BSTT562 Project I

Ruizhe Chen, Hesen Li

November 29, 2018

1 Abstract

In this project, we want to reproduce the results presented in Table 3 Interlaboratory Data for Cadmium from Bhaumik and Gibbons [2005]. # Introduction

A key characteristics of the data is its heteroscedasticity.

2 Data

This data is the experimental data fro cadmium from an interlaboratory study conducted by Ford Motor Company (J. Phillips, personal communication). These data were generated as part of a blind interlaboratory study of laboratories that hold Michigan State Drinking Water Certifications for the parameters tested. Samples were prepared by an independent source, randomized, and submitted on a weekly basis over a 5-week period. Cadmium was analyzed by inductively coupled plasma atomic emissions spectroscopy using EPA method 200.7. The dataset used for this example comprised five replicates at each of threee concentrations (0, 20 and 100 $\mu g/L$) in each of the $q = 5$ laboratories. Using the fisrt replicate from the first three laboratories as the new measurement (ie, $q' = 3$), we would like to reproduce the results of point estimates, variances, confidence intervals and simulated confidence levels for the true concentrations in each of the three cases. The data is displayed as follows:

Table 1: Interlaboratory Data for Cadmium (ug / L)

Lab	Replication	0	20	100
1	1	-3.000	10.000	92.000
1	2	4.000	20.000	100.000
1	3	-4.000	17.200	97.800
1	4	3.000	24.000	100.000
1	5	3.100	19.100	109.000
2	1	-0.060	17.815	90.455
2	2	0.010	17.305	87.610
2	3	0.115	16.570	85.550
2	4	-0.055	17.360	89.925
2	5	0.340	18.120	90.070
3	1	-7.400	27.100	107.400
3	2	-2.100	19.400	108.100
3	3	-11.400	9.000	83.800

Lab	Replication	0	20	100
3	4	-11.100	10.500	81.900
3	5	-1.400	19.300	94.200
4	1	1.000	21.000	96.000
4	2	-2.126	16.049	90.650
4	3	0.523	16.082	89.388
4	4	-2.000	17.000	91.000
4	5	-0.551	15.489	85.867
5	1	0.000	18.000	91.000
5	2	0.000	19.000	101.000
5	3	0.000	19.000	102.000
5	4	-1.000	18.700	92.700
5	5	0.038	19.790	99.884

3 Method

To measure the true concentration of an analyte, x, the traditional way is to propose a simple linear calibration model:

$$y = \alpha + \beta x + e$$

with normal assumption on errors. But as the data is hetergeneous, this models fails to explain the increasing measurement variation with increasing analyte concentration commonly observed in analytical data. To overcome this drawback of this simple model, we propose a log-normal model:

$$y = xe^{\eta}$$

4 Analysis

We choose the Method of Meoments to estimate the model parameters.

Why do we choose the Method ofMoments: A good property of MOM: the estimates obtained by MOM are asymptotically efficient.

4.1 The model

$$y_{ijk} = \alpha_i + \beta_j x_j e^{\eta_{jik}} + e_{jik}$$

5 Table/Discussion

Reference

Dulal K Bhaumik and Robert D Gibbons. Confidence regions for random-effects calibration curves with heteroscedastic errors. *Technometrics*, 47(2):223–231, may 2005.