Slide 1

V0I   2014 Schield Log-Normal Income2 Excel2013   1

Lognormal Distribution of Subjects by Income

Milo Schield

Augsburg College
Editor: www.StatLit.org
US Rep: International Statistical Literacy Project

www.StatLit.org/
pdf/2014-Schield-LogNormal-Income2-Excel2013-Slides.pdf
Excel/2014-Schield-LogNormal-Income2-Excel2013-Data.xlsx

Slide 2

V0I   2014 Schield Log-Normal Income2 Excel2013   2

Log-Normal Distributions

The lognormal is positive, has a single peak and is right-skewed so mode < median < mean.

Examples of **things** distributed log-normally:
- Incomes, assets, size of cities, size of insurance claims
- Weight and blood pressure of humans (by gender)

Money is different, we …
- don't care how total weight is distributed by person weight.
- do care how total income is distributed by HH income. This worksheet distributes both: **things** and **money**.

Note: HH stands for Household.
A Log-Normal distribution is simply the log of the normal.

Slide 3

V0I   2014 Schield Log-Normal Income2 Excel2013   3

Assignment
No graphs!

Enter formulas. Check results against demo.pdf

Shortcut: Rather than type long formulas, do this:
- Copy the formula to the clipboard, and paste it into the desired box.
- Edit the new formula, remove the leading single quote, and press enter. This eliminates errors!

Review definitions of various column headings.
Study the answers to practice questions!!!
You made need to answer this kind of question later.

Slide 4

V0I   2014 Schield Log-Normal Income2 Excel2013   4

Enter data and formulas in the top section of page 1

Enter 50 (median income) and 80 (mean income) in C4 & C5.
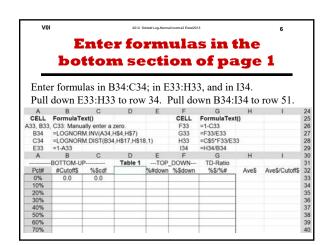Enter formulas in H4:H7.   Enter formulas in C7:C11.
Check values against demo output.

C	D	E	F	G	H	I	
							1
have a Log-Normal Distribution by Income.							2
ds Log-Normal Distribution				mu		=LN(C4)	3
50				mu		=LN(C4)	4
80				mu+S^2/2		=LN(C5)	5
				Sigma^2		=2*(H5-H4)	6
	=EXP(H4-H6)			Sigma		=SQRT(H6)	7
	=LOGNORM.DIST(C7,H4,H7,0)						8
	=SQRT((EXP(H6)-1)*EXP(2*H4+H6))						9
	=NORM.S.DIST(SQRT(LN(C5/C4)/2), 1)						10
	=2*NORM.S.DIST(H7/SQRT(2),1)-1						11

Note: "Households Normal Distribution" heading appears over columns G–I at row 3.

Slide 5

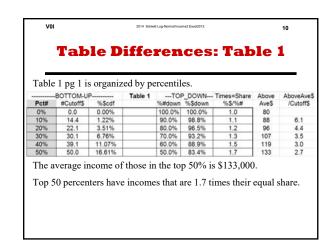V0I   2014 Schield Log-Normal Income2 Excel2013   5

Enter formulas in the middle section of page 1

Enter formulas in H17:H20, C17:C18, and in C20:C23.
Check results against demo output.

C	D	E	F	G	H	I	
							12
e a Log-Normal distribution by Income with mu# and sigma#, it follows that							13
e has a Log-Normal Distribution by HH Income [Aitchinson & Brown (1963)]							14
eters mu$ = (mu# + sigma#^2) and sigma$ = sigma#							15
me Log-Normal Distribution				Total Income Normal Distribution			16
128	=EXP(H17)			mu$	4.852	=H4+H6	17
205	=EXP(H20)			Sigma$	0.970	=H7	18
				Sigma$^2	0.940	=H18^2	19
50.0	=EXP(H17-H19)			mu$+S$^2/2	5.322	=H17+H19/2	20
5.14E-03	=LOGNORM.DIST(C20,H17,H18,0)						21
255.8	=C18*SQRT(((C18/C17)^2)-1)						22
0.686	=1-NORM.DIST(LN(C5),H17,H18,1)						23
C	D	E	F	G	H	I	24

