

eLowQuant-Background, see Fishes: Signal Above the Noise ...

copyright Mary Lesperance

29/05/2023

Contents

| | | |
|----------|--|-----------|
| 1 | Outline: R code for determining limit of detection in the presence of background for low copy number eDNA samples | 2 |
| 1.1 | You will need: | 2 |
| 2 | Definitions | 2 |
| 3 | eLowQuant, Estimate standard curve using Binomial-Poisson model | 3 |
| 3.1 | eLowQuant Instructions | 3 |
| 3.2 | Read in Data | 4 |
| 3.3 | Process/Summarize samples by Target/SQ; Compute the Poisson estimates of SQ | 4 |
| 3.4 | Extract low starting copy number data | 5 |
| 3.5 | Plot the Poisson mean estimates (and CI) of copy number for SQ levels that had non-detects . . . | 6 |
| 3.6 | Estimate Binomial-Poisson models - no intercept model | 7 |
| 3.7 | Estimate predicted SQ given number detects and technical replicates - no intercept model (not shown in knitted document) | 11 |
| 3.8 | Estimate predicted SQ given consecutive numbers of detects from 0 to given number of technical replicates - no intercept model | 11 |
| 4 | Limits of detection/blank/quantification in the case of no background | 20 |
| 4.1 | Determine LOB, LOD, LOQ - no intercept model | 20 |
| 4.2 | Print out LOD, LOQ (no intercept model) with confidence intervals, by number of replicates . . . | 20 |
| 4.3 | Estimates, LOB, LOD, LOQ and confidence limits for a given number of technical reps NN[NNi] . | 21 |
| 5 | eDNA Manuscript Tables and Graphs (not used in <i>Fishes</i> manuscript) | 22 |
| 6 | Background estimation from negative field controls using no intercept model | 22 |
| 7 | Limits of detection/blank given the background estimate, no intercept model | 22 |
| 7.1 | Print out BACKGROUND limits of detection with confidence intervals, by number of replicates, no intercept model | 22 |

1 Outline: R code for determining limit of detection in the presence of background for low copy number eDNA samples

This R markdown file computes the limits of detection and blank in the presence of an ‘estimated’ background as demonstrated in the publication: “Establishing the signal above the noise: Accounting for an environmental background in the detection and quantification of salmonid environmental DNA”, *Fishes*, 2022 by Morgan D. Hocking, Jeffrey C. MacAdams, Michael J. Allison, Lauren C. Bergman, Robert Sneiderman, Ben F. Koop, Brian M. Starzomski, Mary L. Lesperance and Caren C. Helbing; 7: 266. DOI: 10.3390/fishes7050266.

The companion paper, Lesperance, M. L., Allison, Michael J., Bergman, Lauren C., Hocking, Morgan D., Helbing, Caren C. (2021). “A statistical model for calibration and computation of detection and quantification limits for low copy number environmental DNA samples.” *Environmental DNA* 3(5): 970-981, provides eLowQuant details.

1. First eLowQuant is used to estimate standard curve(s) relating the proportion of qPCR-detected technical replicates to copy number for serial dilution data for given assay(s).
2. The standard curve is used to estimate the copy number and the Limits of Blank and Detection in the case of no background.
3. The background is estimated using the investigator’s known field blanks and/or lab NTC detects. The eLowQuant model for the assay converts the proportion of qPCR-technical replicate detections of samples which are known to be negative field controls and/or lab NTCs to the estimated background copy number.
4. The estimated background copy number is then used in the computations of the Limits of Blank and Detection in the presence of background.

1.1 You will need:

- This .Rmd file in your workspace.
- The file **eLowQuant-Functions-V20210407.R** in your workspace.
- Your csv data file. See eLowQuant instructions below and the chunk labelled **READIN** below.
- Create a folder called **Outputs-BG** in your working directory.
- In the chunk **backzeroesBothNoInt**, enter vectors for the number of technical replicates (znum) and the number of detections (zdetect) for the field blanks and/or lab NTCs that you want to treat as Background. Your vectors must be of the same length as the number of unique Targets in your csv data file.

2 Definitions

- Target - character short name of the assay
- Lab - character name of the lab
- Cq - Cycle threshold (also known as Ct)
- SQ - Starting copy number (known) per reaction used in the dilution series
- lambda (λ) - mean copy number, parameter of the Poisson model
- phat - proportion detected, i.e. number detections divided by number technical replicates
- lm Rsq - linear model R-squared value used for descriptive/exploratory purposes to determine if a Binomial-Poisson model is appropriate
- LOB - Limit of Blank
- LOD - Limit of Detection
- LOB-B - Limit of Blank adjusted for estimated background
- LOD-B - Limit of Detection adjusted for estimated background
- LOQ - Limit of Quantification
- Field blank — a Procedural blank that is obtained during the field survey sample collection activities
- Procedural blank — a sample that does not contain the Matrix or the target Analyte of interest that is brought through the entire measurement procedure and analyzed in the same manner as a test sample

- No template control (NTC) — a Negative control in which DNA-free water or buffer is used to replace DNA Template or sample; it is used to detect reagent Contamination on a PCR plate

Reference: CSA Group eDNA standard CSA W219-2023 “Performance criteria for environmental DNA and environmental RNA analyses by targeted quantitative polymerase chain reaction” from which these definitions are taken.

3 eLowQuant, Estimate standard curve using Binomial-Poisson model

3.1 eLowQuant Instructions

When running the code in RStudio ‘live’, the code chunks must be run in the order that they appear.

This section performs the computations in the publication, ‘A Statistical Model for Calibration and Computation of Detection and Quantification Limits for Low Copy Number Environmental DNA samples’ by Lesperance, Allison, Bergman, Hocking, Helbing, *Environmental DNA*, 2021, 3, 970-981 <https://doi.org/10.1002/edn3.220>.

1. Create a folder called **Outputs-BG** in your working directory.
2. Put the file **eLowQuant-Functions-V20210407.R** in your working directory.
3. Put your data file in your working directory and put the name of the file in the chunk labeled ‘READIN’ below.
4. Our sample file in the chunk labeled ‘READIN’ is called “EXAMPLE_eLowQuant.csv”.
5. Data set csv file requirements - include columns named: Target, Lab, Cq, SQ
6. For nondetects, set Cq (Cycle threshold, or Ct) to be empty or NA value.
7. For negative plate controls (no template controls or NTCs), set SQ to be empty or 0 or NA value
8. **Include negative controls in the csv file!**
9. DO NOT DUPLICATE Target names over different Labs!! Target names need to be UNIQUE between labs.
10. This code uses observations with dilution series SQs having less than 100 percent detections.
11. Only the dilution series SQ’s up to the first one with 100 percent detections are used.
12. This code allows for variable numbers of technical replications of SQ levels per Target.
13. It assumes SQs == NA are zero, i.e. are negative controls.
14. The code uses the R function optim. A convergence code 0 indicates successful completion.
15. Ignore warnings if results are sensible.
16. EXECUTE EACH CHUNK LOOKING AT THE OUTPUT. IN PARTICULAR, LOOK AT THE GRAPHS IN THE CHUNK CALLED ‘PlotPois’ TO DETERMINE IF THE MODEL IS APPROPRIATE. IT IS NOT APPROPRIATE IF ‘lm Rsq’ (the linear model R-squared) is small, i.e. near zero!
17. You can send results to files by setting the sink.indicator and Manusink variables to TRUE in the chunk labelled ‘setup’ and running the code in RStudio. This currently does not work for all results files when knitting.
18. If you wish to knit to pdf AND you do NOT have a version of Latex installed on your computer, then run the following in your RStudio console:
install.packages(“tinytex”); tinytex::install_tinytex()

3.2 Read in Data

The data should be in the form of a csv file with column names: Target, Lab, Cq, SQ.

