

Put Your Tables Away

Robert K. Henderson

February 21, 2020

Suppose you have a sample of size 10, and desire to know whether or not the mean of the population from which the sample was obtained is greater than zero.

Most likely, you would calculate a t-statistic $t = \frac{\bar{X}}{\frac{S}{\sqrt{n}}}$, where

\bar{X} = sample mean, S = sample standard deviation, n = the sample size = 10 in this case.

Now, suppose you do this, and you get $t = 3$. You realize this is likely a statistically significant result, but the pesky journal editor requires a p-value.

Hence, it becomes necessary to calculate the probability a student's t random variable with 9 degrees of freedom is greater than equal to $t = 3$, i.e., $P[t \geq 3 \mid t \sim t_{(9)}]$.

In many stat classes, students are asked to do this using a student's t table, such as that at right.

To get $P[t \geq 3 \mid t \sim t_{(9)}]$, the student has to find the value "3" in the row labeled "9" in the "df" column. It is not there, but the values 2.821 and 3.250 bounding it are.

The value "3" is ~41.725% of the distance between these bounds, so to obtain the desired p-value, it is necessary to find the value ~41.725% of the distance between 0.01 & 0.005, which is ~0.0079.

OR ...

t Table

cum. prob one-tail two-tails	$t_{.50}$	$t_{.25}$	$t_{.20}$	$t_{.15}$	$t_{.10}$	$t_{.05}$	$t_{.025}$	$t_{.01}$	$t_{.005}$	$t_{.001}$	$t_{.0005}$
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.898	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300

One could use an Excel utility designed to calculate such probabilities.

Enter desired df = 9

Enter the value of interest = 3

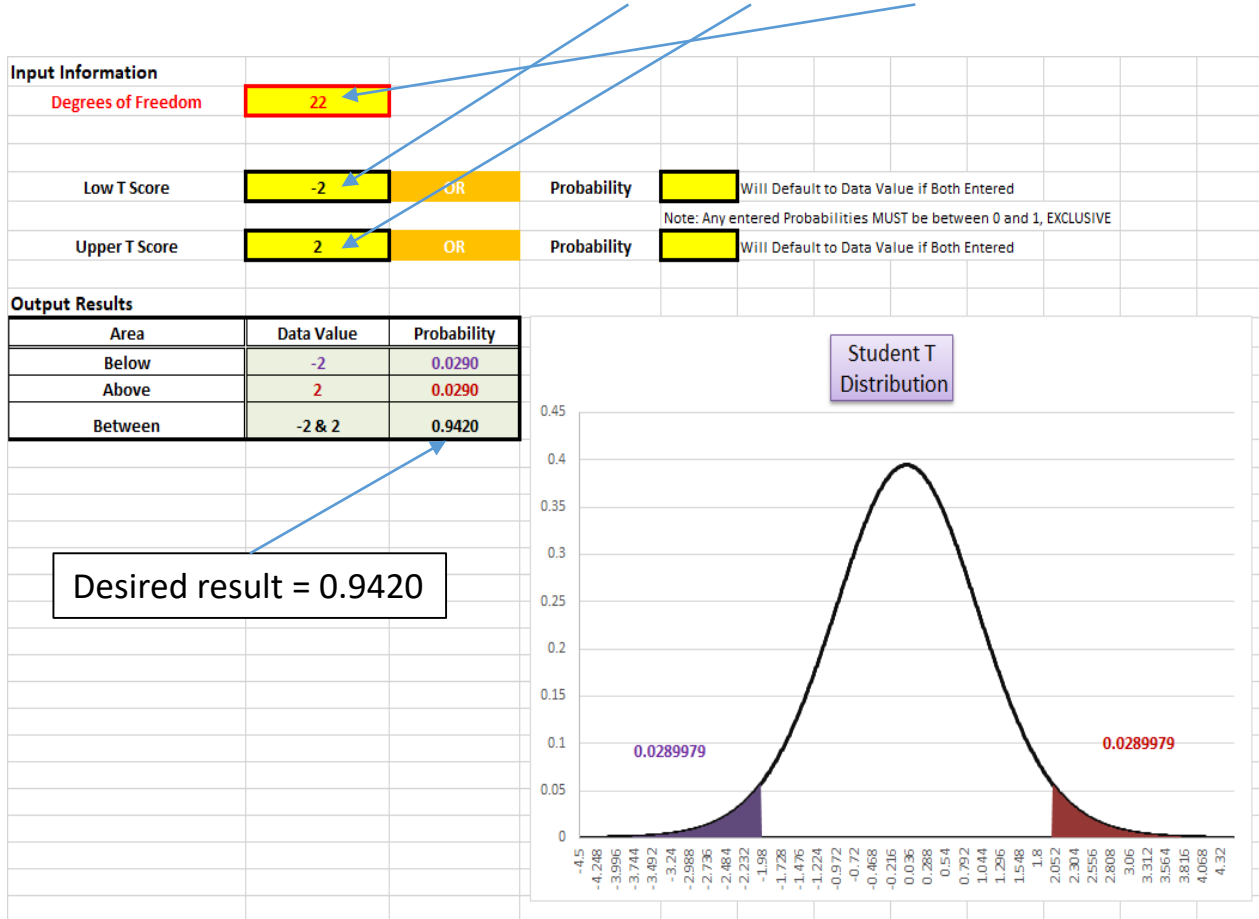
Input Information			
Degrees of Freedom	9		
Low T Score	3	OR	Probability <input type="text"/> Will Default to Data Value if Both Entered
Upper T Score		OR	Probability <input type="text"/> Will Default to Data Value if Both Entered
Note: Any entered Probabilities MUST be between 0 and 1, EXCLUSIVE			
Output Results			
Area	Data Value	Probability	
Below	3	0.9925	
Above	3	0.0075	
Between	3 & 3	0.0000	

