1. Suppose that my continuous distribution was the beta distribution. It has two parameters, α and β, both of which are shape parameters and are constrained to be greater than zero. On the Wikipedia page curves are drawn for the following combinations: $\alpha = 0.5$, $\beta = 0.5$; $\alpha = 5$, $\beta = 1$; $\alpha = 1$, $\beta = 3$; $\alpha = 2$, $\beta = 2$; and $\alpha = 2$, $\beta = 5$.

   This distribution has non-zero density at x only when $0 < x < 1$ and the plot is shown in Fig.~1 and the comparative cdf plot in Fig.~2. (R code to produce the plot is submitted separately.)

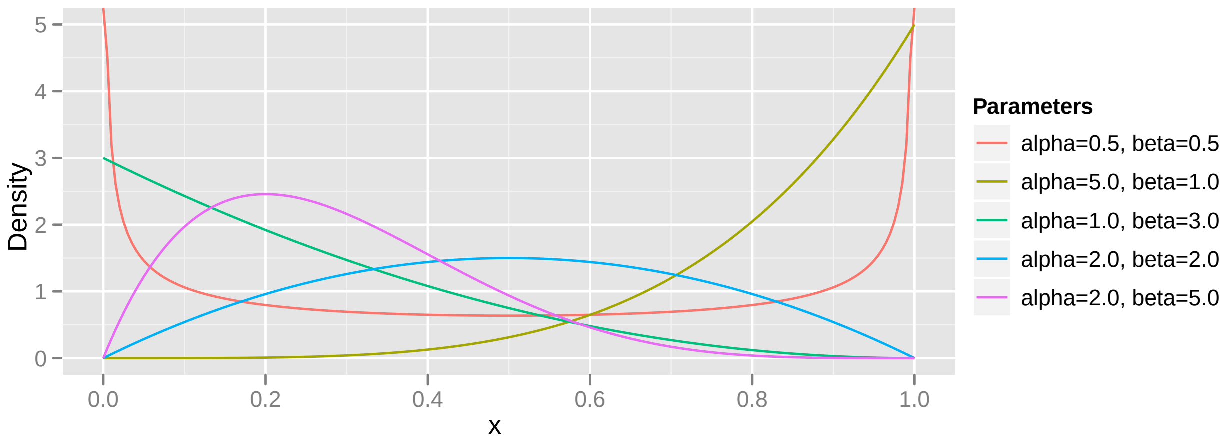

Figure 1: Comparative density plot for the beta distribution. The densities are for the parameter values shown in the legend.

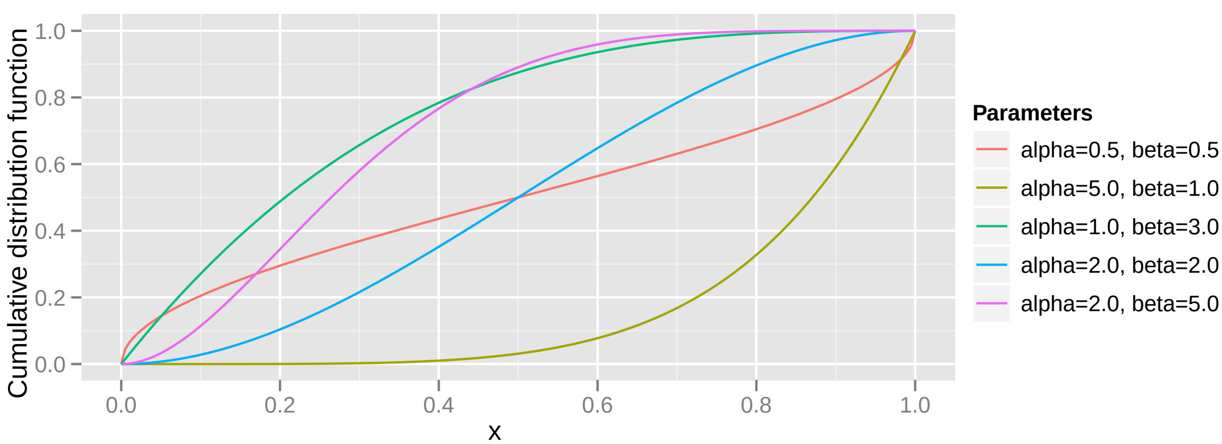

Figure 2: Comparative cumulative distribution function plot for the beta distribution. The densities are for the parameter values shown in the legend.

