1. a) We can calculate the expected utility for taking the \$100 as

$$U(100) = .62\log(.004 \times 100 + 1) = .209$$

While the expected utility for the gamble is

$$\begin{aligned} \mathrm{E}[U(r)] &= \frac{2}{3}.62\log(1) + \frac{1}{3}.62\log(.004 \times 500 + 1) \\ &= .227 \end{aligned}$$

Since the expected utility with the gamble is higher, Mr. Rubin should choose it.

b) The utility for not taking the bet is $U(0) = 0$. By comparison, the utility for taking the bet is

$$\begin{aligned} \mathrm{E}[U(r)] &= \frac{2}{3}.62\log(.004 \times -100 + 1) + \frac{1}{3}.62\log(.004 \times 400 + 1) \\ &= -.014 \end{aligned}$$

So the better choice is not to bet.

2. We can find the optimal q by taking the gradient with respect to q and finding where it is 0. These are all convex functions, so that will be the minimum. Also, I set $q_D = 1 - q_A - q_B - q_C$.

i)

$$\begin{aligned} \mathrm{E}[L(\theta; q)] &= \sum_{i \in \Theta} (1 - p_i)q_i^2 + p_i(q_i - 1)^2 \\ \frac{\partial}{\partial q_i}\mathrm{E}[L(\theta; q)] &= 2(1 - p_i)q_i + 2p_i(q_i - 1) - 2(1 - p_D)q_D - 2p_D(q_D - 1) \end{aligned}$$

Which is 0 only when $p_i = q_i$ for each i, making for a proper scoring rule.

ii)

$$\begin{aligned} \mathrm{E}[L(\theta; q)] &= \sum_{i \in \Theta} -p_i \log(q_i) \\ \frac{\partial}{\partial q_i}\mathrm{E}[L(\theta; q)] &= -\frac{p_i}{q_i} + \frac{p_d}{q_d} \end{aligned}$$

Which is also 0 when $p_i = q_i$ for each i, making for a proper scoring rule.

iii)

$$\begin{aligned} \mathrm{E}[L(\theta; q)] &= \sum_{i \in \Theta} p_i(1 - q_i) \\ \frac{\partial}{\partial q_i}\mathrm{E}[L(\theta; q)] &= -p_i + p_D \end{aligned}$$

Here, there is no minimum, but we are constrained within the probability simplex. This is a linear function, so the minimum is obviously achieved by setting $q_i = 1$ for whatever i has the largest p_i.

3. We can compare these two forecasters by summing up their mean loss over the 10 days, which serves as a frequentest point estimate of their expected loss on any given day. Doing so yields, for the two loss functions,

	Brier Loss	Log Loss
Forecaster 1	.330	.501
Forecaster 2	.524	.709

The first forecaster does better by both loss functions, so he is probably the better forecaster.

4. a) As I showed above, with the L_1 loss, the best strategy is to predict the most likely outcome will occur with probability 1. Doing so would yield the probabilities

$$1\,1\,1\,0\,1\,1\,0\,0\,0\,0$$

   b) The predictions in part a yield an average loss of .3. By comparison, using the original probabilities yields an average loss of .369.

5. We can calculate the risk of $\delta_0(x)$ by taking its expectation given θ:

$$\begin{aligned}
R(\theta, d) &= \mathrm{E}_\theta[L(\theta, d)] \\
&= \sum_{x=0}^{\infty} \frac{\theta^x e^{-\theta}}{x!} \frac{(\theta - x)^2}{\theta} \\
&= \sum_{x=0}^{\infty} \frac{\theta^x e^{-\theta}(\theta - 2x + \frac{x^2}{\theta})}{x!} \\
&= \theta - 2\theta + \sum_{x=0}^{\infty} x^2 \frac{\theta^{x-1} e^{-\theta}}{x!} \\
&= -\theta + 0 + \sum_{x=1}^{\infty} x \frac{\theta^{x-1} e^{-\theta}}{(x-1)!} \\
&= -\theta + \sum_{x=0}^{\infty} (x+1) \frac{\theta^x e^{-\theta}}{x!} \\
&= -\theta + \theta + 1 \\
&= 1
\end{aligned}$$

