

XL5A: 0H 2014 Schield Log-Normal-Income1 1

## Lognormal Distribution of Subjects by Income

---

**Milo Schield**  
 Augsburg College  
 Editor: [www.StatLit.org](http://www.StatLit.org)  
 US Rep: International Statistical Literacy Project

Output, slides and data at [www.StatLit.org/pdf/Excel2013-Schield-LogNormal-Income1-Demo.pdf](http://www.StatLit.org/pdf/Excel2013-Schield-LogNormal-Income1-Demo.pdf)  
[pdf/Excel2013-Schield-LogNormal-Income1-Slides.pdf](http://www.StatLit.org/pdf/Excel2013-Schield-LogNormal-Income1-Slides.pdf)  
[Excel/Excel2013-Schield-LogNormal-Income1-Data.xlsx](http://www.StatLit.org/Excel/Excel2013-Schield-LogNormal-Income1-Data.xlsx)

XL5A: 0H 2014 Schield Log-Normal-Income1 2

## Log-Normal Distributions

---

The log of a Normal distribution is not symmetric. It is never negative and it typically has a long right tail.

Some things are distributed log-normally:

- People by income, assets, weight and blood pressure
- Cities by population; insurance claims by amount

ASSIGNMENT:

- 1) Create the table shown on slide 4 (See demo output)
- 2) Create the graph shown on slide 5. Upload results.
- 3) Review the questions shown on slide 6.

XL5A: 0H 2014 Schield Log-Normal-Income1 3

## Enter data and formulas for top section

---

Enter 50 (median income) and 80 (average income) in B4 and B5.  
 Enter formulas in G2:G5. Enter formulas in B9 & C9. Verify values for all formula.

A	B	C	D	E	F	G	H	
<b>Real-world statistics</b>						$\mu$	3.912	=LN(B4)
Manual Entry						$\mu+S^2/2$	4.382	=LN(B5)
median	50				$\sigma^2$	0.940	=2*(G3-G2)	
mean	80				$\sigma$	0.970	=SQRT(G4)	
						<b>Underlying math statistics</b>		
B9 =Exp(G2-G4)								
	Mode	PDF	C9					
	19.531	1.32E-02	=LOGNORM.DIST(B9,G2,G5,0)					

Sometimes ##### appears in a cell. Either the font is too large, or Excel is trying to show the number with at least 10 digits after decimal. Either way, select the cell. Right-mouse and select Format Cells. Change font size or digits after decimal as needed.

XL5A: 0H 2014 Schield Log-Normal-Income1 4

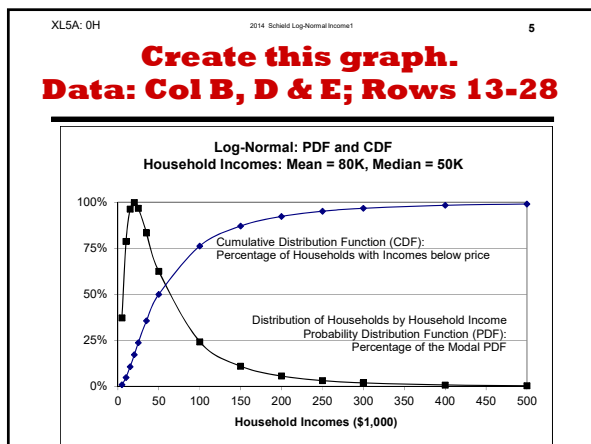
## Enter formula in B14:E14 Pull B14:E14 down to Row 28

---

B	C	D	E	F	G	H	I
B14	C14	D14	E14	C14			
=A14*BS4	Copy F12	=C14/C9	Copy F14	=LOGNORM.DIST(B14,\$G\$2,\$G\$5,0)			
<b>Income</b>	<b>PDF</b>	<b>% of mode</b>	<b>CDF</b>	E14			
5	4.90E-03	37.2%	0.9%	=LOGNORM.DIST(B14,G\$2,G\$5,1)			

**Table: Distribution of subjects by income (\$1,000)**

PDF: Probability Density Function  
 CDF: Cumulative Distribution Function  
 CDF: Percentage of subjects who have income below Col B.



XL5A: 0H 2014 Schield Log-Normal-Income1 6

## CDF Percentages; Practice Problems a and b

---

If X is income, then CDF(X) is the percentage of subjects who have LESS than X thousand dollars in income [Cumulative Distribution Function]

If median = \$50K and mean = \$80K, what percentage of subjects have incomes ...

- a. below 10K? 4.8%.  
 Solution: Find 10K in column B in row 15. Find matching CDF in same row, column E.
- b. ABOVE 10K? 100% - 4.8% = 95.2%

# **Lognormal Distribution of Subjects by Income**

---

**Milo Schield**

**Augsburg College**

**Editor: [www.StatLit.org](http://www.StatLit.org)**

**US Rep: International Statistical Literacy Project**

**Output, slides and data at [www.StatLit.org/  
pdf/Excel2013-Schild-LogNormal-Income1-Demo.pdf](http://www.StatLit.org/pdf/Excel2013-Schild-LogNormal-Income1-Demo.pdf)  
[pdf/Excel2013-Schild-LogNormal-Income1-Slides.pdf](http://www.StatLit.org/pdf/Excel2013-Schild-LogNormal-Income1-Slides.pdf)  
[Excel/Excel2013-Schild-LogNormal-Income1-Data.xlsx](http://www.StatLit.org/Excel/Excel2013-Schild-LogNormal-Income1-Data.xlsx)**

# Log-Normal Distributions

---

The log of a Normal distribution is not symmetric. It is never negative and it typically has a long right tail.

