Kendra Chalkley
CS 559 Machine Learning
4/26/18

HW1 Density Estimation

1 Problem 1: Bishop 2.7

1	$\mu_{post} = \frac{a+m}{a+b+m+l}$	posterior mean given by 2.20
	$= \frac{a}{a+b+m+l} + \frac{m}{a+b+m+l}$ $= \frac{a}{(a+b+m+l)}\frac{(a+b)}{(a+b)} + \frac{m}{(a+b+m+l)}\frac{(m+l)}{(m+l)}$ $= \frac{(a+b)}{(a+b+m+l)}\frac{a}{(a+b)} + \frac{(m+l)}{(a+b+m+l)}\frac{m}{(m+l)}$	
2	let $\lambda = \frac{(a+b)}{(a+b+m+l)}$	$a+b < a+b+m+l \implies 0 \leq \lambda \leq 1$
	it follows: $1 - \lambda = \frac{(a+b+m+l)}{(a+b+m+l)} - \frac{(a+b)}{(a+b+m+l)}$ $= \frac{m+l}{a+b+m+l}$	
6	$\mu_{post} = \lambda\frac{a}{(a+b)} + \frac{(m+l)}{(a+b+m+l)}\frac{m}{(m+l)}$	substitute 5 into 4
9	$\mu_{post} = \lambda\frac{a}{(a+b)} + (1-\lambda)\frac{m}{(m+l)}$	substitute 8 into 6
10	$\mu_{prior} = \frac{a}{a+b}$	prior mean given by 2.15
11	$\mu_{post} = \lambda * \mu_{prior} + (1-\lambda)\frac{m}{(m+l)}$	substitute 10 into 9
12	$\mu_{ML} = \frac{m}{m+l}$	maximum likelihood estimate given by 2.8
13	$\mu_{post} = \lambda * \mu_{prior} + (1-\lambda)\mu_{ML}$	substitute 12 into 10

2 Problem 2: Bishop 2.12

The uniform distribution for a continuous variable x is defined by:

$$U(x|a,b) = \frac{1}{b-a}$$

for $a \leq x \leq b$.

This distribution is normalized, ie integrates to 1 as shown below

$$\begin{aligned}
\int_a^b p(x) &= \int_a^b \tfrac{1}{b-a} \\
&= \tfrac{1}{b-a}(b-a) \\
&= 1
\end{aligned}$$

The mean of the uniform distribution can be expressed as:

$$\mathbb{E}[U] = \frac{b+a}{2}$$

which is derived below

$$\begin{aligned}
\mathbb{E}[f] &= \int p(x)f(x)\mathrm{d}x \\
p(x) &= U(x) = \tfrac{1}{b-a} \\
f(x) &= x \\
\mathbb{E}[f] &= \int_a^b \tfrac{1}{b-a} x\mathrm{d}x \\
&= \tfrac{1}{b-a} \int_a^b x\mathrm{d}x \\
&= \tfrac{1}{b-a} \tfrac{b^2}{2} - \tfrac{a^2}{2} \\
&= \tfrac{1}{b-a} \tfrac{b^2-a^2}{2} \\
&= \tfrac{(b-a)(b+a)}{2(b-a)} \\
&= \tfrac{(b+a)}{2}
\end{aligned}$$

The variance of the uniform distribution can be expressed as:

$$var[U] = \frac{(b-a)^2}{12}$$

which is derived below

$$
\begin{aligned}
var[f] &= \mathbb{E}[(f(x) - \mathbb{E}[f(x)])^2] \\
&= \mathbb{E}[(x - \tfrac{(b+a)}{2})^2] \\
&= \mathbb{E}[x^2 - 2x\tfrac{(b+a)}{2} + \tfrac{(b+a)^2}{4}] \\
&= \int\limits_a^b \tfrac{1}{b-a}(x^2 - 2x\tfrac{(b+a)}{2} + \tfrac{(b+a)^2}{4}) \\
&= \tfrac{1}{b-a}(\int\limits_a^b x^2 - ((b+a)\int\limits_a^b x) + \int\limits_a^b \tfrac{(b+a)^2}{4}) \\
&= \tfrac{1}{b-a}(\tfrac{b^3-a^3}{3} - \tfrac{(b+a)(b^2-a^2)}{2} + \tfrac{(b+a)^2(b-a)}{4}) \\
&= \tfrac{b^3-a^3}{3(b-a)} - \tfrac{(b+a)(b^2-a^2)}{2(b-a)} + \tfrac{(b+a)^2}{4} \\
&= \tfrac{(b^2+ab+a^2)(b-a)}{3(b-a)} - \tfrac{(b+a)(b+a)(b-a)}{2(b-a)} + \tfrac{(b+a)^2}{4} \\
&= \tfrac{(b^2+ab+a^2)}{3} - \tfrac{(b+a)^2}{2} + \tfrac{(b+a)^2}{4} \\
&= \tfrac{(4b^2+4ab+4a^2)}{12} - \tfrac{6(b+a)^2}{12} + \tfrac{3(b+a)^2}{12} \\
&= \tfrac{(4b^2+4ab+4a^2)}{12} - \tfrac{3(b+a)^2}{12} \\
&= \tfrac{(4b^2+4ab+4a^2)}{12} - \tfrac{3b^2+6ba+3a^2}{12} \\
&= \tfrac{(b^2-2ab+a^2)}{12} \\
&= \tfrac{(b-a)^2}{12}
\end{aligned}
$$

3 Problem 3: Poisson

3.1 3a

Using the definition of the Poisson distribution show that the sum of probabilities of all events is 1. (Hint: use the definition of e in terms of a sum).

By definition, the Poison distribution is:

$$p(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$$

The sum of probabilities for all x is then:

$$\sum_{x=0}^{\inf} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \frac{1}{e^\lambda} \sum_{x=0}^{\inf} \frac{\lambda^x}{x!}$$

By definition of power series,

$$\sum_{x=0}^{\inf} \frac{\lambda^x}{x!} = e^\lambda$$

And as such

$$\sum_{x=0}^{\inf} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \frac{1}{e^\lambda} e^\lambda = 1$$

3.2 3b

Derive the mean of the Poisson distribution.

The expected value of a (discrete) function given a distribtuion is given as:

$$\mathbb{E}[f] = \sum_{x=0}^{\inf} p(x) f(x)$$

For $f(x) = x$ and the poisson distribution:

$$\mathbb{E}[f] = \frac{e^{-\lambda}\lambda^x}{x!}$$

The expected value is:

$$\mathbb{E}[f] = \sum_{x=0}^{\inf} \frac{e^{-\lambda}\lambda^x}{x!} x$$

Which can be manipulated to:

$$\begin{aligned}\mathbb{E}[f] &= \tfrac{1}{e^{\lambda}} \sum_{x=0}^{\inf} \tfrac{\lambda^x}{x!} x \\ &= \tfrac{1}{e^{\lambda}} \sum_{x=0}^{\inf} \tfrac{\lambda^{x-1}\lambda x}{(x-1)!x} \\ &= \tfrac{\lambda}{e^{\lambda}} \sum_{x=1}^{\inf} \tfrac{\lambda^x}{x!} \\ &= \tfrac{\lambda}{e^{\lambda}} e^{\lambda} \\ &= \lambda\end{aligned}$$

3.3 3c

$$\lambda_{ML} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

3.4 3d

The posterior distrobution poisson given a prior gamma is:

$$p_{post}(x|\lambda, a, b) = \frac{e^{-\lambda}\lambda^x}{x!} * \frac{\lambda^{a-1} e^{\frac{\lambda}{b}}}{b^a \Gamma(a)}$$

$$= \frac{\lambda^{a-1+x} e^{\frac{\lambda}{b}-\lambda}}{b^a \Gamma(a) x!}$$

Which is of the same form as the gamma distribution, except for the x! in the denominator which I have no idea what to do with.

3.5 3e

See figure 1.

