Bernoulli Trials & Binomial Distributions

Definitions:

- Bernoulli Trial: A random event whose outcome is true (1) or false (0).
- Binomial Distribution: n Bernoulli trials.
- p The probability of a true (1) outcome (also called a success)
- q 1-p. The probability of a false (0) outcome (also called a failure)
- n The number of trials
- $f(x|n,p)$ The probability density function with sample size n and probability of success p.

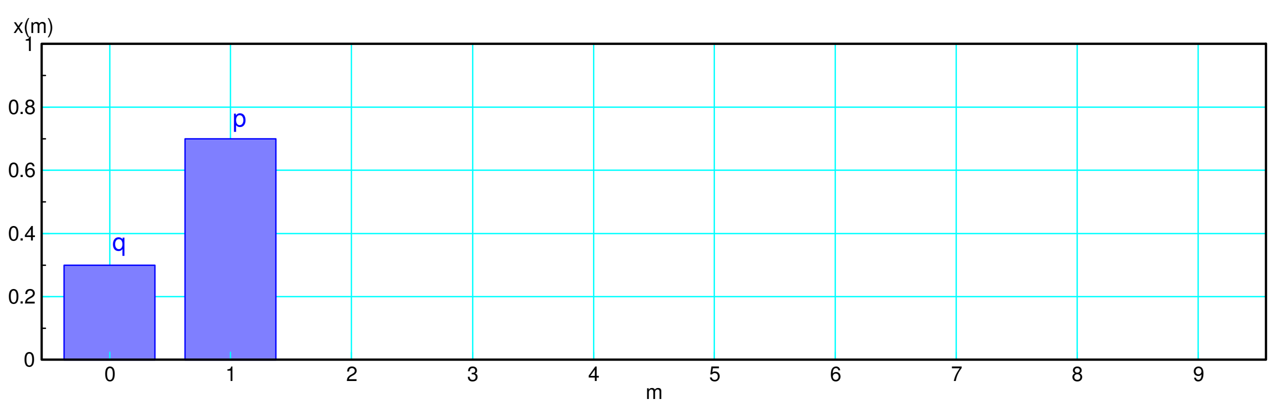
distribution	description	pdf	mgf	mean	variance
Bernoulli trial	flip a coin obtain m heads	$p^m q^{1-m}$	$q+p/z$	p	p(1-p)
Binomial	flip n coins obtain m heads	$\binom{n}{m} p^m q^{n-m}$	$(q+p/z)^n$	np	np(1-p)
Hyper Geometric	Bernoulli trial without replacement	$\frac{\binom{A}{x}\binom{B}{n-x}}{\binom{A+B}{n}}$			

Bernoulli Trial:

A Bernoulli trial is essentially a coin flip:

- There are only two possible outcomes (heads or tails, 1 or 0, success or failure)
- The probability of a success is p
- The probability of a failure is q (which must be 1-p for probabilities to add to 1.000)

Graphically, the probability of m successes with a Bernoulli trial looks like this:



probability density function for a Bernoulli trial. For illustration, q = 0.3, p = 0.7

This is called *the Probability Density Function:* the probability of getting m successes with a Bernoulli trial (single coin toss)

With only two possible outcomes, there are only two bars. You can also write this as

$$x(m) = q \cdot \delta(m) + p \cdot \delta(m-1)$$

Note that the mathematics doesn't care what integer *m* represents:

- In ECE 343 Signals and Systems, m represents time
- In ECE 341 Random Processes, m represents the number of successes

The math doesn't care: it's just an integer, m. Likewise, the same math used in Signals and Systems (i.e. the z-transform) applies to Random Processes.

The z-operator is defined as

$$zx(m) \equiv x(m+1)$$

where *zX* is read as *the next value of x*. If multiplying by z moved forward in time, then dividing by z moves backwards in time

$$z^{-1}x(m) \equiv x(m-1)$$

Likewise, the z-transform for a Bernoulli trial (termed the *moment generating function*) is

$$X(z) = q + z^{-1}p$$

or

$$X(z) = \left(\frac{qz+p}{z}\right)$$

depending upon how you want to write it.

Sidelight: In statistics, the moment generating function uses 's' where s = 1/z. In this class, we'll use z in stead since

- *s is the LaPlace operator, and*
- *z transforms are covered in ECE 343 Signals and Systems, ECE 461 Controls Systems, and ECE 376 Embedded Systems.*

This keeps the notation consistent between these classes.

