Finite Population Agresti-Coull based Binomial Confidence Intervals and Excel VBA Function

William V Harper, Otterbein University, Department of Mathematical Sciences, Towers Hall 139, Westerville, OH 43081 USA, wharper@otterbein.edu

Abstract

Agresti-Coull (1998) developed an improvement over the classical central limit theorem based confidence interval for a single binomial proportion; however, it is based on the assumption of an infinite population. The 2nd edition (2013) of API 1163 used in the oil/gas industry to compare in-line inspection tools to excavated corrosion anomalies introduced the Agresti-Coull approach to the pipeline industry. While the exact Clopper-Pearson (1934) binomial confidence interval approach avoids many of the issues of central limit theorem approximations (including Agresti-Coull) its coverage sometimes exceeds the stated confidence level. For example 95% confidence intervals may average greater than 95% coverage while the Agresti-Coull intervals more closely track to the desired coverage level. In the oil/gas industry small sample sizes are not uncommon and the calling population is often finite. This paper expands the Agresti-Coull method to accommodate finite populations. In addition an Excel VBA function has been developed that creates Agresti-Coull single population binomial confidence intervals as well as expansions beyond the one-tailed upper bound used in API 1163.

Key words: binomial proportion confidence interval, finite population, Agresti-Coull

1. Introduction

This paper discusses the following topics (not in this order) related to the one sample binomial proportion:

- Classical normal approximation confidence intervals
- Exact Clopper-Pearson confidence intervals
- Agresti-Coull normal approximation confidence intervals
- Agresti-Coull extension for finite populations
- American Petroleum Institute API 1163 use of Agresti-Coull
- Pipeline and Hazardous Materials Safety Administration notice of proposed rulemaking
- Excel VBA software for Clopper-Pearson and finite-population Agresti-Coull

2. Clopper-Pearson Binomial Confidence Intervals

Most introductory statistical texts cover a standard normal approximation to the one sample binomial proportion including the development of confidence intervals for the true unknown population proportion. With sufficient sample size and proportions not too close to 0 or 1, this central limit based approximation may be fairly accurate. More advanced text books (e.g., Hollander and Wolfe, 1999) cover exact binomial confidence intervals originally published by Clopper and Pearson (1934) as shown in Equation Set 1 below.

$$p_L^{\alpha/2}(n,x) = \frac{x}{x + (n - x + 1) f_{\alpha/2, 2(n-x+1), 2x}}$$

$$p_U^{\alpha/2}(n,x) = 1 - p_L^{\alpha/2}(n, n - x)$$

Equation Set 1. Exact binomial confidence intervals (Clopper and Pearson, 1934)

In Equation Set 1 x is the number of successes in the n Bernoulli trials and f_{γ,n_1,n_2} is the upper γ^{th} percentile of the F distribution with n_1 and n_2 degrees of freedom. The top equation is the lower bound and the second equation is the upper bound if developing a two-tailed confidence interval. This approach results in coverage of at least 100*(1-α)% and can be overly conservative. One-tailed confidence intervals put the full α into one of the appropriate tails.

3. Agresti-Coull Normal Approximation Binomial Confidence Intervals and Finite-Population Extension

For years I used only Clopper-Pearson exact binomial confidence intervals from Minitab or using Excel VBA functions (Harper, 2005) when developing confidence intervals for a single binomial proportion. In 2013, the American Petroleum Institute in the second edition of API 1163 introduced the Agresti-Coull normal approximation. From an API web page http://www.api.org/about, "The American Petroleum Institute (API) is the only national trade association that represents all aspects of America's oil and natural gas industry." API develops and maintains approximately 700 standards and recommended practices. API 1163 titled "In-line Inspection Systems Qualification" addresses how well an in-line inspection tool has performed. API 1163 has guidelines about qualifying processes based on in-line inspection. Appendix C of the second edition of API 1163 (2013) helps assess how well the in-line inspection depth measurements match corresponding field excavation depths. The first edition of API 1163 (2005) appendix provided similar information; however, it was based on the usual normal approximation to the binomial addressed in the next paragraph.