Target is the character name of the assay,

Lab is the character lab name,

Cq is the numeric cycle threshold where a *missing* Cq is interpreted as a reaction *nondetect*,

SQ is the starting quantity per reaction where missing values are assumed to be zero.

It is important to include the negative controls (i.e. SQ=0)!

3.3 Process/Summarize samples by Target/SQ; Compute the Poisson estimates of SQ

The data are summarized by Lab, Target and SQ, where n is the number of technical replicates, *detect* is the number of detections out of n technical replicates, *phat* is the proportion detections, $(detect/n)$ and *lamhat* is the estimated Poisson mean ($\hat{\lambda}$) which is a function of *phat*. *vlamhat*, *sdlamhat*, *MElamhat* are the estimated variance, standard deviation and margin of error for *lamhat* computed using the delta method. *CIexphat.lower*, *CIexphat.upper* are the lower and upper confidence limits of *phat* computed using an exact method and *Lamhatex.Lower*, *Lamhatex.Upper* are the lower and upper confidence limits of *lamhat* computed using an exact method (Julious 2005, Stat in Medicine).

The literature background for the Poisson part of the model is included here. Hindson et al. 2011 [“High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number”, Anal. Chem.83 (22), pp 8604–8610] use a Poisson approximation for quantification. Before that, Dube et al. 2008, [“Mathematical analysis of copy number variation in a DNA sample using digital PCR on a nanofluidic device”, PloS One, Vol 3, Issue 8, e2876] model the number of molecules in each chamber as a Poisson process, giving the relationship between p and λ .

```
## Dataframe size
## [1] 42 17
##
##
## Table: LabA
##
## Target      SQ    detect    n    phat    lamhat
## -----
## TargetA      0.000         0   24   0.000    0.000
## TargetA      0.244         3   24   0.125    0.134
## TargetA      0.488         4   24   0.167    0.182
## TargetA      0.975         9   24   0.375    0.470
## TargetA      1.950         9   24   0.375    0.470
## TargetA      3.906        22   24   0.917    2.485
## TargetA      7.812        22   24   0.917    2.485
## TargetA     15.625        24   24   1.000     Inf
## TargetA     31.250        24   24   1.000     Inf
## TargetA     62.500        24   24   1.000     Inf
## TargetA    125.000        24   24   1.000     Inf
## TargetA    250.000        24   24   1.000     Inf
## TargetA    500.000        24   24   1.000     Inf
##
##
## Table: LabB
##
##      Target      SQ    detect    n    phat    lamhat
## ---  -----
## 14   TargetB      0.00e+00         0   24   0.000    0.000
```

```
## 15 TargetB 3.20e-02 0 24 0.000 0.000
## 16 TargetB 1.60e-01 4 24 0.167 0.182
## 17 TargetB 8.00e-01 9 24 0.375 0.470
## 18 TargetB 4.00e+00 19 24 0.792 1.569
## 19 TargetB 2.00e+01 24 24 1.000 Inf
## 20 TargetB 1.00e+02 24 24 1.000 Inf
## 21 TargetB 5.00e+02 8 8 1.000 Inf
## 22 TargetB 2.50e+03 8 8 1.000 Inf
## 23 TargetB 1.25e+04 8 8 1.000 Inf
## 24 TargetB 6.25e+04 8 8 1.000 Inf
```

```
##
##
## Table: LabC
```

```
##
##      Target      SQ  detect    n    phat    lamhat
## ---  -
## 25 TargetC 0.000      0    24 0.000 0.000
## 26 TargetC 0.032      0    24 0.000 0.000
## 27 TargetC 0.160      6    24 0.250 0.288
## 28 TargetC 0.800     11    24 0.458 0.613
## 29 TargetC 4.000     22    24 0.917 2.485
## 30 TargetC 20.000    24    24 1.000 Inf
## 31 TargetC 100.000   24    24 1.000 Inf
```

```
##
##
## Table: LabD
```

```
##
##      Target      SQ  detect    n    phat    lamhat
## ---  -
## 32 TargetD 0.00e+00 0    24 0.000 0.000
## 33 TargetD 3.20e-02 3    24 0.125 0.134
## 34 TargetD 1.60e-01 6    24 0.250 0.288
## 35 TargetD 8.00e-01 14   24 0.583 0.875
## 36 TargetD 4.00e+00 21   24 0.875 2.079
## 37 TargetD 2.00e+01 24   24 1.000 Inf
## 38 TargetD 1.00e+02 24   24 1.000 Inf
## 39 TargetD 5.00e+02 8    8 1.000 Inf
## 40 TargetD 2.50e+03 8    8 1.000 Inf
## 41 TargetD 1.25e+04 8    8 1.000 Inf
## 42 TargetD 6.25e+04 8    8 1.000 Inf
```

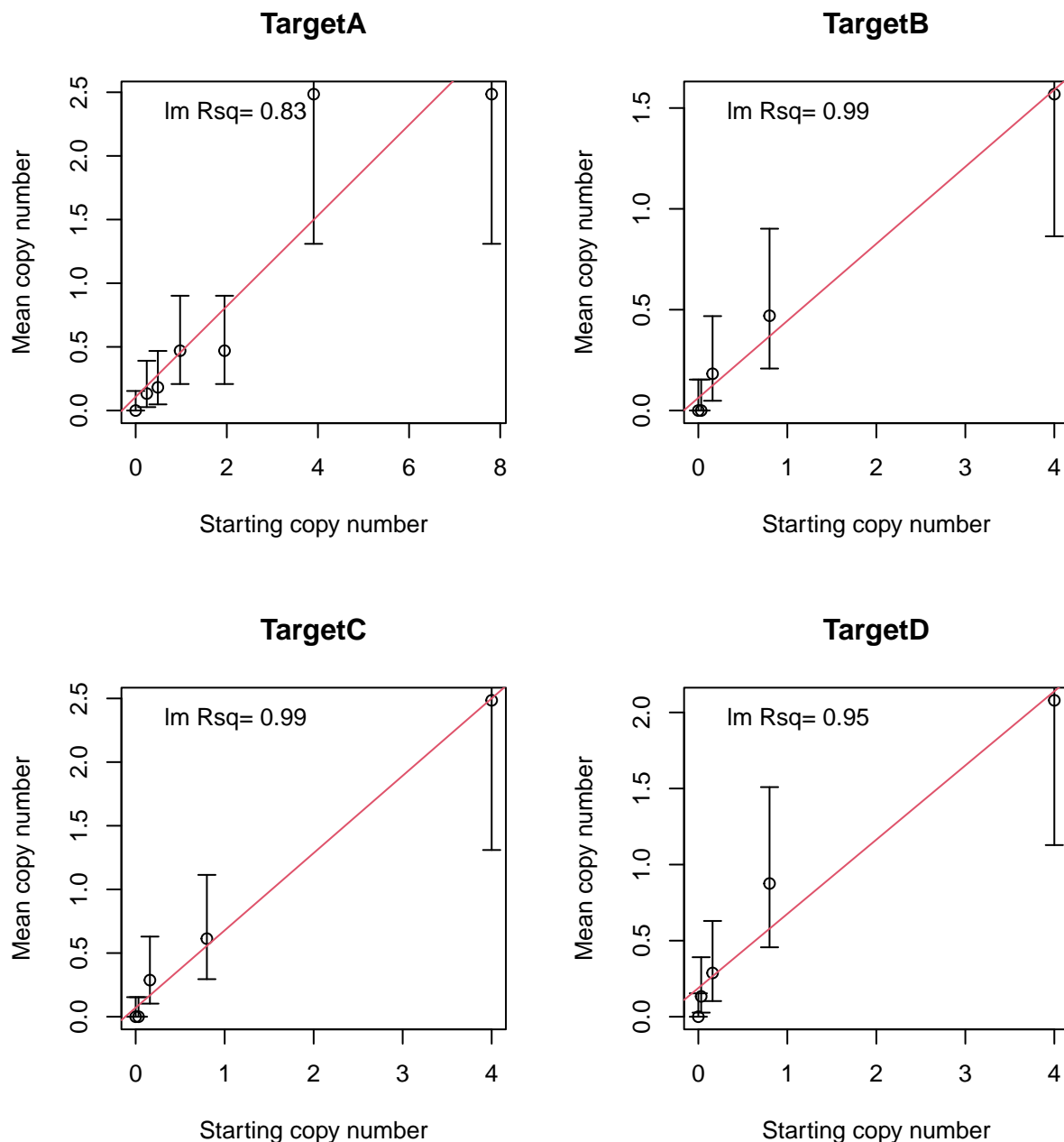
3.4 Extract low starting copy number data

