Assessment of sampling variability on the ability of the 30 day mean DO criterion to serve as an umbrella for the 7 day mean criterion.

Elgin Perry
eperry@chesapeake.net
Feb 26, 2013

Introduction

A sub-committee of TMA reviewed the variability of the weekly mean DO about the monthly mean DO and concluded that in general, that if the 30-day DO criterion is satisfied by the monthly mean, then there is less than a 10% chance that the 7-day DO criterion will be violated by the weekly mean. This conclusion is based on having very accurate estimates of both the monthly mean and the weekly mean derived from near continuous high frequency observations of DO. However, in many parts of the bay, the monthly mean is estimated from as few as one to two point observations per month. Because the uncertainty of a monthly mean of two observations is much greater than the uncertainty of a monthly mean from near continuous data, it is reasonable to expect that effectiveness of the umbrella effect of the 30-day criterion for the 7-day criterion will diminish when the low sample size mean is employed. In this study, we examine the additional uncertainty that is created by the use of small sample size and further evaluate the consequences of this uncertainty for the umbrella concept.

Methods

This study evaluates the additional uncertainty from low sample sizes by resampling from near continuous records in a manner that simulates the twice monthly sampling of routine cruises. The near continuous DO records used are from the Potomac ConMon studies. For each calendar month in the May through September period of each record, a random day between 1 and 15 was chosen as the first sampling day of the month.   To get a second sampling day, a random increment from 10 to 16 was generated and added to the first. In the event that there was no data on this second day, then the last day of the month with data was used.   For each selected day, a random selection from the roughly 24 observations taken between 9:00 a.m. and 3:00 p.m. was chosen as the point estimate. These two estimates were summed and divided by 2 to obtain the monthly mean estimate. This simulation was repeated 20 times to obtain 20 monthly mean estimates for each station and month.

Months were calendar months, and weeks were designated as sequential weeks beginning January 1st of each year.   Weekly means were computed for each unique combination of month and week. Thus if a month terminus divided a week, then the week was divided at this point and the resulting partial weeks were assigned to the two months. Deviations of the weekly means about the monthly mean were computed as (weekly mean DO – monthly mean DO) for weeks that occur within a month. In all cases, the weekly mean DO was computed as the mean of all high frequency observations within a week and is referred to as the near true weekly mean. The monthly mean was computed two ways. A near true estimate of the monthly mean uses all observations in the near continuous record; a small sample

estimate of the monthly mean uses only two observations as described by the resampling methods above. The root mean square error (rmse) was computed across months, years, and stations for both the near true deviations and the small sample deviations. These root mean square estimates quantify the standard deviation of the weekly mean about the monthly mean for both the near true case and the small sample estimate case. The increase in the rmse for small sample case relative to the near true case illustrates the loss of precision in estimating the monthly mean by small samples.

Using these estimates of standard deviation and assuming a normal distribution for these deviations, we estimate the probability that the weekly mean is less than 4.0, the 7-day mean criterion, while the monthly mean is 5.0, the 30 day mean criterion. This probability is a measure of the efficacy of the 30-day criterion as an umbrella for the 7-day criterion.

Results:

Descriptive statistics for the true weekly deviations and the small sample deviations show a negative bias of small sample deviations relative to the true deviations (Table 1). This shows that the resampled monthly means which use daytime data only tend to be biased high, but on average this effect is not large. The range of the mean of the deviations over the resampling experiments is (-0.3428 -0.0133). The variability of the small sample deviations is much greater than that of the near true deviations. The true deviations have a rmse very close to 1.0 while the rmse from the small sample deviations always exceeds 1.6 and in one case exceeds 1.9 indicating a 60 to 90 percent increase in variability.

Summary of Monthly Mean DO vs. Weekly Mean DO resampling study results
Mon Feb 25 16:33:58 2013

Table 1.0 Summary of comparing weekly DO means to Monthly DO means for 'true' means and monthly means from 20 small sample resampling experiments.

