

# Math 445

Exploring the robustness of t-based procedures

# Non-normality

If the population distribution is not normal, are the t-procedures valid?

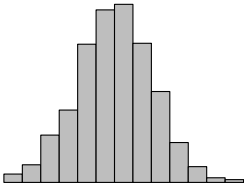
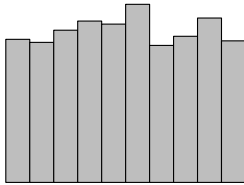
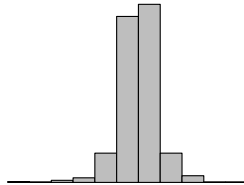
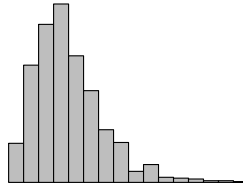
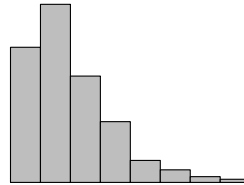
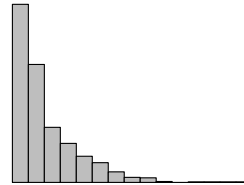
If the t-procedures “perform well” even if some of the assumptions under which they were developed do not hold, then they will be called **robust**.

# Simulation

To check whether a procedure is robust, it is relatively easy to use simulation:

1. Simulate data from a variety of different probability distributions.
2. Run the procedure (e.g., build a one-sample t-interval)
3. Compare the results of the procedure to what should have happened. (e.g., for a large number of CIs, approximately 95% of 95% CIs should capture the parameter value)

% of **one-sample** 95% CIs that are successful when population is non-normal.

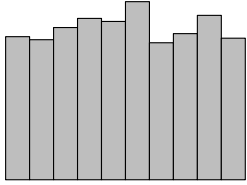
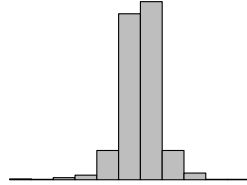
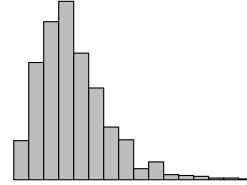
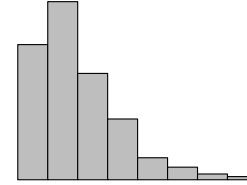
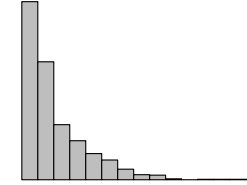
	Bell-shaped	Short-tailed	Long-tailed	Mild Skew	Moderate Skew	Strong Skew
n						
5	95.3	94	96.3	91.6	91.8	89.8
10	95.9	94	96.3	93.3	93.2	90.8
25	95.3	95.4	95.9	93.8	93.5	90.3
50	94.8	94.3	96.3	94.1	94	93.8
100	95.3	95.7	94.9	95.1	95.9	94.6

# One-sample t-procedures

Robustness against departures from Normality:

1. If the population distribution is roughly symmetric and unimodal, then the procedure works well for sample sizes of at least 10–15 (just a rough guide)
2. For skewed population distributions, the t-procedure can be substantially affected, depending on the severity of the skew and the sample size.
3. t-procedures are not resistant to outliers.
4. If observations are not independent, the results can be misleading.

% of 95% **two-sample pooled CIs** that are successful when the two populations are non-normal (with same shape, SD, and sample sizes)

	Short-tailed 	Long-tailed 	Mild Skew 	Moderate Skew 	Strong Skew 
n					
5	94.5	98.3	95.2	95.4	95.5
10	94.6	98.3	95.2	95.4	95.5
25	94.9	98.2	95.1	95.3	95.3
50	95.2	98.1	95.1	95.3	95.1
100	95.6	98	95	95.3	94.8

# Two-sample pooled-procedures

Robustness against departures from Normality:

1. If the populations have equal variances and approx. the same shapes, and if sample sizes approx. equal, then moderately affected by heavy-tails and skew.
2. If the populations have equal variances and approx. same shape, but rather different sample sizes, then moderately affected by heavy-tails and substantially by skew.
3. If the skewness of the two populations differs considerably, the tools can be very misleading with small and moderate sample sizes.

# Unequal Variances

If the sample variances are not equal, how does the pooled t-procedure perform?



% of 95% **two-sample pooled CIs** that are successful when the two populations have different SDs, but are normal with possible different sample sizes.

		$\sigma_2/\sigma_1$				
n1	n2	1/4	1/2	1	2	4
10	10	95.2	94.2	94.7	95.2	94.5
10	20	83	89.3	94.4	98.7	99.1
10	40	71	82.6	95.2	99.5	99.9
100	100	94.8	96.2	95.4	95.3	95.1
100	200	86.5	88.3	94.8	98.8	99.4
100	400	72.6	81.5	95	99.5	99.9