Equation Set 2 has the classical central limit theorem based one-sample two-tailed binomial normal approximation confidence intervals found in standard texts where x is the number of "successes", n is the sample size, $\hat{p}$ = x/n is the sample proportion of "successes", and $z_{(\alpha/2)}$ is the (100 - α/2) percentile of a standard normal z.

$$\hat{p} \pm z_{(\alpha/2)} \sqrt{\hat{p}(1-\hat{p})/n}$$

Equation Set 2. Classical normal approximation of the two-tailed confidence interval for the binomial population proportion p.

Agresti and Coull (1998) discuss the history of various approaches to better normal approximations for the one-sample proportion case. For this particular article we will focus on just the one-tailed upper bound that is used in API 1163 though the Excel VBA macro permits lower, upper, or two-tailed intervals using the Agresti-Coull method. At first glance the upper bound confidence value shown below appears the same as the standard normal approximation. However in place of the typical $\hat{p}$ and n are $\tilde{p}$ and $\tilde{n}$ defined below. In Equation Set 3, n is the sample size, x = number of successes, and $z_{(\alpha)}$

is the (100 - α)th percentile of a standard normal z along with the other variables defined in Equation Set 3. This results in the Agresti-Coull upper confidence bound $\hat{p}_{upper}$.

$$\hat{p}_{upper} = \tilde{p} + z_\alpha \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}} \text{ where}$$

$$\tilde{n} = n + z_\alpha^2, \tilde{p} = \frac{x + \frac{z_\alpha^2}{2}}{\tilde{n}}$$

Equation Set 3. Agresti-Coull upper one-tailed confidence interval in 2nd edition of API 1163

The Agresti-Coull method is a normal approximation and thus requires a reasonable sample size among other limitations. The Agresti-Coull approach provides confidence intervals that are often closer to the intended coverage levels than the exact Clopper-Pearson which can be conservative as mentioned earlier. Equation Set 4 below provides the finite population correction factor to the standard error in the Agresti-Coull infinite population normal approximation. Capital N is the population size. This replaces $\sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}$ shown in Equation Set 3. This finite population adjustment can be important to the pipeline industry. Setting the population size to a large number results in the infinite population Agresti-Coull confidence intervals for all practical purposes

$$\sqrt{\frac{\tilde{p}(1-\tilde{p}) * (N-n)/N}{\tilde{n}}}$$

Equation Set 4. Agresti-Coull standard error adaptation for finite populations

4. API 1163 Application of Agresti-Coull Normal Approximation

An Excel VBA function for the Agresti-Coull infinite population as used in appendix C of the 2nd edition of API 1163 was developed several years ago and has been used to compare in-line inspection (ILI) called depths to field excavation depths. The focus in API 1163 is on assessing if there is sufficient reason to declare at the 95% confidence level that the in-line inspection is not performing adequately in terms of anomaly depth. There are multiple issues such as it is common to ignore uncertainty in the field measured depths, but that is beyond the focus of this paper. In essence for each matched pair of ILI and excavation values, the in-line inspection is found to be within vendor specification (success) or not within specification (failure). Then x is the number of “successes” within specification of the n matched pairs.

Driven by the 2016 Pipeline and Hazardous Materials Safety Administration (PHMSA) notice of proposed rulemaking (NPRM) the mathematical formulation converting the infinite population case to accommodate finite populations was completed and a revised Agresti-Coull finite population Excel VBA function was created. In addition the Agresti-Coull VBA function has been expanded from just the API 1163 illustrated one-tailed upper confidence bound to include a one-tailed lower confidence bound as well as a two-sided (two-tailed) confidence interval lower and upper bounds. One focus of PHMSA’s 2016 NPRM was a concern whether or not sufficient pipeline samples have been collected.

Before illustrating a potential application related to assessing the NPRM issue "have we sampled enough?" using both the exact Clopper-Pearson method and the new finite-population Agresti-Coull method, a brief illustration of the initial VBA function is provided in Table 1 using two examples in appendix C of the 2nd edition of API 1163 (2013). In-line inspection companies (sometimes called pigging vendors as the ILI tool is commonly called a pig) set specifications for how well a particular ILI should perform. There are many aspects to this and brand new technologies may have wider initial specifications than more established technology. All such successful matches may be thought of as a 1 from a binomial distribution perspective while those that do not meet the vendor specification are failures or a 0 for the binomial distribution.

