IS 606: Statistics and Probability for Data Analytics Final Exam

Ben Arancibia

May 22, 2015

Binomial or Hypergeometric

1) Binomial distribution to calculate the probability that there will be three success out of ten trials, probability of success is 0.30

```
dbinom(3, 10, 0.3) #dbinom(k, n, p)
```

```
## [1] 0.2668279
```

2) Expected value and the standard deviation for binomial experiment with ten trials and P = 0.30

Expected value = mean

n*p = expected value

```
10*.30
```

```
## [1] 3
```

Standard deviation is sqrt of variance.

```
sqrt(10*.3*(1-.3))
```

```
## [1] 1.449138
```

3) Hypergeometric distribution to calculate the probability that there will be three success out of ten trials, n = 20 with eight possible successes.

```
dhyper(3, 8, 12, 10) #dhyper(x, success, n-sucess, trials) (x,m,n,k)
```

```
## [1] 0.2400572
```

4) Expected value and the standard deviation for a hypergeometric experiment with ten trials, n=20 with eight possible successes

Expected value = n*(k/ N)

```
10*(8/20)
```

```
## [1] 4
```

Standard deviation

```
sqrt(10*(8/20)*(1-(8/20))*((20-10)/(20-1)))
```

```
## [1] 1.123903
```

5) hypergeometric distribution to calculate the probability that there will be three success out of ten trials, N=60 with 24 possible successes.

```
dhyper(3, 24, 36, 10)
```

```
## [1] 0.2240987
```

6) Expected value and the Standard deviation for a hypergeometric experiment with ten trials, N=60 with 24 possible successes

Expected value = n*(k/ N)

```
10*(24/36)
```

```
## [1] 6.666667
```

Standard deviation

```
sqrt(10*(24/60)*(1-(24/60))*((60-20)/(60-1)))
```

```
## [1] 1.275586
```

7) What population size does the binomial start to make a strong approximation for the hypergeometric distribution?

The Hypergeometric(n, D, M) can be approximated by Binomial(n, D/M). The approximate works well when $n < 0.1$ M

As long as n is less 0.1 of success trials

Conditional Probability and the Naive Bayes Classification Method