DRUG DEVELOPMENT AND INDUSTRIAL PHARMACY, 20(9), 1589-1598 (1994)

EXACT POWER FUNCTION OF COMPENDIAL TEST REQUIREMENTS FOR CONTENT UNIFORMITY

N. R. Bohidar
Philadelphia College of Pharmacy and Science
Villanova University
1530 Bridal Path Road
Lansdale, PA. 19446

Norman R. Bohidar
University of Washington, Seattle, Washington

ABSTRACT

The exact distributions associated with the current compendial test requirements are generated by resorting to the well known Computer Intensive Algorithm method to establish the exact percentage point (limit) for RSD, corresponding to each selected cut-off probability level (confidence level) for each of the four possible experimental outcomes based on the USP-NF test requirements. A table is constructed to present the two-dimensional power function. The similarities between these tabular values and the current compendial RSD limits for 10 and 30 dosage units are extremely remarkable.

Minor differences exist, however. It is suggested that both the theoretical as well as the numerical approaches should be carried out to arrive at a comprehensive solution.

INTRODUCTION

Content uniformity constitutes one of the cardinal properties of a drug product. Since it serves as an index of the degree of homogeneity of the distribution of the active ingredient among the individual dosage units of a drug product, the content uniformity measurement of a sample begins with calculating the sample estimate of the relative standard deviation (RSD) by using the formula 100S/X*, where, S and X* denote the standard deviation and mean respectively. It would be appropriate to briefly outline at the outset the USP-NF requirements(1): (i) The assay value of each of the 10 units in the sample must not only be within the required range of 85% - 115%, but also the RSD of the sample must not exceed 6% and, (ii) if one unit is outside the range of 75% - 125%, then 20 more units are tested and the RSD of the sample of 30(10 + 20) units must not exceed 7.8%.

The routine compliance of the current compendial requirements has been in practice since the late seventies. Presently there are several companies in the industry who would prefer a more stringent limit by requiring that the two-sided upper 95% confidence limit of the sample RSD be less than or equal to 6% (n=10) or 7.8% (n=30), instead of the RSD value itself, as presently required. The rationale for this approach and the necessary statistical tables to facilitate rapid access to the desired RSD confidence limits have been presented in reference (2), and some motivation, discussion and examples have been presented in reference (3). There are also some companies who would like either to have the current limits relaxed or to receive some compendial guidelines about resampling, retesting and repeat testings of batches marginally rejected based on the current requirements. There are, however, a majority of companies who are strongly in favor of the status quo and would like to review from time-to-time the statistical assumptions implicit in the construction of the compendial limits on RSD.

It should be noted that the exact probability distribution associated with the compendial test requirements cannot be accomplished analytically without making simplifying assumptions associated with the classical methods. The primary purpose of this paper is (a) to depict explicitly procedures for generating exact distributions associated with the current compendial test requirements by resorting to the well known Computer Intensive Algorithm (CIA) method (4), (b) to establish the exact percentage point (limit) on RSD for each specific compendial test requirement, and (c) to provide a table of percentage points (critical values) for each of the selected cut-off probability level to have an overview of the possible alternate options.

DISTRIBUTIONS, PROBABILITIES AND PERCENTAGE POINTS

Cumulative distribution function (CDF) of a response variable (e.g. assay value, content uniformity) denoted by $F(X) = P(X \leq X_o)$, and its derivative $dF(X)/dX = f(X)$ are used to determine the cut-off probability level p (1 - p = confidence or assurance level) corresponding to a given percentage point (critical value) X_p by evaluating the integral,

$$p = F(X_p) = \int_{\infty}^{X_p} f(t)dt$$

However for practical applications, one is more interested in determining X_p (e.g. "tabular value") for a given p (e.g. "significance level"), which requires the evaluation of $X_p = F^{-1}(p)$ (4). The problem becomes even more complicated for situations, such as compendial test requirements, in which neither the exact nor an acceptable approximate probability distribution can be derived analytically. Successful solutions under these circumstances can however be accomplished by resorting to suitable numerical methods, such as the CIA method, which can appropriately be streamlined to suit the specific constraints required for the problem at hand. Based on the CIA method, one can generate the exact distribution of a population of vast number of samples of size n, determine the exact CDF of the numerical distribution and obtain the desired critical value (percentage point) pertaining to a given cut-off p-value. The concerns pertaining to the appropriateness of the assumed theoretical distribution, the availability of a closed form expression for $p = F(X_p)$, the availability of an inverse closed form expression for $X_p = F^{-1}(p)$ and the accuracy of the various approximations are non-existent.

