HW1 - Kyle Mondina - 0000151556

~~Q~~ Q1

a)

$$\frac{1}{\cancel{\sqrt{2\pi}}} e^{-\frac{x^2}{2}} = \frac{1}{\cancel{\sqrt{2\pi}}} \frac{1}{\sqrt{\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$$-\frac{x^2}{2} = \ln\left(\frac{1}{\sqrt{\sigma^2}}\right) - \frac{(x-\mu)^2}{2\sigma^2}$$

$$-\frac{x^2}{\cancel{2}} = 2\ln\left(\frac{1}{\sqrt{\sigma^2}}\right) - \frac{x^2}{\cancel{2}\sigma^2} + \frac{2\mu x}{\cancel{2}\sigma^2} - \frac{\mu^2}{\cancel{2}\sigma^2}$$

$$-x^2 = 2\ln\left(\frac{1}{\sqrt{\sigma^2}}\right) - \frac{x^2}{\sigma^2} + \frac{2\mu x}{\sigma^2} - \frac{\mu^2}{\sigma^2}$$

$$\cancel{\left(\sigma^2 x^2 - \frac{x^2}{\sigma^2}\right)} + \frac{2\mu x}{\cancel{\sigma^2}} + \sigma^2 2\ln\left(\frac{1}{\sqrt{\sigma^2}}\right) - \frac{\mu^2}{\cancel{\sigma^2}}$$

$$\frac{\cancel{\sigma^2} x^2(\sigma^2-1)}{\cancel{\sigma^2}}$$

$$x^2(\sigma^2-1) + 2\mu x + 2\sigma^2\ln\left(\frac{1}{\sqrt{\sigma^2}}\right) - \mu^2 = 0$$

$$x = \frac{-2\mu \pm \sqrt{4\mu^2 - 4(\sigma^2-1)\left(2\sigma^2\ln\left(\frac{1}{\sqrt{\sigma^2}}\right) - \mu^2\right)}}{2(\sigma^2-1)}$$

$$= \frac{-\mu \pm \sqrt{\mu^2 - (\sigma^2-1)\left(2\sigma^2\ln\left(\frac{1}{\sqrt{\sigma^2}}\right) - \mu^2\right)}}{\sigma^2-1}$$

Question 1

b) $$N = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$$\text{for } p(x \mid w_1),\ \theta = \begin{array}{c} \mu \quad \sigma^2 \\ [\,0\ ,\ 1\,] \end{array}$$

$$p(x \mid w_2),\ \theta = [\,1\ ,\ 2\,]$$

$$\cancel{\frac{1}{\sqrt{2\pi}}} e^{-\frac{x^2}{2}} = \cancel{\frac{1}{\sqrt{2\pi}}} \frac{1}{\sqrt{2}} e^{\frac{-(x-1)^2}{4}}$$

$$-\frac{x^2}{2} = \ln\left(\frac{1}{\sqrt{2}}\right) + \left(\frac{-(x-1)^2}{4}\right)$$

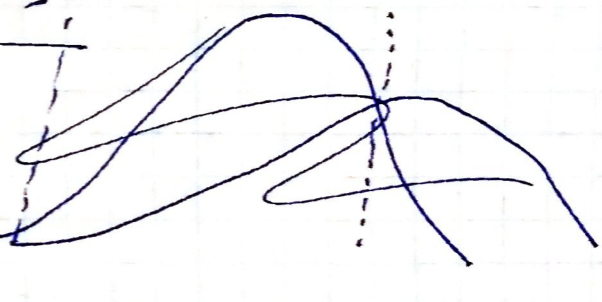
$$-\frac{x^2}{2} = \ln\left(\frac{1}{\sqrt{2}}\right) + \left(\frac{-x^2+2x-1}{4}\right)$$

$$-2x^2 = 4\ln\left(\frac{1}{\sqrt{2}}\right) - x^2 + 2x - 1$$

$$\underset{a}{x^2} + \underset{b}{2x} + \underset{c}{\left[4\ln\left(\frac{1}{\sqrt{2}}\right) - 1\right]}$$

