31102023_Statistical_Learning

Mattia G.

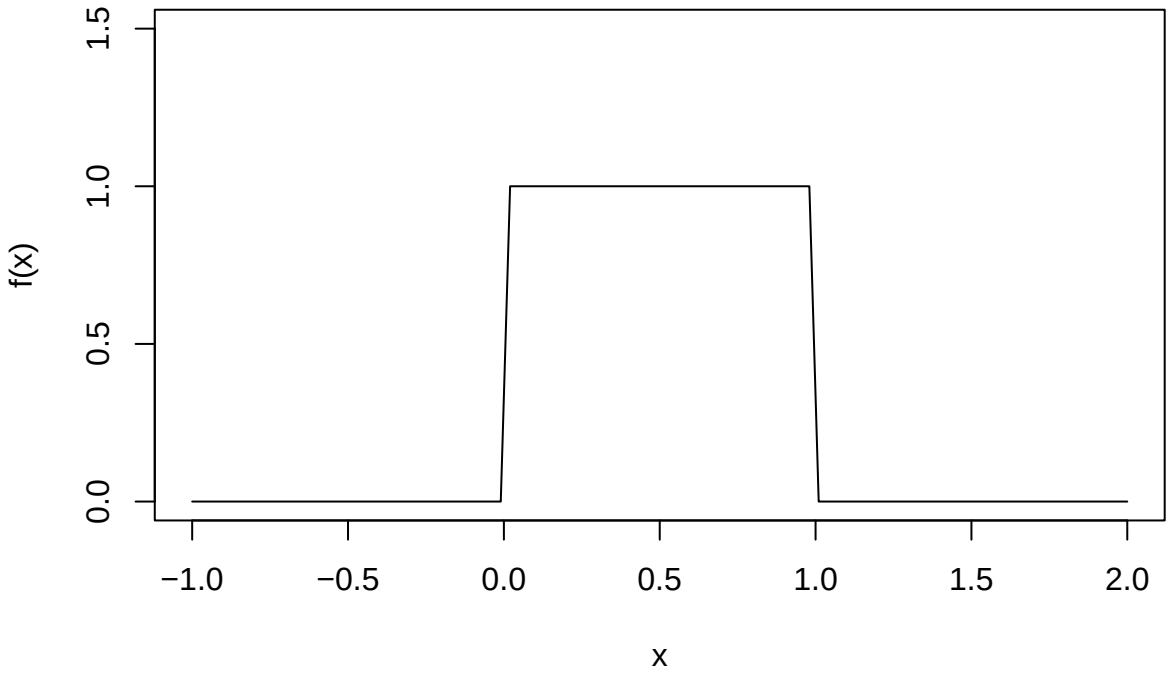
2023-10-31

R Markdown

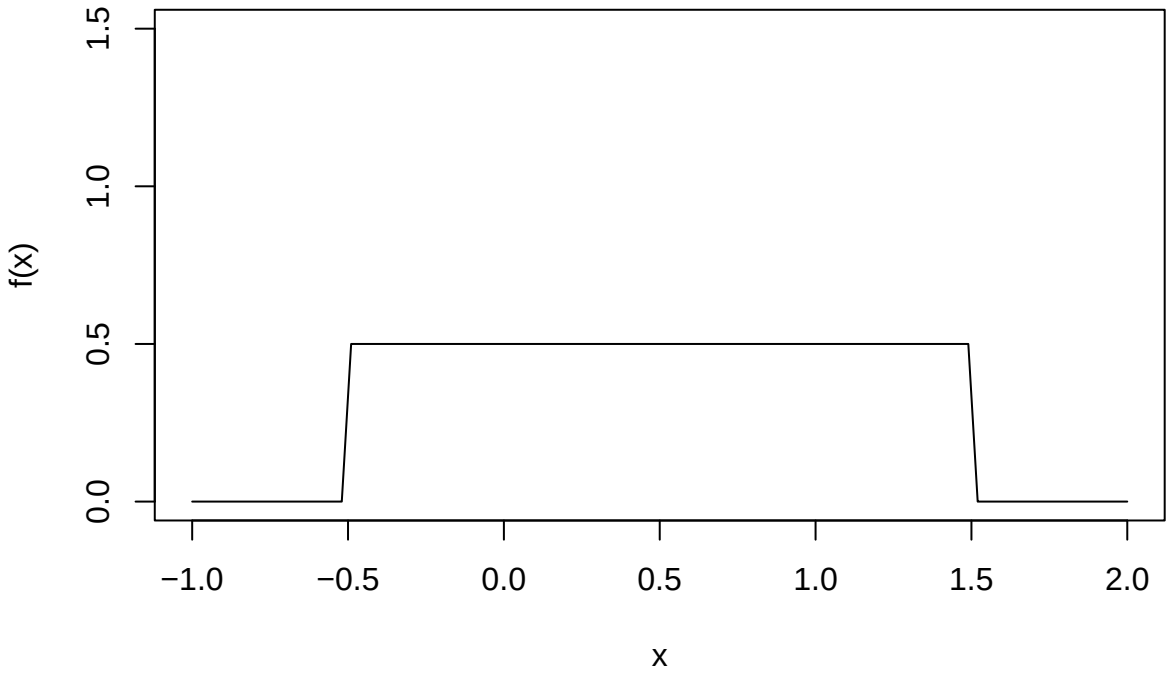
```
########################################
#    PROBABILITY DISTRIBUTIONS
# -- continuous Random Variables --
########################################

# uniform distribution
########################


# density function
dunif(0.2, 0, 1)
```