Consider now the following schematic diagram associated with the current compendial test requirements for dosage units,

125.0 ════════════════════ RU

DU (1/29)

115.0 ════════════════════ AU

D(10) D(30)

85.0 ════════════════════ AL

DL (1/29)

75.0 ════════════════════ RL

The schematic diagram depicts the compendial acceptance (A) region between AL and AU and the two rejection (R) regions, one between AU and RU and the other between AL and RL, where L and U denote lower and upper limits. However, these definitions are applicable if all the 10 units are within the acceptance region. If one unit falls in one of the two rejection regions, then for a sample of 30 units, the rejection regions become the acceptance regions for that one unit and continue to be the rejection regions for the other 29 units. This variant nature of the rejection regions distinctly makes the numerical method more attractive and acceptable.

Based on the compendial requirements, one would need to generate four basic distributions designated by D(10), D(30), DL(29+1), and DU(29+1); D(10) and D(30) represent the distributions of the populations with samples of sizes 10 and 30 respectively from the AL-AU region; DL(29+1) represents the distribution of the population of samples of size 30, with 29 units from the AL-AU region and one unit from the AL-RL region in each sample; and DU(29+1) is the same as DL(29+1) except that one unit comes from the AU-RU region. Since each sample provides three statistics, mean, standard deviation and RSD, each basic distribution (D(10) etc.) would have three parametric distributions. For each of these 9 parametric distributions, a CDF (F(X)) is generated and the critical value (X_p) for each of the several cut-off p values at the two tail areas of the distribution are obtained and tabulated. Since these tables can be used for accepting or rejecting a given RSD based on the specified confidence level and since the tables provide several alternative options for the choices of p's and X_p's for establishing specifications, these tabular values, as a whole, rightfully represent the desired power function for the current compendial test requirements. Indeed the power function is two-dimensional in that the second dimension is defined by the four possible experimental outcomes. Different population sizes (no. of samples in a population) and different populations with the same size are considered here.

Since the observations in a sample must be equally likely to occur (equi-probable) and be independent, the CIA method must involve the generation of standard uniform random numbers by using the efficient linear congrential generator which is determined by four parameters, X_o, A, C and M. Each parameter is a nonnegative integer, and X_o is known as the seed, the constant A is the multiplier, the constant C the increment, and the constant M the modulus. A sequence of nonnegative integers less than M is uniquely determined by the congruence,

$$X_i = AX_{i-1} + C \qquad (\text{Mod } M)$$

where, i = 1,2, ----, $0 \leq X_i < M$ and X_0 is the given starting value. Congruential generators with non-zero C are called mixed congruential generator, which are extremely popular. Generally M is chosen to be 10^d where, d is the maximum significant number of digits retained in a computer, C is chosen to be a number indivisible by 2 or 5 and a = 1 (Mod 20). Since the length of the period (without repetition) is M, X_0 can be chosen to be any arbitrary number. The mean and variance of the generated M values are, (M-1)/2 and $(M^2 - 1)/12$. When these values are scaled into Uniform (0,1) by dividing each value by M, the mean and variance for large M become (1/2) and 1/12), the theoretical first two moments of the uniform probability distribution. The following statistical tests are conducted to confirm if indeed, the generated observations are independent (random): frequency test, serial correlation test, runs test, Kolmogorov-Smirnov test, permutation test, distance test, spectral test, autocorrelation test and lattice test.

The S-plus 3.1 Version (5) computer package is used to accomplish the CIA method for generating all the necessary tables presented in the next section.

RESULTS AND DISCUSSION

The contents of TABLE-I consist of the crucial critical values of RSD based on the current compendial test requirements, D(10), D(30), DL(29+1) and DU(29+1), at the cut-off probability level of p = 0.05 (95% confidence level, generally used) for the various population sizes indicated there. The results of the largest population should be considered as the final results, however the

TABLE-I

CRITICAL VALUES OF RSD FOR COMPENDIAL TEST REQUIREMENTS BASED ON P = 0.05 (5% LEVEL)

Population Size[z]	Compendial Test Requirements[w]			
	D(10)	D(30)	DL(29+1)	DU(29+1)
20* (cc)	6.01 (6.0)	7.39 (7.4)	8.17 (8.2)	8.02 (8.0)
10* (c)	6.00 (6.0)	7.38 (7.4)	8.25 (8.2)	8.07 (8.1)
10* (1)	6.02 (6.0)	7.41 (7.4)	8.11 (8.1)	7.98 (8.0)
5* (1)	5.96 (6.0)	7.36 (7.4)	8.18 (8.2)	7.93 (7.9)
5* (2)	6.05 (6.0)	7.39 (7.4)	8.32 (8.3)	8.22 (8.2)

z * = x1000 (20* = 20,000), c = composite of 5*(1) and 5*(2), cc = composite of 10*(c) and 10*(1), (1) = first seed and (2) = second seed.

