theta)}{\partial\theta\,\partial\theta}] = \frac{r}{\theta^2} + \frac{E[X]}{(1-\theta)^2}
\end{aligned}
$$

The $p(x|r,\theta)$ is negative binomial distribution, which we have

$$E(X) = \frac{r(1-\theta)}{\theta}$$
$$I(\theta) = \frac{r}{\theta^2(1-\theta)}$$

Thus, Jeffery's prior is

$$\pi(\theta) \propto I(\theta)^{1/2} = [\frac{r}{\theta^2(1-\theta)}]^{1/2}$$

(b) Suppose the prior for θ is $\theta \sim beta(\alpha_0, \lambda_0)$, where α_0, λ_0 are specified hyperparameters. Derive $E(X)$ and $Var(X)$.

$$\pi(\theta) = \frac{1}{B(\alpha_0, \lambda_0)}\theta^{\alpha_0 - 1}(1-\theta)^{\lambda_0 - 1}$$

While from part (a), we know that $x|\theta \sim NB(r, \theta)$, then

$$E(X|\theta) = \frac{r(1-\theta)}{\theta}$$
$$Var(X|\theta) = \frac{r(1-\theta)}{\theta^2}$$

So

$$\begin{aligned}
E(X) &= E[E(X|\theta)] = E[\frac{r(1-\theta)}{\theta}] \\
&= \int_0^1 \frac{r(1-\theta)}{\theta}\frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0)\Gamma(\lambda_0)}\theta^{\alpha_0-1}(1-\theta)^{\lambda_0-1}d\theta \\
&= r\frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0)\Gamma(\lambda_0)}\int_0^1 \theta^{\alpha_0-2}(1-\theta)^{\lambda_0}d\theta \\
&= r\frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0)\Gamma(\lambda_0)}\frac{\Gamma(\alpha_0-1)\Gamma(\lambda_0+1)}{\Gamma(\alpha_0+\lambda_0)}\int_0^1 \frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0-1)\Gamma(\lambda_0+1)}\theta^{\alpha_0-2}(1-\theta)^{\lambda_0}d\theta \\
&= r\frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0)\Gamma(\lambda_0)}\frac{\Gamma(\alpha_0-1)\Gamma(\lambda_0+1)}{\Gamma(\alpha_0+\lambda_0)} \\
&= \frac{r\lambda_0}{\alpha_0-1}, \qquad \alpha_0 > 1
\end{aligned}$$

For variance,

$$\begin{aligned}
Var(X) &= E[Var(X|\theta)] + Var[E[X|\theta]] \\
&= E[\frac{r(1-\theta)}{\theta^2}] + Var[\frac{r(1-\theta)}{\theta}] \\
E[\frac{r(1-\theta)}{\theta^2}] &= \int_0^1 \frac{r(1-\theta)}{\theta^2} \frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0)\Gamma(\lambda_0)} \theta^{\alpha_0-1}(1-\theta)^{\lambda_0-1} d\theta \\
&= r\frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0)\Gamma(\lambda_0)} \int_0^1 \theta^{\alpha_0-3}(1-\theta)^{\lambda_0} d\theta \\
&= r\frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0)\Gamma(\lambda_0)} \frac{\Gamma(\alpha_0-2)\Gamma(\lambda_0+1)}{\Gamma(\alpha_0+\lambda_0-1)} \int_0^1 \frac{\Gamma(\alpha_0+\lambda_0-1)}{\Gamma(\alpha_0-2)\Gamma(\lambda_0+1)} \theta^{\alpha_0-3}(1-\theta)^{\lambda_0} d\theta \\
&= r\frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0)\Gamma(\lambda_0)} \frac{\Gamma(\alpha_0-2)\Gamma(\lambda_0+1)}{\Gamma(\alpha_0+\lambda_0-1)} \\
&= \frac{r\lambda_0(\alpha_0+\lambda_0-1)}{(\alpha_0-1)(\alpha_0-2)}, \qquad \alpha_0 > 2 \\
E[[\frac{r(1-\theta)}{\theta}]^2] &= \int_0^1 [\frac{r(1-\theta)}{\theta}]^2 \frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0)\Gamma(\lambda_0)} \theta^{\alpha_0-1}(1-\theta)^{\lambda_0-1} d\theta \\
&= r^2\frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0)\Gamma(\lambda_0)} \int_0^1 \theta^{\alpha_0-3}(1-\theta)^{\lambda_0+1} d\theta \\
&= r^2\frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0)\Gamma(\lambda_0)} \frac{\Gamma(\alpha_0-2)\Gamma(\lambda_0-2)}{\Gamma(\alpha_0+\lambda_0)} \int_0^1 \frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0-2)\Gamma(\lambda_0-2)} \theta^{\alpha_0-3}(1-\theta)^{\lambda_0+1} d\theta \\
&= \frac{r^2\lambda_0(\lambda_0+1)}{(\alpha_0-1)(\alpha_0-2)}, \qquad \alpha_0 > 2
\end{aligned}$$

