Date / /
Page

Exercise 10

1]. $P_1(z = \text{setosa}) = 0.34 \quad = r$
$P_2(z = \text{versicolor}) = 0.33 \quad = s$
$P_3(z = \text{virginica}) = 0.33 \quad = t$
$\sigma^2 = 0.54$

(a). $P(z_i) = r^{\mathbb{1}(z=\text{setosa})} \cdot s^{\mathbb{1}(z=\text{versicolor})} \cdot t^{\mathbb{1}(z=\text{virginica})}$

$$\boxed{P(z_i) = \prod_{c \in (\text{setosa}, \text{versicolor}, \text{virginica})} \pi_c^{\mathbb{1}(z_i=c)}}$$

(b). $$\boxed{P(x_i | z_i) = \prod_{c \in (\text{setosa}, \text{versicolor}, \text{virginica})} \mathcal{N}(x_i; \mu_c, \sigma^2)^{\mathbb{1}(z_i=c)}}$$

(c). $P(x_i) = \sum_{z_i} P(z_i) P(x_i | z_i)$

$$= \sum_{z_i} \pi_c^{\mathbb{1}(z_i=c)} \left[ \mathcal{N}(x_i; \mu_c, \sigma^2)^{\mathbb{1}(z_i=c)} \right]$$

$$\boxed{\begin{aligned} P(x_i) = {} & \pi_{\text{setosa}} \mathcal{N}(x_i, \mu_{\text{setosa}}, \sigma^2) + \\ & \pi_{\text{versicolor}} \mathcal{N}(x_i, \mu_{\text{versicolor}}, \sigma^2) + \\ & \pi_{\text{virginica}} \mathcal{N}(x_i, \mu_{\text{virginica}}, \sigma^2) \end{aligned}}$$

(d). $$\boxed{L = P(D) = \prod_{i=1}^{n=150} \left[ \begin{aligned} & \pi_{\text{setosa}} \mathcal{N}(x_i, \mu_{\text{setosa}}, \sigma^2) + \\ & \pi_{\text{versicolor}} \mathcal{N}(x_i, \mu_{\text{versicolor}}, \sigma^2) + \\ & \pi_{\text{virginica}} \mathcal{N}(x_i, \mu_{\text{virginica}}, \sigma^2) \end{aligned} \right]}$$

(e). $l = \sum_{i=1}^{n=150} \log \left[ P(x_i) \right]$

$$\boxed{\sum_{i=1}^{n=150} \log \left[ \pi_{set} \mathcal{N}(x_i, \mu_{set}, \sigma^2) + \pi_{ver} \mathcal{N}(x_i, \mu_{ver}, \sigma^2) + \pi_{vir} \mathcal{N}(x_i, \mu_{vir}, \sigma^2) \right]}$$

Date / /
Page

(f). $P(z_i | x_i) = \frac{P(x_i | z_i) P(z_i)}{P(x_i)}$

① ← $P(z_i = \text{setosa} | x_i) = \frac{\pi_{set} \mathcal{N}(x_i; \mu_{set}, \sigma^2)}{\pi_{set} \mathcal{N}(x_i; \mu_{set}, \sigma^2) + \pi_{ver} \mathcal{N}(x_i; \mu_{ver}, \sigma^2) + \pi_{vir} \mathcal{N}(x_i; \mu_{vir}, \sigma^2)}$

② ← $P(z_i = \text{versicolor} | x_i) = \frac{\pi_{ver} \mathcal{N}(x_i; \mu_{ver}, \sigma^2)}{\pi_{set} \mathcal{N}(x_i; \mu_{set}, \sigma^2) + \pi_{ver} \mathcal{N}(x_i; \mu_{ver}, \sigma^2) + \pi_{vir} \mathcal{N}(x_i; \mu_{vir}, \sigma^2)}$