A complete ILI to excavation comparison entails more than a mere binomial 0-1 analysis solely on depth; however, the binomial assessment is an important one as it is clearly delineated in API 1163. That does not mean all companies perform such an analysis; however, it is good practice. There are multiple concerns about such comparisons and one that is often overlooked is the quality of the field excavations that are treated as somewhat of a gold standard. In reality there is variability in any measurement and in field measurements this is definitely the case. More can be said about this though it is out of the scope of this paper. API 1163 does allow the incorporation of an estimate of field measurement error prior to assessment of the matched ILI and field measurements though to the best of my knowledge most pipeline companies do not quantify field measurement error nor do they fold such into their API 1163 based analysis.

In the first example the vendor specification is that 80% of the ILI depth calls will be within +/- 10% of the field measured depth. This 80% is seen in the penultimate column of Table 1 titled "Vendor target for proportion in spec". The +/- 10% is not shown in this table but must be used to determine if a given ILI tool call successfully matched the field measurement. For an ILI call of 20% depth, it is considered a success if the excavation depth is included in the interval [10%, 30%]. The "tolerance" of +/- 10% is common but may vary from one tool type to another. Additionally the 80% within the tolerance is vendor specific. For typical pipeline corrosion analysis +/- 10% for 80% of the calls is the current most common combination.

In Table 1 the "Ex 1 API 1163" has x = 5 of the n = 10 matched pairs within +/- 10% between the ILI called and field measured depth or a sample proportion of 5/10 = 0.500 as seen in the p-hat column. The Agresti-Coull binomial assessment in API 1163 is set at a 95% confidence level to develop a one-sided upper confidence value for the true unknown population proportion. Hence the "Conf level" column is set to 0.95 and the Option is set to UpperFullAlpha putting the entire α level of 0.05 in the upper tail. Following the earlier Equation set 3 formulas this results in the "API 1163 Upper Bound 1-tailed" result of 0.731 that is below the vendor specification of 0.80 and thus fails the API 1163 assessment. In the second example using the same vendor specifications and confidence level the 18 of 25 within specification ILI calls results in a sample proportion of 0.72. This results in a corresponding 95% upper confidence bound of 0.842 that exceeds the vendor specification of 0.80 and thus does not fail this API 1163 test.

Li et al (2017) suggest (without any statistical justification in their paper) that a sample size of $n \geq 20$ be required before declaring the API 1163 criteria to be passed. Enbridge, the company all the authors of Li et al (2017) work for has an aggressive excavation program. On some pipelines run by other companies the sample size is often less than 20 as pipelines tend to dig (excavate) only the portions of a given line that appears to have issues as excavations are expensive.

Segment	X, # in spec	n, # matched pairs	p-hat (proportion in spec)	Conf level	Option (using API 1163 choice)	API 1163 Upper Bound 1-tailed	Vendor target for proportion in spec	Passes API 1163?
Ex 1 API 1163	5	10	0.500	0.95	UpperFullAlpha	0.731	0.80	No
Ex 2 API 1163	18	25	0.720	0.95	UpperFullAlpha	0.842	0.80	Yes

Table 1. API 1163 Agresti-Coull examples comparing ILI to field excavations

The initial Excel VBA code associated with this comparison (UpperFullAlpha is the one for API 1163 as seen in Table 1) has been used for about 2-3 years. More recently additional alternatives (LowerAlpha/2, UpperAlpha/2, LowerFullAlpha) have been added to provide two-tailed confidence intervals as well as lower bound confidence intervals. Initially all four options were based on the calling population being infinite.

The Pipeline and Hazardous Materials Safety Administration (PHMSA) notice of proposed rulemaking (NPRM) in 2016 requires sample sizes of 150 – 350 for which no statistical rationale was provided. These samples sizes are based on the number of inconsistencies found in pipeline data. Each observation is binary, i.e., each new data value is either consistent or inconsistent with already collected data for a given pipeline segment. Not addressed but important is how consistency or inconsistency is assessed.