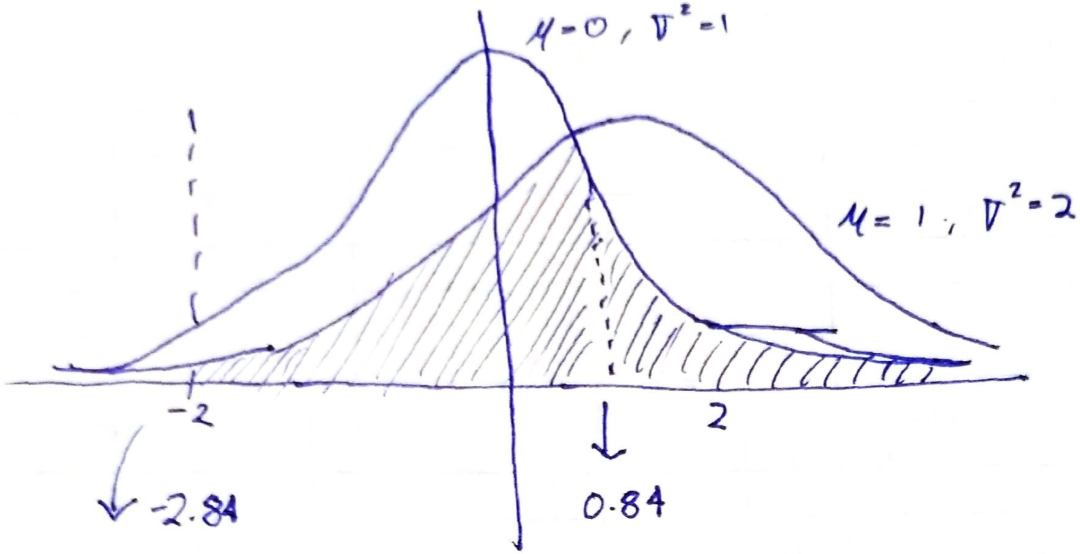
$$\frac{-2 \pm \sqrt{4 - 4\left[4\ln\left(\frac{1}{\sqrt{2}}\right) - 1\right]}}{2a}$$

$$x = 0.84\ ,\ -2.84$$



error rate is

c) estimating error rate,



$$\cancel{\int_{-2.84}^{0.84} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx + \int_{0.84}^{\infty} \frac{1}{\sqrt{4\pi}} e^{\frac{-(x-1)^2}{4}}\,dx}$$

$$= \int_{-2.84}^{0.84} \frac{1}{\sqrt{4\pi}} e^{\frac{-(x-1)^2}{4}}\,dx + \int_{0.84}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$= 0.65$$

COEN Homeworork #1

Q2) $P(X/\theta) = \begin{cases} \theta e^{-\theta x} & x \geq 0 \\ 0 \end{cases}$

a) In jupyter notebook

b) show that $\hat{\theta} = \dfrac{1}{\frac{1}{n}\sum_{1}^{n} X_K}$

$$\ell(\theta) = \ell(P(X/\theta))$$

$$= \ln(\theta e^{-\theta x})$$

$$= \sum_{k}^{n} \ln\theta - \theta X_k$$

taking the derivative & the fact that the sum of functions is the sum of its derivatives,

$$\ell'(\theta) = \sum_{k}^{n} \frac{1}{\theta} - \cancel{\theta} X_k$$

setting to zero

$$\sum_{k}^{n} \frac{1}{\hat{\theta}} - \cancel{\theta} X_k = 0$$

$$\frac{n}{\hat{\theta}} = \sum_{k}^{n} X_K$$

$$\hat{\theta} = \frac{n}{\sum_{k}^{n} X_K}$$

$$= \frac{1}{\frac{1}{n}\sum_{k}^{n} X_K}$$

Q3)

$$P(X|\theta) = \prod_{i=1}^{d} \theta_i^{x_i} (1-\theta_i)^{1-x_i}$$

show $\hat{\theta} = \frac{1}{n} \sum_{k=1}^{n} x_k$

$$f(\theta) = \prod_{i=1}^{n} \prod_{j=1}^{d} \theta_j^{x_{ij}} (1-\theta_j)^{1-x_{ij}}$$

given n samples

$$= \prod_{j=1}^{d} \theta_j^{\sum_{i=1}^{n} x_{ij}} (1-\theta_j)^{n-\sum_{i=1}^{n} x_{ij}}$$

taking the log,

$$= \ln\left(\prod_{j=1}^{d} \theta_j^{\sum_{i=1}^{n} x_{ij}}\right) + \ln\left(\prod_{j=1}^{d} (1-\theta_j)^{n-\sum_{i=1}^{n} x_{ij}}\right)$$

