**RNR/ ENTO 613 – Multiple Regression**

**Simple Linear Regression goal:** Quantify association between a response variable (Y) and a single explanatory variable (X)

**Multiple Linear regression Goal:** Quantify association between a response variable (Y) and a series of explanatory variables ($X_1$, $X_2$, …$X_n$)

Multiple regression analysis is an extension of simple linear regression and ANOVA. It focuses on the mean response (Y) at each *combination* of the explanatory variables. The explanatory variables can be either *continuous* or *nominal* (i.e. *categorical*).

**Terminology and symbols:** (Meadowfoam example)

< Display 9.2>

Notation for multiple regression:

$\mu$ {*flowers* | *light*, *time*}

which reads as "the mean number of flowers, as a function of light intensity and timing."

With specific values for the explanatory variables we get:

$\mu$ {*flowers* | *light* = 300, *time* = 24}

which reads as " the mean number of flowers when light intensity is 300 units and time is 24 days prior to PFI (photoperiodic flower induction).

**Multiple Linear Regression Model**

Simple linear regression model assumes that a straight line describes the relationship between the mean response (Y) and X:

$$\mu \{Y \mid X\} = \beta_0 + \beta_1 X$$

Multiple regression extends that approach to cases with more than one explanatory variable:

$\mu \{Y \mid X_1, X_2\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$ (2 explanatory variables [$X_1$ continuous; $X_2$ categorical]; parallel lines)

$\mu \{Y \mid X_1, X_2\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$ (2 explanatory variables [$X_1$ and $X_2$ as above] with interaction term; non-parallel lines)

$\mu \{Y \mid X_1, X_2\} = \beta_0 + \beta_1 \log X_1 + \beta_2 \log X_2$ (2 explanatory variables [$X_1$ and $X_2$ as above] log transformed; parallel lines)

$\mu\{Y \mid X_1\} = \beta_0 + \beta_1 X_1 + \beta_2 X_1^2$ (1 explanatory variable [continuous] with quadratic term; non-monotonic association)

All those models are *linear regression models*, because the regression coefficients ($\beta_1, \beta_2, \ldots\ldots \beta_n$) represent a linear relationship between a *function* of the explanatory variables (e.g. $X_1^2$) and the mean response Y.

As before, *Least-Squares* are used to estimate the regression coefficients ($\beta_0$, $\beta_1$, $\beta_2 \ldots \beta_n$).

The *least-squares solution* consists of the values $\beta_0$, $\beta_1, \ldots \beta_n$ for which the sum of squares of deviations of the observed Y-values from the corresponding values predicted by the fitted regression model *is a minimum* (i.e. the *Error SS* in the ANOVA table is minimum).

**An example: Effect of two Explanatory Variables without interaction**

Y is the *mean number of flowers* produced by 10 plants
$X_1$ is *light intensity* (6 levels between 150 and 900 units)
$X_2$ is *timing* at which light intensity is changed from a basic value of 195 units. $X_2$ has 2 levels: 24 days before photoperiod floral induction (PFI) or at PFI (0 day). PFI is the time at which photoperiod is increased from 8 to 16 hrs of light per day to induce flowering.

<<Display 9.2>>

The following regression model can be seen as describing a *plane*:

$$\mu\{flowers \mid light, time\} = \beta_0 + \beta_1 light + \beta_2 time$$

$\beta_0$ is the height of the plane when both light and time equal zero.
$\beta_1$ is the slope of the plane as a function of *light* for any <u>fixed</u> value of time
$\beta_2$ is the slope of the plane as a function of *time* for any <u>fixed</u> value of light

<<Fig. 9.5 Sleuth>> In 3-D

But it is more easily visualized in 2-D.

<<Fig. 9.8 Sleuth>> In 2-D

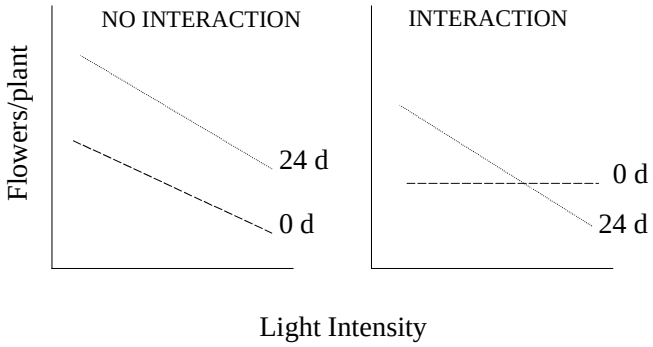
For a parallel line model (no interactions), the effect of an explanatory variable (continuous or nominal) is the *change in mean response associated with a <u>one-unit increase</u> in that explanatory variable, <u>while holding all other explanatory variables fixed</u>.*

[With a significant *interaction* between light and time, the effect of light *depends* on the level of time; see below]

**Interactions**

Two factors "interact" if the effect that one factor has on the mean response (Y) depends on the value of the other factor.

Example: Does the effect of light intensity on flower production depend on timing?




*No interaction* implies that the effect of light intensity is the same at the 2 levels of time. *The lines are parallel.* The relationship is described by a parallel lines model ($\mu \{Y \mid X_1, X_2\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$).

*An interaction* implies that the effect of light depends on timing of the change in light intensity (time). *The lines are not parallel.* The relationship is described by a model with an interaction term ($\mu \{Y \mid X_1, X_2\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$).