③ ← $P(z_i = \text{virginica} | x_i) = \frac{\pi_{vir} \mathcal{N}(x_i; \mu_{vir}, \sigma^2)}{\pi_{set} \mathcal{N}(x_i; \mu_{set}, \sigma^2) + \pi_{ver} \mathcal{N}(x_i; \mu_{ver}, \sigma^2) + \pi_{vir} \mathcal{N}(x_i; \mu_{vir}, \sigma^2)}$

④ ← (g). $\mu_{setosa} = \sum_{i=1}^{n=150} \frac{P(\text{setosa} | x_i) P(x_i)}{\sum_{i=1}^{n=150} P(\text{setosa} | x_i)}$

⑤ ← $\mu_{versicolor} = \sum_{i=1}^{n=150} \frac{P(\text{versicolor} | x_i) P(x_i)}{\sum_{i=1}^{n=150} P(\text{versicolor} | x_i)}$

⑥ ← $\mu_{virginica} = \sum_{i=1}^{n=150} \frac{P(\text{virginica} | x_i) P(x_i)}{\sum_{i=1}^{n=150} P(\text{virginica} | x_i)}$

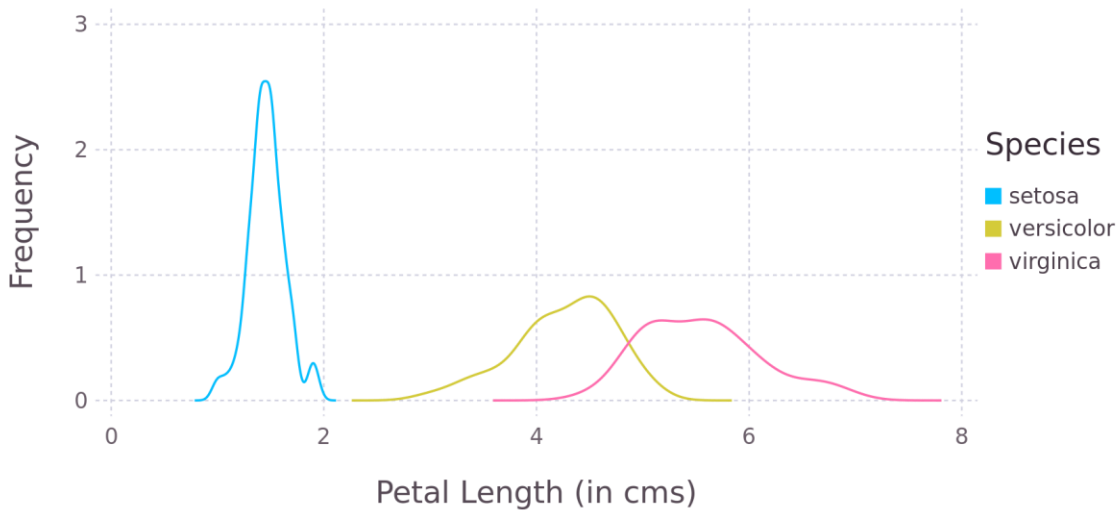
(h). EM algorithm:

1. Pick initial value for $\mu_{setosa}$, $\mu_{versicolor}$, $\mu_{virginica}$.
2. maxIter = 1000
3. for i = 1 : maxIter
4. from equations ①, ②, ③ in (f).:
   Compute $P(z = \text{setosa} | x_i)$, $P(z = \text{versicolor} | x_i)$, $P(z = \text{virginica} | x_i)$
5. from equations ④, ⑤, ⑥ in (g):
   Compute $\mu_{setosa}^{i}$, $\mu_{versicolor}^{i}$, $\mu_{virginica}^{i}$
6. if $|\mu_{set}^{i} - \mu_{set}^{i-1}| < \epsilon$ and $|\mu_{ver}^{i} - \mu_{ver}^{i-1}| < \epsilon$ and $|\mu_{vir}^{i} - \mu_{vir}^{i-1}| < \epsilon$
   terminate; end.
7. end for

