	A	B	C	D	E	F	G	H	I	J	K	5/29/2013

Goal: Model the relationship between rho and sample size for statistical significance   V0b

Step 1   Solve for n_Min: Smallest value of n that is statistically significant given rho.   7/17/2014

Source: www.vassarstats.net/rho.html

Step 2   Model the results using simple formulas.   ROW

6

Given Rho	Vassar nMin	-------------- MODEL #1 --------------				---------------- Model #2 ----------------				7
		Rho = 2/sqr(n)	Error	n = 4/(Rho)^2	Error	Rho = 2/sqr(n-1)	Error	n = 1+ 4/(Rho)^2	Error	8 / 9
0.001	3,840,000	0.001	2.1%	4,000,000	4.2%	0.001	2.1%	4,000,001	4.2%	10
0.002	960,000	0.002	2.1%	1,000,000	4.2%	0.002	2.1%	1,000,001	4.2%	11
0.005	153,000	0.005	2.3%	160,000	4.6%	0.005	2.3%	160,001	4.6%	12
0.01	38,400	0.010	2.1%	40,000	4.2%	0.010	2.1%	40,001	4.2%	13
0.02	9,600	0.020	2.1%	10,000	4.2%	0.020	2.1%	10,001	4.2%	14
0.05	1,530	0.051	2.3%	1,600	4.6%	0.051	2.3%	1,601	4.6%	15
0.1	384	0.102	2.1%	400	4.2%	0.102	2.2%	401	4.4%	16
0.12	267	0.122	2.0%	278	4.1%	0.123	2.2%	279	4.5%	17
0.15	171	0.153	2.0%	178	4.1%	0.153	2.3%	179	4.7%	18
0.2	96	0.204	2.1%	100	4.2%	0.205	2.6%	101	**5.2%**	19
0.3	43	0.305	1.7%	44	2.3%	0.309	2.9%	45	4.7%	20
0.4	25	0.400	0.0%	25	0.0%	0.408	2.1%	26	4.0%	21
0.5	16	0.500	0.0%	16	0.0%	0.516	3.3%	17	**6.3%**	22
0.6	11	0.603	0.5%	11	0.0%	0.632	5.4%	12	**9.1%**	23
0.65	10	0.632	-2.7%	9	**-10.0%**	0.667	2.6%	10	0.0%	24
0.7	9	0.667	-4.8%	8	**-11.1%**	0.707	1.0%	9	0.0%	25
0.8	7	0.756	-5.5%	6	**-14.3%**	0.816	2.1%	7	0.0%	26
0.9	5	0.894	-0.6%	5	0.0%	1.000	11.1%	6	**20.0%**	27

Given   Manual   =2/SQRT(B27)   =ROUND(4/A27^2, 0)

3 signif.digits   =(D27-A27)/A27   =(F27-B27)/B27

ANALYSIS: MODEL STATISTICAL-SIGNIFICANCE CONDITION FOR BIVARIATE CORRELATION

Model #1 nMin =(2/Rho)^2   OK for Rho < 0.6 or n > 10

Over-estimates minimum n (by up to 5%) for rho < 0.3.

Accurate estimates of n for rho between 0.4 and 0.6.

Under-estimates n by one for rho > 0.65.

RhoMin = 2/sqr(n)   Over-estimates minimum rho by up to 2.5% for n > 10

Model #2 nMin -1 =(2/Rho)^2   Over-estimates minimum n for Rho < 0.8

Always over-estimates (up to 5%*) the number required for N > 5.

* When over 5%, N is over-estimated by just one.

RhoMin = 2/sqrt(n-1)   Over-estimates rho by up to 5% for n > 5

CONCLUSIONS:

Model #1: Simplest. Conservative (never under-estimates) for n > 10 or rho < 0.6

Model #2: Second-simplest. Conservative (always over-estimates) for n > 4 or rho < 0.9

RECOMMENDATIONS:

Statistical Literacy for non-Quants: Use model #1 in range for n > 10.

Statistical Literacy for Managers: Use Model #1 or #2 in associated ranges.

Traditional Statistics: Use Model #2 for n > 5