The integrated risk will be

$$\begin{aligned}
r(\pi, \delta) &= \int_{\mathbb{R}_+} R(\theta, \delta)\pi(d\theta) \\
&= \int_{\mathbb{R}_+} \mathrm{E}_\theta[R(x, \delta)] \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} d\theta \\
&= \int_{\mathbb{R}_+} \sum_{x=0}^{\infty} \frac{\theta^x e^{-\theta}(\theta - 2\delta + \frac{\delta^2}{\theta})}{x!} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} d\theta \\
&= \sum_{x=0}^{\infty} \frac{1}{x!} \frac{\beta^\alpha}{\Gamma(\alpha)} \int_{\mathbb{R}_+} (\theta - 2\delta + \frac{\delta^2}{\theta}) \theta^{x+\alpha-1} e^{-(\beta+1)\theta} d\theta \\
&= \sum_{x=0}^{\infty} \frac{1}{x!} \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{\Gamma(x+\alpha+1)}{(\beta+1)^{x+\alpha+1}} + -2\delta \frac{\Gamma(x+\alpha)}{(\beta+1)^{x+\alpha}} + \delta^2 \frac{\Gamma(x+\alpha-1)}{(\beta+1)^{x+\alpha-1}} \right) \\
&= \sum_{x=0}^{\infty} \frac{1}{x!} \frac{\beta^\alpha \Gamma(x+\alpha-1)}{(\beta+1)^{x+\alpha-1}\Gamma(\alpha)} \left(\frac{(x+\alpha)(x+\alpha-1)}{(\beta+1)^2} - 2\delta \frac{(x+\alpha-1)}{\beta+1} + \delta^2 \right) \\
\frac{\partial}{\partial\delta} r(\pi, \delta) &= \sum_{x=0}^{\infty} \frac{1}{x!} \frac{\beta^\alpha \Gamma(x+\alpha-1)}{(\beta+1)^{x+\alpha-1}\Gamma(\alpha)} \left(-2 \frac{(x+\alpha-1)}{\beta+1} + 2\delta \right)
\end{aligned}$$

This equals 0 when

$$\delta = \frac{(x+\alpha-1)}{\beta+1}$$

Since the loss function is convex, this is the optimal value and our Bayes estimator.

6. a)

$$
\begin{aligned}
R(\theta, \delta_c) &= \mathrm{E}_\theta[(\theta - cx)^2] \\
&= \sum_{x=0}^{\infty} (\theta^2 - 2cx\theta + c^2x^2)\frac{\theta^x e^{-\theta}}{x!} \\
&= \theta^2 - 2c\theta^2 + \sum_{x=1}^{\infty} c^2 x\theta \frac{\theta^{x-1}e^{-\theta}}{(x-1)!} \\
&= \theta^2 - 2c\theta^2 + c^2(\theta+1)\theta \\
&= (1 - 2c + c^2)\theta^2 + c^2\theta
\end{aligned}
$$

b) For $c > 1$,

$$
\begin{aligned}
R(\theta, \delta_c) &= (1 - 2c + c^2)\theta^2 + c^2\theta \\
&> \theta \\
&> R(\theta, \delta_1)
\end{aligned}
$$

This result does not depend on θ, so δ_c is strictly better, making δ_c inadmissible.

c) To derive the integrated risk:

$$
\begin{aligned}
r(\pi, \delta_c) &= \int_{\mathbb{R}_+} R(\theta, \delta_c) e^{-\theta} d\theta \\
r(\pi, \delta_c) &= \int_{\mathbb{R}_+} \left((1 - 2c + c^2)\theta^2 + c^2\theta\right) e^{-\theta} d\theta \\
&= -\left((1 - 2c + c^2)\theta^2 + c^2\theta\right) e^{-\theta}\Big]_0^\infty + \int_{\mathbb{R}_+} \left((1 - 2c + c^2)2\theta + c^2\right) e^{-\theta} d\theta \\
&= -0 - \left((1 - 2c + c^2)2\theta + c^2\right) e^{-\theta}\Big]_0^\infty + \int_{\mathbb{R}_+} 2(1 - 2c + c^2) e^{-\theta} d\theta \\
&= c^2 + 2(1 - 2c + c^2) \\
&= 2 - 4c + 3c^2
\end{aligned}
$$

d) Taking the derivative of the integrated risk and setting it to 0,

$$
\begin{aligned}
\frac{\partial}{\partial q_i} r(\pi, \delta_c) &= -4 + 6c \\
0 &= -4 + 6c \\
c &= \frac{2}{3}
\end{aligned}
$$

So we find that the optimal c is $\frac{2}{3}$.