Problem 1.) 3 Doors $= \{D_1, D_2, D_3\}$ with $P(D_i) = \frac{1}{3}$ $\forall i = 1, 2, 3$. ①

~~[illegible]~~ Suppose my friend chooses door 1; then the gameshow host will choose either door 2 or door 3. The probability that he chooses door 2 or door 3 is $\frac{1}{2}$ because the first door was chosen already. So, we can state that

$$P(C_H) = \frac{1}{2}$$

is the probability that the host chooses some door that is available. Then, we would observe the following; assuming we are considering D_3.

$P(C_H | D_1) = \frac{1}{2}$; $P(C_H | D_2) = 1$; $P(C_H | D_3) = 0$.

If my friend's original choice was correct, the host will choose door 3 with probability $\frac{1}{2}$; If the car is in door 2, the host will choose door 3 with probability 1; and, finally, if the car is in door 3, the host will ~~[illegible]~~ choose 3 with probability 0.

Using Bayes Rule, we have the following:

$$P(D_1 | C_H) = \frac{P(C_H | D_1)\, P(D_1)}{P(C_H)} = \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2}} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

$$P(D_2 | C_H) = \frac{P(C_H | D_2)\, P(D_2)}{P(C_H)} = \frac{1 \times \frac{1}{3}}{\frac{1}{2}} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

So, the posterior probability of the prize being in door 2 given that we chose door 1 in the first round of the game and the host revealed door 3 is higher than the posterior odds for door 1.

$P(D_2 | C_H) > P(D_1 | C_H)$.

Problem 2.) $\pi = (\pi_1, \pi_2, \ldots, \pi_k)$; $\pi_j \geq 0$.

Let $X_i \overset{iid}{\sim} \text{Multinomial}(\pi)$ $\forall i = 1, \ldots, N$.

where
$$f(x_i | \pi) = \frac{\Gamma(n+1)}{\prod_{j=1}^{K} \Gamma(x_j + 1)} \prod_{j=1}^{K} \pi_j^{x_j} \quad \forall i = 1, \ldots, N$$

The likelihood for this is then,

$$p(x_{1:N} | \pi) = \prod_{i=1}^{N} \left[\frac{\Gamma(n+1)}{\prod_{j=1}^{K} \Gamma(x_j + 1)} \prod_{j=1}^{K} \pi_j^{x_{ji}} \right]$$

$$\propto \prod_{i=1}^{N} \prod_{j=1}^{K} \pi_j^{x_{ij}} = \prod_{j=1}^{K} \pi_j^{\sum_{i=1}^{N} x_{ij}}$$

A well known conjugate prior distribution for the multinomial is the Dirichlet distribution, which is defined as

$$f(x_i | \alpha_1, \ldots, \alpha_K) = \frac{\Gamma\left(\sum_{\ell=1}^{K} \alpha_\ell\right)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_K)} \prod_{\ell=1}^{K} x_i^{\alpha_\ell - 1}$$

$$\propto \prod_{\ell=1}^{K} x_i^{\alpha_\ell - 1}$$ substituting for π_j yields

$$f(\pi_j | \alpha_1, \ldots, \alpha_K) \propto \prod_{j=1}^{K} \pi_j^{\alpha_j - 1}.$$

Recall the definition of the posterior of $\theta | x$ is:

$$p(\theta | x) = \frac{p(\theta)\, p(x|\theta)}{\int_{\Theta} p(\theta)\, p(x|\theta)\, d\theta}$$

Plugging in our terms yields (continued on next page)...

