Create Discrete Distributions using Excel 2013

by
Milo Schield
Member: International Statistical Institute
US Rep: International Statistical Literacy Project
Director, W. M. Keck Statistical Literacy Project

Slides, output and data at: www.StatLit.org/
pdf/Excel2013-Create-Discrete-Distributions-Slides.pdf
pdf/Excel2013-Create-Discrete-Distributions-Demo.pdf
xls/Excel2013-Create Discrete-Distributions-Data.xlsx

Background & Goals Discrete Outcomes

Three cases:

#1 Binomial: Events are independent; chances are per try.

#2 Hypergeometric: Events are dependent: sample without replacement from a fixed finite population.

#3 Poisson: Events are independent and rare. Only know the average (expected) over a given time or space.

Activity: Enter formulas in Data.

Goal: Create the output on #6, #8 and #10. Upload.

Each Distribution has Certain Assumptions

The key assumption for all three is randomness:

The chance of success is constant in #1 (Binomial) and in #3 (Poisson) and is independent of the past.

The chance of success in #2 (Hyper-geometric) varies. It depends on the past. But given the past, the chance of success on the next try is determined.

Each Distribution can be used two ways

In each case, the data is presented in two ways:

PDF(k); the **probability distribution function:** the chance of getting EXACTLY k successes.

CDF(kmax): the **cumulative distribution function:** the chance of getting UP TO kmax successes: kmax or fewer.

For a discrete outcome, the CDF(kmax) is just the sum of the PDF(k) for *k* between zero and kmax.

#1: Binomial Distribution

The binomial distribution shows the probability of K successes in N tries given a chance of success P per try.

Chance: EXACTLY two heads in three flips of coin:

➢ =BINOM.DIST(2, 3, .5, 0) which returns 0.375.

1st argument is k (# successes); 2nd is N (# tries); 3rd is P (chance of success); 4th is 0 (exactly k successes).

The **cumulative probability for k ≤ 2** is given by:

➢ =BINOM.DIST(2, 3, .5, 1) which returns 0.875: the chance of UP TO two heads (of zero, one OR two heads) in three flips of a fair coin.

Enter Binom.Dist Functions: C7 & D7. Copy down.

	A	B	C	D	E	F	G
2	BINOMIAL DISTRIBUTION						
3	Fixed probability per trial						
4		8	N: # of independent tries				
5		0.5	P: Chance of success per try				
6		k	PDF	CDF	D7		
7		0	0.00	0.00	=BINOM.DIST(B7,B$4,B$5,1)		
8		1	0.03	0.04	C7		
9		2	0.11	0.14	=BINOM.DIST(B7,B$4,B$5,0)		
10		3	0.22	0.36			
11		4	0.27	0.64			

#2: Hyper-Geometric Distribution

This function gives the chance of *k* successes in *n* tries in sampling without replacement from a population of size N containing K successes.

The chance of getting three Aces in a hand of 13 cards from a deck of 52 cards containing four aces is:
=HYPGEOM.DIST(3, 13, 4, 52, 0) which is 0.041.

The cumulative probability for k ≤ 3 is given by:
=HYPGEOM.DIST(3, 13, 4, 52,1) which is 0.997.

First input is kmax; the 2nd is *n*, the 3rd is K, the 4th is N while the 5th is either zero (exact) or one (cumulative).

Enter Hypergeometric in J7 & K7. Copy down

	H	I	J	K	L	M	N
2	HYPERGEOMETRIC DISTRIBUTION						
3		52	Size of Small Population				
4		16	# Successes in Population				
5		13	Size of Sample (no replacement)				
6		k	PDF	CDF	K7		
7		0	0.00	0.00	=HYPGEOM.DIST(I7,I$5,I$4,I$3,1)		
8		1	0.03	0.04	J7		
9		2	0.11	0.15	=HYPGEOM.DIST(I7,I$5,I$4,I$3,0)		
10		3	0.22	0.37			
11		4	0.27	0.64			

#3: Poisson Distribution

Poisson: Gives the chance of EXACTLY k rare events during a given time or space if the expected # is known.

The chance of four deaths per year for kids ages 5-10 when the expected number is two is given by:

> =POISSON.DIST(2, 3, 0) It returns 0.15.

The cumulative probability for UP TO two per year:
> =POISSON.DIST(3, 3, 1) It returns 0.42.

The 1st number is the EXACT or MAXIMUM number that is observed, the 2nd number is what is EXPECTED , the 3rd number is either zero (exact probability) or one (cumulative probability).

