ISyE 6420, Spring 2018

HW2 due Tuesday, February 20, 11:55pm

HW 2

1. Jeremy Mixture. Show that for likelihood $f(x|\theta)$ and mixture prior

$$\pi(\theta) = \epsilon\pi_1(\theta) + (1-\epsilon)\pi_2(\theta),\ \theta \in \Theta,$$

the posterior is a mixture

$$\pi(\theta|x) = \epsilon'\pi_1(\theta|x) + (1-\epsilon')\pi_2(\theta|x)$$

where

$$\begin{aligned}\pi_i(\theta|x) &= \frac{f(x|\theta)\pi_i(\theta)}{m_i(x)},\quad m_i(x) = \int_\Theta f(x|\theta)\pi_i(\theta)d\theta,\ i = 1,2,\ \text{and}\\ \epsilon' &= \frac{\epsilon m_1(x)}{\epsilon m_1(x) + (1-\epsilon)m_2(x)}.\end{aligned}$$

In the Jeremy's IQ example, where $X|\theta \sim \mathcal{N}(\theta, 80)$, assume that the prior for θ is a mixture

$$\theta \sim \pi(\theta) = \frac{2}{3}\mathcal{N}(110, 60) + \frac{1}{3}\mathcal{N}(100, 200).$$

Find the posterior and Bayes estimator for θ if $X = 98$.

2. Maxwell. Sample $y_1, \dots, y_n$, comes from Maxwell distribution with a density

$$f(y|\theta) = \sqrt{\frac{2}{\pi}}\theta^{3/2}y^2e^{-\theta y^2/2},\ y \geq 0, \theta > 0.$$

Assume an exponential prior on θ,

$$\pi(\theta) = \lambda e^{-\lambda\theta},\ \theta > 0, \lambda > 0.$$

(a) Show that posterior belongs to gamma family and depends on data via $\sum_{i=1}^n y_i^2$.

(b) For $\lambda = 1/2$ and $y_1 = 1.4, y_2 = 3.1$, and $y_3 = 2.5$ find Bayes estimator for θ. How the Bayes estimator compares to the MLE and prior mean. The MLE for θ is $\frac{3n}{\sum_{i=1}^n y_i^2} = 3/\overline{y^2}$.

(c) Using MATLAB/Octave/R/Python calculate 95% equitailed credible set for θ.

(d) Find a prediction for a future single observation. For this, you will need the mean of Maxwell, which is $EY = 2\sqrt{\frac{2}{\pi\theta}}$.

3. Times to Failure. Three devices are monitored until failure. The observed lifetimes are 0.9, 1.8, and 0.3 years. If the lifetimes ate modeled as exponential distribution with rate λ,

$$T_i \sim \mathcal{E}xp(\lambda), \quad f(t|\lambda) = \lambda e^{-\lambda t},\ t > 0, \lambda > 0.$$

Assume exponential prior on λ,

$$\lambda \sim \mathcal{E}xp(2), \quad \pi(\lambda) = 2e^{-2\lambda},\ \lambda > 0.$$

(a) Find the posterior distribution of λ
(b) Find the Bayes estimator for λ
(c) Find the MAP estimator for λ
(d) Approximate the posterior median of λ?
(e) Numerically find 95% HPD confidence interval for λ
(f) Numerically find 95% equitailed confidence interval for λ
(g) Find the posterior probability of hypothesis $H_0 : \lambda \leq 1/2$.
(h) Test precise hypothesis $H_0 : \lambda = 1/2$ versus the alternative $H_1 : \lambda \neq 1/2$, assuming equal prior probabilities for H_0 and H_1, and taking the prior

$$\pi(\lambda) = \frac{1}{2} \times \delta_{1/2} + \frac{1}{2} \times \frac{1}{10}\mathbf{1}(0 \leq \lambda \leq 10).$$

