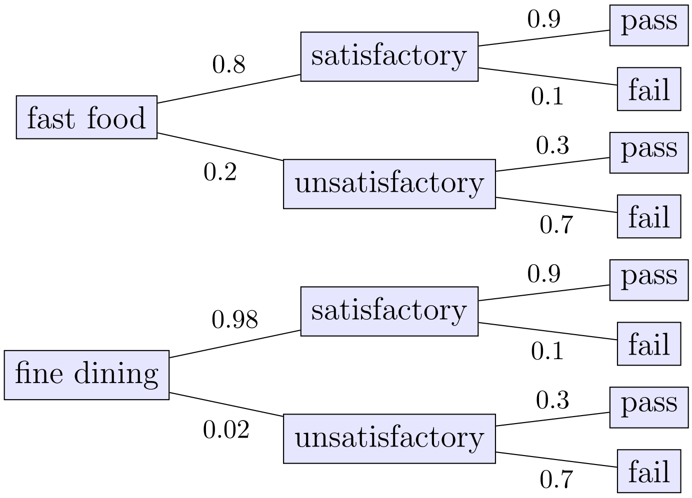
MAST20005/MAST90058: Week 11 Solutions

1. (a) $\Pr(\text{fail}) = 0.8 \times 0.1 + 0.2 \times 0.7 = 0.22$, $\Pr(\text{unsatisfactory} \mid \text{fail}) = \frac{0.2\times 0.7}{0.22} = 0.636$

   (b) $\Pr(\text{fail}) = 0.98 \times 0.1 + 0.02 \times 0.7 = 0.112$, $\Pr(\text{unsatisfactory} \mid \text{fail}) = \frac{0.02\times 0.7}{0.112} = 0.125$

   (c) Tree diagrams:



2. (a) The prior is $\text{Beta}(1, 1)$. We observe 4 successes and 36 failures, so the posterior is $\text{Beta}(1 + 4, 1 + 36) = \text{Beta}(5, 37)$. This gives a posterior mean of $5/42 = 0.12$.

   (b) i. Using a $\text{Beta}(\alpha, \beta)$ prior, we need $\alpha + \beta = 60$ and $\alpha/(\alpha + \beta) = 0.04$. Solving these gives $\alpha = 2.4$ and $\beta = 57.6$.

   ii. The new posterior is $\text{Beta}(6.4, 93.6)$.

   iii. The new posterior mean is $6.4/100 = 0.064$.

3. Let the prior be $\theta \sim \mathrm{N}(\mu_0, \sigma_0^2)$. The posterior pdf is,

$$
\begin{aligned}
f(\theta \mid x) &\propto \exp\left[-\frac{(x-\theta)^2}{2\sigma^2} - \frac{(\theta-\mu_0)^2}{2\sigma_0^2}\right] = \exp\left[-\frac{1}{2}\left(\frac{\theta^2 - 2x\theta + x^2}{\sigma^2} + \frac{\theta^2 - 2\mu_0\theta + \mu_0^2}{\sigma_0^2}\right)\right] \\
&\propto \exp\left[-\frac{1}{2}\left(\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_0^2}\right)\theta^2 - 2\left(\frac{x}{\sigma^2} + \frac{\mu_0}{\sigma_0^2}\right)\theta\right)\right] = \exp\left[-\frac{1}{2\sigma_1^2}\left(\theta^2 - 2\mu_1\theta\right)\right] \\
&\propto \exp\left[-\frac{(\theta-\mu_1)^2}{2\sigma_1^2}\right]
\end{aligned}
$$

where,

$$
\mu_1 = \frac{\frac{\mu_0}{\sigma_0^2} + \frac{x}{\sigma^2}}{\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2}} \quad \text{and} \quad \frac{1}{\sigma_1^2} = \frac{1}{\sigma_0^2} + \frac{1}{\sigma^2}.
$$

We recognise this as being a normal pdf, which means the posterior is $\theta \mid x \sim \mathrm{N}(\mu_1, \sigma_1^2)$.

4. Let the distribution of test scores be $X \sim \mathrm{N}(\theta, 25)$. We observe $\bar{x} = 70$ from a random sample of size $n = 16$. Using an improper uniform prior for θ, we get a posterior $\theta|\bar{x} \sim \mathrm{N}(70, 25/16)$. A central 95% credible interval for θ is $70 \pm 1.96 \times 5/4 = (67.55, 72.45)$.

5. (a) We have $Y \sim \mathrm{Pn}(n\theta)$ so the pdf is $f(y \mid \theta) = e^{-n\theta}(n\theta)^y/y!$. The posterior pdf is,

$$
f(\theta \mid y) \propto f(y \mid \theta) f(\theta) \propto \theta^{\alpha-1}e^{-\theta\beta}e^{-n\theta}\theta^y = \theta^{y+\alpha-1}e^{-\theta(n+\beta)}
$$

   This is a gamma distribution with parameters $y + \alpha$ and $n + \beta$.

