Statistical Estimation	**Midterm 1**
ASEN 5044 Fall 2018	**Due Date: Oct 11, 2018**
Name: Andrew Kramer	**PhD Student**

Problem 1

Inverted pendulum with equations of motion:

$$(M+m)\ddot{z} - ml\ddot{\theta}\cos\theta + ml\dot{\theta}^2\sin\theta = P$$
$$l\ddot{\theta} - g\sin\theta = \ddot{z}\cos\theta$$

Part (a)

The system's state equations can be expressed as follows:

$$\dot{x} = \begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\\ \dot{x}_4\end{bmatrix} = \begin{bmatrix}f_1(x,u,t)\\ f_2(x,u,t)\\ f_3(x,u,t)\\ f_4(x,u,t)\end{bmatrix} = \begin{bmatrix} x_2 \\ \frac{P - g\sin x_3\cos x_3 - mlx_4^2\sin x_3}{M+m\sin^2 x_3} \\ x_4 \\ \frac{P\cos x_3 - mlx_4^2\sin x_3\cos x_3 + (M+m)g\sin x_3}{Ml+ml\sin^2 x_3}\end{bmatrix}$$

To demonstrate the system is in equilibrium at $\dot{z}=0,\ \theta=0,\ \dot{\theta}=0$, and $P(t)=0$ we note first that at the given conditions the equations of motion become

$$(M+m)\ddot{z} - ml\ddot{\theta} = 0$$
$$l\ddot{\theta} = \ddot{z}$$

If we plug the second equation back into the first we get

$$(M+m)\ddot{z} - m\ddot{z} = M\ddot{z} = 0$$

Because we know M is not equal to zero, this means $\ddot{z}$ must be equal to zero. Additionally, because $\ddot{z} = l\ddot{\theta}$ and $l \neq 0$ we can also conclude that $\ddot{\theta} = 0$. This means $\dot{x} = 0$ under the given conditions and the system is therefore in equilibrium.

Part (b)

We'll begin by noting that our measurement function in state space form is

$$y = h(x,u,t) = x_1 - l\sin x_3$$

Next we define $x_{\text{nom}} = [0,0,0,0]^T$, $u_{\text{nom}} = 0$ and $y_{\text{nom}} = 0$. We can then define $\tilde{x}$, $\tilde{y}$, and $\tilde{u}$ where $\tilde{x}(t) = x(t) - x_{\text{nom}}(t)$, and $\tilde{y}$ and $\tilde{u}$ are defined similarly. We can then find $A|_{\text{nom}}$, $B|_{\text{nom}}$, $C|_{\text{nom}}$, and $D|_{\text{nom}}$ such that

$$\dot{\tilde{x}} = A|_{\text{nom}}\tilde{x}(t) + B|_{\text{nom}}\tilde{u}(t)$$
$$\tilde{y} = C|_{\text{nom}}\tilde{x}(t) + D|_{\text{nom}}\tilde{u}(t)$$

The above matrices are defined as

$$A|_{\text{nom}} = \left.\frac{\partial f}{\partial x}\right|_{\text{nom}} \qquad B|_{\text{nom}} = \left.\frac{\partial f}{\partial u}\right|_{\text{nom}}$$
$$C|_{\text{nom}} = \left.\frac{\partial h}{\partial x}\right|_{\text{nom}} \qquad D|_{\text{nom}} = \left.\frac{\partial h}{\partial u}\right|_{\text{nom}}$$

The jacobians are as follows

$$\frac{\partial f}{\partial x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{\partial f_2}{\partial x_3} & \frac{\partial f_2}{\partial x_4} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{\partial f_4}{\partial x_3} & \frac{\partial f_4}{\partial x_4} \end{bmatrix}$$
$$\frac{\partial f}{\partial u} = \begin{bmatrix} 0 \\ \frac{1}{M+m\sin^2 x_3} \\ 0 \\ \frac{\cos x_3}{Ml+ml\sin^2 x_3} \end{bmatrix}$$
$$\frac{\partial h}{\partial x} = \begin{bmatrix} 1 & 0 & -l\cos x_3 & 0 \end{bmatrix}$$
$$\frac{\partial h}{\partial u} = [0]$$