In [1]:
```
### QUESTION 1 ###
```

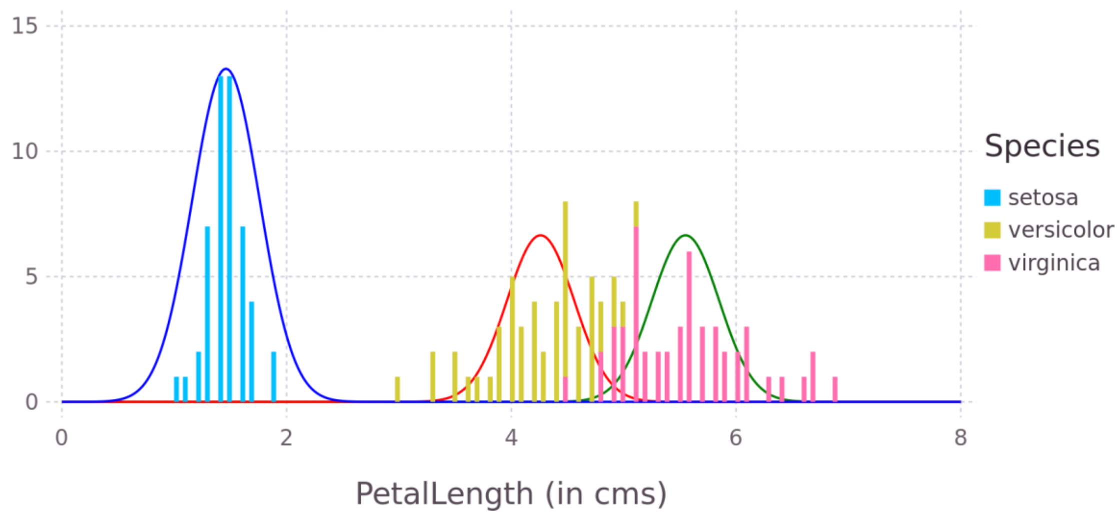
In [2]:
```
using RDatasets, Gadfly, Distributions;
```

In [3]:
```
data = dataset("datasets","iris");
myplot = plot(data,x=:PetalLength,color=:Species,Geom.density,
Guide.xlabel("Petal Length (in cms)"),
Guide.ylabel("Frequency"), major_label_font_size=12pt,
minor_label_font_size=12pt,
key_title_font_size = 12pt,
key_label_font_size = 12pt,
major_label_color=colorant"black",
minor_label_color=colorant"black",
Coord.Cartesian(xmin=0, xmax=8));
draw(PNG(6inch, 3inch,dpi=300), myplot);
```




```
WARNING: Method definition unix2zdt(Real) in module TimeZones at /home/jrun/.julia/v0.6/TimeZones/src/conversions.jl:122 overwritten in module RData at /home/jrun/.julia/v0.6/RData/src/convert.jl:201.
WARNING: key_label_font_size is not a recognized aesthetic. Ignoring.
WARNING: minor_label_color is not a recognized aesthetic. Ignoring.
WARNING: minor_label_font_size is not a recognized aesthetic. Ignoring.
WARNING: major_label_color is not a recognized aesthetic. Ignoring.
WARNING: major_label_font_size is not a recognized aesthetic. Ignoring.
WARNING: key_title_font_size is not a recognized aesthetic. Ignoring.
```

