Within Laboratory Variance Outlier Detection: An Alternative to Cochran's Test

Michael J. Morton

Altria
Altria Client Services

Abstract No. 327279

Interlaboratory Collaborative Studies

- Collaborative studies are an important tool in the characterization of the variation of analytical methods
- An important component of the variation calculation is the determination of outlying laboratories
- Standard practice in the evaluation of collaborative studies is to follow ISO 5725(2)
  - This includes the use of Cochran's Test to determine laboratories with excessively large within laboratory variation

Purpose of this Work

- The standard F-test for comparing variances is known to be sensitive to deviations from normality
- Cochran's Test is built from the F-test using a Bonferroni adjustment, so would be expected to be as or more sensitive to deviations from normality as the standard F-Test

Purpose:

- Gauge the sensitivity of Cochran's Test to deviations from normality
- Evaluate an alternative approach based on Levene's Test for use in the evaluation of collaborative studies

Cochran's Test

- Calculate all of the sample standard deviations for each laboratory

- Calculate $c = \frac{s_{largest}^2}{\sum_1^n s_i^2}$

- If c is larger than the critical value, drop the laboratory with the largest standard deviation as an outlier
  - Frequently this process is run iteratively with the remaining laboratories to determine if any of the remaining laboratories are outliers and is continued until no outliers are found

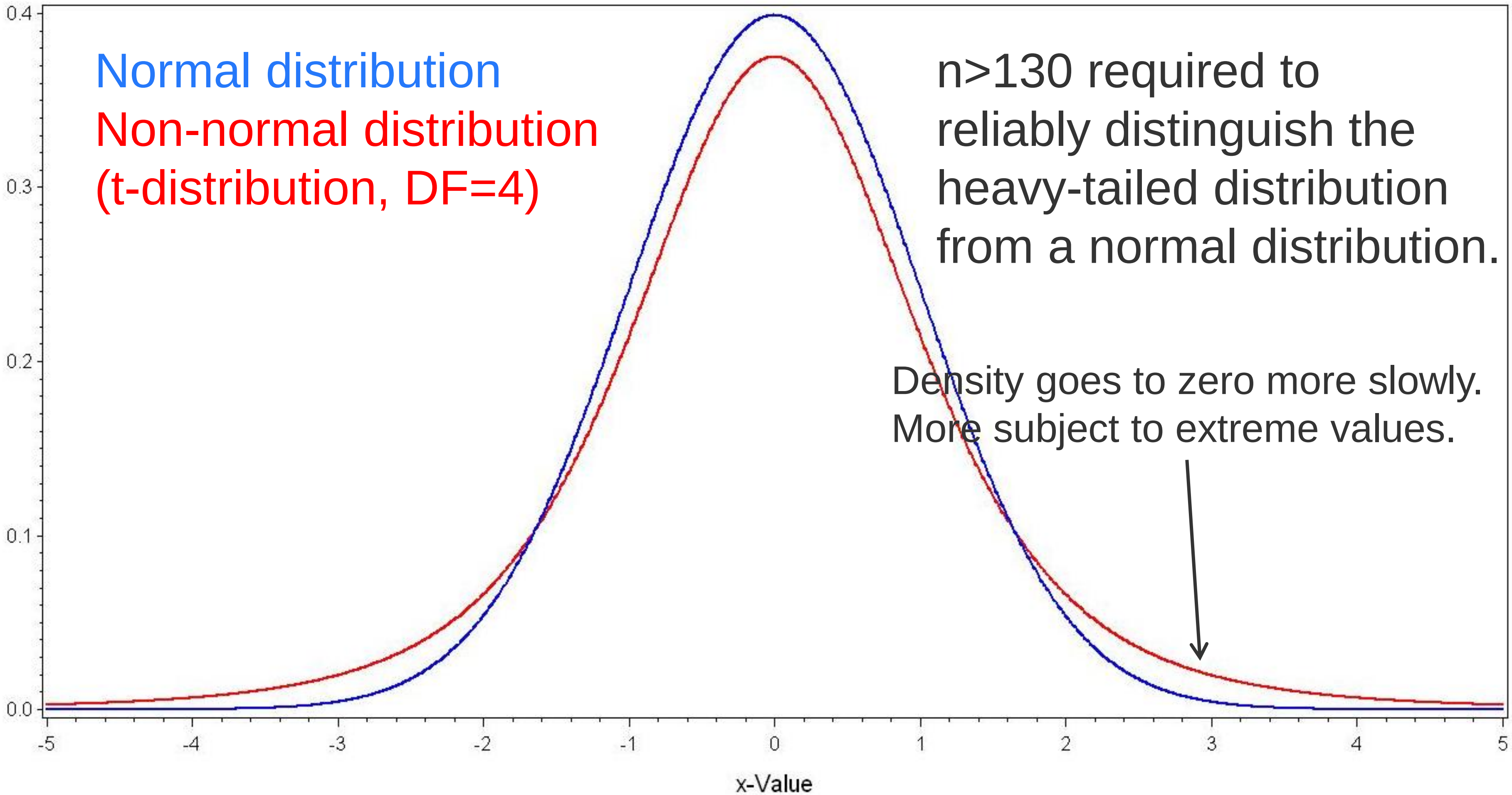
- Cochran's Test is equivalent to a Bonferroni-adjusted maximum F-test, so would be expected to be more sensitive to non-normality than the customary F-test for comparing two variances, since it relies on lower probability tails of the F-distribution.