The mean is the average. You can think about it as getting k dollars if you win. Your winnings from a coin toss are

$$\bar{x} = mean$$

$$\bar{x} = \Sigma\, k \cdot p(k)$$

$$\bar{x} = (0 \cdot q) + (1 \cdot p)$$

$$\bar{x} = p$$

If you get $1 if you win, you expect to make $p every time you flip the coin.

The variance and standard deviation is a measure of the variability

$$s^2 = variance$$

$$s = standard\ deviation$$

These are defined as

$$s^2 = \sum p(k) \cdot (k - \bar{x})^2$$

For a Bernoulli trial

$$s^2 = q(0-p)^2 + p(1-p)^2$$

$$s^2 = (1-p)p^2 + p(1-2p+p^2)$$

$$s^2 = p(1-p)$$

Bernoulli Trials in Matlab:

Matlab has random number functions:

```
rand          generate a random number in the interval of 0..1
rand(10,1)    generate a matrix of 10 random numbers over the range of 0..1
randn         generate a random number with a Gaussian distribution
```

For a Bernoulli trial, use the *rand* function (randn is covered later in this course).

As as example, flip a coin 5 times with the probability of a heads being 0.7.

First, generate 5 random numbers in the interval (0,1):

```
X = rand(5,1)

    0.8147
    0.9058
    0.1270
    0.9134
    0.6324
```

Next, convert this to a binary (1 or 0) number

```
Coin = 1 * (rand(5,1) < 0.7)

     1
     1
     1
     0
     0
```

Suppose you flip a coin 100 times. What is the mean and standard deviation of the result?

Solution: Flip a coin 100 times

```
Coin = 1 * (rand(100,1) < 0.7);
x = mean(Coin)

x =   0.5900

s = std(Coin)

s =   0.4943
```

Note that the mean should be 0.7 (p) and the standard deviation should be 0.4528

$$s = \sqrt{p(1-p)} = 0.4528$$

If you flip a coin 1 million times,

```
Coin = 1 * (rand(1e6,1) < 0.7);
x = mean(Coin)

x =    0.6997

s = std(Coin)

s =    0.4584
```

That is essentially the definition of probability: as the number of coin flips goes to infinity, the number of successes divide by the number of trials approaches p

$$\lim_{trials \to \infty} \left(\frac{\text{\# successes}}{\text{\# trials}} \right) = p$$

Sidelight: Probability is really only defined for repeatable events. Some events are not repeatable, such as the home opener for the Minnesota Vikings. Talking about the probability of such events doesn't make sense from a probability standpoint. Likewise, asking

What is the probability that the Vikings will win their home opener this year?

isn't technically a valid question. What is actually being asked is

What is the betting line on the Vikings winning their home opener?

The betting line is the odds necessary to balance the bets (money) for and against the event. If the money is balanced, the house gets its 5% cut and losers pay the winners: the house takes no risk of losing money.

Binomial Distribution

The binomial distribution is the sum of n Bernoulli trials. The probability of getting m heads with n trials is

$$X(m) = \binom{n}{m} p_m q^{n-m}$$

There are several ways to get this result.

Enumeration

Question: What is the probability of getting m heads in 2 trials?

Answer 1: Enumerate all possibilities and probabilities:

```
result        probability
 1 1            p * p
 1 0            p * q
0 1             q * p
0 0             q * q
```

or

m	X(m)
2	p^2
1	2pq
0	q^2

Problem: What is the probability of flipping m heads in 3 trials (n=3)

Enumerating all combinations (with 1 meaning a heads (success))

n	combinations	probability
3	111	p^3
2	110, 101, 011	$3p^2q$
1	001, 010, 001	$3pq^2$
0	000	q^3

Convolution

The probability density function for a Bernoulli trial is

$$x(m) = q \cdot \delta(m) + p \cdot \delta(m-1)$$

Another way to write this is

$$x(m) = \sum_k x(k) \cdot \delta(m-k)$$

This works since

$$\delta(k) = \begin{cases} 1 & k = 0 \\ 0 & otherwise \end{cases}$$

Likewise, the delta function pick out only those values of x(k) when k = m, giving

$$x(m) = x(m)$$

That seems kind of silly, but it helps when dealing with two coin tosses. If you are tossing two coins, the resulting pdf is