Simu-lation	sample size	mean	rmse	minimum	q25	Median	q75	maximum
true	833	0.0017	1.005	-4.18	-0.4816	0.0125	0.4828	3.2042
1	833	-0.1344	1.6578	-5.1447	-0.9944	-0.0542	0.8052	4.9893
2	833	-0.0247	1.6903	-5.6588	-0.8519	0.0165	0.8543	4.4843
3	833	-0.2745	1.7132	-6.684	-1.1194	-0.1775	0.6852	4.4353
4	833	-0.2187	1.8037	-7.9388	-0.9968	-0.0879	0.7284	5.3265
5	833	-0.1723	1.8766	-8.2638	-0.9699	-0.0726	0.8603	4.9031
6	833	-0.0666	1.6177	-5.379	-0.8897	-0.0173	0.7745	4.6073
7	833	-0.2252	1.7196	-6.8519	-1.066	-0.2264	0.6948	5.3679
8	833	-0.0133	1.6054	-5.4517	-0.7627	0.0211	0.8046	5.1295
9	833	-0.3428	1.7471	-6.3008	-1.1947	-0.2999	0.5542	4.3745
10	833	-0.1639	1.7156	-7.3597	-1.0652	-0.1465	0.8385	4.7042
11	833	-0.0948	1.7555	-5.7288	-1.0169	-0.0054	0.8369	5.0334
12	833	-0.2193	1.9286	-7.2316	-1.0929	-0.0793	0.7621	5.5595
13	833	-0.2014	1.692	-6.5302	-1.0818	-0.0624	0.7351	5.1557

14	833	-0.1747	1.6198	-6.2597	-1.063	-0.1254	0.8021	3.9682
15	833	-0.1424	1.7216	-6.3428	-1.0468	-0.1171	0.8693	4.8051
16	833	-0.1055	1.7055	-6.114	-1.0153	0.0278	0.9094	4.3039
17	833	-0.1663	1.8126	-6.424	-1.1035	-0.107	0.7703	4.6611
18	833	-0.2157	1.8397	-6.3407	-1.1281	-0.1486	0.8262	5.2234
19	833	-0.0624	1.7048	-5.3103	-1.0217	-0.0165	0.8549	4.7011
20	833	-0.2306	1.7493	-8.2242	-1.1226	-0.1713	0.7209	4.2593

