Black WheatEars

mcdevitt

26 mars 2017

Black WheatEars
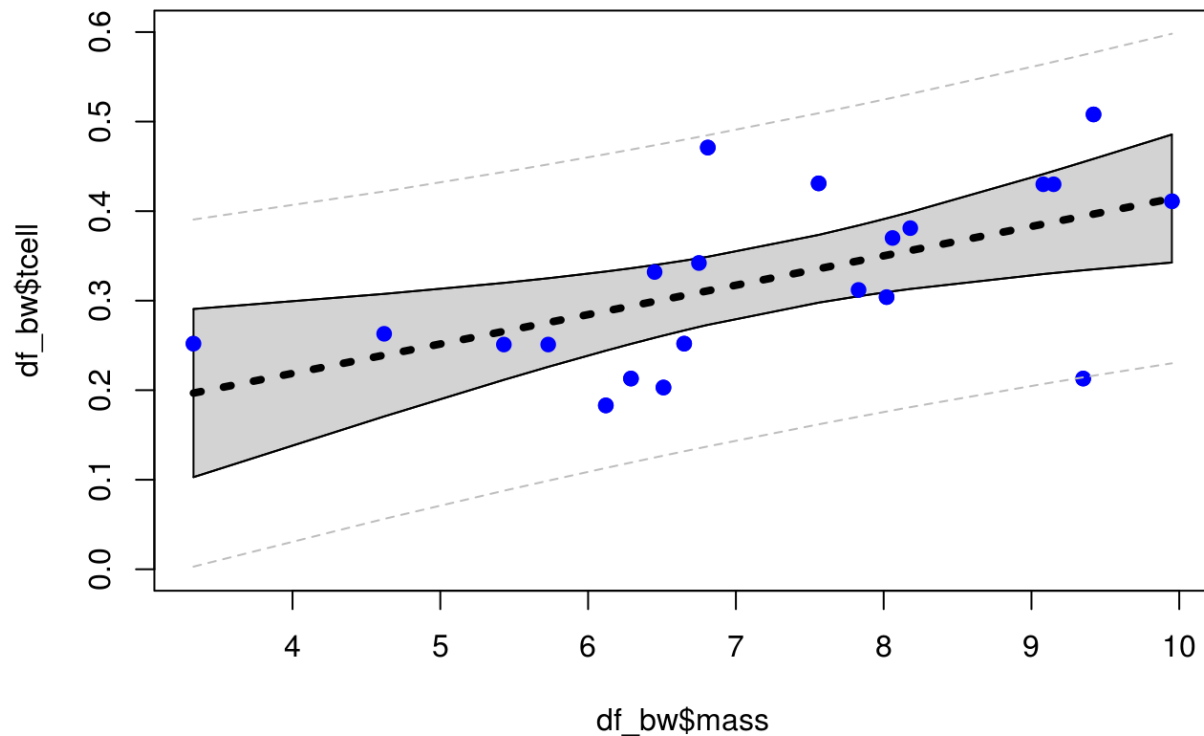

image from Francesco Veronesi - https://www.flickr.com/people/30818542@N04?rb=1
(https://www.flickr.com/people/30818542@N04?rb=1)

Homework 10 - Problem Set 1 - Black Wheatears

b. Analyze the data providing at least the following:

=-=—=-=-
i. A Scatterplot with confidence intervals of the regression line and prediction intervals of the
regression line. Please do in SAS and R!