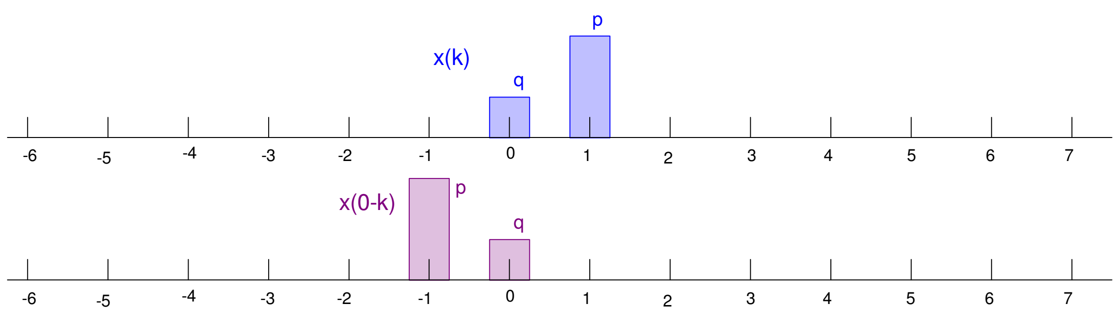
$$y(m) = \sum_k x(k)x(m-k)$$

or

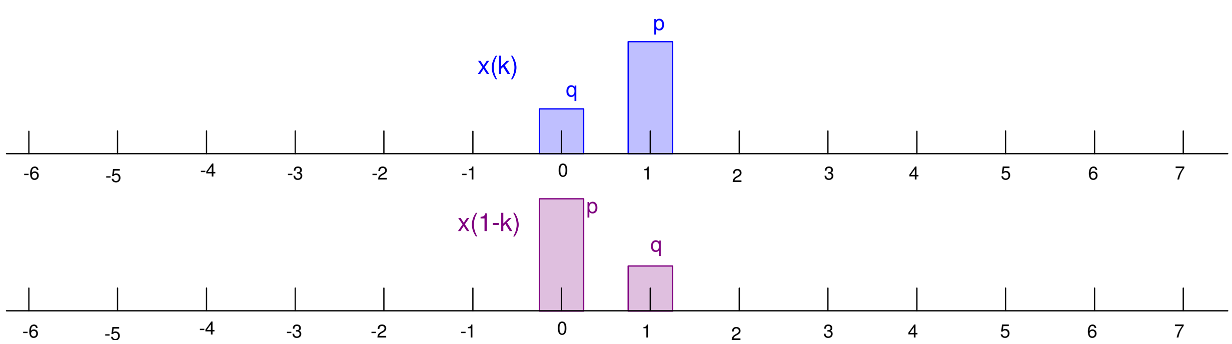
$$y(m) = x(m) ** x(m)$$

where ** denotes convolution. Graphically, what you are doing is

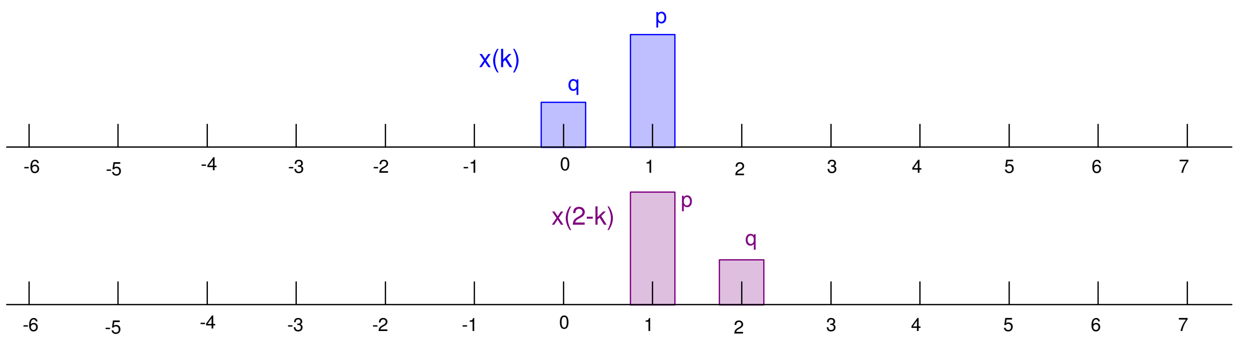
- Flip x(m) to get x(0-m)
- Multiply x(k) with x(-k) and sum the results. This gives $y(0) = q^2$