Slide 6

V0I   2014 Schield Log-Normal Income2 Excel2013   6

Enter formulas in the bottom section of page 1

Enter formulas in B34:C34; in E33:H33, and in I34.
Pull down E33:H33 to row 34. Pull down B34:I34 to row 51.

A	B	C	D	E	F	G	H	I	
									24
CELL	FormulaText()				CELL	FormulaText()			25
A33, B33, C33: Manually enter a zero.					F33	=1-C33			26
B34	=LOGNORM.INV(A34,H$4,H$7)				G33	=F33/E33			27
C34	=LOGNORM.DIST(B34,H$17,H$18,1)				H33	=C$5*F33/E33			28
E33	=1-A33				I34	=H34/B34			29
A	B	C	D	E	F	G	H	I	30
	---------BOTTOM-UP---------			Table 1	---TOP-DOWN---		TD-Ratio		31
Pct#	#Cutoff$	%$cdf		%#down	%$down	%$/%#	Ave$	Ave$/Cutoff$	32
0%	0.0	0.0							33
10%									34
20%									35
30%									36
40%									37
50%									38
60%									39
70%									40

V0I   2014 Schield Log-Normal Income2 Excel2013   **7**

Enter formulas in Row 62: Generate Table 2 on Page 2

Enter formulas in B62:D62; F62:I62. Pull B62:I62 down to row 145.

CELL	FormulaText()		CELL	FormulaText()	
A62	Manual entry (Already entered)		F62	=LOGNORM.DIST(A62,H$17,H$18,0)	53
B62	=LOGNORM.DIST(A62,H4,H7,0)		G62	=F62/C$21	54
C62	=B62/C$8		H62	=F62/C$8	55
D62	=LOGNORM.DIST(A62,H$4,H$7,1)		I62	=LOGNORM.DIST(A62,H$17,H$18,1)	56
					57
					58

A	B	C	D	E	F	G	H	I	
ble 2	Distribution of Subjects by Income				Distribution of Total Income by Amount				60
come	PDF#	%of#mode	CDF#		PDF$	%of $mode	%of#mode	CDF$	61
1									62
2									63
3									64
4									65
5									66
7									67
8									68
9									69

V0I   2014 Schield Log-Normal Income2 Excel2013   **8**

Assignment

1. Complete the worksheet so it matches the demo output.

2. Review the definitions for the various columns

3. Review the "Practice" questions and answers carefully. You may need to answer this kind of question later.

4. Upload your results.

V0I   2014 Schield Log-Normal Income2 Excel2013   **9**

Table Differences: Table 1

Table 1 pg 1 is organized by percentiles.

	----------BOTTOM-UP----------		Table 1	---TOP_DOWN---		Times=Share	Above	AboveAve$
Pct#	#Cutoff$	%$cdf		%#down	%$down	%$/%#	Ave$	/Cutoff$
0%	0.0	0.00%		100.0%	100.0%	1.0	80	
10%	14.4	1.22%		90.0%	98.8%	1.1	88	6.1
20%	22.1	3.51%		80.0%	96.5%	1.2	96	4.4
30%	30.1	6.76%		70.0%	93.2%	1.3	107	3.5
40%	39.1	11.07%		60.0%	88.9%	1.5	119	3.0
50%	50.0	16.61%		50.0%	83.4%	1.7	133	2.7

The bottom 40% of subjects have 11.07% of total income.

The minimum income required to be included in the top 60% of subjects is 39.1 or $39,100.

