Math 50 Stat Inf: Homework 7—selected SOLUTIONS

due Wed Feb 22

A. i) $L(\mu,\sigma) = c\sigma^{-n}e^{-\frac{1}{2\sigma^2}\sum_i (y_i-\mu)^2}$

ii) $g(\mu,\sigma|y) = c.1.L(\mu,\sigma)$ so is identical to the likelihood. Don't bother finding c since as I explained, we won't keep track of such overall constants.

iii) First expand the squared term upstairs then bring out what doesn't depend on μ, and write $\sum_i y_i = n\bar{y}$:

$$\begin{aligned} g(\sigma|y) &= \int_{-\infty}^{\infty} g(\mu,\sigma|y)d\mu \\ &= c\sigma^{-n}\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(\sum_i y_i^2 - 2\mu\sum_i y_i + n\mu^2)}d\mu \\ &= c\sigma^{-n}e^{-\frac{1}{2\sigma^2}\sum_i y_i^2}\int_{-\infty}^{\infty} e^{-\frac{n\mu^2}{2\sigma^2}+\frac{n\bar{y}\mu}{\sigma^2}}d\mu \end{aligned}$$

so to use the integral I gave you we set $a = \sqrt{n}/\sigma$ and $b = -n\bar{y}/\sigma^2$. You then use the Gaussian integral $\int_{-\infty}^{\infty} e^{-(a^2x^2/2)-bx}dx = (\sqrt{2\pi}/a)e^{b^2/2a^2}$, which can be proven simply by completing the square. So marginal posterior becomes

$$\begin{aligned} g(\sigma|y) &= c\sigma^{-n}e^{-\frac{1}{2\sigma^2}\sum_i y_i^2}\cdot c\sigma e^{\frac{n^2\bar{y}^2}{\sigma^4}\cdot\frac{\sigma^2}{2n}} \\ &= c\sigma^{1-n}e^{-\frac{1}{2\sigma^2}(\sum_i y_i^2 - n\bar{y}^2)} \\ &= c\sigma^{1-n}e^{-\frac{1}{2\sigma^2}(\sum_i (y_i-\bar{y})^2)} \end{aligned} \tag{1}$$

In the last step we used $\sum_i (y_i-\bar{y})^2 = (\sum_i y_i^2) - n\bar{y}^2$.

iv)

$$\begin{aligned} \frac{d}{d\sigma}\ln g(\sigma|y) &= \frac{d}{d\sigma}\left(-(n-1)\ln\sigma - \frac{1}{2\sigma^2}\sum_i (y_i-\bar{y})^2\right) \\ &= -\frac{n-1}{\sigma} + \frac{1}{\sigma^3}\sum_i (y_i-\bar{y})^2 \end{aligned} \tag{2}$$

Setting this to zero gives the MAP peak at $\sigma^2 = \frac{1}{n-1}\sum_i (y_i-\bar{y})^2$, the same as the familiar unbiased ML estimate for variance.

B. Bayesian prediction for bus waiting time. [No matlab].

i)

$$h(y|y_1) = \int_0^{\infty} f_Y(y|\theta)g(\theta|y_1)d\theta = \int_0^{\infty} \theta e^{-\theta y}\cdot y_1^2\theta e^{-\theta y_1}d\theta = y_1^2\int_0^{\infty}\theta^2 e^{-(y+y_1)\theta}d\theta$$

which is a gamma integral giving $h(y|y_1) = 2y_1^2/(y+y_1)^3$.

ii) ML estimate is at $\theta_e = 1/y_1$ (by setting θ-deriv of likelihood to zero). So frequentists predictive pdf is $h(y|y_1) = f_Y(y|\theta_e) = -(1/y_1)e^{-y/y_1}$.

iii) $y_1 = 10$ mins. In each case $p(Y \geq 60\text{mins}) = \int_6 0^{\infty} h(y|y_1)dy$. Bayesian gets $y_1^2 \int_7 0^{\infty} u^{-3} du = (1/7)^2 = 0.0204$, whereas frequentist gets $e^{-6} = 0.00248$. [The $\tan^{-1}$ clue was actually wrong - sorry about that!]

iv) The Bayesian allows the chance that the rate is actually much *slower* than one every 10 mins, so its predictive pdf has a 'longer tail' (it's power-law not exponential) than the freuqentist predicitive pdf. This makes long wait times much more likely. I think you'll agree this Bayesian approach is closer to reality—the datum $y_1 = 10$ mins could have been unrepresentitatively short!