- Shift x(-k) to the right to get x(1-k). Multiply by x(k) and take the sum. This gives y(1) = 2pq

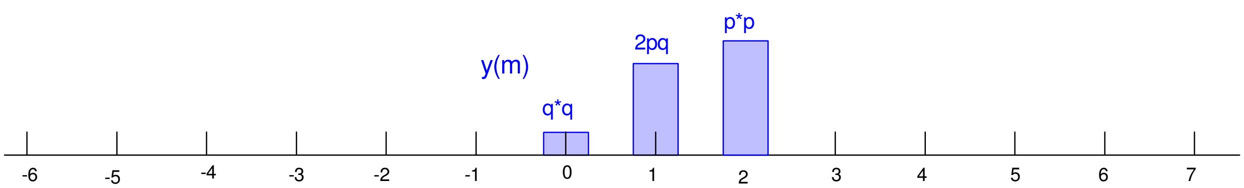


- Shift x(-k) to the right again to get x(2-k). Multiply by x(k) and take the sum. This gives $y(2) = p^2$



From this point onward, the result will be zero. This gives the pdf of two Bernoulli trials being

$$y(m) = q^2 \cdot \delta(m) + 2qp \cdot \delta(m-1) + p^2 \cdot \delta(m-2)$$



In Matlab, this is the convolution function

```
x = [0.3,0.7]

         x(0)        x(1)
x =     0.3000     0.7000

y = conv(x,x)

          y(0)       y(1)        y(2)
y =     0.0900     0.4200     0.4900
```

z-Transform

A property of the z-transform is

Convolution in the time domain (probability domain here) is multiplication in the z-domain

This is essentially because multiplying polynomials is convolution.

Example: Find the product of

$$C = AB = (2x^2 + 3x + 4)(5x^3 + 6x^2 + 7x + 8)$$

Solution using convolution: Flip B(x) and start shifting. The C(0) = 32

$$4 \cdot 8 = 32$$

B(k)	0	0	0	8	7	6	5	0	0	0	0
A(-k)	0	2	3	4	0	0	0	0	0	0	0
A*B	0	0	0	32	0	0	0	0	0	0	0

Shift A by one and multiply. C(1) = 24 + 28 = 52

$$8 \cdot 3x + 7x \cdot 4 = 52x$$

B(k)	0	0	0	8	7	6	5	0	0	0	0
A(1-k)	0	0	2	3	4	0	0	0	0	0	0
A*B	0	0	0	24	28	0	0	0	0	0	0

Shift A by one and multiply. C(2) = 16 + 21 + 24 = 61

$$8 \cdot 2x^2 + 7x \cdot 3x + 6x^2 \cdot 4 = 61x^2$$

B(k)	0	0	0	8	7	6	5	0	0	0	0
A(2-k)	0	0	0	2	3	4	0	0	0	0	0
A*B	0	0	0	16	21	24	0	0	0	0	0

Shift A by one and multiply. C(3) = 14 + 18 + 20 = 52

$$8 \cdot 0x^3 + 7x \cdot 2x^2 + 6x^2 \cdot 3x + 5x^3 \cdot 4 = 52x^3$$

B(k)	0	0	0	8	7	6	5	0	0	0	0
A(3-k)	0	0	0	0	2	3	4	0	0	0	0
A*B	0	0	0	0	14	18	20	0	0	0	0

etc. Again, this is the *conv* operation in Matlab. Inputting the matrices in increasing powers of x:

```
A = [4,3,2];
B = [8,7,6,5];
C = conv(A,B)

        x0    x1    x2    x3    x4    x5
C =     32    52    61    52    27    10
```

This also works if you use decreasing powers of x

```
A = [2,3,4];
B = [5,6,7,8]
C = conv(A,B)

        x5    x4    x3    x2    x1    x0
C =     10    27    52    61    52    32
```

The z-transform of two Bernoulli trials is thus

$$Y(z) = (q + z^{-1}p)(q + z^{-1}p)$$

$$Y(z) = q^2 + 2pqz^{-1} + p^2z^{-2}$$

Problem: Determine the probability distribution of flipping 6 coins.