where

$$\frac{\partial f_2}{\partial x_3} = \frac{-2m(P - g\sin x_3\cos x_3 - lmx_4^2\sin x_3)\sin x_3\cos x_3}{(M+m\sin^2 x_3)^2} + \frac{g\sin^2 x_3 - g\cos^2 x_3 - lmx_4^2\cos x_3}{M+m\sin^2 x_3}$$
$$\frac{\partial f_2}{\partial x_4} = -\frac{2lmx_4\sin x_3}{M+m\sin^2 x_3}$$
$$\frac{\partial f_4}{\partial x_3} = \frac{-2lm(P\cos x_3 + g(M+m)\sin x_3 - lmx_4^2\sin x_3\cos x_3)\sin x_3\cos x_3}{(Ml+ml\sin^2 x_3)^2} + \frac{-P\sin x_3 + g(M+m)\cos x_3 + lmx_4^2\sin^2 x_3 - lmx_4^2\cos^2 x_3}{Ml+ml\sin^2 x_3}$$
$$\frac{\partial f_4}{\partial x_4} = -\frac{2mlx_4\sin x_3\cos x_3}{ML+ml\sin^2 x_3}$$

When these jacobians are evaluated at x_{nom}, u_{nom} we get the following matrices for our linearized state-space equations:

$$A|_{\text{nom}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2m - \frac{g}{M} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{g(M+m)}{Ml} & 0 \end{bmatrix}$$
$$B|_{\text{nom}} = \begin{bmatrix} 0 \\ \frac{1}{M} \\ 0 \\ \frac{1}{Ml} \end{bmatrix}$$
$$C|_{\text{nom}} = \begin{bmatrix} 1 & 0 & -1 & 0 \end{bmatrix}$$
$$D|_{\text{nom}} = [0]$$

Part (c)

We find the discretized F and G matrices by taking $e^{\hat{A}\Delta t}$ where $\hat{A}$ is made by horizontally concatenating the A and B matrices and then adding enough rows of zeros to form a square matrix. F is the upper-left submatrix of $e^{\hat{A}\Delta t}$ with dimensions equal to A and G is the upper-right submatrix with dimensions equal to B:

$$F = \begin{bmatrix} 1 & 0.05 & -0.0036 & -0.0001 \\ 0 & 1 & -0.1461 & -0.0036 \\ 0 & 0 & 1.0185 & 0.0503 \\ 0 & 0 & 0.7403 & 1.0185 \end{bmatrix}, \qquad G = \begin{bmatrix} 0.0006 & 0.025 & 0.0006 & 0.0252 \end{bmatrix}$$

The H and M matrices are simply equal to the C and D matrices for the continuous time system. Because one of the eigenvalues of F lies outside the unit circle in the complex plane we can say the system is asymptotically unstable.
The observability matrix $O = [H, HF, \ldots, HF^{n-1}]$ is full rank when n, the number of observations, is equal to 4. So the system is fully observable from 4 observations.

Part (d)

The system with full-state feedback can be modeled as $\dot{\tilde{x}}(t) = A\tilde{x}(t) - BK\tilde{x}(t)$, which can be rewritten as $\dot{\tilde{x}} = (A - BK)\tilde{x}$. So one could define $A_{CL} = (A - BK)$, and $B_{CL} = B$. Similarly, we can define $C_{CL} = C - DK$ and $D_{CL} = D$ because $y = (C - DK)\tilde{x}$.

$$A_{CL} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.0071 & 0.064 & -19.2471 & -1.364 \\ 0 & 0 & 0 & 1 \\ 0.0071 & 0.0640 & -1.6271 & -1.3640 \end{bmatrix} \qquad B_{CL} = \begin{bmatrix} 0 \\ 0.5 \\ 0 \\ 0.5 \end{bmatrix}$$

$$C_{CL} = \begin{bmatrix} 1 & 0 & -1 & 0 \end{bmatrix} \qquad D_{CL} = [0]$$

Part (e)

$$F_{CL} = \begin{bmatrix} 1 & 0.0501 & -0.024 & -0.0021 \\ 3.41e-4 & 1.0031 & -0.9605 & -0.0895 \\ 8.67e-6 & 7.84e-5 & 0.998 & 0.0483 \\ 3.43e-4 & 0.0031 & -0.0801 & 0.932 \end{bmatrix} \qquad G_{CL} = \begin{bmatrix} 6.092e-4 \\ 0.024 \\ 6.115e-4 \\ 0.0242 \end{bmatrix}$$

$$H_{CL} = \begin{bmatrix} 1 & 0 & -1 & 0 \end{bmatrix} M_{CL} = [0]$$

Because the eigenvalues of the F_{CL} matrix have complex magnitudes less than 1, the system is asymptotically stable. As in the zero input case, the observability matrix is full rank when $n = 4$, so the state is fully observable from 4 measurements.