$$= \sum_{i=1}^{d} \ln\left(\theta_j^{\sum_{i=1}^{n} x_{ij}}\right) + \sum_{j=1}^{d} (1-\theta_j)^{n-\sum_{i=1}^{n} x_{ij}})$$

using $\log(x^y) = y\log(x)$ property

$$\sum_{j=1}^{d}\left(\sum_{i=1}^{n} x_{ij} \ln \theta_j\right) + \sum_{j=1}^{d} (n - \sum_{i=1}^{n} x_{ij}) \ln(1-\theta)$$

lets derive a single θ_j

$$\frac{\sum_{i=1}^{n} x_i}{\theta} + \frac{n - \sum_{i=1}^{n} x \, (-1)}{1-\theta} = 0$$

$$\left(\sum_{i=1}^{n} x_{i}\right)(1-\theta) = n - \sum_{i=1}^{n} x_{i} \theta$$

$$\sum_{i}^{n}$$

$$\sum_{i=1}^{n} X_i - \cancel{\theta \sum_{i=1}^{n} X} = n\theta - \cancel{\theta \sum_{i=1}^{n}} X_i$$

$$\hat{\theta} = \frac{\sum_{i=1}^{n} X_i}{n} \leftarrow \text{for all } \theta_1 \cdots \theta_j$$

thus $\hat{\theta} = \frac{\sum_{i=1}^{n} X_i}{n}$

Question 4

a) from question 3,

we know that

$$P(D|\theta) = \prod_{j=1}^{d} \theta_j^{\overbrace{\sum_1^n x_{ij}}^{s_i}} (1-\theta_j)^{n - \overbrace{\sum_1^n x_{ij}}^{s_i}}$$

$$= \prod_{j=1}^{d} \theta_j^{s_i} (1-\theta_j)^{n-s_i}$$

b) $P(\theta) \leftarrow$ can not worry given uniform distri

$$P(\theta|D) = \frac{P(D|\theta)\,P(\theta)}{P(D)}$$

$$P(D) = \int \theta_i^{s_i} (1-\theta_i)^{n-s_i} = \frac{(s_i)!\,(n-s_i)!}{(\cancel{s_i} + n - \cancel{s_i} + 1)}$$

$$\cancel{P(\theta|D) = \frac{\theta_i^{s_i}(1-\theta_i)^{n-s_i}}{}}$$

$$P(\theta|D) = \frac{\theta^{s}(1-\theta)^{n-s}}{\frac{(s)!\,(n-s)!}{(n+1)!}}$$

$$= \prod_{i=1}^{d} \frac{(n+1)!}{(s_i!)(n-s_i!)} \theta^{s_i} (1-\theta)^{n-s_i}$$

4C:

since $d=1$ and $n=1$,

$$P(\theta_1 | D) = \prod_{i=1}^{1} \frac{(1+1)!}{S_i!\,(1-S_i)!} \theta_1^{S_i} (1-\theta_1)^{1-S_i}$$

$$= \frac{2!}{S_1!\,(1-S_1)!} \theta_1^{S_i} (1-\theta_1)^{n-S_i}$$

$$= \frac{2}{S_1!\,(1-S_1)!} \theta_1^{S_i} (1-\theta_1)^{1-S_i}$$

$S \rightarrow 0 \Rightarrow 2(1-\theta_1)$

$S \rightarrow 1 = 2\theta_1$

graph shown in Jupyter notebook

Ae ~~When~~ The effective Bayesian estimate would be the same ~~if~~ as the one dimensional case. This is because the multivariate case is an extension of the one dimensional case. When you plug in an estimate $\hat{\theta}$ that was learned from $P(X|D)$ to $P(X|\theta)$, the effective would be the same as plugging in θ learned from $P(X|\theta)$

4d)

$$P(X \mid D) = \prod_{1}^{d} \left(\frac{S_i+1}{n+2}\right)^{X_i} \left(1 - \frac{S_i+1}{n+2}\right)^{1-X_i}$$

$$P(X \mid \theta) = \prod_{i=1}^{d} \theta_i^{X_i} (+\theta_i)^{1-X_i}$$

$$P(D \mid \theta) = \prod_{1}^{d} \frac{(n+1)!}{S!(n-S_i)!} \theta_i^{S_i} (1-\theta_i)$$

$$P(X_i \mid D) = \begin{cases} 1 - \frac{S_i+1}{n-2} & X_i = 0 \\ \frac{S_i+1}{n+2} & X_1 = 1 \end{cases}$$

