assignment 1

Norah Alyabs

September 12, 2018

Due date: Thursday, September 20

1. **Textbook problem 1.3** The investigator of a large clinical trial would like to assess factors that might be associated with drop-out over the course of the trial. Describe what would be the event and which observations would be considered censored for such a study.

Answer There are 3 examples in chapter 1, if we focus in the (UMARU) IMPACT Study (UIS) by Drs. Jane McCusker, Carol Bigelow and Anne Stoddard. Multiple regression could be used to determine significant factors that effect the event which is Time to Return to Drug Use (time) which defined as "the number of days from admission to one of the two sites to self-reported return to drug use". Returned to Drug Use (censor) is the censored observation where it is coded 1 for return to drug or lost to follow-up (it considers in the same category because usually those who was lost to follow-up was likely to have returned to drug use based on the study team) and 0 otherwise

2. Let T be a positive continuous random variable, show $E(T) = \int_0^\infty S(t)\,dt$.

Answer Mean survival time is the total area under the survivor curve S(t), for a given time t, as the risk increas,the survival S(t) decrese, and hence the shorter mean survival time E(T), and vice versa.

$$E(T) = \int_0^\infty tf(t)\,dt$$

using integral by parts

$$E(T) = tF(t) - \int_0^\infty F(t)\,dt$$

$$E(T) = t(1 - S(t)) - \int_0^\infty (1 - S(t))\,dt$$

$$E(T) = t - tS(t) - t + \int_0^\infty S(t)\,dt$$

$$E(T) = \int_0^\infty S(t)\,dt$$

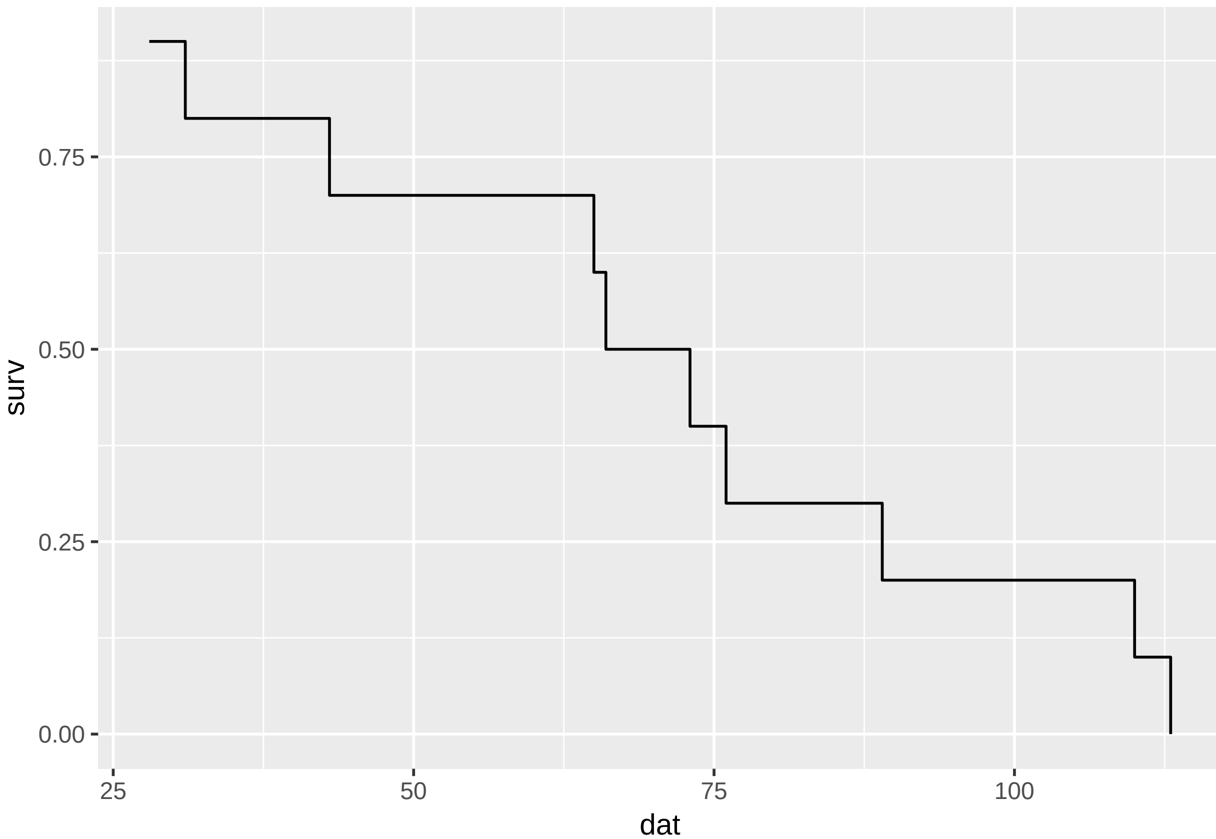
3. Question 2 suggests that the area under the survival curve can be interpreted as the expected survival time. Consider the following hypothetical data set with 10 death times.

```
> dat <- c(43, 110, 113, 28, 73, 31, 89, 65, 66, 76)
```

