STA406 Worksheet 1

Jessica Lam, Tomas Martinek

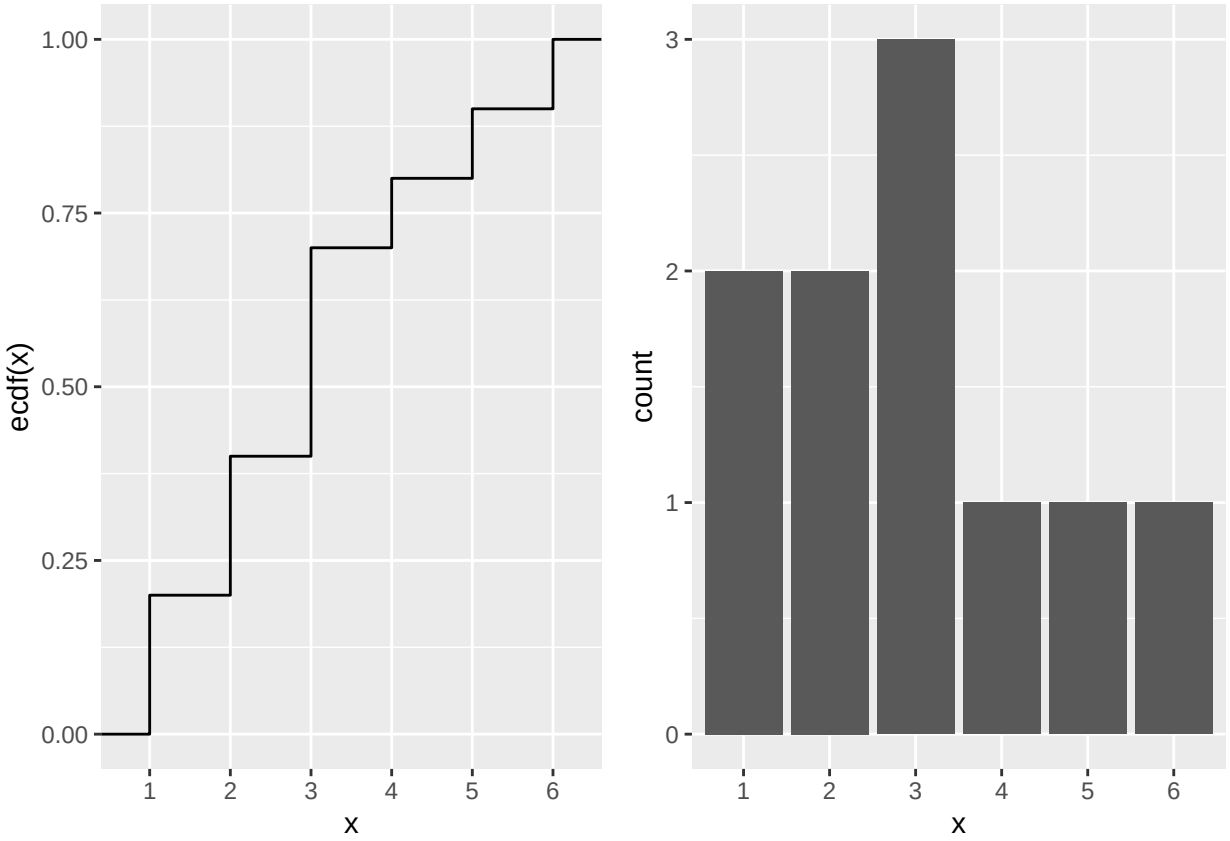
Exercise 1

(a) Roll a die 10 times and compute the ecdf.

```
library(ggplot2)
library(gridExtra)
set.seed(1)

rolls <- data.frame(x=sample(1:6, 10, replace=TRUE))

base <- ggplot(rolls) + scale_x_discrete(limits=factor(1:6)) +
  scale_y_continuous(limits=0:1) + expand_limits(x=c(0.4, 6.6))
ecdf_1a <- base + geom_step(aes(x=x, y=..y..), stat="ecdf") +
  ylab(bquote("ecdf(x)"))
hist_1a <- ggplot() + geom_bar(data=rolls, aes(x=x)) +
  scale_x_discrete(limits=factor(1:6))
grid.arrange(ecdf_1a, hist_1a, ncol=2)
```

