Exercises: Chapter 3

LinMod

Monday, September 15, 2014

3-1 The running proportions plot.

```
# Goal: toss a coin N times and get the running proportion of heads
N <- 500
p.heads <- 0.8 # Assume biased coin

# Heads = 1, Tails=0
flipsequence <- sample(x=c(0,1), prob=c(1-p.heads,p.heads), size=N, replace=TRUE)

r <- cumsum(flipsequence)
n <- 1:N

runprop <- r/n
```

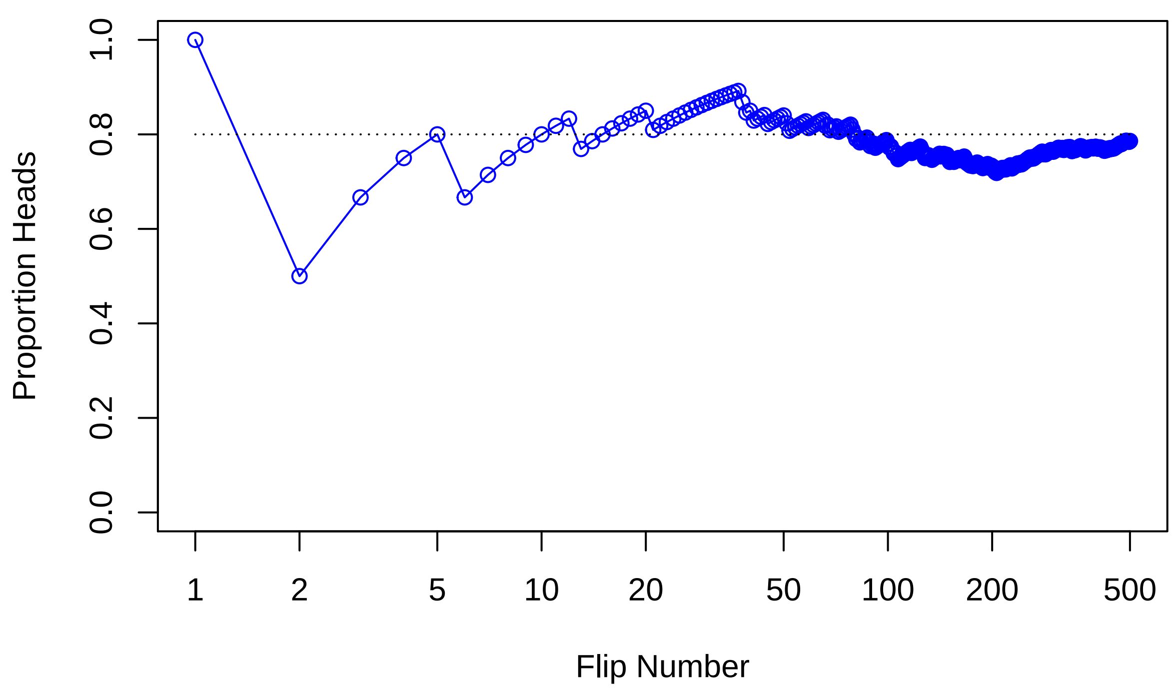

Figure 1: Running Proportions Plot, # of Heads

3-2 Detemine the probability of drawing a 10 from a pinochle deck (9-Ace, 6 of each).

To do this, we will randomly sample from the deck and see what proportion are tens.

```
cards <- c("9", "10", "Jack", "Queen", "King", "Ace")
suits <- c("Hearts", "Diamonds", "Spades", "Clubs")

draws <- cbind.data.frame(sample(cards, size=1000000, replace=TRUE),
                          sample(suits, size=1000000, replace=TRUE))

length(draws[which(draws[,1]=='10'),1])/length(draws[,1])
```

```
## [1] 0.167
```

```
length(draws[which(draws[,1]=='10'|draws[,1]=='Jack'),1])/length(draws[,1])
```

```
## [1] 0.3336
```

Thus, we estimate that the probability of drawing a jack is 0.167 and the probability of drawing a 10 or a jack is 0.334.