So we have

$$\begin{aligned}
Var(\frac{r(1-\theta)}{\theta}) &= E(\frac{r(1-\theta)}{\theta})^2 - [E(\frac{r(1-\theta)}{\theta})]^2 \\
&= \frac{r^2\lambda_0(\lambda_0+1)}{(\alpha_0-1)(\alpha_0-2)} - [\frac{r\lambda_0}{\alpha_0-1}]^2 \\
&= \frac{r^2\lambda_0(\lambda_0+\alpha_0-1)}{(\alpha_0-1)^2(\alpha_0-2)}, \qquad \alpha_0 > 2
\end{aligned}$$

Thus,

$$\begin{aligned}
Var(X) &= E[\frac{r(1-\theta)}{\theta^2}] + Var(\frac{r(1-\theta)}{\theta}) \\
&= \frac{r\lambda_0(\alpha_0+\lambda_0-1)}{(\alpha_0-1)(\alpha_0-2)} + \frac{r^2\lambda_0(\lambda_0+\alpha_0-1)}{(\alpha_0-1)^2(\alpha_0-2)} \\
&= \frac{r\lambda_0(\alpha_0+\lambda_0-1)(\alpha_0+r-1)}{(\alpha_0-1)^2(\alpha_0-2)}, \qquad \alpha_0 > 2
\end{aligned}$$

(c) Suppose the prior for θ is $\theta \sim beta(\alpha_0, \lambda_0)$, where α_0, λ_0 are specified hyperparameters. Derive the posterior distribution of θ.

$$\pi(\theta|X) \propto \frac{\Gamma(\alpha_0+\lambda_0)}{\Gamma(\alpha_0)\Gamma(\lambda_0)}\theta^{\alpha_0-1}(1-\theta)^{\lambda_0-1}\prod_{i=1}^{n}\binom{r+x_i-1}{x_i}\theta^r(1-\theta)^{x_i}, \qquad \theta \in (0,1)$$

because beta distribution is a conjugate of the negative binomial distribution, we have the posterior distribution of $\theta|X$ also a negative binomial

$$\pi(\theta|X) = \binom{nr+\alpha_0+\sum_i x_i+\lambda_0-3}{\sum_i x_i+\lambda_0-1}\theta^{nr+\alpha_0-1}(1-\theta)^{\sum_i x_i+\lambda_0-1}$$

(d) Suppose the prior for θ is $\theta \sim beta(\alpha_0, \lambda_0)$, where α_0, λ_0 are specified hyperparameters. Derive the posterior predictive distribution of $z = (z_1, z_2)$, where conditional on θ, z_1, z_2 are i.i.d with distribution of (1).

The predictive distribution

$$p(z|X) = \int p(z|\theta)p(\theta|X)d\theta$$

As $z = (z_1, z_2)$, where conditional on θ, z_1, z_2 are i.i.d with distribution of (1), so

$$p(z|\theta) = p(z_1|\theta)p(z_2|\theta) = \binom{r+z_1-1}{z_1}\binom{r+z_2-1}{z_2}\theta^{2r}(1-\theta)^{z_1+z_2}$$

$$\begin{aligned}p(z|X) = &\int_0^1 \binom{r+z_1-1}{z_1}\binom{r+z_2-1}{z_2}\theta^{2r}(1-\theta)^{z_1+z_2}\binom{nr+\alpha_0+\sum_i x_i+\lambda_0-3}{\sum_i x_i+\lambda_0-1}\\ &\theta^{nr+\alpha_0-1}(1-\theta)^{\sum_i x_i+\lambda_0-1}d\theta\\ &= \binom{r+z_1-1}{z_1}\binom{r+z_2-1}{z_2}\binom{nr+\alpha_0+\sum_i x_i+\lambda_0-3}{\sum_i x_i+\lambda_0-1}\\ &\int_0^1 \theta^{2r}(1-\theta)^{z_1+z_2}\theta^{nr+\alpha_0-1}(1-\theta)^{\sum_i x_i+\lambda_0-1}d\theta\end{aligned}$$