Extract SQ samples for which there are non-detects. Only the SQ samples before the first one in SQ order with 100 percent detects are extracted.

3.5 Plot the Poisson mean estimates (and CI) of copy number for SQ levels that had non-detects

This section contains plots of the Poisson mean estimates of copy number for the first levels of SQ that had non-detections. The red line is least squares linear regression line and is used for exploratory purposes only. The Binomial-Poisson model should provide an acceptable fit if the least squares line displays an acceptable fit. The R-squared value from the least squares fit is labelled *lm-Rsq* which can be used to judge fit. Note that the least squares line does not accommodate the nonconstant variance that is typical of this data and hence is used for exploratory purposes only.

LOOK AT THE GRAPHS BELOW (created by the chunk called 'PlotPois') TO DETERMINE IF THE MODEL IS APPROPRIATE. IT IS NOT APPROPRIATE IF 'lm Rsq', the linear model R-square, is small, i.e. near zero!

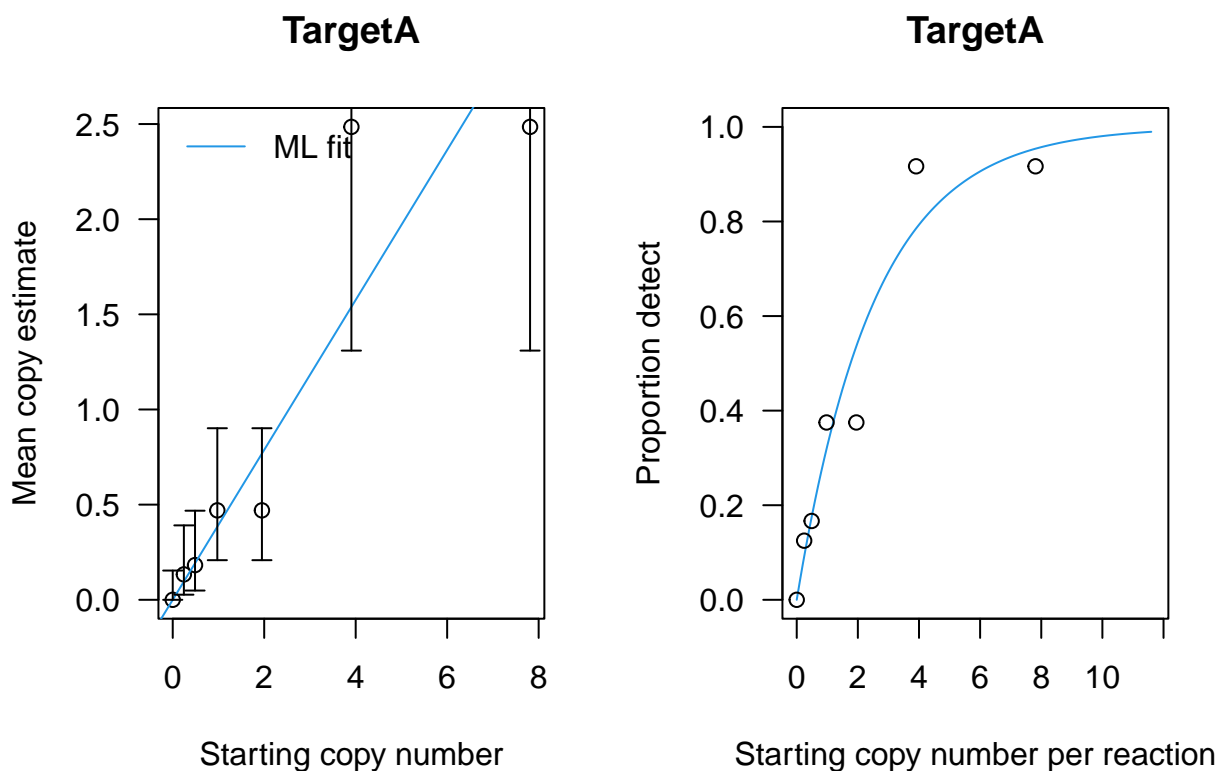


3.6 Estimate Binomial-Poisson models - no intercept model

The no intercept Binomial-Poisson model is fit to the data for which the first levels of SQ had non-detects. The fitted model is graphed on two scales, the Poisson mean copy number scale and the proportion detect scale. The model output contains the beta parameter estimate (β), its standard error and a p-value for the test that beta=0. The LLR test stat is the Likelihood Ratio goodness-of-fit statistic, comparing the Binomial-Poisson model with unconstrained Binomial models. A large p-value (e.g. > 0.05) for this test suggests no evidence against the fit of the Binomial-Poisson model.

