1a.) let $y_i \sim Bin(\theta)$

$$P(y_i \mid \theta) = \prod_{i=1}^{100} \binom{100}{k_i} (\theta)^{k_i} (1-\theta)^{100-k_i}$$

$$= \theta^{\Sigma k} (1-\theta)^{100-\Sigma k} \prod_{i=1}^{100} \binom{100}{k_i}$$

by Factorization Theorem Σk_i is a sufficient statistic

let $T = \Sigma k_i$

$$P(y_i \mid \theta) = \theta^{T} (1-\theta)^{100-T}$$

b.) $$P(\theta \mid y) \propto \theta^{T}(1-\theta)^{100-T}\, \theta^{\alpha-1}(1-\theta)^{\beta-1}$$

$$\propto \theta^{T+\alpha-1} (1-\theta)^{100-T+\beta-1}$$

$$P(\theta \mid T) \sim Beta(T+a,\ 100+b-T)$$

c.) $Unif(0,1) \sim Beta(1,1)$: $a = \beta = 1$; $Beta(T+1,\ 100+1-T)$

$$E[\theta \mid T] = \frac{58}{58+44} = 0.569 \qquad [0.47, 0.66]$$

d.) $Beta(1,3)$; $Beta(3,1)$

$$E[\theta \mid T, a=1, b=3] = \frac{58}{58+52} = 0.569 \qquad [0.48, 0.67]$$

$$E[\theta \mid T, a=3, b=1] = \frac{60}{60+44} = 0.577 \qquad [0.46, 0.65]$$

Notice that the expectation and the 95% quantile based credible interval has not changed drastically and are very similar to one another

e.) $$E\left[\frac{\theta}{1-\theta}\right] = 0.281 \qquad [-0.113, 0.681]$$

The Posterior curve is the updated belief after seeing more data. One can see that the Posterior distribution has a extreme peak and narrow interval indicating high certainty. In contrast, the prior is your prior belief of the parameter theta and thus has a much wider interval and is less certain.

2a.) $P(\theta | y_1, y_2 .. y_{10}) = \prod_{n=1}^{10} \left(\frac{e^{-\theta} \theta^{i}}{i!} \right) P(\theta)$

$$= \frac{e^{-10\theta} \theta^{\Sigma i}}{\prod_{n=1}^{10} i!} \cdot \frac{\beta^{a} \theta^{a-1} e^{-\beta\theta}}{\Gamma a} = \frac{\beta \theta^{T+a-1} \exp[-\beta\theta - 10\theta]}{\prod_{i=1}^{10} i! \, \Gamma a}$$

$\propto \beta \theta^{T+a-1} \exp[-\theta(10+\beta)]$ $P(\theta | y_1 .. y_{10}) \sim \text{Gamma}(T+a, 10+\beta)$

Thus. $A \sim \text{Gamma}(T_A + 120, 20)$ $B \sim \text{Gamma}(T_B + 12, 13)$

Both distributions for θ_A and θ_B are extremely similar. The main difference is that the centrality of the two distributions are in different locations.

b.) $E[\theta_A | T_A, a, b] = \frac{\alpha}{\beta} = \frac{117+120}{20} = 11.85$ $[10.365, 13.407]$

$E[\theta_B | T_B, a, b] = \frac{\ddot{\alpha}}{\beta} = \frac{125}{14} = 8.928$ $[7.41, 10.56]$

c.) $E\left[\frac{\theta_A}{\theta_B}\right] = 1.38$ $[1.076, 1.66]$ I observe that the expectation, as well as the credible intervals are lower than that of question b. Even the plot seems more certain when looking at the ratio of the two campuses to θ_1 and θ_2 separately

d.)

e.) The event that $\theta_A > \theta_B$ is not that sensitive for $n \in [1, 20]$ but as soon as we go $n > 20$ the probability falls at a very high rate to near 0 once $n \approx 25$

3a.) $P(y|\theta) = \theta^{15}(1-\theta)^{28}$

$P(\theta|y) \propto \theta^{15}(1-\theta)^{28} \left[\frac{\omega \theta^{2-1}(1-\theta)^{8-1}}{\beta(2,8)} + \frac{(1-\omega)\theta^{8-1}(1-\theta)^{2-1}}{\beta(8,2)} \right]$

$P(\theta|y) \propto \omega \theta^{17-1}(1-\theta)^{36-1} + (1-\omega)\theta^{23-1}(1-\theta)^{30-1}$

$P(\theta|y) \propto \omega \text{Beta}(17,36) + (1-\omega) \text{Beta}(23,30)$ if we multiply by a constant 7

$P(\theta|y) \propto \omega \, \Gamma 17 \, \Gamma 36 \, \text{Beta}(17,36) + (1-\omega) \, \Gamma 23 \, \Gamma 30 \, \text{Beta}(23,30)$

3 a.) $\int_0^{\infty} P(\theta|y) = 1$

$$C \int_0^{\infty} \left[\omega \Gamma 17 \Gamma 36 \; Beta(17,36) + (1-\omega) \Gamma 23 \Gamma 30 \; Beta(23,30) \right] d\theta = 1$$

$$C = \left(\omega \Gamma 17 \Gamma 36 + (1-\omega) \Gamma 23 \Gamma 30 \right)^{-1} \quad \text{because } \int_0^{1} Beta = 1$$

Thus ; $$\frac{\omega \Gamma 17 \Gamma 36 \; Beta(17,36) + (1-\omega) \; Beta(23,30) \Gamma 23 \Gamma 36}{\omega \Gamma 17 \Gamma 36 + (1-\omega) \Gamma 23 \Gamma 30}$$

$$\omega_* = \frac{\omega \Gamma 17 \Gamma 36}{\omega \Gamma 17 \Gamma 36 + (1-\omega) \Gamma 23 \Gamma 30} \qquad (1-\omega_*) = \frac{(1-\omega) \Gamma 23 \Gamma 30}{\omega \Gamma 17 \Gamma 36 + (1-\omega) \Gamma 23 \Gamma 30}$$

$$\omega_* Beta(17,36) + (1-\omega_*) Beta(23,30)$$

4.) $$P(\alpha,\beta|\gamma) = \frac{P(\alpha,\beta,\gamma)}{P(\gamma)}$$

$$= \frac{P(\alpha|\beta\gamma) P(\cancel{\alpha}\beta)}{P(\gamma)}$$

$$= \frac{P(\alpha|\beta\gamma) P(\beta|\gamma) \cancel{P(\gamma)}}{\cancel{P(\gamma)}}$$

$$= P(\alpha|\beta\gamma) P(\beta|\gamma)$$