2. Suppose that my discrete distribution was the binomial. There are two parameters, p and n, that affect the shape. The Wikipedia page provides only three combinations; $p = 0.5$, $n = 20$; $p = 0.7$, $n = 20$; and $p = 0.5$, $n = 40$ so we will add $p = 0.7$, $n = 40$ to get a total of 4 combinations in Fig.~3.

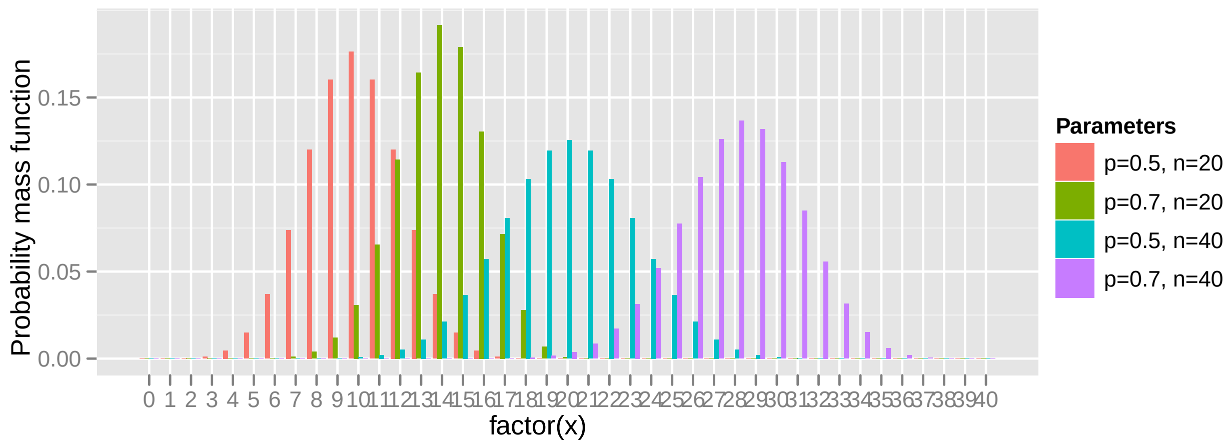

Figure 3: Comparative probability mass functions for the binomial distribution. The p.m.f.s are for the parameter values shown in the legend.

An alternative plot, shown in Fig.˜4, uses faceting instead of colors to distinguish the pmf's for different values of the parameters.

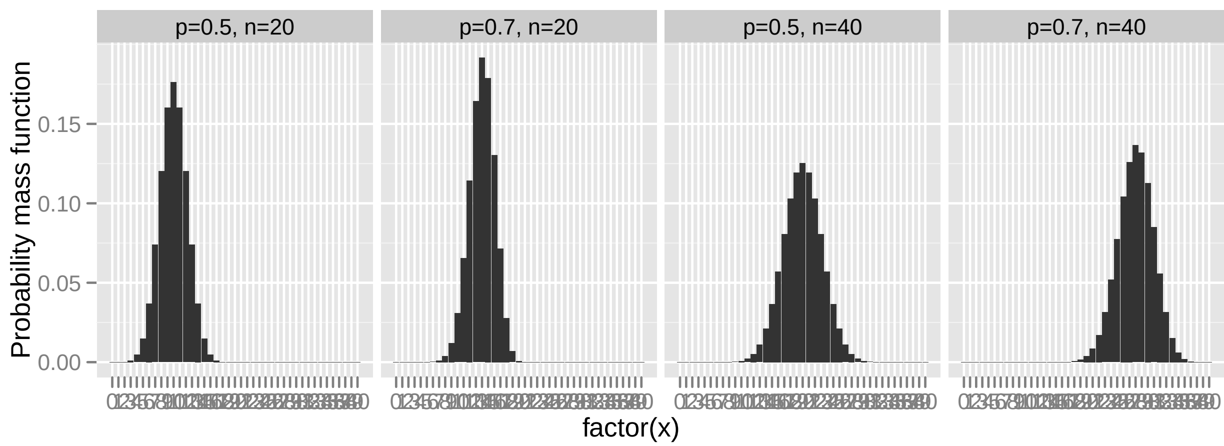

Figure 4: Comparative probability mass functions for the binomial distribution. The p.m.f.s are for the parameter values shown in the strip above each panel.