C57, J57 Enter Poisson function; Copy down

	B	C	3	Expected # success
55	Chance of exactly k successes in time T.			
56	k	PDF		
57	0	0.05	=POISSON.DIST(B57,D$54,0)	
58	1	0.15		
59	2	0.22		

	I	J	
	Chance of kmax or fewer success in time T.		
56	kmax	CDF	
57	0	0.05	=POISSON.DIST(I57,D$54,1)
58	1	0.20	
59	2	0.42	

Process & Outcomes

1. Duplicate the output worksheet.
2. Upload your duplicate worksheet under XL4a.

Create Discrete Distributions using Excel 2013

by
Milo Schield
Member: International Statistical Institute
US Rep: International Statistical Literacy Project
Director, W. M. Keck Statistical Literacy Project

Slides, output and data at: www.StatLit.org/
pdf/Excel2013-Create-Discrete-Distributions-Slides.pdf
pdf/Excel2013-Create-Discrete-Distributions-Demo.pdf
xls/Excel2013-Create Discrete-Distributions-Data.xlsx

Background & Goals
Discrete Outcomes

Three cases:

#1 Binomial: Events are independent; chances are per try.

#2 Hypergeometric: Events are dependent: sample without replacement from a fixed finite population.

#3 Poisson: Events are independent and rare. Only know the average (expected) over a given time or space.

Activity: Enter formulas in Data.

Goal: Create the output on #6, #8 and #10. Upload.

Each Distribution has Certain Assumptions

The key assumption for all three is randomness:

The chance of success is constant in #1 (Binomial) and in #3 (Poisson) and is independent of the past.

The chance of success in #2 (Hyper-geometric) varies. It depends on the past. But given the past, the chance of success on the next try is determined.

Each Distribution can be used two ways

In each case, the data is presented in two ways:

PDF(k); the **probability distribution function:** the chance of getting EXACTLY k successes.

CDF(kmax): the **cumulative distribution function:** the chance of getting UP TO kmax successes: kmax or fewer.

For a discrete outcome, the CDF(kmax) is just the sum of the PDF(k) for *k* between zero and kmax.

#1: Binomial Distribution

The binomial distribution shows the probability of K successes in N tries given a chance of success P per try.

Chance: EXACTLY two heads in three flips of coin:

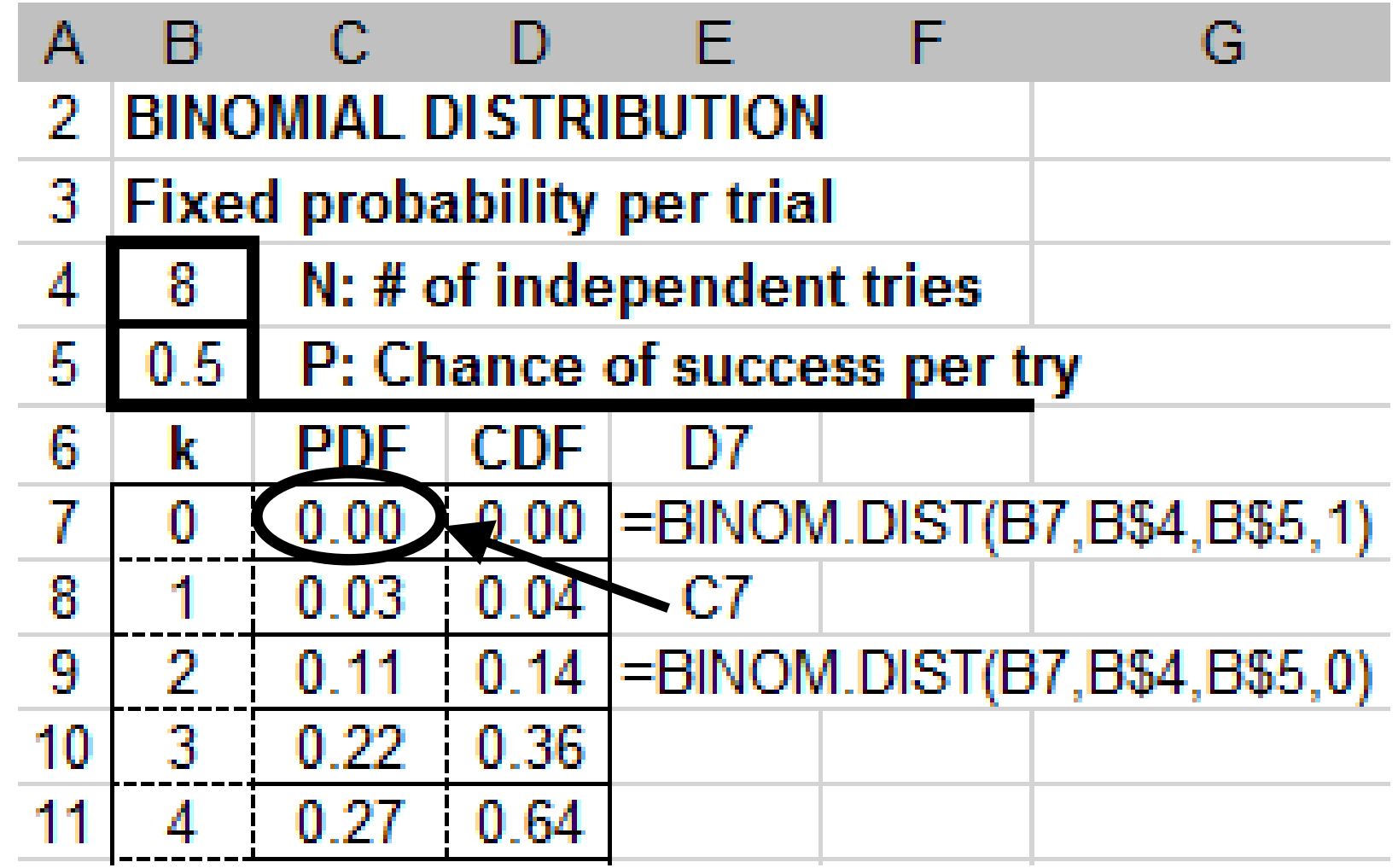
- =BINOM.DIST(2, 3, .5, 0) which returns 0.375.

