EXAMPLES - SAMPLING DISTRIBUTION

EXCEL INSTRUCTIONS

This exercise illustrates the process of the sampling distribution as stated in the Central Limit Theorem. Enter the actual data in Column A in MICROSOFT EXCEL. The data represents a survey of the number of hours that the commuters spent in their cars each day. Use the EXCEL - Tools -Data Analysis. You may need to Add-in the Analysis Tool Pak every time you use the network. This project is to generate 100 observations of n = 5.

Step 1: Enter the actual data and labels

Step 2: Select **Sampling** from **Tools - Data Analysis**. Highlight the **Input Range** with only the data your Column A (do not include your labels or titles)

Type in 100 in the box corresponding to the **Random Number of Samples** Type in the **Output Range** B2:B102 Select **OK**

You will notice on your spreadsheet that EXCEL has generated a column of numbers randomly from your original set of data. If you compare your results in Column B with the person next to you, your sample values may not be the same as your neighbor.

- Step 3: Go to Column C. Use the function AVERAGE to calculate the mean values of n = 5. Column C is the column of the means.
- Step 4: Delete the selected cells in the Column C. Notice the cells in the AVERAGE function do not contain any data values.
- Step 5: Select Histogram from Tools Data Analysis. You will be making two different histograms. The first histogram is your actual data in Column A. So highlight your actual data in the Input Range. Type in Histogram for Actual Data under New Worksheet.

The second histogram is your distribution of the means in Column C. Highlight the data in the **Input Range**. Type in Histogram for Sampling under **New Worksheet**.

Step 6: Use the Chart Wizard to plot out these two histograms.

Step 7: Save your file

🗙 Microsoft Excel - Book2																								
8] <u>F</u> ile	<u>E</u> dit <u>V</u> i	ew <u>I</u> ns	sert	F <u>o</u> rma	it <u>T</u> o	ools <u>D</u> a	ata <u>\</u>	<u>M</u> indow	/ <u>H</u> elp)											_ 6	۶×	1
) 🖻		5 🗟	ABC.	*	B	2 🗸	1) - 0	1 - I	a (8	Σ	f*	<u></u> ≹↓	Zļ	h 🧕		100%	- 🧟				
Ar	ial			- 11) v	B	7 1	T I	= =	=		\$	%		+ .0	.00		. 0		Α.				-
	··	2	T		_	-		2 =		_		Ψ	/0	,	.00	+. 0								off
		Δ					(•		n		F			F		G		н	1			=	İce
1	ACT		ATA	SA	AMPL	ING	SAM	- PLIN	G DIS	TRIB	UTIC)N	n = 2	5								,	-	
2	1.01	1.00							Ĩ					-		-								
3		0.50							•		-		-			-								Ď
4		2.00																						
5		0.50		-																				빌
6		1.00																						
7		1.00																						
8		1.20																						
9		1.40																						6
10		2.10														_							- 11	
11		3.00																					- 11	<u></u>
12	_	3.20																_					- 11	W
13	-	1.20														_		_					- 11	$\overline{\mathbb{R}}$
14	-	1.50		_							_					_		_					- 11	<u> </u>
15	-	4.10														_							-11	
16	-	4.50		_							_					_		_					- 11	٩,
17	-	1.00	1	_							-					_		_					-11	
10	-	1.60									-					-		_					- 11	X
20	-										-												- 11	16R8
20	-			-									-			-		_						
22				-												-								3
		I) Shee	at 1 / C	haar	2 / c⊨	0012	1		-							4							J Č	ic.
			- A - D	neetz oSbar		\ \	, L		<u>الم</u>	1 2	. .	8	. A		= =		≓∎					,		SO
Dei																								
Rea	асу			1					- 1-	_		1											<u> </u>	
	Start	📃 sya	s on	19	Entire	• N	۸*٦.	letWa	are 🤅	Use	rs	3	ja Ne	etwor	'k	\	Quick		Micros.	🛛 🚺 Mi	icrosoft	∢:♥	3:25	PM

Step 1. Type in the data in the ACTUAL DATA column

Step 2: Select **Sampling** from **Tools - Data Analysis**. Highlight the **Input Range** with only the data your Column A (do not include your labels or titles)

Type in 100 in the box corresponding to the **Random Number of Samples**

Type in the Output Range B2:B108

Select OK

You will notice on your spreadsheet that EXCEL has generated a column of numbers randomly from your original set of data. If you compare your results in Column B with the person next to you, your sample values may not be the same as your neighbor.

