

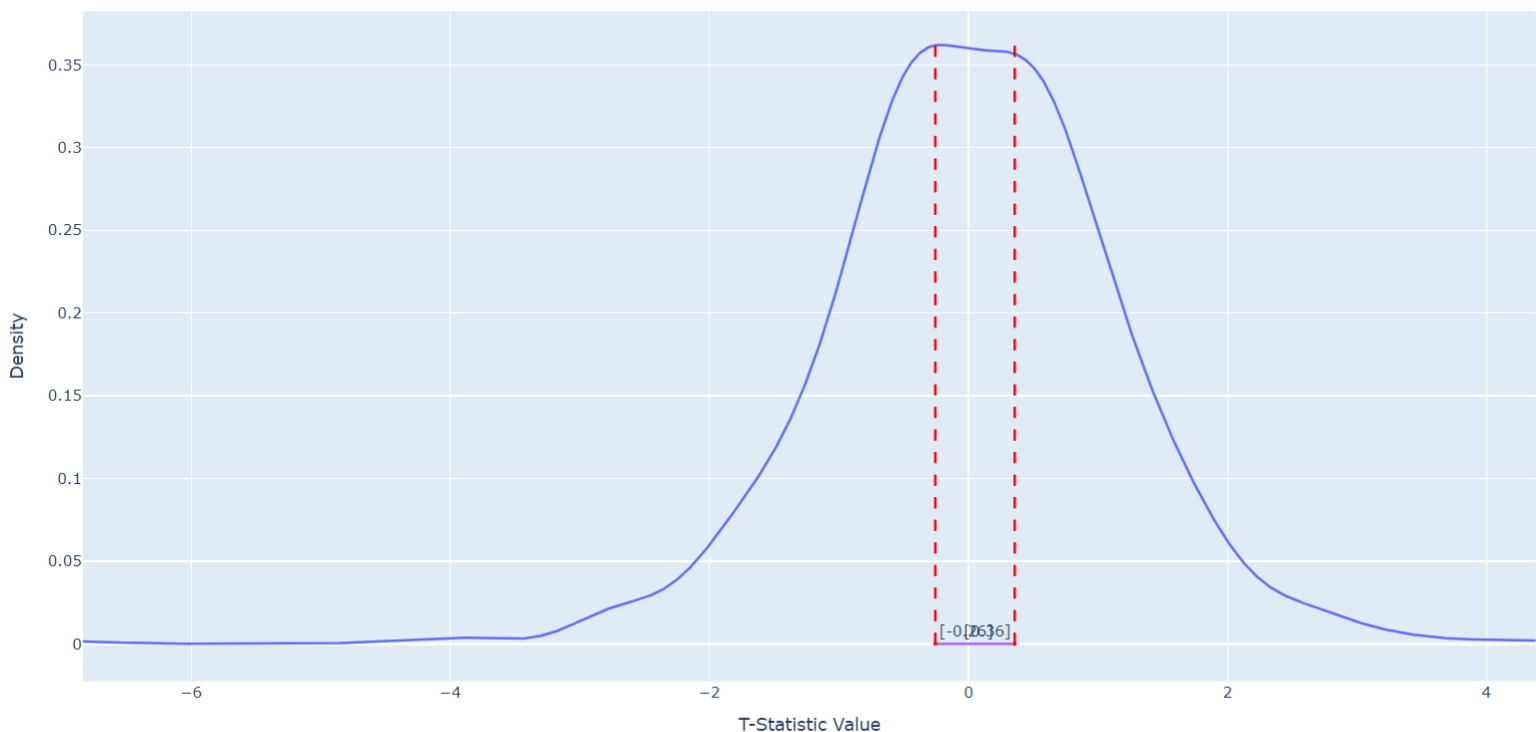
Plots of bootstrap-t distributions, along with the confidence intervals for the true mean of the unknown underlying distributions:

Provided the following details along each distribution:

- A. 5th and 95th percentile: These percentile points are w.r.t. the bootstrap-t table formed from the given sample.
 - B. Underlying distribution's statistical measures, **to understand the parent distribution from where the data has been sampled.**
 - C. The sample's statistical measures show the quality of the sample, i.e., **how appropriate the representation of the given sample is of the true distribution.**
-