(i) Test precise hypothesis $H_0 : \lambda = 1/2$ versus the alternative $H_1 : \lambda > 1/2$, assuming equal prior probabilities for H_0 and H_1, and taking the prior

$$\pi(\lambda) = \frac{1}{2} \times \delta_{1/2} + \frac{1}{2} \times 2e^{-2(\lambda-1/2)}\mathbf{1}(\lambda > 1/2).$$

(j) Find posterior predictive distribution for a new T
(k) Predict a future T in Bayesian fashion
(l) Numerically approximate 95% HPrD and equitailed prediction intervals a future T

If you choose #3, for HW2 SOLVE ONLY PART (h)

Hints and Discussion:

Assume that $T_1, \dots, T_n$ are iid exponential $\mathcal{E}xp(\lambda)$ (λ rate parameter), then $S = \sum_{i=1}^n T_i$ is gamma $\mathcal{G}a(n, \lambda)$.

So, the likelihood is

$$f(s|\lambda) = \frac{s^{n-1}\lambda^n}{\Gamma(n)} e^{-\lambda s},\ s > 0, \lambda > 0,$$

where $s = \sum_{i=1}^n t_i$ is the variable corresponding to the sum of individual lifetimes t_i.

When the prior on λ is also exponential $\mathcal{E}xp(\beta)$,

$$\lambda \sim \beta e^{-\beta\lambda},\ \lambda > 0, \beta > 0,$$

then the posterior is gamma $\mathcal{G}a(n+1, s+\beta)$

$$\pi(\lambda|s) = \frac{(s+\beta)^{n+1}\lambda^n}{\Gamma(n+1)} e^{-(s+\beta)\lambda},$$

The posterior mean is $\hat{\lambda}_B = E^{\pi(\cdot|s)}\lambda = \frac{n+1}{s+\beta}$, and the posterior mode is $n/(s+\beta)$.

For the observed data and priot hyperparameter $\beta = 2$, the $\hat{\lambda}_B = 4/5$ and MAP $= 3/5$. The median is

```
>> gaminv(0.5, 4, 1/5)  % 0.734412149770179
```

. Note that MATLAB/R uses scale parameter 1/5 instead of rate parameter 5.

The credible sets are also found numerically using MATLAB/Octave.

Both HPD and equitailed credibkle sets are found.

```
%CREDIBLE SETS
%HPD
k=0.14782221790488
a1=fzero(@(x) gampdf(x, 4, 1/5)-k,0) %0.142500168176045
a2=fzero(@(x) gampdf(x, 4, 1/5)-k,2) %1.589659231823644
gamcdf(a2,4,1/5)-gamcdf(a1,4,1/5)     %0.950000000000000
%[0.1425, 1.59]
a2-a1 %1.447159063647599
%%
%EQUITAILED
a1=gaminv(0.025,4,1/5)  %0.217973074725265
a2=gaminv(0.975,4,1/5)  %1.753454613948465
%[0.2181.447159063647599, 1.753]
a2-a1  %1.535481539223200
```

For the posterior predictive distribution for a single future lifetime T we find

$$f(t|s) = \int_0^\infty \lambda e^{-\lambda t} \pi(\lambda|s) d\lambda.$$

This integral is

$$f(t|s) = \frac{(n+1)(\beta+s)^{n+1}}{(\beta+s+t)^{n+2}}.$$

The prediction is the mean of T wrt distribution $f(t|s)$.

$$\hat{T} = E^{f(\cdot|s)}T = \int_0^t t \times \frac{(n+1)(\beta+s)^{n+1}}{(\beta+s+t)^{n+2}} dt = \frac{\beta+s}{n}.$$

Alternatively, the prediction can be obtained by integrating the $ET = 1/\lambda$ with respect to the posterior distribution,

$$\int_0^\infty \frac{1}{\lambda}\pi(\lambda|s)d\lambda = \frac{\beta+s}{n}.$$

Thus in our specific case the posterior predictive distribution is

$$f(t|s=5) = \frac{2500}{(5+t)^5},$$

and a prediction for future onservation is $\hat{T} = 5/3$.