---

obtaining the integral

$$\int \prod_{i=1}^{d} \theta_i^{X_i} (1-\theta_i)^{1-X_i} \prod_{i=1}^{d} \frac{(n+1)!}{S_i!(n-S_i)!} \theta_i^{S_i} (1-\theta_i)^{n-S_i}$$

$$\int \prod_{i=1}^{d} \theta_i^{X_i+S_i} (1-\theta_i)^{n+1-X_i-S_i} \frac{(n+1)!}{S_i!(n-S_i)!} d\theta$$

$$\Downarrow$$

$$\cancel{\prod^{d}} \prod_{1}^{d} \int \theta_i^{X_i+S_i} (1-\theta_i)^{n+1-X_i-S_i} \frac{(n+1)}{S_i!(n-S_i)!}$$

applying integration by parts

$$\frac{\cancel{(n+1)!}}{S_i!(n-S_i)!} \cdot \frac{(X_i+S_i)!(1+n-X_i-S_i)!}{\cancel{(n+1)}!(n+2)}$$

when $X_i = 1 \rightarrow \frac{(1+S_i)!(1+n-1-S_i)!}{S_i!(n-S_i)!(n+2)}$

$$= \frac{1+S_i}{n+2}$$

When $x_i = 0$

$$\frac{1}{s!\,(n-s_i)} \quad \frac{(s_i)!\,(1+n-0-s_i)!}{(n+2)!}$$

$$= \frac{(1+n-s_i)!\,\cancel{(n-s_i)!}}{\cancel{(n-s_i)!}\,(n+2)}$$

$$= \frac{1+n-s_i}{n+2} = \frac{1+n-s_i-1+1}{n+2}$$

$$= \frac{n+2-s_i-1}{n+2} = \frac{1-s_i+1}{n+2}$$

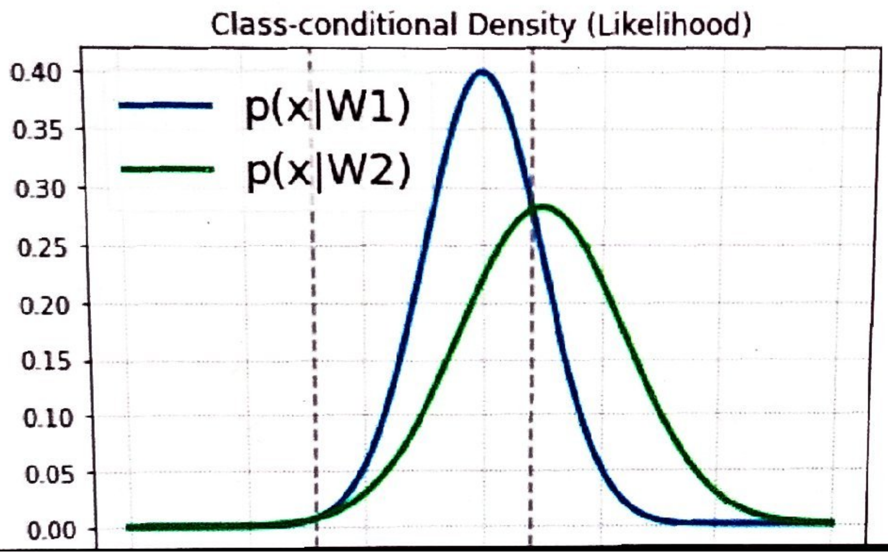
$$X_1 = \frac{1+s_i}{n+2}$$

$$X_0 = 1 - \frac{s_i+1}{n+2}$$

put into pernouli form and you get

$$P(X_i \mid D) = \begin{cases} \dfrac{1-s_i+1}{n+2} & x_i = 0 \\ \dfrac{s_i+1}{n+2} & x_i = 1 \end{cases}$$