Part (f)

We first construct an observation matrix for the system as $O = [H, HF, \ldots HF^{n-1}]$ for $n = 301$ because we have 301 observations. We find the initial state $x_0 = (O^T O)^{-1} O^T Y = [0.1776, -1.4567, -0.0155, 0.0435]^T$ where Y is the vector of stacked observations.

Part (g)

The predicted and actual measurements are plotted in figure 1. We can see the predicted measurements track the actual measurements fairly well. However, the changes in the predicted measurements clearly lag behind the actual measurements, and they seem to lag more as time goes on. Also, the predicted and actual measurements at $t = 0$ do not match, so our predicted initial state is not completely accurate.

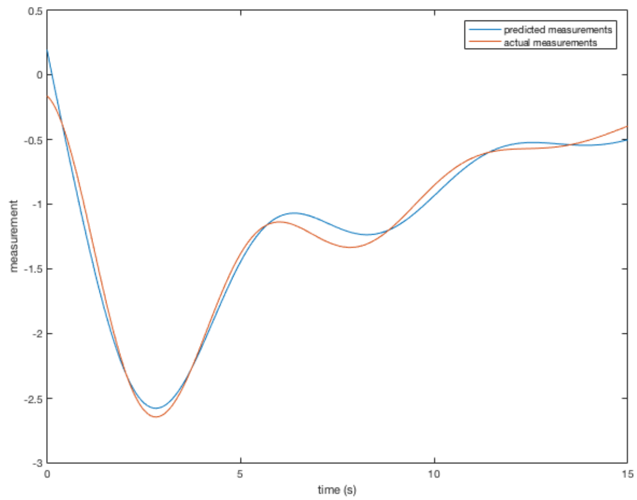

Figure 1: predicted vs actual measurements for x_0 from all observations

If we use only the first four observations to calculate x_0 we find $x_0 = [-0.0708, -0.2894, 0.0923, -0.0346]^T$. Interestingly this x_0 gives us a $y(0)$ that matches the actual measurement much better, as seen in figure 2. However, the predicted measurement deviate from the actual measurements much more than in the previous case, and we still see the increasing time lag between the actual and predicted measurements.

Problem 2

Two 6-sided dice rolls with R_1 and R_2 denoting the outcome of the first and second die, respectively.

Part (a)

$P(R_1) = P(R_2) = \frac{1}{6}$ for all R_1 and R_2. Because the outcomes R_1 and R_2 are independent, $P(R_1, R_2) = P(R_1) * P(R_2)$. So

$$P(R_1, R_2) = \frac{1}{36},\ \forall R_1, R_2$$

Part (b)

The joint probabilities for X and Y are shown in table 1 below

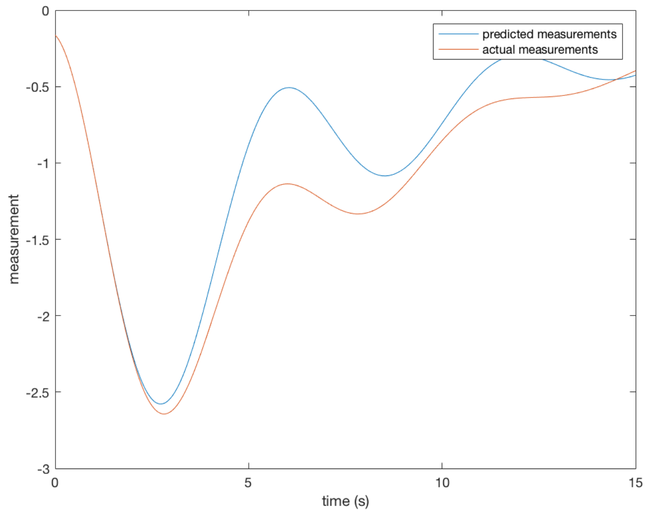

Figure 2: predicted vs actual measurements for x_0 from first 4 observations

Table 1: Joint Probabilities

X	Y=1	Y=2	Y=3	Y=4	Y=5	Y=6
1	1/36	2/36	2/36	2/36	2/36	2/36
2	0	1/36	2/36	2/36	2/36	2/36
3	0	0	1/36	2/36	2/36	2/36
4	0	0	0	1/36	2/36	2/36
5	0	0	0	0	1/36	2/36
6	0	0	0	0	0	1/36

Part (c)

The marginal probabilities of X obtained from the sum $\sum_y P(X = x, Y = y)$ and the marginal probabilities of Y obtained from the sum $\sum_x P(X = x, Y = y)$ are shown below in table 2.

