Paper JM01-2013

Dealing with Data below the Detection Limit: Limit Estimation and Data Modeling

Mark Bailey, SAS Institute Inc., Haddonfield, NJ
Diane K. Michelson, SAS Institute Inc., Austin, TX

ABSTRACT

Measuring trace levels of contaminants in chemicals or gases can be difficult. When the signal is very small, it can be lost in the noise. Chemical analyses are characterized by their accuracy, precision, and linear range. A detection limit is the smallest amount the can be reliably detected by the procedure. The procedure can be used on a blank, with no amount of the substance, or on a sample containing the substance to be measured. Various methods of estimating a detection limit are compared. We will examine the use of prediction intervals on measurements on blanks, as well as regression approaches, including the use of linear regression by ordinary and weighted least squares. A contamination example will be demonstrated, using Fit Y by X, and Fit Model in JMP. Responses below the detection limit can be included in your data analysis. Ad hoc approaches produce biased estimates and should be avoided. Such responses are censored data, and likelihood methods exist to handle censoring. An example of a designed experiment is handled using the Parametric Survival personality in Fit Model in JMP.

INTRODUCTION

To learn about a system or a process, there must be variation. If the characteristics or outcomes never change, then it is impossible to learn anything. Thus, we design experiments to provoke a large change in the response, in the hope that the analysis will be more informative, both in kind (factor effects) and degree (precision). The determination of the response requires a measurement that is accurate (unbiased) and precise over a useful range. Many physical quantities are bounded by zero and all measurements are limited by noise. The background signal, when the response is absent or zero, can be translated in various ways into an upper bound on measurement, or limit of detection (LOD). How do you estimate the limit of detection? What value should you use in your analysis for a response reported to be below the limit of detection?

PART 1: LIMIT ESTIMATION

The concentration of an element in a substance (an *analyte*) may be a key characteristic of the quality of a chemical process. Analytical chemistry provides information about the concentration of the analyte. To estimate the detection limit of the measurement method, we would like to find the lowest level of concentration where the results become indistinguishable from a zero reading. This is the point at which we would start to see signals for zero concentrations. Different methods for finding the LOD exist in the literature. Industry standards, including SEMI C10 (Guide for Determination of Method Detection Limits), ISO:11843 (Capability of Detection), and the IUPAC Compendium of Analytical Nomenclature, discuss methodologies for determining the detection limit. These and others are listed in the References.

Most authors advise the use of the calibration curve to fit a linear model. The levels of the known concentrations are spaced over the range of interest, including a zero level, or blank. After measuring concentrations for these known samples, a regression line is fit to the data.

A common method of estimating the detection limit is to first estimate the standard deviation of the response at the zero level. Estimation methods include the standard deviation of the data at the zero level, the square root of the mean square error of the regression line, or the standard deviation of the *y*-intercept of the regression line. Then use a multiple of this estimated standard deviation of the response at the zero level divided by the slope of the calibration curve to find the detection limit.

In this paper, we recommend a different approach. We will examine various methods to fit a linear model to a calibration curve, followed by estimating a prediction interval for future observations at the zero level, and finally using inverse prediction of the upper bound to estimate the LOD. Our recommendation is to use *ordinary least squares* (OLS) *regression* to fit a linear model to the calibration curve. Other methods popular in the literature are based on weighting schemes.

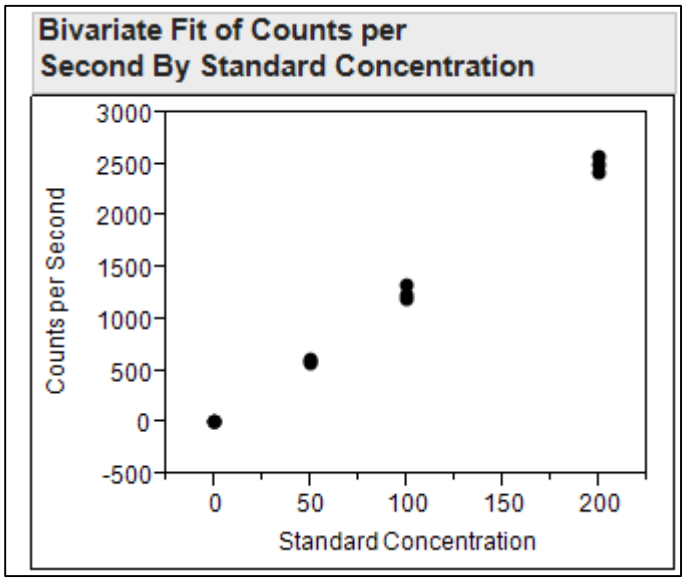
A $100(1-\alpha)\%$ prediction interval is a range of values of a variable with the property that if many such intervals are calculated for many samples, $100(1-\alpha)\%$ of them will contain one observation from a future realization of the process. A prediction interval is different from a confidence interval. A confidence interval gives limits in which we expect a population parameter, such as a mean or variance, to lie. A prediction interval gives limits in which we expect a future individual observation to lie.

Formulas for prediction interval calculations can be found in the literature. The formula depends on the number of future observations to be predicted. See Ramirez[1] for an accessible discussion of statistical intervals.

As an example, consider an inductively coupled plasma-mass spectrometry (ICP-MS) system that performs trace element analysis for the purpose of determining the amount of contaminants in a sample. A sample containing a known amount of a contaminant is measured, and the response is signal intensity, measured in counts per second (cps). The experiment is repeated for four known levels, including no contaminant at all.

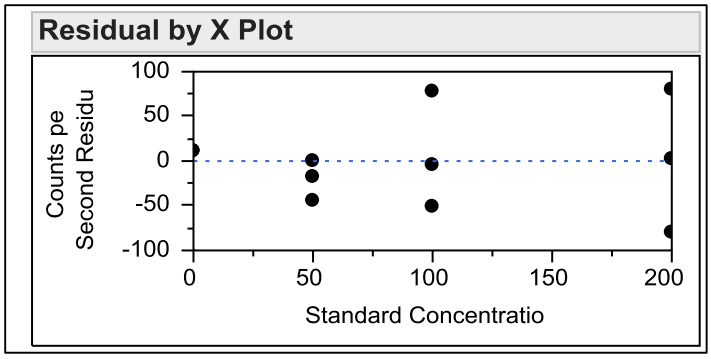
	Standard Concentration	Counts per Second
1	0	0.01
2	0	0.02
3	0	0.01
4	50	597.027
5	50	570.511
6	50	615.541
7	100	1319.319
8	100	1237.294
9	100	1188.903
10	200	2496.103
11	200	2574.509
12	200	2413.543

A graph of the measured values against the known values shows that the data seem to follow a linear trend. It is notable that the variability at the zero concentration level is much less than that at other levels.



It can be hard to see the difference in variability at the zero level by using a scatterplot of the data. Instead, fit a line to the data and examine a plot of the residuals by the predictor variable. This residual plot clearly shows the variance is much higher where the element exists than where it doesn't.

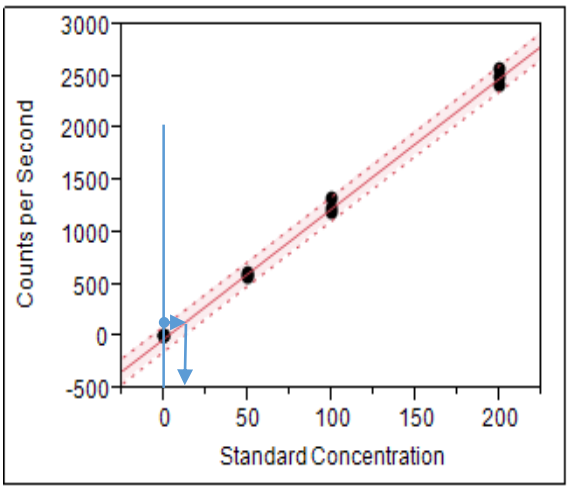
[1] Ramírez, José G., "Statistical Intervals: Confidence, Prediction, Enclosure," retrieved from http://www.jmp.com/software/whitepapers/



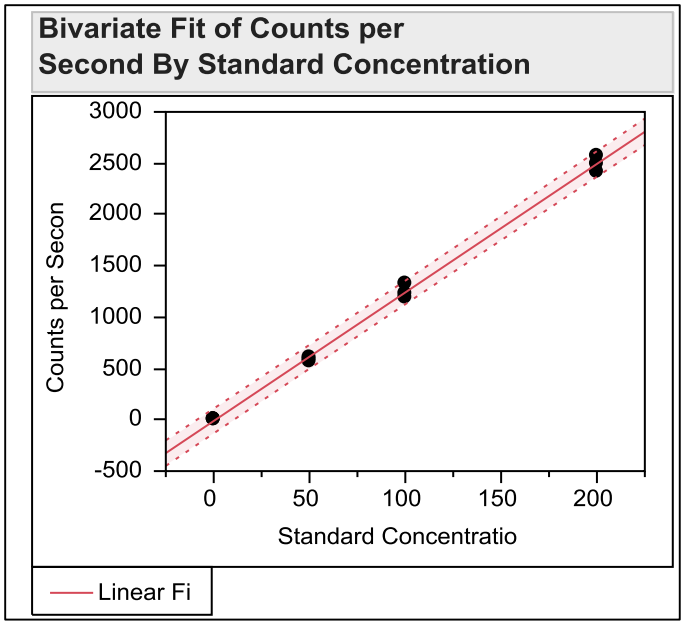
Many physical characteristics show *heteroskedasticity*, or the variance of the response changing over the range of the data. The variability of physical data often increases as the response increases. Ordinary least squares regression makes the assumption of *homoskedasticity*, or equal variances across the range of data. Therefore, naively, it might make sense to pursue methods to stabilize the variance.

One method of variance stabilization is to use *weighted least squares regression*, which weights the data by the inverse of the variance within each group. Most calibration problems do not use samples that are large enough to estimate group variances with low uncertainty. Therefore, another popular method is to actually model the variance as well as the mean. JMP implements this model using the **LogLinear Variance** personality of **Fit Model**. We will demonstrate OLS and the two methods of WLS regression.