Deviations from the Normal Distribution

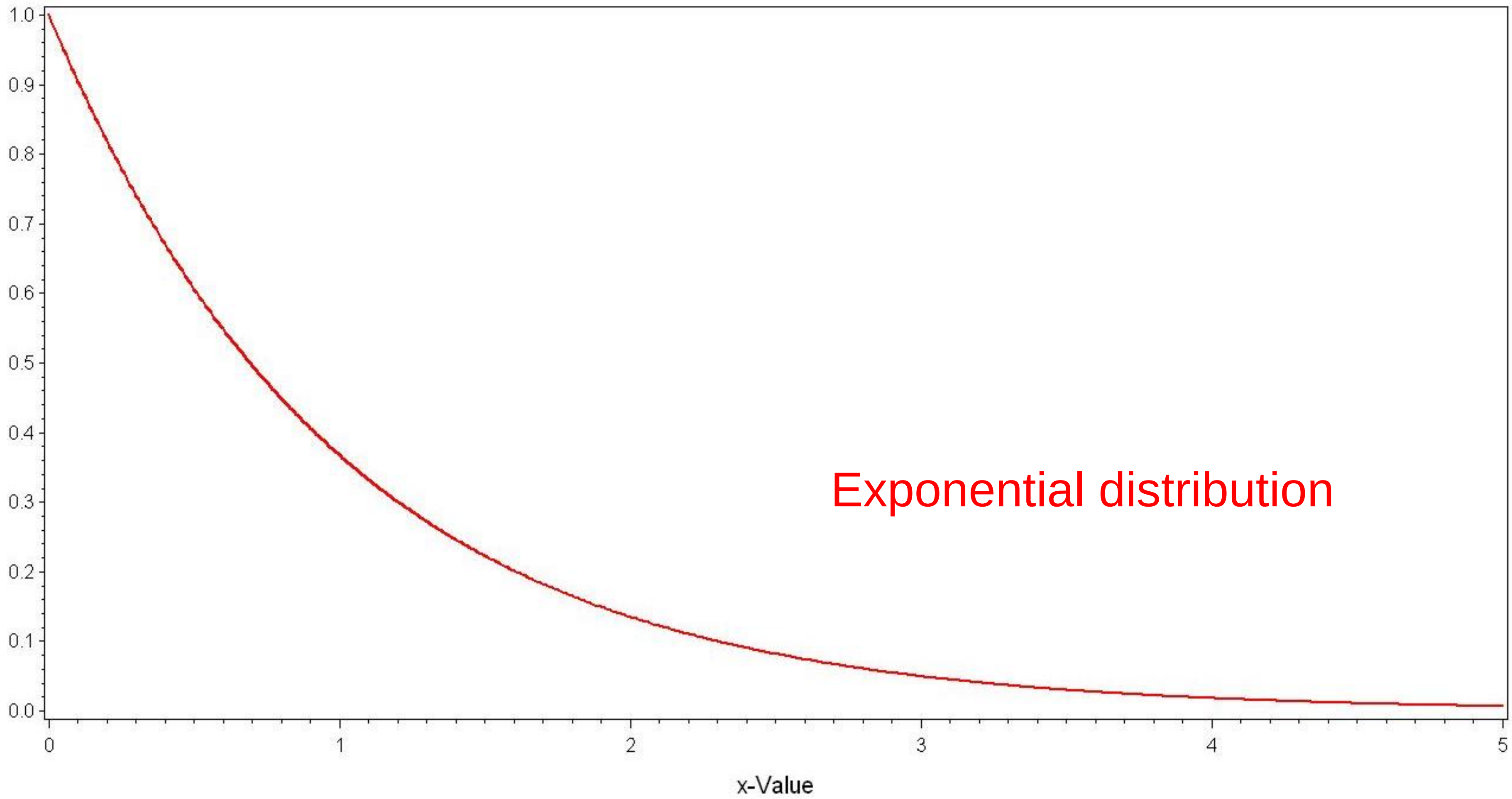
- "Heavy tailed" distribution
  - Subject to large deviations from the mean
  - t-distribution with 4 degrees of freedom

