Extra Bayes Exercises

Solution

Note: The order of the questions is as on the handout on the course web-page.

Q.1 $$L = \prod_i p(y_i|\lambda) = \lambda^{\sum_i y_i} e^{-n\lambda} / \prod_i y_i!$$

$$\propto \lambda^{\sum_i y_i} e^{-n\lambda}$$

$$p(\lambda) \propto 1/\lambda \quad ; \quad \lambda > 0$$

Bayes' Theorem:

$$p(\lambda|y) \propto \left(\frac{1}{\lambda}\right) \lambda^{\sum y_i} e^{-n\lambda}$$

$$\propto \lambda^{\sum y_i - 1} e^{-n\lambda}$$

$$\propto \text{Gamma}\left(\sum_i y_i, n\right)$$

Under quadratic loss,

$$\hat{\lambda} = E(\lambda|y) = \left(\sum_i y_i / n\right) = \bar{y}.$$

Q.2. $$L = (1/\theta)^n = \theta^{-n}$$

$$p(\theta) = a k^a \theta^{-(a+1)}$$

(a) By Bayes' Theorem:

$$p(\theta|y) \propto a k^a \theta^{-(a+1)} \theta^{-n}$$

$$\propto \theta^{-(a+n+1)}.$$

This is the kernel of a Pareto distribution with parameter $(a+n)$ instead of 'a'. So, the Pareto prior is Conjugate.

(b) Let $c = \int_k^\infty \theta^{-(a+n+1)} d\theta$

$$= \left(\frac{-1}{a+n}\right)\left[\frac{1}{\theta^{a+n}}\right]_k^\infty$$

$$= -\frac{1}{(a+n)}\left[0 - k^{-(a+n)}\right]$$

$$= \frac{k^{-(a+n)}}{(a+n)}$$

So, $p(\theta|y) = \frac{1}{c}\theta^{-(a+n+1)}$

$$= (a+n)k^{-(a+n)}\theta^{-(a+n+1)}$$

(c) Mean of prior:

$$E(\theta) = \int_k^\infty \theta\, p(\theta)\, d\theta$$

$$= \int_k^\infty a k^a \theta^{-(a+1)} \theta\, d\theta$$

$$= ak^a \int_k^\infty \theta^{-a}\, d\theta$$

$$= \frac{-ak^a}{(a-1)} \left[\theta^{-(a-1)} \right]_k^\infty \quad ; a > 1$$

$$= \frac{-ak^a}{(a-1)} \left[0 - k^{-(a-1)} \right] \quad ; a > 1$$

$$= \frac{ak}{(a-1)} \quad ; \quad a > 1 .$$

(Need $a > 1$ for mean to be +ve.)

(d) Under quadratic loss the Bayes' estimator is the mean of the posterior p.d.f. So,

$$\hat{\theta} = E(\theta | y) = \frac{(a+n)k}{(a+n-1)} .$$

Q.5. (a) $p(y_i | \theta, \lambda) = (2\lambda)^{-1} \exp\{-|y_i - \theta|/\lambda\}$

$$L(\theta, \lambda | \underset{\sim}{y}) = \prod_{i=1}^{n} p(y_i | \theta, \lambda)$$

$$= (2\lambda)^{-n} \exp\left[-\frac{1}{\lambda} \sum_i |y_i - \theta| \right]$$

$$\ell = \log L = -n \log(2\lambda) - \frac{1}{\lambda} \sum_i |y_i - \theta|$$

Now, to maximize ℓ w.r.t. θ, for any λ, we need to minimize $\sum_i |y_i - \theta|$. So, $\hat{\theta} = M$

where M = median of sample.

Also, $\frac{\partial \ell}{\partial \lambda} = -n\left(\frac{1}{2\lambda}\right)\cdot 2 + \frac{1}{\lambda^2}\sum_i |y_i - \theta|$

$= 0$

$\Rightarrow n/\hat{\lambda} = \frac{1}{\hat{\lambda}^2}\sum_i |y_i - \hat{\theta}|$

$\Rightarrow \hat{\lambda} = \frac{1}{n}\sum_i |y_i - \hat{\theta}|$; $\hat{\theta} = M$.

(b) $\text{var}(y_i) = E(y_i^2) - [E(y_i)]^2 = 2\lambda^2 + \theta^2 - \theta^2$

$= 2\lambda^2$.

So, $CV = \theta / \sqrt{2\lambda^2} = \frac{\theta}{\lambda\sqrt{2}}$.

By Slutsky's Theorem, a consistent estimator of CV is $\widehat{CV} = \frac{\hat{\theta}}{\hat{\lambda}\sqrt{2}}$, because MLE's are consistent. (In fact this is actually the MLE for CV, by invariance, so it is also asymptotically Normal & asymptotically efficient.)

(c) Note that θ is known, & $\lambda > 0$.

So, $p(\lambda) \propto \lambda^{-1}$

By Bayes' Theorem:

$p(\lambda|y) \propto \lambda^{-1}(2\lambda)^{-n} \exp\left[-\frac{1}{\lambda}\sum_i |y_i - \theta|\right]$

$\propto \lambda^{-(n+1)} \exp\left[-\frac{c}{\lambda}\right]$

where $c = \sum_i |y_i - \theta|$ is known.

For a 0-1 loss function we need the mode of $p(\lambda|y)$:

$$\frac{\partial p(\lambda|y)}{\partial \lambda} = -(n+1)\lambda^{-(n+2)} e^{-c/\lambda} + \lambda^{-(n+1)} e^{-c/\lambda} (-c)(-1)\lambda^{-2}$$

$$= 0$$

$$\Rightarrow (n+1)\hat{\lambda}^{-(n+2)} e^{-c/\hat{\lambda}} = c\,\hat{\lambda}^{-(n+3)} e^{-c/\hat{\lambda}}$$

$$\Rightarrow (n+1)\hat{\lambda} = c$$

$$\Rightarrow \hat{\lambda} = \left(\frac{c}{n+1}\right) = \frac{1}{n+1}\sum_i |y_i - \theta|.$$

So, this differs from the MLE only by the use of $(n+1)$ rather than 'n' in the denominator. Because 'n' & '$n+1$' change at the same rate, the Bayes' & MLE will converge in probability to λ at the same rate — the usual rate of $n^{-1/2}$.