PharmaSUG China 2015 - Paper 57

The Application of Tolerance Interval in Defining Drug Response for Biomarker

Jing Pan, Roche Product Development Center in Asia Pacific, Shanghai, China

ABSTRACT

This paper intends to introduce the tolerance intervals and how the tolerance limit can be generated using the CAPABILITY Procedure. Also it uses an example to describe how tolerance intervals can be applied to define a drug response by calculating a specific cutpoint for a certain biomarker.

INTRODUCTION

The point estimates provide a concise summary of the sample results, but they give no information about their precision. We need to quantify the uncertainty associated with our estimate. An understanding of this uncertainly is an important input for decision making.

The Confidence Intervals are the most well-known and most often used statistical intervals in the clinical trials. Confidence Intervals are used to express the uncertainty associated with a population parameter such as the population mean, μ, or the population standard deviation, s.

However, Confidence Intervals are not always appropriate. There are other statistical intervals to characterize the results. One type is to cover a specified proportion of a population distribution with a given confidence. They are Tolerance Intervals which are the least-known.

TOLERANCE INTERVALS

A tolerance interval consists of two data values expected to contain a prespecified proportion of the underlying data population, with a specified level of confidence. As an example, for a 95% tolerance interval with a 90% confidence level, there is 90% confidence that, on average, 95% of the data population is contained within the interval. The tolerance interval is said to have 95% coverage and a confidence level of 90%.

Figure 1.1 depicts a Tolerance Interval graphically. Notice in figure 1.1 that the interval range extends beyond the tail areas of the actual population distribution (solid line). This is because the Tolerance Interval must take into account the uncertainty of knowing the true location of the mean of the population distribution. This uncertainty is represented by the confidence level associated with the interval. The confidence level indicates the likelihood that the interval covers the desired proportion of the population. The dotted line distributions result from the uncertainty of knowing the true location of the population mean. The difference in location of the mean due to this uncertainty is defined by the confidence interval.

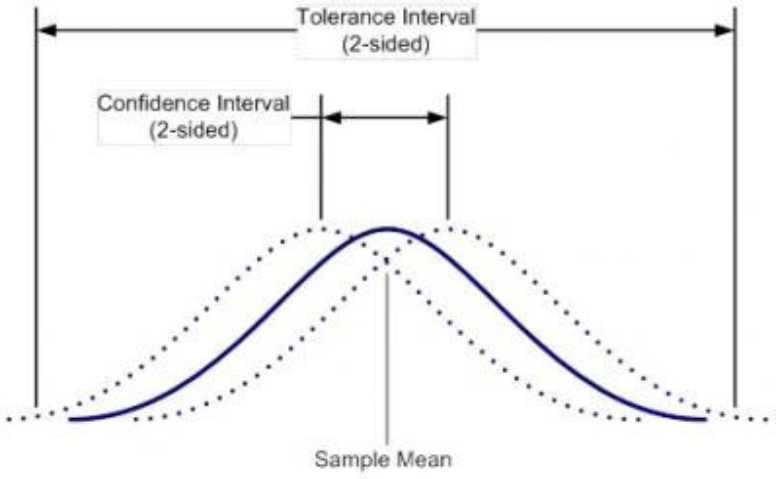

Figure 1.1 Graphical Depiction of a Tolerance Interval

THE CAPABILITY PROCEDURE

OVERVIEW

The CAPABILITY procedure provides the following and not limit to:

- descriptive statistics based on moments, including skewness and kurtosis. Other descriptive information provided includes quantiles or percentiles (such as the median), frequency tables, and details on extreme values.
- histograms and comparative histograms. Optionally, these can be superimposed with specification limits, fitted probability density curves for various distributions, and kernel density estimates.
- cumulative distribution function plots (cdf plots). Optionally, these can be superimposed with specification limits and probability distribution curves for various distributions.
- quantile-quantile plots (Q-Q plots), probability plots, and probability-probability plots (P-P plots). These plots facilitate the comparison of a data distribution with various theoretical distributions. Optionally, Q-Q plots and probability plots can be superimposed with specification limits.
- goodness-of-fit tests for a variety of distributions including the normal. The assumption of normality is critical to the interpretation of capability indices.
- statistical intervals (prediction, tolerance, and confidence intervals) for a normal population
- the ability to produce plots either as traditional graphics, ODS Graphics output, or legacy line printer plots. Traditional graphics can be saved, replayed, and annotated.

SYNTAX

The following are the primary statements that control the CAPABILITY procedure:

```
PROC CAPABILITY <options> ;
VAR variables ;
BY variables ;
HISTOGRAM <variables> </ options> ;
INTERVALS <variables> </ options> ;
OUTPUT <OUT= SAS-data-set> <keyword1=names ...keywordk=names> ;
```

Notes:

The BY statement specifies variables in the input data set that are used for BY processing. A separate analysis is done for each group of observations defined by the levels of the BY variables.