The distribution of the true weekly deviations tends to follow the normal distribution closely for the bulk of the observations (Figure 1.0). However, there is a small percentage of outliers at both the upper end and the lower end of the distribution that are more extreme than are expected for the normal distribution. Because of this heavy tailed feature of the true deviations, when the normal distribution is used to compute probabilities for this problem, these probabilities may be a slight underestimate of the true probabilities. There appear to be 10 to 15 extreme outliers in the lower tail of the distribution and thus the probability bias may be 1.2 to 1.8 percent.
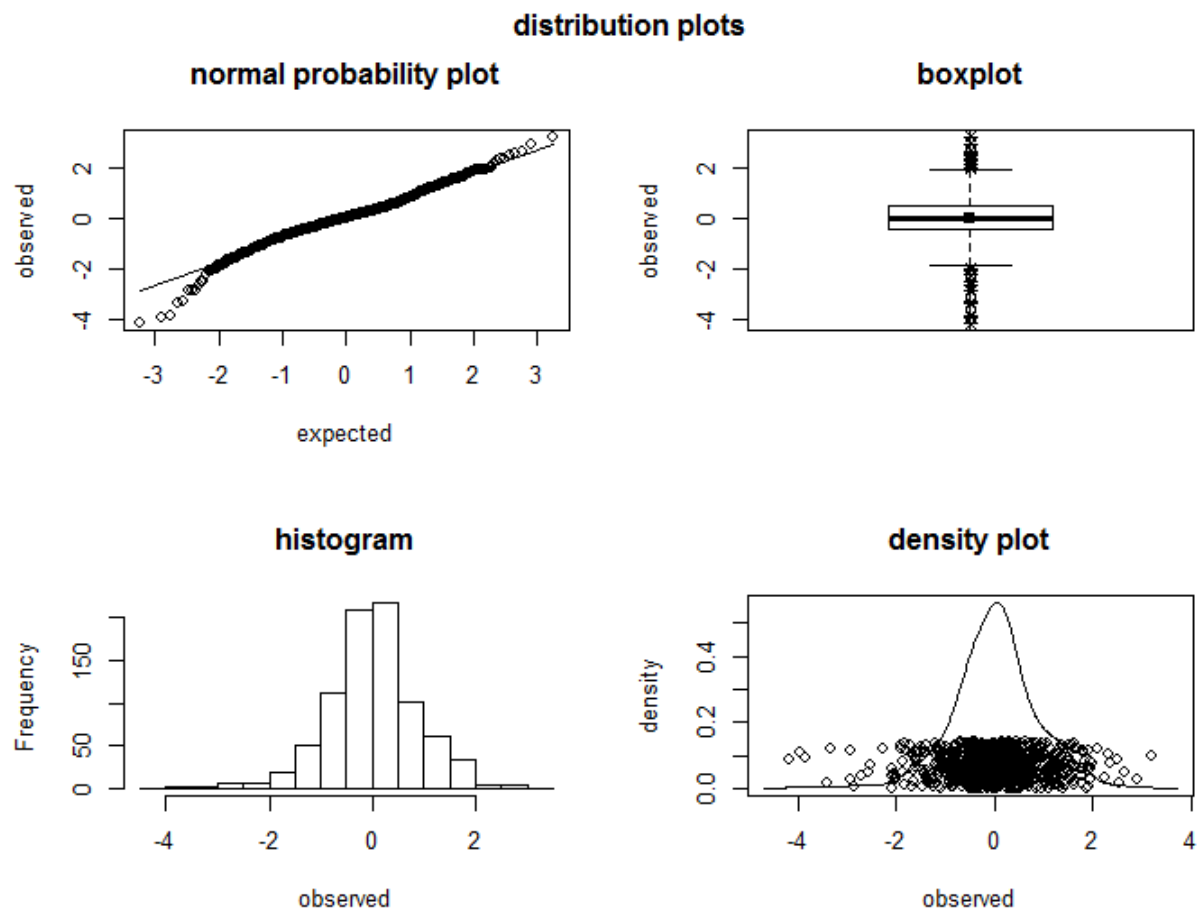

Figure 1.0 Distribution plots for the true weekly deviations.