To find HPrD and equitailed predictive sets we use MATLAB/Octave

```
%%
%PREDICTION SETS
%HPrD
q=0.0189148600581184
g = @(t) 2500./(5+t)^5  - q
u = fzero(g, 5) % 5.573709517547107
conf = quad(@(t) 2500./(5+t).^5, 0, u)  %  0.950000000000000
%[0, 5.57]
%%
%EQUITAILED
quad(@(t) 2500./(5+t).^5, 0, 0.0317476266) %0.025000000000
quad(@(t) 2500./(5+t).^5, 0, 7.57432140202851) %0.975000000000000

%[0.0317, 7.574]
```

The ξ- part of the mixture prior

$$\pi(\lambda) = \pi_0\delta_{1/2} + \pi_1\xi(\lambda) = \frac{1}{2}\times\delta_{1/2} + \frac{1}{2}\times\beta e^{-\beta(\lambda-1/2)}\mathbf{1}(\lambda > 1/2).$$

spreads the mass over the range of H_1, the set $(1/2,\infty)$. To find $P(H_0)$ we need to integrate the product of likelihood and the spread part of prior, with respect to λ. The solution involves the incomplete gamma function and has to be evaluated numerically. Here MATHAMATICA was used for symbolic calculation.

$$m_1(s) = \int_{1/2}^\infty f(\lambda|s)\xi(\lambda)d\lambda = \beta e^{-\beta/2}\frac{s^{n-1}}{(\beta+s)^{n+1}}\,\frac{\Gamma(n+1,(\beta+s)/2)}{\Gamma(n)},$$

which, for $s=3, n=3$, and $\beta=2$ gives 0.177924.

Thus

$$P(H_0) = \left(1+\frac{1/2}{1/2}\,\frac{m_1}{f(1/2|s=3)}\right)^{-1} = \left(1+\frac{0.177924}{0.125511}\right)^{-1} = 0.413633$$

Since the prior probabilities of hypotheses are equal, $\pi_0 = \pi_1$, the Bayes Factor BF_{10} is $p_1/p_0 = 1.4176$, and $\log_{10} BF_{10} = 0.151554$. According to Jeffreys' chart, the evidence in favor of H_1 is very weak.

4. Beyond Conjugate Pairs – Logistic Prior. Let

$$\begin{aligned} X|\theta &\sim \mathcal{N}(\theta, 1) \\ \theta &\sim \mathcal{L}(0, 1), \end{aligned}$$

where $\mathcal{L}(0, 1)$ is standard logistic distribution with density

$$\pi(\theta) = \frac{e^\theta}{(1+e^\theta)^2} = \frac{e^{-\theta}}{(1+e^{-\theta})^2},$$

and CDF

$$\Pi(\theta) = \frac{1}{1+e^{-\theta}} = 1/2 + 1/2 \tanh(\theta/2).$$

Let $X = 2$ was observed. Estimate θ in Bayesian fashion.

(a) Solve the needed integrals numerically. The estimator should be close to 1.4444269.

(b) Simulate from the normal distribution and approximate integrals by Monte Carlo method

(c) Simulate from the logistic distribution and approximate integrals by Monte Carlo method. You can simulate standard logistic by invese CDF method. Also, if U, V are two independent exponential random variables with rate 1, the logarithm of their ratio is logistic.

5. Beyond Conjugate Pairs – Logistic Prior via Metropolis.

For model in Exercise 4, approximate Bayes rule using Metropolis Algorithm.

(a) Use Metropolis algorithm using normal distribution as the proposal distribution, $\theta_{new} \sim \mathcal{N}(\theta_{old}, 2^2)$. (This is Random walk Metropolis)

(b) Use Metropolis algorithm using normal distribution as the proposal distribution, $\theta_{new} \sim \mathcal{N}(x, 1)$. (This is Independence Metropolis)

(c) Since you would be sampling from the posterior, show the histogram of the posterior and propose empirical 95% equitailed credible set using simulations from (b).