Attachment 1

Calculation of critical values for applying the exact binomial test

Summary

States are required to establish a list of impaired waterbodies under Section 303(d) of the Clean Water Act. The Oregon DEQ is proposing to evaluate attainment of numeric water quality criteria for aquatic life toxics and conventional pollutants using the exact binomial hypothesis test. DEQ is not proposing to apply the binomial statistical test for assessment of acute standards or human health criteria. This will enable Oregon to determine if data indicates waterbodies are exceeding criteria with a reasonable level of confidence.

The State of Oregon currently assesses compliance with numeric biologically-based aquatic life criteria for impairment using a simple 2 sample threshold for toxic substances. If any two observed samples within a waterbody exceed the criteria value, it is grounds for listing a waterbody as impaired (Category 5) on the 303(d) list. This matched the >1-in-3-year frequency of exceedance used for acute toxic assessment. Conventional pollutants are assessed for compliance using the “10% rule,” a raw score method that considers a waterbody to be impaired if more than 10% of the observed samples are above the criterion threshold.

DEQ has received significant input from stakeholders that the current assessment method for toxic substances is likely to overestimate the number of impaired waters, particularly in cases where sample size is large and the vast majority of samples are below the criteria thresholds. Additionally, there has not been a clear set of requirements for removal of a waterbody from the 303(d) list when improvements in water quality indicate the waterbody is currently attaining the criteria.

DEQ is proposing to update the listing and delisting methodology for assessing toxic substances and conventional pollutants using a one-sample test on binomial proportions, the exact binomial test. The binomial test allows for a statistically valid assessment of the proportion of samples in a waterbody that would be expected to exceed the criterion while accounting for uncertainty in the assessment procedure.

The proposed assessment methodology attempts to balance desired data requirements with the practical realities affecting the availability of information and the strength of the available evidence. In order to address the propensity for small data sets, DEQ will continue to apply the >1-in-3-year critical exceedance rate to all data sets less than a minimum number of samples.

Background

The exact binomial test directly limits the type-I error probability (the probability of incorrectly rejecting a null hypothesis) at or below a chosen nominal significance level (α). Using the binomial distribution, the critical value (the number of samples that exceed the threshold set by the criterion) can be calculated for a given sample size such that α is ≥ the desired confidence level. The type-II error rate is indirectly determined by sample size. However, balancing of α and β error rates can be accomplished by selection of complimentary critical rates for simultaneous evaluation of attainment and impairment in the same data set.

This attachment details the methodology and probability equations used to derive tables of critical values and both α and β rates for each proposed formulation of the test.

The binomial test determines the likelihood that the number of values exceeding the criterion threshold, *k*, observed in a sample of the population indicates that the true exceedance rate in the population, *r*, would be greater than the regulatory critical exceedance rate, *p*.

A water body is determined to be impaired if the number of excursions, *k,* in *n* samples equals or exceeds the critical value for impairment, k_i. A water body is determined to be attaining if $k \leq k_i$ for a given *n.* An alternative critical value for attainment, k_a, can be used to determine attainment if the desired error probability is different from that for impairment. In this case, a waterbody would be considered attaining if $k \leq k_a$.

The value of k_i and k_a are determined iteratively as the largest number of exceedances, respectively, such that 1-α is within the desired confidence interval for a sample of size n, with a regulatory critical exceedance rate threshold, *p*. EPA recommends an effect size of 15% for determining the difference between attaining and impaired waters. Therefore, the difference between the acceptable critical exceedance proportion, p_1, and the unacceptable critical exceedance proportion, p_2, should be such that $p_2 - p_1 = 0.15$.

Note that β is not considered in choosing the value of either k_i or k_a. The following procedures for listing and delisting are based on keeping α at or below a pre-determined limit but do not directly limit β. In this proposal, p_1 is used to calculate α, and p_2 is used to calculate β. However, only α, and consequently p_1, has an effect on the value of k_i or k_a. Tables for listing and delisting that show the value of k for a range of samples are produced below.

Alternately, selection of *k* could be based on optimization of the minimum achievable α and β, without requiring either to be held below a maximum level. This process is known as error balancing.

Generalized Listing Procedure

The listing procedure is based on a default assumption that the true exceedance rate, *r*, is less than or equal to the regulatory exceedance rate, p_1.