We can construct the negative binomial distribution for

$$\begin{aligned}&\int_0^1 \theta^{2r}(1-\theta)^{z_1+z_2}\theta^{nr+\alpha_0-1}(1-\theta)^{\sum_i x_i+\lambda_0-1}d\theta\\ &= \frac{1}{A}\int_0^1 \binom{(n+2)r+\alpha_0+\sum_i x_i+\lambda_0+z_1+z_2-3}{\sum_i x_i+\lambda_0-1}\theta^{(n+2)r+\alpha_0-1}(1-\theta)^{\sum_i x_i+\lambda_0+z_1+z_2-1}d\theta\end{aligned}$$

let $A = \binom{(n+2)r+\alpha_0+\sum_i x_i+\lambda_0+z_1+z_2-3}{\sum_i x_i+\lambda_0-1}$, so

$$p(z|\theta) = \binom{r+z_1-1}{z_1}\binom{r+z_2-1}{z_2}\binom{nr+\alpha_0+\sum_i x_i+\lambda_0-3}{\sum_i x_i+\lambda_0-1}\frac{1}{\binom{(n+2)r+\alpha_0+\sum_i x_i+\lambda_0+z_1+z_2-3}{\sum_i x_i+\lambda_0-1}}$$

3 Problem 3

We have distribution

$$\pi(\theta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-(\alpha+1)}exp(-\frac{\beta}{\theta}), & \theta > 0 \\ 0 & \theta \leq 0 \end{cases} \tag{2}$$

(a) Verify that $\pi(\theta)$ is a probability density function.

We need to show that $\int \pi(\theta)d\theta = 1$.

$$\begin{aligned}
\int_0^\infty \pi(\theta)d\theta &= \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-(\alpha+1)}exp(-\frac{\beta}{\theta})d\theta \\
\text{let} x &= \frac{1}{\theta}, J = |\frac{\partial\theta}{\partial x}| = \frac{1}{x^2} \\
&= \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha+1}exp(-\beta x)\frac{1}{x^2}dx \\
&= \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}exp(-\beta x)dx \\
&= 1
\end{aligned}$$

(b) Consider family of prior distribution, show it is a conjugate prior for normal family with a known value of mean μ, and unknown value of variance θ.

We need to show that the posterior distribution of $p(\theta|x)$ is also a gamma distribution, which follows the distribution (2).

$$\begin{aligned}
p(x_i|\mu,\theta) &= \frac{1}{\sqrt{2\pi\theta}}exp(-\frac{(x_i-\mu)^2}{2\theta}) \\
p(\theta|X) &= \frac{\prod_{i=1}^n f(x_i|\mu,\theta)f(\theta)}{f(X)} \\
&= \frac{(2\pi\theta)^{-n/2}exp(-\frac{\sum_{i=1}(x_i-\mu)^2}{2\theta})\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-(\alpha+1)}exp(-\frac{\beta}{\theta})}{f(X)}
\end{aligned}$$

To calculate $f(X) = \int f(x|\theta)f(\theta)d\theta$,

$$f(X) = \int_0^\infty (2\pi\theta)^{-n/2} exp(-\frac{\sum_{i=1}(x_i-\mu)^2}{2\theta})\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-(\alpha+1)}exp(-\frac{\beta}{\theta})d\theta$$
$$= \int_0^\infty (2\pi)^{-n/2}\frac{\beta^\alpha}{\Gamma(\alpha)}exp(-\frac{\sum_{i=1}(x_i-\mu)^2+2\beta}{2\theta})\theta^{-(\alpha+1+\frac{n}{2})}d\theta$$
$$\text{let} y = \frac{1}{\theta}, \qquad J = |\frac{\partial\theta}{\partial y}| = \frac{1}{y^2}$$
$$= \int_0^\infty (2\pi)^{-n/2}\frac{\beta^\alpha}{\Gamma(\alpha)}exp(-\frac{\sum_{i=1}(x_i-\mu)^2+2\beta}{2}y)y^{(\alpha+1+\frac{n}{2})}y^{-2}dy$$
$$= \int_0^\infty (2\pi)^{-n/2}\frac{\beta^\alpha}{\Gamma(\alpha)}exp(-\frac{\sum_{i=1}(x_i-\mu)^2+2\beta}{2}y)y^{(\alpha+\frac{n}{2})-1}dy$$