1st argument is k (# successes); 2nd is N (# tries); 3rd is P (chance of success); 4th is 0 (exactly k successes).

The **cumulative probability for k ≤ 2** is given by:

- =BINOM.DIST(2, 3, .5, 1) which returns 0.875: the chance of UP TO two heads (of zero, one OR two heads) in three flips of a fair coin.

Enter Binom.Dist Functions: C7 & D7. Copy down.



A	B	C	D	E	F	G
2	BINOMIAL DISTRIBUTION					
3	Fixed probability per trial					
4	8	N: # of independent tries				
5	0.5	P: Chance of success per try				
6	k	PDF	CDF	D7		
7	0	0.00	0.00	=BINOM.DIST(B7,B$4,B$5,1)		
8	1	0.03	0.04	C7		
9	2	0.11	0.14	=BINOM.DIST(B7,B$4,B$5,0)		
10	3	0.22	0.36			
11	4	0.27	0.64			

#2: Hyper-Geometric Distribution

This function gives the chance of *k* successes in *n* tries in sampling without replacement from a population of size N containing K successes.

The chance of getting three Aces in a hand of 13 cards from a deck of 52 cards containing four aces is:

=HYPGEOM.DIST(3, 13, 4, 52, 0) which is 0.041.

The cumulative probability for $k \leq 3$ is given by:

=HYPGEOM.DIST(3, 13, 4, 52,1) which is 0.997.

First input is kmax; the 2nd is *n*, the 3rd is K, the 4th is N while the 5th is either zero (exact) or one (cumulative).

Enter Hypergeometric in J7 & K7. Copy down

	I	J	K	L	M	N
2	HYPERGEOMETRIC DISTRIBUTION					
3	52	Size of Small Population				
4	16	# Successes in Population				
5	13	Size of Sample (no replacement)				
6	k	PDF	CDF	K7		
7	0	0.00	0.00	=HYPGEOM.DIST(I7,I$5,I$4,I$3,1)		
8	1	0.03	0.04	J7		
9	2	0.11	0.15	=HYPGEOM.DIST(I7,I$5,I$4,I$3,0)		
10	3	0.22	0.37			
11	4	0.27	0.64			

#3: Poisson Distribution

Poisson: Gives the chance of EXACTLY k rare events during a given time or space if the expected # is known.

The chance of four deaths per year for kids ages 5-10 when the expected number is two is given by:

> =POISSON.DIST(2, 3, 0) It returns 0.15.

The cumulative probability for UP TO two per year:
> =POISSON.DIST(3, 3, 1) It returns 0.42.

The 1st number is the EXACT or MAXIMUM number that is observed, the 2nd number is what is EXPECTED , the 3rd number is either zero (exact probability) or one (cumulative probability).

C57, J57 Enter Poisson function; Copy down

	B	C	3	Expected # success
55	Chance of exactly k successes in time T.			
56	k	PDF		
57	0	0.05	=POISSON.DIST(B57,D$54,0)	
58	1	0.15		
59	2	0.22		

	I	J	
	Chance of kmax or fewer success in time T.		
56	kmax	CDF	
57	0	0.05	=POISSON.DIST(I57,D$54,1)
58	1	0.20	
59	2	0.42	

Process & Outcomes

1. Duplicate the output worksheet.
2. Upload your duplicate worksheet under XL4a.