Solution 1: Convolve 6 times

```
X = [0.3, 0.7]

X =    0.3000    0.7000

Y2 = conv(X,X)

Y2 =    0.0900    0.4200    0.4900

Y3 = conv(Y2,X)

Y3 =    0.0270    0.1890    0.4410    0.3430

Y4 = conv(Y3,X)

Y4 =    0.0081    0.0756    0.2646    0.4116    0.2401

Y5 = conv(Y4,X)

Y5 =    0.0024    0.0284    0.1323    0.3087    0.3601    0.1681

Y6 = conv(Y5,X)

Y6 =    0.0007    0.0102    0.0595    0.1852    0.3241    0.3025    0.1176
```

Solution 2: The probability of m heads in n tosses is

$$y(m) = \binom{n}{m} p^m q^{n-m}$$

For example, the probability of 4 heads is

$$y(4) = \binom{6}{4} (0.7)^4 (0.3)^2$$

$$y(4) = 0.324135$$

which matches the Matlab results

Pascal's Triangle

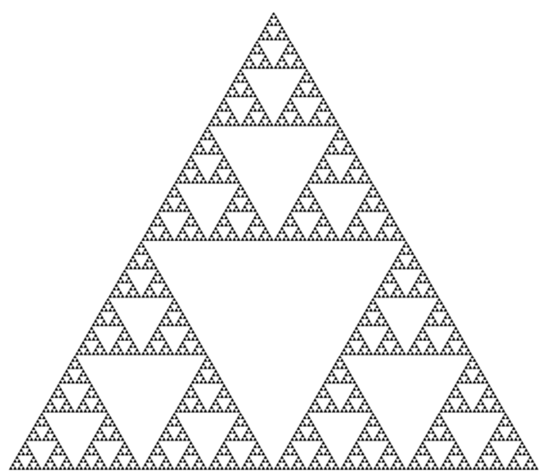
A kind of neat pattern for the combination $\binom{n}{m}$ is as follows:

- Start with the number 1 (or 0 1 0) in row #1
- Offset row #2 by 1/2 a digit. Add the numbers to the left and right of each spot in row #1 to generate row 2.
- Offset row #3 by 1/2 a digit. Add the numbers to the left and right of each spot in row #2 to generate row 3.

The result is as follows:

```
            1
          1   1
        1   2   1
      1   3   3   1
    1   4   6   4   1
  1   5   10  10  5   1
1   6   15  20  15  6   1
```

Each row is the value of $\binom{n}{m}$ as m goes from zero to n - which is one way of computing combinatorics. If you shade in the odd entries, you get a pretty picture as well: (in the limit becoming Sierpinski Triangle)



SierpinskiTriangle / Pascal's Triangle - source Wikipedia.com

Central Limit Theorem (take 1)

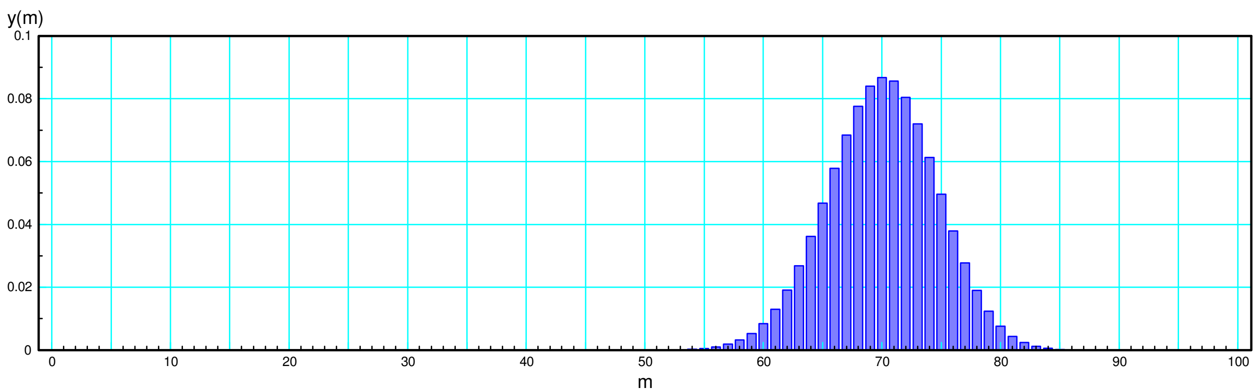
Determine the pdf of flipping 100 coins.