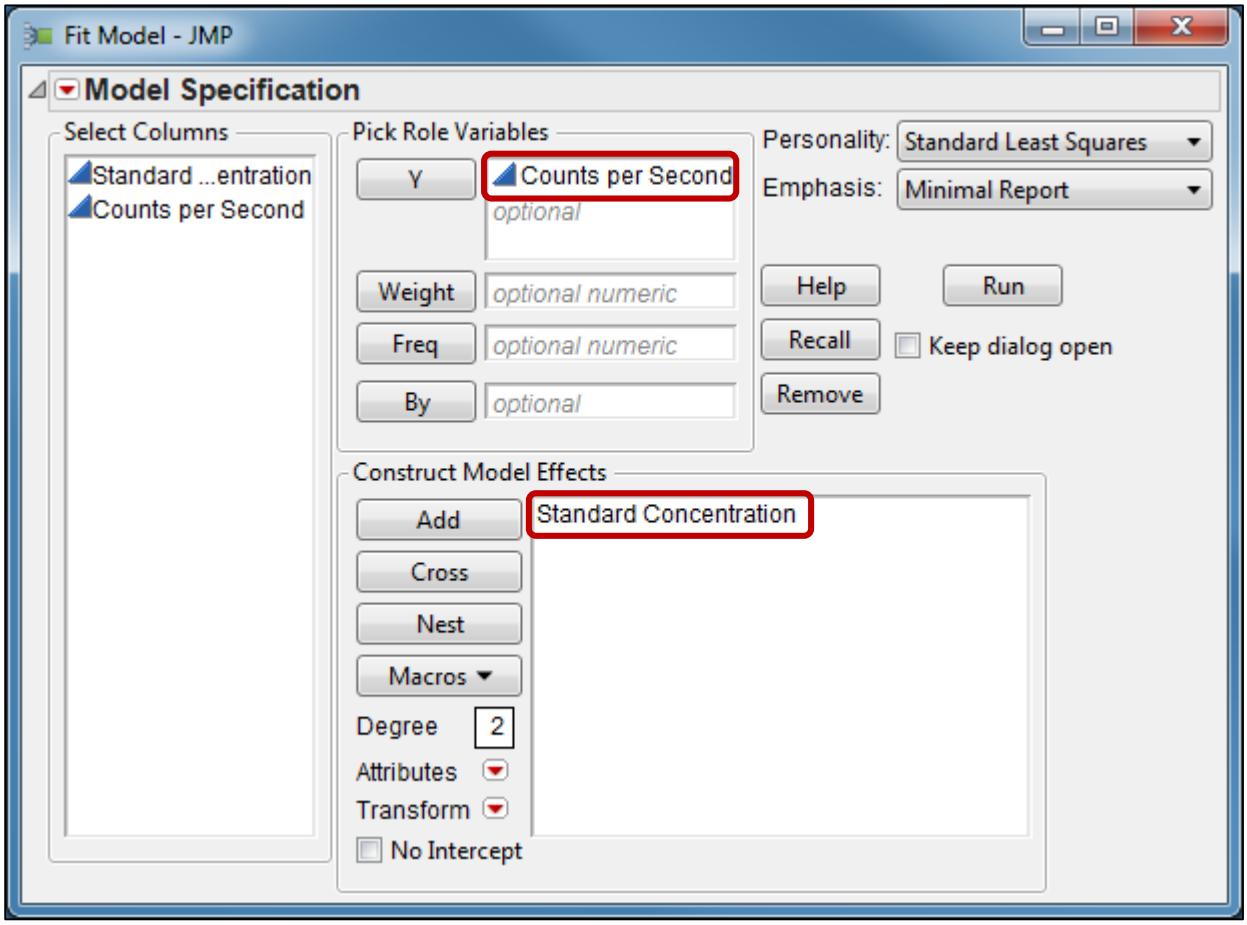
Heteroskedasticity does not affect the bias of the regression coefficients (slope and intercept) using ordinary least squares. It does affect the standard errors of the regression coefficients. However, for typical data like our simulated data, the variance of the cps for the blank is much less than that for the non-blanks, and the variance of the non-blanks is what we would like to use to find prediction intervals for the detection limit. Therefore, we recommend the use of ordinary least squares regression to find the prediction interval, followed by inverse prediction of the upper prediction limit to the regression line as shown in the following graphic.



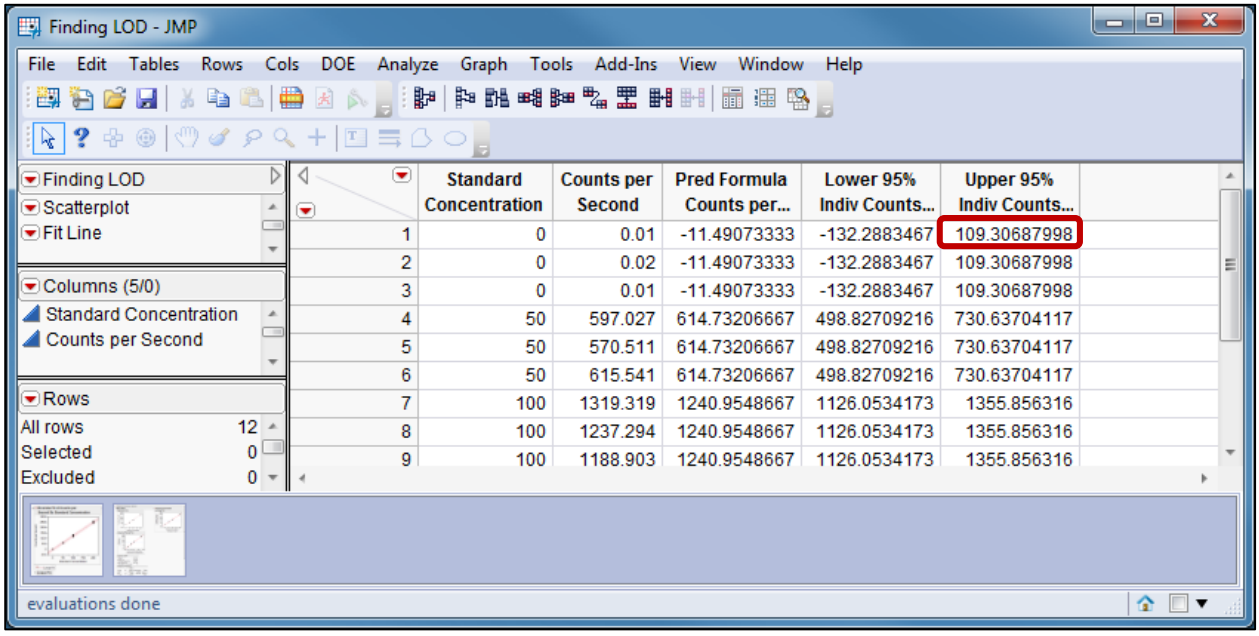
The **Fit Y by X** platform can be used to visualize the prediction intervals. From the red triangle next to **Bivariate Fit**, select **Fit Line**. From the red triangle next to **Linear Fit**, select **Confidence Curves Indiv** and **Confidence Shaded Indiv**.



JMP does not provide the means to save the prediction interval from the **Bivariate** platform, so we will use it for visualization only, and use **Fit Model** to gain more information about the fit.

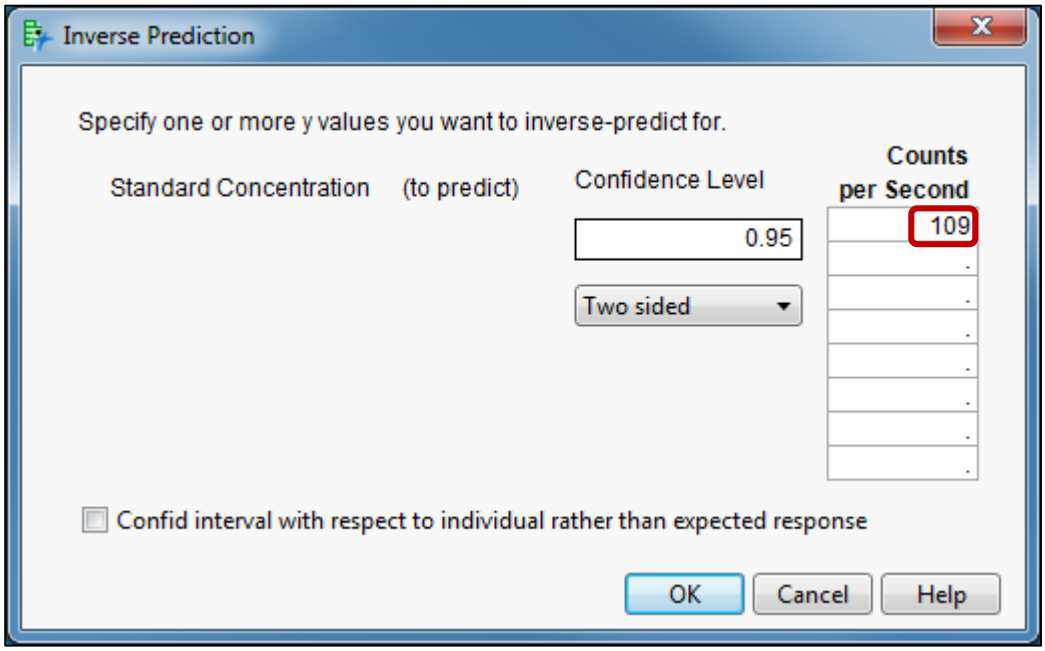


After running the model, save the prediction formula and prediction interval. From the red triangle next to **Response Counts per Second**, select **Save Columns → Prediction Formula** and **Save Columns → Individual Confidence Limit Formula**.