The HISTOGRAM statement displays the density curves for fitted theoretical distributions (beta, exponential, gamma, Johnson , Johnson , lognormal, normal, and Weibull) on histograms.

- The NORMAL option displays a fitted normal curve on the histogram.

The INTERVALS statement tabulates various statistical intervals for selected process variables.

- The METHODS= option specifies which intervals are computed.
- The TYPE= option specifies the type of intervals (one-sided lower, one-sided upper, or two-sided)

METHOD FOR COMPUTING TOLERANCE INTERVALS

You can specify options after the slash (/) in the INTERVALS statement to control the computation and printing of intervals.

The METHODS=3 option requests an approximate statistical tolerance interval that contains at least proportion of the population. AND the TYPE= UPPER option requests one-sided upper limit.

AN REAL EXAMPLE

A drug response was defined in those subjects who exhibited a fold-change from baseline at any time post dose that was greater than a specific cutpoint for each biomarker IP-10, and Neopterin, respectively.

The cutpoints utilized are upper tolerance limits calculated from the placebo measurements taken in the 8 subjects. In this example, 90% of the population measurements will be less than the cutpoint shown with 90% confidence.

Back transform the upper limit using 10**(upper limit)

STEP 1 INPUT DATA

The following statements create the data set which contains log transformed values for the Maximum Fold Change from Baseline of IP10 and Neopterin for 8 subjects who received placebo tablets.

```
data ip10neop;
  input PARAMCD $ FoldChgfromBL  log10var;
  cards;
  IP10        0.59672       -0.22423
  IP10        0.50329       -0.29818
  IP10        0.05546       -1.25604
  IP10        0.75974       -0.11933
  IP10        0.7841        -0.10563
  IP10        0.6098        -0.21481
  IP10        0.61757       -0.20931
  IP10        0.65616       -0.18299
  NEOPT       1.11392       0.04686
  NEOPT       0.79167       -0.10146
  NEOPT       0.80597       -0.09368
  NEOPT       0.91837       -0.03698
  NEOPT       0.84127       -0.07506
  NEOPT       0.80645       -0.09342
  NEOPT       0.85714       -0.06695
  NEOPT       0.78873       -0.10307
  ;
run;

proc sort data=ip10neop;
  by paramcd;
run;
```

STEP 2 THE UPPER TOLERANCE LIMIT

The following statements compute the upper tolerance limits calculated from the placebo measurements, i.e., 90% of population measurements will be less than the cutpoint shown with 90% confidence. The NORMAL option summarizes the fitted distribution in the printed output shown in Output X.1 and Output X.2 for IP-10 and Neopterin, respectively.

```
ods graphics on;
proc capability;
  var log10var;
  by paramcd;
  intervals  log10var/type=upper method=3;
  histogram/normal;
  title 'Log10 Max fold change ';
run;
ods graphics off;
```

Log10 Max fold change 16:50 Sunday, August 2, 2015 11

PARAMCD=IP10

The CAPABILITY Procedure
Fitted Normal Distribution for log10var

Parameters for Normal Distribution

Parameter	Symbol	Estimate
Mean	Mu	-0.32632
Std Dev	Sigma	0.380566

Goodness-of-Fit Tests for Normal Distribution

Test	----Statistic-----		DF	------p Value------	
Kolmogorov-Smirnov	D	0.4044667		Pr > D	<0.010
Cramer-von Mises	W-Sq	0.3119676		Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	1.6207839		Pr > A-Sq	<0.005
Chi-Square	Chi-Sq	25.2758989	2	Pr > Chi-Sq	<0.001

Quantiles for Normal Distribution

Percent	------Quantile------ Observed	Estimated
1.0	-1.25604	-1.21164
5.0	-1.25604	-0.95229
10.0	-1.25604	-0.81403
25.0	-0.26121	-0.58300
50.0	-0.21206	-0.32632
75.0	-0.15116	-0.06963
90.0	-0.10563	0.16140
95.0	-0.10563	0.29966
99.0	-0.10563	0.55901

Output X.1 Summary for Fitted Normal Distribution for IP-10

Log10 Max fold change 16:50 Sunday, August 2, 2015 15

PARAMCD=NEOPT

The CAPABILITY Procedure
Fitted Normal Distribution for log10var

Parameters for Normal Distribution

Parameter	Symbol	Estimate
Mean	Mu	-0.06547
Std Dev	Sigma	0.050441

Goodness-of-Fit Tests for Normal Distribution

Test	----Statistic-----		DF	------p Value------	
Kolmogorov-Smirnov	D	0.26170382		Pr > D	0.103
Cramer-von Mises	W-Sq	0.13363766		Pr > W-Sq	0.032
Anderson-Darling	A-Sq	0.79991559		Pr > A-Sq	0.022
Chi-Square	Chi-Sq	8.86380656	2	Pr > Chi-Sq	0.012

