

HyperGeometric Distribution

Young W. Lim

2018-02-22 Thr

1 HyperGeometric Distribution

- Based on
- HyperGeometric Random Variables
- Cumulative Distributive Function
- Lottery
- Hypergeometric vs. Binomial

"Probability with R: An Introduction with Computer Science Applications" Jane Horgan

I, the copyright holder of this work, hereby publish it under the following licenses: GNU head Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled GNU Free Documentation License.

CC BY SA This file is licensed under the Creative Commons Attribution ShareAlike 3.0 Unported License. In short: you are free to share and make derivative works of the file under the conditions that you appropriately attribute it, and that you distribute it only under a license compatible with this one.

Calculating hypergeometric pdfs

```
x <- 0:5  
dhyper(x, 13, 39, 5)
```

```
x <- 0:3  
dhyper(x, 3, 3, 3)
```

```
x <- 0:4  
dhyper(x, 4, 16, 10)
```

```
x <- 0:10  
dhyper(x, 30, 70, 20)
```

Plotting hypergeometric pdfs

```
par (mfrow = c(2,2))
```

```
x<-0:5
```

```
plot(x, dhyper(x, 13, 39, 5),  
      xlab="X= number of trials", ylab="P(X=x)",  
      type="h", main="N=52, M=13, n=5");
```

```
x<-0:3
```

```
plot(x+1, dhyper(x, 3, 3, 3),  
      xlab="X= number of trials", ylab="P(X=x)",  
      type="h", main="N=6, M=3, n=3");
```

```
x<-0:4
```

```
plot(x+1, dhyper(x, 4, 16, 10),  
      xlab="X= number of trials", ylab="P(X=x)",  
      type="h", main="N=20, M=4, n=10");
```

```
x<-0:20
```

```
plot(x+1, dhyper(x, 30, 70, 20),  
      xlab="X= number of trials", ylab="P(X=x)",  
      type="h", main="N=100, M=30, n=20");
```

Calculating hypergeometric cdfs

`phyper(q, m, n, k)`

- `q` : vector of quantiles representing the number of white balls drawn without replacement from an urn which contains both black and white balls
- `m` : the number of white balls in the urn
- `n` : the number of black balls in the urn
- `k` : the number of balls drawn from the urn

`phyper(2, 13, 39, 5)`

Plotting geometric cdfs

```
par (mfrow = c(2,2))
```

```
x<-0:5
```

```
plot(x, phyper(x, 13, 39, 5),  
      xlab="X= number of trials", ylab="P(X<=x)",  
      type="h", main="p=.95");
```

```
x<-0:3
```

```
plot(x, phyper(x, 3, 3, 3),  
      xlab="X= number of trials", ylab="P(X<=x)",  
      type="h", main="p=.5");
```

```
x<-0:10
```

```
plot(x, phyper(x, 4, 16, 10),  
      xlab="X= number of trials", ylab="P(X<=x)",  
      type="h", main="p=.2");
```

```
x<-0:20
```

```
plot(x, phyper(x, 30, 70, 20),  
      xlab="X= number of trials", ylab="P(X<=x)",  
      type="h", main="p=.01");
```

```
n random numbers are selected from the N  
n=6 favorable  
N-n=30 unfavorable
```

```
x<- 0:6  
round( dhyper(x, 6, 30, 6), 7)
```

```
the length of time expected
```

```
x <- choose(36, 6)  
x/52  
x/(2*52)  
x/(100*52)
```


Winning

```
1-phyper(2, 6, 30, 6)
1-phyper(2, 6, 33, 6)
1-phyper(2, 6, 36, 6)
1-phyper(2, 6, 39, 6)
```

n random numbers are selected from the N
n=6 favorable
N-n=30 unfavorable

```
x<- 0:6
round( dhyper(x, 6, 30, 6), 7)
```

the length of time expected

```
x <- choose(36, 6)
x/52
x/(2*52)
x/(100*52)
```

Hypergeometric vs. Binomial

```
M <- 2
L <- 18
y <- dhyper(x, M, L, 10)
yround = round(y, 3)
names <- c(yround)
plot(x, y, type="h", ylim= c(0, .51), xlab= " ",
      main= Hyper(N=20, p=0.1, axes=FALSE)
text(x, dhyper(x, M, L, 10), names)
axis(1)
axis(2)
```