Student T Distribution			
Area	Data Value	Probability	
Below	3	0.9925218	
Above	3	0.0074782	
Between	3 & 3	0.0000	

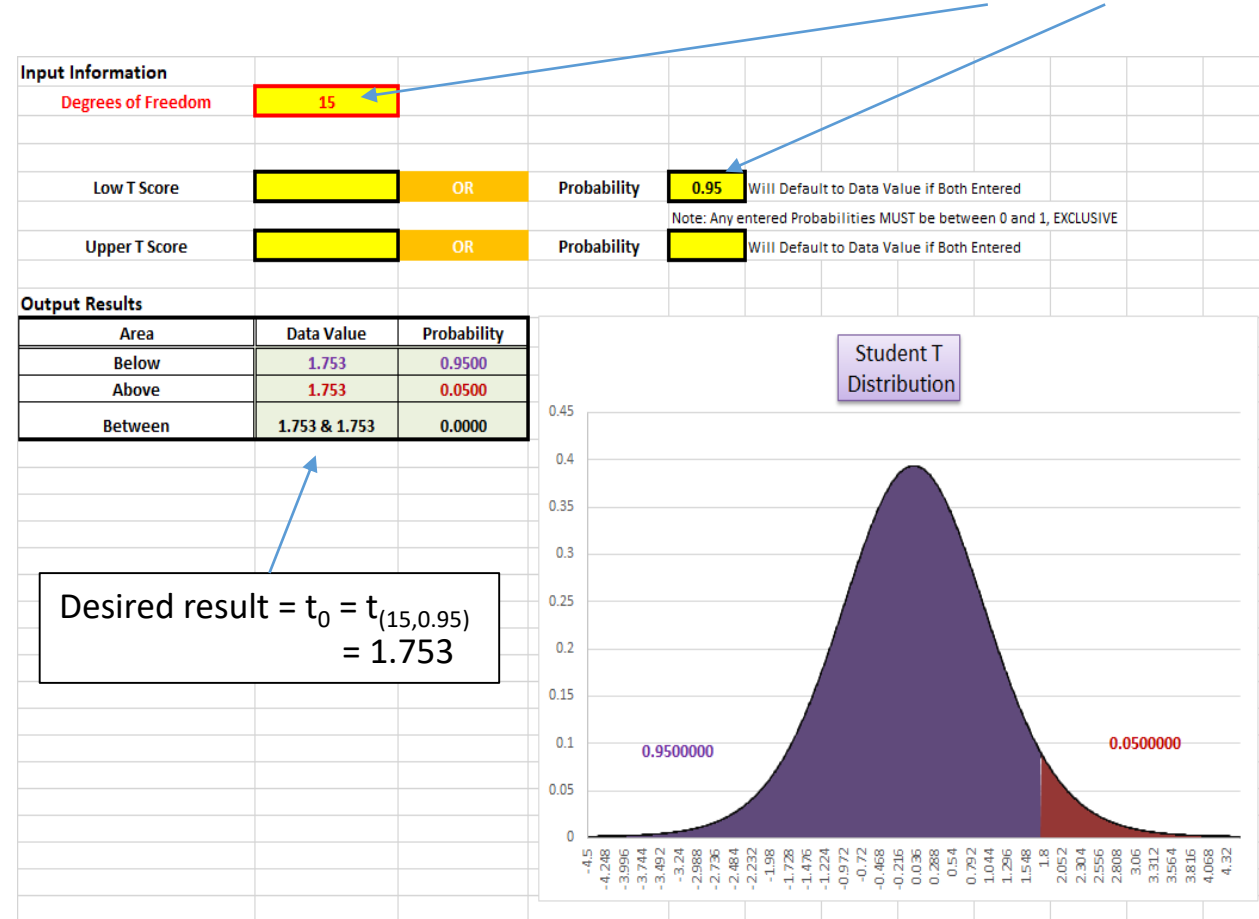
Output includes $P[t_{(9)} > 3]$

Note: Result is less than the 0.0079 obtained with the table.

Utility can handle two percentiles as well.
 For example, can find $P[-2 < t < 2 \mid t \sim t_{(22)}]$



Also, can find percentiles given desired probabilities.
 For example, can find t_0 for $P[t < t_0 \mid t \sim t_{(15)}] = 0.95$



How about finding a binomial probability?

Might be necessary when using data to test if a population proportion is smaller or larger than some reference value?

Suppose we had results from an exit poll of 20 randomly selected people voting in the March primary here in Nacogdoches where the respondents were asked who they intended to vote for for president in November.

Suppose further that 14 of the 20 indicated their intent to vote for Donald Trump. Does this result provide sufficient evidence to indicate President Trump is likely to carry Nacogdoches in November?

To answer this, we would need to evaluate the probability of obtaining this observation or one more extreme if the true proportion of Trump voters is ≤ 0.5 . This probability can be estimated using the binomial probability model as

$P[\geq 14 \text{ Trump voters of } 20 | \% \text{ of all Nacogdoches Trump voters is } \leq 50\%]$

This requires the summation of 7 binomial probabilities:

$$P_i = \binom{20}{i} 0.5^{20}, i = 14, \dots, 20$$

Or, using a binomial table at right adding the values

$$.037 + .015 + .005 + .001 + .000 + .000 + .000 = .058$$

Table 4 continued

n	r	p																			
		.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
16	2	.010	.146	.275	.277	.211	.134	.073	.035	.015	.006	.002	.001	.000	.000	.000	.000	.000	.000	.000	.000
	3	.000	.036	.142	.229	.246	.208	.146	.089	.047	.022	.009	.003	.001	.000	.000	.000	.000	.000	.000	.000
	4	.000	.006	.051	.131	.200	.225	.204	.155	.101	.057	.028	.011	.004	.001	.000	.000	.000	.000	.000	.000
	5	.000	.001	.014	.056	.120	.180	.210	.201	.162	.112	.067	.034	.014	.005	.001	.000	.000	.000	.000	.000
	6	.000	.000	.003	.018	.055	.110	.165	.198	.198	.168	.122	.075	.039	.017	.006	.001	.000	.000	.000	.000
	7	.000	.000	.000	.005	.020	.052	.101	.152	.189	.197	.175	.132	.084	.044	.019	.006	.001	.000	.000	.000
	8	.000	.000	.000	.001	.006	.028	.049	.092	.142	.181	.196	.181	.142	.092	.049	.020	.006	.001	.000	.000
	9	.000	.000	.000	.000	.001	.006	.019	.044	.084	.132	.175	.197	.189	.152	.101	.052	.020	.005	.000	.000
	10	.000	.000	.000	.000	.000	.001	.006	.017	.039	.075	.122	.168	.198	.198	.165	.110	.055	.018	.003	.000
	11	.000	.000	.000	.000	.000	.000	.001	.005	.014	.034	.067	.112	.162	.201	.210	.180	.120	.056	.014	.001
	12	.000	.000	.000	.000	.000	.000	.000	.001	.004	.011	.028	.057	.101	.155	.204	.225	.200	.131	.051	.006
	13	.000	.000	.000	.000	.000	.000	.000	.000	.001	.003	.009	.022	.047	.089	.146	.208	.246	.229	.142	.036
	14	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001	.002	.006	.015	.035	.073	.134	.211	.277	.275	.146
	15	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001	.003	.009	.023	.053	.113	.210	.329	.371
	16	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001	.003	.010	.028	.074	.185	.440
	20	.818	.358	.122	.039	.012	.003	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1	.165	.377	.270	.137	.058	.021	.007	.002	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	2	.016	.189	.285	.229	.137	.067	.028	.010	.003	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	3	.001	.060	.190	.243	.205	.134	.072	.032	.012	.004	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000
	4	.000	.013	.090	.182	.218	.190	.130	.074	.035	.014	.005	.001	.000	.000	.000	.000	.000	.000	.000	.000
	5	.000	.002	.032	.103	.175	.202	.179	.127	.075	.036	.015	.005	.001	.000	.000	.000	.000	.000	.000	.000
	6	.000	.000	.009	.045	.109	.169	.192	.171	.124	.075	.037	.015	.005	.001	.000	.000	.000	.000	.000	.000