Quantiles for Normal Distribution

Percent	------Quantile------ Observed	Estimated
1.0	-0.10307	-0.18281
5.0	-0.10307	-0.14844
10.0	-0.10307	-0.13011
25.0	-0.09757	-0.09949
50.0	-0.08424	-0.06547
75.0	-0.05197	-0.03145
90.0	0.04686	-0.00083
95.0	0.04686	0.01750
99.0	0.04686	0.05187

Output X.2 Summary for Fitted Normal Distribution for Neopterin

STEP 3 THE CUTPOINT

Copy the results from the Output X.3 of above program and back transform the upper limit using 10**(upper limit) in the following statements.

```
                          Log10 Max fold change       16:50 Sunday, August 2, 2015  12

---------------------------------- PARAMCD=IP10 ----------------------------------------

                          The CAPABILITY Procedure
     One-Sided Upper Statistical Intervals for log10var Assuming Normality

                            Tolerance Limit For
                            At Least Proportion
                            p of the Population

                       Confidence      p      Upper Limit

                          99.00%     0.900     1.005
                          99.00%     0.950     1.305
                          99.00%     0.990     1.885

                          95.00%     0.900     0.656
                          95.00%     0.950     0.887
                          95.00%     0.990     1.331

                          90.00%     0.900     0.518
                          90.00%     0.950     0.722
                          90.00%     0.990     1.113
```

Output X.3 One-Sided Statistical Intervals for IP-10

```
data ip10;
       input Confidence $ p ul;
       upperlimit=10**ul;
       cards;
                                        99.00%     0.900     1.005
                                        99.00%     0.950     1.305
                                        99.00%     0.990     1.885
                                        95.00%     0.900     0.656
                                        95.00%     0.950     0.887
                                        95.00%     0.990     1.331
                                        90.00%     0.900     0.518
                                        90.00%     0.950     0.722
                                        90.00%     0.990     1.113
;
run;

proc print;
  title 'IP10';
  label p='Proportion of Population';
  var Confidence p upperlimit;
run;
```

Read the upper limit in the 90% of population shown with 90% confidence from the Output X.4 and Output X.5 for IP-10 and Neopterin, respectively.

IP10 16:50 Sunday, August 2, 2015 17

Obs	Confidence	p	upperlimit
1	99.00%	0.90	10.1158
2	99.00%	0.95	20.1837
3	99.00%	0.99	76.7361
4	95.00%	0.90	4.5290
5	95.00%	0.95	7.7090
6	95.00%	0.99	21.4289
7	90.00%	0.90	3.2961
8	90.00%	0.95	5.2723
9	90.00%	0.99	12.9718

Output X.4 Proportion of Population for IP-10

Neopterin 16:50 Sunday, August 2, 2015 18

Obs	Confidence	p	upperlimit
1	99.00%	0.90	1.29122
2	99.00%	0.95	1.41579
3	99.00%	0.99	1.69044
4	95.00%	0.90	1.16145
5	95.00%	0.95	1.24451
6	95.00%	0.99	1.42561
7	90.00%	0.90	1.11173
8	90.00%	0.95	1.18304
9	90.00%	0.99	1.33352

Output X.5 Proportion of Population for Neopterin

IP-10 and Neopterin response defined as fold change from baseline greater than 3.30 and 1.11, respectively.

CONCLUSION

The upper one-sided tolerance interval (or more appropriately, the upper tolerance bound) is known as the upper tolerance limit (UTL), and is frequently used in environmental site monitoring. In this application, the UTL is constructed using a data sample from placebo measurements taken in the 8 subjects.

For subject who takes actual treatment, if no biomarker measurements value exceeds the, say, 90% coverage UTL, or less stringently, if no more than 10% of data values exceed the 90% coverage UTL, it is inferred that the data are essentially identical to the subjects who took placebo measurements. It is drug no-response. On the other hand, if subjects who exhibited a fold-change from baseline at any time post dose that was greater than the cutpoint for each biomarker, it is drug response.

REFERENCES

Ofungwu, Joseph. (May 27, 2014). Statistical Applications for Environmental Analysis and Risk Assessment. 648. John Wiley & Sons

NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/

GARAJ, I and JANIGA, I. (2002). Two-sided tolerance limits of normal distribution for unknown mean and variability. 147. Bratislava: Vydavatelstvo STU.

Boyles, R. A. (1992), Cpm for Asymmetrical Tolerances, Technical report, Precision Castparts Corp., Portland, OR.

Boyles, R. A. (1994), “Process Capability with Asymmetric Tolerances,” Communication and Statistics, Part B—Simulation and Computation, 23, 615–643.

Odeh, R. E. and Owen, D. B. (1980), Tables for Normal Tolerance Limits, Sampling Plans, and Screening, New York: Marcel Dekker.

RECOMMENDED READING

- Base SAS® Procedures Guide
- SAS® For Dummies®

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Jing Pan
Enterprise: Roche Product Development Center in Asia Pacific
Address: 4th Floor, Building 11,1100 Long Dong Avenue
City, State ZIP: Shanghai
Work Phone: +86-21-2892 2809
E-mail: gen.pan@roche.com