5. Pipeline and Hazardous Materials Safety Administration notice of proposed rulemaking

Uses of the Excel VBA software addressing the NPRM sample size concerns provides an upper consistency bound based on binomial confidence intervals. One worksheet uses the exact Clopper-Pearson methodology that is based on an infinite calling population and cannot be easily changed to address a finite population. A second Excel worksheet uses a finite population Agresti-Coull approach developed in this paper. Accounting for a finite population may reduce the number of needed samples and save pipeline operators time and expense. Care is needed when using the Agresti-Coull approach as it is still a central limit theorem based normal approximation to the binomial and thus will not be applicable for small sample sizes. Related to this to some extent is the Li et al (2017) recommended that the sample size be a minimum of 20. At this point in time, the Agresti-Coull worksheet starts with a sample size of ten.

The binomial work in this paper is one approach to calculate the number of measurements needed to verify that a material property such as pipe grade, or wall thickness, is consistent with existing measurements to a given statistical confidencel level on a given pipeline. Such a framework provides a basis to assess needed sample size as opposed to the NPRM sample sizes of at least 150 measurements. For the NPRM a lower bound confidence interval can be used as a reliability metric when comparing verifiability for prioritization between different pipe populations.

Table 2 summarizes the number of measurements needed to achieve a given confidence level and lower bound consistency for an infinite population using the exact Clopper-Pearson solution, assuming no inconsistencies (n - x = 0 inconsistencies or failures; thus x = n successes) have been found. If an operator wants to have 95% confidence that 95% of the pipe joints on a pipeline are the same grade, he must perform 59 digs, all of which must be consistent with the target grade. If sufficient data is available, the finite population approach covered later may require fewer samples.

Table 2. Required Number of Measurements as a Function of Confidence Level and Lower Bound Consistency with the Exact Clopper-Pearson Solution when the Sample has no inconsistencies (x = n, i.e., all successes)

		Consistency			
		80%	**90%**	**95%**	**99%**
Confidence Level	**80%**	8	16	32	160
	90%	11	22	45	228
	95%	14	29	59	297
	99%	21	44	90	456

6. Application of Excel VBA software for Clopper-Pearson and Finite Population Agresti-Coull

6.1 Clopper-Pearson Application

Table 3 based on a 95% confidence level provides an example to aid in the interpretation of lower confidence bounds for the binomial proportion using the Clopper-Pearson approach. There are numerous applications for such binomial confidence intervals though the wording that follows focused on PHMSA's NPRM concern with consistency of pipeline material property assessments. Consistency may be viewed as either a proportion in the range [0, 1] or a percentage in the range [0%, 100%]. The computed consistency is a function of:

- Desired confidence level (80%, 90% or 95% are suggested)
- # observations (labeled "# joints inspected"): row labels in the first column
- # inconsistencies: 0 – 9 shown here

Table 3. Lower-Bound Clopper-Pearson Consistency (95% Confidence)