- Skewed distribution
  - Larger deviations in one direction than the other
  - Exponential Distribution

Comparative graph of distributions ("heavy-tailed" alternative)



Skewed distribution as an alternative



Cochran's Test Performance (α=1%)

% of cases Cochran's Test is Significant Based on Simulation (# Labs =10)			
# Reps	Normal	Heavy Tailed	Skewed
10	1.0%	28%	34%
5	1.0%	19%	24%
3	1.0%	12%	16%

Levels should be near 1%. t-distribution represents "heavy tails." The exponential distribution represents skewness.

Different Flavors of Levene's Test

- Looks at the deviations within each lab $d_i = |x_i - \hat{x}|$
  - d_i^2 and $\hat{x}$ the mean – **L2**
  - d_i and $\hat{x}$ the mean – **L1**
  - d_i and $\hat{x}$ the median – **Lmed** (Brown and Forsythe)
  - d_i and $\hat{x}$ the median but drop the median point when number of reps is odd ("structural zero") – **L0** (Hines and Hines)

- Use one way analysis of variance to determine if there are lab-to-lab differences in magnitude of deviations

Comparison of Levene's Test Alternatives

		L2	L1	Lmed	L0
No. Labs	No. Reps				
		1%	1%	1%	1%
Normal Distribution					
10	3	0.1%	5.9%	0.0%	0.6%
10	4	4.2%	5.3%	1.9%	1.9%
10	5	3.2%	3.5%	0.0%	0.6%
10	10	1.7%	2.1%	0.5%	0.5%
Heavy-tailed distribution (t with 4 degrees of freedom)					
10	3	4.5%	22.6%	0.0%	1.2%
10	4	4.5%	13.1%	2.3%	2.3%
10	5	3.4%	8.6%	0.0%	1.0%
10	10	2.0%	3.9%	0.7%	0.7%
Skewed distribution (exponential)					
10	3	5.7%	47.6%	0.0%	0.9%
10	4	9.3%	33.2%	5.6%	5.6%
10	5	8.5%	26.8%	0.2%	1.3%
10	10	4.4%	22.2%	1.1%	1.1%

Comparison of Levene's Test Alternatives

		L2	L1	Lmed	L0
No. Labs	No. Reps				
		1%	1%	1%	1%
Normal Distribution					
10	3	0.1%	5.9%	0.0%	0.6%
10	4	4.2%	5.3%	1.9%	1.9%
10	5	3.2%	3.5%	0.0%	0.6%
10	10	1.7%	2.1%	0.5%	0.5%
Heavy-tailed distribution (t with 4 degrees of freedom)					
10	3	4.5%	22.6%	0.0%	1.2%
10	4	4.5%	13.1%	2.3%	2.3%
10	5	3.4%	8.6%	0.0%	1.0%
10	10	2.0%	3.9%	0.7%	0.7%
Skewed distribution (exponential)					
10	3	5.7%	47.6%	0.0%	0.9%
10	4	9.3%	33.2%	5.6%	5.6%
10	5	8.5%	26.8%	0.2%	1.3%
10	10	4.4%	22.2%	1.1%	1.1%

