**Solutions**

1. Let $A = \{\text{Person is a male}\}$, $B = \{\text{Person is above 190cm}\}$. We want to compute $\mathbb{P}(A^c|B)$. First

$$\begin{aligned}\mathbb{P}(A^c|B) &= 1 - \mathbb{P}(A|B) = 1 - \frac{P(A\cap B)}{P(B)} = 1 - \frac{P(B|A)P(A)}{P(B)} \\ &= 1 - \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} = 1 - \frac{\frac{4}{100}\cdot\frac{3}{5}}{\frac{4}{100}\cdot\frac{3}{5} + \frac{1}{100}\cdot\frac{2}{5}}.\end{aligned}$$

2. Using that $\frac{1}{N}\sum_{i=1}^{n} X_i \to \mathbb{E}[X]$. One can generate $N$ samples and use the average value as an estimate. The uncertainty is given by the variance of $\frac{1}{N}\sum_{i=1}^{n} X_i$ which is $\frac{1}{N}\mathbb{V}[X]$.

3. The middle figures has the most uncertainty there for it is $p(\mu|y_1,\ldots y_{10})$, the figure to right has the least therefore it is $p(\mu|y_1,\ldots y_{70})$, and thus the figure to the left $p(\mu|y_1,\ldots y_{20})$

4. a) It is used to measure how well a model fits the data. Where the lowest WAIC indicates the best model.
   b) The first term measures how well the model predicts the data. The second term measures how flexible the model is (basically how many parameters there is in the model).
   c) The WAIC can be used to weighting different models against each other it done through the following equation:
   $$w_i = \frac{\exp(-\frac{1}{2}dWAIC_i)}{\sum_{j=1}^{m}\exp(-\frac{1}{2}dWAIC_j)}$$
   where $dWAIC_j = WAIC_i - WAIC_{min}$.

5. a) The 90% interval is approximately $[8.8, 11]$ and the map is approx 9.8.
   b) Yes the MCMC samples appears to be in a stationary state.

6. 

$$\begin{aligned} y_i &\sim Bin(1, p_i),\ i = 1, 2, \ldots, n_1 + n_2 \\ logit(p_i) &= \alpha + drug_i\beta_1 + diabetes_i\beta_2. \\ \alpha &\sim N(0, 100) \\ \beta_1 &\sim N(0, 100) \\ \beta_1 &\sim N(0, 100)\end{aligned}$$

   Here $drug_i$ is an indication if person $i$ has got the drug and $diabetes_i$ is if the person has diabetes or not. To check if there is an effect of the drug one checks if the posterior PI of $\beta_1$ contains zero.

7. 
```
data{
int<lower=1> n;
int<lower=0> y[n];
vector[n] x;
}
parameters{
real alpha;
real<lower=0> beta;
}
model{
```

```
beta ~ N(0,100);
alpha ~ N(0,100);
y ~ poisson_log(alpha + x*beta);
}
```

8. Using Bayes theorem we derive

$$p(z_1 = 1|...) \propto p(z_1 = 1|\theta)p(y_1|z_1 = 1, ...) = \theta N(y_1; \mu_1, \sigma_1),$$

and

$$p(z_1 = 0|...) \propto p(z_1 = 1|\theta)p(y_1|z_1 = 0, ...) = (1 - \theta)N(y_1; \mu_2, \sigma_2).$$

Since $z_i$ only takes two values

$$\begin{aligned} p(z_1 = 1|...) &= \frac{p(z_1 = 1|...)}{p(z_1 = 1|...) + p(z_1 = 0|...)} \\ p(z_1 = 0|...) &= 1 - p(z_1 = 1|...). \end{aligned}$$