The top 60% of subjects have 88.9% of total income.
The top 60% of subjects have 1.5 times [as much as] their equal share

---

V0I   2014 Schield Log-Normal Income2 Excel2013   **10**

Table Differences: Table 1

Table 1 pg 1 is organized by percentiles.

	----------BOTTOM-UP----------		Table 1	---TOP_DOWN---		Times=Share	Above	AboveAve$
Pct#	#Cutoff$	%$cdf		%#down	%$down	%$/%#	Ave$	/Cutoff$
0%	0.0	0.00%		100.0%	100.0%	1.0	80	
10%	14.4	1.22%		90.0%	98.8%	1.1	88	6.1
20%	22.1	3.51%		80.0%	96.5%	1.2	96	4.4
30%	30.1	6.76%		70.0%	93.2%	1.3	107	3.5
40%	39.1	11.07%		60.0%	88.9%	1.5	119	3.0
50%	50.0	16.61%		50.0%	83.4%	1.7	133	2.7

The average income of those in the top 50% is $133,000.

Top 50 percenters have incomes that are 1.7 times their equal share.

V0I   2014 Schield Log-Normal Income2 Excel2013   **11**

Table Differences: Table 2

Table 2 pg 2 is organized by income. CDF# & CDF$ are cumulative.

Table 2	Distribution of Subjects by Income			Distribution of Total Income by Amount			
Income	PDF#	% of mode	CDF#	PDF$	% of $mode	%of#mode	CDF$
1	1.20E-04	0.91%	0.00%	1.50E-06	0.03%	0.01%	0.00%
2	8.31E-04	6.31%	0.05%	2.08E-05	0.40%	0.16%	0.00%
3	2.04E-03	15.46%	0.19%	7.63E-05	1.48%	0.58%	0.01%
4	3.46E-03	26.25%	0.46%	1.73E-04	3.36%	1.31%	0.02%
5	4.90E-03	37.25%	0.88%	3.07E-04	5.96%	2.33%	0.04%
7	7.52E-03	57.12%	2.13%	6.58E-04	12.79%	5.00%	0.14%

#1: The 2.13% of subjects that have incomes BELOW 7K have 0.14% of the total income.

The 97.87% of subjects that have incomes ABOVE 7K have 99.86% of the total income.

Note; #2 is obtained from #1 by subtraction #1 from 100%.

V0I   2014 Schield Log-Normal Income2 Excel2013   **12**

Conclusion

The log-normal is a very useful distribution. It and the exponential are the most common asymmetric (skewed) distributions.

The distribution of subjects by HH income is related to -- but different from the distribution of total income by HH income.
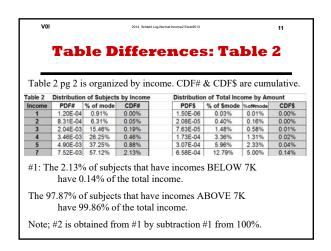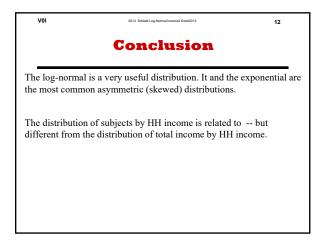

Lognormal Distribution of Subjects by Income

Milo Schield

Augsburg College
Editor: www.StatLit.org
US Rep: International Statistical Literacy Project

www.StatLit.org/
pdf/2014-Schield-LogNormal-Income2-Excel2013-Slides.pdf
Excel/2014-Schield-LogNormal-Income2-Excel2013-Data.xlsx

Log-Normal Distributions

The lognormal is positive, has a single peak and is right-skewed so mode < median < mean.

Examples of **things** distributed log-normally:
- Incomes, assets, size of cities, size of insurance claims
- Weight and blood pressure of humans (by gender)

Money is different, we …
- don't care how total weight is distributed by person weight.
- do care how total income is distributed by HH income.
  This worksheet distributes both: **things** and **money**.

Note: HH stands for Household.
A Log-Normal distribution is simply the log of the normal.

Assignment
No graphs!

