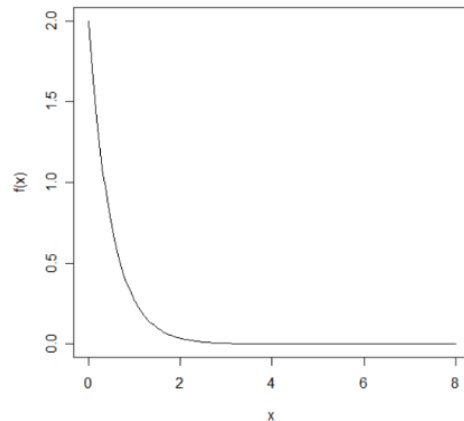


## Statistics 251: Lab 3 Exercises – Distribution Functions

The teaching assistant in the STAT lab noticed that students raised questions at a rate of 2 per minute, and the time between questions was Exponentially distributed.

Let  $X$  = the time between questions

Figure 3.1: The probability density function of the time between questions.



In addition, recall the R commands related to an exponential distribution mentioned in the pre-reading:

```
<d, p, q, r>exp( x, rate),    where  
pexp:      CDF function:  $F(x) = P(X \leq x)$   
dexp:      PDF function.  
rexp:      Random draws from an Exponential distribution.  
qexp:      Gives the quantile function.
```

### Let's find some probabilities and quantiles: (10 min)

- Find the probability that the *time between questions* is
  - One minute or less.
  - Between one and two minutes.
  - At least two minutes.

[Hint: An expression for a probability can look like one of the following:

$P(a < X < b)$ ,  $P(a \leq X \leq b)$ ,  $P(a \leq X < b)$ ,  $P(a < X)$ ,  $P(a \leq X)$ ,  $P(X > b)$ ,  $P(X \geq b)$ , etc.]

- Find the first quartile, median, and third quartile of  $X$ . That is, find a value  $q$ , such that  $F(q) = P(X \leq q) = 0.25, 0.5$  and  $0.75$ .

**Let's do a simulation: (25 min)**

3. Simulate *times between questions* (from the Exponential distribution with the same parameter of interest, i.e. the rate).
  - a. Generate a random sample of six *times between questions*. Draw a histogram representing the probability density of the sample. On top of the histogram, draw the probability density function of an exponential distribution with rate = 2.
  - b. Generate a random sample of 6,000 *times between questions*. Draw a histogram representing the probability density of the sample. On top of the histogram, draw the probability density function of an exponential distribution with rate = 2.
  - c. Which sample appears more representative of its true distribution?

[Hint: When calling the function `hist( )`, specify `freq=FALSE` or `probability=TRUE` as an argument to represent probability densities, rather than frequencies. Use `?hist` to check the help document for details.]