However, the data in the SAMPLING and SAMPLING DISTRIBUTION columns in your spreadsheet may not be identical to what is printed here.

X 1	licros	oft Excel	- Book	2												_ 8	× 🖪
1	<u>)</u> <u>F</u> ile	<u>E</u> dit <u>V</u> iew	Insert	F <u>o</u> rma	t <u>T</u> o	iols <u>D</u> a	ta <u>W</u>	indow <u>H</u> el	lp							_ 8	× 🕅
			B. V	1	1 2 (2 🖋	N	• CH +	🤹 😤	Σ	$f_{\mathcal{R}} \stackrel{A}{\underset{Z}{\downarrow}} \stackrel{Z}{\underset{A}{\downarrow}}$	l 🏨 🧶	100%	- 🧟			
II Ari	ial		 	10 -	в	 	r I 🖃			B %	+.0 .0	2 硅 柱		- A -			4
	02	>		=		~ 2			- H == 1	p 70	J ,00 ≁.	0 == ==	<u> </u>				- Off
		- A		B		C		D	F	-	F	G	Н				
1	ACTU	UAL DAT.	A S	AMPL	ING	SAMP	PLING		BUTION	. n =	5						<u> </u>
2	·	1.00															
3		0.50															
4		2.00															
5	[0.50															
6	[1.00	_														
7	ļ	1.00					Sa	mpling						? ×	1		
8	È	1.20					Ir	nput									<u>=</u>
10		1.40					I	nput Range	е:		\$A\$2:\$/	4\$18	<u>.</u>	ок	<u> </u>		R C
11	-	3.00					- Ur	Labels						Iancel			
12		3.20					- 14	<u>E</u> apois									
13		1.20					S	ampling Me	thod					Help			<u>189</u>
14	1	1.50					- (🔿 P <u>e</u> riodic									X
15		4.10						Period:									•
16	E	4.50					6	Pandom									
17	[1.00						Number	of Compl	~~.	100						
18	Ļ	1.60					_	Number	u sampi	es;	1100						
19							0	utput optic	ns								凝
20							- (Output F	Rande:		\$B\$2:\$E	\$108	3		L		<u> 8</u>
21							- 6	New Wo	rkchoot (olo-			=				1
22		ikas ir	1 =1		1.5	,	=		noncet j	-77.	1					B	10
		\Sheet1	∫ Shee	t2 <u>/</u> SH	eet3	/	_	_ New <u>w</u> o	TKDOOK						<u> </u>		- TOS
Dr	aw 👻	66	A <u>u</u> toSha	apes 👻	1		4				-	· · · · · · ·	_		J		of
Cal	Calculating Sampling																
	Start	📄 sys or	n 5	Entire	N	N ^{لا} مر	etWare	e 🔄 Us	ers	🛃 N	etwork 🔊	i Quick	Kicro	s 🕎 Mici	rosoft	🍕 😥 🛛 3	:29 PM

Step 3: Go to Column C. Use the function AVERAGE to calculate the mean values of n = 5. Column C is the column of the means.

X	🗙 Microsoft Excel - Book2 📃 🖉 🗙 🖪									
125] <u>File E</u> dit <u>V</u> iew Inse	ert F <u>o</u> rmat <u>T</u> o	ools <u>D</u> ata <u>W</u> i	indow <u>H</u> elp						_ 뭔 ㅗ 🕅
Î) 🚅 🔲 🖨 🖪 :	ا 🛍 🎸	🏦 🝼 🗠	+ Ci + 🙆	😨 Σ	f≈ <u>A</u>	KI 🛍 🧶	🚜 100% 🔹	2	
به ا	rial .	• 10 • B	 7 Π ≡		¢ %	+,0	.00 (#] = #]	- A -	A -	
		0)			Ψ %	9 .00	•.0	<u> </u>	•• ·	S S
	<u> </u>	= = AV		. 66)	F			11		
4						г г	6	н		J
1		SAMPLING			ON, n =	5				
2	1.00	4.50	2.04							
5	0.00	3.00	2.00							
4	2.00	4.50	2.20							É
с а	1.00	1.00	2.00							
7	1.00	3.20	2.22				_			
8	1.00	1.40	2.02				_			
q	1.20	1.40	2.40							
10	2 10	4 10	2.32							<u>1</u> 2
11	3.00	3.20	2.02							
12	3.20	2.10	1.70							
13	1.20	1.00	1.48							
14	1.50	1.20	1.70							<u>×</u>
15	4.10	3.00	1.70							•
16	4.50	1.20	2.00							
17	1.00	1.00	1.86							
18	1.60	2.10	1.90							
19		1.20	2.30							
20		4.50	2.26							
21		0.50	2.18							
22		1.20	2.28							
N.	🕩 🕨 Sheet1 / Sh	eet2 / Sheet3	/				•			S DE
D	raw 🗸 🔓 🙆 Autos	5hapes 👻 🔨		🗐 📣 🔌 ·	- 🚄 - 🛕	- = :	■ ☴ 🗖 (los
Re	ady				[NUM	
	Start 📃 sys on	🔮 Entire N	NetWare 4م	e 🔄 Users	🛛 🛛 🙀	uick 🍴	🔀 Micros	W Microsoft		🍕 🕅 3:46 PM

