7. (a) The likelihood for the data is:

$$L(\theta|\beta) = P(\beta|\theta)$$
$$= P(\beta_1|\theta) \times \cdots \times P(\beta_n|\theta)$$
$$= \prod_{i=1}^{n} P(\beta_i|\theta)$$
$$= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{Y_i - (\beta_0 + \sum_{j=1}^{p} \beta_j X_{ij})}{2\sigma^2}\right)$$
$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} \left[Y_i - (\beta_0 + \sum_{j=1}^{p} \beta_j X_{ij})\right]^2\right)$$

(b) The posterior is

$$f(\beta|X, Y) \propto f(Y|X, \beta) P(\beta|X) = f(Y|X, \beta) p(\beta)$$
$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} \left[Y_i - (\beta_0 + \sum_{j=1}^{\beta} \beta_j X_{ij})\right]^2\right)\left(\frac{1}{2b} exp\left(-\frac{|\beta|}{b}\right)\right)$$
$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \left(\frac{1}{2b}\right) exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} \left[Y_i - (\beta_0 + \sum_{j=1}^{p} \beta_j X_{ij})\right]^2 - \frac{|\beta|}{b}\right)$$

(c) The mode of the posterior distribution can be found by maximizing the logarithm of the posterior, which is the logarithm of the product of the likelihood and the prior. The mode represent the most likely values of the coefficients β given the data and the prior.

By taking the derivative of the log posterior with respect to β_j and setting it equal to 0, we have: $\sum \frac{1}{\sigma^2}(y_i - \sum X_{ij}\beta_j)X_{ij} - \frac{|\beta_j|}{b} = 0$

This equation is equivalent to the soft-threshdding rule that is used to calculate the lasso estimate. Hence, the lasso estimate is the mode of the posterior distribution under the double exponential prior.

(d) The posterior distributed according to Normal distribution with mean $=0$ and variance $=c$ is:

$$f(\beta|X,Y) \propto f(Y|X,\beta)P(\beta|X) = f(Y|X,\beta)P(\beta)$$

$$P(\beta) = \prod_{i=1}^{p} P(\beta_i) = \prod_{i=1}^{p} \frac{1}{\sqrt{2c\pi}} \exp\left(-\frac{\beta_i^2}{2c}\right) = \left(\frac{1}{\sqrt{2c\pi}}\right)^p \exp\left(-\frac{1}{2c}\sum_{i=1}^{p}\beta_i^2\right)$$

$$\therefore f(Y|X)P(\beta) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[Y_i-\left(\beta_0+\sum_{j=1}^{p}\beta_j X_{ij}\right)\right]^2\right)\left(\frac{1}{\sqrt{2c\pi}}\right)^p \exp\left(-\frac{1}{2c}\sum_{i=1}^{p}\beta_i^2\right)$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \left(\frac{1}{\sqrt{2c\pi}}\right)^p \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[Y_i-\left(\beta_0+\sum_{j=1}^{p}\beta_j X_{ij}\right)\right]^2-\frac{1}{2c}\sum_{j=1}^{p}\beta_j^2\right)$$

(e) The mode of the posterior distribution can be found by maximizing the logrithm of the posterior. The mode corresponds to the most likely values of the coefficients β given the data and the prior.

By taking the derivative of the log posterior with respect to β_j and setting it equal to 0 we have: $\frac{1}{\sigma^2}\sum(y_i - \sum X_{ij}\beta_j)X_{ij} - \frac{\beta_j}{c} = 0$

This equation is equalent to the normal equation for ridge regression. Hence, the ridge regression estimate is both mode and the mean of the posterior distribution under the normal prior.