	7	.000	.000	.002	.016	.055	.112	.164	.184	.166	.122	.074	.037	.015	.005	.001	.000	.000	.000	.000	.000
	8	.000	.000	.000	.005	.022	.061	.114	.161	.180	.162	.120	.073	.035	.014	.004	.001	.000	.000	.000	.000
	9	.000	.000	.000	.001	.007	.027	.065	.116	.160	.177	.160	.119	.071	.034	.012	.003	.000	.000	.000	.000
	10	.000	.000	.000	.000	.002	.010	.031	.069	.117	.159	.176	.159	.117	.069	.031	.010	.002	.000	.000	.000
	11	.000	.000	.000	.000	.000	.003	.012	.034	.071	.119	.160	.177	.160	.116	.065	.027	.007	.001	.000	.000
	12	.000	.000	.000	.000	.000	.001	.004	.014	.035	.073	.120	.162	.180	.161	.114	.061	.022	.005	.000	.000
	13	.000	.000	.000	.000	.000	.000	.001	.005	.015	.037	.074	.122	.166	.184	.164	.112	.055	.016	.002	.000
	14	.000	.000	.000	.000	.000	.000	.000	.001	.005	.015	.037	.075	.124	.171	.192	.169	.109	.045	.009	.000
	15	.000	.000	.000	.000	.000	.000	.000	.000	.001	.005	.015	.036	.075	.127	.179	.202	.175	.103	.032	.002
	16	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001	.005	.014	.035	.074	.130	.190	.218	.182	.090	.013
	17	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001	.004	.012	.032	.072	.134	.205	.243	.190	.060
	18	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001	.003	.010	.028	.067	.137	.229	.285	.189
	19	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.002	.007	.021	.058	.137	.270	.377
	20	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001	.003	.012	.039	.122	.358	.818

OR ...

One could use an Excel utility designed to calculate such probabilities.

$$P[\geq 14 \text{ Trump voters of } 20 \mid \% \text{ of all Nacogdoches Trump voters is } \leq 50\%] = 0.057659$$

Binomial Probabilities				Binomial Parameters			
Exact	n	20	Input	Number	n	20	Input
Number	n(Success)	14	Input	of	n(Success-Lo)	6	Input
of	P[Success]	0.5	Input	Successes	n(Success-Hi)	14	Input
Successes	P[X=14 in 20]	0.036964	Output	Between	P[Success]	0.5	Input
					P[6<=X<=14 in 20]	0.958611	Output
At Least	n	20	Input				
Number	n(Success)	14	Input				
of	P[Success]	0.5	Input				
Successes	P[X>=14 in 20]	0.057659	Output				
At Most	n	20	Input				
Number	n(Success)	13	Input				
of	P[Success]	0.5	Input				
Successes	P[X<=13 in 20]	0.942341	Output				

Often, Binomial Tables are difficult to find for $n > 20$; however, most real polls have well beyond $n = 20$ respondents.

What if the primary exit poll obtained responses from 200 voters and 118 of them indicated they would vote for president Trump in November?

