One- and Two-Sided Tolerance Limits for Ratios Adapted for Cost-Effectiveness Analysis

Marilena Flouri

Mathematical Statistician
Division of Biometrics III
CDER/OTS/OB
U.S. Food and Drug Administration

International Conference on Health Policy Statistics
January 11, 2018

FDA Disclaimer

U.S. FOOD & DRUG ADMINISTRATION

This presentation reflects the views of the authors and should not be construed to represent the FDA's views or policies.

This work was done as part of my doctoral dissertation.

Outline

- **Introduction**
  - Motivation
  - Application: Cost-Effectiveness Analysis (CEA)
- **Tolerance limits for ratios of random variables**
  - The bivariate normal case
  - The lognormal/normal case
- **Simulation Study**
- **Example**
- **Discussion**
- **References**

Motivation

U.S. FOOD & DRUG ADMINISTRATION

Statistical inference for ratios is relevant in many applications:

- In health economics: Incremental Cost-Effectiveness Ratio (ICER), Average Cost-Effectiveness Ratio (ACER)

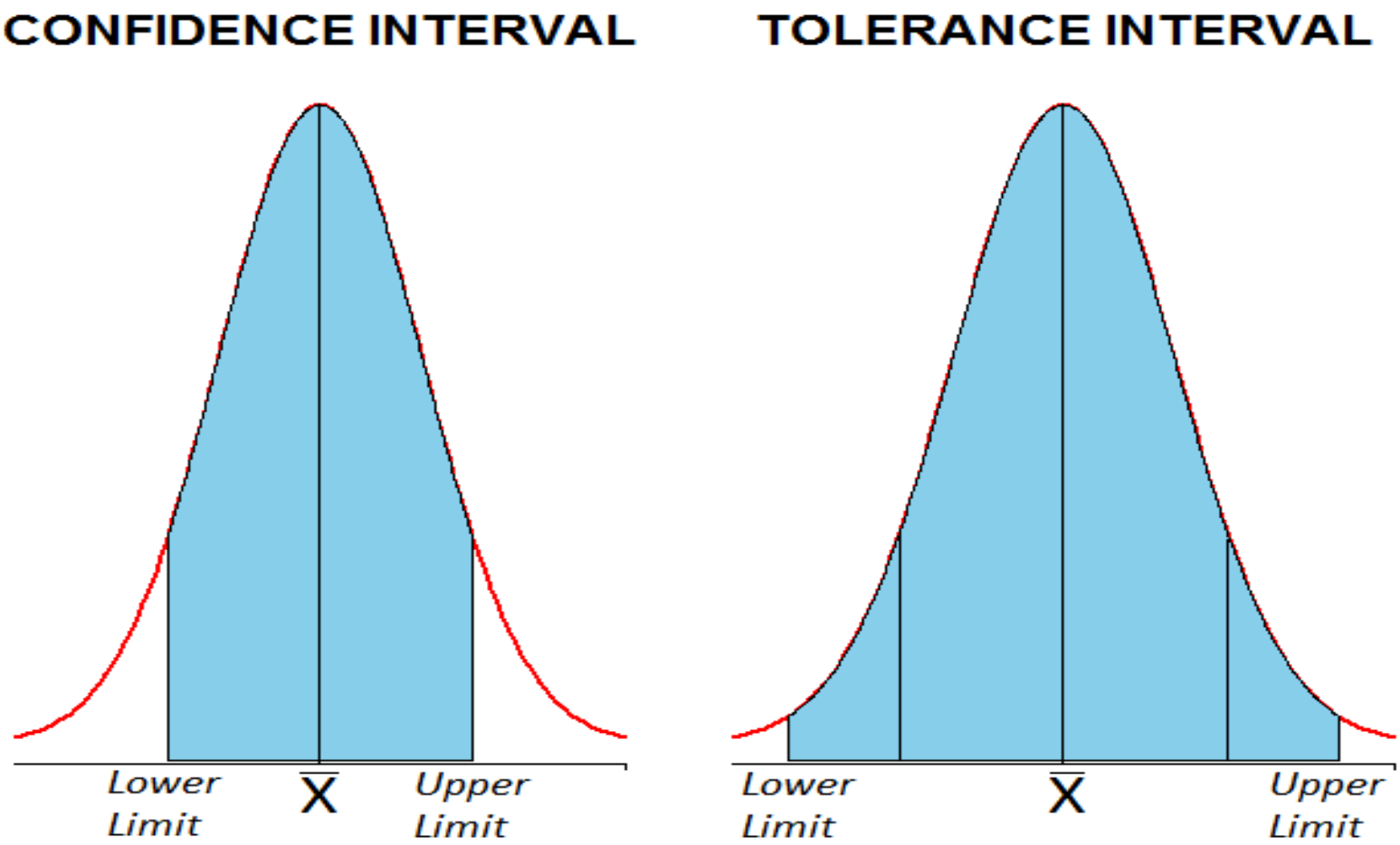
Methodology for Confidence Intervals:

- Fieller's Theorem [see Zerbe (1978)]
- Bootstrap [see Chen (2007)]
- Generalized pivotal quantity (GPQ) [see Bebu et al. (2009, 2016)]

Problem of Interest:
Computation of tolerance intervals for the ratio of two random variables.