And the SAS plot …
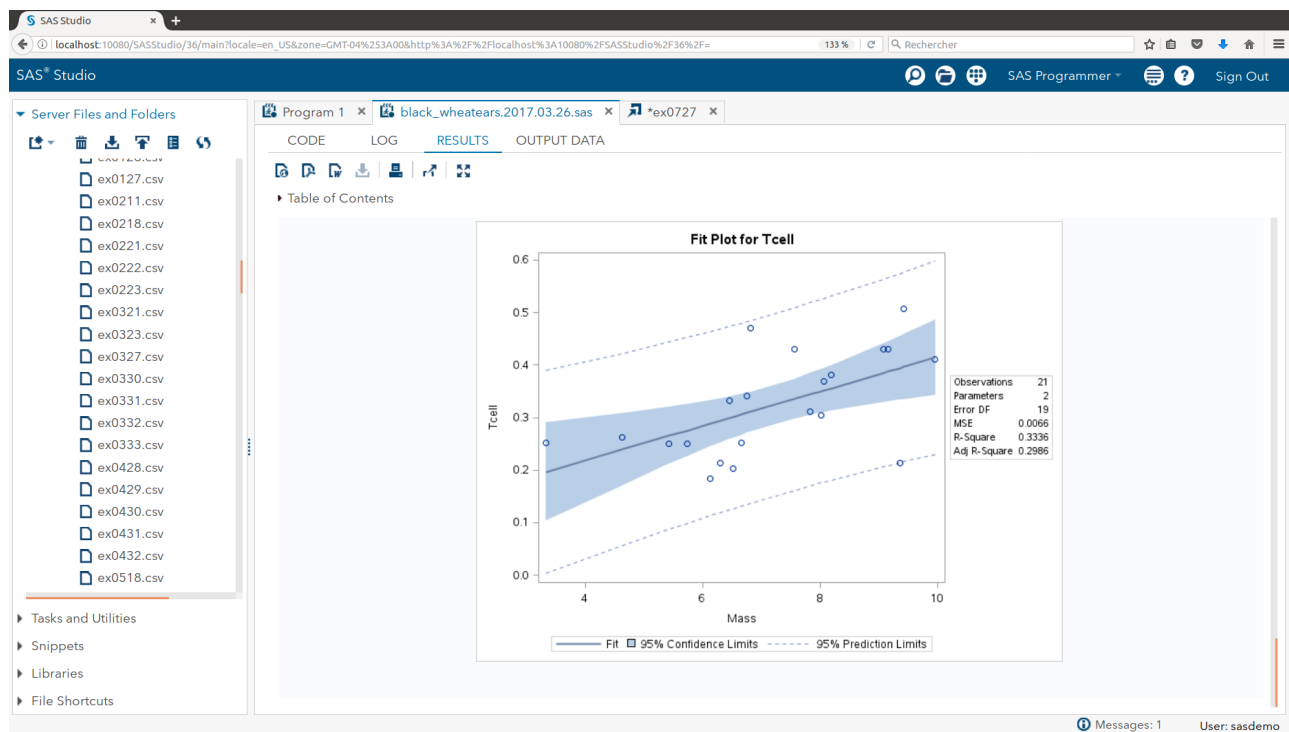
Tcell vs. Stone Mass from SAS PROC REG

=-

ii. A table showing the t-statistics and pvalues for the significance of the regression parameters: .
Please do in SAS and R!

```
##
## Call:
## lm(formula = tcell ~ mass, data = df_bw)
##
## Residuals:
##       Min       1Q   Median       3Q      Max
## -0.18138 -0.04673  0.01796  0.04219  0.15999
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.08750    0.07868   1.112  0.27996
## mass         0.03282    0.01064   3.084  0.00611 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08102 on 19 degrees of freedom
## Multiple R-squared:  0.3336, Adjusted R-squared:  0.2986
## F-statistic: 9.513 on 1 and 19 DF,  p-value: 0.006105
```

And the SAS table …

Tcell vs. Stone Mass from SAS PROC REG

=-

iii. Using the data in ii show all 6 steps of each hypothesis test.

=-=—=-=-=-

Six-Step Hypothesis Test - Slope

1: Ho : b1 = 0; Ha ; b1 != 0
2 : Critical Value : t(0.975, df = 19) = +/- 2.0930241
3 : t = 3.0842956
4 : p_value = 0.006105 < 0.05 ?
5: Reject Ho
6 : There is sufficient evidence to suggest at the alpha = 0.05 level of significance (p-value = 0.006105) that the slope of the regression line that estimates the T cell for Black Wheatears based on the mass of the birds is equal to zero. A 95% confidence interval for the slope is (0.011, 0.055), which is an interval that does not contain the value zero

Six-Step Hypothesis Test - Intercept

1 : Ho : b0 = 0; Ha ; b0 != 0
2 : Critical Value : t(0.975, df = 19) = +/- 2.0930241
3 : t = 1.1121104
4 : p_value = 0.2799648 < 0.05 ?
5: Do not Reject Ho
6 : There is not sufficient evidence to suggest at the alpha = 0.05 level of significance (p-value = 0.2799648) that the intercept of the regression line that estimates the T cell for Black Wheatears based on the mass of the birds is not equal to zero. A 95% confidence interval for the slope is (-0.077, 0.252), which is an interval that does contain the value zero

=-=—=-=-=-
iv. The regression equation.