Solution (use Matlab for this):

$$y(m) = \binom{100}{m} p^m q^{100-m}$$

In Matlab

```
Y = zeros(101,1);
for m=0:100
   Y(m+1) = factorial(100) / (factorial(m) * factorial(100-m)) * 0.7^m  * 0.3^(100-m);
end
bar(Y)
```



Note that this is a bell-shaped curve, called a *Normal Distribution*. The Central Limit Theorem states that all distributions converge to this shape. It's probably the most important distribution in statistics - we'll be covering it later.

In addition, all probabilities add to one:

```
sum(Y)

ans =   1.0000
```

The mean of a binomial distribution is np (70)

```
sum(Y .* m)

ans =   70.0000
```

The variance is npq (21)

```
sum(Y .* (m - 70).^2)

ans =   21.0000
```

Binomial Distributions with Multiple Outcomes:

Problem: Two people are playing tennis. Person A has a 60% chance of winning a given point. The match is over when the first person wins 4 points (best of 7 series). What's the chance person A wins the match?

Solution: Assume all 7 games are played. Person A wins if he/she wins 4, 5, 6, or 7 games. (The match will actually end once you get to 4 wins, but this covers all combinations thereof.) The probability of player A winning is likewise the sum of probability of these four outcomes.

The probability of each case is:

4 wins: $f(4) = \binom{7}{4}(0.6)^4(0.4)^3 = 0.2903$

5 wins: $f(5) = \binom{7}{5}(0.6)^5(0.4)^2 = 0.2613$

6 wins $f(6) = \binom{7}{6}(0.6)^6(0.4)^1 = 0.1306$

7 wins: $f(7) = \binom{7}{7}(0.6)^7(0.4)^0 = 0.0280$

The total is 0.7102.

- Person A has a 71.02% chance of winning the match.
- The 'better' player will win the match 71.02% of the time.

An interesting question is how to set up a tournament so that the best player wins.

Problem: Two people are playing tennis. Person A has a 60% chance of winning a given point. The match is over when the first person wins by 4 points. What's the chance person A wins the match?

Solution: This is a *totally* different problem. If person A wins followed by a loss, you're back where you started. The net result is potentially an infinite series. This is a Markov chain (coming later this semester).

Hypergeometric Distribution:

Problem: Suppose a box contains A white balls and B black balls. Each trial, you take one ball out of the bin and then put it back into the bin. Find the probability distribution function for drawing n white balls.

Solution: This is sampling with replacement. Each trial has the same probability of success (drawing a white ball)

$$p = \frac{A}{A+B}$$

The probability of n white balls is then a binomial distribution

$$f(x|n,p) = \binom{n}{x} p^x q^{n-x}$$

Problem: Suppose you do *not* replace the ball after you select it. This changes the problem considerably.

First, you cannot draw more white balls than there are balls in the bin and you can't draw more white balls than the total number of balls you draw:

$$x \le \min(n, A)$$

You also cant draw less than zero balls:

$$x \ge \max(0, n-B)$$

The probability of drawing x balls in n draws is from the following:

- There are $\binom{A}{x}$ ways of drawing x white balls
- There are $\binom{B}{n-x}$ ways of drawing the remaining n-x black balls
- The total number of ways you can draw n balls is $\binom{A+B}{n}$

So, the pdf for a Hypergeometric distribution is:

$$\boxed{f(x|A,B,n) = \frac{\binom{A}{x}\binom{B}{n-x}}{\binom{A+B}{n}}}$$

Problem: A bin has 8 white balls and 10 black balls. You draw 5 balls without replacement. Find the probability that three of the balls are white:

$$f(3|8,10,6) = \frac{\binom{8}{3}\binom{10}{2}}{\binom{18}{5}} = \frac{(56)(45)}{8568} = 0.2941$$

Problem: Find the probability of drawing three or more white balls.

This is the sum of the probability of drawing 3, 4, or 5 white balls.

$$f(4|8,10,6) = \frac{\binom{8}{4}\binom{10}{1}}{\binom{18}{5}} = \frac{(70)(10)}{8568} = 0.0817$$

$$f(5|8,10,6) = \frac{\binom{8}{5}\binom{10}{0}}{\binom{18}{5}} = \frac{(56)(1)}{8568} = 0.0065$$

The sum of these three is 0.3823.