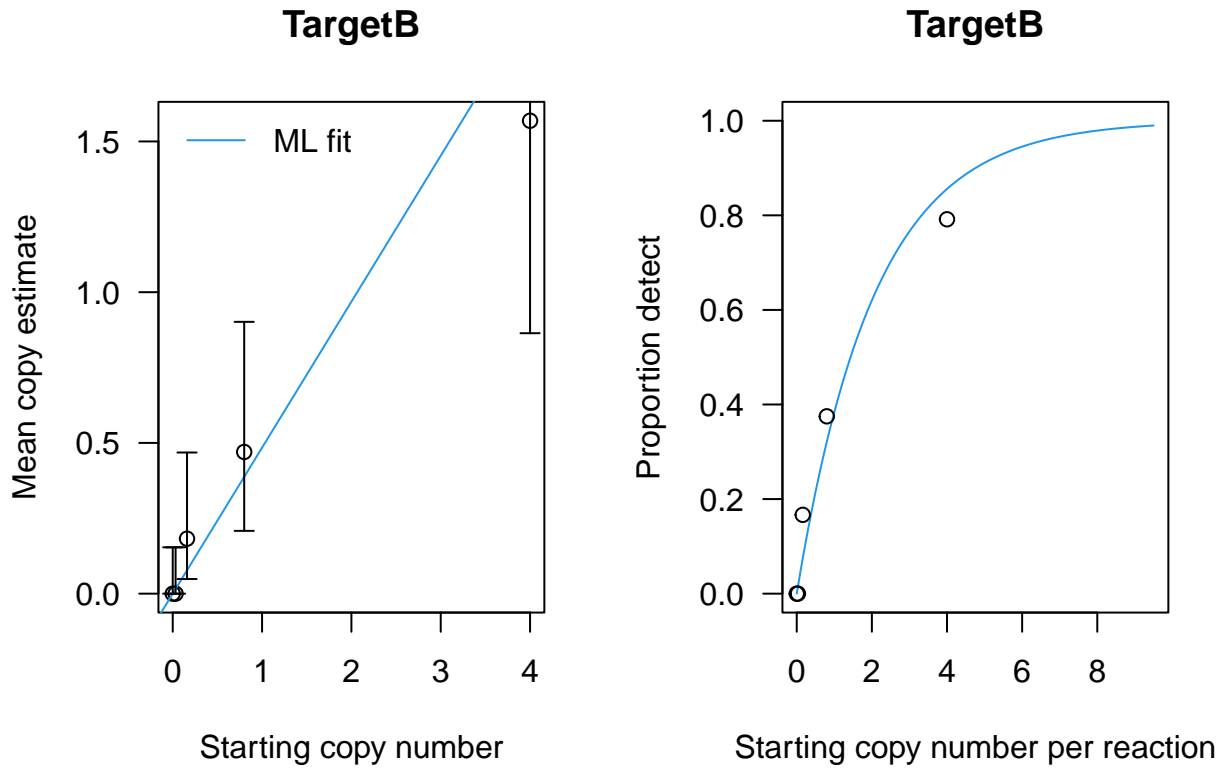
The results of the model fits are saved in RDS files, *assay_name.rds*, one for each target. Use readRDS to read in the results. The results can be used for Background estimates and LOD-B estimates, etc. The following quantities are saved:

- * Target.dat - the data used in the model
- * Calib.fit - output from optim fit
- * Calib.tbl - table of estimates, se, z, pvalue
- * Calib.fitted - fitted values on Poisson scale, i.e. beta times SQ
- * Calib.LLR - log likelihood ratio statistic
- * Calib.LLR.pv - pvalue for LLR
- * Calib.degf - degrees of freedom for LLR



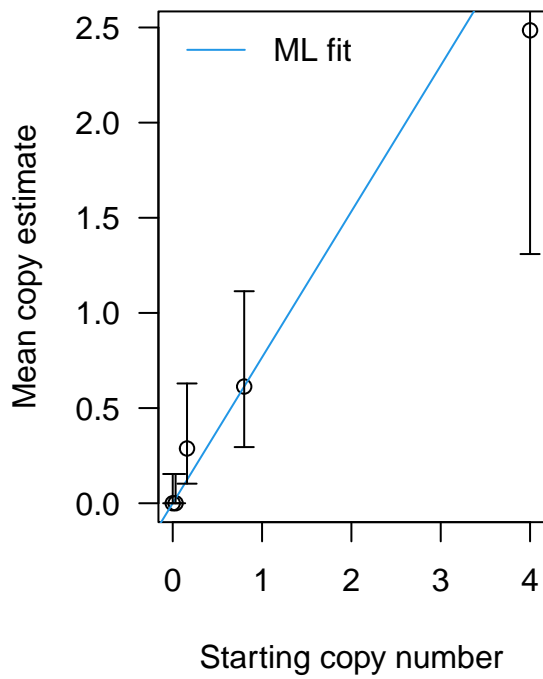
```
##
##
##
##
## TargetA
## Convergence= 0
##      Estimate Std.Err Z value Pr(>z)
## beta    0.394   0.054    7.29 3.1e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## LLR test stat= 6.787625 , df= 6 , p-value= 0.340935
```

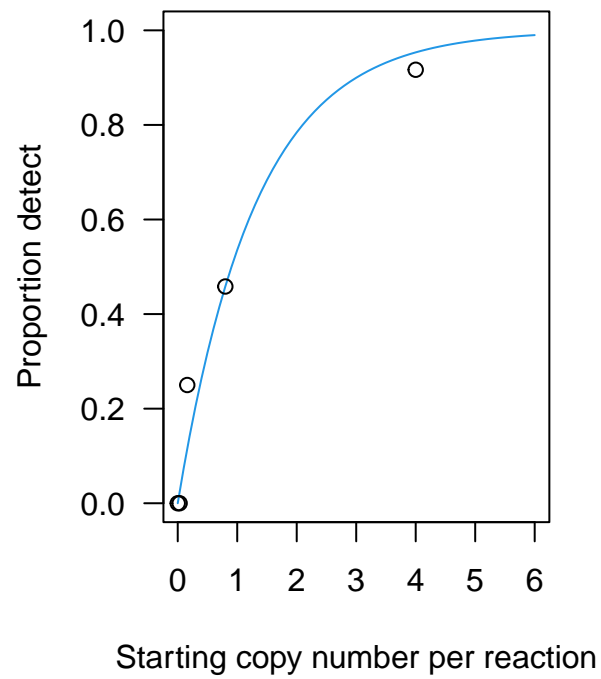


```
##
##
##
##
## TargetB
## Convergence= 0
##      Estimate Std.Err Z value Pr(>z)
## beta   0.4842  0.0933   5.19 2.1e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## LLR test stat= 4.017301 , df= 4 , p-value= 0.4036694
```


TargetC

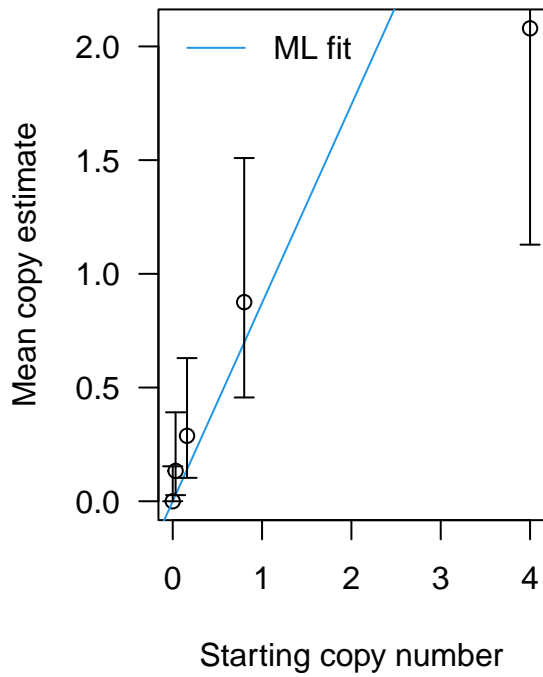


TargetC

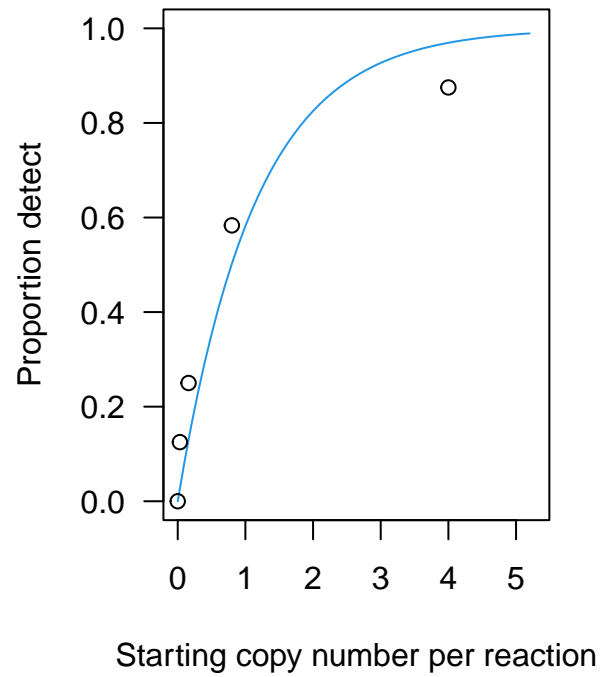


```
##
##
##
##
## TargetC
## Convergence= 0
##      Estimate Std.Err Z value Pr(>z)
## beta    0.767   0.147   5.22 1.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## LLR test stat= 5.106559 , df= 4 , p-value= 0.2765376
```