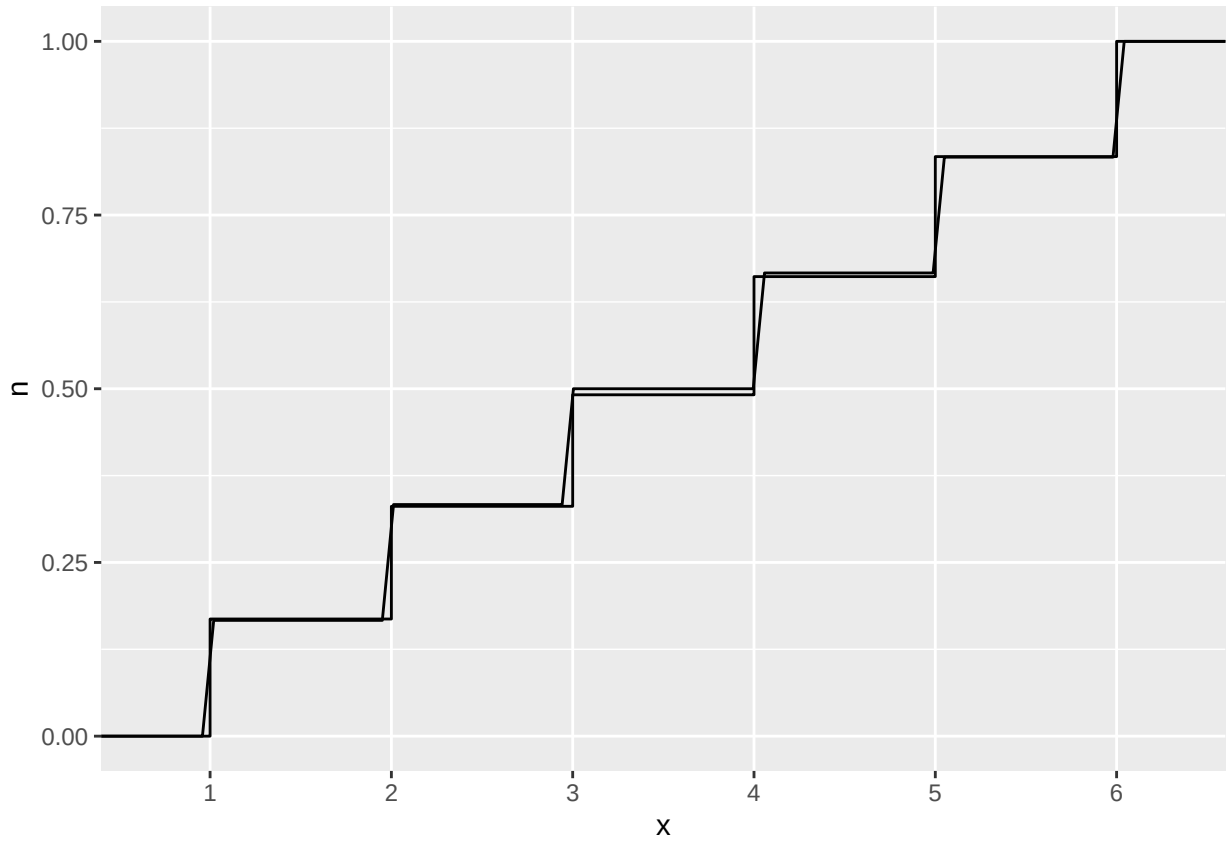


(b) Roll the die 10,000 times. Compare the ecdf and the cdf.

```
set.seed(1)

rolls <- data.frame(x=sample(1:6, 10000, replace=TRUE))

cdf_die <- function(x) {
  return(pmin(1, pmax(0, floor(x)/6)))
}
plot_1b <- base +
  geom_step(data=rolls, aes(x=x, y=..y..), stat="ecdf") +
  stat_function(fun=cdf_die) + ylab("n")
plot_1b
```


1.00
0.75
0.50
0.25
0.00
n
1
2
3
4
5
6
x