1. Uniform Distribution

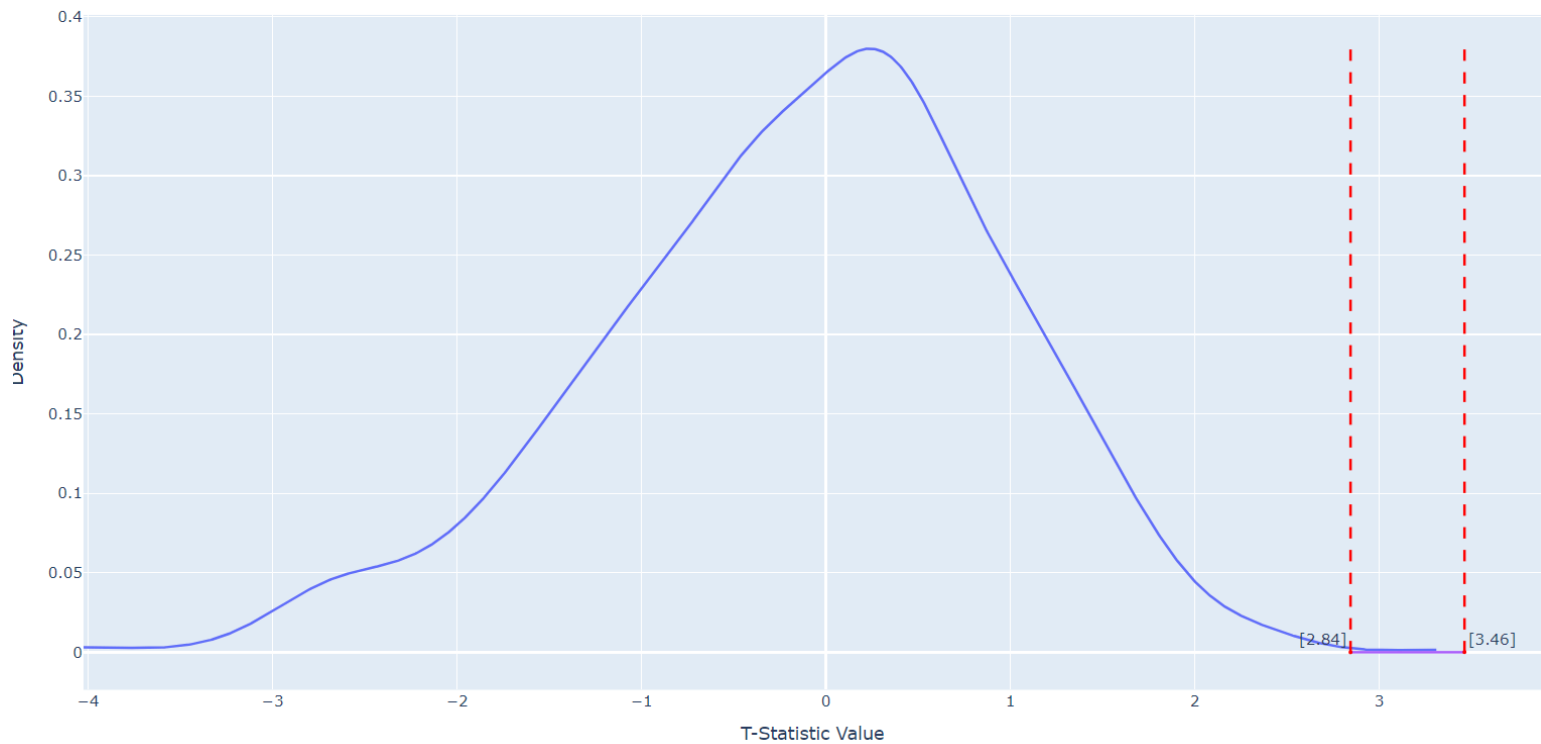
Bootstrap T-Statistic Distribution with 95% Confidence Interval



- a. 5th percentile at -1.80 and 95th percentile at 1.74.
- b. 95% Confidence Interval (Uniform): `[np.float64(0.3575694198028659), np.float64(-0.255137324924904)]`
- c. Underlying Distribution's Mean = 0.
Range = $[-\pi/3, +\pi/3]$.
- d. Sample's [Mean, Median, Variance] = [0.04, 0.09, 0.51]