The Regression equation : * Tcell = 0.0328215 * mass + 0.087497

=-=—=-=-=-
v. Interpretation of the slope and intercept in the model (regression equation.)

The slope represents the rate of change in Tcell estimated for a unit change in stone mass transported by this populations of birds sampled. I.e., for each increase in 1 gram of stone mass there is a corresponding increase in Tcell count of 0.0328215

The intercept represents the estimated average level of Tcell counts associated to a transported stone mass of 0 grams by this population of birds. From a practical point of view, since each bird evalauted did actually transport stones of some mass, the intercept provides an estimate of the lower bound of Tcell counts in this population of birds.

=-=—=-=-=-
vi. Find and interpret the 95% confidence interval for the mean t-cell response conditional on a stone mass of 4.5 grams. Please do in SAS and R!

The confidence interval is the upper and lower bound for the expected mean value at the given indpendent value (e.g., mass = 4.5 grams) for the current regression relationship.

For this particular regression, the confidence interval at stone mass = 4.5 : Tcell count in the range (0.165, 0.306)

=-=—=-=-=-

vii. Find and interpret the 95% prediction interval for the predicted t-cell response given a stone mass of 4.5 grams. Please do in SAS and R!

The prediction interval is the upper and lower bound for an 'next' observation of dependent value at the given independent value, based on the current regression relationshhip. That is to say, for an observation of mass and tcell not included in this analysis, the prediction interval bounds the range of future observations that are expected.

For this particular case, the prediction interval at stone mass = 4.5 : Tcell count in the range (0.051, 0.419)

=-=—=-=-=-

And the SAS table …

Dependent Variable: Tcell

Output Statistics

Obs	Tcell	Mass	Dependent Variable	Predicted Value	Std Error Mean Predict	95% CL Mean		95% CL Predict		Residual
1	0.252	3.33	0.252	0.1968	0.0449	0.1029	0.2907	0.002959	0.3906	0.0552
2	0.263	4.62	0.263	0.2391	0.0327	0.1707	0.3076	0.0563	0.4220	0.0239
3	0.251	5.43	0.251	0.2657	0.0259	0.2116	0.3199	0.0877	0.4437	-0.0147
4	0.251	5.73	0.251	0.2756	0.0236	0.2261	0.3250	0.0989	0.4522	-0.0246
5	0.183	6.12	0.183	0.2884	0.0211	0.2442	0.3326	0.1131	0.4636	-0.1054
6	0.213	6.29	0.213	0.2939	0.0202	0.2517	0.3362	0.1192	0.4687	-0.0809
7	0.332	6.45	0.332	0.2992	0.0194	0.2586	0.3398	0.1248	0.4736	0.0328
8	0.203	6.51	0.203	0.3012	0.0192	0.2611	0.3413	0.1269	0.4754	-0.0982
9	0.252	6.65	0.252	0.3058	0.0186	0.2668	0.3448	0.1318	0.4798	-0.0538
10	0.342	6.75	0.342	0.3090	0.0183	0.2707	0.3474	0.1352	0.4829	0.0330
11	0.471	6.81	0.471	0.3110	0.0182	0.2730	0.3490	0.1372	0.4848	0.1600
12	0.431	7.56	0.431	0.3356	0.0181	0.2978	0.3735	0.1619	0.5094	0.0954
13	0.312	7.83	0.312	0.3445	0.0189	0.3049	0.3840	0.1704	0.5186	-0.0325
14	0.304	8.02	0.304	0.3507	0.0197	0.3095	0.3919	0.1762	0.5252	-0.0467
15	0.37	8.06	0.370	0.3520	0.0199	0.3104	0.3937	0.1774	0.5266	0.0180
16	0.381	8.18	0.381	0.3560	0.0205	0.3131	0.3989	0.1811	0.5309	0.0250
17	0.43	9.08	0.430	0.3855	0.0267	0.3297	0.4413	0.2070	0.5640	0.0445
18	0.43	9.15	0.430	0.3878	0.0272	0.3308	0.4448	0.2089	0.5667	0.0422
19	0.213	9.35	0.213	0.3944	0.0289	0.3339	0.4548	0.2144	0.5744	-0.1814
20	0.508	9.42	0.508	0.3967	0.0295	0.3350	0.4584	0.2162	0.5771	0.1113
21	0.411	9.95	0.411	0.4141	0.0342	0.3426	0.4855	0.2300	0.5981	-0.003071
22	.	4.5	.	0.2352	0.0338	0.1645	0.3059	0.0515	0.4189	.