3. For the beta distribution the theoretical mean is $\frac{\alpha}{\alpha+\beta}$ and the theoretical variance is $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$, according to Wikipedia. I simulated 1,000,000 values from beta distributions with parameter values shown in Table˜1 to compare the theoretical and empirical mean and variance. I used samples of size 1,000,000 to get empirical values that should be very close to the theoretical values. Because only one sample is being simulated for each set of parameter values, it is not a great burden to make that a large sample.

   There is very good agreement between the empirical and the theoretical values.

4. For the binomial distribution the theoretical mean is np and the theoretical variance is $np(1-p)$, according to Wikipedia. I simulated 1,000,000 values from binomial distributions with parameter

		mean		variance	
α	β	theoretical	empirical	theoretical	empirical
0.5	0.5	0.50000	0.50009	0.12500	0.12499
5.0	1.0	0.83333	0.83303	0.01984	0.01991
1.0	3.0	0.25000	0.24998	0.03750	0.03749
1.0	5.0	0.16667	0.16639	0.01984	0.01976
2.0	2.0	0.50000	0.49997	0.05000	0.04991
2.0	5.0	0.28571	0.28562	0.02551	0.02545

Table 1: Theoretical and empirical means and variances for the beta distribution with different values of the parameters, α and β. Each set of empirical values are based on a sample of size 1,000,000.

values shown in Table˜2 to compare the theoretical and empirical mean and variance. Again,

		mean		variance	
n	p	theoretical	empirical	theoretical	empirical
20	0.3	6.00000	6.00141	4.20000	4.19943
20	0.5	10.00000	10.00189	5.00000	4.99342
20	0.7	14.00000	13.99856	4.20000	4.19359
40	0.3	12.00000	12.00250	8.40000	8.40743
40	0.5	20.00000	19.99622	10.00000	9.99561
40	0.7	28.00000	27.99579	8.40000	8.38292

Table 2: Theoretical and empirical means and variances for the binomial distribution with different values of the parameters, n and p. Each set of empirical values are based on a sample of size 1,000,000.

there is very good agreement between the empirical and the theoretical values.

5. I simulated 100,000 replications of the means of samples of sizes 9, 16, 25, and 36 from a beta distribution with $\alpha = 2$ and $\beta = 5$. A comparative density plot of the sample means is shown in Fig.˜5.

   The theoretical and empirical means and standard deviations of the sample means are shown in Table˜3.

	mean		standard deviation	
n	theoretical	empirical	theoretical	empirical
9	0.28571	0.28575	0.05324	0.05336
16	0.28571	0.28573	0.03993	0.03992
25	0.28571	0.28557	0.03194	0.03190
36	0.28571	0.28576	0.02662	0.02651

Table 3: Theoretical and empirical means and standard deviations of the sample mean from samples of sizes 9, 16, 25, and 36 from a beta distribution with $\alpha = 2$ and $\beta = 5$.