TargetD



TargetD



```
##
##
##
##
## TargetD
## Convergence= 0
##      Estimate Std.Err Z value Pr(>z)
## beta    0.872   0.157   5.54 2.9e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## LLR test stat= 11.91839 , df= 4 , p-value= 0.01796854
```

3.7 Estimate predicted SQ given number detects and technical replicates - no intercept model (not shown in knitted document)

The estimated SQ (SQ0) is easily obtained for given new *user-specified* values of nn0=number of replicates, nd0=number detected, and the estimated slope β as:

$$SQ_0 = \frac{-\log((nn0 - nd0)/nn0)}{\beta}$$

3.8 Estimate predicted SQ given consecutive numbers of detects from 0 to given number of technical replicates - no intercept model

The estimated SQ (SQ0) is easily obtained for given new values of nn0=number of replicates, nd0=number detected, and the estimated slope β as:

$$SQ_0 = \frac{-\log((nn0 - nd0)/nn0)}{\beta}$$

Estimates of SQ0 are given for the sequence of number detected, from 0 to the number of technical replicates in the vector nn0vec0. SE_SQ0 is the standard error of the estimated SQ0.

```
##
## TargetA
## ML estimate of SQ for numbers of detects and 8 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ---  -
## 0          0      0.0000      NaN
## 1          1      0.3390      0.3424
## 2          2      0.7303      0.5278
## 3          3      1.1931      0.7142
## 4          4      1.7596      0.9294
## 5          5      2.4913      1.2087
## 6          6      3.5120      1.6237
## 7          7      5.2789      2.4826
##
## TargetA
## ML estimate of SQ for numbers of detects and 16 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ---  -
## 0          0      0.0000      NaN
## 1          1      0.1639      0.1655
## 2          2      0.3390      0.2443
## 3          3      0.5271      0.3133
## 4          4      0.7303      0.3799
## 5          5      0.9512      0.4473
## 6          6      1.1931      0.5181
## 7          7      1.4606      0.5945
## 8          8      1.7596      0.6790
## 9          9      2.0993      0.7754
```

```

## 10      10    2.4900    0.8877
## 11      11    2.9527    1.0248
## 12      12    3.5191    1.2006
## 13      13    4.2495    1.4441
## 14      14    5.2789    1.8287
## 15      15    7.0387    2.6410
##
## TargetA
## ML estimate of SQ for numbers of detects and 24 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ---- -
## 0           0    0.0000      NaN
## 1           1    0.1080    0.1090
## 2           2    0.2209    0.1592
## 3           3    0.3390    0.2013
## 4           4    0.4628    0.2403
## 5           5    0.5928    0.2779
## 6           6    0.7303    0.3155
## 7           7    0.8754    0.3535
## 8           8    1.0293    0.3927
## 9           9    1.1931    0.4335
## 10          10    1.3683    0.4765
## 11          11    1.5564    0.5223
## 12          12    1.7596    0.5717
## 13          13    1.9805    0.6254
## 14          14    2.2225    0.6848
## 15          15    2.4899    0.7512
## 16          16    2.7889    0.8267
## 17          17    3.1279    0.9145
## 18          18    3.5180    1.0187
## 19          19    3.9821    1.1484
## 20          20    4.5485    1.3161
## 21          21    5.2788    1.5505
## 22          22    6.3082    1.9243
## 23          23    8.0695    2.7215
##
## TargetA
## ML estimate of SQ for numbers of detects and 32 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ---- -
## 0           0    0.0000      NaN
## 1           1    0.0807    0.0815
## 2           2    0.1638    0.1180
## 3           3    0.2499    0.1484
## 4           4    0.3390    0.1759
## 5           5    0.4313    0.2020
## 6           6    0.5271    0.2274
## 7           7    0.6267    0.2525
## 8           8    0.7303    0.2778

```

```

## 9      9      0.8383      0.3034
## 10     10     0.9512      0.3295
## 11     11     1.0693      0.3564
## 12     12     1.1931      0.3842
## 13     13     1.3234      0.4133
## 14     14     1.4606      0.4436
## 15     15     1.6057      0.4756
## 16     16     1.7596      0.5096
## 17     17     1.9235      0.5458
## 18     18     2.0986      0.5847
## 19     19     2.2867      0.6267
## 20     20     2.4899      0.6726
## 21     21     2.7108      0.7230
## 22     22     2.9528      0.7792
## 23     23     3.2202      0.8425
## 24     24     3.5192      0.9150
## 25     25     3.8582      0.9997
## 26     26     4.2496      1.1012
## 27     27     4.7124      1.2270
## 28     28     5.2789      1.3908
## 29     29     6.0092      1.6206
## 30     30     7.0383      1.9882
## 31     31     8.7980      2.7748

```

```
##
```

```
## TargetB
```

```
## ML estimate of SQ for numbers of detects and 8 replicates
```

```
##
```

```
##
```

| ## | NumDetects | SQ0 | SE_SQ0 |
|------|------------|--------|--------|
| ## 0 | 0 | 0.0000 | NaN |
| ## 1 | 1 | 0.2758 | 0.2811 |
| ## 2 | 2 | 0.5942 | 0.4369 |
| ## 3 | 3 | 0.9712 | 0.5961 |
| ## 4 | 4 | 1.4316 | 0.7806 |
| ## 5 | 5 | 2.0257 | 1.0204 |
| ## 6 | 6 | 2.8631 | 1.3799 |
| ## 7 | 7 | 4.2947 | 2.1018 |

```
##
```

```
## TargetB
```

```
## ML estimate of SQ for numbers of detects and 16 replicates
```

```
##
```

```
##
```