Tcell vs. Stone Mass from SAS PROC REG

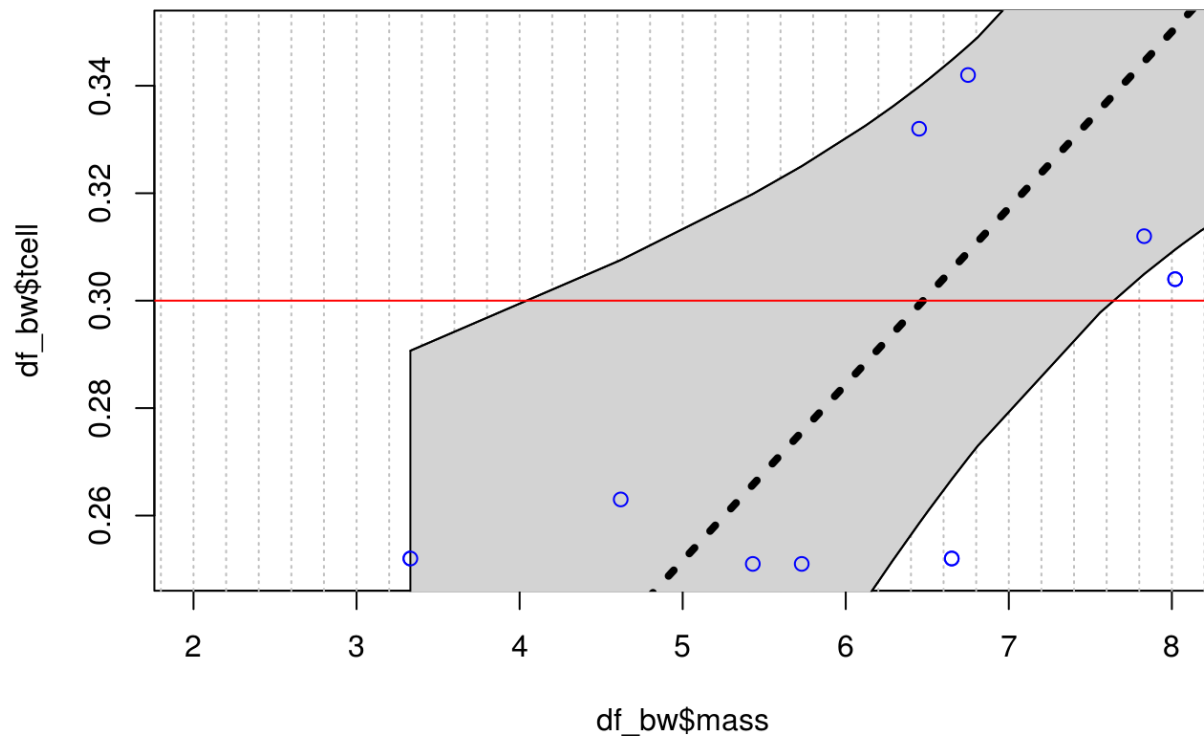=-=—=-=-=-

viii. Using the graphical method, find and interpret the calibration intervals