Effluent Guidelines (ELG) Parameter Estimation

Modified Delta Log-Normal Approach
Versus
Censoring Techniques

1

Introduction To ELG

2

ELG Basics
(www.epa.gov/waterscience/guide)

- CWA - Clean Water Act (1972)
  - **www.epa.gov/watertrain/cwa/**
- BPT - Best Practicable Control Technology
  - Setting Pollutant Specific Limits
- BAT - Best Available Technology
  - Setting Pollutant Specific Limits

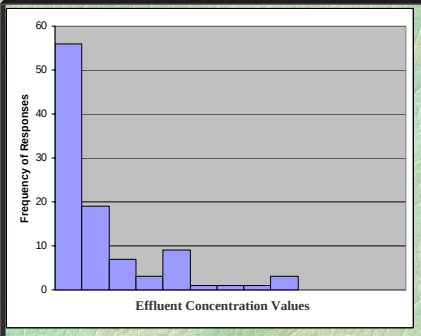
3

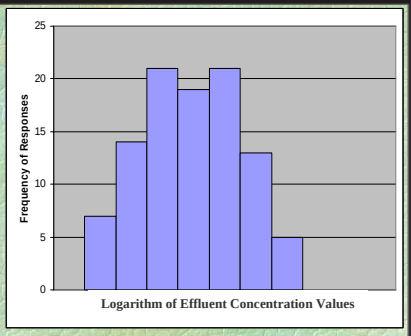
Purpose/Goal
(Comparing Parameter Estimation Approaches)

- Assume LogNormal Underlying Distribution
- Three Parameter Estimation Techniques
  - Modified Delta LogNormal Approach
  - Maximum Likelihood Estimation - Left Censored
  - Least Squares - Left Censored
- Hypothesis: Approaches Are Equivalent
  - Use Previous ELG Pollutant Data

4

Estimation Of Parameters (Log-Normal Distribution)



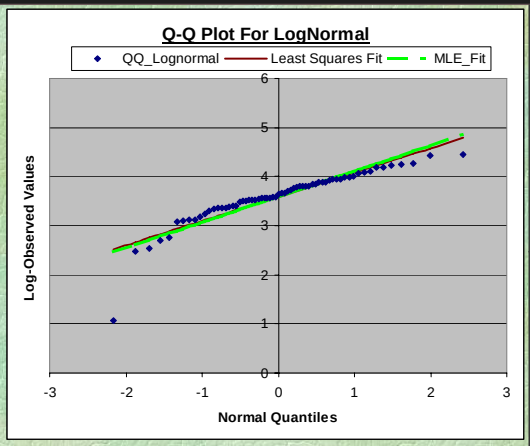


Basic Assumption

Effluent Concentration is Log-Normally Distributed

5

Estimation Of Parameters (Quantile-Quantile Plotting)



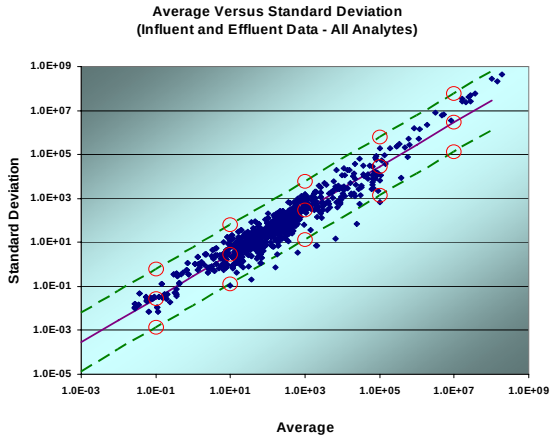
LeastSquares Estimates(Log)	
std. Dev.	0.4939
Mean	3.5880
R-Square	0.8400
Exp(Mean)	36.1632
Uppr 99%	114.1045

Max. Likelihood Estimates(Log)	
std. Dev.	0.5216
Mean	3.5979
Exp(Mean)	36.5211
Uppr 99%	122.8788

Least Square and Maximum Likelihood Approaches

6

Estimation Of Parameters (Population/Design Of Experiment)



Mean	Standard Dev.
0.1	0.0013
0.1	0.0288
0.1	0.6189
10	0.1343
10	2.883
10	61.885
1000	13.434
1000	288
1000	6188
100000	1343
100000	28833
100000	618850
10000000	134338
10000000	2883312
10000000	61884974

7

Estimation Of Parameters (Random Number Of Censored Point)

Mean	0.1	0.1	0.1	10	10	10
Std. Dev.	0.0013	0.029	0.62	0.134	2.88	61.9
Censored	8	6	31	16	19	34
Observed	44	46	21	36	33	18

Mean	10^3	10^3	10^3	10^5	10^5	10^5
Std. Dev.	13.43	288	6188	1343	28833	618850
Censored	17	14	37	22	19	35
Observed	35	38	15	30	33	17

Mean	10^7	10^7	10^7
Std. Dev.	134338	2883312	61884974
Censored	20	10	33
Observed	32	42	19

8

Using MS Excel™ To Randomly Generate Log-Normal Variables

- $SD^2 = Ln(1 + Std.Dev^2/Mean^2)$
- $M = Ln(Mean) - SD^2/2$
- L-N = Exp(NormInv(Rand(), M, SD))