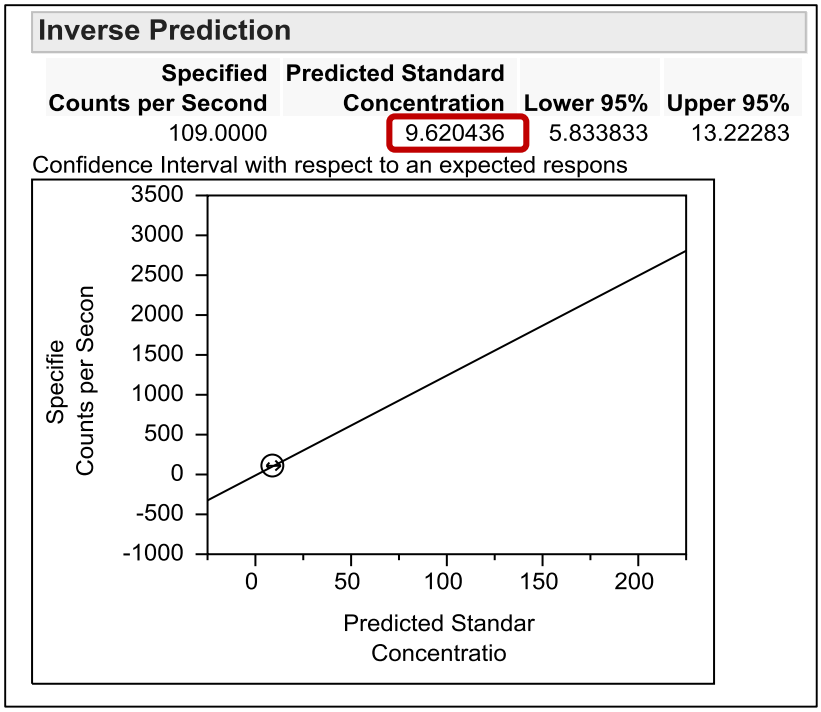


	Standard Concentration	Counts per Second	Pred Formula Counts per...	Lower 95% Indiv Counts...	Upper 95% Indiv Counts...
1	0	0.01	-11.49073333	-132.2883467	109.30687998
2	0	0.02	-11.49073333	-132.2883467	109.30687998
3	0	0.01	-11.49073333	-132.2883467	109.30687998
4	50	597.027	614.73206667	498.82709216	730.63704117
5	50	570.511	614.73206667	498.82709216	730.63704117
6	50	615.541	614.73206667	498.82709216	730.63704117
7	100	1319.319	1240.9548667	1126.0534173	1355.856316
8	100	1237.294	1240.9548667	1126.0534173	1355.856316
9	100	1188.903	1240.9548667	1126.0534173	1355.856316

The estimate of the upper 95% prediction limit at zero is 109. Use inverse prediction to estimate the true concentration at this value. Return to the **Fit Least Squares** report. From the red triangle next to **Response Counts per Second**, select **Estimates → Inverse Prediction…** . Enter 109 in the first blank space and click **OK**.

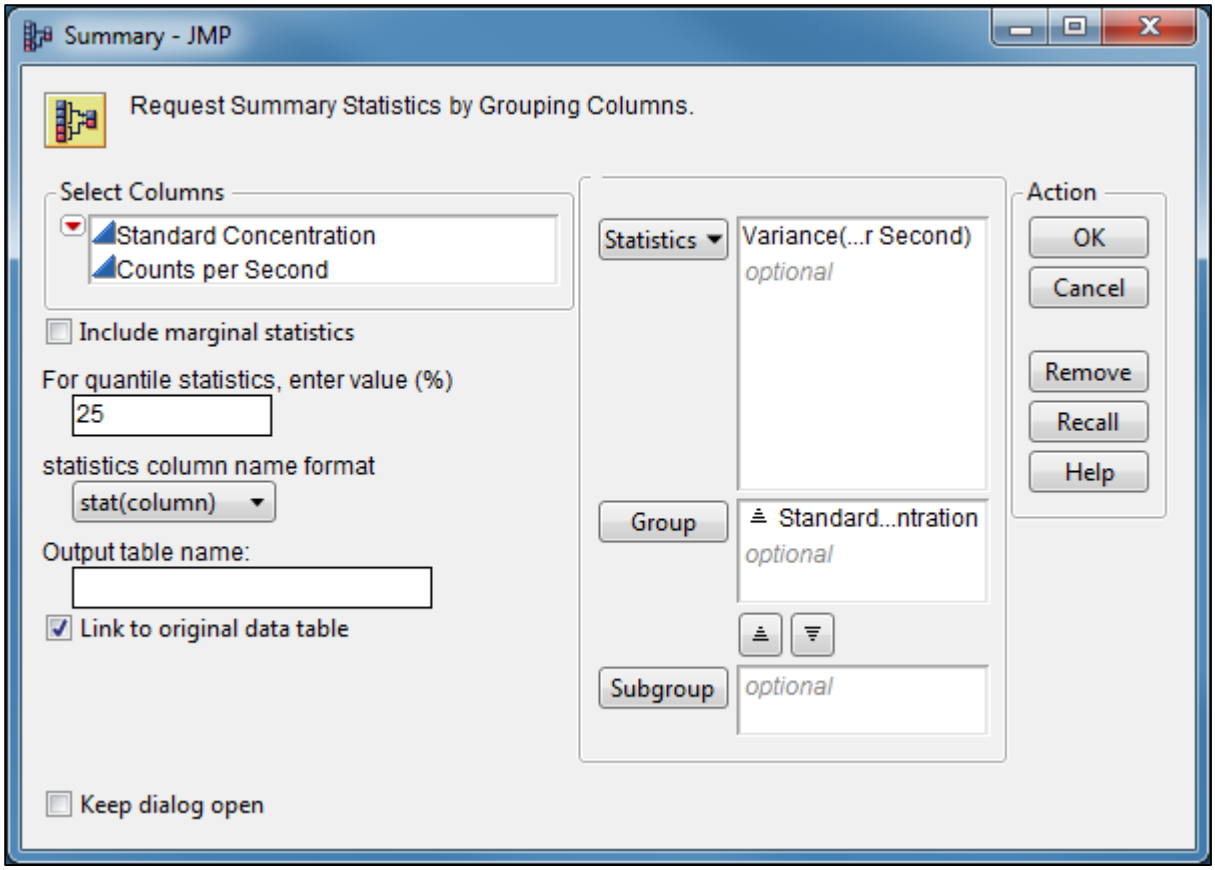


The predicted value of the response is 9.6, and that is our estimate of the detection limit.

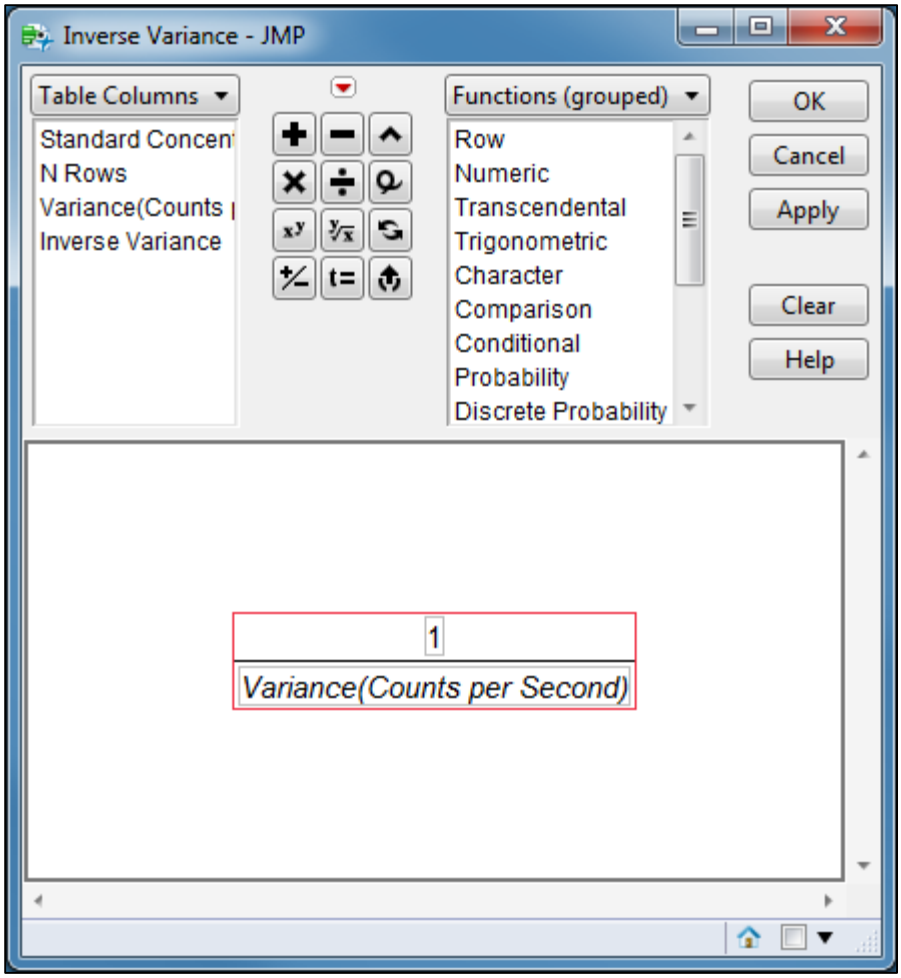


Some industry standard documents, notably SEMI C10, recommend weighted least squares (WLS) regression. Let's compare the inverse prediction from the OLS prediction interval with that derived from WLS.

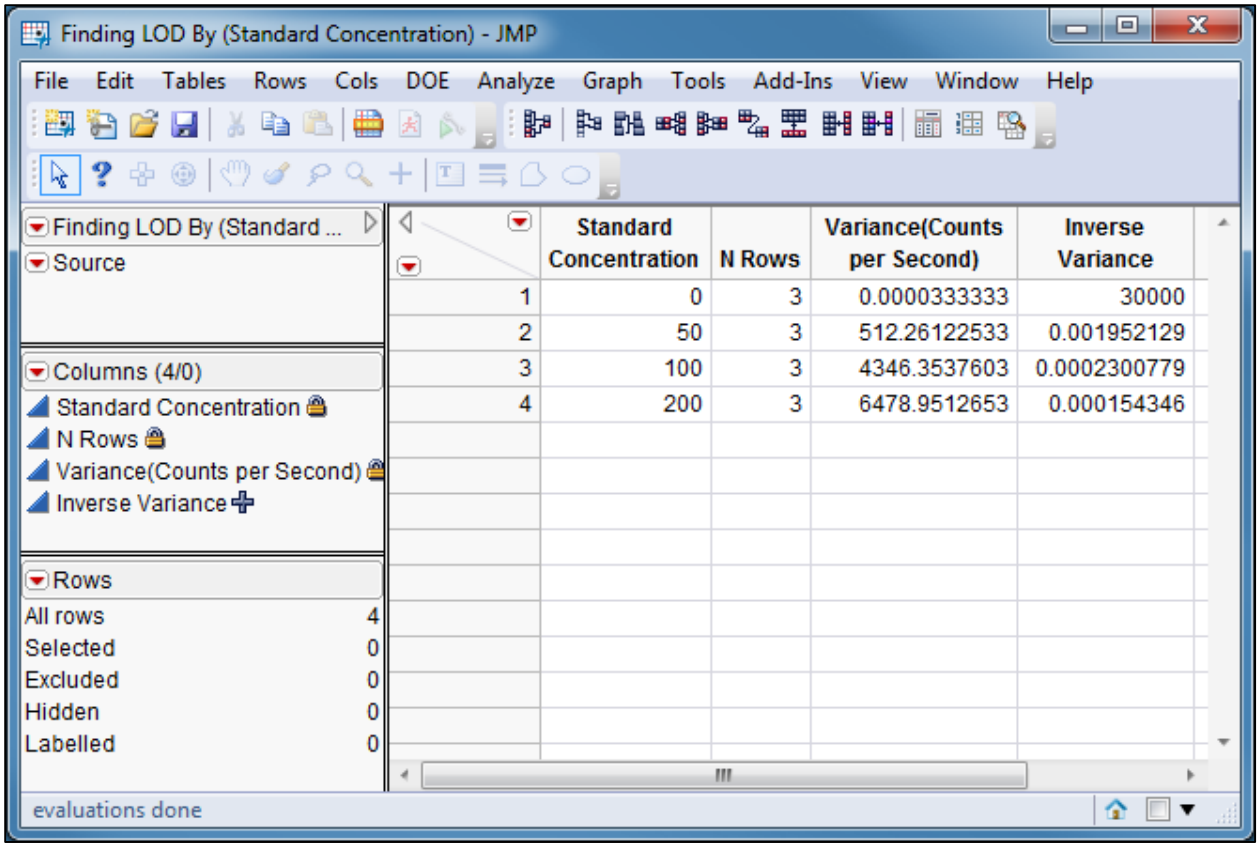
Weighted least squares uses weights on the data in order to stabilize the variance. An *ad hoc* method for weighting is to use the inverse variance of groups as the weights. To find the variance of each group, summarize the data using **Tables** → **Summary**. Add the **Standard Concentration** as a grouping column and select **Counts per Second** then click **Statistics** → **Variance** to ask for the variance of cps by group. Click **OK**.