What is a Tolerance Interval?

An interval that captures a specified proportion or more of a population, with given confidence level.



Notation

A (p, $1-\alpha$) tolerance interval is a tolerance interval with content p and confidence level $1-\alpha$.

Why Tolerance Intervals?

- Some applications do call for the computation of tolerance limits for the ratio of two random variables
- A single parameter summary is inadequate to describe the distribution of a random variable.

Application: Cost-Effectiveness Analysis (CEA)

U.S. FOOD & DRUG ADMINISTRATION

- ACER $= \frac{\mu_C}{\mu_E}$, μ_C= average cost and μ_E= average effectiveness.
- $X_{\text{ACER}} = \frac{C}{E}$: Random variable for the individual cost-effectiveness ratio.
- One-sided confidence limits for the percentiles can provide information concerning the extremes of the distribution of X_{ACER}.
- Such confidence limits are referred to as **tolerance limits**.

Example (Pharmacological Agents Comparison) *Gardiner et al. (2000)*

- Test drug: 150 patients
- Standard drug: 150 patients

Effectiveness: Measured in Quality-Adjusted-Life-Years (QALYs).

Cost: Measured in U.S. dollars.

Parameter estimates:

	Test drug ($n_1 = 150$)	Standard drug ($n_2 = 150$)
Mean cost	200000	80000
Mean effectiveness	8	5
SD cost	78400	27343
SD effectiveness	2.1	2

Tolerance intervals for ratios of random variables

U.S. FOOD & DRUG ADMINISTRATION

Background history

- Hall and Sampson (1973): ratio of independent normal random variables
- Zhang et al. (2010): ratio of correlated normal random variables
  - Based on the concept of generalized pivotal quantity.
  - Applicable only for computing one-sided tolerance limits.
  - Performance is not satisfactory in terms of maintaining the coverage probability.

The bivariate normal case

Let $\boldsymbol{X} = (X_1, X_2)' \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Need upper tolerance limit for $Y = X_1/X_2$.

Proposed Methodology:

- Generate B parametric bootstrap samples $(X_{1i}^*, X_{2i}^*)' \sim N_2(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$, i=1, 2, ..., B and compute

$$Y_i^* = \frac{X_{1i}^*}{X_{2i}^*}.$$

- Compute a nonparametric upper tolerance limit (based on order statistics) for the distribution of Y, using Y_i^* , i = 1, 2, ..., B.

- Use a bootstrap calibration on the content p in order to improve accuracy.

Basic idea of bootstrap calibration: Use a modified content p, say p_0. To specify p_0, try a sequence of candidate content values and choose the one that gives final coverage closest to $1 - \alpha$.

Proposed Methodology (cont.):

- A second level bootstrap sample is required to implement calibration.
- The calibration also requires the cdf of $Y = X_1/X_2$.
- Use a representation of the exact cdf due to Hinkley (1969).

The methodology can be adapted for one-sided and two-sided tolerance intervals.

The lognormal/normal case

- Cost data are often highly skewed.
- $\boldsymbol{X} = (\ln C, E)' \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
- The proposed methodology can still be applied with some modifications.
- The CDF of Y is estimated via simulation.

Simulation Study

FDA U.S. FOOD & DRUG ADMINISTRATION

Bivariate normal

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ 10 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} 500 & \sigma_{12} \\ \sigma_{21} & 8 \end{pmatrix},$$

where $\sigma_{12} = \rho\sqrt{500 \times 8}$, $\mu_1 = 500, 525$, and $\rho = -0.5, 0.1, 0.5$.

Lognormal-Normal

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ 10 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} 1.6 & \sigma_{12} \\ \sigma_{21} & 8 \end{pmatrix},$$

where $\sigma_{12} = \rho\sqrt{1.6 \times 8}$, $\mu_1 = 8, 10$, and $\rho = -0.5, 0.1, 0.5$.

Pre-specified confidence level $1 - \alpha = 0.95$ and content $p = 0.9$.