Table 2: Marginal Probabilities

	X	Y
1	11/36	1/36
2	9/36	3/36
3	7/36	5/36
4	5/36	7/36
5	3/36	9/36
6	1/36	11/36

Part (d)

X and Y are not independent. Two variables are considered independent if the realization of one variable does not affect the probability of the other. This is not the case for X and Y. By the definitions of X and Y, the value of Y cannot be less than the value of X, since the maximum of R_1 and R_2 cannot be less than the minimum. So the $P(Y = 3, X = 5) = 0$, while $P(Y = 3, X = 1) = 2/36$.

Problem 3

A random variable X has the pdf $p(x) = k(1 - x^4)$ for $-1 \leq x \leq 1$ and $p(x) = 0$ elsewhere.

Part (a)

Because $\int_{-\infty}^{\infty} p(x)dx = 1$ we can find k as follows:

$$\begin{aligned}
\int_{-1}^{1} k(1 - x^4)dx &= 1 \\
k\int_{-1}^{1} (1 - x^4)dx &= \\
k\big[x - \frac{1}{5}x^5\big|_{-1}^{1} &= \\
k\Big(1 - \frac{1}{5} + 1 - \frac{1}{5}\Big) &= \\
\frac{8k}{5} &= 1 \\
k &= \frac{5}{8}
\end{aligned}$$

Now we can calculate $E[x] = \int_{-\infty}^{\infty} xp(x)dx$ as follows:

$$\begin{aligned}
\int_{-\infty}^{\infty} xp(x)dx &= \frac{5}{8}\int_{-1}^{1} x(1 - x^4)dx \\
&= \frac{5}{8}\int_{-1}^{1} (x - x^5)dx \\
&= \frac{5}{8}\left[\frac{1}{2}x^2 - \frac{1}{6}x^6\right|_{-1}^{1} \\
&= \frac{5}{8}\left(\frac{1}{2} - \frac{1}{6} - \frac{1}{2} + \frac{1}{6}\right) \\
&= 0
\end{aligned}$$

Next we find $E[x^2] = \int_{-\infty}^{\infty} x^2 p(x)dx$ as

$$\begin{aligned}
\int_{-\infty}^{\infty} x^2 p(x)dx &= \frac{5}{8}\int_{-1}^{1} x^2(1-x^4)dx \\
&= \frac{5}{8}\int_{-1}^{1} (x^2 - x^6)dx \\
&= \frac{5}{8}\left[\frac{1}{3}x^3 - \frac{1}{7}x^7\right|_{-1}^{1} \\
&= \frac{5}{8}\left(\frac{1}{3} - \frac{1}{7} + \frac{1}{3} - \frac{1}{7}\right) \\
&= \frac{5}{21}
\end{aligned}$$

Finally we can find $\mathrm{var}(x) = E[x^2] - (E[x])^2 = \frac{5}{21} - 0 = \frac{5}{21}$.

Part (b)

The cumulative distribution function is defined as $P_X(\zeta) = \int_{-\infty}^{\zeta} p(x)dx$. So the cdf is:

$$\begin{aligned}
P_X(\zeta) &= \int_{-\infty}^{\zeta} p(x)dx \\
&= \frac{5}{8}\int_{-1}^{\zeta} (1-x^4)dx \\
&= \frac{5}{8}\left[x - \frac{1}{5}x^5\right|_{-1}^{\zeta} \\
&= \frac{5}{8}\left(\zeta - \frac{1}{5}\zeta^5 + \frac{4}{5}\right)
\end{aligned}$$

Part (c)

Because the pdf is symmetric about zero, $P(|X| < 0.5)$ is equivalent to $P(-0.5 < X < 0.5)$, which can be found as

$$P(-0.5 < X < 0.5) = P_X(0.5) - P_X(-0.5) = 0.7895$$

Problem 4

Blood alcohol tests on drivers given the conditional probabilities given in table 3:

Table 3: Conditional Probabilities

$P(T\|A)$	$A =$ drunk	$A =$ sober
$T =$ positive	0.99	0.001
$T =$ negative	0.01	0.999

Part (a)

We can find $P(A = \text{drunk}|T = \text{positive})$ through a straightforward application of Bayes' rule:

$$P(A = \text{drunk}|T = \text{positive}) = \frac{P(T = \text{positive}|A = \text{drunk}) * P(A = \text{drunk})}{P(T = \text{positive})}$$

Since we aren't given a number for $P(T = \text{positive})$ we can find it as $P(T = \text{positive}) = P(T = \text{positive}|A = \text{drunk})P(A = \text{drunk}) + P(T = \text{positive}|A = \text{sober})P(A = \text{sober}) = 0.99 * 0.2 + 0.001 * 0.8 = 0.1988$. So the conditional probability is:

$$P(A = \text{drunk}|T = \text{positive}) = \frac{0.99 * 0.2}{0.1988} = 0.996$$

Part (b)

If $P(A = \text{drunk}) = 0.001$ the probability $P(T = \text{positive}) = 0.99 * 0.001 + .001 * 0.999 = 0.002$. So the conditional probability becomes

$$P(A = \text{drunk}|T = \text{positive}) = \frac{0.99 * 0.001}{0.002} = 0.495$$

Problem AQ1

Probability of abiogenesis occurring on an Earth-like planet is given by

$$P(\lambda, n, t) = \text{Poisson}[\lambda, n, t] = e^{-\lambda(t-t_{\text{min}})} \frac{\{\lambda(t - t_{\text{min}}\}^n}{n!}$$

Part (a)

The probability of life arising at least once ($n \geq 1$) could be calculated from the infinite sum

$$P(\lambda, n \geq 1, t) = \sum_{n=1}^{\infty} P[\lambda, n, t]$$

This sum would be difficult to evaluate, however, and we can make things easier for ourselves by noting that the probability life arises at least once is equivalent to the probability that it doesn't arise zero times. Because $\sum_{n=0}^{\infty} P(\lambda, n, t) = 1$, we can calculate the probability life arises at least once as:

$$P(\lambda, n \geq 1, t) = 1 - P(\lambda, 0, t) = 1 - e^{-\lambda(t-t_{\text{min}})} \frac{\{\lambda(t - t_{\text{min}}\}^0}{0!} = 1 - e^{-\lambda(t-t_{\text{min}})}$$

Part (b)

The probability life arises at least once within t_{emerge} is

$$P(E = 1) = 1 - e^{-\lambda(t_{\text{emerge}}-t_{\text{min}})}$$

Similarly, the probability life arises at least once within t_{req} is

$$P(R = 1) = 1 - e^{-\lambda(t_{\text{req}}-t_{\text{min}})}$$

Lastly, because $t_{\text{min}} < t_{\text{emerge}} < t_{\text{req}}$, we know that if life has arisen within t_{emerge} then it definitely has also arisen within t_{required}, so $P(R=1|E=1) = 1$. Using these probabilities we can apply Bayes rule to get

$$P(E=1|R=1) = \frac{P(R=1|E=1)P(E=1)}{P(R=1)} = \frac{1-e^{-\lambda(t_{\text{emerge}}-t_{\text{min}})}}{1-e^{-\lambda(t_{\text{req}}-t_{\text{min}})}}$$

Part (c)

The conditional probability $P(E=1|R=1,\lambda,t_{\text{min}})$ can be expressed as the following likelihood function given $y = \log_{10}\lambda$:

$$P(E=1|R=1,y,t_{\text{min}}) = \frac{1-e^{-(10^y)(t_{\text{emerge}}-t_{\text{min}})}}{1-e^{-(10^y)(t_{\text{req}}-t_{\text{min}})}}$$

Given this likelihood function, we can find the posterior for y via Bayes' rule

$$P(y|E=1,R=1,t_{\text{min}}) = \frac{P(y)P(E=1|R=1,y,t_{\text{min}})}{\int_{-\infty}^{\infty} P(y)P(E=1|R=1,y,t_{\text{min}})dy}$$

For the given cases, $P(y)$ was replaced with the appropriate density function, $P(E=1|R=1,y,t_{\text{min}})$ was replaced with the likelihood function given above and $P(y)P(E=1|R=1,y,t_{\text{min}})$ was evaluated at 10^4 y values in the range $[-3,3]$. These samples were normalized with the integral $\int_{-\infty}^{\infty} P(y)P(E=1|R=1,y,t_{\text{min}})dy$, numerically evaluated using a right Reimann sum over the same range of y values. The posteriors $P(y|R=1,E=1,t_{\text{min}})$ for cases 1, 2, and 3 are plotted in figures 3, 4, and 5, respectively. On each figure the posterior is plotted in blue, while the likelihood function is plotted in red and the prior is plotted in yellow.