Enter formulas. Check results against demo.pdf

Shortcut: Rather than type long formulas, do this:

- Copy the formula to the clipboard, and paste it into the desired box.

- Edit the new formula, remove the leading single quote, and press enter. This eliminates errors!

Review definitions of various column headings.
Study the answers to practice questions!!!
You made need to answer this kind of question later.

Enter data and formulas in the top section of page 1

Enter 50 (median income) and 80 (mean income) in C4 & C5.
Enter formulas in H4:H7.   Enter formulas in C7:C11.
Check values against demo output.

	C	D	E	F	G	H	I	
								1
	have a Log-Normal Distribution by Income.							2
	ds Log-Normal Distribution				Households Normal Distribution			3
	50				mu		=LN(C4)	4
	80				mu+S^2/2		=LN(C5)	5
					Sigma^2		=2*(H5-H4)	6
		=EXP(H4-H6)			Sigma		=SQRT(H6)	7
		=LOGNORM.DIST(C7,H4,H7,0)						8
		=SQRT((EXP(H6)-1)*EXP(2*H4+H6))						9
		=NORM.S.DIST(SQRT(LN(C5/C4)/2), 1)						10
		=2*NORM.S.DIST(H7/SQRT(2),1)-1						11

Enter formulas in the middle section of page 1

Enter formulas in H17:H20, C17:C18, and in C20:C23.
Check results against demo output.

C	D	E	F	G	H	I	
							12
e a Log-Normal distribution by Income with mu# and sigma#, it follows that							13
e has a Log-Normal Distribution by HH Income [Aitchinson & Brown (1963)]							14
eters mu$ = (mu# + sigma#^2) and sigma$ = sigma#							15
me Log-Normal Distribution				Total Income Normal Distribution			16
128	=EXP(H17)			mu$	4.852	=H4+H6	17
205	=EXP(H20)			Sigma$	0.970	=H7	18
				Sigma$^2	0.940	=H18^2	19
50.0	=EXP(H17-H19)			mu$+S$^2/2	5.322	=H17+H19/2	20
5.14E-03	=LOGNORM.DIST(C20,H17,H18,0)						21
255.8	=C18*SQRT(((C18/C17)^2)-1)						22
0.686	=1-NORM.DIST(LN(C5),H17,H18,1)						23
C	D	E	F	G	H	I	24

Enter formulas in the bottom section of page 1

Enter formulas in B34:C34; in E33:H33, and in I34.
Pull down E33:H33 to row 34. Pull down B34:I34 to row 51.

A	B	C	D	E	F	G	H	I	
CELL	**FormulaText()**				**CELL**	**FormulaText()**			25
A33, B33, C33: Manually enter a zero.					F33	=1-C33			26
B34	=LOGNORM.INV(A34,H$4,H$7)				G33	=F33/E33			27
C34	=LOGNORM.DIST(B34,H$17,H$18,1)				H33	=C$5*F33/E33			28
E33	=1-A33				I34	=H34/B34			29
A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	30
----------BOTTOM-UP----------			**Table 1**	---TOP_DOWN---		TD-Ratio			31
Pct#	#Cutoff$	%$cdf		%#down	%$down	%$/%#	Ave$	Ave$/Cutoff$	32
0%	0.0	0.0							33
10%									34
20%									35
30%									36
40%									37
50%									38
60%									39
70%									40

Enter formulas in Row 62: Generate Table 2 on Page 2

Enter formulas in B62:D62; F62:I62. Pull B62:I62 down to row 145.