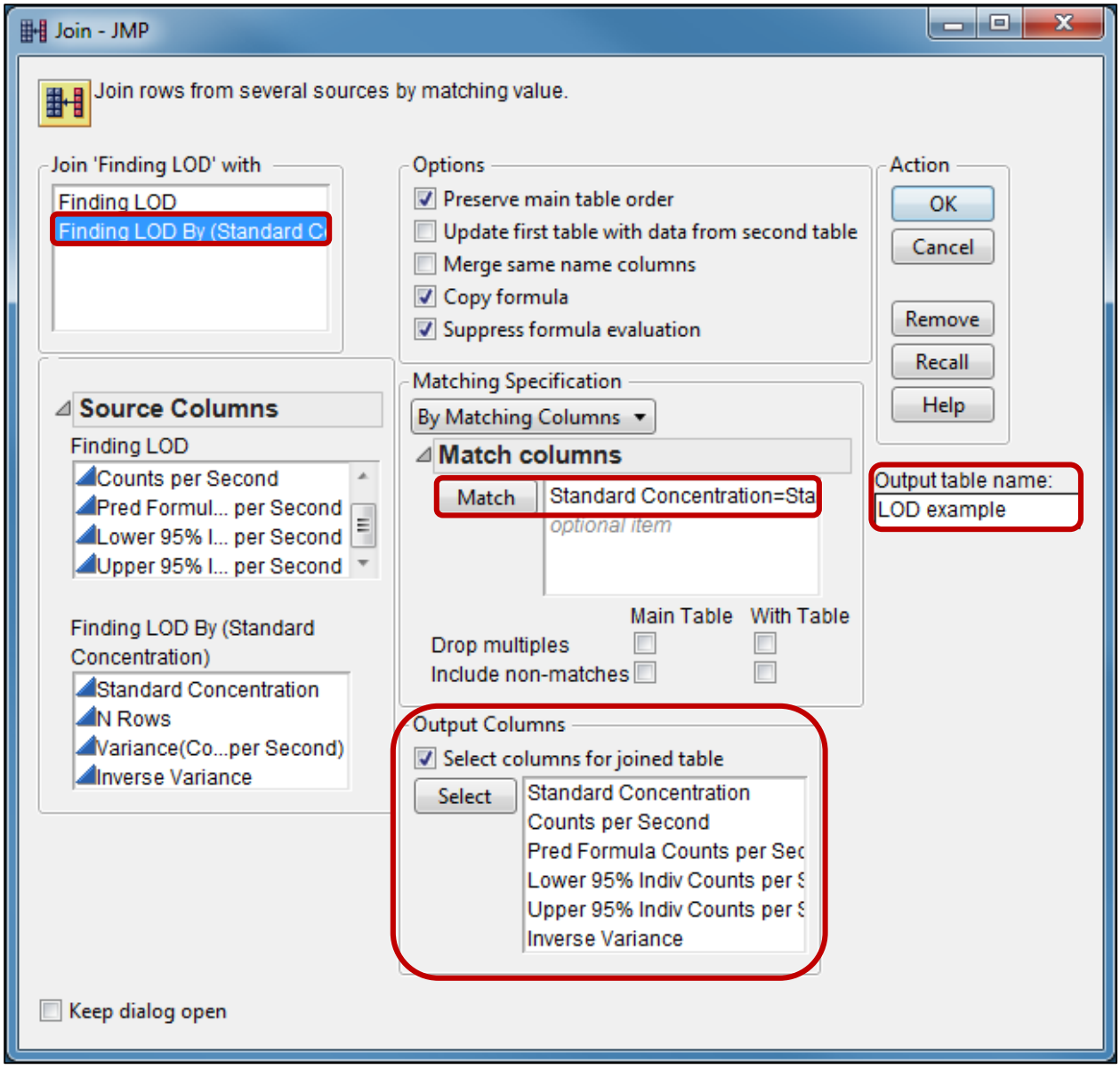
Next, add a new column containing a formula to the summary table. This formula will contain the reciprocal of the variance of each group.



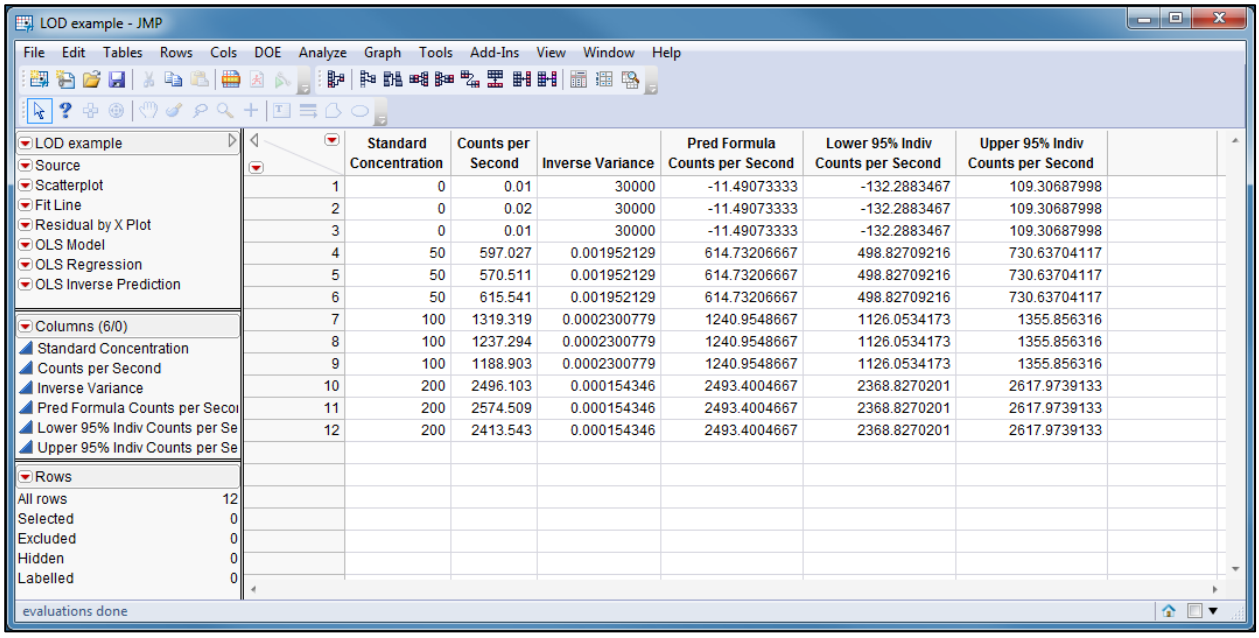
The summary table contains the grouping variable, the sample size and variance for each group, and the inverse variance for each group.



The inverse variance needs to go back into the original data table. Return to the original table and use **Tables** → **Join** to add the column of inverse variances.

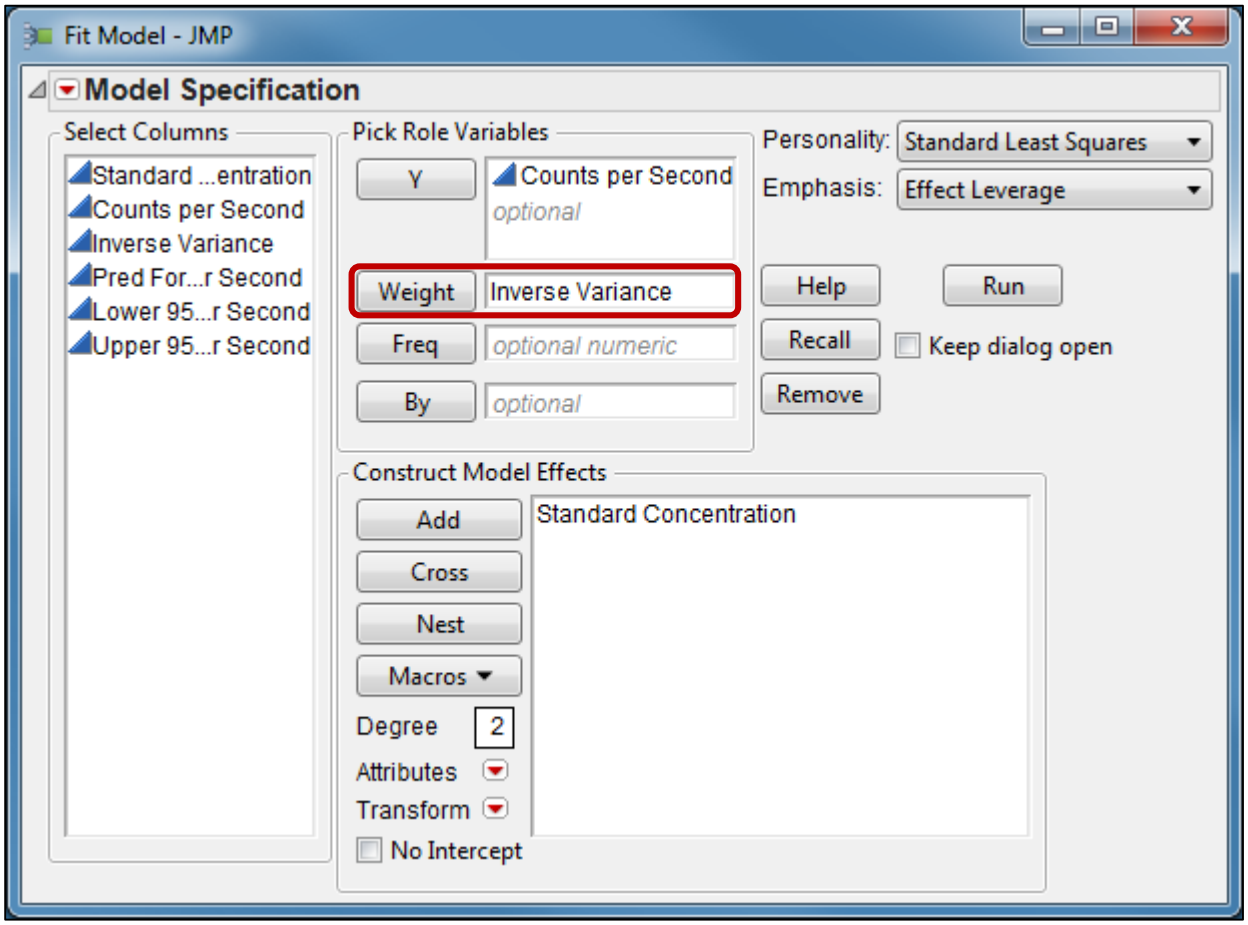


The resulting table contains the columns from the original table in addition to the new weighting column.

