Jin Young Jeon

Lab 3 Gaussian Distribution

1) f(x) = $\frac{1}{\sqrt{2\pi}\sigma_f} e^{-\frac{(x-\mu_f)^2}{2\sigma_f^2}}$ and g(x) = $\frac{1}{\sqrt{2\pi}\sigma_g} e^{-\frac{(x-\mu_g)^2}{2\sigma_g^2}}$

f(x)g(x) = $\frac{1}{2\pi\sigma_f\sigma_g} e^{-(\frac{(x-\mu_f)^2}{2\sigma_f^2}+\frac{(x-\mu_g)^2}{2\sigma_g^2})}$

let $\beta = -(\frac{(x-\mu_f)^2}{2\sigma_f^2} + \frac{(x-\mu_g)^2}{2\sigma_g^2})$

$\beta = \frac{(\sigma_f^2 + \sigma_g^2)x^2 - 2(\mu_f\sigma_g^2 + \mu_g\sigma_f^2)x + \mu_f^2\sigma_g^2 + \mu_g^2\sigma_f^2}{2\sigma_f^2\sigma_g^2}$

$\beta \div (\sigma_f^2 + \sigma_g^2)$

$\beta = \frac{x^2 - 2\frac{\mu_f\sigma_g^2 + \mu_g\sigma_f^2}{\sigma_f^2 + \sigma_g^2}x + \frac{\mu_f^2\sigma_g^2 + \mu_g^2\sigma_f^2}{\sigma_f^2 + \sigma_g^2}}{2\frac{\sigma_f^2\sigma_g^2}{\sigma_f^2 + \sigma_g^2}}$

β is again in quadratic form in x so f(x)g(x) from above is a Gaussian function.

2) Prior x ~ N(μ, σ^2) and Likelihood p ~ N(μ, s^2)
Posterior ∝ Prior x Likelihood
P = P(x)P(p|x) ~ N(μ', σ'^2)

$=\left(\frac{\sigma'}{2\pi}\right)^2 exp\left[-\frac{1}{2}\sigma'(\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x}-\mu)^2)\right]\sqrt{\frac{\sigma'}{2\pi}} exp\left(-\frac{1}{2}\sigma'(\mu - \mu_0)^2\right)$

$= exp\left(-\frac{1}{2}\left(\sigma'(\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x}-\mu)^2)\right) + \sigma_0(\mu-\mu_0)^2\right)$

$= exp\left(-\frac{1}{2}(n\sigma'(\bar{x}-\mu)^2 + \sigma_0(\mu-\mu_0)^2)\right)$

$= exp\left(-\frac{1}{2}\left(n\sigma' + \sigma_0\right)\left(\mu - \frac{n\sigma'\bar{x}+\sigma_0\mu_0}{n\sigma'+\sigma_0}\right)^2\right)$

Posterior ~ $N(\frac{n\sigma'\bar{x}+\sigma_0\mu_0}{n\sigma'+\sigma_0}, \frac{1}{n\sigma'+\sigma_0})$

Task2) We can conclude that the greater the standard deviation, the greater the entropy. Because STD is big, the range of assumption becomes greater and thus entropy rises. Example 3 and 4 are both higher than example 1 and 2 because their standard deviations are greater.
It seems example 3 has greater entropy because example 4 has expected mean of 10 and because the upper boundary is limited to 20, the standard deviation of example 4 is limited upwards.

(Appendix)

```
mu1, sigma1 = 0, 1
mu2, sigma2 = 10, 1
mu3, sigma3 = 0, 10
mu4, sigma4 = 10, 10

s1 = np.random.normal(mu1, sigma1, 100000)
s2 = np.random.normal(mu2, sigma2, 100000)
s3 = np.random.normal(mu3, sigma3, 100000)
s4 = np.random.normal(mu4, sigma4, 100000)

hist1 = np.histogram(s1, bins=100, range=(-20,20), density=True)
# print(hist1) #lists array of data that adds up to 1
data1 = np.trim_zeros(hist1[0]) #trims out the zero values
ent1 = -(data1*np.log(np.abs(data1))).sum()
print("entropy for example 1 is:", ent1) #3.56108941738

hist2 = np.histogram(s2, bins=100, range=(-20,20), density=True)
# print(hist2) #lists array of data that adds up to 1
data2 = np.trim_zeros(hist2[0]) #trims out the zero values
ent2 = -(data2*np.log(np.abs(data2))).sum()
print("entropy for example 2 is:", ent2) #3.55136926537

hist3 = np.histogram(s3, bins=100, range=(-20,20), density=True)
# print(hist3) #lists array of data that adds up to 1
data3 = np.trim_zeros(hist3[0]) #trims out the zero values
ent3 = -(data3*np.log(np.abs(data3))).sum()
print("entropy for example 3 is:", ent3) #8.90450371361

hist4 = np.histogram(s4, bins=100, range=(-20,20), density=True)
# print(hist4) #lists array of data that adds up to 1
data4 = np.trim_zeros(hist4[0]) #trims out the zero values
ent4 = -(data4*np.log(np.abs(data4))).sum()
print("entropy for example 4 is:", ent4) #8.49200221924

table1 = [["example1",ent1],["example2",ent2],["example3",ent3], ["example4",ent4]]
print (tabulate(table1))
#--------  -------
#example1  3.56109
#example2  3.55137
#example3  8.9045
#example4  8.492
#--------  -------
```