$$p(\pi \mid x_{1:N}, \alpha_{1:K}) = \frac{p(\pi)\, p(x_{1:N} \mid \pi)}{\int_{\Theta} p(\pi)\, p(x_{1:N} \mid \pi)\, d\pi}$$

$$= \frac{\left[\frac{\Gamma(\sum^{K} \alpha_j)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_K)} \prod_{j=1}^{K} \pi_j^{\alpha_j - 1}\right] \times \left[\prod_{i=1}^{N} \prod_{j=1}^{K} \pi_j^{x_{ij}} \frac{\Gamma(n+1)}{\prod_{j=1} \Gamma(x_j+1)}\right]}{\int \left[\frac{\Gamma(\sum^{K} \alpha_j)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_K)} \prod_{j=1}^{K} \pi_j^{\alpha_j - 1}\right] \times \left[\prod_{j=1}^{K} \pi_j^{\sum_i^N x_{ij}} \frac{\Gamma(n+1)}{\prod_{j=1}^{K} \Gamma(x_j+1)}\right] d\pi}$$

$$\propto \prod_{j=1}^{K} \pi_j^{\alpha_j - 1} \times \prod_{j=1}^{K} \pi_j^{\sum_{i=1}^{N} x_{ij}} = \prod_{j=1}^{K} \pi_j^{\alpha_j + \sum_{i=1}^{N} x_{ij} - 1}$$

which is, indeed, a Dirichlet distribution with parameters.

$$Dir\left(\alpha_1 + \sum_{\substack{i=1 \\ \forall i \in j}}^{N} x_{ij}, \ldots, \alpha_K + \sum_{\substack{i=1 \\ \forall i \in k}}^{N} x_{ij}\right)$$

The most obvious features about the parameters of this distribution are: (1) the hyperparameters in the prior correspond to prior sample sizes, and the posterior incorporates this into its expectation. (2) It is also the multivariate extension of the Beta Distribution. (3) It is the same distribution as the prior.

~~Calculate~~ $\cancel{E_Q[\ln \pi_j] = \ln\left(\frac{\alpha_j}{\sum_{j=1}^{K} \alpha_k}\right) = \ln \alpha_j - \ln \sum_{j=1}^{K} \alpha_j.}$

Detailed on next page.

Problem 3) a) Let $\{x_1, \ldots, x_n\}$ where $x \in \mathbb{R}$

$x_i \overset{iid}{\sim} Normal(\mu, \lambda^{-1})$

$\mu|\lambda \sim Normal(0, a\lambda^{-1})$

$\lambda \sim Gamma(b, c)$

$$p(x_{1:n}|\mu, \lambda^{-1}) = \prod_{i=1}^{n} (2\pi \tfrac{1}{\lambda})^{-\frac{n}{2}} \exp\left[-\frac{\lambda}{2} \sum_{i=1}^{n} (x_i - \mu)^2\right]$$

$$p(\mu|\lambda) = (2\pi \tfrac{a}{\lambda})^{-1/2} \exp\left[\frac{-\lambda}{2a}(\mu^2)\right]$$

$$p(\lambda) = \frac{c^b}{\Gamma(b)} \lambda^{b-1} e^{-c\lambda}$$

$$p(\mu|\lambda, x_i) \propto \frac{1}{\sqrt{2\pi \frac{a}{\lambda}}} \exp\left[\frac{-\lambda}{2a}\mu^2\right] \times \left[(2\pi \tfrac{1}{\lambda})^{-\frac{n}{2}} \exp\left[\frac{-\lambda}{2} \sum_{i=1}^{n} (x_i - \mu)^2\right]\right]$$

$$= (2\pi \tfrac{a}{\lambda})^{-\frac{1}{2}} (2\pi \tfrac{1}{\lambda})^{-\frac{n}{2}} \exp\left[\frac{-\lambda}{2a}\mu^2 - \frac{\lambda}{2} \sum (x_i - \mu)^2\right]$$

$$= (2\pi)^{-(\frac{n+1}{2})} a^{-\frac{1}{2}} \lambda^{\frac{n}{2}} \exp\left[\frac{-\lambda}{2}\left(a\mu^2 + \sum_{i=1}^{n} (x_i - \mu)^2\right)\right]$$

$$(*) = (2\pi)^{-(\frac{n+1}{2})} a^{-\frac{1}{2}} \lambda^{\frac{n}{2}} \exp\left[\frac{-\lambda}{2}\left(\mu^2(n+a) - 2n\bar{x}\mu + (n\bar{x})^2\right)\right]$$