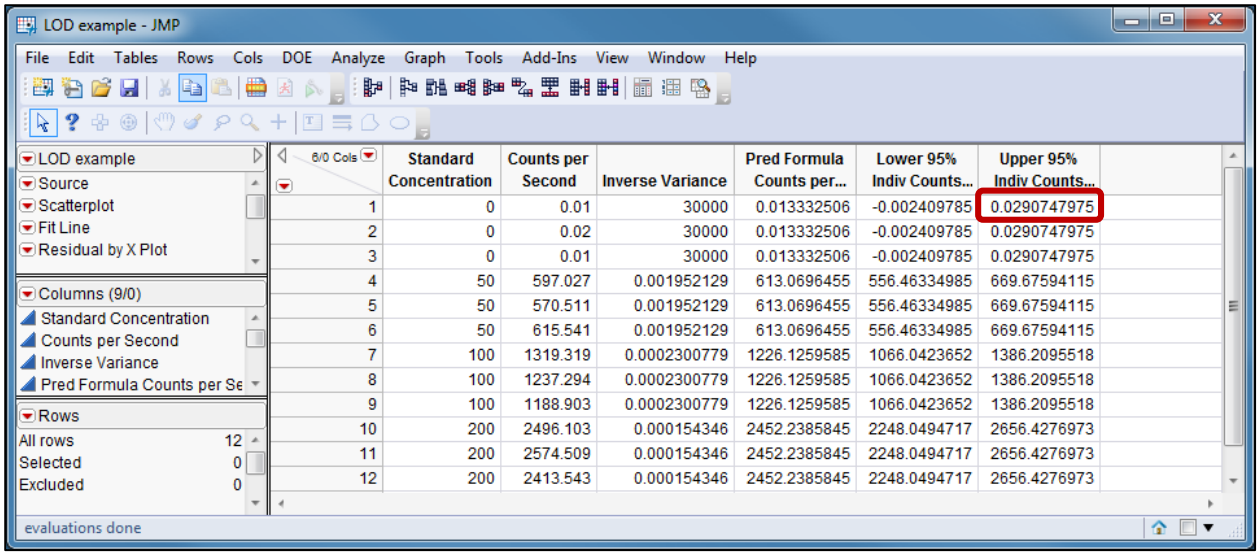


	Standard Concentration	Counts per Second	Inverse Variance	Pred Formula Counts per Second	Lower 95% Indiv Counts per Second	Upper 95% Indiv Counts per Second
1	0	0.01	30000	-11.49073333	-132.2883467	109.30687998
2	0	0.02	30000	-11.49073333	-132.2883467	109.30687998
3	0	0.01	30000	-11.49073333	-132.2883467	109.30687998
4	50	597.027	0.001952129	614.73206667	498.82709216	730.63704117
5	50	570.511	0.001952129	614.73206667	498.82709216	730.63704117
6	50	615.541	0.001952129	614.73206667	498.82709216	730.63704117
7	100	1319.319	0.0002300779	1240.9548667	1126.0534173	1355.856316
8	100	1237.294	0.0002300779	1240.9548667	1126.0534173	1355.856316
9	100	1188.903	0.0002300779	1240.9548667	1126.0534173	1355.856316
10	200	2496.103	0.000154346	2493.4004667	2368.8270201	2617.9739133
11	200	2574.509	0.000154346	2493.4004667	2368.8270201	2617.9739133
12	200	2413.543	0.000154346	2493.4004667	2368.8270201	2617.9739133

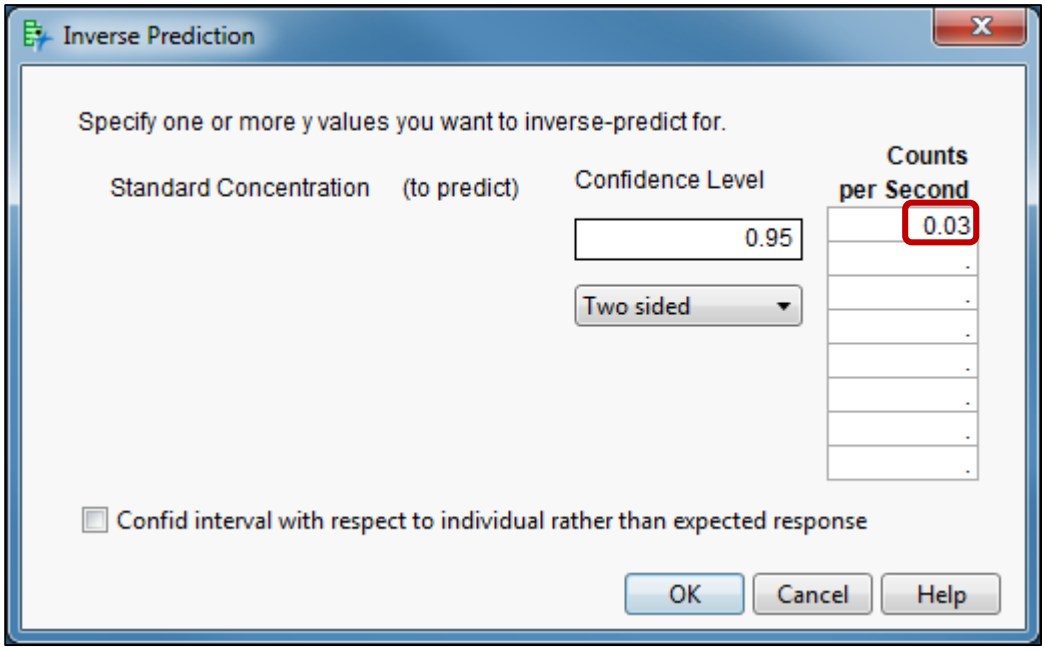
The regression model can now be fit again, this time using the **Inverse Variance** column in the **Weight** role.



Again, save the prediction formula and prediction interval to the data table by clicking the red triangle and selecting **Save Columns → Prediction Formula** and **Save Columns → Indiv Confidence Limit Formula**. The OLS regression prediction columns have been hidden in the screenshot below.



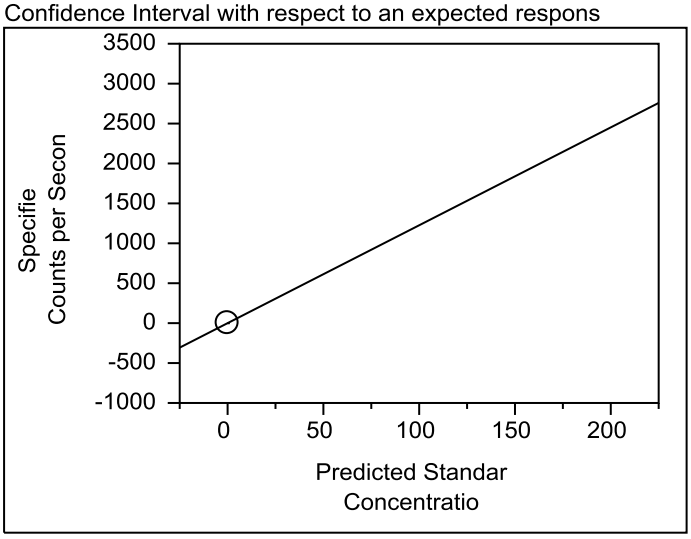
The estimate of the upper 95% prediction limit at zero is 0.03, four orders of magnitude smaller than that found by using OLS regression! Use inverse prediction once again to find the estimate of the detection limit. From the red triangle, select **Estimates → Inverse Prediction…** . Enter 0.03 and click **OK**.



Inverse Prediction

Specified Counts per Second	Predicted Standard Concentration	Lower 95%	Upper 95%
0.0300000	0.0013594	0.0007171	0.0020042