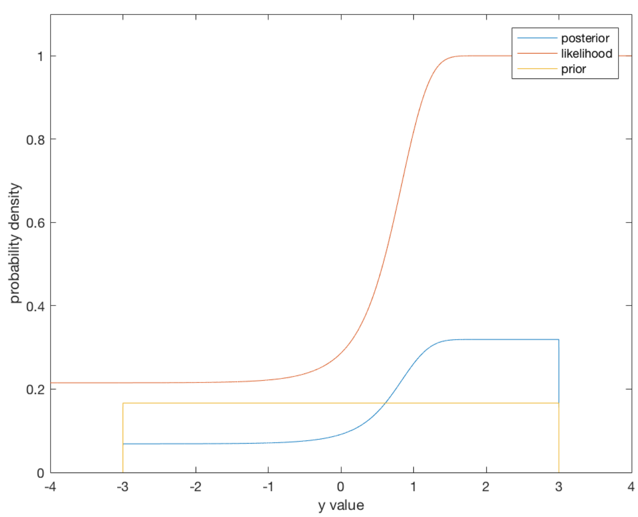

Figure 3: pdfs for case 1

Part (d)

Case 1 shows the greatest change from the prior. As can be seen in figures 4 and 5, the prior essentially overrides the likelihood function in cases 2 and 3, and there is little difference between

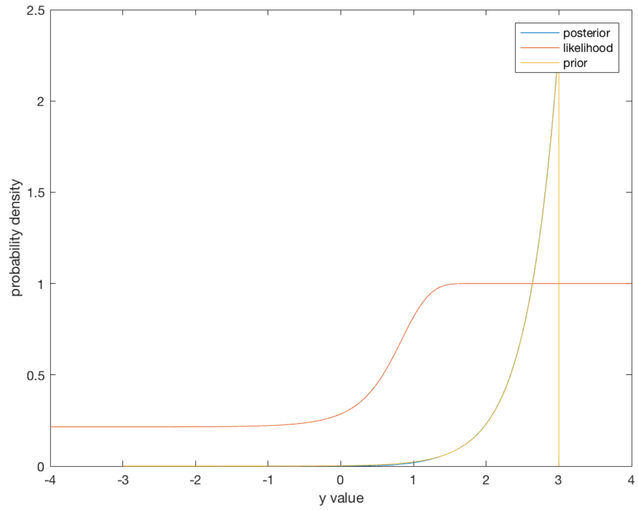

Figure 4: pdfs for case 2

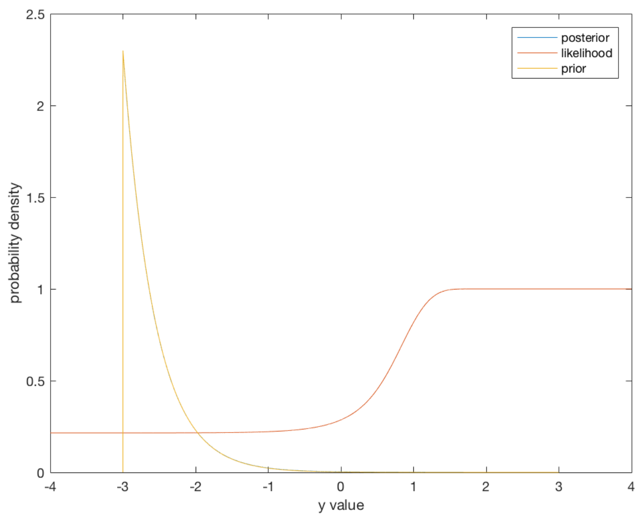

Figure 5: pdfs for case 3

the prior and posterior in these cases. If the intention of conditioning the prior using the likelihood function is to update the prior given information from the likelihood function, then the prior from case 1 is the best. This is because the posterior in case 1 is clearly a fusion of the prior and the likelihood function. In cases 2, and 3, it is difficult to tell that the likelihood function has had any affect at all on the posterior.