Which is a Normal with $N(\mu^*, \frac{a^*}{\lambda^*})$ which reduces $(*)$ to:

$$= (2\pi)^{-(\frac{n+1}{2})} a^{-\frac{1}{2}} \lambda^{\frac{n}{2}} \exp\left[\frac{-\lambda}{2a}\left(\frac{(n\bar{x} - \mu)^2}{a} + \frac{2\mu}{\lambda}\right)\right]$$

$$= (2\pi)^{-(\frac{n+1}{2})} a^{-\frac{1}{2}} \lambda^{\frac{n}{2}} \exp\left[\frac{-\lambda}{a}\left(\frac{n\bar{x} - \mu^2}{2a} + \frac{\mu}{\lambda}\right)\right]$$

$$\mu|x, \lambda \sim N\left(\frac{n\bar{x} - \mu^2}{2a}, \frac{a}{\lambda}\right).$$

Problem 3) a. Find $p(\lambda|x)$, recall that, 6

$$p(\lambda|x) = \frac{\int_{-\infty}^{\infty} p(\lambda)p(\mu|\lambda)p(x|\lambda,\mu)d\mu}{\int_0^{\infty}\int_{-\infty}^{\infty} p(\lambda)p(\mu|\lambda)p(x|\lambda,\mu)d\mu d\lambda}$$

So, $$p(\lambda|x) \propto \int_{-\infty}^{\infty} p(\lambda)p(\mu|\lambda)p(x|\lambda,\mu)d\mu.$$

Using the definitions of our priors & likelihood we have:

$$p(\lambda|x) \propto \int_{-\infty}^{\infty} \frac{c^b}{\Gamma(b)} \lambda^{b-1} \exp[-c\lambda]\left(2\pi\frac{a}{\lambda}\right)^{-\frac{1}{2}}\left(\frac{2\pi}{\lambda}\right)^{-\frac{n}{2}} \exp\left[\frac{-\lambda}{2a}\mu^2\right]\exp\left[\frac{-\lambda}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right]d\mu.$$

$$= \frac{c^b}{\Gamma(b)}(2\pi)^{-\left(\frac{n+1}{2}\right)} a^{-\frac{1}{2}} \lambda^{\frac{n}{2}+b-1+\frac{1}{2}} \exp[-c\lambda]\int_{-\infty}^{\infty}\exp\left[\frac{-\lambda}{2}\left((n\bar{x})^2-2n\bar{x}\mu+n\mu^2+\frac{1}{a}\mu^2\right)\right]d\mu.$$

$$= \frac{c^b}{\Gamma(b)}(2\pi)^{-\left(\frac{n+1}{2}\right)} a^{-\frac{1}{2}} \lambda^{\frac{n}{2}+b-\frac{1}{2}} \exp[-c\lambda]\int_{-\infty}^{\infty}\exp\left[\frac{-\lambda}{2}\left(\mu^2\left(n+\frac{1}{a}\right)-2n\bar{x}+(n\bar{x})^2\right)\right]d\mu$$

Recall the general gaussian integral identity

$$\int_{-\infty}^{\infty}\exp[-(ax^2+bx+c)]dx = \sqrt{\frac{\pi}{a}}\exp\left[\frac{b^2-4ac}{4a}\right]$$

Then let $a = \frac{\lambda\left(n+\frac{1}{a}\right)}{2}$; $b = [\lambda(n\bar{x})]^2$; $c = \frac{\lambda}{2}(n\bar{x})^2$

Which yields

$$p(\lambda|x) \propto \frac{c^b}{\Gamma(b)}\lambda^{\frac{n}{2}+b-\frac{1}{2}}\exp[-c\lambda](2\pi)^{-\left(\frac{n+1}{2}\right)}a^{-\frac{1}{2}}\left(\frac{\sqrt{2\pi}}{\sqrt{\lambda\left(n+\frac{1}{a}\right)}}\right)\exp\left[\frac{\lambda^2(n\bar{x})^2-4\left(\frac{\lambda}{2}\right)^2\left(n+\frac{1}{a}\right)(n\bar{x})^2}{4\frac{\lambda}{2}\left(n+\frac{1}{a}\right)}\right]$$