(b) $E(\theta \mid y) = (y + \alpha)/(n + \beta)$

(c) The MLE is y/n and the prior mean is α/β. We can see that,

$$\frac{y+\alpha}{n+\beta} = \frac{n}{n+\beta} \cdot \frac{y}{n} + \frac{\beta}{n+\beta} \cdot \frac{\alpha}{\beta}.$$

6. The likelihood is,

$$f(x_1, \ldots, x_n \mid \theta) = (3\theta)^n \left(\prod_{i=1}^{n} x_i^2 \right) e^{-\theta \sum_{i=1}^{n} x_i^3}$$

and the prior pdf is,

$$f(\theta) = \frac{4^4}{\Gamma(4)} \theta^3 e^{-4\theta}, \quad 0 \leqslant \theta < \infty,$$

so the posterior pdf is,

$$f(\theta \mid x_1, \ldots, x_n) \propto \theta^n e^{-\theta \sum_{i=1}^{n} x_i^3} \theta^3 e^{-4\theta} = \theta^{n+3} e^{-\theta(4 + \sum_{i=1}^{n} x_i^3)},$$

which we recognise as being a gamma distribution with parameters $n+4$ and $4+\sum_{i=1} x_i^3$. Therefore,

$$E(\theta \mid x_1, \ldots, x_n) = \frac{n+4}{4 + \sum_{i=1}^{n} x_i^3}$$

7. Let $y = \sum_{i=1}^{n} x_i$ where the x_i are the individual times. Then the likelihood is,

$$f(x_1, \ldots, x_n \mid \theta) = \theta^n e^{-\theta y}$$

and the prior pdf is,

$$f(\theta) \propto \theta^{\alpha-1} e^{-\theta\beta}$$

which gives the posterior pdf,

$$f(\theta \mid x_1, \ldots, x_n) \propto \theta^{n+\alpha-1} e^{-\theta(y+\beta)}$$

which we recognise as being a gamma distribution with parameters $n + \alpha$ and $y + \beta$. Moreover, $\alpha/\beta = 0.2$ and $\sqrt{\alpha/\beta^2} = 0.1$ gives $\alpha = 4$ and $\beta = 20$, and from the data we have $n = 20$ and $y = n \times 3.8 = 76$ so that the posterior is a gamma distribution with parameters 24 and 96.

8. (a) The pdf for a single observation is $f(x \mid \theta) = \theta^{-1}$ with $0 < x < \theta$. Therefore, the likelihood is,

$$f(x_1, \ldots, x_n \mid \theta) = \theta^{-n}, \quad x_{(n)} < \theta.$$

Using an improper uniform prior for θ then gives the following for the posterior,

$$f(\theta \mid x_1, \ldots, x_n) \propto \theta^{-n}, \quad x_{(n)} < \theta.$$

We need to integrate this to calculate the normalising constant,

$$\int_{x_{(n)}}^{\infty} \theta^{-n} \, d\theta = \frac{1}{(n-1)x_{(n)}^{n-1}}.$$

Therefore, the posterior pdf is,

$$f(\theta \mid x_1, \ldots, x_n) = (n-1)x_{(n)}^{n-1}\theta^{-n}, \quad x_{(n)} < \theta.$$

(b) The posterior is a decreasing function of θ, so its maximum occurs at the smallest possible value of θ, which is $x_{(n)}$.

(c) By integrating the posterior pdf we can show the posterior cdf is,

$$F(\theta \mid x_1, \ldots, x_n) = 1 - \left(\frac{x_{(n)}}{\theta}\right)^{n-1}, \quad x_{(n)} < \theta.$$

Solving for $F(\theta_u \mid x_1, \ldots, x_n) = 0.95$ will give us the upper bound of the desired credible interval.

$$0.95 = 1 - \left(\frac{x_{(n)}}{\theta_u}\right)^{n-1} \quad \Rightarrow \quad \theta_u = \frac{x_{(n)}}{\sqrt[n-1]{0.05}}$$

(d) Follow the steps outlined in the lectures (Module 9, slide 23 or page 6 of the notes), but do it in general for a sample size of n. For step 1 you need to use the cdf of $X_{(n)}$; the general formula for this is given earlier in the same module. The resulting confidence interval is,

$$\left(x_{(n)}, \frac{x_{(n)}}{\sqrt[n]{0.05}}\right).$$

Note that the denominator in the upper bound differs to that of the credible interval.

(e) If instead of a uniform prior we used one of the form $f(\theta) = \theta^{-1}$ we would get the two intervals to be the same. Note that this prior is also improper.