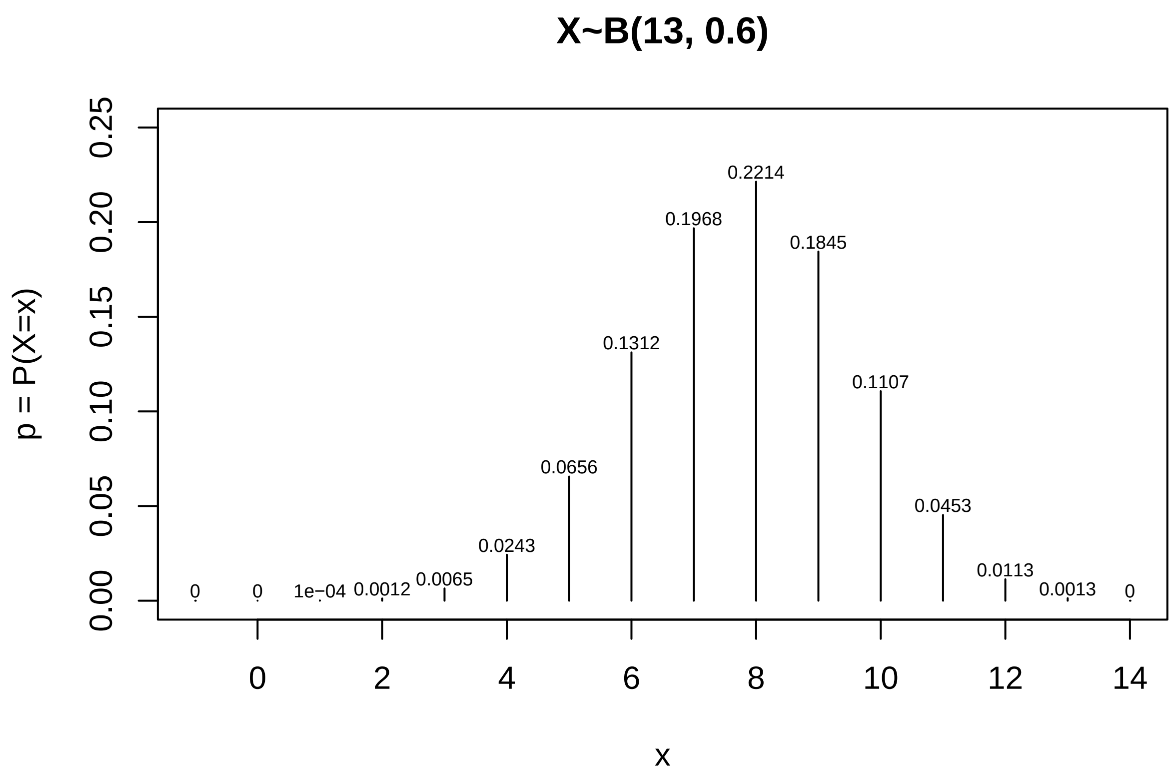
Binomial pmf and CDF with Most Probable Value

Oliver d'Pug

pmf

We can use the base plot functions in R to create a plot of the pmf for a binomial random variable X with $n = 13$ and $p = 0.6$ — i.e. $X\ \tilde{}B(13,\ 0.6)$.

```
n <- 13
p <- 0.6
x <- -1:14
pmf <- dbinom(x, n, p) ### dbinom() is the pmf
plot(x, pmf, type="h", xlab="x", ylab="p = P(X=x)", main="X~B(13, 0.6)",
     ylim=c(0, 0.25), xaxt="n")
axis(1,at=seq(0,14,by=2))
text(x, pmf+0.005, round(pmf, digits=4), cex=0.6)
```