Further,

$$f(X) = (2\pi)^{-n/2}\frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha+\frac{n}{2})}{[\frac{\sum_{i=1}(x_i-\mu)^2+2\beta}{2}]^{\alpha+\frac{n}{2}}}$$
$$\int_0^\infty \frac{[\frac{\sum_{i=1}(x_i-\mu)^2+2\beta}{2}]^{\alpha+\frac{n}{2}}}{\Gamma(\alpha+\frac{n}{2})}exp(-\frac{\sum_{i=1}(x_i-\mu)^2+2\beta}{2}y)y^{(\alpha+\frac{n}{2})-1}dy$$
$$= (2\pi)^{-n/2}\frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha+\frac{n}{2})}{[\frac{\sum_{i=1}(x_i-\mu)^2+2\beta}{2}]^{\alpha+\frac{n}{2}}}$$

Then the posterior distribution

$$p(x_i|\mu,\theta) = \frac{(2\pi)^{-n/2}\frac{\beta^\alpha}{\Gamma(\alpha)}exp(-\frac{\sum_{i=1}(x_i-\mu)^2+2\beta}{2\theta})\theta^{-(\alpha+1+\frac{n}{2})}}{(2\pi)^{-n/2}\frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha+\frac{n}{2})}{[\frac{\sum_{i=1}(x_i-\mu)^2+2\beta}{2}]^{\alpha+\frac{n}{2}}}}$$
$$= \frac{[\frac{\sum_{i=1}(x_i-\mu)^2+2\beta}{2}]^{\alpha+\frac{n}{2}}}{\Gamma(\alpha+\frac{n}{2})}exp(-\frac{\sum_{i=1}(x_i-\mu)^2+2\beta}{2\theta})\theta^{-(\alpha+1+\frac{n}{2})}$$

The above distribution follows the distribution (2), so gamma distribution is a conjugate prior for a normal family with unknown variance.

4 Problem 4

Suppose that the lifetime X of a medical device follows a shifted exponential distribution

$$p(x|\theta,\mu) = \theta exp(-\theta(x-\mu))I(x>\mu)$$

where $\theta > 0, -\infty < \mu < \infty$. Suppose that $X_1, X_2, ...X_n$ is a random sample from X.

(a) Derive the posterior mean and variance.

As we know that the gamma distribution is a conjugate prior for exponential distribution. So the posterior distribution is also a gamma distribution.

$$
\begin{aligned}
\theta \sim gamma(a_0, b_0) &= \frac{b_0^{a_0}}{\Gamma(a_0)}\theta^{a_0-1}exp(-b_0 x) \\
p(x_i|\theta,\mu) &= \theta exp(-\theta(x_i-\mu))I(x_i>\mu) \\
p(\theta|X) &\propto p(X|\theta)p(\theta) \\
&= \prod_{i=1}^{n}\theta exp(-\theta(x_i-\mu))I(x_i>\mu)\frac{b_0^{a_0}}{\Gamma(a_0)}\theta^{a_0-1}exp(-b_0\theta) \\
&= \theta^{a_0+n-1}exp[-\theta\sum_{i=1}^{n}(x_i-\mu)-b_0\theta]
\end{aligned}
$$

Thus, the posterior distribution of θ

$$
p(\theta|X) \sim Gamma(a_0+n, \sum_{i=1}^{n}(x_i-\mu)+b_0), \qquad X_{(1)}>\mu
$$

Then the

$$
\begin{aligned}
E(\theta|X) &= \frac{a_0+n}{\sum_{i=1}^{n}(x_i-\mu)+b_0} \\
Var(\theta|X) &= \frac{a_0+n}{(\sum_{i=1}^{n}(x_i-\mu)+b_0)^2}
\end{aligned}
$$

(b) Suppose that (μ,θ) are both unknown, $\mu \sim N(\mu_0,\sigma_0^2), \theta \sim gamma(a_0,b_0), (\mu,\theta)$, are assumed independent a priori.

(i) Derive the joint posterior distribution of (μ,θ) and express it in the simplest possible form. Note that the normalizing constant of the joint posterior does not have an analytic closed form, but can be expressed as an expectation.
Joint posterior distribution of (μ,θ), in which (μ,θ) are independent priori.

$$
\begin{aligned}
p(\mu,\theta|X) &= \frac{p(X|\mu,\theta)p(\theta)p(\mu)}{p(X)} \\
p(X|\mu,\theta)p(\theta)p(\mu) &= \prod_{i=1}^{n}\theta exp(-\theta(x_i-\mu))I(x_i>\mu)\frac{b_0^{a_0}}{\Gamma(a_0)}\theta^{a_0-1}exp(-b_0\theta)\frac{1}{\sqrt{2\pi}\sigma_0}exp(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2} \\
&= \frac{1}{\sqrt{2\pi}\sigma_0}\frac{b_0^{a_0}}{\Gamma(a_0)}\theta^{a_0+n-1}exp[-\theta(\sum_{i=1}^{n}(x_i-\mu)+b_0)]exp(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2})
\end{aligned}
$$