   a. Plot the empirical survival curve.
   b. Find the expected survival time for the hypothetical data set.

Answer a. The empirical survival curve:

```
> library(ggplot2)
> surv = 1 - ecdf(dat)(dat)
> dat2<-data.frame(dat,surv)
> ggplot(data=dat2,mapping= aes(dat, surv)) + geom_step()
```



b. The expected survival times for dat observations respectively are:

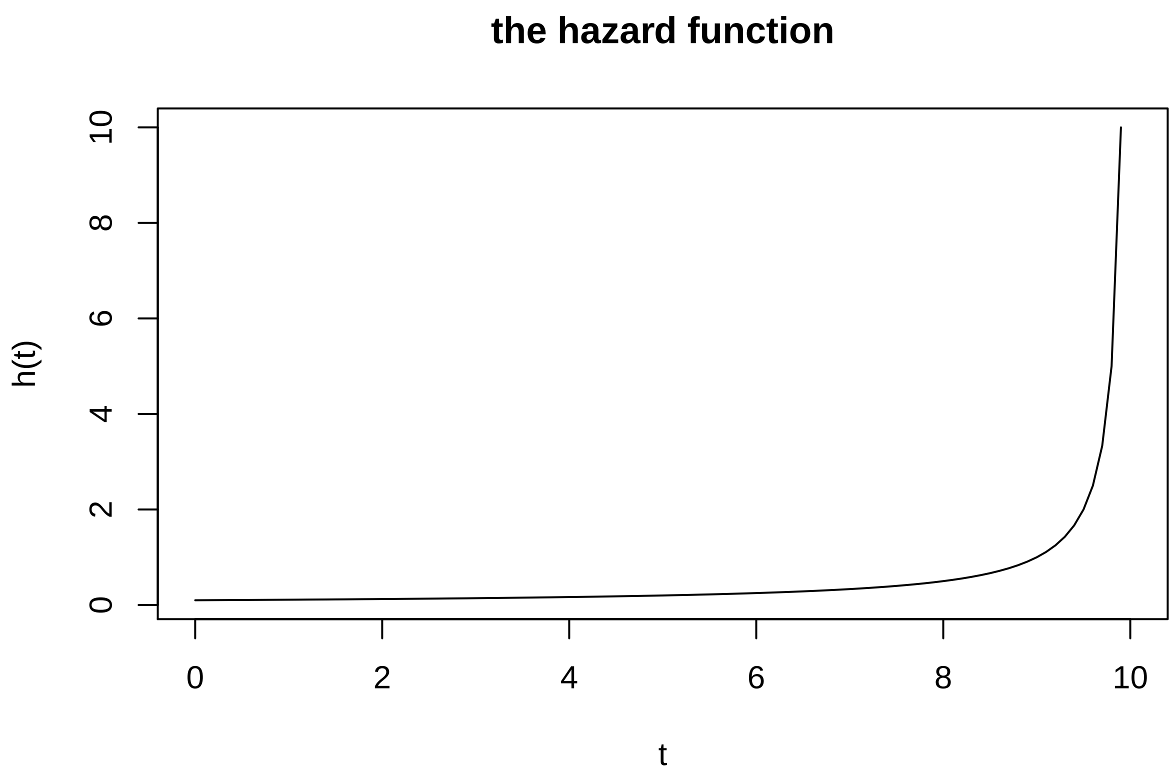
```
> surv
```

```
[1] 0.7 0.1 0.0 0.9 0.4 0.8 0.2 0.6 0.5 0.3
```