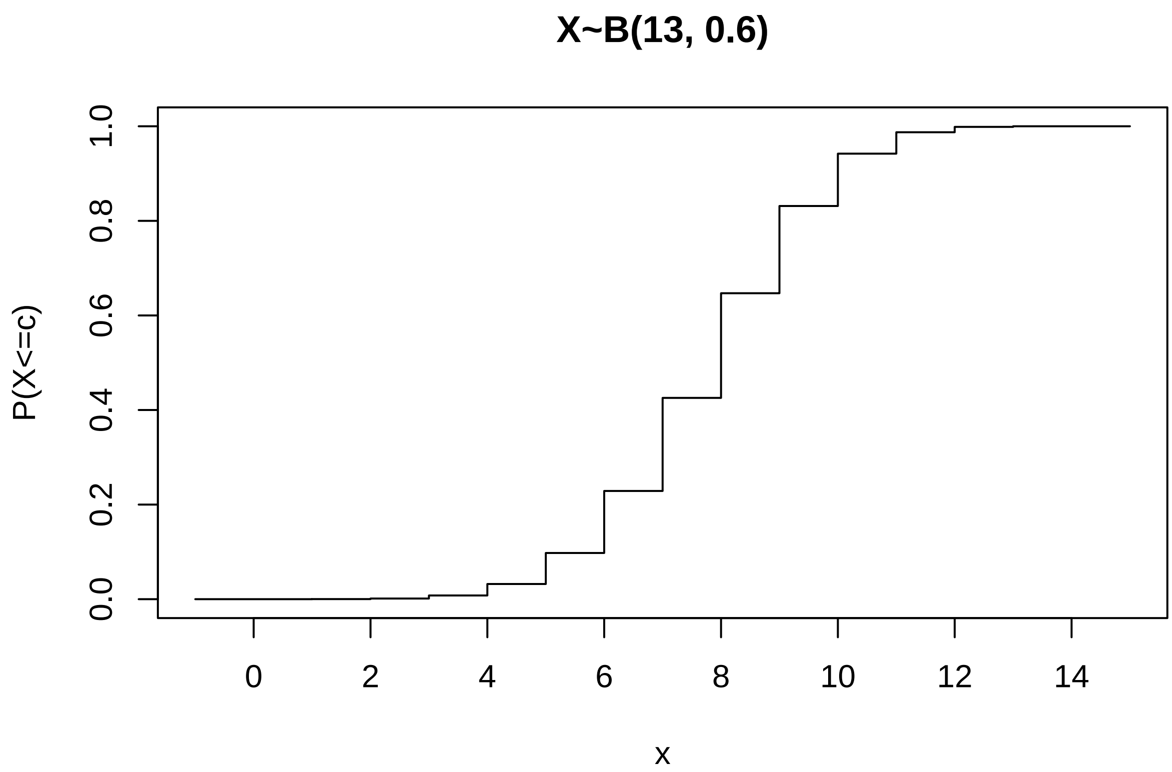


We note that the maximum is at 8 .

CDF

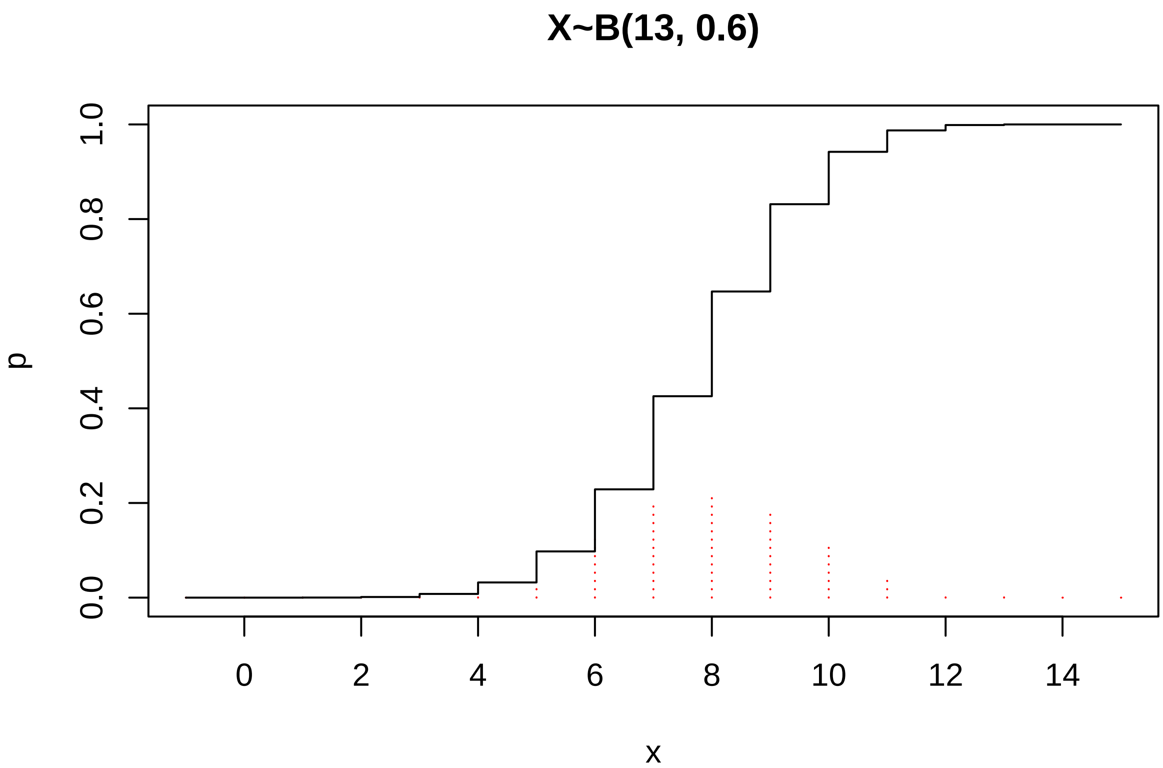
The CDF may be plotted analogously.

```
x <- -1:15
cdf <- pbinom(x, n, p) ### pbinom() is the CDF
plot(x, cdf, type="s", xlab="x", ylab="P(X<=c)",
     main="X~B(13, 0.6)", xaxt="n")
axis(1, at=seq(0,14,by=2))
```