The weekly deviations computed using the small sample monthly mean estimates appear to fit the normal distribution better than the true week deviations (Figure 2.0). The variability of deviations in the small sample experiment is clearly greater than variability for the true deviations. Compare for example the frequency of observation where the weekly mean is greater than 2 units below the monthly mean between Figures 1 and 2.
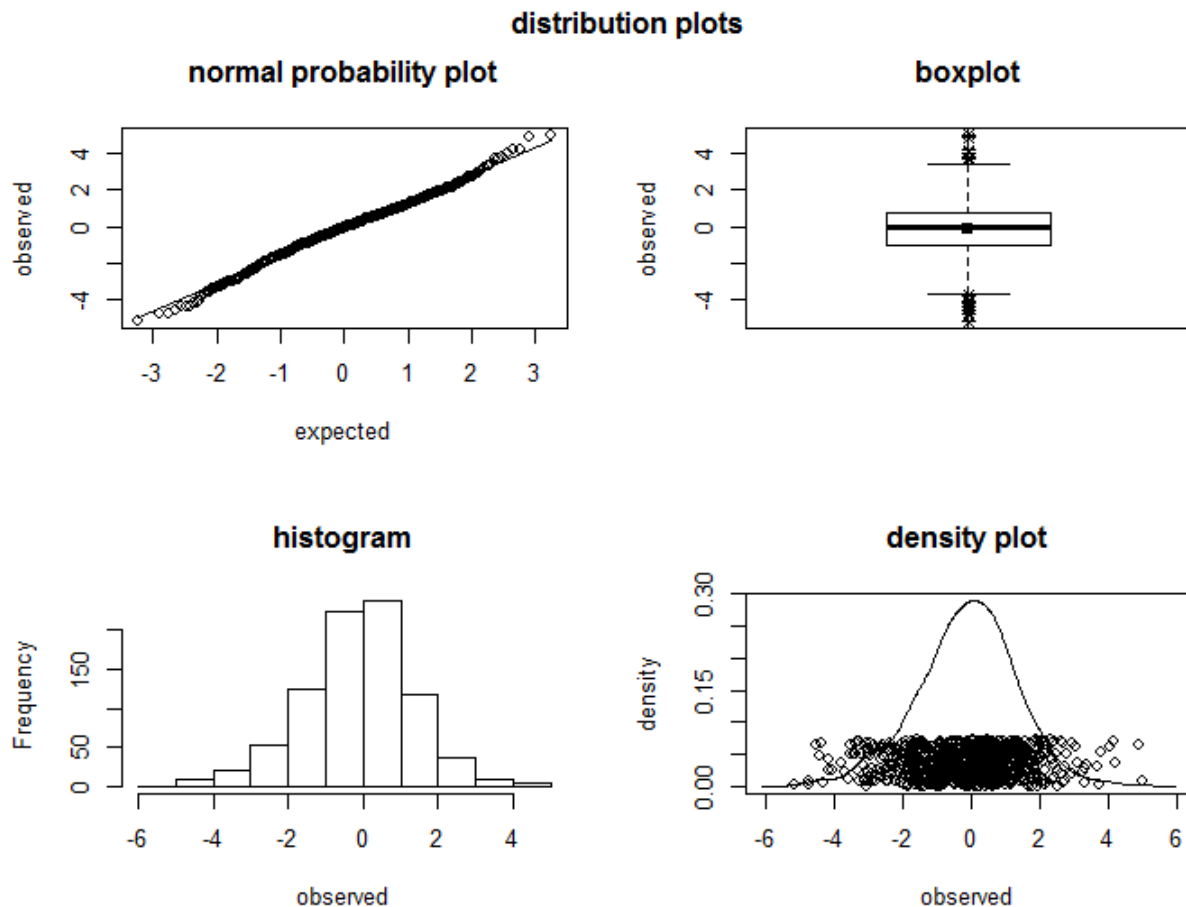

Figure 2. Distribution plots for weekly deviations computed for the first resampling experiment.