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

```
# plot the density function
curve(dunif, -1, 2, ylab="f(x)", ylim=c(0,1.5))
```



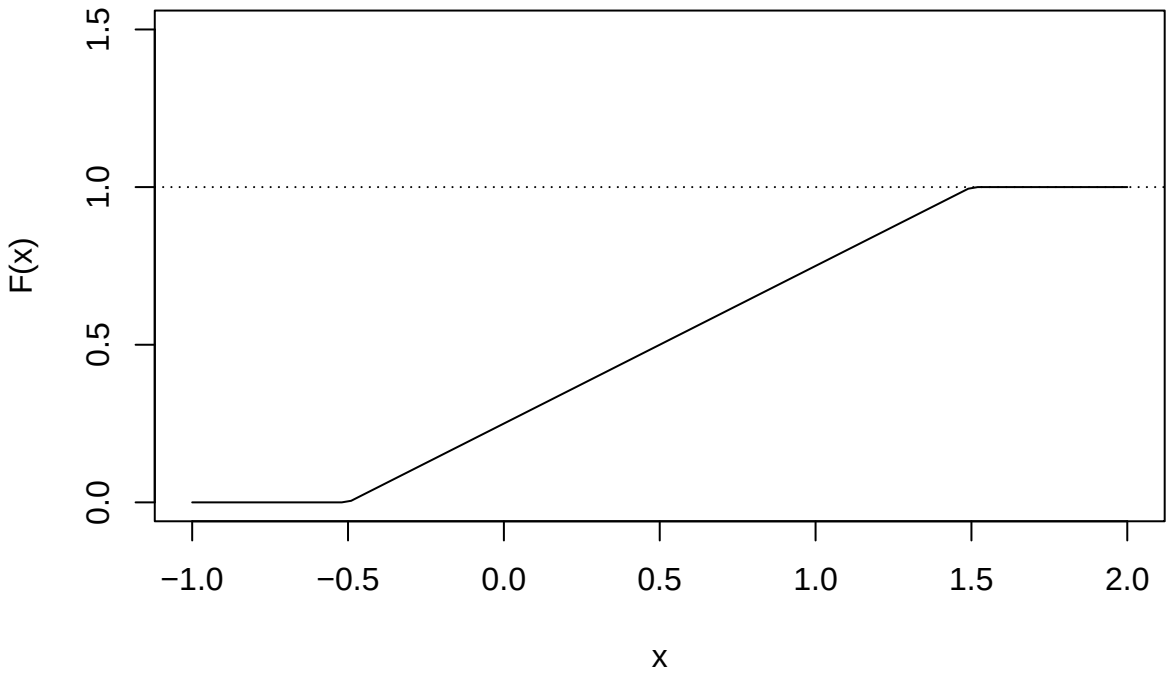

```
curve(dunif(x, min=-0.5, max=1.5), -1, 2, ylab="f(x)", ylim=c(0,1.5))
```




```
# cumulative distribution function
punif(0.2, 0, 1)
```

```
## [1] 0.2
```

```
# plot the cumulative distribution function
curve(punif(x, min=-0.5, max=1.5),-1, 2, ylab="F(x)", ylim=c(0,1.5))
abline(h=1, lty=3)
```




```
# uniformly distributed random values
runif(3, -0.5, 1.5)
```

```
## [1] -0.08483095  1.21656393  1.05333277
```