for the t-cell response of 0.3. (Both for mean t-cell response and for a single t-cell response. Please do in SAS and R! (R: package investr)
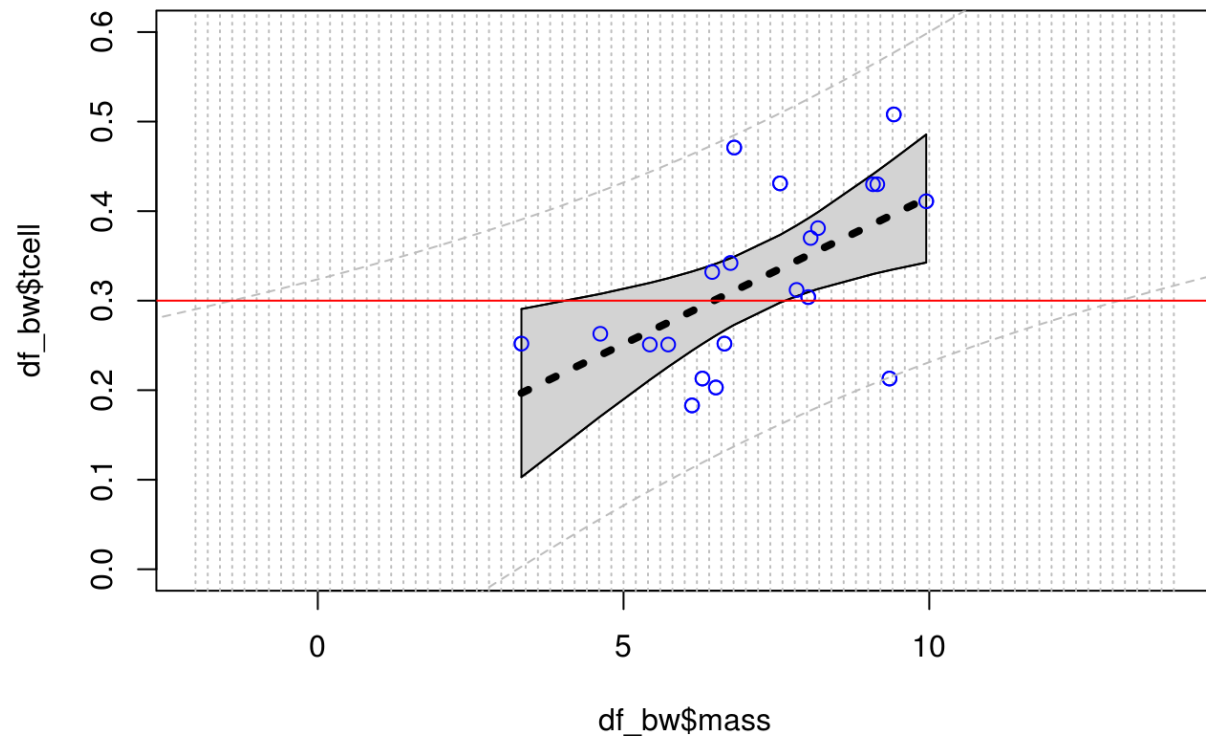