| ## | NumDetects | SQ0 | SE_SQ0 |
|------|------------|--------|--------|
| ## 0 | 0 | 0.0000 | NaN |
| ## 1 | 1 | 0.1333 | 0.1358 |
| ## 2 | 2 | 0.2758 | 0.2023 |
| ## 3 | 3 | 0.4288 | 0.2614 |
| ## 4 | 4 | 0.5942 | 0.3193 |
| ## 5 | 5 | 0.7739 | 0.3787 |
| ## 6 | 6 | 0.9707 | 0.4416 |
| ## 7 | 7 | 1.1883 | 0.5097 |

```

## 8      8      1.4316      0.5854
## 9      9      1.7074      0.6716
## 10     10      2.0257      0.7725
## 11     11      2.4023      0.8949
## 12     12      2.8631      1.0509
## 13     13      3.4573      1.2647
## 14     14      4.2947      1.5973
## 15     15      5.7253      2.2835
##
## TargetB
## ML estimate of SQ for numbers of detects and 24 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ---- -
## 0      0      0.0000      NaN
## 1      1      0.0879      0.0895
## 2      2      0.1797      0.1317
## 3      3      0.2758      0.1680
## 4      4      0.3765      0.2020
## 5      5      0.4825      0.2354
## 6      6      0.5942      0.2690
## 7      7      0.7122      0.3034
## 8      8      0.8372      0.3389
## 9      9      0.9707      0.3764
## 10     10      1.1132      0.4159
## 11     11      1.2663      0.4583
## 12     12      1.4317      0.5039
## 13     13      1.6113      0.5536
## 14     14      1.8081      0.6085
## 15     15      2.0257      0.6698
## 16     16      2.2690      0.7394
## 17     17      2.5448      0.8199
## 18     18      2.8631      0.9153
## 19     19      3.2397      1.0321
## 20     20      3.7006      1.1821
## 21     21      4.2947      1.3890
## 22     22      5.1321      1.7128
## 23     23      6.5623      2.3843
##
## TargetB
## ML estimate of SQ for numbers of detects and 32 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ---- -
## 0      0      0.0000      NaN
## 1      1      0.0656      0.0668
## 2      2      0.1333      0.0977
## 3      3      0.2033      0.1238
## 4      4      0.2758      0.1479
## 5      5      0.3509      0.1711
## 6      6      0.4288      0.1939

```

```

## 7      7      0.5098      0.2168
## 8      8      0.5942      0.2399
## 9      9      0.6821      0.2635
## 10     10      0.7739      0.2878
## 11     11      0.8699      0.3130
## 12     12      0.9707      0.3391
## 13     13      1.0766      0.3664
## 14     14      1.1883      0.3952
## 15     15      1.3064      0.4255
## 16     16      1.4318      0.4578
## 17     17      1.5649      0.4920
## 18     18      1.7074      0.5289
## 19     19      1.8604      0.5687
## 20     20      2.0257      0.6121
## 21     21      2.2054      0.6597
## 22     22      2.4023      0.7125
## 23     23      2.6199      0.7718
## 24     24      2.8631      0.8393
## 25     25      3.1389      0.9177
## 26     26      3.4573      1.0108
## 27     27      3.8338      1.1251
## 28     28      4.2947      1.2722
## 29     29      4.8889      1.4753
## 30     30      5.7262      1.7938
## 31     31      7.1566      2.4561
##
## TargetC
## ML estimate of SQ for numbers of detects and 8 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ---- -
## 0      0      0.0000      NaN
## 1      1      0.1742      0.1775
## 2      2      0.3753      0.2758
## 3      3      0.6130      0.3760
## 4      4      0.9041      0.4926
## 5      5      1.2793      0.6438
## 6      6      1.8082      0.8706
## 7      7      2.7142      1.3274
##
## TargetC
## ML estimate of SQ for numbers of detects and 16 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ---- -
## 0      0      0.0000      NaN
## 1      1      0.0842      0.0857
## 2      2      0.1742      0.1277
## 3      3      0.2708      0.1650
## 4      4      0.3752      0.2015
## 5      5      0.4887      0.2389

```

```

## 6          6    0.6132    0.2786
## 7          7    0.7504    0.3215
## 8          8    0.9041    0.3692
## 9          9    1.0782    0.4235
## 10         10    1.2793    0.4871
## 11         11    1.5171    0.5642
## 12         12    1.8081    0.6625
## 13         13    2.1833    0.7973
## 14         14    2.7122    1.0071
## 15         15    3.6162    1.4404
##
## TargetC
## ML estimate of SQ for numbers of detects and 24 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ---- -
## 0          0    0.0000      NaN
## 1          1    0.0555    0.0565
## 2          2    0.1135    0.0832
## 3          3    0.1741    0.1060
## 4          4    0.2378    0.1275
## 5          5    0.3047    0.1485
## 6          6    0.3752    0.1697
## 7          7    0.4500    0.1915
## 8          8    0.5290    0.2139
## 9          9    0.6130    0.2373
## 10         10    0.7029    0.2622
## 11         11    0.7997    0.2889
## 12         12    0.9041    0.3176
## 13         13    1.0176    0.3489
## 14         14    1.1419    0.3835
## 15         15    1.2793    0.4221
## 16         16    1.4329    0.4659
## 17         17    1.6071    0.5166
## 18         18    1.8081    0.5767
## 19         19    2.0459    0.6503
## 20         20    2.3370    0.7448
## 21         21    2.7121    0.8752
## 22         22    3.2410    1.0794
## 23         23    4.1452    1.5037
##
## TargetC
## ML estimate of SQ for numbers of detects and 32 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ---- -
## 0          0    0.0000      NaN
## 1          1    0.0414    0.0422
## 2          2    0.0842    0.0617
## 3          3    0.1284    0.0781
## 4          4    0.1742    0.0933

```



```

## 5      5      0.2216      0.1079
## 6      6      0.2709      0.1223
## 7      7      0.3220      0.1367
## 8      8      0.3752      0.1513
## 9      9      0.4307      0.1662
## 10     10     0.4887      0.1815
## 11     11     0.5494      0.1973
## 12     12     0.6130      0.2137
## 13     13     0.6799      0.2309
## 14     14     0.7504      0.2490
## 15     15     0.8250      0.2681
## 16     16     0.9041      0.2884
## 17     17     0.9882      0.3100
## 18     18     1.0776      0.3329
## 19     19     1.1749      0.3583
## 20     20     1.2793      0.3856
## 21     21     1.3928      0.4155
## 22     22     1.5171      0.4488
## 23     23     1.6545      0.4861
## 24     24     1.8081      0.5286
## 25     25     1.9823      0.5780
## 26     26     2.1833      0.6366
## 27     27     2.4211      0.7086
## 28     28     2.7122      0.8013
## 29     29     3.0874      0.9293
## 30     30     3.6163      1.1302
## 31     31     4.5197      1.5480
##
## TargetD
## ML estimate of SQ for numbers of detects and 8 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ---- -
## 0      0      0.0000      NaN
## 1      1      0.1531      0.1557
## 2      2      0.3299      0.2415
## 3      3      0.5390      0.3287
## 4      4      0.7949      0.4300
## 5      5      1.1248      0.5613
## 6      6      1.5897      0.7585
## 7      7      2.3848      1.1558
##
## TargetD
## ML estimate of SQ for numbers of detects and 16 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ---- -
## 0      0      0.0000      NaN
## 1      1      0.0740      0.0752
## 2      2      0.1531      0.1118
## 3      3      0.2380      0.1442