6. The comparative quantile-quantile plot is shown in Fig.˜6

7. The comparative density plots are shown in Fig.˜7.

   The table of theoretical and empirical values is Table˜5

   The comparative quantile-quantile plot is shown in Fig.˜8

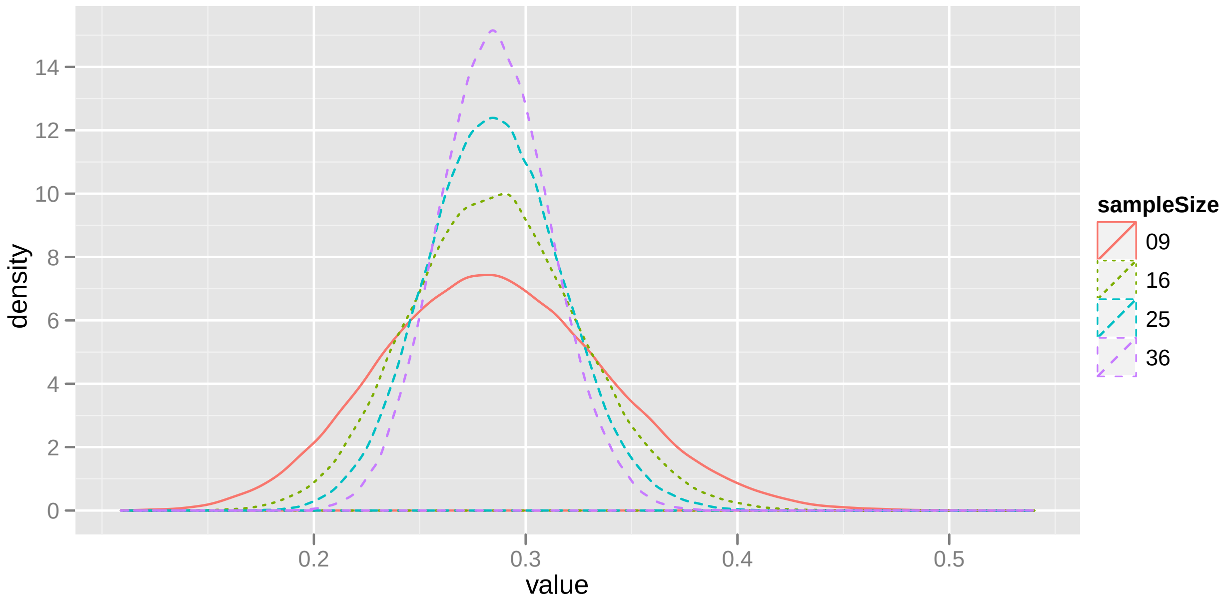

Figure 5: Empirical density plot of the sample means from samples of sizes 9, 16, 25 and 36 from a beta distribution with $\alpha = 2$ and $\beta = 5$. Each empirical density is based on 100,000 replications.

	mean		standard deviation	
n	theoretical	empirical	theoretical	empirical
9	6.00000	6.00107	0.68313	0.68513
16	6.00000	5.99935	0.51235	0.51349
25	6.00000	6.00327	0.40988	0.40985
36	6.00000	5.99970	0.34157	0.34062

Table 4: Theoretical and empirical means and standard deviations of the sample mean from samples of sizes 9, 16, 25, and 36 from a binomial distribution with $n = 20$ and $p = 0.3$.

8. 20,000 replications of the sample mean and sample median from samples of size 9, 16, 25, and 36 from a standard normal distribution were created. Comparative density plots are shown in Figure˜9

   The table of means and standard deviations for the different sample sizes and statistics is Table˜**??**

   The table of the middle 80%, 90%, 95% and 99% of the samples is Table˜6

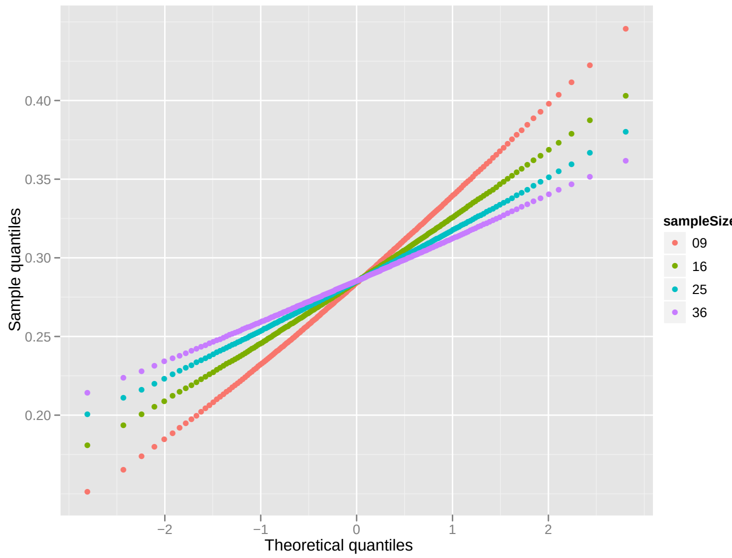

Figure 6: Comparative normal quantile-quantile plots based on 200 equally spaced quantiles of the means of samples of size 9, 16, 25 and 36 from a beta distribution with $\alpha = 2$ and $\beta = 5$.

	sample mean		sample median	
n	mean	std.dev.	mean	std.dev
9	-0.00018	0.33343	0.00178	0.40586
16	0.00052	0.24853	-0.00014	0.29752
25	-0.00148	0.20093	-0.00171	0.24944
36	0.00118	0.16674	0.00173	0.20444

Table 5: Theoretical and empirical means and standard deviations of the sample mean from samples of sizes 9, 16, 25, and 36 from a binomial distribution with $n = 20$ and $p = 0.3$.