The tested one-sided hypotheses are:
H_O: $r \leq p_1 = 0.05$**,**
H_A: $r > p_2 = 0.20$

Then calculate α from the right tail probability of the cumulative binomial distribution:

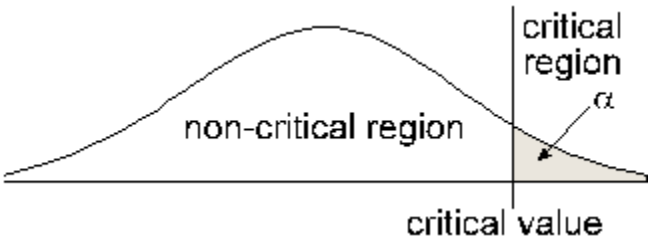
Where, n = the number of samples,
k_i = the critical value of the minimum number of sample excursions needed to place a water on the section 303(d) list,
p_1 = acceptable critical exceedance rate, and
p_2 = unacceptable critical exceedance rate.
BINOMDIST( ) is an Excel® software function that returns cumulative left tail binomial probabilities.

For sample sizes of n, the minimum number of measured exceedances is established where $\alpha < 0.1$, and where $|\alpha - \beta|$ is minimized.

α = Excel® Function BINOMDIST($n-k_i$, n, $1 - p_1$, TRUE)
β = Excel® Function BINOMDIST(k_i-1, n, p_2, TRUE)

The minimal sample size for listing and delisting was selected based on the sample size needed to achieve sufficient power with a minimum of 2 excursions, where $1- \beta > 0.90$.
The value of $k_i=2$ for n=2, and increased by 1, where $1-\alpha < 0.90$

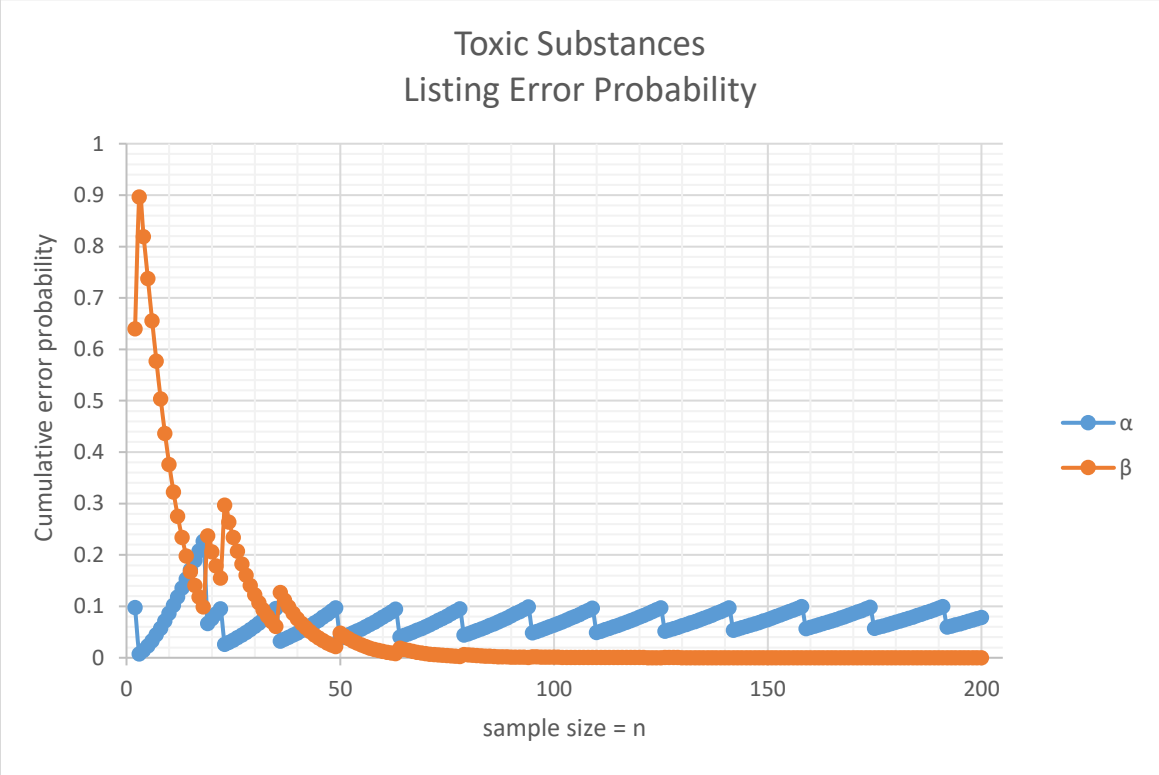
(See attached Spreadsheet)



Critical Values for Listing Chronic Toxic Substances

Null Hypothesis: Actual exceedance proportion is ≤5%
Alternate hypothesis: Actual exceedance proportion is >20%
Minimum confidence level is 90%
The minimum sample size necessary to attain 90% confidence in H_A with a minimum of 2 excursions is 18, but this critical rate is extended to smaller samples sizes.