	A	B	C	D	E	F	G	H	I	J	K
1	**Lower bound on reliability shown in matrix below**										
2		95%	confidence level *(change as desired)*								
3		0.80	Reliability level for conditional formatting *(change as desired)*								
4	References										
5	Hollander, Myles and Douglas A., Wolfe Nonparametric Statistical Methods,										
6	2nd Edition, ISBN 978-0471190455, 1999, John Wiley & Sons, New York										
7	Harper, William V., "Excel Functions to Compute Exact Binomial Confidence Intervals"										
8	Proceedings of the 25th European Meeting of Statisticians, July 2005, Oslo, Norway, Paper P-031										
9	*This method works for any binary assessment and is not limited to bends*										
10		# Inconsistencies									
11	# Joints inspected	0	1	2	3	4	5	6	7	8	9
12	1	0.050	0.000								
13	2	0.224	0.025	0.000							
14	3	0.368	0.135	0.017	0.000						
15	4	0.473	0.249	0.098	0.013	0.000					
16	5	0.549	0.343	0.189	0.076	0.010	0.000				
17	6	0.607	0.418	0.271	0.153	0.063	0.009	0.000			
18	7	0.652	0.479	0.341	0.225	0.129	0.053	0.007	0.000		
19	8	0.688	0.529	0.400	0.289	0.193	0.111	0.046	0.006	0.000	
20	9	0.717	0.571	0.450	0.345	0.251	0.169	0.098	0.041	0.006	0.000
21	10	0.741	0.606	0.493	0.393	0.304	0.222	0.150	0.087	0.037	0.005
22	11	0.762	0.636	0.530	0.436	0.350	0.271	0.200	0.135	0.079	0.033
23	12	0.779	0.661	0.562	0.473	0.391	0.315	0.245	0.181	0.123	0.072
24	13	0.794	0.684	0.590	0.505	0.427	0.355	0.287	0.224	0.166	0.113
25	14	0.807	0.703	0.615	0.534	0.460	0.390	0.325	0.264	0.206	0.153
26	15	0.819	0.721	0.637	0.560	0.489	0.423	0.360	0.300	0.244	0.191
27	16	0.829	0.736	0.656	0.583	0.516	0.452	0.391	0.333	0.279	0.227
28	17	0.838	0.750	0.674	0.604	0.539	0.478	0.420	0.364	0.311	0.260
29	18	0.847	0.762	0.690	0.623	0.561	0.502	0.446	0.392	0.341	0.291
30	19	0.854	0.774	0.704	0.641	0.581	0.524	0.470	0.418	0.368	0.320
31	20	0.861	0.784	0.717	0.656	0.599	0.544	0.492	0.442	0.394	0.347
32	21	0.867	0.793	0.729	0.671	0.616	0.563	0.513	0.464	0.417	0.372
33	22	0.873	0.802	0.741	0.684	0.631	0.580	0.532	0.485	0.439	0.395
34	23	0.878	0.810	0.751	0.696	0.645	0.596	0.549	0.504	0.460	0.417
35	24	0.883	0.817	0.760	0.708	0.658	0.611	0.565	0.521	0.479	0.437
36	25	0.887	0.824	0.769	0.718	0.670	0.625	0.580	0.538	0.496	0.456
37	26	0.891	0.830	0.777	0.728	0.682	0.637	0.595	0.553	0.513	0.474
38	27	0.895	0.836	0.785	0.737	0.692	0.649	0.608	0.568	0.529	0.491
39	28	0.899	0.841	0.792	0.746	0.702	0.661	0.620	0.581	0.543	0.506
40	29	0.902	0.847	0.798	0.754	0.712	0.671	0.632	0.594	0.557	0.521
41	30	0.905	0.851	0.805	0.761	0.720	0.681	0.643	0.606	0.570	0.535

As seen in Table 3, a lower bound consistency of 0.807 (80.7%) is obtained at the 95% confidence level when zero inconsistencies are found in 14 observations. Of course, for this example consistency could be as high as 100% since no inconsistencies have been found. If the operator wants a higher statistical consistency, more samples are required. For example, if the operator wants 90% of the population to be the same grade, 29 consistent samples are needed. What happens if the operator finds an inconsistent measurement? The number of required digs to achieve the same statistical consistency increases. If there is one inconsistency, Table shows that 22 samples are needed to conclude that 80% of the population is the same grade. As the number of inconsistencies increases, the lower confidence bound for consistency drops for a given number of samples.

Table 4 has similar lower-bound consistency values at the 99% confidence levels. Instead of 14 samples with no inconsistencies for an 80% lower bound as the 95% confidence level in Table 3, it now requires 21 samples to obtain the 80% lower bound with 99% confidence seen in Table 4.

Table 4. Lower-Bound Clopper-Pearson Consistency (99% Confidence)

Lower bound on reliability shown in matrix below

	99%	confidence level *(change as desired)*
	0.80	Reliability level for conditional formatting *(change as desired)*

References

Hollander, Myles and Douglas A., Wolfe Nonparametric Statistical Methods,
2nd Edition, ISBN 978-0471190455, 1999, John Wiley & Sons, New York

Harper, William V., "Excel Functions to Compute Exact Binomial Confidence Intervals"
Proceedings of the 25th European Meeting of Statisticians, July 2005, Oslo, Norway, Paper P-031

This method works for any binary assessment and is not limited to bends.