Step 4: Delete the selected cells in the Column C. Notice the cells in the AVERAGE function do not contain any data values.

🗙 Microsoft Excel - Book2																					
8	Eile	<u>E</u> dit	⊻iev	v <u>I</u> nse	ert F <u>o</u> rm	iat <u>T</u> i	ools <u>D</u> ata	a <u>W</u> in	dow <u>H</u> e	lp										_ 8]	×1 🕅
D	é	; 🔒	6	Q.	ABC X	Ē	e 🗸	кл +	C2 +	ا 🏩	*	Σ	f∗ A	↓Z↓		🤵 [100%	• 🧟			
Δri	al				· 10 ·	B	<u>л</u> п		= =		¢.	%	+	.0 .00		, ≓≡ ∣	- A	- A -			4
		101		ΨĪ	-	- 63		(B10	1·B105)		Ψ	/0	, .	JU +. U	, =;		ш · <u>~</u>	· ••• ·]			- 13
		101	۵	<u> </u>	<u>–</u> В				n. D100) D		F		F		9		н	1		1	티렸
83			<u> </u>		16	Π	2.12								- 0					,	
84					0.5	n	1.90														
85					1.2	0	2.22			_						_					- 🖻
86					4.1	0	2.88														
87					3.2	0	2.48														
88					0.5	0	2.48														
89					2.1	0	2.66														
90					4.5	0	2.84														9
91					2.1	0	2.36			_											02
92					3.2	U O	2.84			_											
93					1.4	0	2.30														
94					3.0	0	2.20			_											127
80					Z.1	n n	1.70			_											\sim
97					0.5	n	0.94			-											
98					1.2	0	1.05														
99					0.5	0	1.00									_					
100					2.0	0	1.25														
101					0.5	0	0.50														*
102																					
103																					
104										_											- Mic
) N∖S	heet	1 / Sh	ieet2 / S	iheet3	1														5
Dr	aw 🔻	ß	Ġ	A <u>u</u> to:	5hapes 👻	\mathbf{i}	\mathbf{X}		I 📣	🕭 -	4	- <mark>A</mark>	- E		≓ ⊑		0				Sol
Rea	idy																		NUM		
18	Star	t 🗀	sys o	in	🔮 Enti	re N	Nel 4م.	Ware	. 🔁 Us	ers	1	i Qu	uick		Micros	s	🕎 Microsof	it		∢ ∈ ♥ 3	:51 PM

Step 5: Select Histogram from Tools - Data Analysis. You will be making two different histograms. The first histogram is your actual data in Column A. So highlight your actual data in the Input Range. Type in Histogram for Actual Data under New Worksheet.

The second histogram is your distribution of the means in Column C. Highlight the data in the **Input Range**. Type in Histogram for Sampling under **New Worksheet**.