The normalizing constant

$$
\begin{aligned}
p(X) &= \int\int p(\theta,\mu,x)d\theta d\mu \\
&= \int_0^\infty \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}\sigma_0}\frac{b_0^{a_0}}{\Gamma(a_0)}\theta^{a_0+n-1}exp[-\theta(\sum_{i=1}^n(x_i-\mu)+b_0)]exp(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2})d\mu d\theta \\
&= \frac{1}{\sqrt{2\pi}\sigma_0}\frac{b_0^{a_0}}{\Gamma(a_0)}\int_{-\infty}^\infty\Big[\int_0^\infty \theta^{a_0+n-1}exp[-\theta(\sum_{i=1}^n(x_i-\mu)+b_0)]d\theta\Big]exp(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2})d\mu \\
&= \frac{1}{\sqrt{2\pi}\sigma_0}\frac{b_0^{a_0}}{\Gamma(a_0)}\frac{\Gamma(a_0+n)}{[\sum(x_i-\mu)+b_0]^{a_0+n}}\int_{-\infty}^\infty exp(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2})d\mu \\
&= \frac{b_0^{a_0}\Gamma(a_0+n)}{\Gamma(a_0)}\int_{-\infty}^\infty \frac{1}{[\sum(x_i-\mu)+b_0]^{a_0+n}}\frac{1}{\sqrt{2\pi}\sigma_0}exp(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2})d\mu \\
&= \frac{b_0^{a_0}\Gamma(a_0+n)}{\Gamma(a_0)}E_\mu\Big[\frac{1}{[\sum(x_i-\mu)+b_0]^{a_0+n}}\Big]
\end{aligned}
$$

Thus the posterior distribution of $(\mu,\theta|X)$

$$
\begin{aligned}
p(\mu,\theta|X) &= \frac{\frac{1}{\sqrt{2\pi}\sigma_0}\frac{b_0^{a_0}}{\Gamma(a_0)}\theta^{a_0+n-1}exp[-\theta(\sum_{i=1}^n(x_i-\mu)+b_0)]exp(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2})}{\frac{b_0^{a_0}\Gamma(a_0+n)}{\Gamma(a_0)}E_\mu\Big[\frac{1}{[\sum(x_i-\mu)+b_0]^{a_0+n}}\Big]} \\
&= \frac{1}{\sqrt{2\pi}\sigma_0}\frac{\theta^{a_0+n-1}}{\Gamma(a_0+n)E_\mu\Big[\frac{1}{[\sum(x_i-\mu)+b_0]^{a_0+n}}\Big]}exp[-\theta(\sum_{i=1}^n(x_i-\mu)+b_0)]exp(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2})
\end{aligned}
$$

(ii) Obtain the limiting joint posterior distribution $\sigma_0\to\infty$.
As $\sigma_0\to\infty$, the posterior distribution will be a gamma distribution,

$$
\begin{aligned}
p(\mu,\theta|X) \to & \frac{\theta^{a_0+n-1}}{\Gamma(a_0+n)E_\mu\Big[\frac{1}{[\sum(x_i-\mu)+b_0]^{a_0+n}}\Big]}exp[-\theta(\sum_{i=1}^n(x_i-\mu)+b_0)] \\
& \sim Gamma(a_0+n,\sum(x_i-\mu)+b_0)
\end{aligned}
$$

5 Problem 5

Let $U(a,b)$ denote the uniform distribution on the interval (a,b).

(a) $U(\theta,\theta+1)$, find posterior distribution of θ

$$\theta < X_1 ... X_n < \theta + 1$$

$$\begin{aligned}
p(\theta) &= \frac{1}{b_0 - a_0} \\
p(\theta|X) &= \frac{p(x|\theta)p(\theta)}{p(X)} \\
p(x|\theta)p(\theta) &= \prod_{i=1}^{n} I(x_i \in (\theta, \theta+1)) \frac{1}{b_0 - a_0} \\
&= \frac{1}{b_0 - a_0} I(X_i \in (\theta, \theta+1)) \\
&= \frac{1}{b_0 - a_0} I(X_{(1)} > \theta) I(X_{(n)} < (\theta+1))
\end{aligned}$$

where $X_{(1)}, X_{(n)}$ are the minimum and maximum of the $X_1, ... X_n$. While the normalizing constant term,

$$\begin{aligned}
p(X) &= \frac{1}{b_0 - a_0} \int_{X_{(n)}-1}^{X_{(1)}} d\theta \\
&= \frac{1 + X_{(1)} - X_{(n)}}{(b_0 - a_0)}
\end{aligned}$$