Bivariate normal

			With calibration			Without calibration		
n	μ_1	ρ	LTL	UTL	2-sided TL	LTL	UTL	2-sided TL
50	500	-0.5	0.956	0.945	0.955	0.671	0.665	0.633
50	525	-0.5	0.955	0.952	0.938	0.691	0.650	0.628
150	500	-0.5	0.947	0.945	0.948	0.771	0.760	0.731
150	525	-0.5	0.949	0.944	0.939	0.764	0.763	0.755
50	500	0.1	0.945	0.949	0.952	0.642	0.678	0.645
50	525	0.1	0.947	0.953	0.946	0.681	0.660	0.640
150	500	0.1	0.951	0.940	0.960	0.762	0.761	0.776
150	525	0.1	0.948	0.951	0.944	0.789	0.765	0.744
50	525	0.5	0.951	0.958	0.948	0.649	0.655	0.633
50	525	0.5	0.939	0.948	0.954	0.652	0.645	0.622
150	525	0.5	0.957	0.949	0.960	0.756	0.759	0.750
150	525	0.5	0.955	0.946	0.958	0.747	0.766	0.767

Table 1: Estimated coverage probabilities of lower tolerance limits (LTLs), upper tolerance limits (UTLs) and two-sided tolerance intervals (2-sided TLs) corresponding to 90% content and 95% confidence level with and without calibration.

Example: Pharmacological agents comparison *Gardiner et al. (2000)*

U.S. FOOD & DRUG ADMINISTRATION

	Tolerance limits	Confidence limits
Test drug (lower limit)	14300	24900
Standard drug (lower limit)	10000	15900
Test drug (upper limit)	35600	25100
Standard drug (upper limit)	27800	16000

Table 2: (0.9,0.95) one-sided tolerance limits for the C/E ratio and 95% confidence limits for μ_C/μ_E.

Test Drug vs. Standard drug

- Larger tolerance and confidence limits for test drug
- C/E stochastically larger under the test drug (see Figure 1)