Discussion

If the 30-day criterion is to serve as an umbrella for the 7-day criterion, we would like to show that if the 30-day criterion is satisfied, there is a small probability that the 7-day criterion is violated. If we accept less than 10% as an acceptable risk of wrongly concluding that the 7-day criterion is satisfied when it is in fact violated, then this is satisfied when the standard deviation of the weekly mean from the monthly mean is .7805 or smaller. At this level of variability in the weekly deviations, excursions of the weekly mean below the 7-day criterion of 4.0 while the monthly mean is at the 30-day criterion of 5.0 would be about 10% (Figure 3). This scenario would be strong evidence that the 30-day criterion is an effective umbrella for the 7-day criterion.
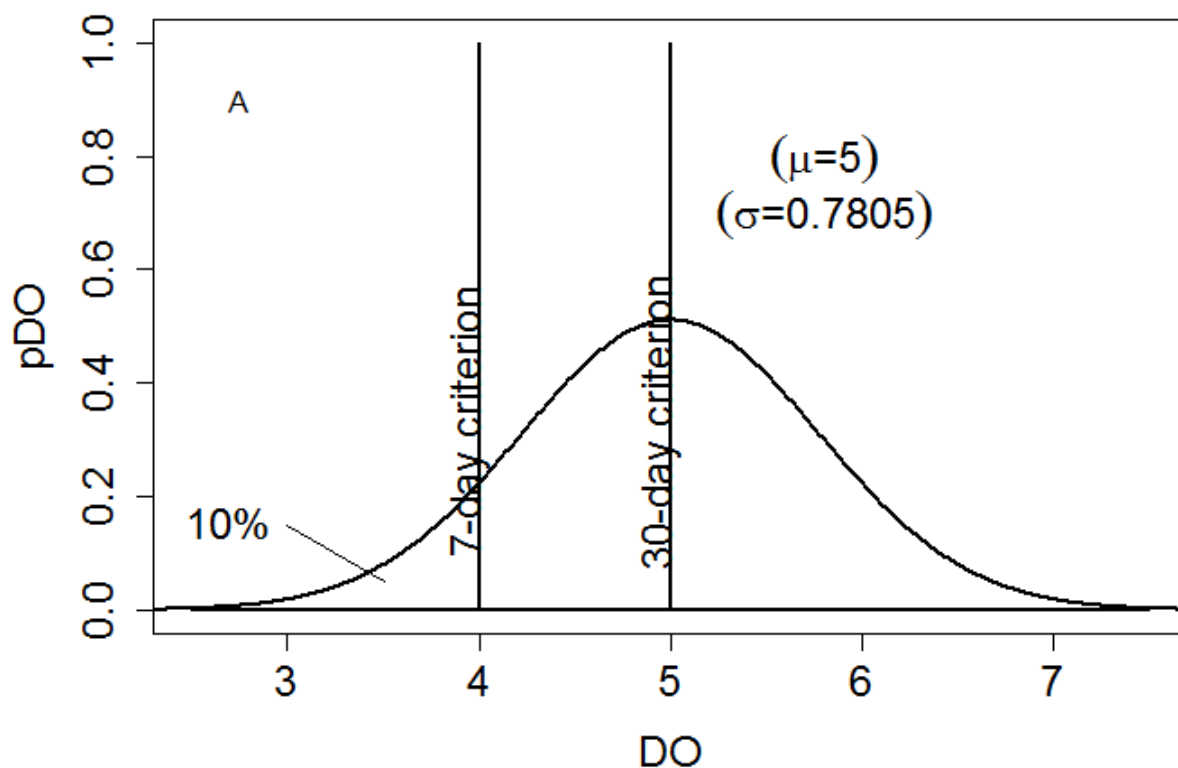

Figure 3. Illustration of the 30-day criterion as an umbrella for the 7-day criterion.

In the Potomac ConMon data, the standard deviation of the weekly mean from the monthly mean exceeds this ideal 0.7805 value and is estimated to be 1.005 or very close to 1.0. At this level of variability, the risk of violating the 7-day criterion when the 30-day criterion is satisfied exactly is about 16% (table 2., Figure 4.) However, increasing the monthly mean to 5.285 again brings the risk of violations of the 7-day criterion to an acceptable level of 10%. Because it is unlikely that the monthly mean will hover in this narrow window of (5.0, 5.285) for an extended time then it seems reasonable to consider that the 7-day criterion is satisfied if the 30-day criterion is