Now interested in $P[\geq 118 \text{ Trump voters in } 200 \mid \% \text{ of all Nacogdoches Trump voters} \leq 50\%] = 0.006565$

Binomial Probabilities							
Exact	n	20	Input	Number	n	20	Input
Number	n(Success)	14	Input	of	n(Success-Lo)	6	Input
of	P[Success]	0.5	Input	Successes	n(Success-Hi)	14	Input
Successes	P[X=14 in 20]	0.036964	Output	Between	P[Success]	0.5	Input
					P[6<=X<=14 in 20]	0.958611	Output
At Least	n	200	Input				
Number	n(Success)	118	Input				
of	P[Success]	0.5	Input				
Successes	P[X>=118 in 200]	0.006565	Output				
At Most	n	20	Input				
Number	n(Success)	13	Input				
of	P[Success]	0.5	Input				
Successes	P[X<=13 in 20]	0.942341	Output				

Binomial Parameters			
	n	20	Input
	P[Success]	0.5	Input
	Mean	10	Output
	Variance	5	Output
	Std Deviation	2.236068	Output

Binomial Utility can also produce 100(1- α)% exact Binomial Confidence Intervals for a population proportion.

95% Confidence intervals for the proportion of all Nacogdoches Trump voters based on the two polls would be

Binomial Confidence Intervals for Population Proportion				
Inputs				
Alpha	0.05	2.9277	This interval more appropriate than interval based on Normal approximation	
Obs Successes	14			
Sample Size	20			
Exact Interval				
	Lower		Upper	
	Limit	Estimate	Limit	
	0.457211	0.7	0.881068	

So with n = 20 & 14 Trump voters, we could be 95% confident the true % of all Nacogdoches Trump voters is between ~45.7% & ~88.1%.

Binomial Confidence Intervals for Population Proportion				
Inputs				
Alpha	0.05	11.7891	Normal approximation interval should be similar	
Obs Successes	118			
Sample Size	200			
Exact Interval				
	Lower		Upper	
	Limit	Estimate	Limit	
	0.518422	0.59	0.658869	
Approximate Interval				
	Lower		Upper	
	Limit	Estimate	Limit	
	0.521837	0.59	0.658163	

But with n = 200 & 118 Trump voters, we could be 95% confident the true % of all Nacogdoches Trump voters is between ~51.8% & ~65.9%.

So the potential of a Normal approximation for these confidence intervals suggests how the

$P[\geq 118 \text{ Trump voters in } 200 \mid \% \text{ of all Nacogdoches Trump voters} \leq 50\%]$

would need to be estimated **WITHOUT** a Binomial Utility & No Binomial Table w/ $n = 200$

Step 1: Find Expected Count = $200 * 0.50 = 100$

Step 2: Find Standard Deviation of Count = $\sqrt{200 * 0.5 * 0.5}$
 ≈ 7.07107

Step 3: Calculate a Z-score for the Count:

$$Z = \frac{117.5 - 100}{7.07107} \approx 2.475$$

Step 4: Use a Standard Normal Table to Determine Desired Probability

Here, need to find the value 60% of the way between

.00554 & .00539

which is .00545

Note this approximation is not very good
 Approximately 17% < Actual = .006565.