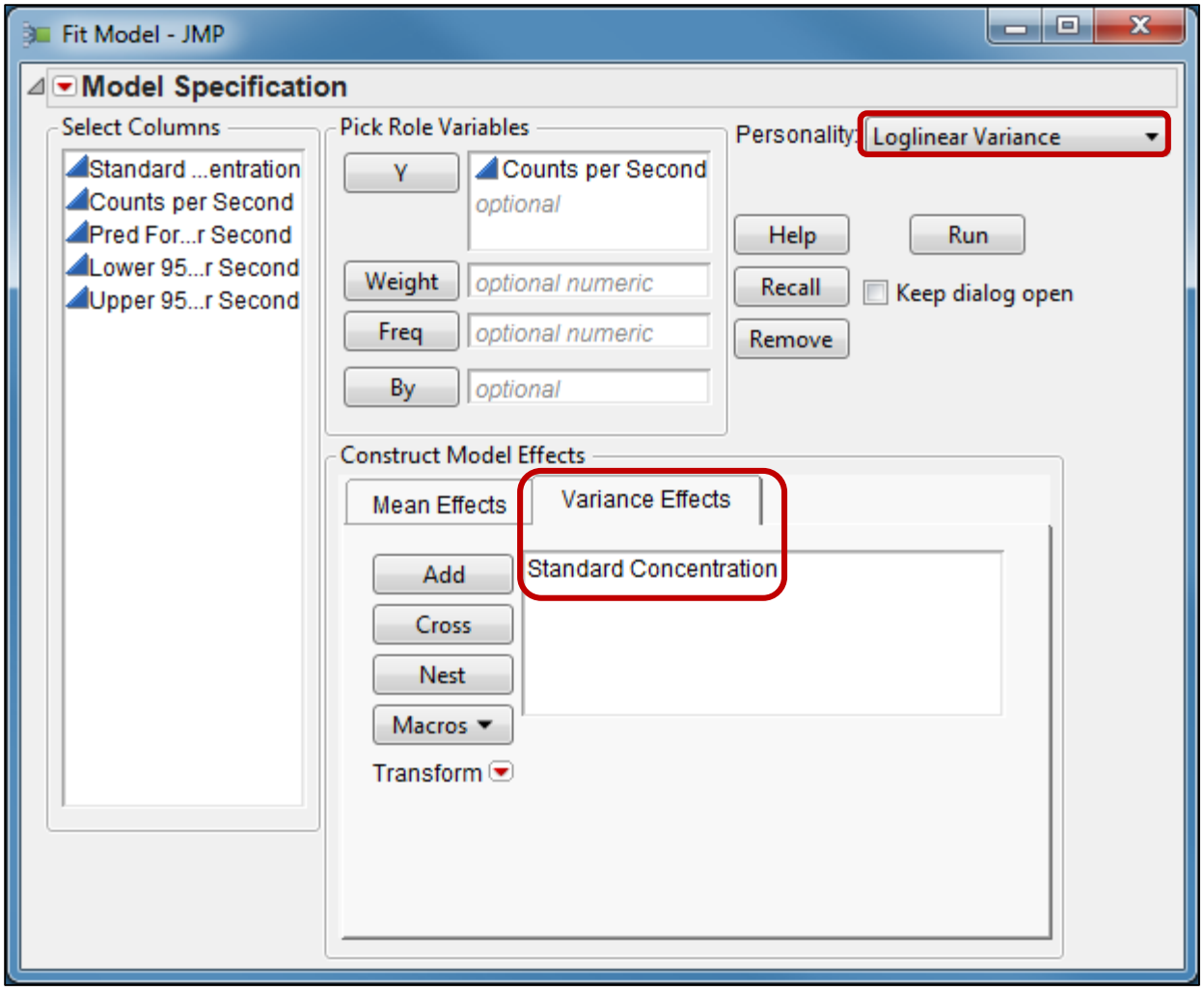
Confidence Interval with respect to an expected respons



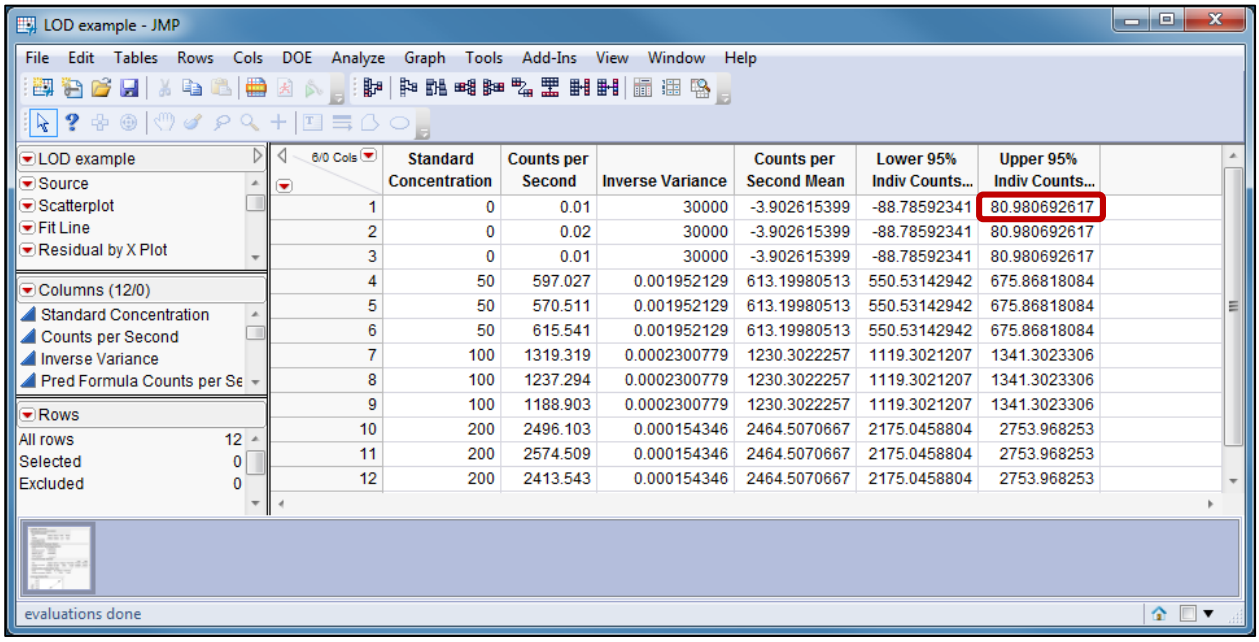
The prediction is 0.0014, which is the estimate of the detection limit. This number is much smaller than that given by OLS regression (9.6). It may seem like a smaller number would be better, but WLS can lead to estimates that are known to be ridiculously small by the practitioner. For example, matching systems is sometimes done using the LOD. LODs that are too small can lead to systems that are not matched statistically, but for all practical purposes, can be considered to measure the same. This can lead to problems justifying the use of good systems with auditors.

One reason the prediction limit is so small, and thus so is the estimated detection limit, is that the group variance for the zero level is much less than the group variance for the other levels. Therefore, the inverse variance of the zero level is huge compared with the others. One remedy for this is to model not only the mean response but the variance of the response as well. This method is particularly useful if you can assume the variance is proportional to the mean, for example.

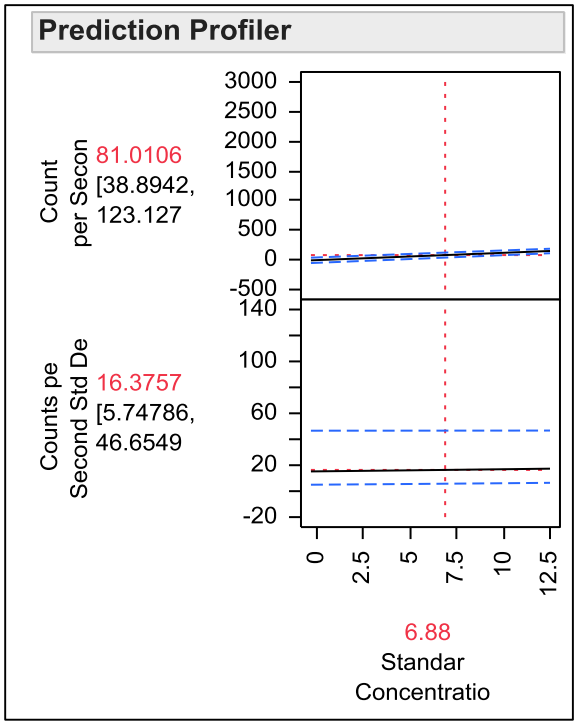
Using **Fit Model**, specify the **Loglinear Variance** personality. Select the **Variance Effects** tab and add the **Standard Concentration**, then run the model.



Save the prediction formula and prediction interval by clicking the red triangle next to **Loglinear Variance Fit** and selecting **Save Columns** → **Prediction Formula**, then again, **Save Columns** → **Indiv Confidence Interval**.



The upper prediction limit is 81. The **Fit LogVariance** report does not allow for inverse prediction, but it is simple to do with the **Profiler**. Return to the **Fit LogVariance** window. From the red triangle, select **Profilers** → **Profiler**. Drag the **Standard Concentration** slider to the left until the prediction of **Counts per Second** is near 81. You can adjust the scale of the *x* axis if necessary to zoom in on the region of interest.



The inverse prediction is around 6.9, still a smaller number than 9.6 that was found using OLS regression, but much more reasonable than 0.0014 found using inverse variance weights.

In any case, we recommend OLS regression simply because we don't want the variance of the zero level to overwhelm the calculation of the prediction interval at the zero level. We have shown that the variance of the zero level lead to very small estimated detection limits when using the inverse variance as weights. This even happens

when modeling the variance explicitly, with the **LogLinear Variance** personality. Using OLS regression leads to a more conservative estimate of the detection limit.

PART 2: DATA MODELING

There are many intuitive practices for selecting the value for analysis when the response is below this threshold. Some analysts use 0, other analysts use the LOD itself, and still others split the difference and use half of the LOD. Finally, some analysts regard such a case as indeterminate and so leave the value missing. These *ad hoc* approaches unfortunately do not address the central problem but instead introduce bias in any estimates, such as model parameters. A missing value reduces the sample size and, therefore, the power of the analysis, as if nothing is known about the response when, in fact, there is information available. Using 0 biases your estimates downward, using the LOD biases your estimates upward. Using half the LOD might average out the bias, if you are optimistic and tend to be lucky. Isn't there a better way? Isn't there a rigorous approach based on statistical theory that eliminates this bias and allows you to use all of the data?

The solution is found in an unrelated field of study that has nothing to do with chemistry or any other physical science. Investigators encountered the same problem in the beginning of formal *survival analysis*. In this analysis, the response is the *life time* or the *time-to-event* where the event is death or the onset of disease. Subjects often survive or never incur the disease during the study period. What do to with their data? Ignoring it or using an arbitrary value would introduce bias as described above. Analysts realized that two kinds of data existed in these studies: for one kind, the actual life time is known, and for the other kind, it is a lower bound on the actual life time. The second kind is called *censored data*. These life times are *right-censored* because the actual life time is greater than the observation and would appear farther to the right on a number line. In the same way, the responses that are below the LOD are called *left-censored* data.

Ordinary least squares regression is not able to analyze censored responses but *maximum likelihood estimation* accommodates censoring directly through the likelihood function.

JMP provides a *parametric analysis* of survival models with multiple factors, which suits the case of our experiment. The normal distribution is not available for the likelihood function but the log-normal distribution is available to model the statistical errors. We merely transform the response by exponentiating the response as e^y for the analysis and then transform back using the natural logarithm after fitting the model for prediction. The following fictitious example illustrates the points above and shows how to use JMP for such an analysis.

A hypothetical experiment will be used to demonstrate censoring with responses below the LOD. The experiment includes four continuous factors, **X1**-**X4**. The response **Y** is simulated without censoring and then an arbitrary LOD is applied. Perhaps Y is the level of a chemical impurity that you intend to minimize through judicious selection of levels for the factors in a purification process. An arbitrary LOD (15) was selected to cause a few responses to become censored. A thorough study of this matter would involve many simulated data sets. The single simulated set of responses here is presented only to illustrate the problem and how to deal with it.