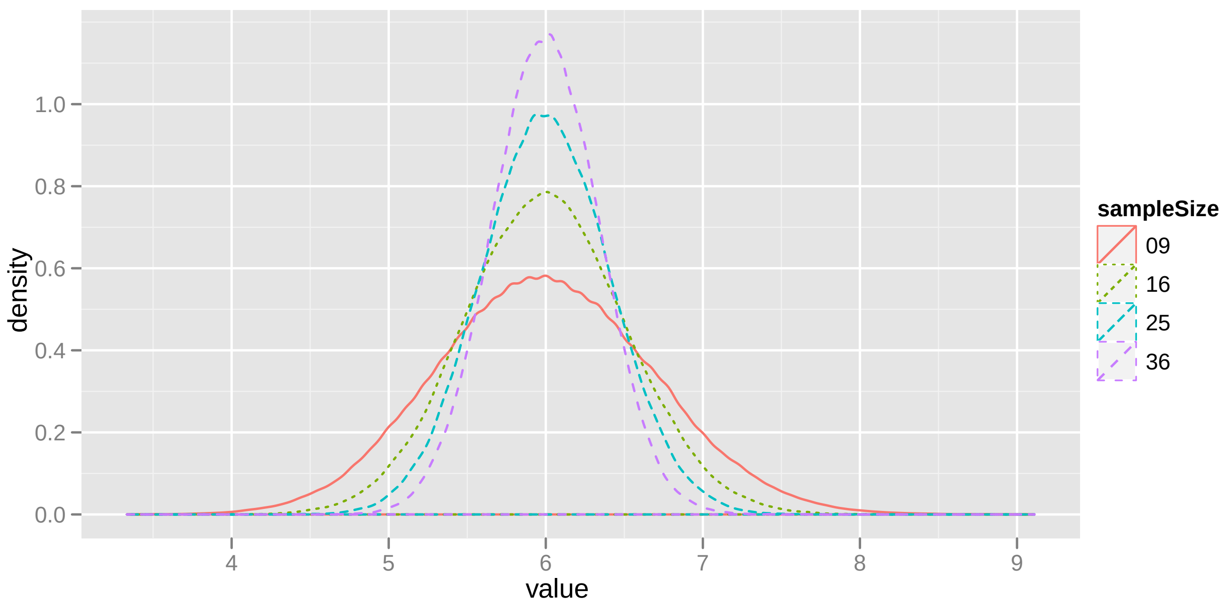

Figure 7: Empirical density plot of the sample means from samples of sizes 9, 16, 25 and 36 from a binomial distribution with $n = 20$ and $p = 0.3$. Each empirical density is based on 100,000 replications.

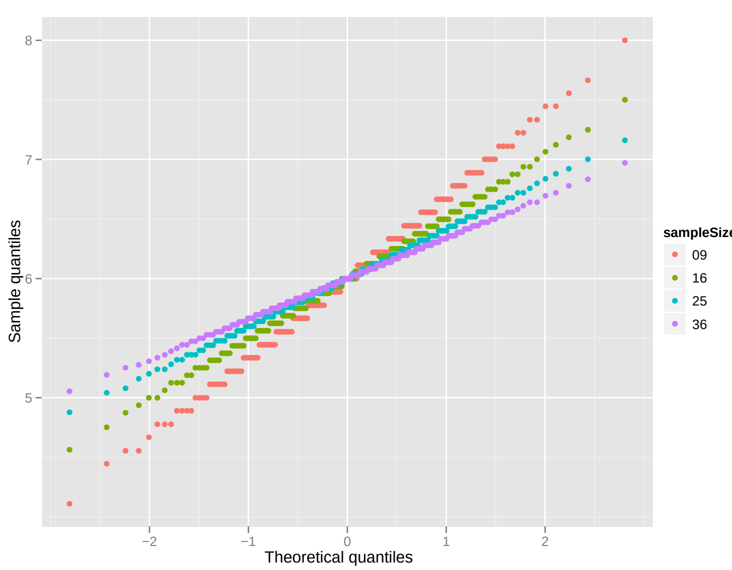

Figure 8: Comparative normal quantile-quantile plots based on 200 equally spaced quantiles of the means of samples of size 9, 16, 25 and 36 from a binomial distribution with $n = 20$ and $p = 0.3$.