	# Inconsistencies									
# Joints inspected	0	1	2	3	4	5	6	7	8	9
1	0.010	0.000								
2	0.100	0.005	0.000							
3	0.215	0.059	0.003	0.000						
4	0.316	0.141	0.042	0.003	0.000					
5	0.398	0.222	0.106	0.033	0.002	0.000				
6	0.464	0.294	0.173	0.085	0.027	0.002	0.000			
7	0.518	0.357	0.236	0.142	0.071	0.023	0.001	0.000		
8	0.562	0.410	0.293	0.198	0.121	0.061	0.020	0.001	0.000	
9	0.599	0.456	0.344	0.250	0.171	0.105	0.053	0.017	0.001	0.000
10	0.631	0.496	0.388	0.297	0.218	0.150	0.093	0.048	0.016	0.001
11	0.658	0.530	0.428	0.340	0.262	0.194	0.134	0.084	0.043	0.014
12	0.681	0.560	0.463	0.378	0.302	0.235	0.175	0.121	0.076	0.039
13	0.702	0.587	0.494	0.412	0.339	0.273	0.213	0.159	0.111	0.069
14	0.720	0.611	0.522	0.443	0.373	0.308	0.249	0.195	0.146	0.102
15	0.736	0.632	0.547	0.471	0.403	0.340	0.282	0.229	0.179	0.135
16	0.750	0.651	0.570	0.497	0.431	0.370	0.313	0.261	0.212	0.166
17	0.763	0.668	0.590	0.520	0.457	0.397	0.342	0.291	0.242	0.197
18	0.774	0.684	0.609	0.542	0.480	0.423	0.369	0.319	0.271	0.226
19	0.785	0.698	0.626	0.561	0.502	0.446	0.394	0.345	0.298	0.254
20	0.794	0.711	0.642	0.579	0.522	0.468	0.417	0.369	0.323	0.280
21	0.803	0.723	0.656	0.596	0.540	0.488	0.439	0.392	0.347	0.305
22	0.811	0.734	0.670	0.611	0.557	0.507	0.459	0.413	0.370	0.328
23	0.819	0.744	0.682	0.626	0.573	0.524	0.478	0.433	0.391	0.350
24	0.825	0.754	0.693	0.639	0.588	0.540	0.495	0.452	0.410	0.370
25	0.832	0.763	0.704	0.651	0.602	0.556	0.512	0.469	0.429	0.390
26	0.838	0.771	0.714	0.663	0.615	0.570	0.527	0.486	0.446	0.408
27	0.843	0.778	0.723	0.674	0.627	0.583	0.542	0.502	0.463	0.426
28	0.848	0.785	0.732	0.684	0.639	0.596	0.555	0.516	0.479	0.442
29	0.853	0.792	0.740	0.693	0.650	0.608	0.568	0.530	0.493	0.458
30	0.858	0.798	0.748	0.702	0.660	0.619	0.580	0.543	0.507	0.473

6.2 Finite Population Agresti-Coull Application

Samples sizes of at least 10 are shown in Table 5 for the finite-population N = 50 Agresti-Coull normal approximation to the binomial. While not specifically tied directly to this issue, keep in mind for pipeline applications that Li et al (2017) suggest the sample size be at least 20. The three circled results in Table 5 can be compared to the Clopper-Pearson infinite population 95% confidence level results in Table 2. Doing so it is seen that less samples are needed to obtain the same lower bounds using the Agresti-Coull finite population methodology than the Clopper-Pearson exact solution; however, a finite population size of N = 50 is small. As the finite population N is increased in size the two results will become closer.

Table 5. Lower-Bound Agresti-Coull Consistency (95% Confidence) for a Finite Population of 50