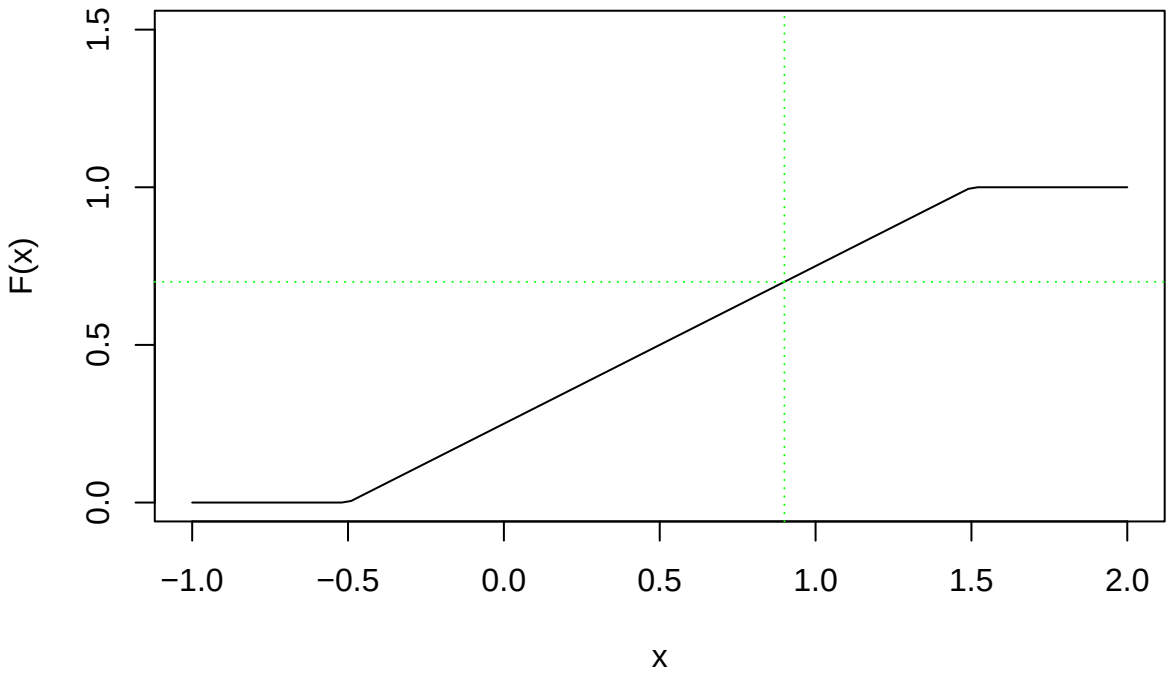
```r
# quantiles of the uniform distribution
qunif(.3, -0.5, 1.5)
```

```
## [1] 0.1
```

```r
# how to see quantiles on the cumulative distribution function

pr <- 0.7
qunif(pr, -0.5, 1.5)
```

```
## [1] 0.9
```

```r
curve(punif(x, min=-0.5, max=1.5), -1, 2, ylab="F(x)", ylim=c(0,1.5))
abline(h=pr, v=qunif(pr, -0.5, 1.5), lty=3, col="green")
```




```r
# exponential distribution
############################

f <- function(x, lambda=1) lambda*exp(-lambda*x)

f(0.5)
```

```
## [1] 0.6065307
```

```r
f(0.5, lambda=1/3)
```

```
## [1] 0.2821606
```

```r
# or equivalently

dexp(0.5)
```

```
## [1] 0.6065307
```