2. Gamma Distribution

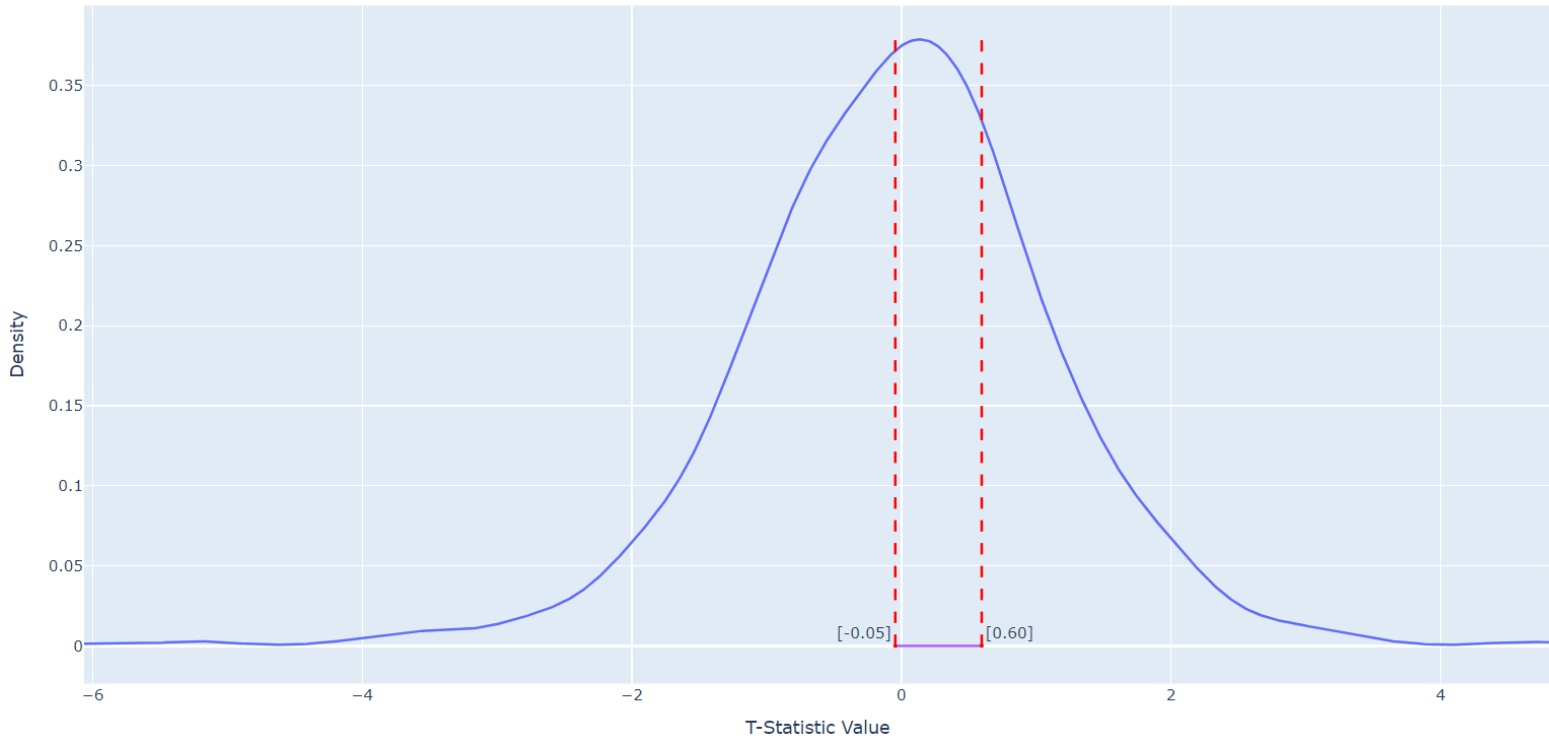
Bootstrap T-Statistic Distribution with 95% Confidence Interval (Input File : input_files/gamma_df_15.txt)



- a. 5th percentile at -2.07 and 95th percentile at 1.50.
- b. 95% Confidence Interval (Gamma): `[np.float64(2.8443684741263158), np.float64(3.462088516151728)]`
- c. *Underlying Distribution's* `[shape, scale] = [2, 2]`;
- d. *Sample's* `[Mean, Median, Variance] = [3.2, 2.9, 3.9]`