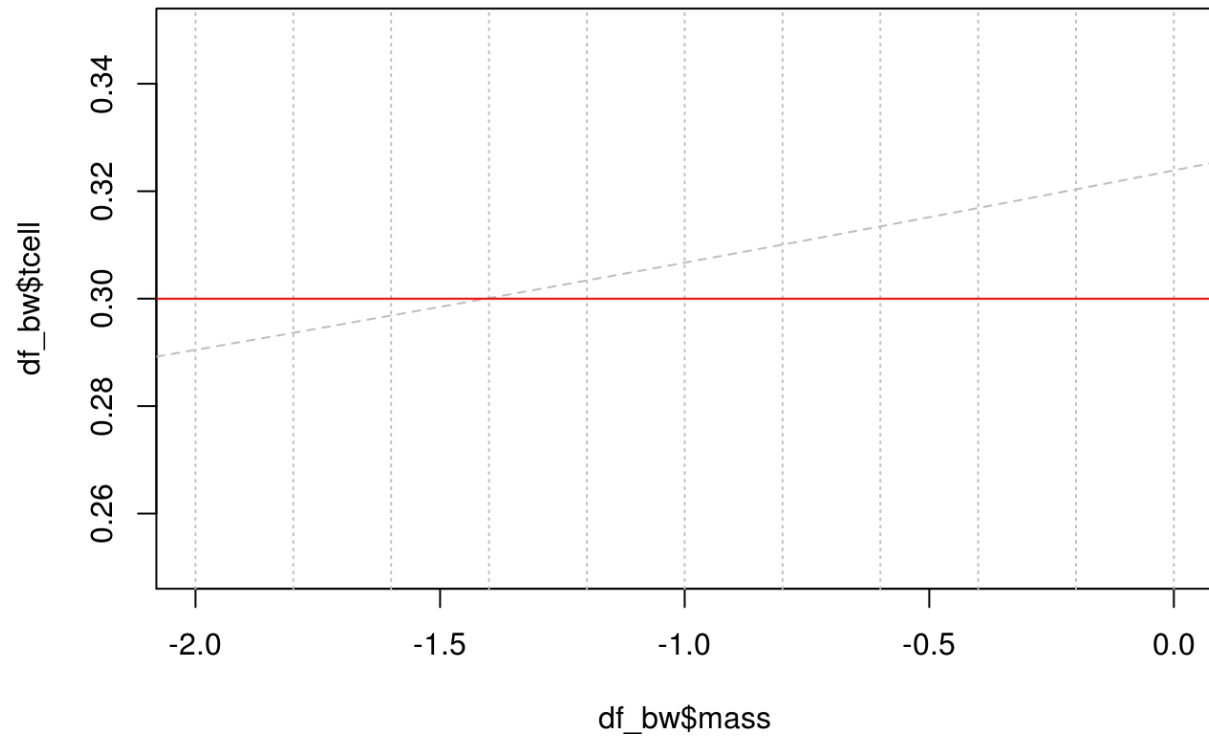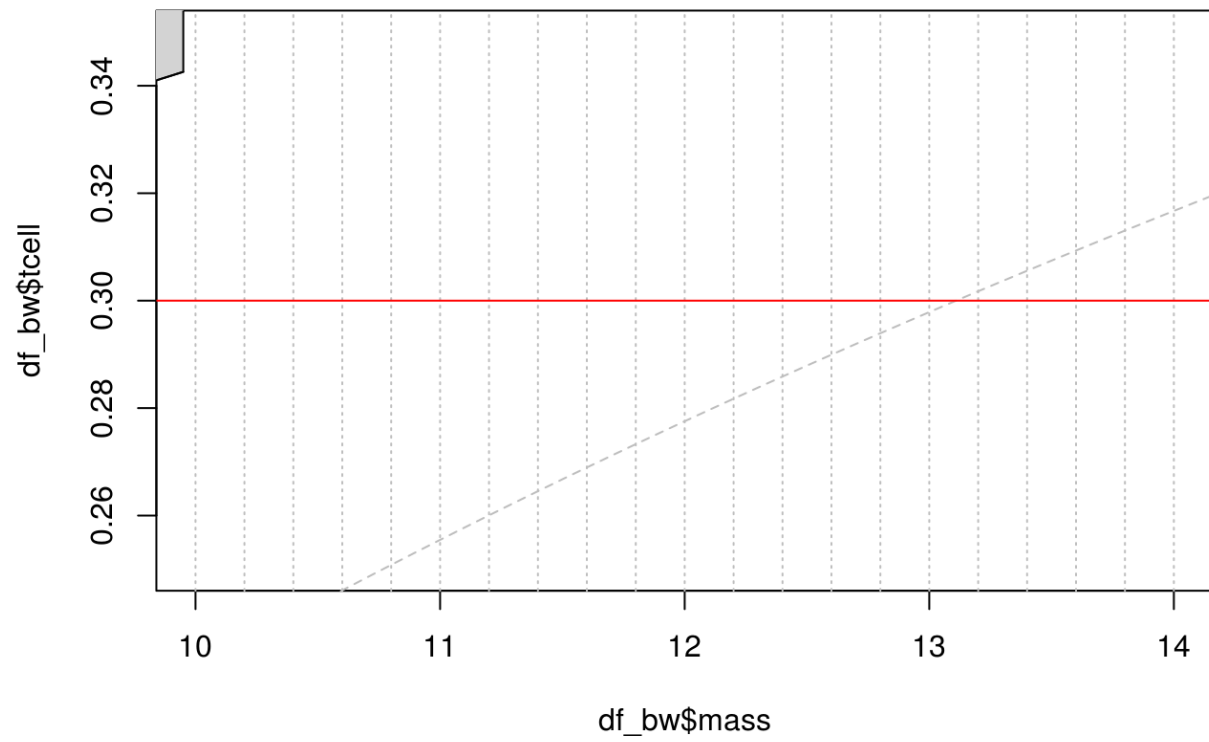
The calibration estimates based on graphical inspection are as follows :