3.6 3f

$$\lambda_{ML} = \frac{calls}{time}$$

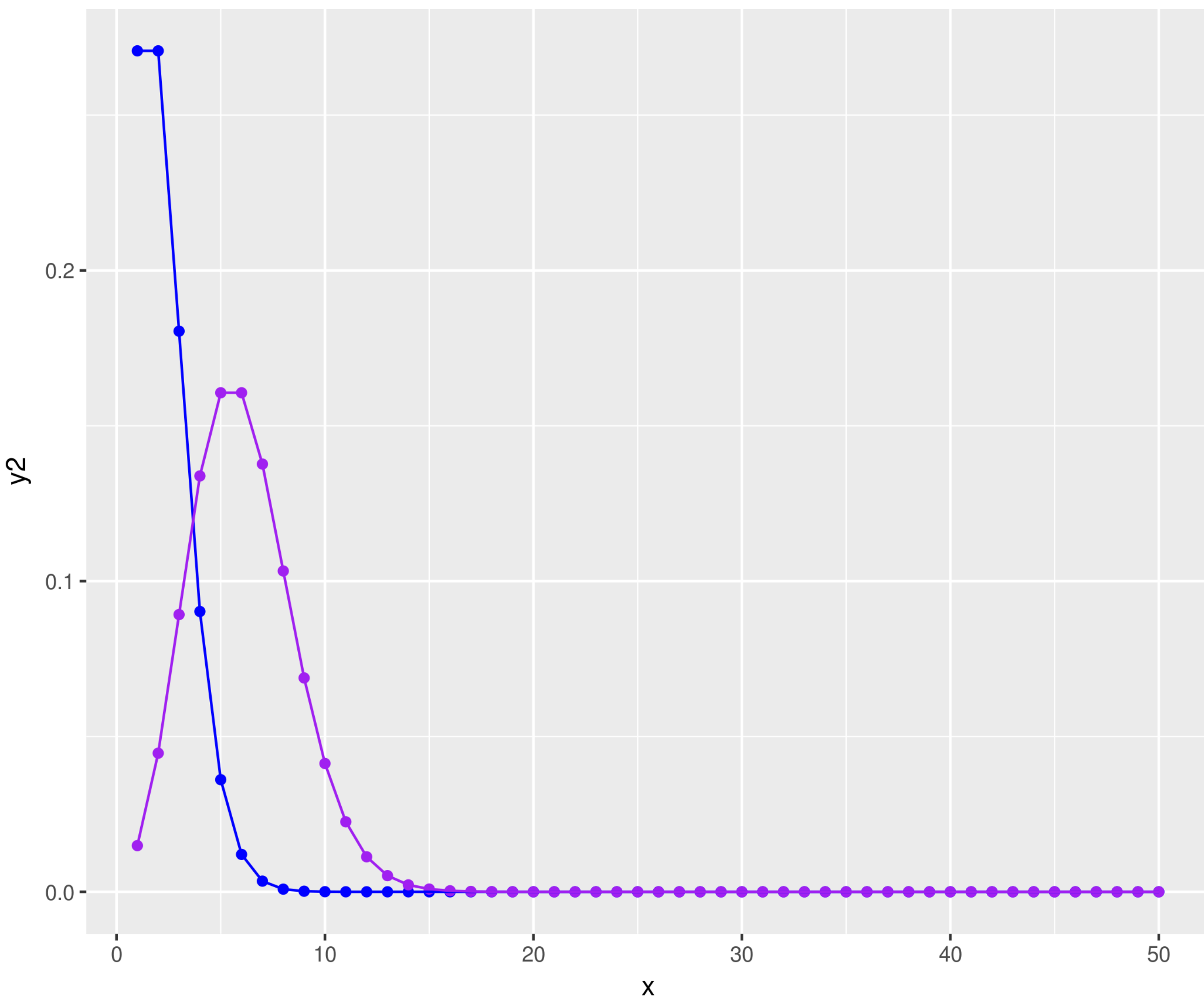

Figure 1: 3e: Combined plot showing difference between $\lambda = 2$ and $\lambda = 6$

3.7 3g

See figure 2.