3. Beta Distribution

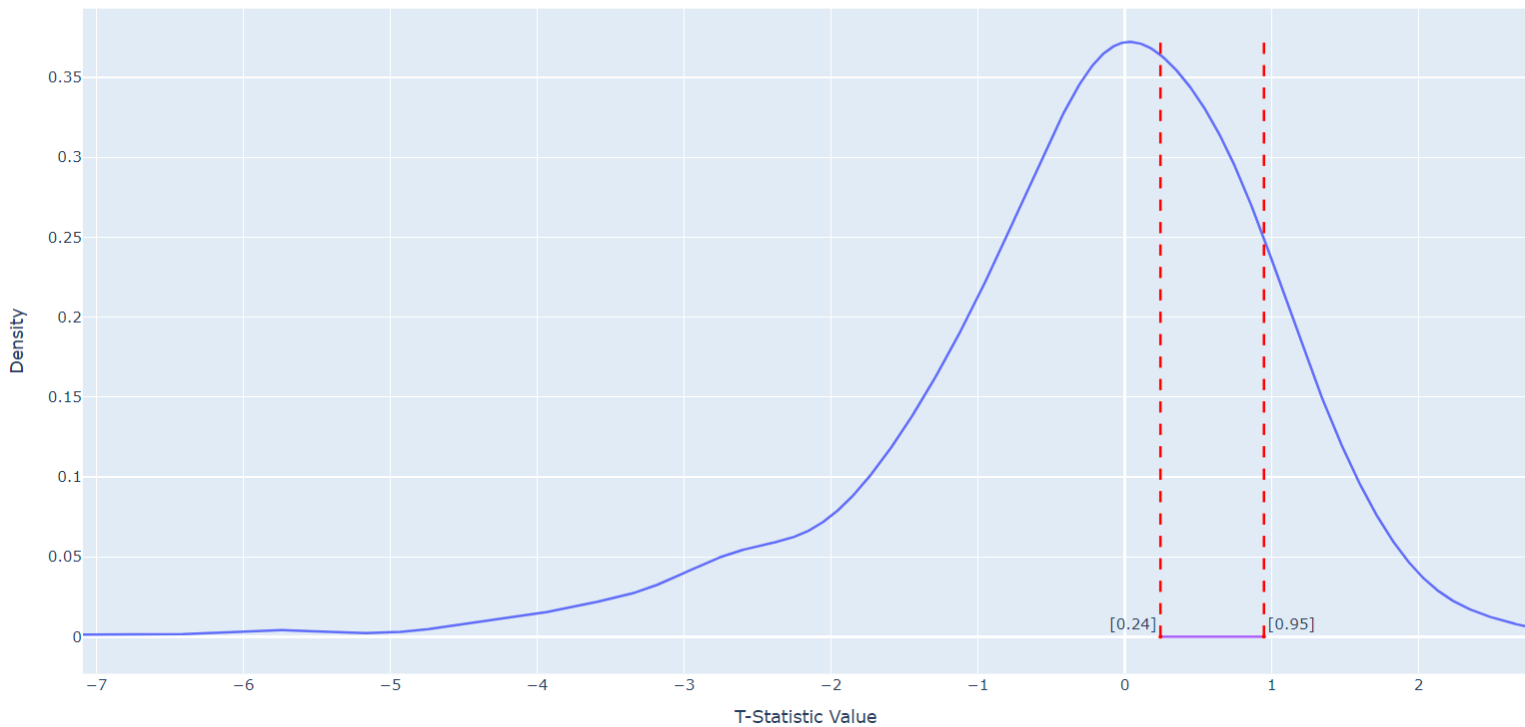
Bootstrap T-Statistic Distribution with 95% Confidence Interval (Input File : input_files/beta_df_15.txt)



- a. 5th percentile at -1.87 and 95th percentile at 1.83.
- b. 95% Confidence Interval (Beta): [np.float64(-0.04530380710104531), np.float64(0.5955989796652122)]
- c. Underlying Distribution's $[\alpha, \beta] = [2, 5]$;
- d. Sample's $[\text{Mean}, \text{Median}, \text{Variance}] = [0.27, 0.3, 0.02]$