Comparison of Levene's Test Alternatives

		L2	L1	Lmed	L0
No. Labs	No. Reps				
		1%	1%	1%	1%
Normal Distribution					
10	3	0.1%	5.9%	0.0%	0.6%
10	4	4.2%	5.3%	1.9%	1.9%
10	5	3.2%	3.5%	0.0%	0.6%
10	10	1.7%	2.1%	0.5%	0.5%
Heavy-tailed distribution (t with 4 degrees of freedom)					
10	3	4.5%	22.6%	0.0%	1.2%
10	4	4.5%	13.1%	2.3%	2.3%
10	5	3.4%	8.6%	0.0%	1.0%
10	10	2.0%	3.9%	0.7%	0.7%
Skewed distribution (exponential)					
10	3	5.7%	47.6%	0.0%	0.9%
10	4	9.3%	33.2%	5.6%	5.6%
10	5	8.5%	26.8%	0.2%	1.3%
10	10	4.4%	22.2%	1.1%	1.1%

Alternative to Cochran's Test

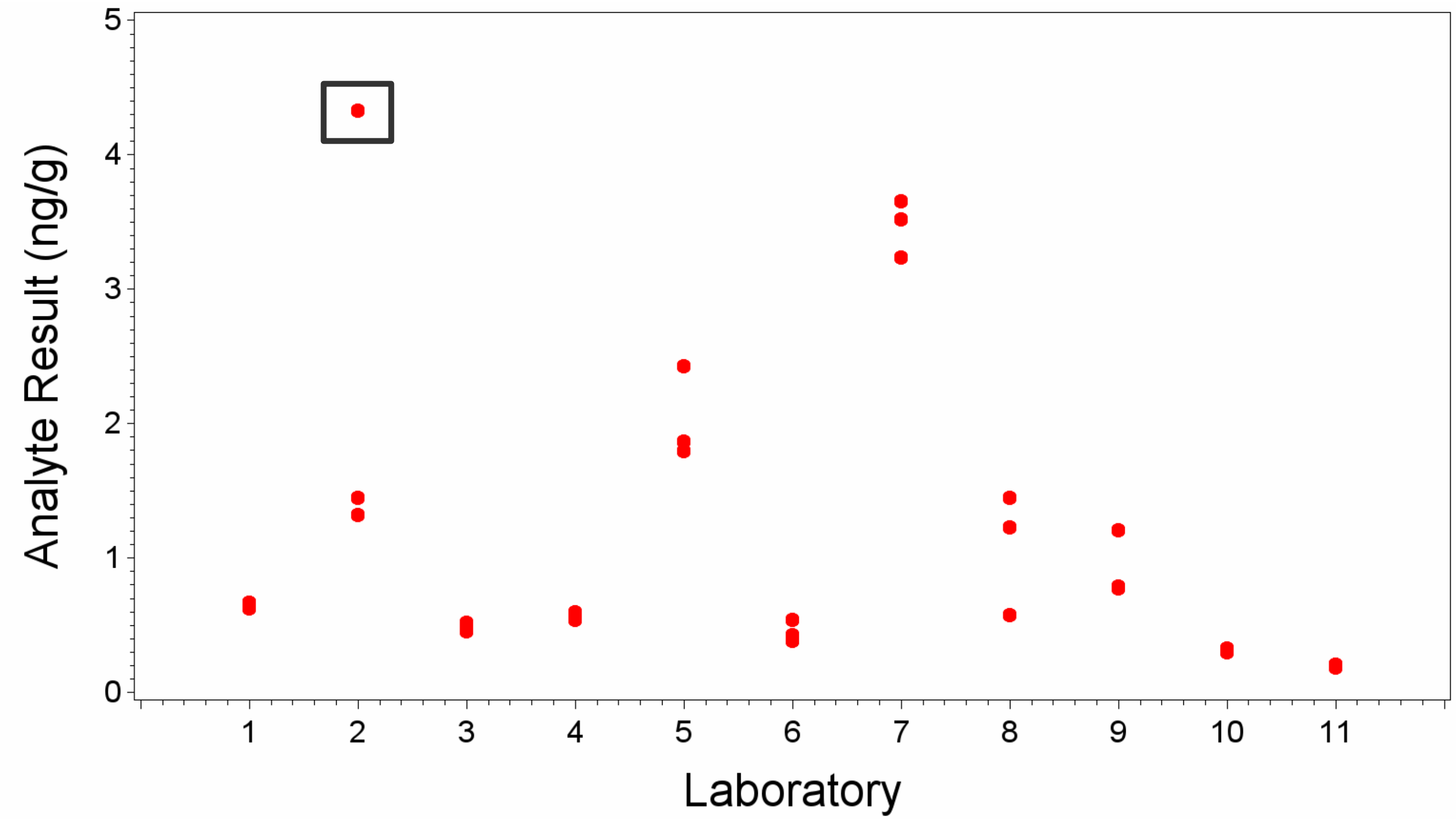
- Based on L0, the best performing of the Levene's Test alternatives