```

```

## 4      4      0.3299      0.1759
## 5      5      0.4296      0.2082
## 6      6      0.5390      0.2424
## 7      7      0.6598      0.2794
## 8      8      0.7949      0.3205
## 9      9      0.9480      0.3673
## 10     10     1.1248      0.4221
## 11     11     1.3338      0.4886
## 12     12     1.5897      0.5734
## 13     13     1.9196      0.6899
## 14     14     2.3848      0.8720
## 15     15     3.1786      1.2491
##
## TargetD
## ML estimate of SQ for numbers of detects and 24 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ----
## 0      0      0.0000      NaN
## 1      1      0.0488      0.0496
## 2      2      0.0998      0.0728
## 3      3      0.1530      0.0926
## 4      4      0.2091      0.1113
## 5      5      0.2679      0.1294
## 6      6      0.3299      0.1477
## 7      7      0.3954      0.1663
## 8      8      0.4650      0.1855
## 9      9      0.5390      0.2057
## 10     10     0.6181      0.2271
## 11     11     0.7031      0.2499
## 12     12     0.7949      0.2745
## 13     13     0.8946      0.3013
## 14     14     1.0039      0.3309
## 15     15     1.1248      0.3640
## 16     16     1.2598      0.4015
## 17     17     1.4129      0.4450
## 18     18     1.5897      0.4966
## 19     19     1.7988      0.5599
## 20     20     2.0547      0.6413
## 21     21     2.3846      0.7540
## 22     22     2.8494      0.9310
## 23     23     3.6444      1.3008
##
## TargetD
## ML estimate of SQ for numbers of detects and 32 replicates
##
##
##      NumDetects      SQ0      SE_SQ0
## ----
## 0      0      0.0000      NaN
## 1      1      0.0364      0.0370
## 2      2      0.0740      0.0540

```

| | | | |
|-------|----|--------|--------|
| ## 3 | 3 | 0.1128 | 0.0683 |
| ## 4 | 4 | 0.1531 | 0.0814 |
| ## 5 | 5 | 0.1948 | 0.0940 |
| ## 6 | 6 | 0.2381 | 0.1064 |
| ## 7 | 7 | 0.2831 | 0.1188 |
| ## 8 | 8 | 0.3299 | 0.1313 |
| ## 9 | 9 | 0.3787 | 0.1440 |
| ## 10 | 10 | 0.4296 | 0.1571 |
| ## 11 | 11 | 0.4830 | 0.1706 |
| ## 12 | 12 | 0.5390 | 0.1847 |
| ## 13 | 13 | 0.5978 | 0.1994 |
| ## 14 | 14 | 0.6598 | 0.2148 |
| ## 15 | 15 | 0.7253 | 0.2310 |
| ## 16 | 16 | 0.7949 | 0.2483 |
| ## 17 | 17 | 0.8689 | 0.2667 |
| ## 18 | 18 | 0.9480 | 0.2865 |
| ## 19 | 19 | 1.0330 | 0.3078 |
| ## 20 | 20 | 1.1248 | 0.3311 |
| ## 21 | 21 | 1.2245 | 0.3567 |
| ## 22 | 22 | 1.3338 | 0.3851 |
| ## 23 | 23 | 1.4546 | 0.4169 |
| ## 24 | 24 | 1.5897 | 0.4533 |
| ## 25 | 25 | 1.7428 | 0.4955 |
| ## 26 | 26 | 1.9196 | 0.5458 |
| ## 27 | 27 | 2.1287 | 0.6077 |
| ## 28 | 28 | 2.3842 | 0.6873 |
| ## 29 | 29 | 2.7140 | 0.7978 |
| ## 30 | 30 | 3.1794 | 0.9722 |
| ## 31 | 31 | 3.9727 | 1.3360 |

4 Limits of detection/blank/quantification in the case of no background

4.1 Determine LOB, LOD, LOQ - no intercept model

This section follows Lavagnini and Magno 2007, Mass Spectrometry Reviews, which uses a hypothesis testing approach to develop Limits of Blank and Detection. See Lesperance et al. (2021) for further details.

The notation in our eDNA paper is different from the Lavagnini 2007 paper and is shown in brackets in the definitions below. I also include the variable names used in the R code in this description.

The following definitions are extracted from Lavagnini and Magno 2007.

- Lc (*LOB Limit of blank*) = critical level is the assay signal above which a response is reliably attributed to the presence of analyte.
- Ld (*Ld = expected number detects out of NN replicates at concentration LOD*) = signal corresponding to an analyte concentration x_d (=LOD *Limit of Detection*) level which may be a priori expected to be recognized.
- Lq (=LOQ = *Limit of Quantification*) = is a signal with a precision which satisfies an expected value (= γ_Q).

We interpret the above in the context of eDNA studies where S =concentration/copy number and NN is the number of technical replicates at concentration S . Y is the number of detections out of NN replicates. *LOB* corresponds to the critical number of detections out of NN replicates, above which we would reject the null hypothesis that S is zero with α level (False Positive Rate) $\alpha = \alpha_{Lc}$ (= γ_{FP}). Essentially, the test is positive if $Y > LOB$ and negative if $Y \leq LOB$. The probability that $Y > LOB$ if the true $S=0$ (true concentration is 0) is at most the False Positive Rate, i.e. $P(Y > LOB \text{ given that } S=0) \leq \alpha_{Lc}$ (= γ_{FP}). If $S=0$ truly, then we expect in repeated sampling that at most $100 * \gamma_{FP}$ percent samples to test positive.

The *LOD* is computed to correspond to the False Negative Rate, $\beta = \beta_{Ld}$ (= γ_{FN}) here. It is computed so that if the true target DNA concentration is equal to *LOD* or higher, we expect in repeated sampling that $100 * \gamma_{FN}$ or fewer samples would test negative. The probability of observing *LOB* or less detections if the concentration is *LOD* or more is less than or equal to β_{Ld} (= γ_{FN}).

The values of *LOB* depend on the number of replicates, NN , so *LOD* does as well.

The $P(Y \leq LOB \text{ given that } S=LOD) \leq \beta_{Ld}$, (= γ_{FN}). L_d is the expected number of detections at concentration/copy number $S=LOD$ and NN replicates.

LOQ is less well defined. Some authors suggest using the “analyte concentration x_q (=LOQ *Limit of Quantification*) for which the experimental relative standard deviation of the responses reaches a fixed level (= γ_Q), for example, the level 0.1.” [Lavagnini and Magno 2007] I interpret the term “relative standard deviation” to mean the coefficient of variation, $CV = sd/\text{mean}$.

In the code below, we use the fits from the ML models to estimate the *LOB*, *LOD* and *LOQ*, for various values of NN replicates for a new observation, i.e. a new (unknown concentration) response number of detections. Only the no intercept model is considered here.

4.2 Print out LOD, LOQ (no intercept model) with confidence intervals, by number of replicates

```
##
## No intercept model TargetA
##      3      8     16     24     32     48     64     96
## LOD      2.5 0.95 0.48 0.32 0.24 0.16 0.12 0.08
## LOD_lower 2.0 0.75 0.37 0.25 0.19 0.12 0.09 0.06
## LOD_upper 3.5 1.30 0.65 0.43 0.33 0.22 0.16 0.11
## LOQ      5.7 3.60 2.39 1.81 1.47 1.06 0.84 0.59
## LOQ_lower 4.5 2.83 1.88 1.43 1.15 0.84 0.66 0.46
## LOQ_upper 7.8 4.92 3.27 2.48 2.00 1.46 1.15 0.80
##
## No intercept model TargetB
##      3      8     16     24     32     48     64     96
```

```

## LOD      2.1 0.77 0.39 0.26 0.19 0.13 0.10 0.06
## LOD_lower 1.5 0.56 0.28 0.19 0.14 0.09 0.07 0.05
## LOD_upper 3.3 1.24 0.62 0.41 0.31 0.21 0.16 0.10
## LOQ      4.6 2.93 1.94 1.47 1.19 0.87 0.68 0.48
## LOQ_lower 3.4 2.12 1.41 1.07 0.87 0.63 0.49 0.35
## LOQ_upper 7.4 4.70 3.12 2.37 1.92 1.39 1.09 0.77
##
## No intercept model TargetC
##          3    8   16   24   32   48   64   96
## LOD      1.30 0.49 0.24 0.16 0.12 0.08 0.06 0.04
## LOD_lower 0.95 0.36 0.18 0.12 0.09 0.06 0.04 0.03
## LOD_upper 2.09 0.78 0.39 0.26 0.20 0.13 0.10 0.07
## LOQ      2.91 1.85 1.23 0.93 0.75 0.55 0.43 0.30
## LOQ_lower 2.12 1.34 0.89 0.68 0.55 0.40 0.31 0.22
## LOQ_upper 4.66 2.96 1.97 1.49 1.21 0.88 0.69 0.48
##
## No intercept model TargetD
##          3    8   16   24   32   48   64   96
## LOD      1.15 0.43 0.21 0.14 0.11 0.07 0.05 0.04
## LOD_lower 0.85 0.32 0.16 0.11 0.08 0.05 0.04 0.03
## LOD_upper 1.77 0.66 0.33 0.22 0.17 0.11 0.08 0.06
## LOQ      2.56 1.63 1.08 0.82 0.66 0.48 0.38 0.27
## LOQ_lower 1.89 1.20 0.80 0.60 0.49 0.36 0.28 0.20
## LOQ_upper 3.96 2.51 1.67 1.27 1.02 0.74 0.58 0.41