In [4]:
```
myplot = plot(layer(data,x=:PetalLength, color=:Species, Geom.histogram,
              Theme(default_color=colorant"purple")),
              layer(x=0:0.02:8,y=pdf.(Normal(1.46,0.3),0:0.02:8)*10,Geom.line,
              Theme(default_color=colorant"blue")),
              layer(x=0:0.02:8,y=pdf.(Normal(4.26,0.3),0:0.02:8)*5,Geom.line,
              Theme(default_color=colorant"red")),
              layer(x=0:0.02:8,y=pdf.(Normal(5.55,0.3),0:0.02:8)*5,Geom.line,
              Theme(default_color=colorant"green")),
              Guide.xlabel("PetalLength (in cms)"),Guide.ylabel(""), major_lab
el_font_size=18pt,
              minor_label_font_size=14pt,
              key_title_font_size = 18pt,
              key_label_font_size = 14pt,
              major_label_color=colorant"black",
              minor_label_color=colorant"black",Coord.Cartesian(xmin=0, xmax=8
));
draw(PNG(6inch, 3inch, dpi=300), myplot);
```




```
WARNING: key_label_font_size is not a recognized aesthetic. Ignoring.
WARNING: minor_label_color is not a recognized aesthetic. Ignoring.
WARNING: minor_label_font_size is not a recognized aesthetic. Ignoring.
WARNING: major_label_color is not a recognized aesthetic. Ignoring.
WARNING: major_label_font_size is not a recognized aesthetic. Ignoring.
WARNING: key_title_font_size is not a recognized aesthetic. Ignoring.
```

In [9]:
```
function E_step(x,mu_S,mu_V,mu_R,sigma,p_S, p_V, p_R)
    numerator1 = p_S*pdf.(Normal(mu_S,sigma),x);
    numerator2 = p_V*pdf.(Normal(mu_V,sigma),x);
    numerator3 = p_R*pdf.(Normal(mu_R,sigma),x);
    denom = numerator1.+numerator2.+numerator3;
    post_S = numerator1./denom;
    post_V = numerator2./denom;
    post_R = numerator3./denom;
    return post_S, post_V, post_R;
end
```

Out[9]: E_step (generic function with 1 method)

In [10]:
```
function M_step(x,post_S, post_V, post_R)
    mu_S = (post_S'*x)./sum(post_S);
    mu_V = (post_V'*x)./sum(post_V);
    mu_R = (post_R'*x)./sum(post_R);
    return mu_S, mu_V, mu_R;
end
```

Out[10]: M_step (generic function with 1 method)

In [12]:
```
function EM(x,mu_S,mu_V,mu_R,p_S, p_V, p_R,sigma)
    maxIter = 1000;
    for i=1:maxIter
        print(i,"\n");
        post_S, post_V, post_R = E_step(x,mu_S,mu_V,mu_R,sigma,p_S, p_V, p_R);
 #print(post_x,"\n");
        mu_S_new, mu_V_new, mu_R_new = M_step(x,post_S, post_V, post_R); #prin
t(mu_S_new," ",mu_V_new,"\n");
        if(abs(mu_S-mu_S_new)<0.0001 && abs(mu_V-mu_V_new)<0.0001 && abs(mu_R-
mu_R_new)<0.0001)
            break;
        end;
        mu_S = mu_S_new;
        mu_V = mu_V_new;
        mu_R = mu_R_new;
    end
    return mu_S, mu_V, mu_R;
end
```

Out[12]: EM (generic function with 1 method)

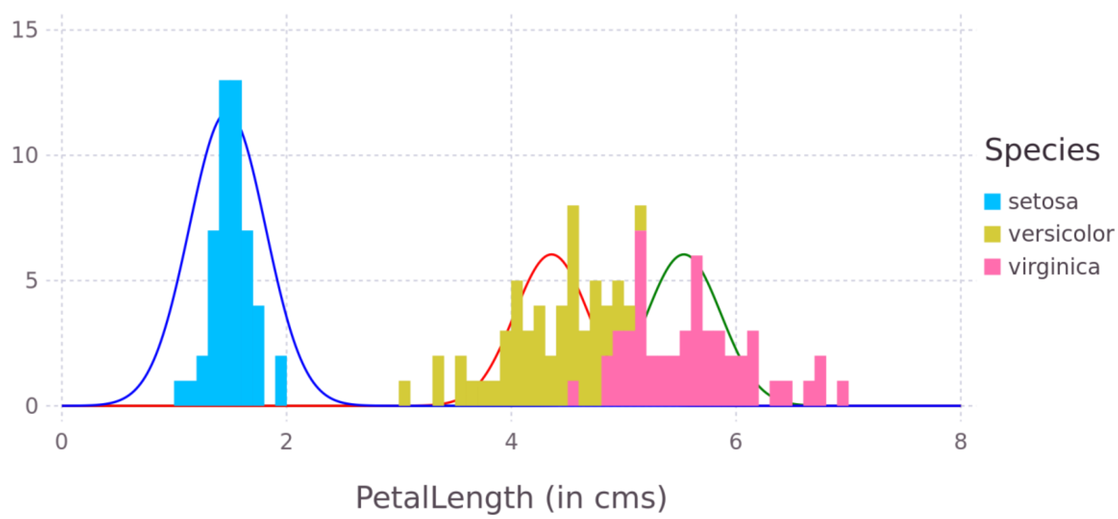
In [13]:
```
data = dataset("datasets","iris");
x = data[:PetalLength];
mu_S=2;
mu_V=4;
mu_R=6;
p_S = 0.34;
p_V = 0.33;
p_R = 0.33;
sigma=0.54;
mu_S, mu_V, mu_R = EM(x,mu_S,mu_V,mu_R,p_S,p_V,p_R,sigma)
```

```
1
2
```

