Solution to Exam 1
STAT 638, Fall 2016

1. Let M be the event that a person has malaria, and D the event that the test indicates the person has malaria. We are asked to find $P(M|D)$, which by Bayes rule is

$$\begin{aligned}
P(M|D) &= \frac{P(D|M)P(M)}{P(D|M)P(M)+P(D|M^c)P(M^c)} \\
&= \frac{0.98(0.01)}{0.98(0.01)+0.01(0.99)} \\
&= 98/(98+99) \\
&= 0.497.
\end{aligned}$$

2. The likelihood function is

$$p(5|\theta) = \frac{1}{\theta} I_{(0,\theta)}(5) = \frac{1}{\theta} I_{(5,\infty)}(\theta)$$

and so

$$\begin{aligned}
p(\theta|5) &\propto \theta e^{-\theta} \cdot \frac{1}{\theta} I_{(5,\infty)}(\theta) \\
&= e^{-\theta} I_{(5,\infty)}(\theta).
\end{aligned}$$

To find $P(\theta > 6|5)$, we must determine the multiplier for the last expression that will make it a density function. We have

$$\int_5^\infty e^{-\theta}\,d\theta = -e^{-\theta}\Big|_5^\infty = e^{-5}.$$

So, $p(\theta|5) = e^5 e^{-\theta} I_{(5,\infty)}(\theta) = e^{-(\theta-5)} I_{(5,\infty)}(\theta)$. Integrating the density from 6 to ∞ gives $P(\theta > 6|5) = e^5 e^{-6} = e^{-1}$.

3. Frequentist inference is based on **repeated sampling from the same population.**

4. Response (a) is the definition of Y and Z being independent given θ, while (b) is a consequence of (a). Response (b) is analogous to the result at the top of p. 14 of the notes.

5. The posterior is beta$(7+1/2, 13+1/2)$, or beta$(7.5, 13.5)$. The mode of a beta(a,b) distribution is $(a-1)/(a+b-2)$, and so the posterior mode is $(7.5-1)/(21-2) = 6.5/19$. The MLE for a binomial experiment is the sample proportion, which in this case is $7/20$.

6. The density of $Y_1, \ldots, Y_n$ given σ is

$$p(y_1,\ldots,y_n|\sigma) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \sigma^{-n} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n y_i^2\right),$$

and so

$$\log p(y_1,\ldots,y_n|\sigma) = -(n/2)\log(2\pi) - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^n y_i^2.$$

We now have

$$\frac{\partial}{\partial\sigma}\log p(y_1,\ldots,y_n|\sigma) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^n y_i^2,$$

and

$$\frac{\partial^2}{\partial\sigma^2}\log p(y_1,\ldots,y_n|\sigma) = \frac{n}{\sigma^2} - \frac{3}{\sigma^4}\sum_{i=1}^n y_i^2.$$

Since $E(Y_i^2) = \sigma^2$, it follows that the Jeffreys prior is proportional to $\left[-(n/\sigma^2 - 3n/\sigma^2)\right]^{1/2}$, which is proportional to $1/\sigma$. This prior is improper since $1/\sigma$ is not integrable over $(0, \infty)$.

7. The likelihood is proportional to

$$\theta^{n/2} \exp\left(-\frac{\theta}{2}\sum_{i=1}^{n} y_i^2\right).$$

If this is multiplied by any gamma(a, b) density, we get a function that is proportional to another gamma density, and so the gamma(a, b) family is a conjugate family in this case.

8. By definition of a credible interval, (a) is correct, and by definition of an HPD region and the fact that the posterior is normal, (b) is also correct.

9. The posterior odds ratio is $0.03/0.97 = 3/97$. The prior odds ratio is $0.50/0.50 = 1$, and so the Bayes factor is $(3/97)/1 = 3/97$.

10. Generally speaking, when the number of observed data increases **the prior distribution becomes less influential**.

11. Because of the invariance property of the Jeffreys prior, we know that g, the prior for τ, is

$$g(\tau) = p(\log\tau)\left|\frac{d\theta}{d\tau}\right| = p(\log\tau)\left|\frac{d\log\tau}{d\tau}\right| = p(\log\tau)/\tau.$$

12. The answer is $p(y|\hat{\theta})$, as discussed on p. 91 of the notes.

13. When testing a point null hypothesis against a two-sided alternative, a frequentist P-value **often overstates the significance of evidence against the null hypothesis.** This is **Lindley's paradox.**

14. Using the principle discussed on pp. 84-86 of the notes, the correct answer is (b).