Discussion 2: Numeric and Graphical Summaries and Probability(I)

1. Two methods were studied for the recovery of protein. Thirteen runs were made using each method, and the fraction of protein recovered was recorded for each run. The results are as follows:

Method 1	0.32	0.35	0.37	0.39	0.42	0.47	0.51	0.58	0.60	0.62	0.65	0.68	0.75
Method 2	0.25	0.40	0.48	0.55	0.56	0.58	0.60	0.65	0.70	0.76	0.80	0.91	0.99

   (a) Construct comparative boxplots for the two methods (do this in R and see if values computed by hand match). By hand: Method 1: Q1: 0.39, Median: 0.51, Q3: 0.62, Method 2: Q1: 0.55, Median: 0.60, Q3: 0.76. These values match those calculated by R's summary function.

   (b) Compare the shape, center, and spread of the two methods' fraction of protein recovered. Both roughly symmetric; Median of Method 2 > Median of Method 1; Range of Method 2 > Range of Method 1; IQRs similar with Method 1 IQR slightly larger. Mean Method 2 > Mean Method 1; sd(Method2) > sd(Method1)

2. Four young lab puppies from different litters are available for a new method of training.

Dog	Sex	Age (weeks)
1	M	10
2	M	15
3	F	10
4	F	10

   Two dogs will be selected by lottery to receive the training. Assuming that each pair of dogs are equally likely to receive the training calculate the following probabilities:

   (a) P(A) where A is the event that the dogs are the same gender.12 or 34, so 2/6=1/3

   (b) P(B) where B is the event that the dogs are the same age.13, or 14, or 34, so 3/6=1/2

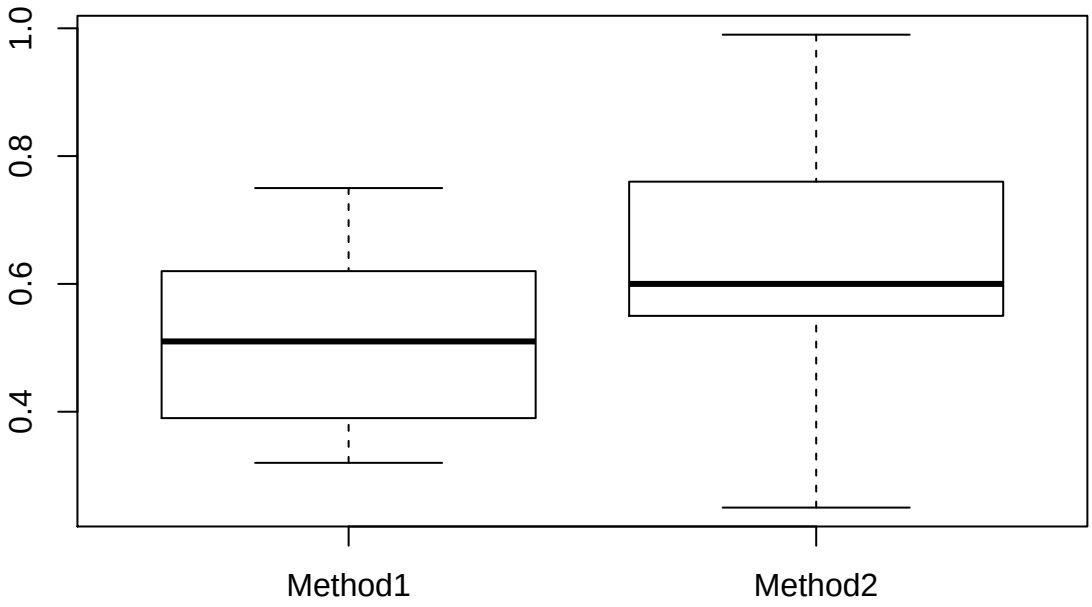
   (c) P(A or B) 12,34,13,14=4/6

   (d) P(A and B) 34=1/6

   (e) Are A and B mutually exclusive? no since they both contain 34 (they have an overlap)

   (f) Are A and B independent? yes. P(A and B)=1/6 and $P(A) * P(B) = 1/3 * 1/2$ or P(A|B)=P(A); 1/3=1/3

```
Method1<-c(0.32 , 0.35 , 0.37 , 0.39 , 0.42 , 0.47 , 0.51 , 0.58 , 0.60 , 0.62 , 0.65 , 0.68 , 0.75)
Method2<-c(0.25 , 0.40 , 0.48 , 0.55 , 0.56 , 0.58 , 0.60 , 0.65 , 0.70 , 0.76 , 0.80 , 0.91 , 0.99)
data<-c(Method1, Method2)
Method=rep(c("Method1", "Method2"), each=13)
boxplot(data~Method)
```




```
summary(Method1)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.3200  0.3900  0.5100  0.5162  0.6200  0.7500
```

```
mean(Method1); sd(Method1)
```

```
## [1] 0.5161538
```

```
## [1] 0.1402699
```

```
summary(Method2)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.2500  0.5500  0.6000  0.6331  0.7600  0.9900
```

```
mean(Method2); sd(Method2)
```

```
## [1] 0.6330769
```

```
## [1] 0.2027883
```