Some things are distributed log-normally:

- People by income, assets, weight and blood pressure
- Cities by population; insurance claims by amount

## ASSIGNMENT:

- 1) Create the table shown on slide 4 (See demo output)
- 2) Create the graph shown on slide 5. Upload results.
- 3) Review the questions shown on slide 6.

# Enter data and formulas for top section

Enter 50 (median income) and 80 (average income) in B4 and B5.

Enter formulas in G2:G5. Enter formulas in B9 & C9. Verify values for all formula.

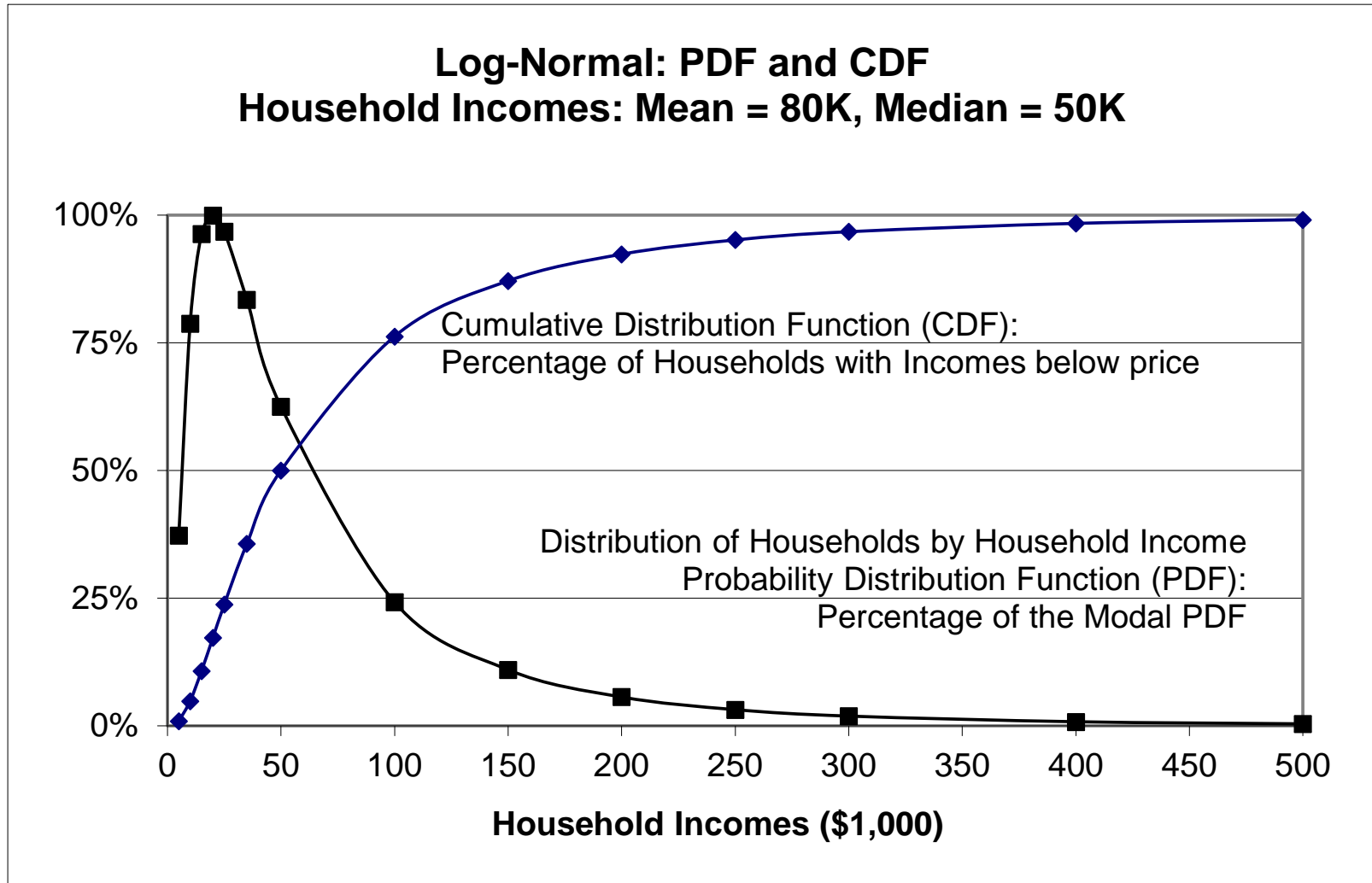
A	B	C	D	E	F	G	H	
	<b>Real-world statistics</b>				mu	3.912	=LN(B4)	
	Manual Entry				mu+S^2/2	4.382	=LN(B5)	
median	<b>50</b>				Sigma^2	0.940	=2*(G3-G2)	
mean	<b>80</b>				Sigma	0.970	=SQRT(G4)	
					<b>Underlying math statistics</b>			
B9	=Exp(G2-G4)							
	Mode	PDF	C9					
	19.531	1.32E-02	=LOGNORM.DIST(B9,G\$2,G\$5,0)					
A	B	C	D	E	F	G	H	

Sometimes ##### appears in a cell. Either the font is too large, or Excel is trying to show the number with at least 10 digits after decimal. Either way, select the cell. Right-mouse and select Format Cells. Change font size or digits after decimal as needed.



# Create this graph.

## Data: Col B, D & E; Rows 13-28



# **CDF Percentages; Practice Problems a and b**

---

If  $X$  is income, then  $CDF(X)$  is the percentage of subjects who have LESS than  $X$  thousand dollars in income [Cumulative Distribution Function]

---

If median = \$50K and mean = \$80K,  
what percentage of subjects have incomes ...

a. below 10K? 4.8%.

Solution: Find 10K in column B in row 15.

Find matching CDF in same row, column E,

b. ABOVE 10K?  $100\% - 4.8\% = 95.2\%$