🗙 Microsoft Excel - Book2 📃 🗗 🗙 🗖										
125	Eile Edit View Ins	ert F <u>o</u> rmat <u>T</u> o	ools <u>D</u> ata <u>W</u>	<u>V</u> indow <u>H</u> elp	_8× 🕅					
Шг		🤊 👗 🗈 I	🔁 🛷 🖍	· · · · · · · · · · · · · · · · · · ·						
J A	rial	- 10 - B	1 1	= = = ፼ ₩ % , % % ₽ ₽ <u>□ • ♥</u> • ▲ •	9					
	C23 🔽	= =A\	/ERAGE(B2	23:B27)						
	A	В	C	DEFGHIJ	ă ă					
1	ACTUAL DATA	SAMPLING	SAMPLING	G DISTRIBUTION, n = 5						
2	1.00	4.50	2.84							
3	0.50	3.00	2.58							
4	2.00	4.50	2.26		J B					
5	0.50	1.00	1.60		의 종등					
6	1.00	1.20	2.22	Input OK						
H	1.00	3.20	2.62	Input Range: \$4\$2:\$4\$18						
8	1.20	1.40	2.40	Bin Range: Cancel	<u> </u>					
9	1.40	1.20	2.32	Help	1					
	2.10	4.10	2.32							
	1 3.00	3.20	2.10	-Output options						
12	2 3.20	2.10	1.70		W7					
1/	1.20	1.00	1.40							
14	4 10	3.00	1.70	New Worksheet Ply: Histogram for Actual Da						
16	4.10 4.50	1.20	2.00	O New Workbook						
17	1.00	1.00	1.86	Pareto (sorted bistogram)						
18	1.60	2.10	1.90							
19	3	1.20	2.30	Chart Output						
20)	4.50	2.26							
21		0.50	2.18							
22	2	1.20	2.28		- M					
IÎ I	↓ ▶ ▶ Sheet1 / Sh	eet2 / Sheet3	7		I FIE					
	Praw + 🗟 🍐 Auto:	Shapes 👻 🔪		≝ ┩ 🏼 • 🚣 • ≡ ☴ ☵ 🖬 🗃	loso					
Re	ady			NUM						
1	Start 📄 sys on	🔮 Entire N	NetWar الحم.	re 🔄 Users 🛛 🙀 Quick 🛛 🔀 Micros 💯 Microsoft	∢ € ♥ 4:01 PM					

Step 6: Use the Chart Wizard to plot out these two histograms. The histogram for the Actual

Data is shown here.

🔀 Microsoft Excel - Book2 📃 🖪 💌										
🖹 Eile Edit View Insert Format Iools Data Window Help										
🛛 🖆 🖬 🍜 🖪 🖤 👗 🖻 🖻 🚿	ютат 🍓 🏶 Σ.	🆟 👌 🏭 🛍 🖤 🚜 100%	- 🛛 🔤							
Arial • 10 • B Z I		. *# # 佳 信 • 🕭	· A ·							
	D F F	G H I	К Н							
1 Bin Frequency										
2 0.5 2										
3 1.5 8										
4 2.5 3										
5 3.5 2										
6 More 2										
7	Chart Wizard - Step 3 of 4 - C	hart Options	? ×							
8										
9	Titles Axes Gridlines	Legend Data Labels Data Tal								
11	Chart <u>ti</u> tle:	Bar Chart for An	mal Data							
12	Bar Chart for Actual Data	Bai Chait roi Ac								
13	<u>Category (X) axis:</u>									
14	Number of Commuting Hours	, ÷ ;								
15	Value (Y) axis:									
16	Number of Drivers		E Frequency							
17	Second category (X) axis:									
18										
19										
20	Second value (Y) axis:	Number of Computing	.5 More							
21	1									
22			₹							
Histogram for Actual Data / St	2	Cancel < Back Ne>	t > Finish							
📗 Draw 👻 😓 🍪 🛛 AutoShapes 👻 📐 🔌 🗖										
Ready Sum=25 NUM NUM										
😭 Start 📄 sys on 🛛 👰 Entire N 💭 NetWare 🔄 Users 🛛 🏹 Quick 🛛 🔀 Micros 🕎 Microsoft										

EXAMPLES - SAMPLING DISTRIBUTION (CONT.)

Step 6: Use the Chart Wizard to plot out these two histograms. The histogram for the Sampling Distribution is shown here. Your sampling distribution will be different from the chart below.



Income Analysis Project

Based on past experience, the annual pay for a person with a Bachelor degree in one Business school with less than three year of experience from the sample taken by a recruiting officer is as follows:

78000, 65000, 58500, 67000, 51250, 54000, 69500, 58000, 72000, 76850, 73500, 79000, 72000, 68000, 61000, 66000, 64750, 61500, 75500, 64000.

Use EXCEL to generate 250 random numbers from this sample. Form the distribution of means with size = 10.

- (i) Print the entire sheet with the actual data and the numerical data of the sampling distribution using the font size of 5. What is the mean value and the standard deviation.
- (ii) Print the histogram of the actual data and the column chart (the bar graph)
- (iii) Print the histogram of the sampling distribution and the column chart (the bar graph).
- (iv) With the mean value and standard deviation obtained in (i),
 apply the command NORMINV(probability,mean,standard_dev) to generate 20 normal distributed numbers with two decimal places. Based on the generated numbers, do (i) (ii) and (iii) accordingly.