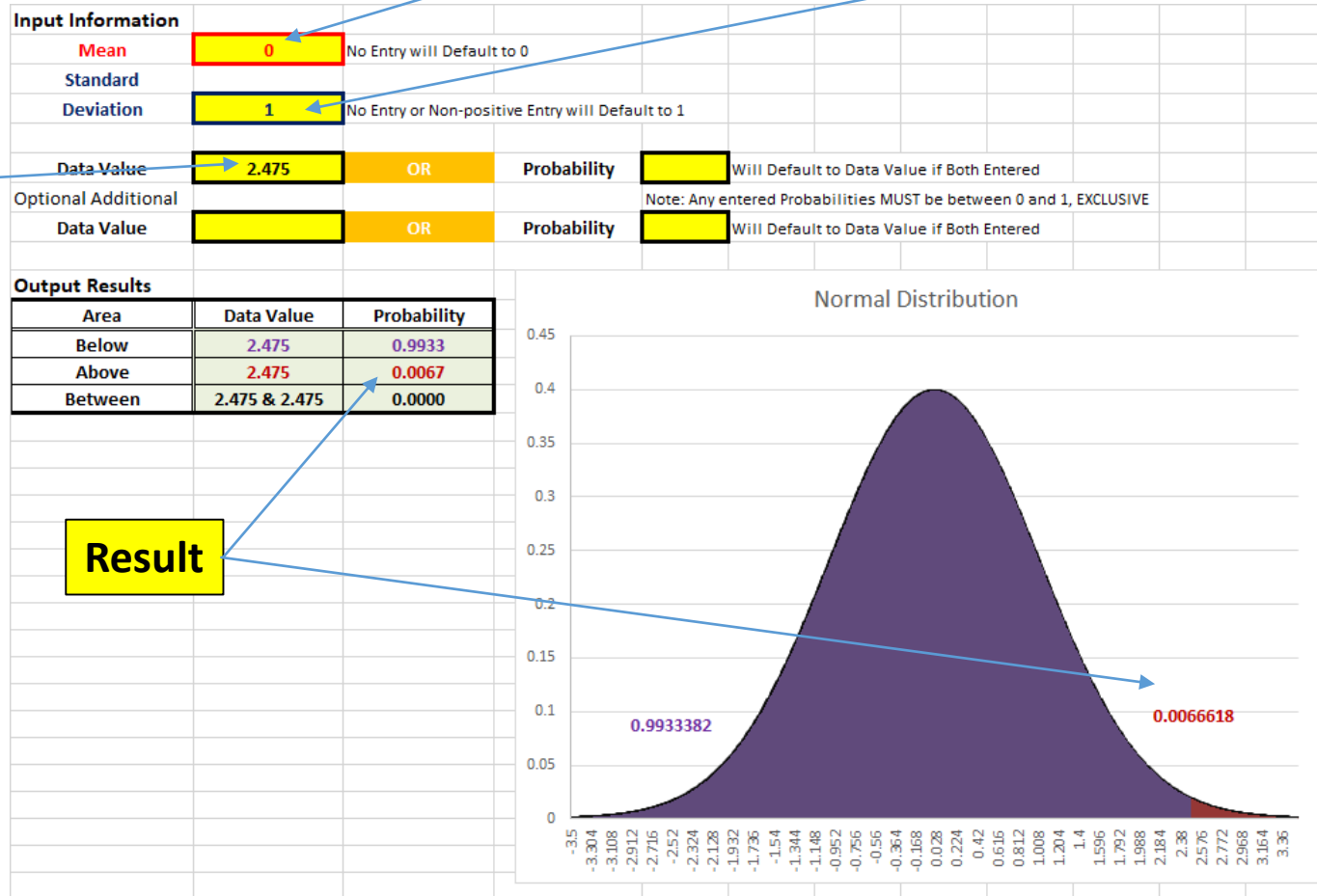
Better with Continuity Correction
 Approximately 1.5% > Actual = .006565.

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414
-0.1	.46017	.45620	.45224	.44828	.44433	.44034	.43640	.43251	.42858	.42465
-0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
-0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
-0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
-0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
-0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
-0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
-0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
-0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
-1	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
-1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08692	.08534	.08379	.08226
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
-1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
-1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
-1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
-1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
-2	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
-2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
-2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
-2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
-2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
-2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
-2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
-2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
-2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
-2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
-3	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
-3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
-3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
-3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
-3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
-3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
-3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
-3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
-3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
-3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
-4	.00003	.00003	.00003	.00003	.00003	.00003	.00002	.00002	.00002	.00002

However, with a suitable Normal Distribution Excel utility:

Standard Normal has Mean = 0 and Standard Deviation = 1

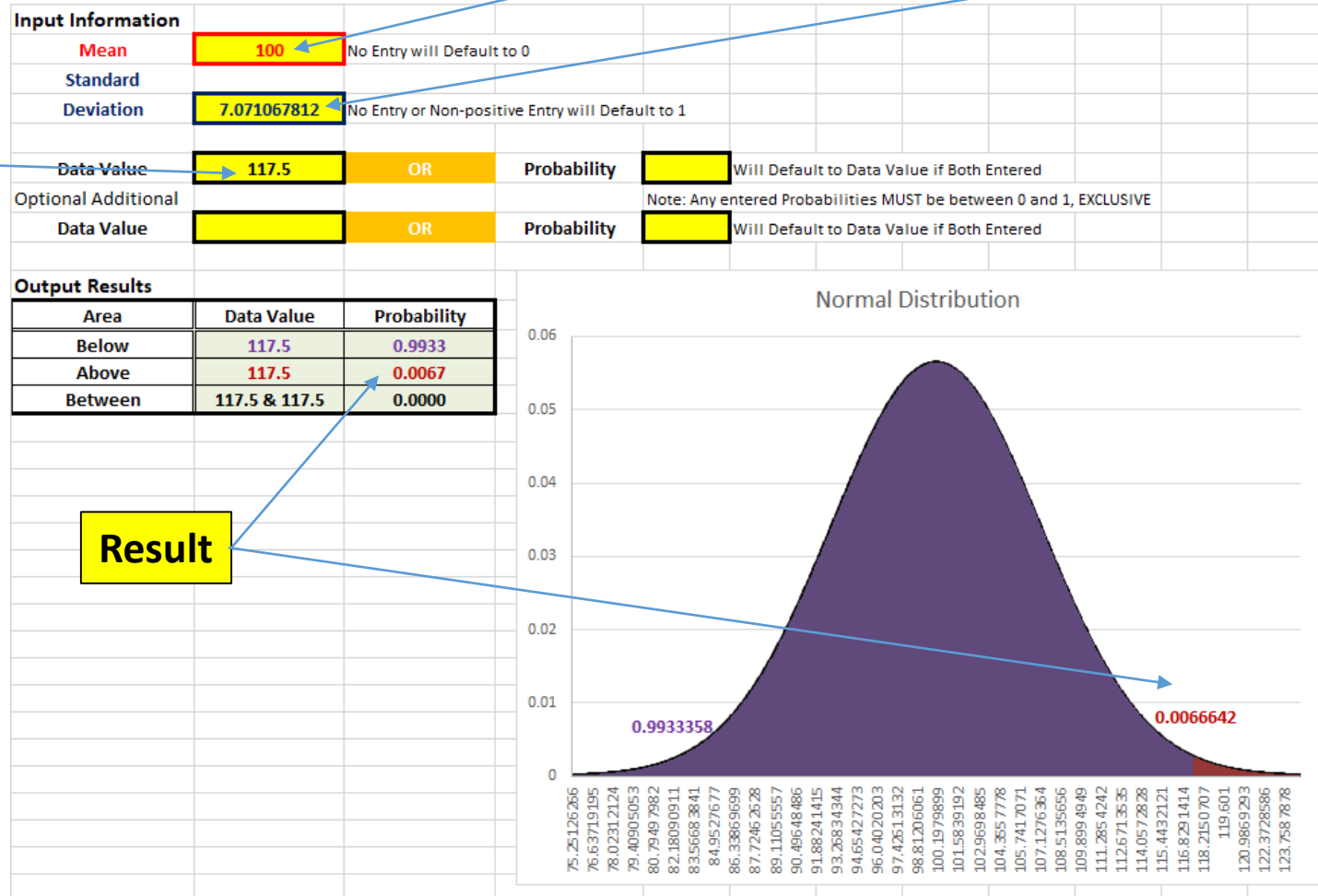
Desired
Z-score



Of course, with this utility, there is no need to even calculate the Z-score:

Approximate Count Distribution is Normal with Mean = 100 and Standard Deviation = 7.07107