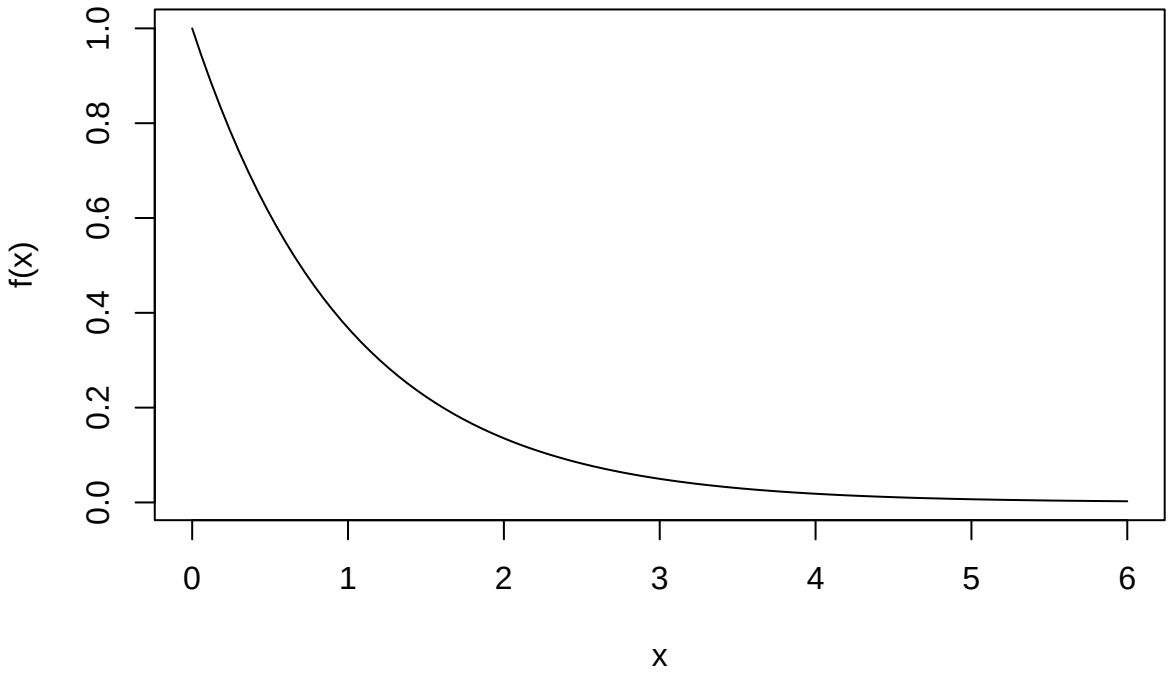
```r
dexp(0.5, rate=1/3)
```

```
## [1] 0.2821606
```

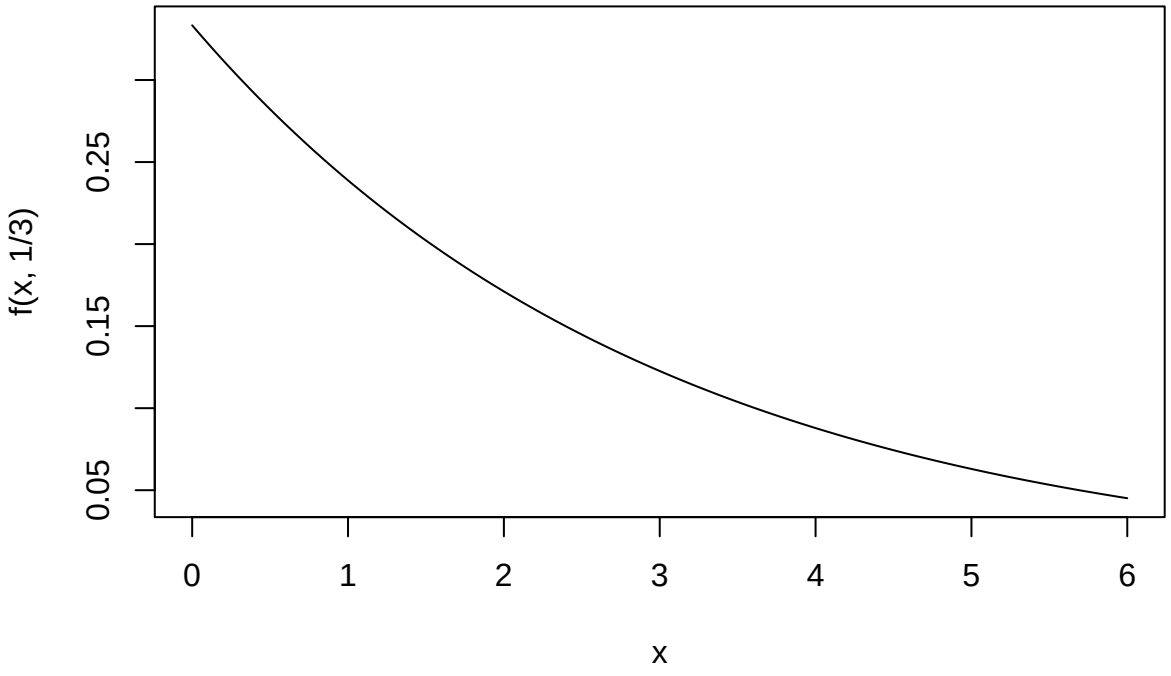
```r
# numerical integration
# check the area under the curve is equal to 1

integrate(f, 0, Inf)
```

```
## 1 with absolute error < 5.7e-05
```