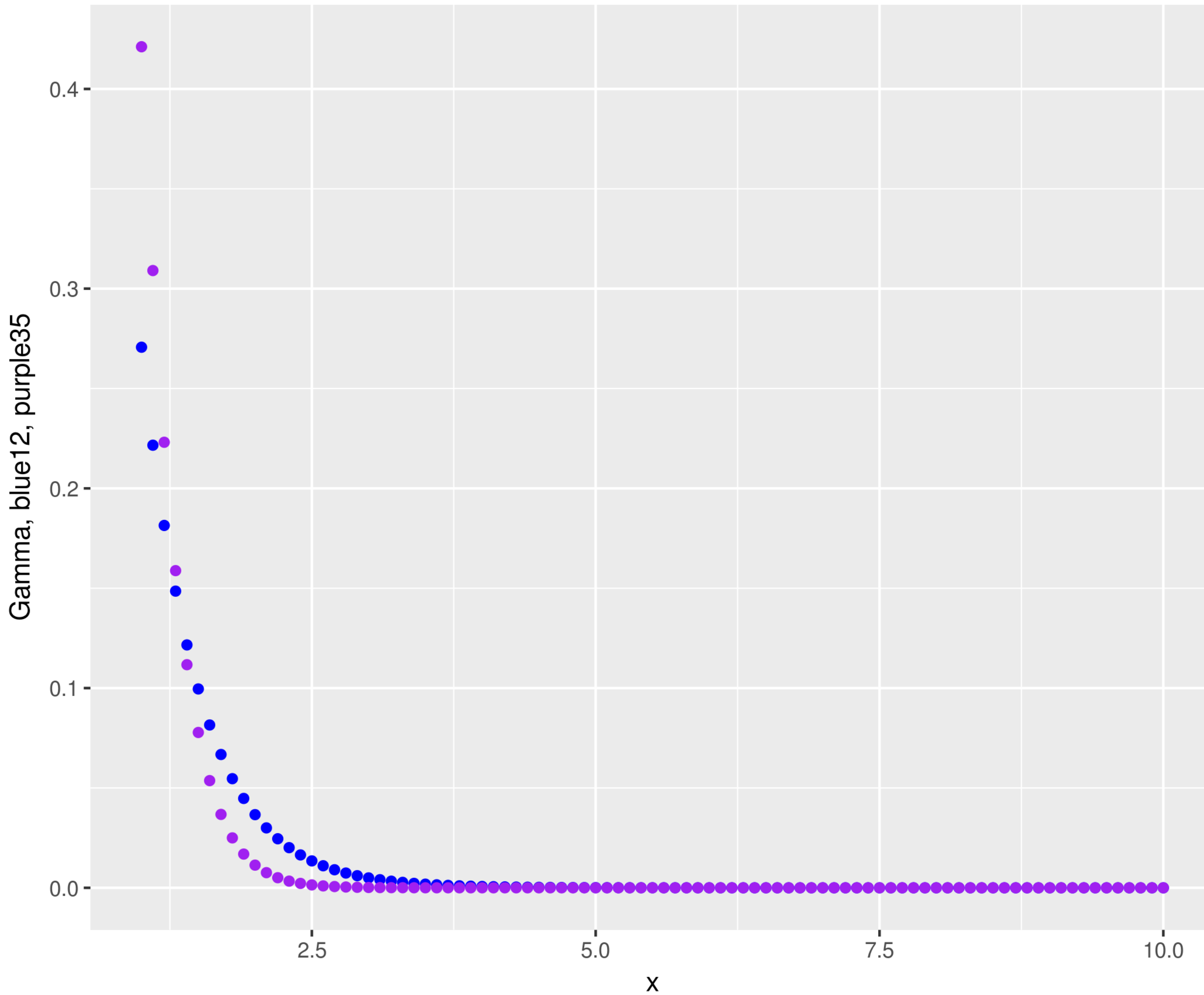

Figure 2: 3g. Plot of gamma distributions. Blue points show when a=1 b=2, and purple points show when a=3 and b=5.

3.8 3h

4 Problem 4: Exponential Family

The Poisson distribution:

$$p(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$$

Can be manipulated as:

$$p(x|\lambda) = \frac{e^{-\lambda} exp(log(\lambda^x))}{x!}$$

$$= \frac{e^{-\lambda} exp(xlog(\lambda))}{x!}$$

Which can be represented by the general form of the exponential family:

$$f(x|\eta) = \frac{1}{Z(\eta)} h(x) exp[\eta^T t(x)]$$

where

$$\eta = log(\lambda)$$

$$Z(\eta) = \frac{1}{e^{-\lambda}}$$

$$t(x) = x$$

$$h(x) = \frac{1}{x!}$$