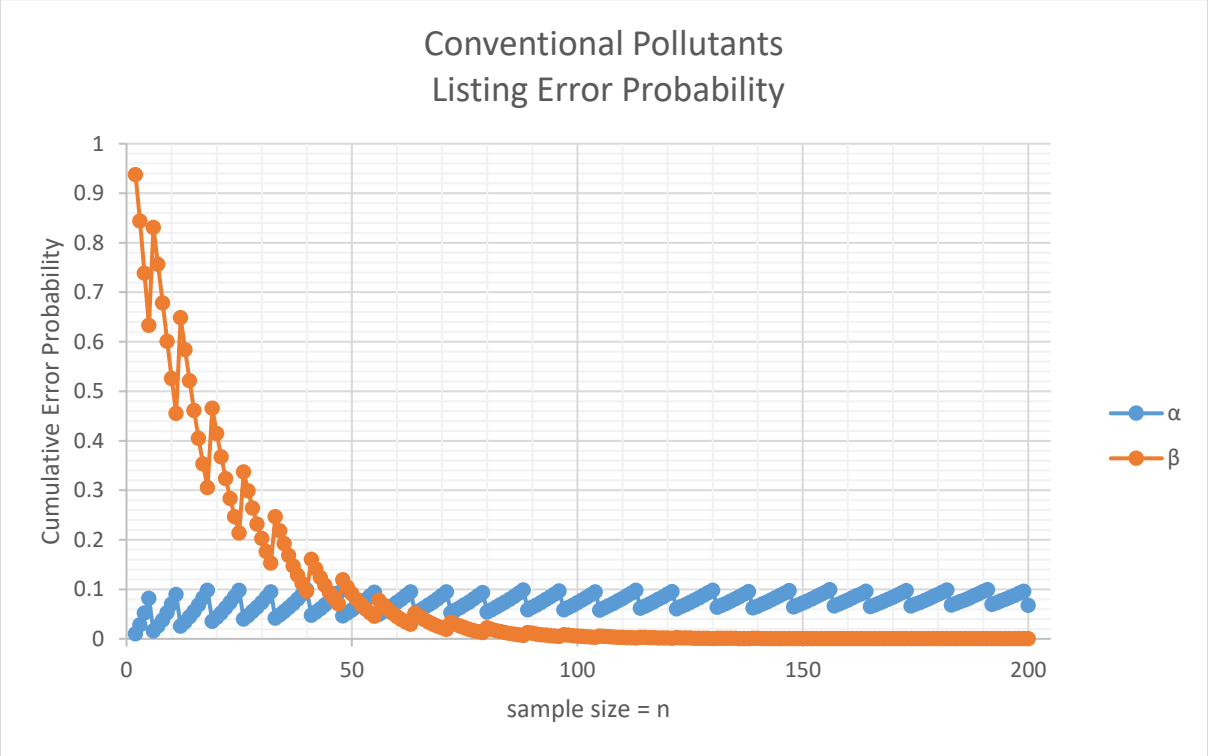
Minimum number of sample excursions required to list as impaired for toxic substances	
Sample Size	**Minimum number of excursions**
2-18	2*
19-22	3
23-35	4
36-49	5
50-63	6
64-78	7
79-92	8
93-109	9
110-125	10
126-141	11
142-158	12
159-171	13
179-191	14
192-200	15
* The use of 2 excursions to list is extended for sample sizes <10 in order to achieve >90% confidence in H_A for small sample sizes.	


Toxic Substances
Listing Error Probability
Cumulative error probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
50
100
150
200
sample size = n
α
β

Critical Values for Listing Conventional Pollutants

Null Hypothesis: Actual exceedance proportion is ≤10%
Alternate hypothesis: Actual exceedance proportion is >25%
Minimum confidence level is 90%
The minimum sample size necessary to attain 90% confidence in H_A with a minimum of 2 excursions is 14, but this critical rate is extended to smaller samples sizes.

Minimum number of sample excursions required to list as impaired for conventional pollutants	
Sample Size	**Minimum number of excursions**
2 - 11	2*
12-18	4
19-25	5
26-32	6
33-40	7
41-47	8
48-55	9
56-63	10
64-71	11
72-79	12
80-88	13
89-96	14
97-100	15
* The use of 2 excursions to list is extended for sample sizes <10 in order to achieve >90% confidence in H_A for small sample sizes.	