4. Exponential Distribution

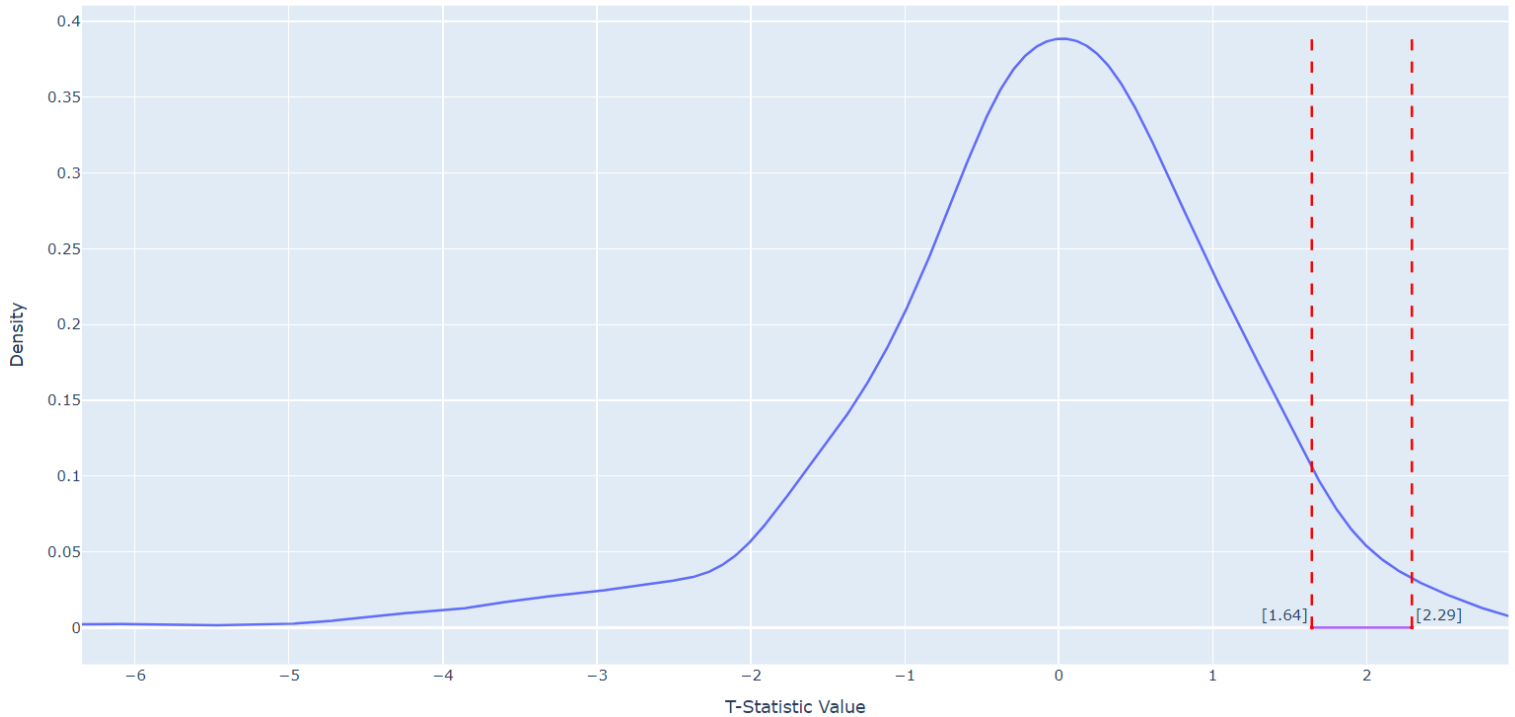
Bootstrap T-Statistic Distribution with 95% Confidence Interval (Input File : input_files/exponential_df_15.txt)



- a. 5th percentile at -2.65 and 95th percentile at 1.42.
- b. 95% Confidence Interval (Beta): `[np.float64(0.24341994753786467), np.float64(0.9479336020248816)]`
- c. *Underlying Distribution's $[\lambda] = [1]$;*
- d. *Sample's $[\text{Mean}, \text{Median}, \text{Variance}] = [0.7, 0.58, 0.42]$*
- e. *Mean can't be -ve for exponential functions.*

5. Log-Normal Distribution

Bootstrap T-Statistic Distribution with 95% Confidence Interval (Input File : input_files/log_normal_df_15.txt)



- a. 5th percentile at -2.19 and 95th percentile at 1.57.
 - b. 95% Confidence Interval (Beta): `[np.float64(1.643682488681918), np.float64(2.293945481030995)]`
 - c. *Underlying Distribution's [Mean, Variance] = [0, 1];*
 - d. *Sample's [Mean, Median, Variance] = [2.01, 1.404, 2.8]*
 - e. *Mean can't be -ve for log-normal functions.*
-

RESULTS :

1. The bootstrap-t gives somewhat erratic results and can be heavily influenced by a few outlying data points.
2. Bootstrap-t pdf is not the best representation of the true unknown underlying distribution due to the small sample size and outliers present in it.
3. The confidence interval is about the original sample's mean, which is greater than 0 in most of the cases above; hence, the confidence interval dominates towards the right side in the bootstrap-t pdf.
4. The bootstrap-t method produces unreliable plots for complex underlying distributions (which can be seen above) and complicated statistics (ex - correlation coefficient).
5. The formula for the confidence interval of the true mean is:

$$(\hat{\theta} - \hat{t}^{(1-\alpha)} \cdot \hat{se}, \hat{\theta} - \hat{t}^{(\alpha)} \cdot \hat{se}).$$

where,

$\hat{\theta}$ represents the original sample's mean

\hat{t} represents the derived bootstrap-t distribution

α represents the $(100 \times \alpha)$ th percentile point of the bootstrap-t distribution.

\hat{se} represents the standard error associated with the bootstrapped samples