Out[13]: (1.4734289595257766, 4.3579060135505046, 5.5355429945989085)

```
3
4
5
6
7
8
9
10
11
```

In [14]:
```
myplot = plot(layer(data,x=:PetalLength, color=:Species, Geom.histogram,
              Theme(default_color=colorant"purple")),
              layer(x=0:0.02:8,y=pdf.(Normal(mu_S,0.34),0:0.02:8)*10,Geom.line
,
              Theme(default_color=colorant"blue")),
              layer(x=0:0.02:8,y=pdf.(Normal(mu_V,0.33),0:0.02:8)*5,Geom.line,
              Theme(default_color=colorant"red")),
              layer(x=0:0.02:8,y=pdf.(Normal(mu_R,0.33),0:0.02:8)*5,Geom.line,
              Theme(default_color=colorant"green")),
              Guide.xlabel("PetalLength (in cms)"),Guide.ylabel(""), major_lab
el_font_size=18pt,
              minor_label_font_size=14pt,
              key_title_font_size = 18pt,
              key_label_font_size = 14pt,
              major_label_color=colorant"black",
              minor_label_color=colorant"black",Coord.Cartesian(xmin=0, xmax=8
));
draw(PNG(6inch, 3inch, dpi=300), myplot);
```




```
WARNING: key_label_font_size is not a recognized aesthetic. Ignoring.
WARNING: minor_label_color is not a recognized aesthetic. Ignoring.
WARNING: minor_label_font_size is not a recognized aesthetic. Ignoring.
WARNING: major_label_color is not a recognized aesthetic. Ignoring.
WARNING: major_label_font_size is not a recognized aesthetic. Ignoring.
WARNING: key_title_font_size is not a recognized aesthetic. Ignoring.
```

Based on visual inspection we see that setosa has a good fit as compared to versicolor and virginica. Since it does not have any overlapping section with the 2. So if given a petal length we can certainly identify if the species is setosa or not. But between verscicolor and virginica it is difficult because or overlapping between them.

In [ ]:

In [ ]:

In [ ]:

In [ ]:

```
```

In [15]:

```
### BONUS QUESTION ###
```

## To estimate σ using Mixture Model

Differentiating $N(x;\mu,\sigma^2)$ w.r.t. $\sigma^2$.

$$\frac{d}{d\sigma^2} N(x;\mu,\sigma^2) = \frac{d}{d\sigma^2}\left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{\sigma^2}}\right]$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}}\left[\frac{d}{d\sigma^2} e^{-(x-\mu)^2\sigma^{-2}}\right] + e^{-\frac{(x-\mu)^2}{\sigma^2}} \frac{d}{d\sigma^2}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\right]$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{\sigma^2}} \cdot \left[\frac{2(x-\mu)^2}{\sigma^3}\right] + e^{-\frac{(x-\mu)^2}{\sigma^2}} \frac{d}{d\sigma^2}\left[\frac{(\sigma^2)^{-1/2}}{\sqrt{2\pi}}\right]$$