Thus, the posterior distribution

$$\begin{aligned}
p(\theta|X) &= \frac{\frac{1}{b_0 - a_0} I(X_{(1)} > \theta) I(X_{(n)} < (\theta+1))}{\frac{1 + X_{(1)} - X_{(n)}}{(b_0 - a_0)}} \\
&= \frac{I(X_{(1)} > \theta) I(X_{(n)} < (\theta+1))}{1 + X_{(1)} - X_{(n)}}
\end{aligned}$$

The posterior mean of θ

$$\begin{aligned}
E(\theta|X) &= \int_{X_{(n)}-1}^{X_{(1)}} \theta \frac{1}{1 + X_{(1)} - X_{(n)}} d\theta \\
&= \frac{X_{(1)}^2 - (X_{(n)} - 1)^2}{2(1 + X_{(1)} - X_{(n)})}
\end{aligned}$$

(b) Suppose $X_1 ... X_n$ are i.i.d. $U(\theta_1, \theta_2)$, where (θ_1, θ_2) are both unknown. Suppose that the joint prior for (θ_1, θ_2), where $\theta_1 \sim U(a_0, b_0), \theta_2 \sim U(a_1, b_1)$ for which $\theta_2 > \theta_1$. Derive joint posterior distribution of (θ_1, θ_2) and derive the posterior mean and variance of $\theta_2 - \theta_1$.

As $\theta_1 \sim U(a_0, b_0), \theta_2 \sim U(a_1, b_1)$ for which $\theta_2 > \theta_1$. The a_0, a_1, b_0, b_1 are specified hyperparamters, while $a_1 > b_0$. So we have θ_2, θ_1 independent priori.

$$
\begin{aligned}
\pi(\theta_1, \theta_2) &= \pi(\theta_2|\theta_1)\pi(\theta_1) \\
\pi(\theta_2, \theta_1) &= \pi(\theta_1)\pi(\theta_2), \qquad \theta_1 \sim U(a_0, b_0), \theta_2 \sim U(a_1, b_1)
\end{aligned}
$$

The joint posterior distribution

$$
\begin{aligned}
p(X|\theta_1, \theta_2) &= \prod_{i=1}^{n} \frac{1}{\theta_2 - \theta_1} I(\theta_1 < x_i < \theta_2) \\
p(\theta_2, \theta_1|X) &= \frac{p(X|\theta_1, \theta_2)p(\theta_1)p(\theta_2)}{p(X)} \\
p(X|\theta_1, \theta_2)p(\theta_1)p(\theta_2) &= [\frac{1}{\theta_2 - \theta_1}]^n \frac{1}{b_0 - a_0} \frac{1}{b_1 - a_1} I(X_{(1)} > \theta_1) I(X_{(n)} < \theta_2), \qquad \theta_1 \in (a_0, b_0), \theta_2 \in (a_1,
\end{aligned}
$$

The normalizing constant term,

$$
p(X) = \int_{a_0}^{X_{(1)}} \int_{X_{(n)}}^{b_1} [\frac{1}{\theta_2 - \theta_1}]^n \frac{1}{b_0 - a_0} \frac{1}{b_1 - a_1} d\theta_1 d\theta_2, \qquad X_{(1)} < b_0, X_{(n)} > a_1
$$

let $y = \theta_2 - \theta_1, \theta_2 = y + \theta_1$. The Jacobian transformation matrix

$$
|\frac{\partial(\theta_2, \theta_1)}{\partial(y, \theta_1)}| = \left| \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right| = 1
$$

So ,

$$
p(X) = \frac{1}{b_0 - a_0} \frac{1}{b_1 - a_1} \int_{a_0}^{X_{(1)}} \Big[\int_{X_{(n)} - X_{(1)}}^{b_1 - a_0} y^{-n} dy \Big] d\theta_1
$$

6 Problem 6

Find the density $p(x)$ from the Exercise on page 51 of the notes. The joint posterior density of $p(\mu, \tau|x)$ can also be obtained in this case by recognizing

$$
p(\mu, \tau|x) = p(\mu|\tau, x)p(\tau|x), \qquad \text{normal} \times \text{gamma}
$$

Find $p(\mu|\tau, x), p(\tau|x), p(x)$.