```

4.3 Estimates, LOB, LOD, LOQ and confidence limits for a given number of technical reps NN[NNi]

A table of values for all assays is printed. beta is the estimate of beta (β), bSE is the standard error of beta, LOD_Low and LOD_Up are the lower and upper confidence limits for LOD and LOQ_Low and LOQ_Up are the lower and upper confidence limits for LOQ.

The user can set the value for NNi which is the index into NN, the vector number of technical replicates.

Limits for no intercept model for N= 8

```

##      beta  bSE LOB LOD_Low  LOD LOD_Up LOQ_Low  LOQ LOQ_Up
## TargetA 0.39 0.05  0   0.75 0.95   1.30   2.83 3.60   4.92
## TargetB 0.48 0.09  0   0.56 0.77   1.24   2.12 2.93   4.70
## TargetC 0.77 0.15  0   0.36 0.49   0.78   1.34 1.85   2.96
## TargetD 0.87 0.16  0   0.32 0.43   0.66   1.20 1.63   2.51

```

5 eDNA Manuscript Tables and Graphs (not used in *Fishes* manuscript)

Revised for general use to use all eligible targets.

Revised to output results for no intercept models only.

6 Background estimation from negative field controls using no intercept model

Here is an example of a background scenario. The user needs to provide the number of field and/or lab NTC detects (vector `zdetect`) and number of technical replicates (vector `znum`) for the assays. These vectors must have the same length as the number of Targets. `background.cp` is the estimate of the background copy number.

```
## Background num detects (zdetect), num tech reps
##      (znum), No intercept model copy number estimate (background.cp):

##              TargetA      TargetB      TargetC      TargetD
## zdetect      2.00000000  3.00000000  2.00000000  3.00000000
## znum          96.00000000 96.00000000 96.00000000 96.00000000
## background.cp 0.05344608 0.06557117 0.0274597  0.03640747
```

7 Limits of detection/blank given the background estimate, no intercept model

For the ordinary Limit of detection, *LOB* (or *Lc* in code) is the critical value for the test of the null hypothesis that $SQ=0$. Here the null hypothesis is that $SQ=S_Background$ (using `SQBboth0` computed above) and using the no intercept model, we compute a critical value *LOB-B*, Limit of Blank Background corrected. We also compute *LOD-B*, Limit of Detection Background corrected.

```
## TargetA Background estimate (no intercept model) and confidence interval
## 0.05344608 ( 0.04211966 0.0731047 )
##
## TargetB Background estimate (no intercept model) and confidence interval
## 0.06557117 ( 0.04759126 0.1053857 )
##
## TargetC Background estimate (no intercept model) and confidence interval
## 0.0274597 ( 0.01996434 0.0439663 )
##
## TargetD Background estimate (no intercept model) and confidence interval
## 0.03640747 ( 0.02689893 0.05631398 )
##
```

7.1 Print out BACKGROUND limits of detection with confidence intervals, by number of replicates, no intercept model

LOB-B = Limit of Blank, adjusted for background on binomial scale

LOB-Bcn = Limit of Blank, adjusted for background on copy number scale

LOD-B, *LOD-B_lower*, *LOD-B_upper* = Limit of Detection (and confidence interval) adjusted for background on copy number scale.

In the *Fishes* article, we made the following conclusions. If the estimated copy number is less than *LOB-Bcn*, we concluded **No Detect**. If the estimated copy number is between *LOB-Bcn* and *LOD-B*, we concluded **Detect***, that is, eDNA is detected with uncertainty and more sampling is recommended to confirm the result. If the estimated copy number is above *LOD-B*, we concluded **Detect**.

```

## No intercept model, Background Adjusted TargetA
##          3      8      16      24      32      40      48      64      96
## LOB-B      1.0 1.00 1.00 2.00 2.00 3.00 3.00 3.00 5.00
## LOB-Bcn     1.0 0.34 0.16 0.22 0.16 0.20 0.16 0.12 0.14
## LOD-B       5.1 1.61 0.78 0.70 0.52 0.51 0.42 0.32 0.29
## LOD-B_lower 4.0 1.27 0.61 0.55 0.41 0.40 0.33 0.25 0.23
## LOD-B_upper 6.9 2.21 1.06 0.95 0.71 0.70 0.58 0.43 0.39
##
## No intercept model, Background Adjusted TargetB
##          3      8      16      24      32      40      48      64      96
## LOB-B      1.00 1.00 2.00 2.00 3.00 3.00 4.00 4.00 6.00
## LOB-Bcn     0.84 0.28 0.28 0.18 0.20 0.16 0.18 0.13 0.13
## LOD-B       4.13 1.31 0.87 0.57 0.53 0.42 0.41 0.31 0.26
## LOD-B_lower 3.00 0.95 0.63 0.41 0.38 0.30 0.30 0.22 0.19
## LOD-B_upper 6.64 2.11 1.40 0.91 0.85 0.67 0.66 0.49 0.42
##
## No intercept model, Background Adjusted TargetC
##          3      8      16      24      32      40      48      64      96
## LOB-B      1.00 1.00 1.00 2.00 2.00 3.00 3.00 3.00 5.00
## LOB-Bcn     0.53 0.17 0.08 0.11 0.08 0.10 0.08 0.06 0.07
## LOD-B       2.61 0.83 0.40 0.36 0.27 0.26 0.22 0.16 0.15
## LOD-B_lower 1.90 0.60 0.29 0.26 0.19 0.19 0.16 0.12 0.11
## LOD-B_upper 4.18 1.33 0.64 0.57 0.42 0.42 0.35 0.26 0.23
##
## No intercept model, Background Adjusted TargetD
##          3      8      16      24      32      40      48      64      96
## LOB-B      1.00 1.00 2.00 2.00 3.00 3.00 4.00 4.00 6.00
## LOB-Bcn     0.46 0.15 0.15 0.10 0.11 0.09 0.10 0.07 0.07
## LOD-B       2.29 0.73 0.48 0.31 0.29 0.23 0.23 0.17 0.15
## LOD-B_lower 1.69 0.54 0.36 0.23 0.22 0.17 0.17 0.13 0.11
## LOD-B_upper 3.55 1.13 0.75 0.49 0.45 0.36 0.35 0.26 0.23

```