satisfied. This evidence is one supporting fact for the TMAW conclusion that the 30-day criterion is an effective umbrella for the 7-day criterion. It is important to recognize that this conclusion assumes that both the true monthly mean and the true weekly mean can be estimated with great precision. This high level of precision is obtained here by using a near continuous record of DO.
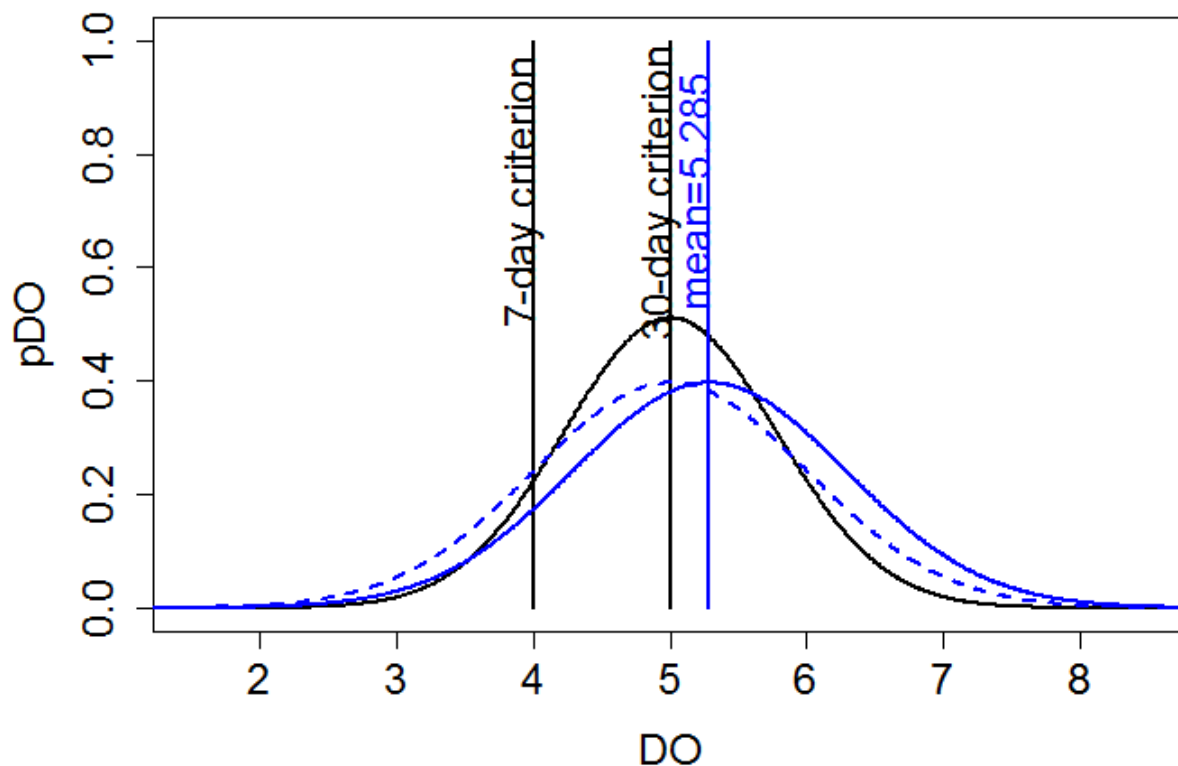

Figure 4. Illustration of the shift in the monthly mean required to meet 10% risk tolerance for the 7-day criterion when the weekly mean deviation is 1.005.

When the monthly mean is estimated by a sample size of two observations, then the variability of the deviations between the monthly mean estimate and the weekly means increases by 60 to 90 percent (table 3.0, Figure 5). At this higher level of variability, satisfying the 30-day criterion exactly results in a 28% risk of violating the 7-day criterion. Estimates of the monthly mean would have to exceed a threshold of 6.22 to insure that the risk of violating the 7-day criterion is 10% or less.

Table 2.0 Estimates of risk of violating the 7-day criterion given the monthly mean estimate (column 1) and four levels of sampling variation (columns 2-5). Column 1 assumes near true weekly deviations, column 2 assumes variation the average of 20 small sample estimates, column 3 assumes variation at the minimum of 20 small sample estimates and column 4 assumes variation at the maximum of 20 small sample estimtes.

Monthly Mean DO	Risk of violating 7-day criterion			
		$SD=1.7358^2$	$SD=1.6054^3$	$SD=1.9287^4$
5.0	0.1598	0.2822	0.2666	0.3020
5.1	0.1368	0.2631	0.2466	0.2842
5.2	0.1162	0.2446	0.2273	0.2669
5.3	0.0979	0.2269	0.2090	0.2501
5.4	0.0818	0.2099	0.1915	0.2339
5.5	0.0677	0.1937	0.1750	0.2183
5.6	0.0556	0.1783	0.1594	0.2033
5.7	0.0453	0.1636	0.1448	0.1890
5.8	0.0366	0.1498	0.1311	0.1753
5.9	0.0293	0.1368	0.1183	0.1622
6.0	0.0232	0.1246	0.1064	0.1498
6.1	0.0183	0.1131	0.0954	0.1381
6.2	0.0142	0.1024	0.0852	0.1269
6.3	0.0110	0.0925	0.0759	0.1165
6.4	0.0084	0.0833	0.0674	0.1066
6.5	0.0064	0.0748	0.0597	0.0974