We could recognize the posterior distribution is a product of normal and gamma distribution.

$$p(\mu|\tau) \sim N(\mu_0, \tau^{-1}\sigma_0^2) = \frac{1}{\sqrt{2\pi\tau^{-1}\sigma_0^2}} exp(-\frac{(\mu-\mu_0)^2}{2\tau^{-1}\sigma_0^2})$$

$$p(\mu,\tau|x) = \tau^{\frac{n+\delta_0+1}{2}-1} exp(-\frac{\tau}{2}(\gamma_0 + \frac{\mu_0^2}{\sigma_0^2} + \sum_{i=1}^n x_i^2)) exp(-\frac{\tau}{2}(\mu^2(n+\frac{1}{\sigma_0^2}) - 2\mu(\sum_{i=1}^n x_i + \frac{\mu_0}{\sigma_0^2})))$$

So we have

$$\begin{aligned}
p(\mu|\tau,x) &\propto \tau^{\frac{n+\delta_0+1}{2}-1} exp(-\frac{\tau}{2}(\gamma_0 + \frac{\mu_0^2}{\sigma_0^2} + \sum_{i=1}^n x_i^2)) \\
p(\mu|\tau,x) &= \frac{[\frac{\gamma_0 + \frac{\mu_0^2}{\sigma_0^2} + \sum_{i=1}^n x_i^2}{2}]^{\frac{n+\delta_0+1}{2}}}{\Gamma(\frac{n+\delta_0+1}{2})} \tau^{\frac{n+\delta_0+1}{2}-1} exp(-\frac{\tau}{2}(\gamma_0 + \frac{\mu_0^2}{\sigma_0^2} + \sum_{i=1}^n x_i^2)) \\
p(\tau|x) &\propto exp(-\frac{\tau}{2}(\mu^2(n+\frac{1}{\sigma_0^2}) - 2\mu(\sum_{i=1}^n x_i + \frac{\mu_0}{\sigma_0^2}))) \\
p(\tau|x) &=
\end{aligned}$$

So

$$p(X) =$$

7 Problem 7

Consider the model

$$y_i = \beta x_i + \epsilon_i$$

where the $\epsilon_i, i = 1, ..n$ are i.i.d. $N(0, \sigma^2)$ random variables and (β, σ^2) are both unknown. Let $\tau = 1/\sigma^2$. Consider the joint improper prior

$$\pi(\beta, \tau) \propto \tau^{-1}$$

(a) Derive the joint posterior density of (β, τ).

$$\begin{aligned}
\epsilon_i &\sim N(0, \sigma^2) \\
y_i &\sim N(\beta x_i, \sigma^2) \\
p(y_i|\beta, \sigma^2) &= \frac{1}{\sqrt{2\pi}\sigma} exp[-\frac{(y_i - \beta x_i)^2}{2\sigma^2}] \\
\tau = 1/\sigma^2, &\qquad \pi(\beta, \tau) = \tau^{-1}
\end{aligned}$$

The joint posterior distribution of (β, τ)

$$p(\beta, \tau|X) = \frac{p(x|\beta,\tau)p(\beta,\tau)}{p(x)}$$
$$p(x|\beta,\tau)p(\beta,\tau) = (\frac{\tau^{1/2}}{\sqrt{2\pi}})^n exp(-\frac{\tau\sum_{i=1}^n (y_i - \beta x_i)^2}{2})\tau^{-1}$$

The normalizing constant

$$\begin{aligned}
p(X) &= \int_0^\infty \int_{-\infty}^\infty (\frac{\tau^{1/2}}{\sqrt{2\pi}})^n exp(-\frac{\tau\sum_{i=1}^n (y_i - \beta x_i)^2}{2})\tau^{-1} d\beta d\tau \\
&= \int_0^\infty \int_{-\infty}^\infty \frac{\tau^{n/2-1}}{(2\pi)^{n/2}} exp[-\frac{\tau\sum_{i=1}^n x_i^2}{2}(\beta^2 - 2\beta\frac{\sum x_i y_i}{\sum x_i^2}) - \frac{\tau}{2}\sum y_i^2] d\beta d\tau \\
&= \int_0^\infty \frac{\tau^{n/2-1-1/2}}{(2\pi)^{n/2-1}[\sum x_i^2]^{1/2}} exp[-\frac{\tau}{2}\sum y_i^2 + \frac{\tau(\sum x_i y_i)^2}{2\sum x_i^2}] \\
&\int_{-\infty}^\infty \frac{(\tau\sum x_i^2)^{1/2}}{\sqrt{2\pi}} exp[-\frac{\tau\sum x_i^2(\beta - \frac{\sum x_i y_i}{\sum x_i^2})^2}{2}] d\beta d\tau \\
&= \int_0^\infty \frac{\tau^{n/2-1-1/2}}{(2\pi)^{n/2-1}[\sum x_i^2]^{1/2}} exp[-\frac{\tau}{2}\sum y_i^2 + \frac{\tau(\sum x_i y_i)^2}{2\sum x_i^2}] d\tau \\
&= \frac{1}{(2\pi)^{n/2-1}[\sum x_i^2]^{1/2}} \int_0^\infty \tau^{n/2-1-1/2} exp[-\frac{\tau}{2}(\sum y_i^2 - \frac{(\sum x_i y_i)^2}{\sum x_i^2})] d\tau \\
&= \frac{\Gamma((n-1)/2)}{(2\pi)^{n/2-1}[\sum x_i^2]^{1/2}[\frac{(\sum y_i^2 - \frac{(\sum x_i y_i)^2}{\sum x_i^2})}{2}]^{(n-1)/2}} \\
&\int_0^\infty \frac{[\frac{(\sum y_i^2 - \frac{(\sum x_i y_i)^2}{\sum x_i^2})}{2}]^{(n-1)/2}}{\Gamma((n-1)/2)} \tau^{n/2-1-1/2} exp[-\frac{\tau}{2}(\sum y_i^2 - \frac{(\sum x_i y_i)^2}{\sum x_i^2})] d\tau \\
&= \frac{\Gamma((n-1)/2)}{(2\pi)^{n/2-1}[\sum x_i^2]^{1/2}[\frac{(\sum y_i^2 - \frac{(\sum x_i y_i)^2}{\sum x_i^2})}{2}]^{(n-1)/2}}
\end{aligned}$$