9

Estimation Of Parameters Results - Least Squares Approach

Mean	Standard Dev.	Actual Uppr 99%	L.S. Mean	L.S. Std. Dev.	L.S. Uppr 99%	L.S. Err. %
0.1	0.0013	0.1032	0.10002	0.001356	0.1032	0.0417%
0.1	0.0288	0.1854	0.10057	0.026853	0.1789	3.5134%
0.1	0.6189	1.376	0.11737	0.5034	1.4636	6.3663%
10	0.1343	10.32	10.01175	0.165667	10.4032	0.8403%
10	2.883	18.54	9.63214	2.470976	16.786	9.4728%
10	61.885	138	18.31491	166.9009	267.5166	94.4161%
1000	13.434	1032	1001.352	12.70864	1031.2731	0.0366%
1000	288	1854	1035.46	302.823	1935.2751	4.3694%
1000	6188	13760	733.197	3964.251	9780.59	28.9201%
100000	1343	103165	99685.4	1279.45	102698.2	0.4525%
100000	28833	185425	98325.5	38780.4	221531.3	19.4719%
100000	618850	1376000	439871.9	12805197	63477509	361.3191%
10000000	134338	10316506	99940279	111331.5	10255764	0.5888%
10000000	2883312	18542546	97527163	19062678	15018044	19.0076%
10000000	61884974	137599990	11454163	69518753	156961782	14.0711%

Censor -34 Obs. -18

Censor - 35 Obs. - 17

10

Estimation Of Parameters
Results - Maximum Likelihood Approach

Mean	Standard Dev.	Actual Uppr 99%	MLE Mean	MLE Std. Dev.	MLE Uppr 99%	MLE Err. %
0.1	0.0013	0.1032	0.100107	0.001319	0.1032	0.0417%
0.1	0.0288	0.1854	0.096009	0.025514	0.1704	8.0977%
0.1	0.6189	1.376	0.1158	0.59102	1.5217	10.5887%
10	0.1343	10.32	10.03587	0.158348	10.4098	0.9043%
10	2.883	18.54	9.45247	2.3904	16.3543	11.8010%
10	61.885	138	17.70743	228.5424	264.5232	92.2407%
1000	13.434	1032	1002.409	13.55731	1034.3533	0.2620%
1000	288	1854	1073.05	294.0674	1935.4182	4.3772%
1000	6188	13760	661.2836	3152.494	8523.69	38.0546%
100000	1343	103165	99556.2	1345.73	102727.04	0.4246%
100000	28833	185425	102203	36134.9	214106.6	15.4677%
100000	618850	1376000	438062.2	17255693	6090329	342.6112%
10000000	134338	10316506	9969986.3	128662.97	10272971	0.4220%
10000000	2883312	18542546	9584963.9	1745706.8	14355374	22.5814%
10000000	61884974	137599990	10566825	49704389	135662089	1.4084%

Censor -34
Obs. -18

Censor - 35
Obs. - 17

11

Estimation Of Parameters
Results - Modified Delta LogNormal Approach

Mean	Standard Dev.	Actual Uppr 99%	Mod Delt Mean	Mod Delt Std. Dev.	Mod Delt Uppr 99%	Mod Delt Err. %
0.1	0.0013	0.1032	0.10014	0.001175	0.1029	0.2491%
0.1	0.0288	0.1854	0.101725	0.0244907	0.1694	8.6370%
0.1	0.6189	1.376	0.11318	0.20099	0.9443	31.3735%
10	0.1343	10.32	10.04624	0.12326	10.3548	0.3711%
10	2.883	18.54	10.11104	1.89408	15.572	16.0200%
10	61.885	138	14.7039	36.09944	157.7541	14.6469%
1000	13.434	1032	1003.3403	10.1107	1027.6791	0.3850%
1000	288	1854	1076.676	246.5648	1788.9328	3.5228%
1000	6188	13760	802.275	1351.351	6526.98	52.5655%
100000	1343	103165	100038.1	872.76	102311.98	0.8269%
100000	28833	185425	106563.9	29276.1	199008.1	7.3251%
100000	618850	1376000	247726.6	964263.3	3457789.5	151.2929%
10000000	134338	10316506	10013335	86587.4	10224668	0.8902%
10000000	2883312	18542546	10007766	1568795.7	14325898	22.7404%
10000000	61884974	137599990	10699754	21044740	96892077	29.5842%

Censor - 35
Obs. - 17

12

Comparing Approaches
Using Error For 99th Percentile

	# Lowest Percents	Averages
MLE	3.5	37.53
LS	6.5*	36.62
MDL	5	22.70**

* LS is ranked most accurate using lowest errors
** MDL ranked most accurate using average error

13

Comparing Approaches
Using Error For 99th Percentile

Fit Error Data To Best Distribution Function

Probability Error Percentage Will Not Exceed Threshold						
Threshold	0.1%	1 %	10 %	100 %	200 %	300 %
Least Squares	0.12	0.41	0.76	0.95	0.97	0.98
MLE	0.11	0.39	0.75	0.95	0.97	0.98
Modified Delta	0.02	0.22	0.68	0.96	0.98	0.99

14

Conclusions

- When few Censored Points LS does best
- MDL Does Best When many Censored Pts.
  - Protects Against Large Errors Best
- No One Approach Is Superior To Another
  - Each has utility
  - Performance depends on data set involved
- Analysis Is Restricted; Could Broaden

15