4. Consider a survival time random variable with hazard $\lambda(t) = \frac{1}{10-x}$ in $[0, 10)$.
   a. Plot the hazard function.
   b. Plot the survival function.

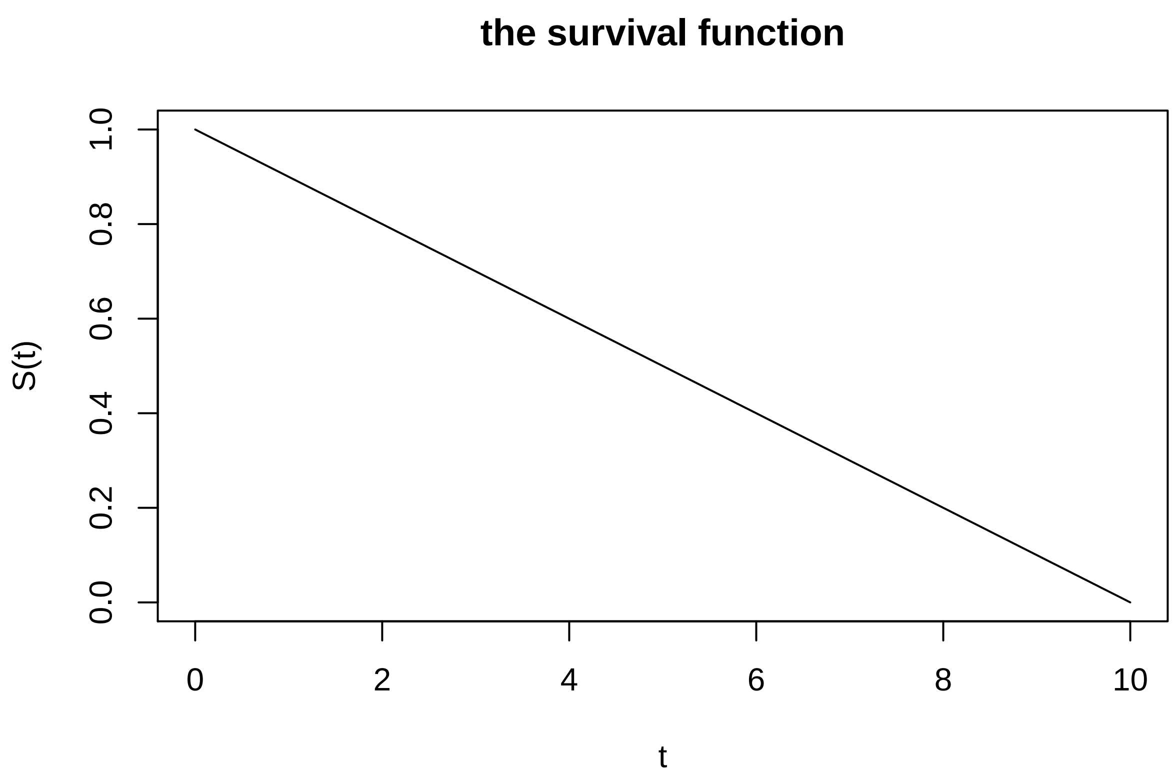
Answer a.

```
> curve(1/(10-x), from=0, to=10,  type = "l", ylab ="h(t)", xlab = "t", main="the hazard function")
```



b. $S(t) = \frac{10-t}{10}, (0 =< t < 10)$

```
> curve((10-x)/10, from=0, to=10, , add = FALSE, type = "l", ylab ="S(t)", xlab = "t", main="the surviva
```