Conventional Pollutants
Listing Error Probability
Cumulative Error Probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
50
100
150
200
sample size = n
α
β

Generalized Delisting Procedure

For finding a waterbody is attaining, we want to know the critical value for samples that indicate the true proportion in the population, r, is greater than the regulatory critical exceedance rate, p_1.
H_O: $r > p_1 = 0.05$
H_A: $r \leq p_2 = 0.20$

Then calculate α from the left tail probability of the cumulative binomial distribution:

Where n = the number of samples,
K_a = maximum number of measured exceedances to determine a water is attaining, or should be removed from the 303(d) list,

p_1 = unacceptable exceedance proportion, and
p_2 = acceptable exceedance proportion.
BINOMDIST( ) is an Excel software function that returns cumulative left tail binomial probabilities.

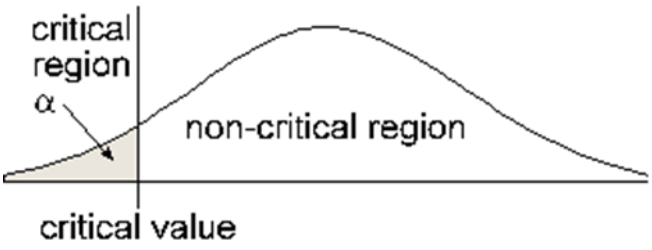
For sample sizes of n, the minimum number of measured exceedances is established where $\alpha < 0.10$ and $1-\alpha > 0.90$, and where $|\alpha - \beta|$ is minimized.

$\alpha = 1 -$ Excel® Function BINOMDIST (k_a-1, n, p_1, TRUE)
$\beta = 1 -$ Excel® Function BINOMDIST (n-k_a-1, n, 1-p_2, TRUE)

The minimal sample size for delisting was selected based on the sample size needed to achieve sufficient power with a minimum of 1 excursion, where $1- \beta > 0.90$.

The value of k_a=1 for n=10, and increased by 1, where $1-\alpha < 0.90$

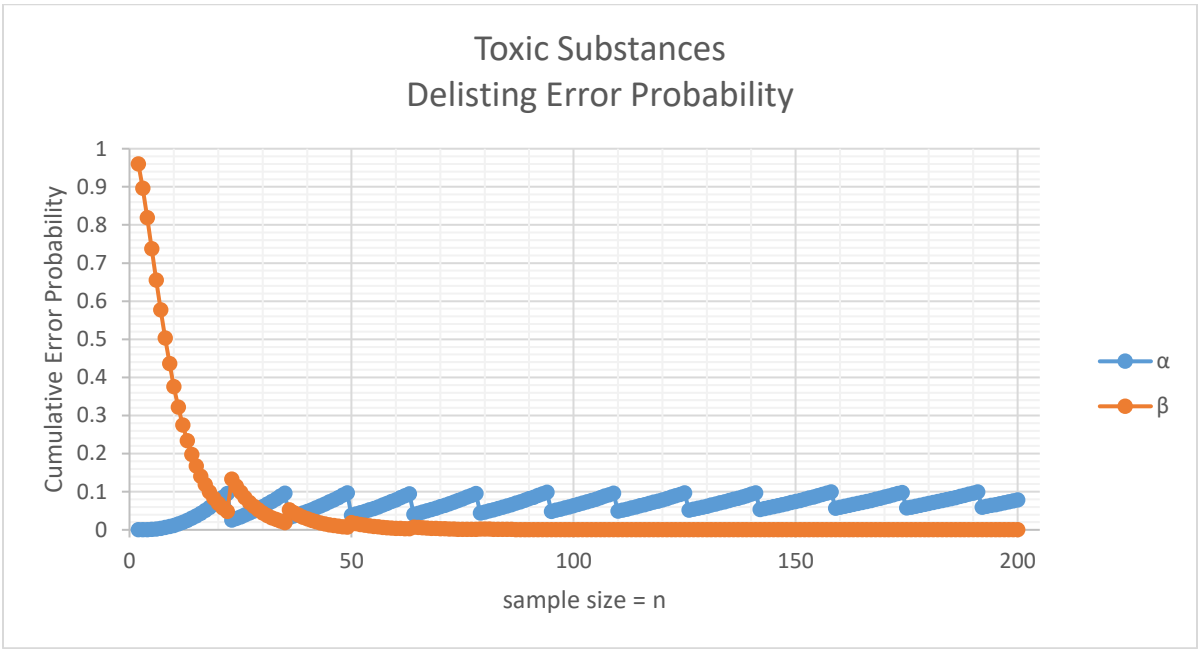
(See attached Spreadsheet)



Critical Values for Delisting Chronic Toxic Substances

Null Hypothesis: Actual exceedance proportion is >5%
Alternate hypothesis: Actual exceedance proportion is ≤20%
Minimum confidence level is 90%
The minimum sample size necessary to achieve 90% confidence for H_A with 1 exceedance is 18.