	A	B	C	D	E	F	G	H	I	J	K
1	**Lower bound on reliability shown in matrix below**										
2		95%	confidence level *(change as desired)*								
3		0.80	Reliability level for conditional formatting *(change as desired)*								
4		50	Finite Population Size N								
5	References										
6	Hollander, Myles and Douglas A., Wolfe Nonparametric Statistical Methods,										
7	2nd Edition, ISBN 978-0471190455, 1999, John Wiley & Sons, New York										
8	Harper, William V., “Excel Functions to Compute Exact Binomial C							Agresti_Coull_FinitePop_CI(C67, D67, F67,			
9	Proceedings of the 25th European Meeting of Statisticians, July 20							Function Agresti_Coull_FinitePop_CI(X, N,			
10	*This method works for any binary assessment and is not limited to*							Function Agresti_Coull_FinitePop_CI(X, N,			
11		# Inconsistencies									
12	# Joints inspected	0	1	2	3	4	5	6	7	8	9
13	10	0.766	0.654	0.554	0.393	0.304	0.222	0.150	0.087	0.037	0.005
14	11	0.784	0.680	0.587	0.436	0.350	0.271	0.200	0.135	0.079	0.033
15	12	0.800	0.703	0.615	0.473	0.391	0.315	0.245	0.181	0.123	0.072
16	13	0.814	0.723	0.640	0.505	0.427	0.355	0.287	0.224	0.166	0.113
17	14	0.826	0.740	0.663	0.534	0.460	0.390	0.325	0.264	0.206	0.153
18	15	0.837	0.756	0.682	0.560	0.489	0.423	0.360	0.300	0.244	0.191
19	16	0.846	0.770	0.700	0.583	0.516	0.452	0.391	0.333	0.279	0.227
20	17	0.855	0.783	0.717	0.604	0.539	0.478	0.420	0.364	0.311	0.260
21	18	0.863	0.795	0.732	0.623	0.561	0.502	0.446	0.392	0.341	0.291
22	19	0.870	0.805	0.745	0.641	0.581	0.524	0.470	0.418	0.368	0.320
23	20	0.877	0.815	0.757	0.656	0.599	0.544	0.492	0.442	0.394	0.347
24	21	0.883	0.824	0.769	0.671	0.616	0.563	0.513	0.464	0.417	0.372
25	22	0.889	0.832	0.779	0.684	0.631	0.580	0.532	0.485	0.439	0.395
26	23	0.894	0.840	0.789	0.696	0.645	0.596	0.549	0.504	0.460	0.417
27	24	0.899	0.847	0.798	0.708	0.658	0.611	0.565	0.521	0.479	0.437
28	25	0.904	0.853	0.807	0.718	0.670	0.625	0.580	0.538	0.496	0.456
29	26	0.908	0.860	0.815	0.728	0.682	0.637	0.595	0.553	0.513	0.474
30	27	0.912	0.866	0.822	0.737	0.692	0.649	0.608	0.568	0.529	0.491
31	28	0.916	0.871	0.829	0.746	0.702	0.661	0.620	0.581	0.543	0.506
32	29	0.919	0.876	0.836	0.754	0.712	0.671	0.632	0.594	0.557	0.521
33	30	0.922	0.881	0.842	0.761	0.720	0.681	0.643	0.606	0.570	0.535

7. Summary

This paper examines one-sample binomial confidence intervals including the standard textbook normal approximation, the exact Clopper-Pearson, and the Agresti-Coull normal approximation. An extension of the Agresti-Could normal approximation to finite populations is made. Discussion involving the API 1163 use of the infinite population Agresti-Coull is providing showing how it is used as part of an evaluation of in-line inspection tools to actual field excavations for oil and gas pipelines. Finally Excel VBA software to implement exact Clopper-Pearson and finite population Agresti-Coull normal approximations to binomial confidence intervals is shown along with some discussion of PHMSA’s recent NPRM.

8. References

Agresti, Alan & Brent A. Coull (1998), 'Approximate is Better than "Exact" for Interval Estimation of Binomial Proportions', The American Statistician, 52:2, 119-126.

API Standard 1163, "In-line Inspection Systems Qualification", 1st edition, August 2005, American Petroleum Institute.

API Standard 1163, "In-line Inspection Systems Qualification", 2nd edition, April 2013, American Petroleum Institute.

API web page http://www.api.org/about.

Clopper, C. J., and E.S. Pearson, 1934, "The use of confidence or fiducial limits illustrated in the case of the binomial", Biometrika, 26, 404-413.

Harper, William V., "Excel Functions to Compute Exact Binomial Confidence Intervals", proceedings of the 25th European Meeting of Statisticians, July 2005, Oslo, Norway, Paper P-031.

Hollander, Myles, and Douglas A. Wolfe, Nonparametric Statistical Methods, 2nd edition, John Wiley & Sons, Inc., New York, 1999, pp. 32, note 20, Interval Endpoints.

Li, Yanping, Gordon Fredine, Yvan Hubert, Vasily Vorontsov, Sherif Hassanien, and Janaine Woo, "Evaluating ILI tool performance using a validation process", Pipeline pigging and integrity management conference, Houston, March 2017.

Pipeline and Hazardous Materials Safety Administration (PHMSA) notice of proposed rulemaking (NPRM), Pipeline Safety: Safety of Gas Transmission and Gathering Pipelines, Docket No. PHMSA-2011-0023, published on April 8, 2016.