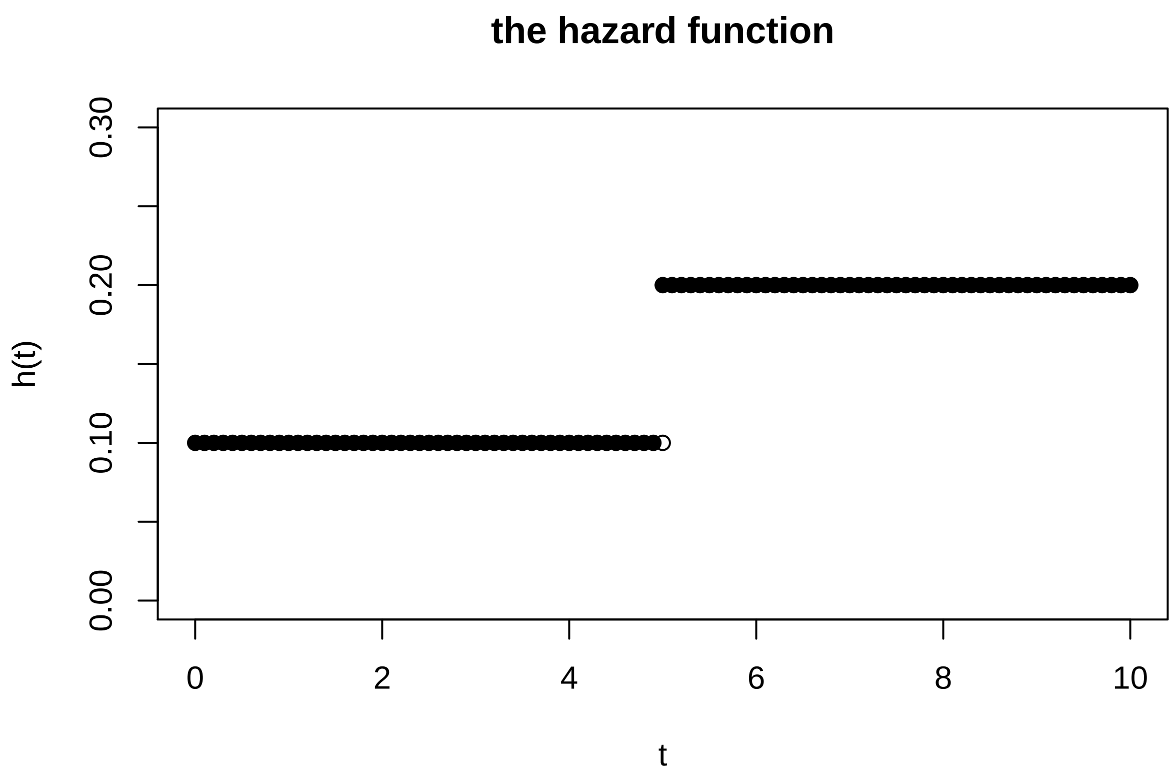
5. Consider a survival time random variable with constant hazard $\lambda = 0.1$ in $[0, 5)$, and $\lambda = 0.2$ in $[5, \infty)$. This is known as a piece-wise constant hazard.

a. Plot the hazard function.
b. Plot the survival function.