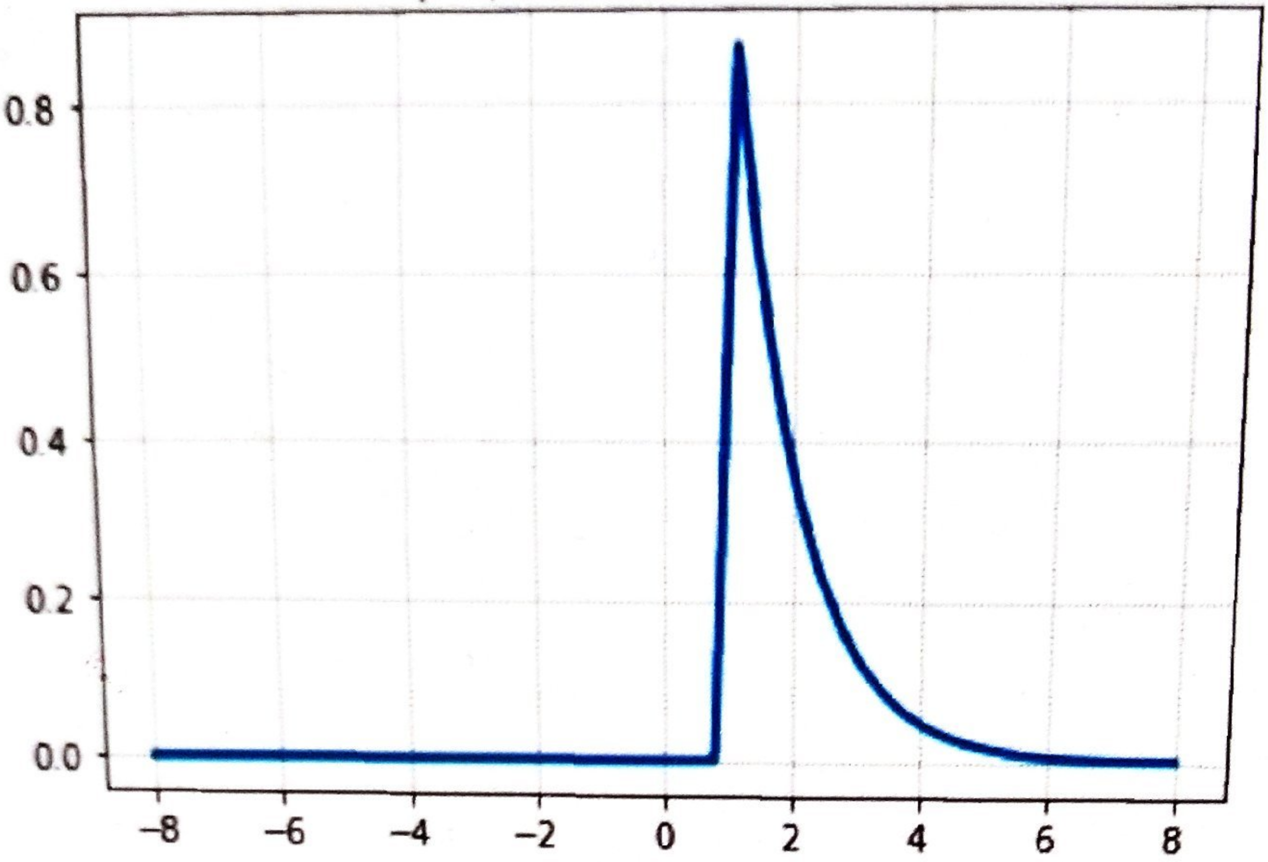
```
%matplotlib inline

import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm
from scipy.stats import expon
from pdb import set_trace as st


x = np.linspace(-6, 6, 5000)
# x1 and x2 denote p(x|w_1) and p(x|w_2)
x1 = norm.pdf(x, 0,1)
x2 = norm.pdf(x, 1,1.41)


plt.grid(which='both',axis='both', color='grey',linestyle=':')
plt.plot(x, x1, label='p(x|W1)', color='tab:blue', linewidth=3)
plt.plot(x, x2, label='p(x|W2)', color='tab:green', linewidth=3)
plt.title('Class-conditional Density (Likelihood)')
plt.legend(fancybox=True, framealpha=0.5, fontsize=18)

intersections = [0.84018,-2.84]
plt.axvline(x=0.84,color='gray',linestyle='--')
plt.axvline(x=-2.84018867541,color='gray',linestyle='--')

plt.show()
```



