Stat 352: Solutions to Assignment #1

1. Using a Riemann midpoint sum with four partitions of equal width gives

$$\int_0^1 x^2\,\mathrm{d}x \approx \frac{1}{4}\left[\left(\frac{1}{8}\right)^2+\left(\frac{3}{8}\right)^2+\left(\frac{5}{8}\right)^2+\left(\frac{7}{8}\right)^2\right]=\frac{21}{64}=0.328125.$$

2. Note that a special case of Bayes' Theorem follows from the definition of conditional probability, namely that if $P(A)>0$ and $P(B)>0$, then $P(A|B)\cdot P(B)=P(B|A)\cdot P(A)$ since both of these expression equal $P(A\cap B)$. Hence,

$$\begin{aligned}&P(\text{Elvis was an identical twin}|\text{Elvis had a twin brother})\\&\quad=\frac{P(\text{Elvis had a twin brother}|\text{Elvis was an identical twin})\cdot P(\text{Elvis was an identical twin})}{P(\text{Elvis had a twin brother})}.\end{aligned}$$

We are told that

$$P(\text{Elvis was an identical twin})=\frac{1}{300}.$$

Furthermore, it follows immediately that

$$P(\text{Elvis had a twin brother}|\text{Elvis was an identical twin})=1.$$

Therefore, we must calculate $P(\text{Elvis had a twin brother})$ using the law of total probability. Thus,

$$\begin{aligned}&P(\text{Elvis had a twin brother})\\&\quad=P(\text{Elvis had a twin brother}|\text{Elvis was an identical twin})\cdot P(\text{Elvis was an identical twin})\\&\qquad+P(\text{Elvis had a twin brother}|\text{Elvis was a fraternal twin})\cdot P(\text{Elvis was a fraternal twin})\\&\qquad\quad+P(\text{Elvis had a twin brother}|\text{Elvis was NOT a twin})\cdot P(\text{Elvis was NOT a twin})\\&\quad=1\cdot\frac{1}{300}+\frac{1}{2}\cdot\frac{1}{125}+0\end{aligned}$$

so that

$$P(\text{Elvis was an identical twin}|\text{Elvis had a twin brother})=\frac{1\cdot\frac{1}{300}}{1\cdot\frac{1}{300}+\frac{1}{2}\cdot\frac{1}{125}}=\frac{5}{11}.$$

3. Suppose that $Y|\theta\sim\text{Bin}(n,\theta)$ with $\theta\sim\beta(a,b)$ so that

$$g(\theta)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1},\quad 0\le\theta\le 1.$$

As shown in class,

$$f(\theta|y)=\frac{\theta^y(1-\theta)^{n-y}\cdot g(\theta)}{\displaystyle\int_{-\infty}^{\infty}\theta^y(1-\theta)^{n-y}\cdot g(\theta)\,\mathrm{d}\theta}.$$

Thus,

$$f(\theta|y)\propto\theta^y(1-\theta)^{n-y}\theta^{a-1}(1-\theta)^{b-1}=\theta^{y+a-1}(1-\theta)^{n-y+b-1}$$

from which we conclude that the posterior distribution of θ given y is $\beta(y+a,n-y+b)$.

4. (a) If $Y_1, \ldots Y_n$ are i.i.d. Poisson(θ) random variables, then

$$f(y_1, \ldots, y_n|\theta) = \prod_{i=1}^{n} f(y_i|\theta) = \frac{1}{\prod y_i!} e^{-n\theta} \theta^{\sum y_i}$$

where the sum and product both run from $i = 1$ to n.

4. (b) If the prior distribution of θ is $\Gamma(\alpha, \beta)$, then

$$f(\theta|y_1, \ldots, y_n) \propto e^{-n\theta} \theta^{\sum y_i} \cdot \theta^{\alpha-1} e^{-\theta/\beta} = e^{-\theta(n+1/\beta)} \theta^{\alpha-1+\sum y_i}$$

which implies that the posterior distribution of θ given $(y_1, \ldots, y_n)$ is

$$\Gamma\left(\alpha + \sum_{i=1}^{n} y_i \ , \ \frac{1}{n + 1/\beta}\right).$$

5. (a) As shown in class, if $Y_1, \ldots, Y_n$ are a random sample of $\mathcal{N}(\theta, \sigma^2)$ randopm variables with prior density $g(\theta) \sim \mathcal{N}(\mu, \tau^2)$, then the posterior of θ given $(y_1, \ldots, y_n)$ is

$$\mathcal{N}\left(\frac{\sigma^2\mu + n\tau^2\overline{y}}{\sigma^2 + n\tau^2}, \frac{\sigma^2\tau^2}{\sigma^2 + n\tau^2}\right).$$

Thus, using the data in the problem, we have

$$\sigma^2 = 20^2, \quad \mu = 180, \quad \tau^2 = 40^2$$

which implies that the posterior of θ given $\overline{y} = 150$ is

$$\mathcal{N}\left(\frac{20^2 \cdot 180 + n40^2 \cdot 150}{20^2 + n40^2}, \frac{20^2 \cdot 40^2}{20^2 + n40^2}\right) = \mathcal{N}\left(\frac{72000 + 240000n}{400 + 1600n}, \frac{640000}{400 + 1600n}\right)$$
$$= \mathcal{N}\left(\frac{180 + 600n}{1 + 4n}, \frac{1600}{1 + 4n}\right).$$

5. (b) As shown in class, the posterior predictive distribution for $\tilde{y}$ given y is

$$\mathcal{N}(\nu, \sigma^2 + \phi^2)$$

where

$$\nu = \frac{\sigma^2\mu + n\tau^2\overline{y}}{\sigma^2 + n\tau^2} \quad \text{and} \quad \phi^2 = \frac{\sigma^2\tau^2}{\sigma^2 + n\tau^2}.$$

Thus, since $\sigma^2 = 20^2$, we find

$$f(\tilde{y}|y) \sim \mathcal{N}\left(\frac{180 + 600n}{1 + 4n}, 400 + \frac{1600}{1 + 4n}\right).$$