- Judge a laboratory an outlier if:
  1. L0 is significant at the nominal α=0.02
  2. The lab with the biggest deviations is significantly greater than zero as a one-sided Bonferroni adjusted linear contrast (α=0.01)
     - i.e., $\max_i \left\{ \overline{d}_i - \frac{1}{n-1} \sum_{j \neq i} \overline{d}_j \right\}$ is significantly greater than zero with a one-sided Bonferroni adjustment

Comparison of Cochran's and Alternative Approach

% of cases Cochran's Test is Significant Based on Simulation (# Labs =10)			
# Reps	Normal	Heavy Tailed	Skewed
10	1.0%	28%	34%
5	1.0%	19%	24%
3	1.0%	12%	16%
Alternative Approach			
# Reps	Normal	Heavy Tailed	Skewed
10	0.5%	1.0%	1.5%
5	0.7%	1.4%	1.9%
3	0.8%	1.9%	1.1%

Caveat with single-point outliers



The proposed approach is subject to error when a single outlier is "too large" because the analysis of variance tends to lose statistical significance. This is a general problem with Levene's Test.

Summary

- Estimating method variability is an important aspect of characterizing an analytical method
- Estimated variability is typically calculated from collaborative studies and is strongly affected by outlier determination
  - Unnecessarily dropping laboratories may give an unrealistic picture of intra-lab variation and sacrifices valuable information on lab-to-lab reproducibility
- Deviations from normality that are very difficult to detect have a dramatic effect on Cochran’s Test
- An alternative approach adapted from Levene’s Test was proposed and shown to be much more robust than Cochran’s Test
  - A caveat is that it is subject to not detect a single large outlying observation
  - Single large outliers must be dealt with “manually” when present

References

- Morton, Michael J., 2017, Beitrage zur Tabakforschung International, Vol 27, No 7, September 2017: 135-144.
- Levene, H. (1960), “Robust Tests for the Equality of Variance,” in I. Olkin, ed., *Contributions to Probability and Statistics*, 278–292, Palo Alto, CA: Stanford University Press.
- Brown, Morton B.; Forsythe, Alan B. (1974), "Robust tests for equality of variances," *Journal of the American Statistical Association* **69**: 364–367.
- Conover, W.J.; Johnson, M.E.; Johnson, M.M. (1981). “A comparative study of tests for homogeneity of variances, with applications to the Outer Continental Shelf bidding data.” Technometrics 23: 351-361.
- Milliken, George A. and Johnson, Dallas E. (1992). Analysis of Messy Data Volume I: Designed Experiments. Chapman & Hall, London, pp 18 ff.
- Hines, W.G.S. and Hines, R.J.O., 2000, “Increased power with modified forms of the Levene (med) test for heterogeneity of variance,” Biometrics, 56, 451-454.