Enter
Observed
Count
w/
Continuity
Correction



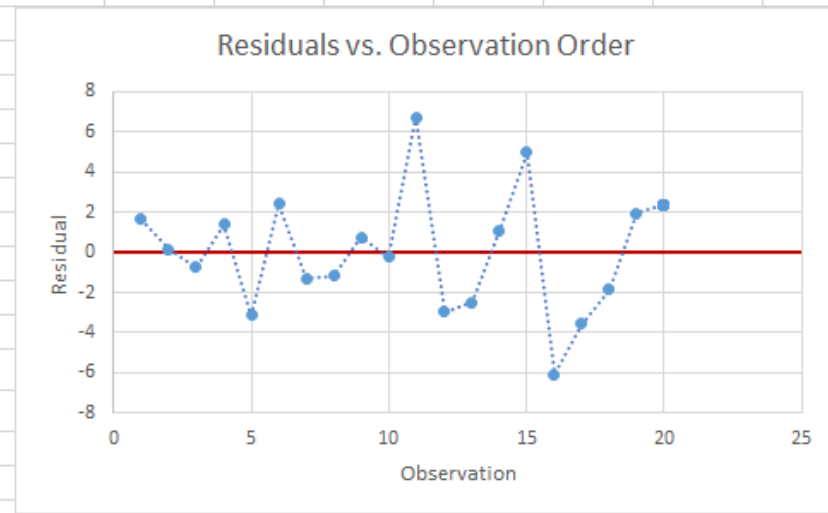
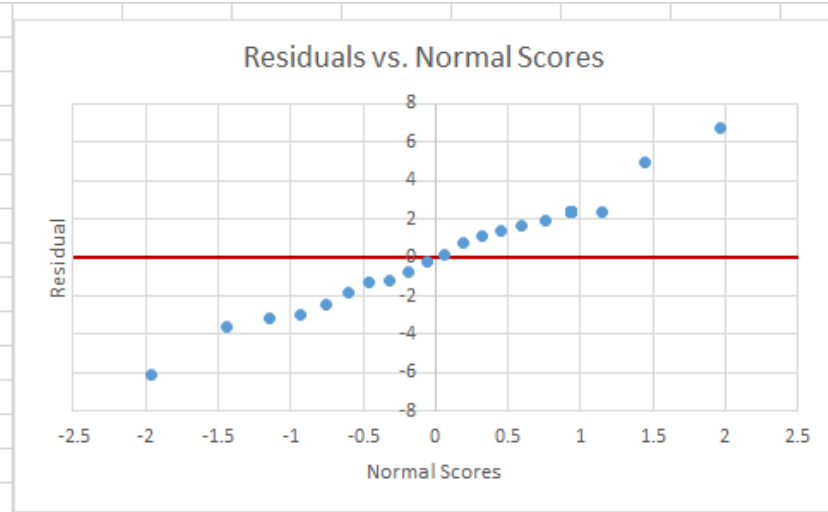
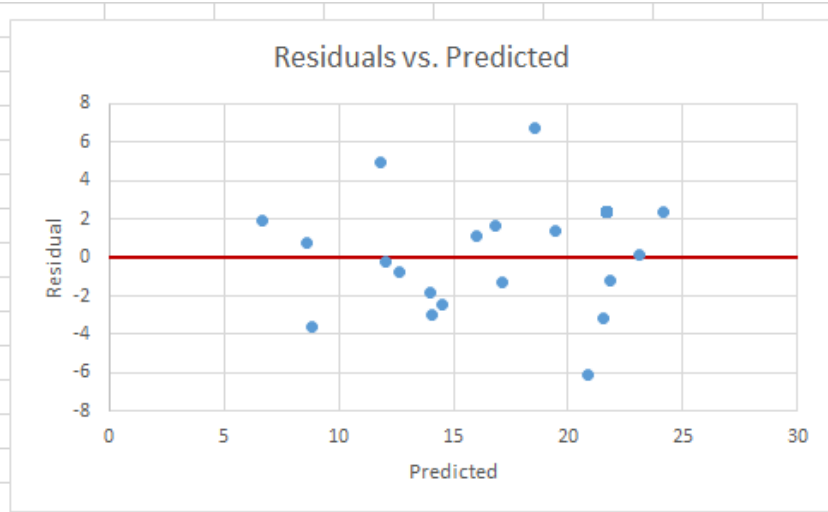
Result

Also, there is a Simple Regression Excel utility available:

Basic Worksheet

Diagnostics Worksheet

Only input is ordered pairs of Predictor (X) and Response (Y) Starting in Row 2. Currently, can accommodate up to n = 1000 data pairs.



Test for Constant Error Variance			
	Breusch-Pagan Chi-Square	0.01406	
	p-value	0.905612	Assumption of Constant Error Variance is REASONABLE
Test for Normally Distributed Errors			
	Ryan-Joiner Correlation	0.989923	
	p-value	0.541667	Assumption of Normally Distributed Errors is REASONABLE
Test for Independent Errors			
	Durbin-Watson Statistic	2.272635	
	Critical	d_L	0.95
	Values	d_U	1.15
			Assumption of Independent Errors is REASONABLE