- calibration interval for the mean response of 0.3 = (4, 7.6)
- calibration interval for single point response of 0.3 = (-1.4, 13.2)

=-

ix. Find the same calibration intervals analytically using the SE equations given in class and in the book (Version 3 page 194).

```
y_est <- 0.3
x_at_y_est <- (y_est - b0) / b1

x_at_y_est
```

```
## (Intercept)
##    6.474508
```

```r
se_at_y_est <- predict(fit_all, data.frame(mass = x_at_y_est), se.fit = TRUE)

se_ci_at_y_est <- se_at_y_est$se.fit / abs(b1)

alpha <- 0.05
dof <- dim(df_bw)[1] - 2
crit_value <- qt(1 - alpha/2, dim(df_bw)[1] - 2)

cal_est_at_y_est_upr <- x_at_y_est + crit_value * se_ci_at_y_est
cal_est_at_y_est_lwr <- x_at_y_est - crit_value * se_ci_at_y_est


# ...   calibration of predicted value

res_std_dev <- sigma(fit_all) * sigma(fit_all)

se_pi_at_y_est <- (sqrt (res_std_dev + se_at_y_est$se.fit * se_at_y_est$se.fi
t) ) / abs(b1)

cal_pi_est_at_y_est_upr <- x_at_y_est + crit_value * se_pi_at_y_est
cal_pi_est_at_y_est_lwr <- x_at_y_est - crit_value * se_pi_at_y_est
```

The calibration estimates are as follows :

- calibration interval for the mean response of 0.3 = (5.243091, 7.7059257)
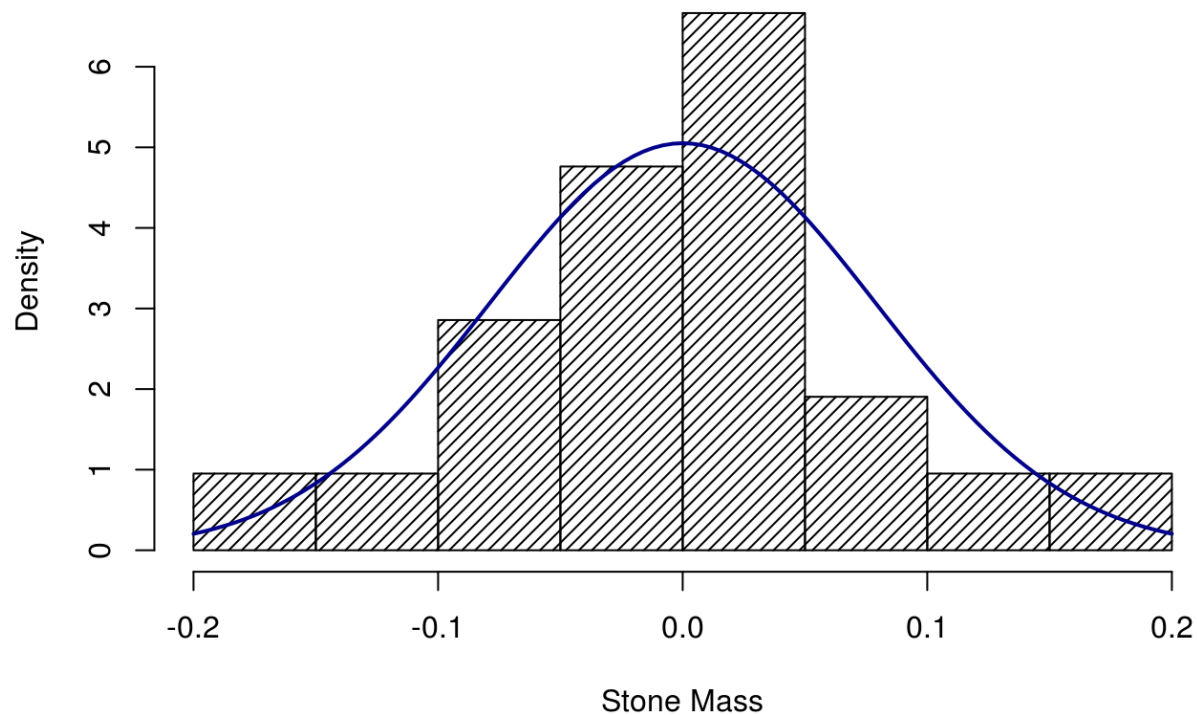- calibration interval for single point response of 0.3 = (1.1631805, 11.7858362)