$$\propto \frac{c^b}{\Gamma(b)}\lambda^{\frac{n}{2}+b-\frac{1}{2}-\frac{1}{2}}(2\pi)^{-n}\left(n+\frac{1}{a}\right)^{-\frac{1}{2}}a^{-\frac{1}{2}}\exp[-c\lambda]\exp\left[\frac{\lambda(n\bar{x})^2-\lambda\left(n+\frac{1}{a}\right)(n\bar{x})^2}{2\left(n+\frac{1}{a}\right)}\right]$$

$$\propto \frac{c^b}{\Gamma(b)}\lambda^{\frac{n}{2}+b-1}(2\pi)^{-n}\left(n+\frac{1}{a}\right)^{-\frac{1}{2}-1}a^{-\frac{1}{2}}\exp\left[-\lambda\left(\frac{(n\bar{x})^2\left(n+\frac{1}{a}-1\right)}{2\left(n+\frac{1}{a}\right)}+c\right)\right]$$

$$\propto \frac{c^b}{\Gamma(b)}\frac{1}{2\pi^n\sqrt{n+\frac{1}{a}}\times\sqrt{a}}\lambda^{\frac{n}{2}+b-1}\exp\left[-\lambda\left(\frac{(n\bar{x})^2\left(n+\frac{1}{a}-1\right)}{2\left(n+\frac{1}{a}\right)}+c\right)\right]$$

$$p(\lambda|x) \sim \Gamma\left(b+\frac{n}{2},\ \frac{(n\bar{x})^2\left(n+\frac{1}{a}-1\right)}{2\left(n+\frac{1}{a}\right)}+c\right)$$

Problem 3) b) $p(x^*|x) = \iint p(x^*|\mu,\lambda^{-1})\, p(x|\mu,\lambda)\, p(\mu|\lambda)\, p(\lambda)\, d\mu\, d\lambda$

$$p(x^*|x) = \int_0^{\infty}\int_{-\infty}^{\infty} \frac{c^b}{\Gamma(b)} \lambda^{b-1} e^{-c\lambda} \lambda^{\frac{n}{2}+\frac{1}{2}} (2\pi)^{-\frac{n}{2}-\frac{1}{2}-\frac{1}{2}} a^{-\frac{1}{2}} \exp\left[\frac{-\lambda}{2}\sum(x_i-\mu)^2\right] \exp\left[-(x^*-\mu)^2\frac{\lambda}{2}\right]$$

$$\times \exp\left[\frac{-\lambda}{2a}\mu^2\right] d\mu\, d\lambda.$$

$$= \frac{c^b}{\Gamma(b)} (2\pi)^{-\left(\frac{n+3}{2}\right)} a^{-1/2} \int_0^{\infty} \lambda^{\frac{n-1}{2}+b} \exp[-c\lambda] \int_{-\infty}^{\infty} \exp\left[\frac{-\lambda}{2}\left(\underbrace{\sum(x_i-\mu)^2}_{\text{step 1}} + (x^*-\mu)^2 + \frac{\mu^2}{a}\right)\right] d\mu\, d\lambda.$$

$$\text{step 1} = \int_{-\infty}^{\infty} \exp\left[\frac{-\lambda}{2}\left[n^2\bar{x}^2 - 2\mu n\bar{x} + n\mu^2 + x^{*2} - 2\mu x^* + \mu^2 + \frac{\mu^2}{a}\right]\right] d\mu.$$

$$= \int_{-\infty}^{\infty} \exp\left[\frac{-\lambda}{2}\left(n^2\bar{x}^2 + x^{*2} + \left(\frac{na+a+1}{a}\right)\mu^2 - 2\mu(n\bar{x}+x^*)\right)\right]$$

$$= \sqrt{\frac{2\pi a}{an+a+1}} \exp\left[-\lambda\left(\frac{\left(\frac{na+a+1}{a}\right)(n\bar{x})^2 + \left(\frac{na+a+1}{a}\right)x^{*2} - (n\bar{x})^2 - x^{*2} - 2n\bar{x}x^*}{2\left(\frac{na+a+1}{a}\right)}\right)\right]$$