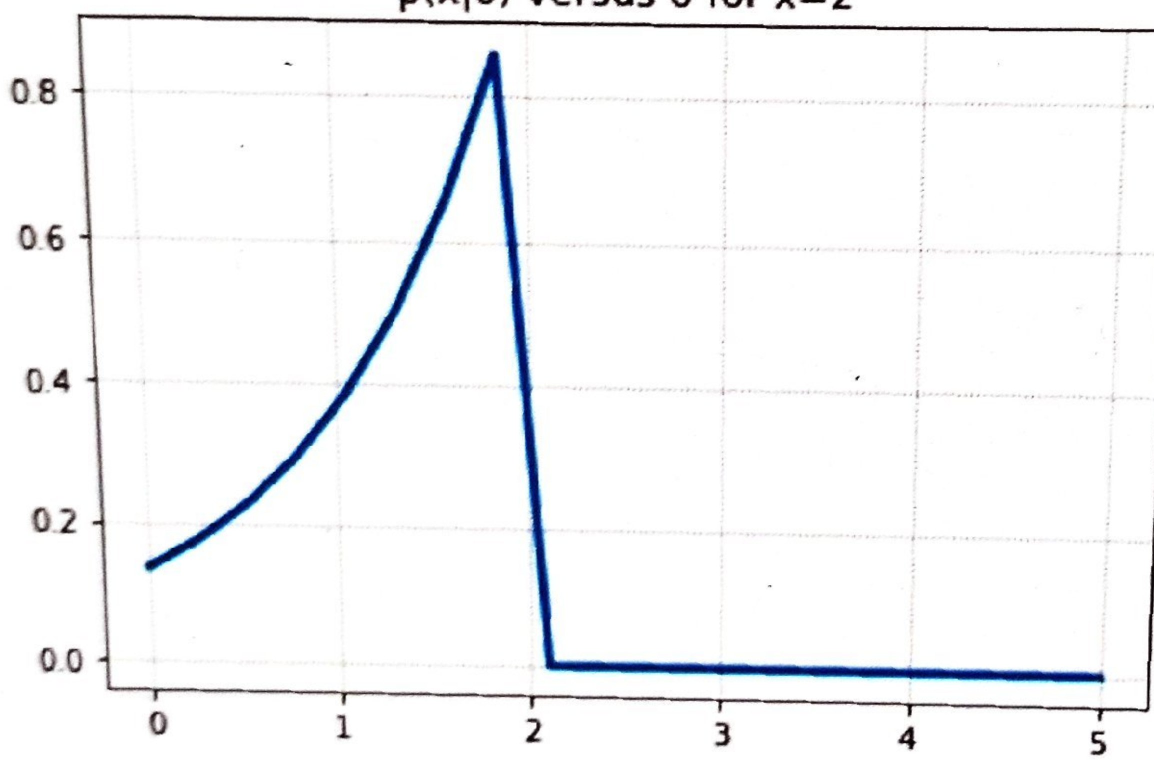
```python
###Question 2 a
%matplotlib inline

import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm
from scipy.stats import expon
from pdb import set_trace as st


### p(x|θ) versus x for θ=1
theta1 = 1
x1 = np.linspace(-8,8,50)
expo1 = expon.pdf(x1,theta1)

plt.title('p(x|θ) versus x for θ=1')
plt.grid(which='both',axis='both', color='grey',linestyle=':')
plt.plot(x1,expo1,label='p(x|θ)',color="tab:blue",linewidth=3)
plt.show()


### p(x|θ) versus θ for x = 2
theta2 = np.linspace(0,5,20)
x2 = 2
expo2 = expon.pdf(x2,theta2)


plt.title('p(x|θ) versus θ for x=2')
plt.grid(which='both',axis='both', color='grey',linestyle=':')
plt.plot(theta2,expo2,label='p(x|θ)',color="tab:blue",linewidth=3)
plt.show()
```

```
plt.show()
```

p(x|0) versus x for 0=1



p(x|0) versus 0 for x=2



```python
import matplotlib.pyplot as plt
from scipy.stats import norm
from scipy.stats import expon
from pdb import set_trace as st


### p(x|θ) versus x for θ=1

x = np.linspace(-2,2,50)

s0 = 2*(1-x)
s1 = 2*x


plt.grid(which='both',axis='both', color='grey',linestyle=':')
plt.plot(x,s0,label='s0',color="tab:blue",linewidth=3)
plt.plot(x,s1,label='s1',color="tab:green",linewidth=3)
plt.show()


```