The setup for this problem in JMP is simple. The original columns for the experiment are in the left red box in the figure below. Three columns were added to facilitate the analysis as seen in the right box in the figure below. We use *interval censoring* for this analysis. That is, we specify a lower and upper bound for each response in two new columns, here called **Right Censored Y** and **Left Censored Y**, respectively. These values are the same as the exponentiated **Y** when it is above the LOD as seen in the first row. The right-censored value is missing and the left-censored value is the exponentiated LOD for censored responses as seen in the second row. Finally, we add a new indicator variable for censoring, here called **Below LOD**. It is set to 0 if **Y** is the actual response or 1 if it is censored. Here are the first five rows in this example

X1	X2	X3	X4	Y	Right Censored Y	Left Censored Y	Below LOD
-1	1	1	1	23.1863128	11740530594	11740530594	0
-1	1	-1	0	9.31947015	•	22026.465795	1
0	0	0	1	35.7669323	3.414926e+15	3.414926e+15	0
0	1	1	0	15.5181448	5488386.3693	5488386.3693	0
-1	-1	0	1	39.5386891	1.484002e+17	1.484002e+17	0

Examine the bias caused by using one of the ad hoc corrections before examining the results of using the correct analysis. The data were fit to a second order polynomial function with all two factor interaction terms using ordinary least squares regression. The parameter estimates in column **Fit Y** used the simulated response levels before

imposing the LOD. This regression represents the best we could do (unbiased estimates of model parameters) if there was no limit of detection. It is our benchmark for comparison with different ways of handling a LOD.

Then the OLS regression analysis is repeated with three different versions of the response **Y**.

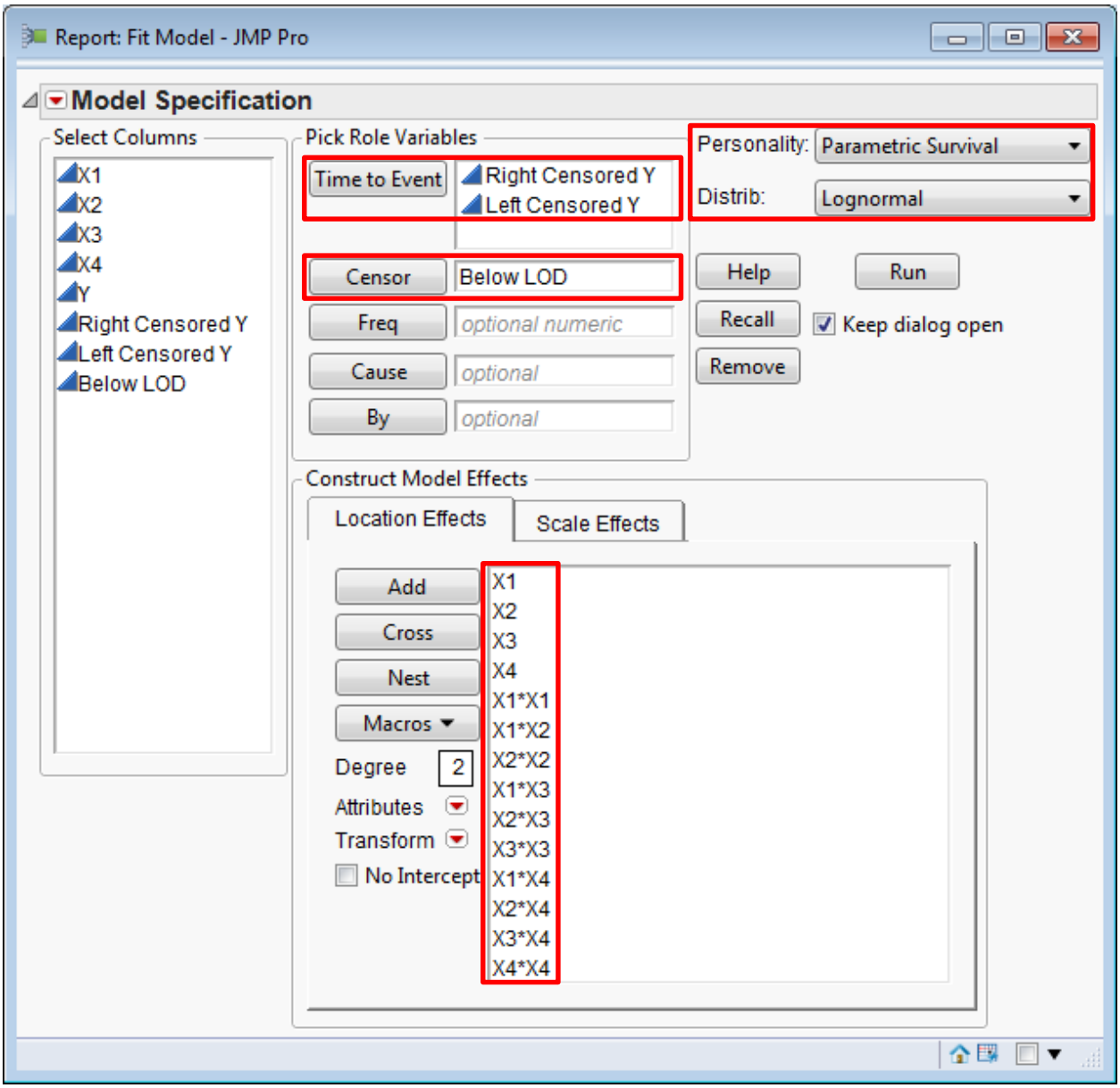
1. The parameter estimates in column Fit **Y2** used 0 for **Y** if the simulated response was below the LOD.
2. The parameter estimates in column Fit **Y3** used half way between 0 and LOD (7.5 in this case) for **Y** if the simulated response was below the LOD.
3. The parameter estimates in column Fit **Y4** used the LOD (15 in this case) for **Y** if the simulated response was below the LOD.

The prediction formulae from all four OLS regressions are saved to the data table. The estimates for the parameters that are active in the simulated response are compiled from each of these four regressions in the following table along with the true value from the simulation in the following table.

Parameter	True Value	Fit Y	Fit Y2	Fit Y3	Fit Y4
Intercept	30	29.4077	28.982	29.2763	29.57
X1	5	4.85517	5.60608	4.887	4.1679
X2	-8	-8.31617	-9.3121	-8.4699	-7.6277
X4	7	6.55928	7.14137	6.4476	5.7539
$X2^2$	-6	-6.84832	-7.9906	-7.013	-6.0356
X1*X4	-5	-5.46217	-4.7999	-4.7129	-4.626

Notice that the estimates are close to the true value in the absence of a LOD (**Y**). On the other hand, the application of one of the three *ad hoc* methods when the response is below the LOD results in estimates that are not as close to the true value.

Now try the parametric survival model. It is initiated in Fit Model by changing the fitting Personality from **Standard Least Squares** to **Parametric Survival**, selecting the **Lognormal** for the Distribution, using the **Right Censored Y** and **Left Censored Y** in the Time to Event role, and using **Below LOD** in the Censor role. The terms in this model (Location Effects) are the same as the terms used in the OLS regression above. There are no terms for the Scale Effects so this situation treats the variance as a constant, independent of factor levels.



The statistics and model estimates are in terms of the exponentiated response and the results are phrased or labeled in terms from survival or reliability analysis, but they can be translated as follows. The linear model determines the mean of the log-normal distribution and sigma estimates the standard deviation of the same distribution. The response is supposed to be the time to event, so the results are in terms of survival or failure time (*time quantile*), not the continuous response that we analyzed. That is not what we would call it, but it still works.

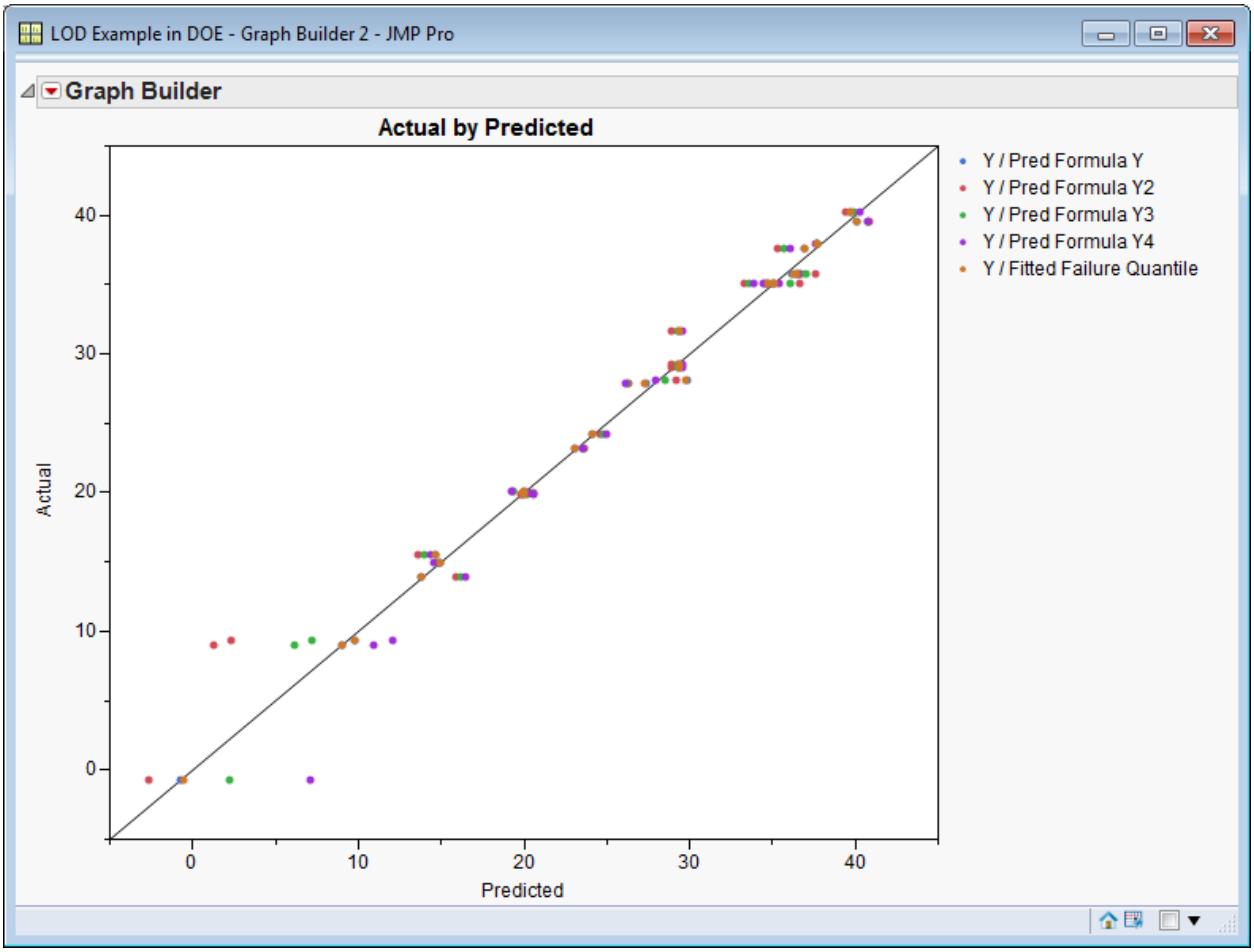
The second aspect that must be translated is the *survival probability* and *time quantile*. In survival analysis, you might ask a question such as, "What is the probability of survival after 100 days?" or "How many days before the probability of survival reaches 0.5?" For our original purpose, we are interested in the mean, so we will use a probability of 0.5. We can save the prediction formula for the time quantile, which is the response: click the red triangle at the top of the platform and select **Save Quantile Function** and enter **0.5** for the probability. A new column called **Fitted Failure Quantile** is added to the data table. You can rename this column, of course, to reflect the original response.