$$p(x^*|x) = \frac{c^b}{\Gamma(b)} (2\pi)^{-\left(\frac{n+3}{2}\right)} a^{-1/2} \int_0^{\infty} \sqrt{\frac{2\pi a}{an+a+1}}\, \underbrace{\exp\left[-\lambda\left(\frac{\left(\frac{na+a+1}{a}-1\right)(n\bar{x})^2 + \left(\frac{na+a+1}{a}-1\right)x^{*2} - 2n\bar{x}x^*}{2\left(\frac{na+a+1}{a}\right)}\right)\right]}_{\text{Step 2}} \lambda^{\frac{n-1}{2}+b} d\lambda$$

$$\text{Step 2} = \int_0^{\infty} \lambda^{\frac{n-1}{2}+b} \exp\left[-\lambda\left(\frac{(na+1)x^{*2} - 2an\bar{x}x^*}{2(na+a+1)} + \frac{(an+1)(n\bar{x})^2}{2(na+a+1)} + c - 1 + 1\right)\right] d\lambda$$

$$\int_0^{\infty} \lambda^{\frac{n-1}{2}+b} \exp\left[\frac{-\lambda}{2(na+a+1)}\left((na+1)x^{*2} - 2x^*(\bar{x}an) + \underbrace{c + (na+1)(n\bar{x})^2 + 1}_{k} - 1\right)\right] d\lambda$$

$$\int_0^{\infty} \lambda^{\frac{n-1}{2}+b} \exp\left[\frac{-\lambda \cdot na}{2(na+a+1)}\left(x^{*2}\left(1+\frac{1}{na}\right) - 2x^*\bar{x} + k\right) + 1\right] d\lambda$$

$$\int_0^{\infty} \lambda^{\frac{n-1}{2}+b} \exp\left[\frac{-\lambda}{2(na+a+1)}\left((1+an)\left(x^* - \frac{n\bar{x}}{\frac{1}{a}+n}\right)^2\right) + 1\right] d\lambda$$

$$P(x^*|x) = \frac{c^b}{\Gamma(b)}(2\pi)^{-\left(\frac{n+1}{2}\right)}(na+a+1)^{-\frac{1}{2}}\int_0^\infty \lambda^{\frac{n-1}{2}+b} \exp\left[\frac{-\lambda}{2(na+a+1)}\left((1+an)\left(x^* - \frac{n\bar{x}}{\frac{1}{a}+n}\right)^2\right)+1\right]d\lambda$$

Recall that $\int_0^\infty x^{\alpha-1}\exp[-\beta x]dx = \frac{\Gamma(\alpha)}{\beta^\alpha}$, which yields

$$P(x^*|x) = \frac{c^b}{\Gamma(b)}(2\pi)^{-\left(\frac{n+1}{2}\right)}(na+a+1)^{-\frac{1}{2}}\Gamma\left(\frac{n+1}{2}+c\right)\left(\frac{(1+an)\left(x^* - \frac{n\bar{x}}{\frac{1}{a}+n}\right)}{2(na+a+1)}+1\right)^{-\left(\frac{n+1}{2}+b\right)}$$

which is a nonstandardized student t.

Problem 4.) a) See Code

b.) See Output

	Predicted 4	Predicted 9
Actual 4	926	56
Actual 9	156	853

Accuracy = 0.89352084

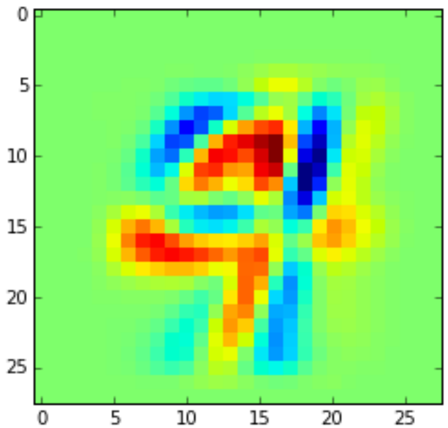
c.) See Output

d.) See Output.