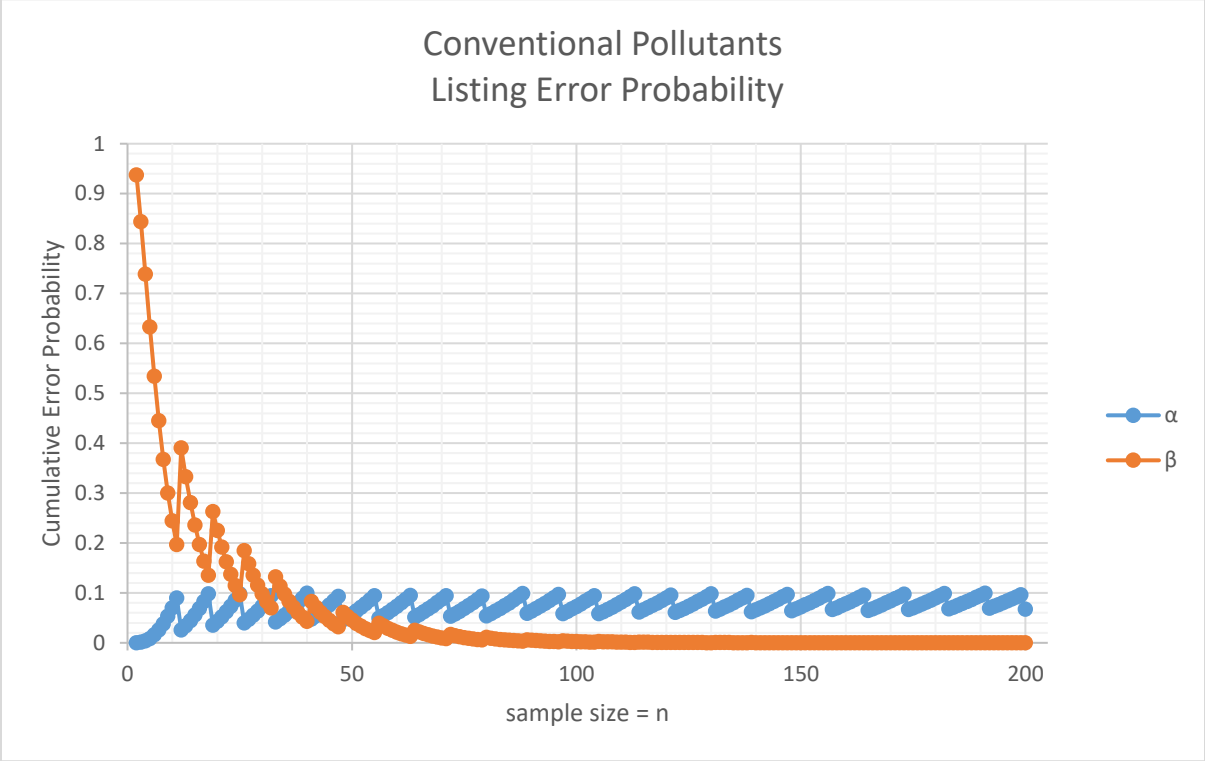
Maximum number of sample excursions to delist as impaired for toxic substances	
Sample Size	**Maximum number of excursions**
18-22	1
23-35	2
36-49	3
50-63	4
64-78	5
79-94	6
95-109	7
110-125	8
126-141	9
142-158	10
159-174	11
175-191	12
192-200	13


Toxic Substances
Delisting Error Probability
Cumulative Error Probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
50
100
150
200
sample size = n
α
β

Critical Values for Delisting Conventional Pollutants

Null Hypothesis: Actual exceedance proportion is >10%
Alternate hypothesis: Actual exceedance proportion is ≤25%
Minimum confidence level is 90%
The minimum sample size necessary to achieve 90% confidence for H_A with 1 exceedance is 15.

Maximum number of sample excursions to delist as impaired for toxic substances	
Sample Size	**Excursions must be less than**
15	1
16-25	2
26-33	3
34-42	4
43-51	5
52-59	6
60-67	7
68-76	8
77-84	9
85-92	10
93-99	11
100-107	12
108-115	13
116-123	14


Conventional Pollutants
Listing Error Probability
Cumulative Error Probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
50
100
150
200
sample size = n
α
β