The last step is to edit the prediction formula to transform back to the original response. This step is easy. When you open the formula editor for the new column, the entire formula is already selected. Simply click on the **Transcendental** group of functions and select **Log**, then save the new formula.

Here are the first five rows of the original response **Y** and the columns creating by saving the prediction formula for the four OLS regressions and the parametric survival regression. Notice that the **Fitted Failure Quantile** prediction is much closer to that from **Pred Formula Y**, the first OLS regression or our benchmark.

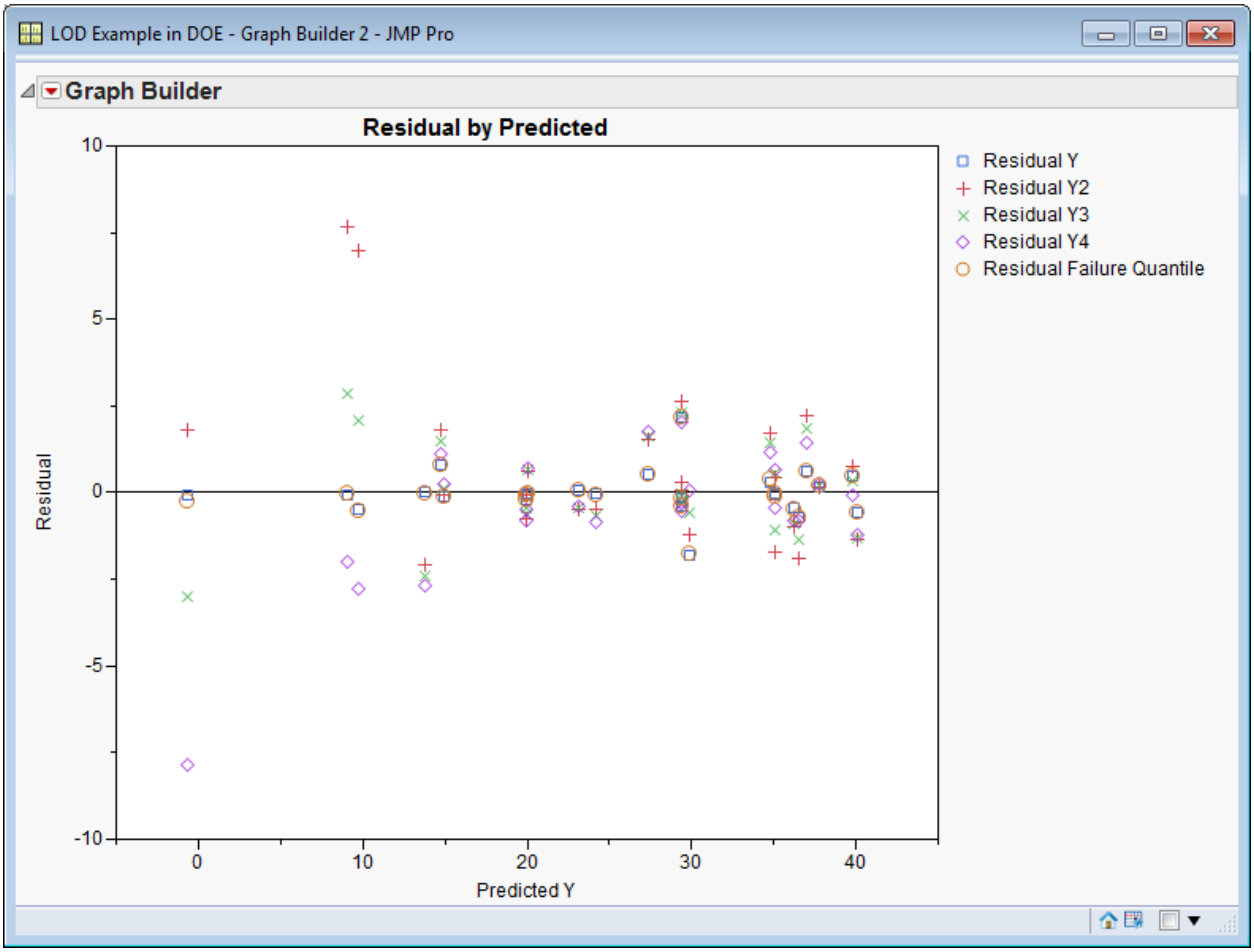
Y	Pred Formula Y	Pred Formula Y2	Pred Formula Y3	Pred Formula Y4	Fitted Failure Quantile
23.1863128	23.120913319	23.684590797	23.628355238	23.57211968	23.139741208
9.31947015	9.7803716958	2.3401707614	7.2157265386	12.091282316	9.8467086644
35.7669323	36.490068108	37.676755566	37.111051607	36.545347647	36.483442238
15.5181448	14.717950223	13.689135167	14.052597165	14.416059163	14.707552447
39.5386891	40.119295847	40.883040338	40.828308183	40.773576027	40.131839786

Examine the correspondence between the observed response and the predicted response from all of the models so far. Make a plot of these predicted responses over-laid in a scatter plot and add the identity line (Y=X) for reference.



The markers closest to the Y=X line are those from the first OLS regression (no LOD imposed) and the parametric survival regression. The discrepancy becomes worse as the response approaches the LOD.

Perhaps a better way to see the distinctions between these different approaches is with a residual plot.



The residuals closest to 0 are those from the first OLS regression (no LOD imposed) and the parametric survival regression. The large residuals from the models using ad hoc adjustments to the response (increased bias) indicate that these models are less valuable if you need predictions near the LOD, as in the case of optimizing factor levels to reduce an impurity.

CONCLUSION

In conclusion, we have demonstrated the use of inverse prediction from a prediction interval from an ordinary least squares regression as an analytical tool to estimate a limit of detection. This method is superior to more advanced methods that seek to stabilize the variance of a response over a grouping variable.

We have also demonstrated that ad hoc methods for using a response below the detection limit result in biased parameter estimates and model predictions. On the other hand, using interval censoring with a parametric survival model avoids these problems. The survival model is easy to fit, the transformations allowing the log normal distribution to model the errors is easy, and the prediction formula may be used with tools such as the Prediction Profiler to find optimal settings. Further note that the same approach could be used when the limiting response is an upper bound by substituting a missing value for the Left Censored Y value.

REFERENCES

Shrivastava A, Gupta VB. Methods for the determination of limit of detection and limit of quantitation of the analytical methods. Chron Young Sci [serial online] 2011 [cited 2013 Aug 16];2:21-5. Available from: http://www.cysonline.org/text.asp?2011/2/1/21/79345

EPA SW-846, *Test Methods for Evaluating Solid Waste, Physical/Chemical Methods*, chapter 3, “Inorganic Analytes,” retrieved from http://www.epa.gov/osw/hazard/testmethods/sw846/pdfs/chap3.pdf

USP <1225>, *Validation of Compendial Procedures*, retrieved from http://www.geinstruments.com/GetLibraryDoc.aspx?id=b9eeb287-49e7-47fd-a90c-89024009a146

SEMI C10-1109, *Guide for Determination of Method Detection Limits*, available (for purchase) from http://ams.semi.org/ebusiness/standards/SEMIStandardDetail.aspx?ProductID=211&DownloadID=1500

ISO:11843-7.2012, *Capability of Detection – Part 7: Methodology based on stochastic properties of instrumental noise*, available (for purchase) from http://www.iso.org/iso/catalogue_detail.htm?csnumber=53678

IUAPC, *Compendium of Analytical Nomenclature*, retrieved from http://iupac.org/publications/analytical_compendium/

SAS Data Analysis Examples. *Tobit Analysis*, UCLA Statistical Consulting Group, from http://www.ats.ucla.edu/stat/sas/dae/tobit.htm (accessed August 14, 2013).

Carry W. Croghan and Peter P. Egeghy*, Methods of Dealing with Values Below the Limit of Detection using SAS*, U.S. Environmental Protection Agency (no date given).

Assigning Values to Non-detected/Non-quantified Pesticide Residue in Human Health Food Exposure Assessments, Office of Pesticide Programs, U.S. Environmental Protection Agency, Washington, DC 20460, March 23, 2000

Dennis R. Helsel, "Fabricating data: How substituting values for nondetects can ruin results, and what can be done about it," *Chemosphere* 65 (2006) 2434-2439.

Dennis R. Helsel, *Nondetects and data analysis: statistics for censored environmental data*, Hoboken, NJ, John Wiley & Sons, 2005.

Wisconsin Department of Natural Resources Laboratory Certification Program, *Analytical Detection Limit Guidance & Laboratory Guide for Determining Method Detection Limits*, PUBL-TS-056-96, April, 1996.

Samuel Young Annan, Piaomu Liu, and Yuang Zhang, *Comparison of Kaplan-Meier, Maximum Likelihood, and ROS Estimators for Left-Censored Data Using Simulation Studies*, December 7, 2009.

Paul Hewett and Gary H. Ganser, "A Comparison of Several Methods for Analyzing Censored Data," *Annals of Occupational Hygiene*, 51(7)611-632, 2007.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Mark Bailey
SAS Education
919-531-9041
mark.bailey@sas.com

Di Michelson
SAS Education
919-531-9869
di.michelson@sas.com