Just for fun we can overlay the two.

```
pmf <- dbinom(x, n, p)
plot(x, pmf, type="h", xlab="x", ylab="p", main="X~B(13, 0.6)",
     ylim=c(0, 1), xaxt="n", col="red", lty=3)
  lines(x, cdf, type="s", xlab="x", ylab="P(X<=c)",
        main="X~B(13, 0.6)", xaxt="n")
axis(1,at=seq(0,14,by=2))
```



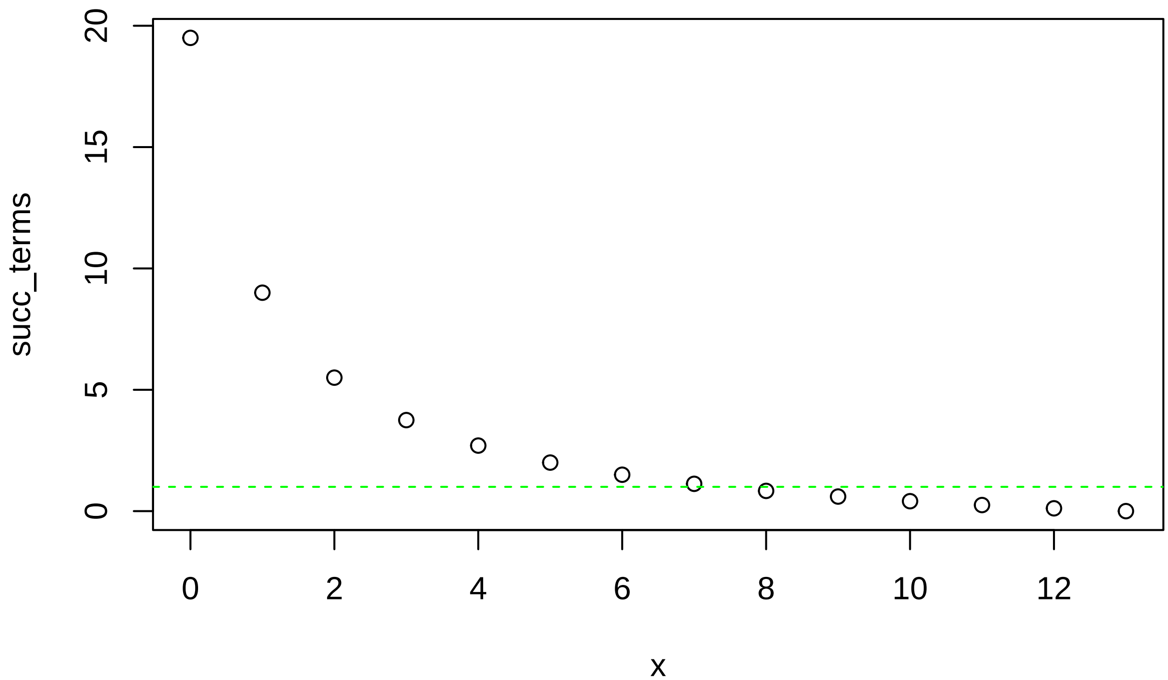
Most Probable Value

Looking at successive terms allows us to find the most probable values of X. We note that

$$\frac{p(x+1)}{p(x)} = \frac{n-x}{x+1} \cdot \frac{p}{1-p}$$

We can look at where the ratio is greater/less than one.

```
x <- 0:n
succ_terms = ((n-x)/(x+1))*(p/(1-p))
plot(x, succ_terms)
abline(h=1, lty=2, col="green")
```




```
cbind(x, succ_terms)
```

```
##        x succ_terms
##  [1,]  0 19.5000000
##  [2,]  1  9.0000000
##  [3,]  2  5.5000000
##  [4,]  3  3.7500000
##  [5,]  4  2.7000000
##  [6,]  5  2.0000000
##  [7,]  6  1.5000000
##  [8,]  7  1.1250000
##  [9,]  8  0.8333333
## [10,]  9  0.6000000
## [11,] 10  0.4090909
## [12,] 11  0.2500000
## [13,] 12  0.1153846
## [14,] 13  0.0000000
```