CELL	FormulaText()				CELL	FormulaText()	
A62	Manual entry (Already entered)				F62	=LOGNORM.DIST(A62,H$17,H$18,0)	54
B62	=LOGNORM.DIST(A62,H4,H7,0)				G62	=F62/C$21	55
C62	=B62/C$8				H62	=F62/C$8	56
D62	=LOGNORM.DIST(A62,H$4,H$7,1)				I62	=LOGNORM.DIST(A62,H$17,H$18,1)	57

A	B	C	D	E	F	G	H	I	
ble 2	**Distribution of Subjects by Income**				**Distribution of Total Income by Amount**				60
come	PDF#	%of#mode	CDF#		PDF$	%of $mode	%of#mode	CDF$	61
1									62
2									63
3									64
4									65
5									66
7									67
8									68
9									69

Assignment

---

1. Complete the worksheet so it matches the demo output.

2. Review the definitions for the various columns

3. Review the "Practice" questions and answers carefully. You may need to answer this kind of question later.

4. Upload your results.

Table Differences: Table 1

Table 1 pg 1 is organized by percentiles.

| | ----------BOTTOM-UP---------- | | Table 1 | ---TOP_DOWN--- | | Times=Share | Above | AboveAve$ |
Pct#	#Cutoff$	%$cdf		%#down	%$down	%$/%#	Ave$	/Cutoff$
0%	0.0	0.00%		100.0%	100.0%	1.0	80	
10%	14.4	1.22%		90.0%	98.8%	1.1	88	6.1
20%	22.1	3.51%		80.0%	96.5%	1.2	96	4.4
30%	30.1	6.76%		70.0%	93.2%	1.3	107	3.5
40%	39.1	11.07%		60.0%	88.9%	1.5	119	3.0
50%	50.0	16.61%		50.0%	83.4%	1.7	133	2.7

The bottom 40% of subjects have 11.07% of total income.

The minimum income required to be included in the top 60% of subjects is 39.1 or $39,100.

The top 60% of subjects have 88.9% of total income.
The top 60% of subjects have 1.5 times [as much as] their equal share

Table Differences: Table 1

Table 1 pg 1 is organized by percentiles.

	----------BOTTOM-UP----------		**Table 1**	---TOP_DOWN---		Times=Share	Above	AboveAve$
Pct#	#Cutoff$	%$cdf		%#down	%$down	%$/%#	Ave$	/Cutoff$
0%	0.0	0.00%		100.0%	100.0%	1.0	80	
10%	14.4	1.22%		90.0%	98.8%	1.1	88	6.1
20%	22.1	3.51%		80.0%	96.5%	1.2	96	4.4
30%	30.1	6.76%		70.0%	93.2%	1.3	107	3.5
40%	39.1	11.07%		60.0%	88.9%	1.5	119	3.0
50%	50.0	16.61%		50.0%	83.4%	1.7	133	2.7

The average income of those in the top 50% is $133,000.

Top 50 percenters have incomes that are 1.7 times their equal share.

Table Differences: Table 2

Table 2 pg 2 is organized by income. CDF# & CDF$ are cumulative.

Table 2	Distribution of Subjects by Income				Distribution of Total Income by Amount			
Income	PDF#	% of mode	CDF#		PDF$	% of $mode	%of#mode	CDF$
1	1.20E-04	0.91%	0.00%		1.50E-06	0.03%	0.01%	0.00%
2	8.31E-04	6.31%	0.05%		2.08E-05	0.40%	0.16%	0.00%
3	2.04E-03	15.46%	0.19%		7.63E-05	1.48%	0.58%	0.01%
4	3.46E-03	26.25%	0.46%		1.73E-04	3.36%	1.31%	0.02%
5	4.90E-03	37.25%	0.88%		3.07E-04	5.96%	2.33%	0.04%
7	7.52E-03	57.12%	2.13%		6.58E-04	12.79%	5.00%	0.14%

#1: The 2.13% of subjects that have incomes BELOW 7K
have 0.14% of the total income.

The 97.87% of subjects that have incomes ABOVE 7K
have 99.86% of the total income.

Note; #2 is obtained from #1 by subtraction #1 from 100%.

Conclusion

The log-normal is a very useful distribution. It and the exponential are the most common asymmetric (skewed) distributions.

The distribution of subjects by HH income is related to -- but different from the distribution of total income by HH income.