Problem Statement 1 and 2

For $x_1, x_2, \ldots, x_n$

(i) Mean = $\frac{\sum_{i=1}^{n} x_i}{n}$

(ii) Median if you arrange $x_i's$ in order for n odd number, median is the $x_{\frac{n}{2}}$ and for even n median is $(x_{\frac{n}{2}} + x_{\frac{n+1}{2}})/2$

(iii) Mode is the distinct value of x_i with the highest frequency given that the frequency is greater than 1.

(iv) Standard deviation is the square root of ($\frac{\sum_{i=1}^{n} x_i^2}{n}$ - $mean^2$)

Using R to calculate the above

```
> #Assignment Statistics 1
> #
> #Problem Statement 1
> #
>   # The marks of the 20 students are stored in a vector marks
> marks = c(6,7,5,7,7,8,7,6,9,7, 4,10,6,8,8,9,5,6,4,8)
>
>   # Using the mean function, mean of marks is calcualte and stored in th
e mean_marks
> mean_marks = mean(x = marks)
> mean_marks
[1] 6.85
>
>   # Using the median function, median of marks is calcualte and stored i
n the median_marks
> median_marks = median(x = marks)
> median_marks
[1] 7
>
>   # creating a function that returns the mode of a vector v called rt_mo
de.
> mode_fun <- function(w){
+
+   table_w = table(w)
+
+   if(which.max(table_w) == 1){
+     return('No mode')
+   }else{
+    as.numeric(names(table_w[which.max(table_w)]))
+   }
+
+ }
>   # Mode using the above function
> marks_mode = mode_fun(marks)
> marks_mode
[1] 7
>
>   # Standard devition
> std_marks = sqrt(var(marks))
> std_marks
[1] 1.631112
>
>
>
```

```
> #Problem Statement 2
> #Repeating the same steps as on problem statement 1
>
> calls = c(28, 122, 217, 130, 120, 86, 80, 90, 140, 120, 70, 40, 145,
+           113, 90, 68, 174, 194, 170, 100, 75, 104, 97, 75,123, 100,
+           75, 104, 97, 75, 123, 100, 89, 120, 109)
>   #Mean
> mean_calls = mean(x = calls)
> mean_calls
[1] 107.5143
>
>   #Median
> median_calls = median(x = calls)
> median_calls
[1] 100
>
> # Mode
> calls_mode = mode_fun(calls)
> calls_mode
[1] 75
>
> # Standard devition
> std_calls = sqrt(var(calls))
> std_calls
[1] 39.33893
```

Solution for Problem 1

(i) Mean = 6.85
(ii) Median = 7
(iii) Mode = 7
(iv) Standard deviation = 1.6331112

Solution for Problem 2

(i) Mean = 107.5143
(ii) Median = 100
(iii) Mode = 75
(iv) Standard deviation = 39.33893

Problem Statement 3

For $x_1, x_2, \ldots, x_n$ with probability distribution function $f(x_i)$

(i) Mean = $\sum_{i=1}^{n} x_i * f(x_i)$
(ii) Variance = $\sum_{i=1}^{n} x_i^2 * f(x_i) - mean^2$

Using R to calculate the above

```
#Problem statement 3
>
> x = c(0,1,2,3,4,5)
> f_x = c(0.09,0.15,0.4,0.25,0.1,0.01)
>
> #Mean
> mean_x = sum(x*f_x)
> mean_x
```

```
[1] 2.15
>
> #variance
> var_x = sum((x^2)*f_x)-mean_x^2
> var_x
[1] 1.2275


>
```

Solution Problem 3

(i) Mean = 2.15
(ii) Variance = 1.2275

Problem Statement 4

The PDF does not contain the random variable