Another way of looking at the problem is to compute the ratios by using the pmf.

```
pmf <- dbinom(x,n,p)
cbind(x, pmf)
```

```
##        x          pmf
##  [1,]  0 6.710886e-06
##  [2,]  1 1.308623e-04
##  [3,]  2 1.177761e-03
##  [4,]  3 6.477683e-03
```

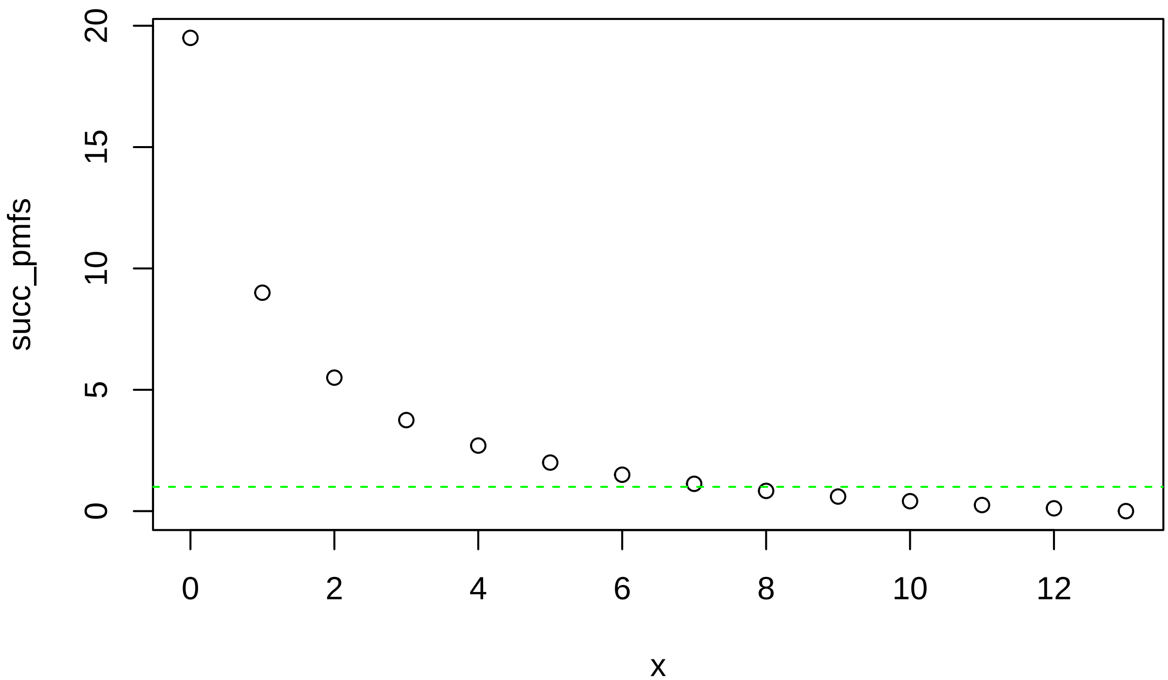
```
## [5,]  4 2.429131e-02
## [6,]  5 6.558654e-02
## [7,]  6 1.311731e-01
## [8,]  7 1.967596e-01
## [9,]  8 2.213546e-01
## [10,]  9 1.844621e-01
## [11,] 10 1.106773e-01
## [12,] 11 4.527707e-02
## [13,] 12 1.131927e-02
## [14,] 13 1.306069e-03
```

```
succ_pmfs <- pmf[2:(n+1)]/pmf[1:n]
succ_pmfs <- c(succ_pmfs,0) ### Last term is 0 because p(n+1) = 0
cbind(x,succ_pmfs)
```

```
##        x  succ_pmfs
##  [1,]  0 19.5000000
##  [2,]  1  9.0000000
##  [3,]  2  5.5000000
##  [4,]  3  3.7500000
##  [5,]  4  2.7000000
##  [6,]  5  2.0000000
##  [7,]  6  1.5000000
##  [8,]  7  1.1250000
##  [9,]  8  0.8333333
## [10,]  9  0.6000000
## [11,] 10  0.4090909
## [12,] 11  0.2500000
## [13,] 12  0.1153846
## [14,] 13  0.0000000
```

Plotting the ratios generates the same plot as above.

```
x <- 0:n
plot(x, succ_pmfs)
abline(h=1, lty=2, col="green")
```



In either case, we see that the function goes from increasing to decreasing at the integer part of $(n+1)p$. This can be computed as $\text{trunc}((n+1)p) = 8$ or $\lceil np \rceil = 8$.

```
(max_x <- trunc((n+1)*p))
```

```
## [1] 8
```

```
(max_x <- ceiling(n*p))
```

```
## [1] 8
```