1 standard deviation of true weekly mean from true monthly mean
2 standard deviation base on pooling 20 resampling estimates
3 standard deviation based on minimum of 20 resampling estimates
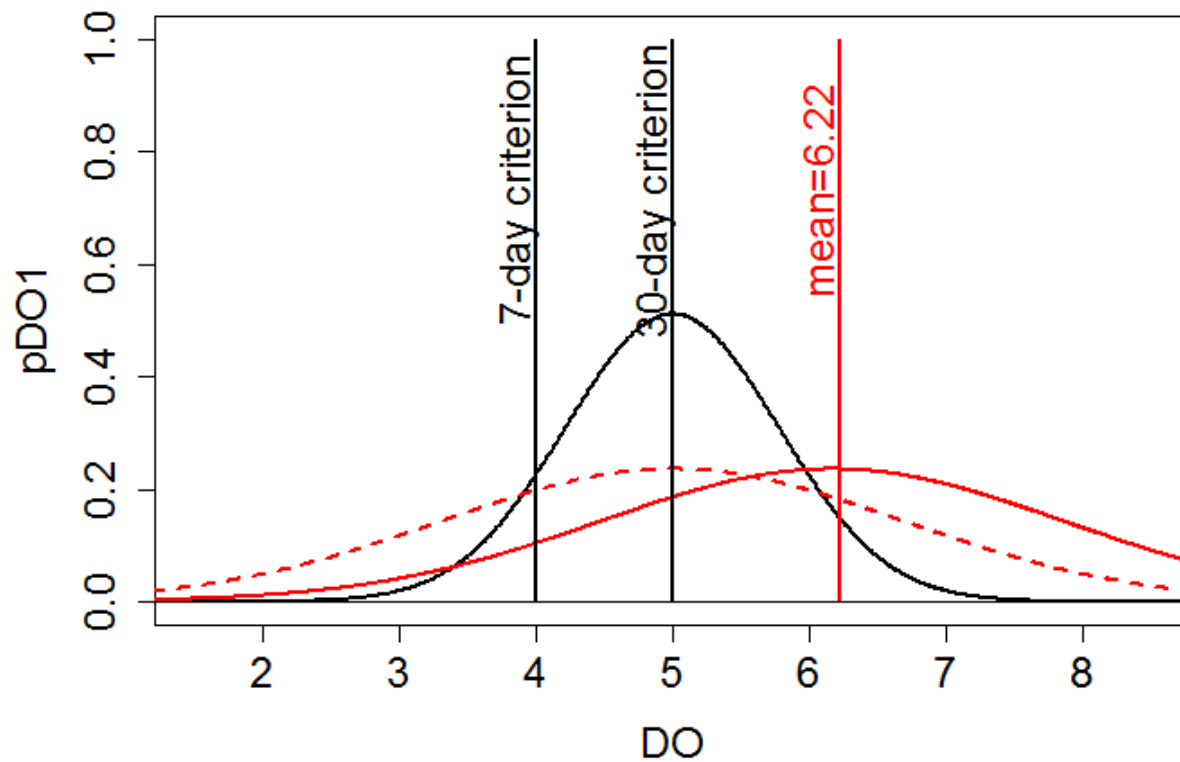4 standard deviation based on maximum of 20 resampling estimates

Figure 5. Illustration of the shift (from red dashed line to red solid line) in the monthly mean required to meet 10% risk tolerance for the 7-day criterion when the weekly mean deviation is at the small sample level of 1.74.

Conclusion:

Based on the evidence that the monthly mean threshold of 6.22 which insures that violations of the 7-day criterion is less than 10% is far above the 30-day criterion of 5.0, the 30-day criterion is not an umbrella for the 7-day criterion when the monthly mean is estimated by a sample size as small a two.