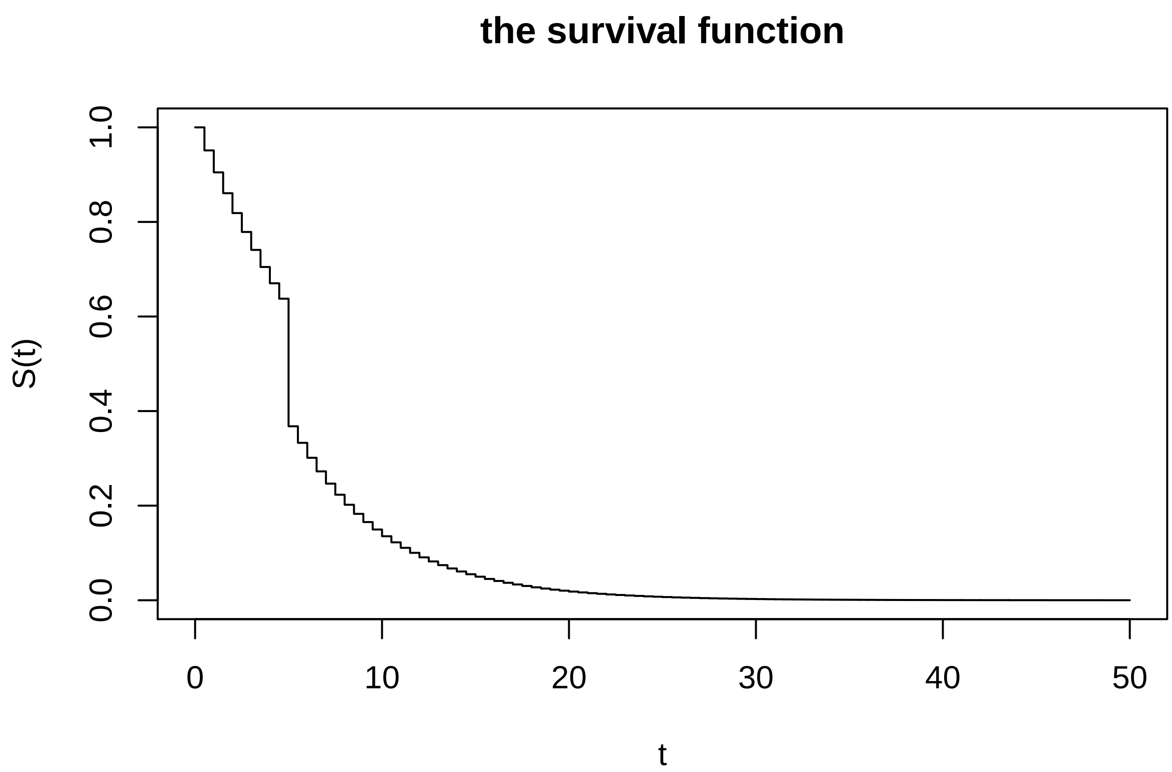
Answer

```
> #a
> library( mosaicCore)
> f1=makeFun(0.1~x)
> f2=makeFun(0.2~x)
> hazard <- function(x) {(x >= 0 & x < 5)*f1(x) + (x >= 5)*f2(x)}
> curve(hazard, xlim = c(0,10), ylim=c(0,0.3) , xlab = "t", ylab = "h(t)", type = "p", pch=19, main="the
> points(5,0.1, pch=1, add=TRUE)
```

```
Warning in plot.xy(xy.coords(x, y), type = type, ...): "add" is not a
graphical parameter
```




```
> #b
> f11=makeFun(exp(-0.1*x)~x)
> f22=makeFun(exp(-0.2*x)~x)
> Survival<- function(x) {(x >= 0 & x < 5)*f11(x) + (x >= 5)*f22(x)}
> curve(Survival, xlim = c(0,50), ylim=c(0,1) , xlab = "t", ylab = "S(t)", type = "s", main="the surviva
```


the survival function
S(t)
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
t