$$= N(x;\mu,\sigma^2) \cdot \left[\frac{2(x-\mu)^2}{\sigma^3}\right] + e^{-\frac{(x-\mu)^2}{\sigma^2}} \frac{1}{\sqrt{2\pi}}\left[\frac{-1}{2}(\sigma^{2})^{-3/2}\right]$$

$$= N(x;\mu,\sigma^2)\left[\frac{2(x-\mu)^2}{\sigma^3}\right] + e^{-\frac{(x-\mu)^2}{\sigma^{2}}} \frac{1}{\sqrt{2\pi}}\left[-\frac{1}{2\sigma^3}\right]$$

$$= N(x;\mu,\sigma^2)\left[\frac{2(x-\mu)^2}{\sigma^3}\right] + \left(e^{-\frac{(x-\mu)^2}{\sigma^{2}}} \frac{1}{\sqrt{2\pi\sigma^2}}\right)\left(\frac{-1}{2\sigma^2}\right)$$

$$= N(x;\mu,\sigma^2)\left[\frac{2(x-\mu)^2}{\sigma^3}\right] + N(x;\mu,\sigma^2)\left(-\frac{1}{2\sigma^2}\right)$$

$$\frac{d}{d\sigma^2} N(x; \mu, \sigma^2) = N(x; \mu, \sigma^2)\left[\frac{2(x-\mu)^2}{\sigma^3} - \frac{1}{2\sigma^2}\right]$$

$\sigma^2_{setosa}$

$$P(\text{Setosa} \mid x_i) = \frac{\pi_s N(x_i, \mu_s, \sigma^2_{setosa})}{\pi_s N(x_i, \mu_s, \sigma^2_{setosa}) + \pi_R N(x_i, \mu_R, \sigma^2_{virginica}) + \pi_V N(x_i, \mu_V, \sigma^2_{versi})}$$

$$P(\text{versicolor} \mid x_i) = \frac{\pi_V N(x_i, \mu_V, \sigma^2_{versicolor})}{\pi_s N(x_i, \mu_s, \sigma^2_{setosa}) + \pi_V N(x_i, \mu_V, \sigma^2_{versicolor}) + \pi_R N(x_i, \mu_r, \sigma^2_{virginica})}.$$

$$P(\text{virginica} \mid x_i) = \frac{\pi_R N(x_i, \mu_R, \sigma^2_{virginica})}{\pi_s N(x_i, \mu_s, \sigma^2_{setosa}) + \pi_V N(x_i, \mu_V, \sigma^2_{versicolor}) + \pi_R N(x_i, \mu_R, \sigma^2_{virginica})}.$$

$\pi_s N = \pi_s N(x_i, \mu_s, \sigma^2_{setosa})$

For Calculating $\sigma_s^2$ :

$$\sum_{i=1}^{150} \frac{\pi_s N}{\pi_s N + \pi_v N + \pi_R N} \left[ \frac{2(x-\mu)^2}{\sigma^3} - \frac{1}{2\sigma^2} \right] = 0$$

$$P(S|x_i) = \frac{\pi_s N}{\pi_s N + \pi_v N + \pi_R N}$$

$$\sum_{i=1}^{n=150} P(S|x_i) \left[ \frac{2(x-\mu)^2}{\sigma^3} - \frac{1}{2\sigma^2} \right] = 0.$$

$$\sum_{i=1}^{150} p(S|x_i) \frac{1}{\cancel{\sigma^2}} \left[ \frac{4(x-\mu)^2 - \sigma}{\cancel{2}\sigma} \right] = 0.$$

$$\sigma_s = \frac{\sum_{i=1}^{n=150} 4(x-\mu_s)^2 P(S|x_i)}{\sum_{i=1}^{150} P(S|x_i)}$$

