# HW1 Answer Key

Jan 14 2016

**Problem 1**

**(a)** $f(y;k) = k\frac{y^{k-1}}{y_{max}^k}$, taking log and we have $\log f(y;k) = (k-1)\log y + \log\left(\frac{k}{y_{max}^k}\right)$. Therefore, $a(y) = \log(y)$, $b(k) = k-1$, $c(k) = \log\left(\frac{k}{y_{max}^k}\right)$, and thus it belongs to exponential family.

**(b)**

$$\mathrm{E}[Y] = \int_0^{y_{max}} y f(y) dy = \frac{k}{y_{max}^k}\int_0^{y_{max}} y^k dy = \frac{k}{y_{max}^k}\frac{y_{max}^{k+1}}{k+1} = \frac{k}{k+1}y_{max}$$

$$\mathrm{Var}[Y] = \mathrm{E}[Y^2] - (\mathrm{E}[Y])^2 = \int_0^{y_{max}} y^2 f(y) dy - (\mathrm{E}[Y])^2 = \frac{k}{k+2}y_{max}^2 - \frac{k^2}{(k+1)^2}y_{max}^2 = \frac{k}{(k+1)^2(k+2)}y_{max}^2$$

**(c)**

```
dpareto<-function(x,ymax,k){k*(x^(k-1))/(ymax^k)}
ppareto<-function(x,ymax,k){(x/ymax)^k}
x<-seq(0,10,0.001)
plot(x,dpareto(x,10,0.5),type="l",col=2,ylab="pdf of x",main="Density of
    Pareto Distributions with Different Parameters")
lines(x,dpareto(x,10,1),col=3)
lines(x,dpareto(x,10,2),col=4)
legend("topright",pch=0,c("k=0.5", "k=1", "k=2"), col = 2:4)

plot(x,ppareto(x,10,0.5),type="l",col=2,ylab="cdf of x",main="Cumulative
    Density of Pareto Distributions with Different Parameters")
lines(x,ppareto(x,10,1),col=3)
lines(x,ppareto(x,10,2),col=4)
legend("bottomright",pch=0,c("k=0.5", "k=1", "k=2"), col = 2:4)
```

Notes: we first define the pdf and cdf function of Pareto distribution, then, generate a grid (`x`) and plot corresponding functions. Use `lines` to append another line on existing plot, and use `legend` to attach legend.

**Problem 2**

**(a)**

```
d<-read.dta("your directory/org_example.dta")
d<-subset(d,year==2008)
```

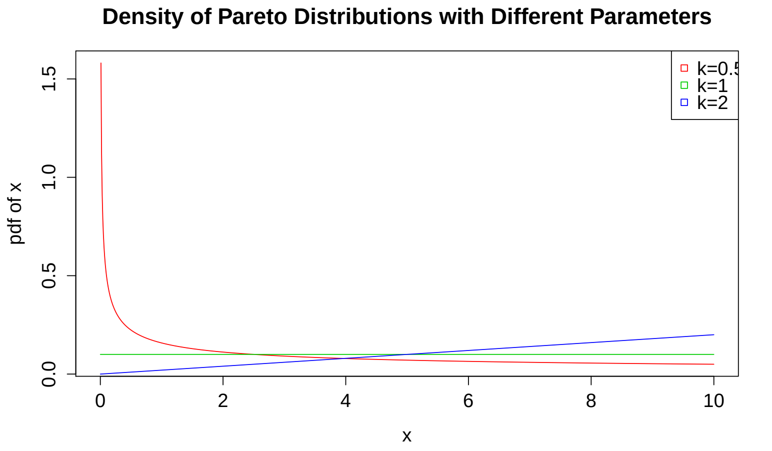


Figure 1: Plot

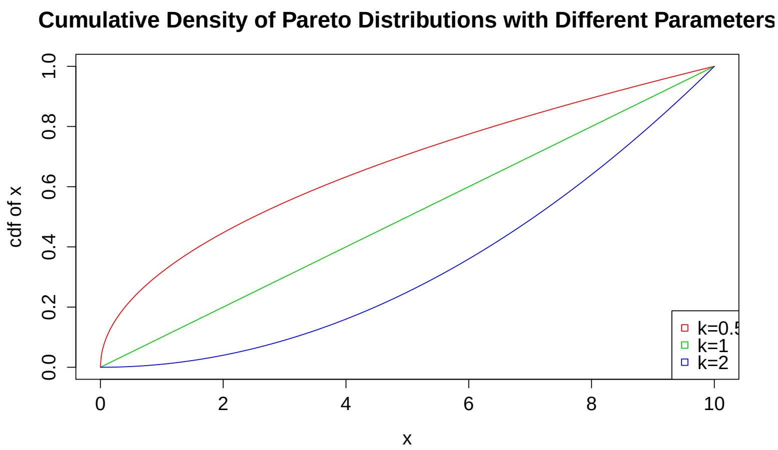


Figure 2: Plot

**(b)**

```
d<-subset(d,nilf==0)
```

**(c)** By typing

```
glm_probit<-glm(unem~educ,d,family=binomial(link="probit"))
summary(glm_probit)
```

you will get the following result

```
Call:
glm(formula = unem ~ educ, family = binomial(link = "probit"),
    data = d)

Deviance Residuals:
```

```
    Min       1Q   Median       3Q      Max
-0.5008  -0.3705  -0.3095  -0.2427   2.7574

Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)      -1.18579    0.01653  -71.75   <2e-16 ***
educHS           -0.31792    0.01976  -16.09   <2e-16 ***
educSome college -0.49126    0.02069  -23.75   <2e-16 ***
educCollege      -0.70958    0.02397  -29.60   <2e-16 ***
educAdvanced     -0.82203    0.03122  -26.33   <2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 43241  on 104043  degrees of freedom
Residual deviance: 41944  on 104039  degrees of freedom
AIC: 41954

Number of Fisher Scoring iterations: 6
```

A person without any degree has the expectation of unemployment rate of 11.79% (`pnorm(-1.18579)`).
A high school degree decreases the unemployment rate by 5.15% (`pnorm(-1.18579)-pnorm(-1.18579-0.31792)`).
Having some college decreases the unemployment rate by 7.11% (`pnorm(-1.18579)-pnorm(-1.18579-0.49126)`).
A college degree decreases the unemployment rate by 8.88% (`pnorm(-1.18579)-pnorm(-1.18579-0.70958)`).
Finally, if you have an advanced degree (Master or higher), your expected unemployment rate will decrease by 9.55% (`pnorm(-1.18579)-pnorm(-1.18579-0.82203)`). These effects are all statistically significant.

**(d)** The marginal effect of an advanced degree will reduce the unemployment rate by 0.67% (`pnorm(-1.18579-0.70958)-pnorm(-1.18579-0.82203)`).