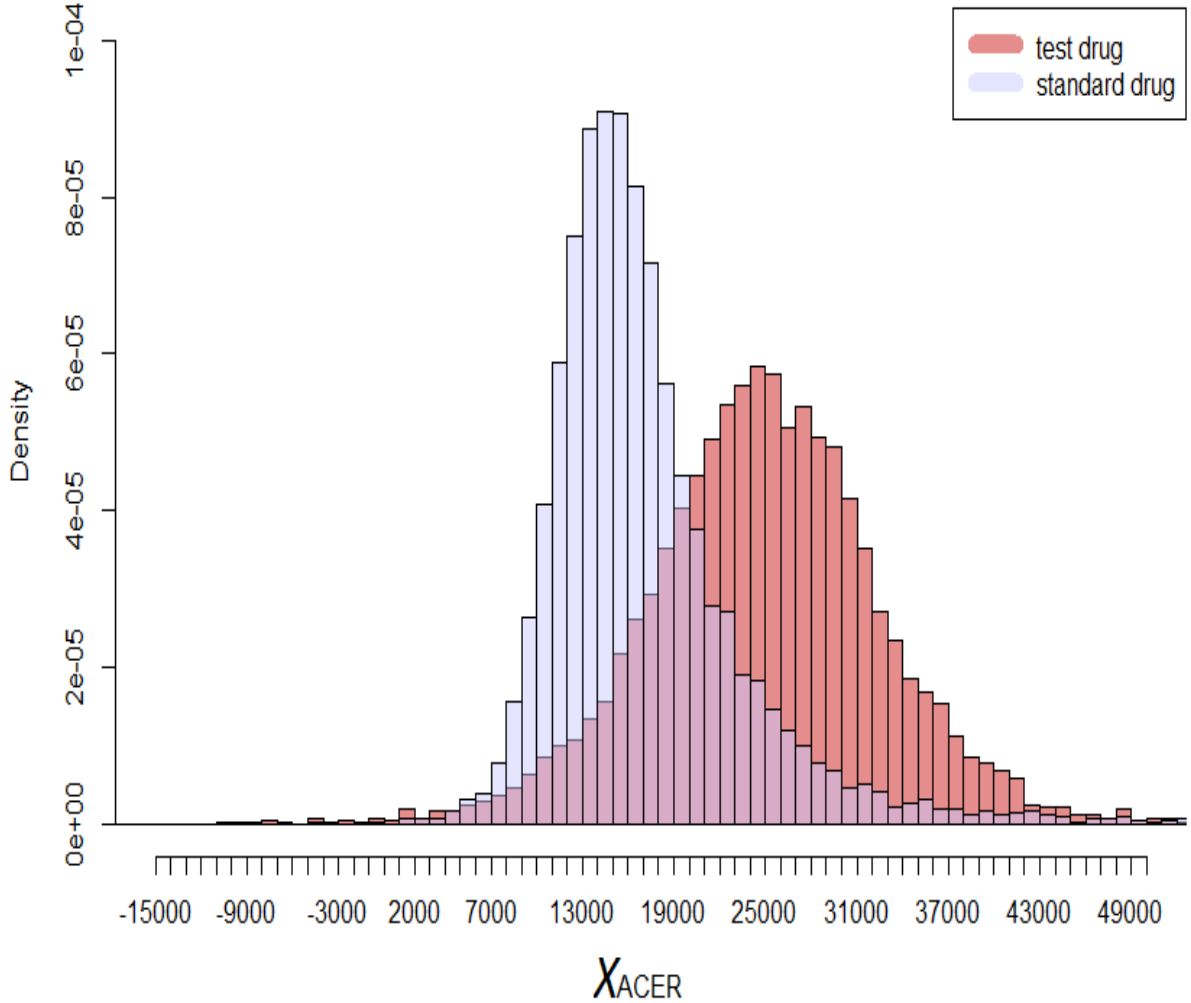


Figure 1: A plot of the distributions of C/E using simulated data from the bivariate normal distributions for $(C, E)'$ with the unknown parameters replaced by the estimates.

	Tolerance intervals	Confidence intervals
Test drug	(9400, 40000)	(24900, 25000)
Standard drug	(8500, 39700)	(15900, 16000)

Table 3: (0.9, 0.95) two-sided tolerance intervals for the C/E ratio and 95% two-sided confidence intervals for μ_C/μ_E.

Tolerance interval (TI) vs. Confidence Interval (CI)

- TI: We are 95% confident that at least 90% of the population will have cost-effectiveness ratio between $\sim$\$9400/QALY and $\sim$\$40000/QALY when taking the test drug.
- CI: We are 95% confident that the average cost-effectiveness ratio for the test drug will be between $\sim$\$24900/QALY and $\sim$\$25000/QALY.

*Not exactly the correct scientific interpretation but a simple way to compare the two intervals.

Suppose the willingness-to-pay amount is $25000/QALY

Tolerance limits	Test Drug	Standard Drug
(**0.015**,0.95) lower tolerance limit	-	35400
(**0.050**,0.95) lower tolerance limit	35600	-

- At least 5% of the population taking the **test drug** are expected to have cost $>$\$35000/QALY.
- At least 1.5% of the population taking the **standard drug** are expected to have cost $>$\$35000/QALY.

Discussion

U.S. FOOD & DRUG ADMINISTRATION

Tolerance limits and tolerance intervals for ratios of random variables:

- have not been explored in Cost-Effectiveness Analysis;
- can supplement the traditional analysis based on confidence limits.

Proposed Methodology

- Satisfactory performance based on simulations.
- Applicable even when the distribution of the ratio random variable is not available in tractable form.

What next?

- Explore more scenarios in Cost-Effectiveness Analysis (e.g., discrete effectiveness and zero-inflated costs).

Marilena Flouri, Shuyan Zhai, Thomas Mathew, Ionut Bebu. Tolerance limits and tolerance intervals for ratios of normal random variables using a bootstrap calibration. Biometrical journal **59** (2017) 3, 550-566

DOI: 10.1002/bimj.201600117

References

I. Bebu, F. Seillier-Moiseiwitsch, and T. Mathew. Generalized confidence intervals for ratios of regression coefficients with applications to bioassays. *Biometrical Journal*, 1:1047–1058, 2009.

I. Bebu, G. Luta, T. Mathew, P. A. Kennedy, and B. K. Agan. Parametric cost-effectiveness inference with skewed data. *Computational Statistics and Data Analysis*, 94:210–220, 2016.

D.G. Chen. Bootstrapping estimation for estimating relative potency in combinations of bioassays. *Computational Statistics & Data Analysis*, 51:4597 – 4604, 2007.

B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. CRC Press, 1993.

J. C. Gardiner, M. Huebner, J. Jetton, and C. J. Bradley. Power and sample size assessments for test of hypotheses on cost-effectiveness ratios. *Health Economics*, 9:227–234, 2000.

I. J. Hall and C. B. Sampson. Tolerance limits for the distribution of the product and quotient of normal variates. *Biometrics*, 29:109–119, 1973.

D. V. Hinkley. On the ratio of two correlated normal random variables. *Biometrika*, 56:635–639, 1969.

K. Krishnamoorthy and T. Mathew. *Statistical Tolerance Regions: Theory, Applications, and Computation*. John Wiley & Sons, 2009.

G. O. Zerbe. On fieller's theorem and the general linear model. *The American Statistician*, 32: 103–105, 1978.

L. Zhang, T. Mathew, H. Yang, K. Krishnamoorthy, and I. Cho. Tolerance limits for a ratio of normal random variables. *Journal of Biopharmaceutical Statistics*, 20:172–184, 2010.