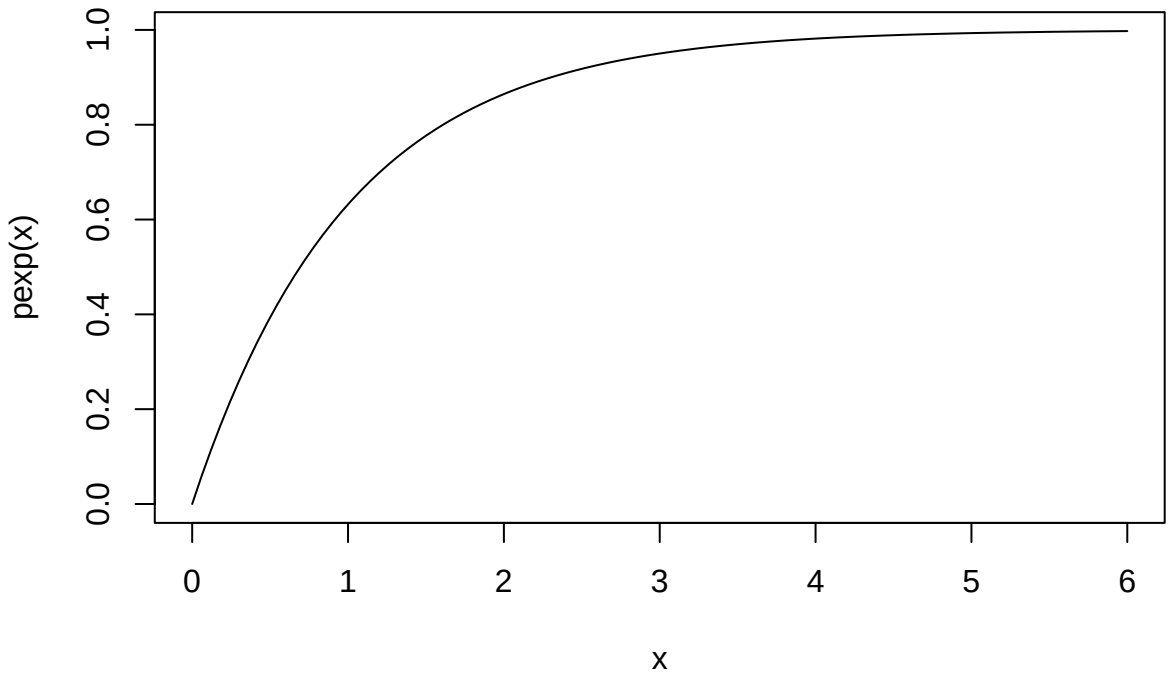
```r
# plot the pdf

curve(f, from=0, to=6)
```