Similarly for $\sigma_v$ and $\sigma_r$ :

$$\sigma_v = \frac{\sum_{i=1}^{n=150} 4(x-\mu_v)^2 P(V|x_i)}{\sum_{i=1}^{n=150} P(V|x_i)}$$

$$\sigma_R = \frac{\sum_{i=1}^{150} 4(x-\mu_R)^2 P(R|x_i)}{\sum_{i=1}^{n=150} P(R|x_i)}$$

In [17]:
```
function E_modified_step(x,mu_S,mu_V,mu_R,sigma_S, sigma_V, sigma_R,p_S, p_V, 
p_R)
    numerator1 = p_S*pdf.(Normal(mu_S,sigma_S),x);
    numerator2 = p_V*pdf.(Normal(mu_V,sigma_V),x);
    numerator3 = p_R*pdf.(Normal(mu_R,sigma_R),x);
    denom = numerator1.+numerator2.+numerator3;
    post_S = numerator1./denom;
    post_V = numerator2./denom;
    post_R = numerator3./denom;
    return post_S, post_V, post_R;
end
```

Out[17]: E_modified_step (generic function with 1 method)

In [26]:
```
function M_modified_step(x,post_S, post_V, post_R, mu_S, mu_V, mu_R)
    sigma_S = ((post_S)''*(4*(x.-mu_S)^2))./sum(post_S);
    sigma_V = ((post_V)''*(4*(x.-mu_V)^2))./sum(post_V);
    sigma_R = ((post_R)''*(4*(x.-mu_R)^2))./sum(post_R);
    return sigma_S, sigma_V, sigma_R;
end
```

Out[26]: M_modified_step (generic function with 1 method)

In [27]:
```
function EM_modified(x,mu_S,mu_V,mu_R,p_S, p_V, p_R,sigma_S, sigma_V, sigma_R)
    maxIter = 1000;
    for i=1:maxIter
        print(i,"\n");
        post_S, post_V, post_R = E_modified_step(x,mu_S,mu_V,mu_R,sigma_S, sig
ma_V, sigma_R,p_S, p_V, p_R); #print(post_x,"\n");
        sigma_S_new, sigma_V_new, sigma_R_new = M_modified_step(x,post_S, post
_V, post_R, mu_S, mu_V, mu_R); #print(mu_S_new," ",mu_V_new,"\n");
        if(abs(sigma_S-sigma_S_new)<0.0001 && abs(sigma_V-sigma_V_new)<0.0001 
&& abs(sigma_R-sigma_R_new)<0.0001)
            break;
        end
        sigma_S = sigma_S_new;
        sigma_V = sigma_V_new;
        sigma_R = sigma_R_new;
    end
    return sigma_S, sigma_V, sigma_R;
end
```

Out[27]: EM_modified (generic function with 1 method)

In [28]:
```
using RDatasets, Gadfly, Distributions;
data = dataset("datasets","iris");
x = data[:PetalLength];
mu_S=1.473;
mu_V=4.358;
mu_R=5.536;
p_S = 0.34;
p_V = 0.33;
p_R = 0.33;
sigma_S=0.54;
sigma_V=0.54;
sigma_R=0.54;
sigma_S, sigma_V, sigma_R = EM_modified(x,mu_S,mu_V,mu_R,p_S, p_V, p_R,sigma_S
, sigma_V, sigma_R)
```

```
DimensionMismatch("Cannot multiply two vectors")

Stacktrace:
 [1] power_by_squaring(::Array{Float64,1}, ::Int64) at ./intfuncs.jl:169
 [2] M_modified_step(::Array{Float64,1}, ::Array{Float64,1}, ::Array{Float64,
1}, ::Array{Float64,1}, ::Float64, ::Float64, ::Float64) at ./In[26]:2
 [3] EM_modified(::Array{Float64,1}, ::Float64, ::Float64, ::Float64, ::Float
64, ::Float64, ::Float64, ::Float64, ::Float64, ::Float64) at ./In[27]:6
 [4] include_string(::String, ::String) at ./loading.jl:522

1
```

In [ ]:
```
```