So the joint posterior distribution

$$p(\beta,\tau|X) = \frac{(\tau^{n/2-1}) exp(-\frac{\tau\sum_{i=1}^n (y_i - \beta x_i)^2}{2})[\sum x_i^2]^{1/2}[\frac{(\sum y_i^2 - \frac{(\sum x_i y_i)^2}{\sum x_i^2})}{2}]^{(n-1)/2}}{2\pi\Gamma((n-1)/2)}$$

(b) Derive the marginal posterior density of β and mean and variance.

The marginal distribution of β

$$p(\beta|X) = \int_0^\infty p(\beta,\tau|X)d\tau$$

$$= \int_0^\infty \frac{(\tau^{n/2-1})exp(-\frac{\tau\sum_{i=1}^n(y_i-\beta x_i)^2}{2})[\sum x_i^2]^{1/2}[\frac{(\sum y_i^2-\frac{(\sum x_iy_i)^2}{\sum x_i^2})}{2}]^{(n-1)/2}}{2\pi\Gamma((n-1)/2)}d\tau$$

$$= \frac{[\sum x_i^2]^{1/2}[\frac{(\sum y_i^2-\frac{(\sum x_iy_i)^2}{\sum x_i^2})}{2}]^{\frac{n-1}{2}}\Gamma(\frac{n}{2})}{2\pi\Gamma(\frac{n-1}{2})[\frac{\sum(y_i-\beta x_i)^2}{2}]^{\frac{n}{2}}}$$

$$\int_0^\infty \frac{[\frac{\sum(y_i-\beta x_i)^2}{2}]^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\tau^{\frac{n}{2}-1}exp(-\frac{\tau\sum_{i=1}^n(y_i-\beta x_i)^2}{2})d\tau$$

$$= \frac{[\sum x_i^2]^{1/2}[\frac{(\sum y_i^2-\frac{(\sum x_iy_i)^2}{\sum x_i^2})}{2}]^{\frac{n-1}{2}}\Gamma(\frac{n}{2})}{2\pi\Gamma(\frac{n-1}{2})[\frac{\sum(y_i-\beta x_i)^2}{2}]^{\frac{n}{2}}}$$

posterior mean

$$E(\beta|X) = \int_{-\infty}^\infty \beta\frac{[\sum x_i^2]^{1/2}[\frac{(\sum y_i^2-\frac{(\sum x_iy_i)^2}{\sum x_i^2})}{2}]^{\frac{n-1}{2}}\Gamma(\frac{n}{2})}{2\pi\Gamma(\frac{n-1}{2})[\frac{\sum(y_i-\beta x_i)^2}{2}]^{\frac{n}{2}}}d\beta$$

$$= \frac{[\sum x_i^2]^{1/2}[\frac{(\sum y_i^2-\frac{(\sum x_iy_i)^2}{\sum x_i^2})}{2}]^{\frac{n-1}{2}}\Gamma(\frac{n}{2})}{2\pi\Gamma(\frac{n-1}{2})}\int_{-\infty}^\infty \beta[\frac{\sum(y_i-\beta x_i)^2}{2}]^{\frac{n}{2}}d\beta$$

(c) Derive the marginal posterior density and calculate the posterior mean andvariance.

The marginal distribution of τ