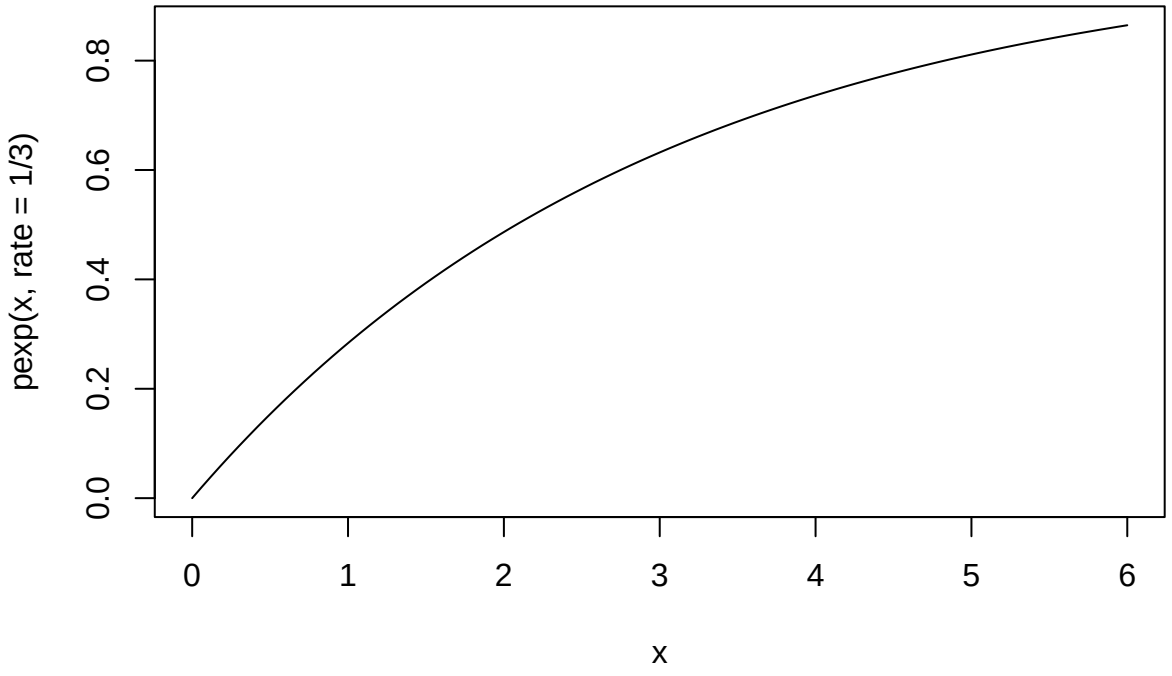

```r
curve(f(x, 1/3), 0, 6)
```



```r
# plot the cdf

curve(pexp, 0, 6)
```



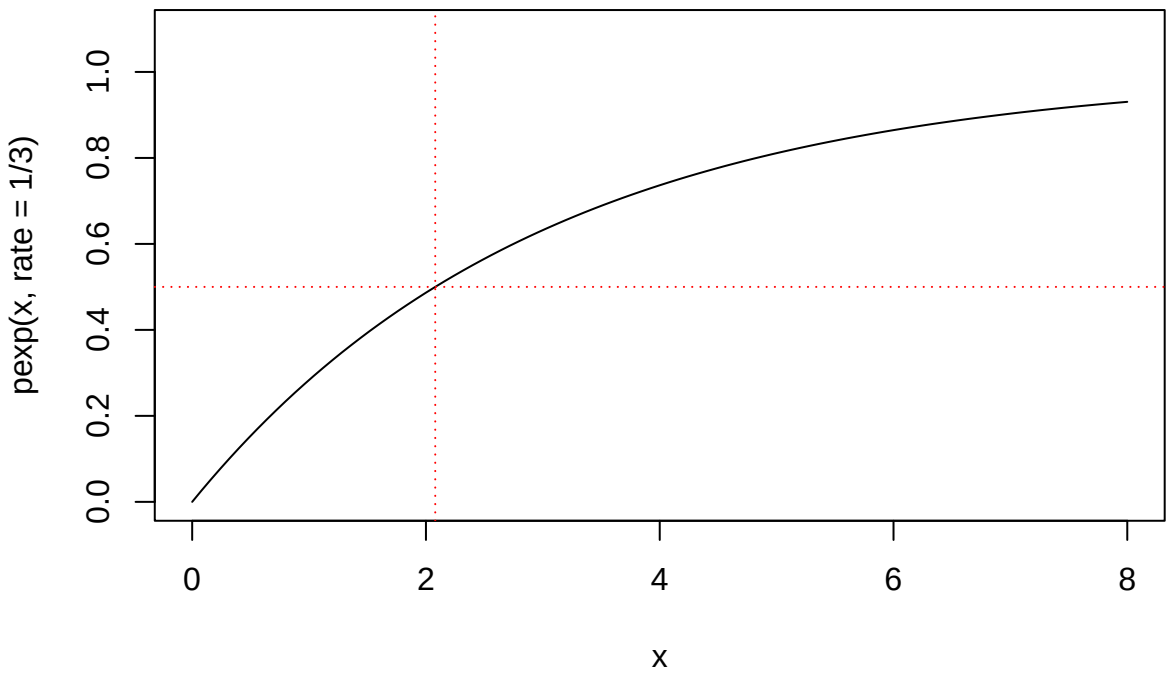

```r
curve(pexp(x, rate=1/3), 0, 6)
```




```r
# quantiles of the exponential distribution


pr <- 0.5
qexp(pr, rate=1/3)
```

```
## [1] 2.079442
```

```
curve(pexp(x, rate=1/3), 0, 8, ylim=c(0, 1.1))
abline(h=pr, v=qexp(pr, rate=1/3), col="red",lty=3)
```




```
# E(X)

f <- function(y) y*exp(-y)
integrate(f, 0, Inf)
```

```
## 1 with absolute error < 6.4e-06
```