	Predicted Class = 4	Predicted Class = 9
Actual Class = 4	926	56
Actual Class = 9	156	853

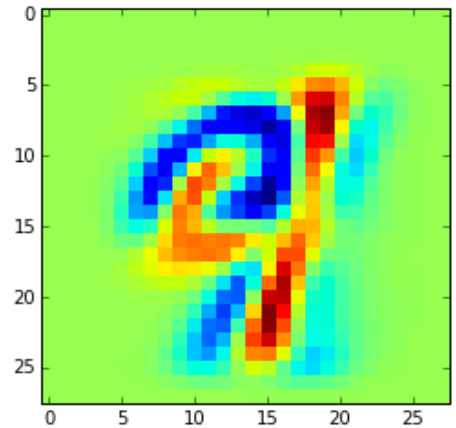
Misclassified

Observation 16
[Probability=9] = 0.50024



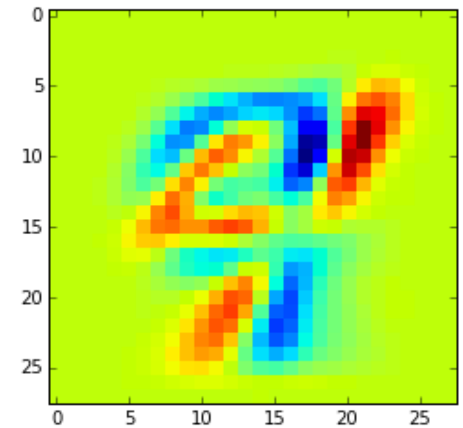
Actual Class = 4
Predicted Class = 9

Observation 21
[Probability=9] = 0.5132275



Actual Class = 4
Predicted Class = 9

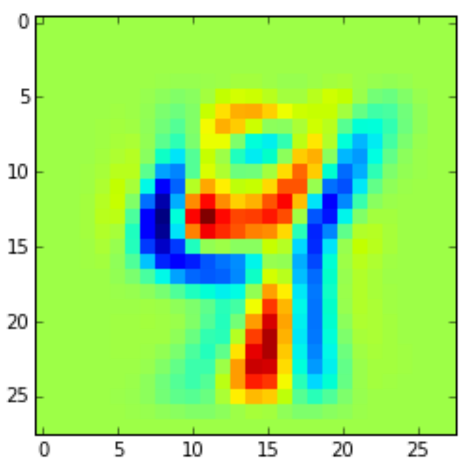
Observation 57
[Probability=9] = 0.5056573



Actual Class = 4
Predicted Class = 9

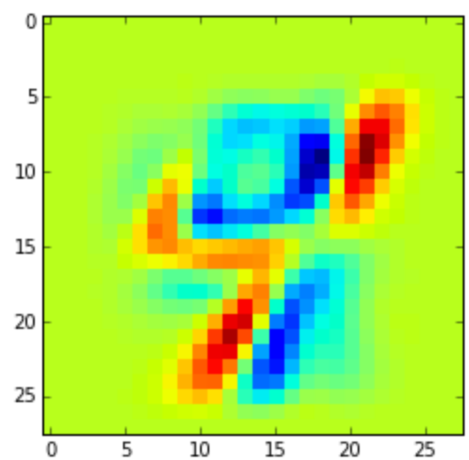
Most Ambiguous

Observation 1353
[Probability=9] = 0.5000133



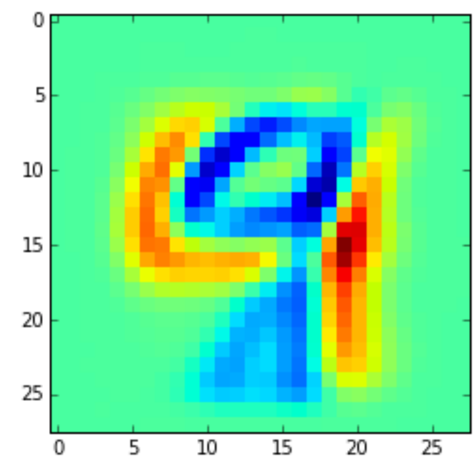
Actual Class = 9
Predicted Class = 4

Observation 1687
[Probability=9] = 0.5000153



Actual Class = 9
Predicted Class = 9

Observation 301
[Probability=9] = 0.5001167



Actual Class = 4
Predicted Class = 9