=-
x. A scatterplot of residuals. Please do in SAS and R!

Black Wheatears - Residual Plot

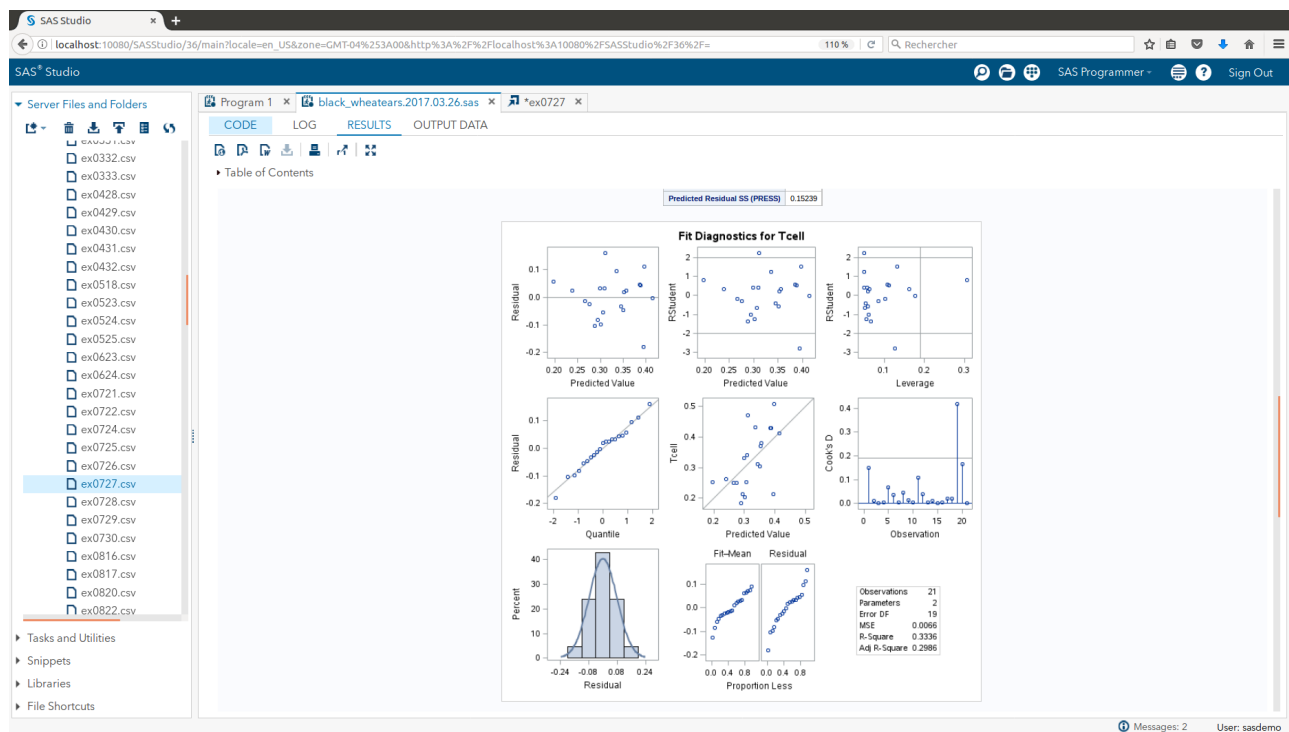=-=—=-=-=-

xi. A histogram of residuals with normal distribution superimposed. (from SAS).

Black Wheatears - Residuals

And the SAS plots …

Tcell vs. Stone Mass from SAS PROC REG