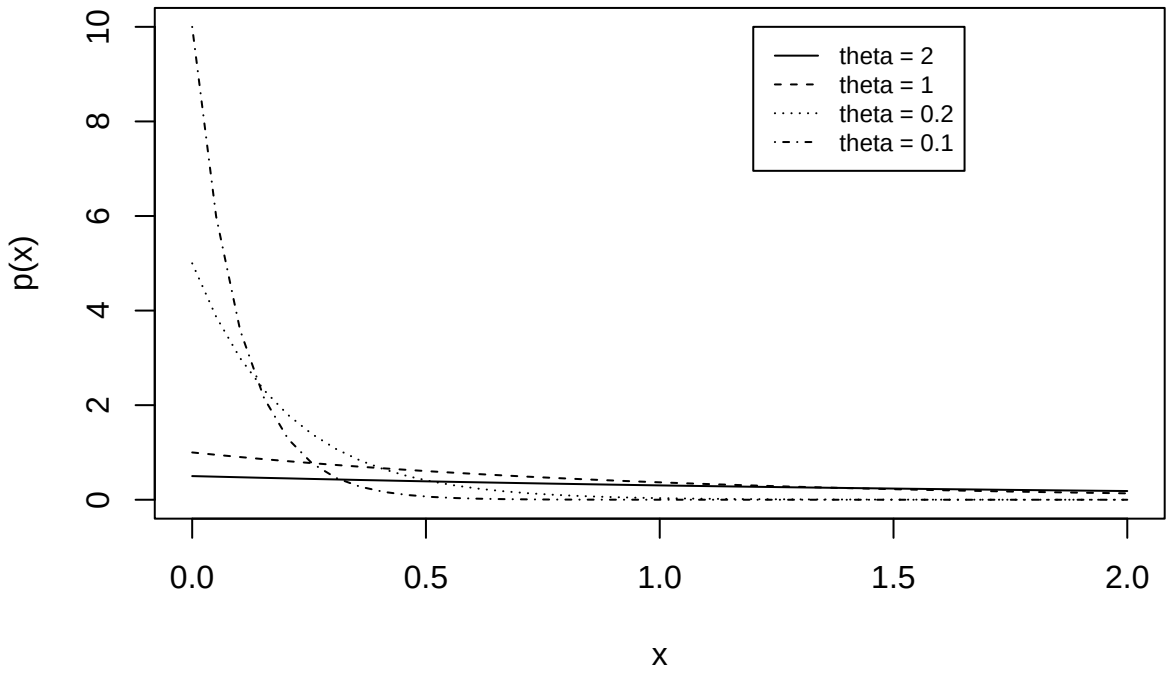
```
# E(Y^2)

f <- function(y) y^2*exp(-y)
integrate(f, 0, Inf)
```

```
## 2 with absolute error < 7.1e-05
```

```
# variance

f <- function(y) (y-1)^2*exp(-y)
integrate(f, 0, Inf)
```

```
## 1 with absolute error < 5.8e-05
```

```
# comparison of exponential densities

x <- seq(0, 2, length=40)
theta <- c(2, 1, .2, .1) # mean of distribution
y <- matrix(NA, 40, 4)
for (i in 1:4) {
  y[,i] <- dexp(x, 1/theta[i]) # parameter is the rate
}
matplot(x, y, type="l", xlab="x", ylab="p(x)", lty=1:4, col=1)
legend(1.2, 10, paste("theta =", theta), lty=1:4, cex=.75)
```


theta = 2
theta = 1
theta = 0.2
theta = 0.1
p(x)
0
2
4
6
8
10
0.0
0.5
1.0
1.5
2.0
x