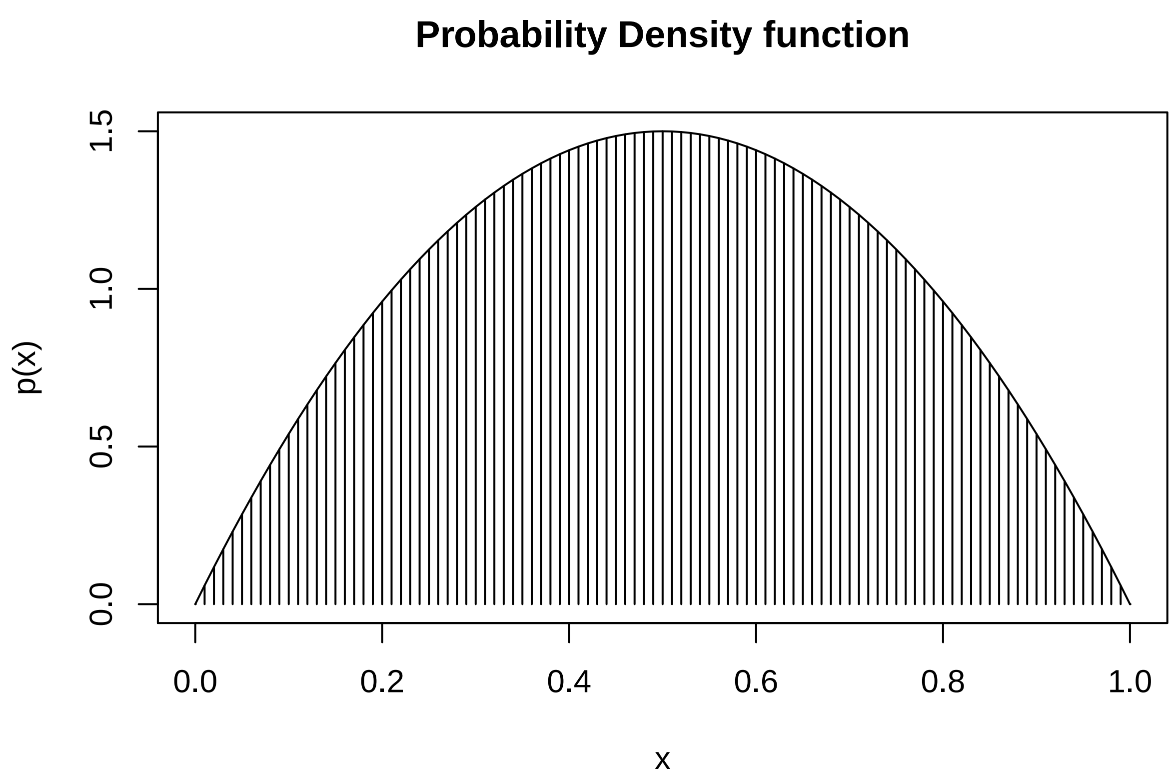

Figure 2: PDF - $6x(1-x)$

3-3 From this plot, we estimate that the area underneath the curve is

$$\sum_{i}^{n} p(x_i)\delta x$$

which is estimated numerically by `R` as 0.9999. Analytically, the solution is

$$\begin{aligned}
\int 6x(1-x)\delta x &= \int (6x - 6x^2)\delta x \\
&= \int 6x\delta x - \int 6x^2\delta x \\
&= 3x^2|_0^1 - 2x^3|_0^1 \\
&= (3-2) - (0-0) \\
&= 1
\end{aligned}$$

Thus, because the area under the curve is 1, $p(x)$ is a probability density function, and satisfies $Equation 3.3$.

3-4

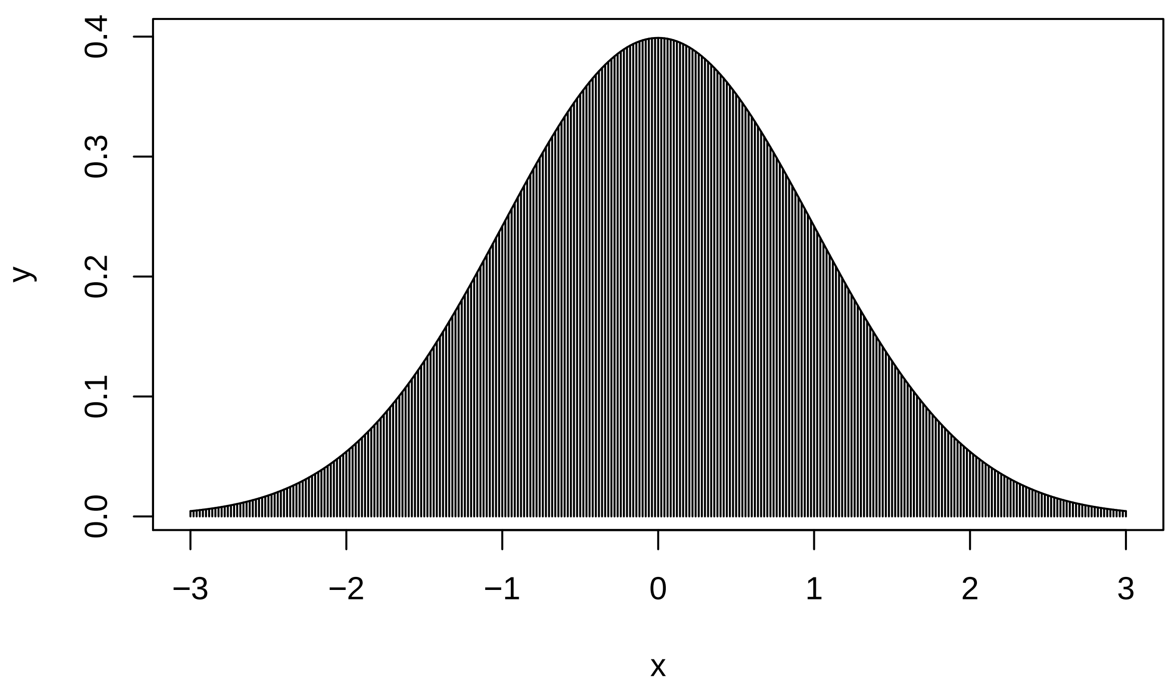

Figure 3: plot of chunk normal

A) This is the distribution $Y \sim N(0, 1)$.

B) This statement is essentially that we believe 68% of women will fall within the range of 147cm and 177cm in height, with the average woman's height at 162cm.

	Ice Cream	Fruit	French Fries
1st	0.3	0.6	0.1
6th	0.6	0.3	0.1
11th	0.3	0.1	0.6

3-5 This table represents $p(snack|grade)$. We want to find $p(grade|snack)$. In order to find this, we need $p(y, x)$. Using some algebraic manipulation, we see that

$$p(x|y) = \frac{p(x, y)}{p(y)}$$
$$p(x|y)p(y) = p(x, y)$$

Multiplying our table by the proportion of students within each grade gives us this joint distribution $p(x, y)$. Thus, our new table becomes

	Ice Cream	Fruit	French Fries
1st	0.06	0.12	0.02
6th	0.12	0.06	0.02
11th	0.18	0.06	0.36

We know that these probabilities are not independent because we cannot multiply the marginal distributions together to get each cell.