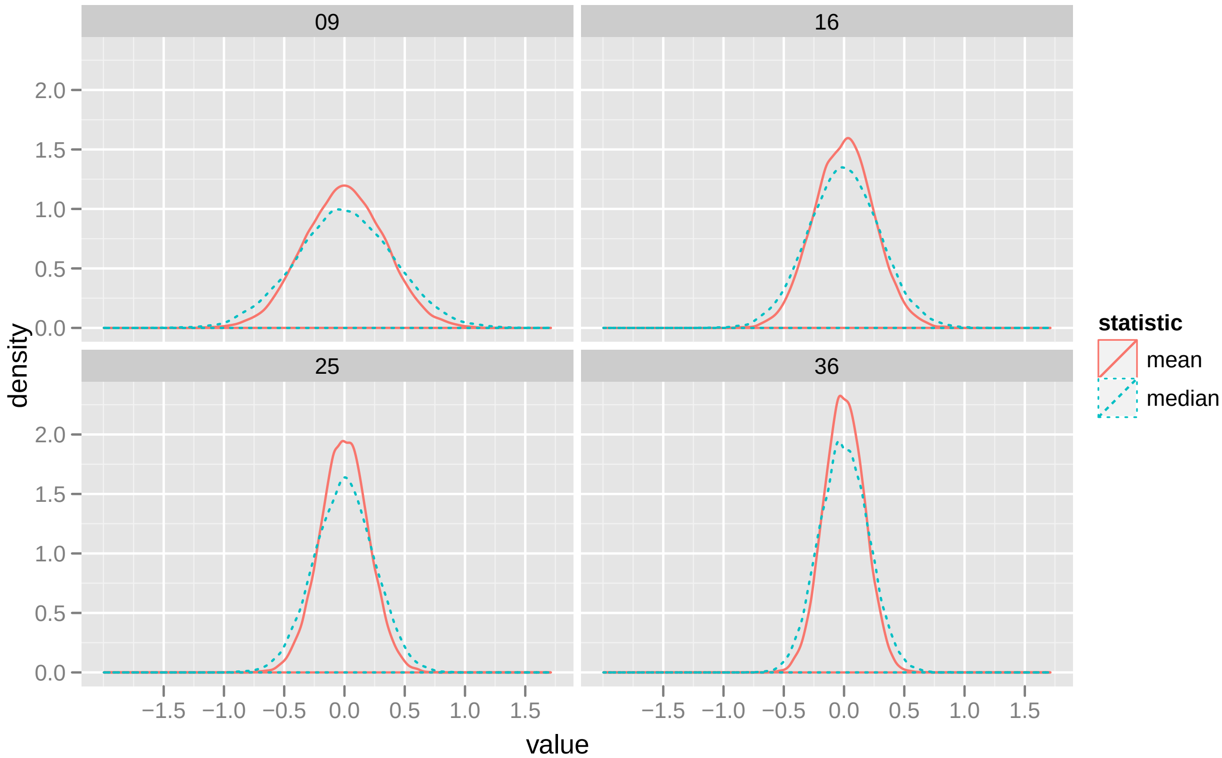

Figure 9: Comparative density plots of the distribution of the sample mean and the sample standard deviation for samples of size 9, 16, 25 and 36 from a standard normal distribution. Each empirical density curve is based on 20,000 replications.

		sample mean		sample median	
	n	lower	upper	lower	upper
80%	9	-0.42761	0.42357	-0.51914	0.52274
	16	-0.31999	0.31702	-0.37951	0.38037
	25	-0.25803	0.25514	-0.32082	0.32049
	36	-0.21216	0.21592	-0.25956	0.26412
90%	9	-0.54564	0.54530	-0.67009	0.66976
	16	-0.40619	0.41237	-0.48649	0.48980
	25	-0.33236	0.32732	-0.41548	0.40726
	36	-0.27052	0.27750	-0.32914	0.34058
95%	9	-0.64786	0.64833	-0.79846	0.79941
	16	-0.48075	0.49193	-0.58238	0.59419
	25	-0.40202	0.39309	-0.49137	0.48181
	36	-0.32380	0.32683	-0.39232	0.40795
99%	9	-0.85702	0.86405	-1.03449	1.07465
	16	-0.63180	0.65044	-0.74862	0.77873
	25	-0.52758	0.51456	-0.65155	0.64146
	36	-0.42058	0.42392	-0.50900	0.54059

Table 6: Empirical confidence intervals for the population mean using the sample mean or the sample median as the estimator.