w These symbols are defined in the text. Placed in the parenthesis are the rounded values for direct comparison with the current compendial limits.

results of the other population sizes are presented here only to show the excellent rate of convergence and the close similarities among the results associated with the various populations. A cursory glance at the contents of the table clearly demonstrates the remarkable resemblance of D(10) and D(30) values to their counterparts of the current compendial RSD limits. The D(30) limit is

TABLE-II

TWO-DIMENSIONAL POWER FUNCTION OF COMPENDIAL TEST REQUIREMENTS FOR RSDa

Compendial Requirements	Statistics	Cut-off Probability Level (p)							
		1%	2.5%	5%	10%	90%	95%	97.5%	99%
D(10)	RSD	4.90	5.46	6.01	6.65	10.35	10.8	11.22	11.67
	MEAN	93.60	94.54	95.46	96.47	103.55	104.54	105.40	106.37
D(30)	RSD	6.82	7.14	7.39	7.69	9.61	9.88	10.09	10.34
	MEAN	96.33	96.87	97.36	97.94	102.06	102.58	103.06	103.64
DU(29+1)	RSD	7.44	7.76	8.02	8.31	10.13	10.36	10.59	10.84
	MEAN	97.06	97.61	98.10	98.65	102.65	103.21	103.71	104.31
DL(29+1)	RSD	7.64	7.92	8.17	8.45	10.21	10.45	10.64	10.88
	MEAN	95.78	96.28	96.75	97.31	101.28	101.83	102.31	102.83

a For the total population of 20*(cc), defined in the text

slightly smaller than the compendial limit because the tabular values are generated based on the exact compendial range and distributions, whereas the classical theoretical distributions are non-specific and cover a wide range of values. There are no theoretical distributions for DL(29+1) and DU(29+1) exist, and their critical limits are higher than the current compendial limit of 7.8, because the limits here are obtained based on the exact distributions and the exact required compendial range. The contents of TABLE-II depicts the two-dimensional power function associated with the current compendial test requirements for RSD. One would be able to visualize the entire spectrum of critical values associated with the selected tail probabilities. The critical values assuming a p-value of 0.01 (99% confidence) are 4.9, 6.8, 7.4 and

7.6 respectively for D(10), D(30), DL(29+1) and DU(29+1) which are stringent and extremely conservative. It should be noted here that the mean value corresponding to each selected cut-off probability, derived from the mean distribution, also appears in the table. It not only shows the broad critical range but also that (a) the observations in D(10) and D(30) are symmetrically distributed, (b) the distribution of DL(29+1) is skewed to the left and (c) the distribution of DU(29+1) is skewed to the right, showing that the assumption of symmetricity is not tenable. It should be noted here that these derivations are neither based on the distributional assumption of normality nor the parametric assumption of $\mu = 100$, $\sigma = 10$ and RSD = 10%. However, the analytical (theoretical) methods as well as the CIA method should be carried out for the construction of the limits. And the reconciliation of the results of these methods should be the prime consideration for a comprehensive solution.

ACKNOWLEDGEMENT

Grateful thanks are due to Mrs. Barbara J. Tomlinson for her meticulous efforts in wordprocessing the manuscript with utmost rapidity.

REFERENCES

1. USP-NF United States Pharmacopeia-National Formulary. USP-22nd Revision, NF-17th Revision. USP Convention, Inc. Rockville, Md. Pg. 1618-1619 (1989).
2. N. R. Bohidar and N. R. Bohidar, Construction of Upper Confidence Limit of Coefficient of Variation for Content Uniformity. Drug Development and Industrial Pharmacy, Vol. 18, No.1, 21-37 (1992).
3. N. R. Bohidar Norman R. Bohidar and Nicholas R. Bohidar, Detection of Discordant Content Uniformity

Observations and Compendial Compliance. Drug Development and Industrial Pharmacy, Accepted for Publication in 1993.

4. W.J. Kennedy Jr. and J.E. Gentle, "Statistical Computing", Chapter 5 and 6. Marcel Dekker, Inc. N.Y. (1980).

5. S-Plus (3.1 Version), Statistical Sciences, Inc. Seattle, Wash. (1992).