<< Display 9.8>>

The presence of a significant interaction in a regression model affects the interpretation of main effects (here light and time). *Those main effects <u>can no longer be interpreted independently of each other</u>*.

**Indicator variables (or Dummy variables) to take into account nominal factors**

In the meadowfoam example, time is treated as a *categorical* variable with 2 levels (typical of ANOVA framework). How do we deal with categorical variables in Multiple <u>regression</u> models?

The methods of regression analysis can be generalized to treat categorical explanatory variables with the use of *Indicator* or *Dummy* variables. Dummy variables allow analyses of categorical variables by *comparison of several regression equations originating from a single multiple regression model*.

In the flower production example, light can be treated as a *continuous* variable (simple linear regression framework; valid if mean responses are linear).

The effect of a continuous variable is estimated by $\beta$ (least squares method)

To estimate the effect of categorical variables in the multiple regression framework, a binary *indicator variable* (also called *Dummy variable*) is created to represent the presence (coded as 1) or absence (coded as 0) of each *level* of the categorical variable.

<<Display 9.7 >>

**Rule for coding indicator variables**

k-1 indicator variables are used to describe the set of levels of a particular nominal variable (a *factor*). These indicator variables are included as explanatory variables in the regression model. The level corresponding to the indicator variable that is <u>*not included*</u> is called the *reference level*.

The coefficient of an indicator variable is the *difference* between the mean response for the indicated category (= 1) and the mean response for the reference level (=0), *at fixed values of the other explanatory variables*.

**Example:** A parallel line model for the flower VS light / time problem.

Light is treated as *continuous* variable
Time is *nominal* (categorical with 2 levels).

Because you code time as nominal, JMP defines a dummy variable to be 0 when time = 0 day prior to PFI and 1 when time = 24 days. Let's designate that dummy variable as *day24.*

The single multiple regression model with the dummy variable *day24* is:

$\mu\{\text{flowers} \mid \text{light, time}\} = \beta_0 + \beta_1 \text{ light} + \beta_2 \text{ day24}$

This model, when treated by a computer program, yields the following <u>*2 models*</u>, which are used to estimate $\beta_2$:

day24= 0: $\mu\{\text{flowers} \mid \text{light, time=0}\} = \beta_0 + \beta_1 \text{ light}$ (reference level)
day24=1: $\mu\{\text{flowers} \mid \text{light, time=1}\} = (\beta_0 + \beta_2) + \beta_1 \text{ light}$

Only the Y's corresponding to day24=0 are used to calculate $\mu\{\text{flowers} \mid \text{light, time=0}\}$; only the observations matching day24=1 are used to calculate $\mu\{\text{flowers} \mid \text{light, time=1}\}$. The difference between the 2 means provides an estimate of $\beta_2$ .

The coefficient corresponding to the effect of time is thus estimated as:

$$time\ effect = \mu\{flowers \mid light, day24 = 1\} - \mu\{flowers \mid light, day24 = 0\}$$

$= [\beta_0 + \beta_1\ light + \beta_2] - [\beta_0 + \beta_1\ light]$

$= \beta_2$ (where $\beta_2$ is the difference between 2 means)

Thus, the multiple regression procedure with the dummy variable day24 *incorporates the two parallel line models within a single model*: $\mu\{\text{flowers} \mid \text{light, time}\} = \beta_0 + \beta_1\ \text{light} + \beta_2\ \text{day24}$

a) This multiple regression model states that the mean number of flowers is a straight- line function of light intensity for both levels of timing, the slope of the line being $\beta_1$.
b) The intercept for the line when day24=0 is $\beta_0$.
c) The intercept for the line when day24 =1 is $\beta_0 + \beta_2$.

The coefficient of the *indicator variable,* $\beta_2$, is the amount by which the mean number of flowers with timing =24 days exceeds that with timing = 0 day, *after accounting for (i.e. averaging) the effect of light intensity differences.*

<<Fig 9.8 Sleuth>>

**Analysis of the parallel line model in JMP: flower VS light and timing**

Light is continuous
Time is nominal

Response: Flowers

Summary of Fit

| | |
|---|---|
| RSquare | 0.799159 |
| RSquare Adj | 0.780031 |
| Root Mean Square Error | 6.441073 |
| Mean of Response | 56.1375 |
| Observations (or Sum Wgts) | 2 |

Lack of Fit

| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Lack of Fit | 9 | 215.31077 | 23.9234 | 0.4377 |
| Pure Error | 12 | 655.92510 | 54.6604 | Prob>F |
| Total Error | 21 | 871.23588 | | 0.8894 |
| | | | | Max RSq |
| | | | | 0.8488 |

Parameter Estimates

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | 77.385 | 2.998156 | 25.81 | <.0001 |
| Time[1-2] | -6.079166 | 1.314779 | -4.62 | 0.0001 |
| Intensity | -0.040471 | 0.005132 | -7.89 | <.0001 |

Sleuth reports a coefficient of 12.2 for the effect of Time, but JMP a coefficient of – 6.1. Why?

1- The sign of the coefficient for a categorical variable depends on which *reference level* is chosen in the analysis. Sleuth used time = 0 as the reference, but JMP used time= 24.

2- JMP *does not* report and test the effect of a categorical variable by taking the difference between a given level and the *reference level*.

The parameters estimated for each level are obtained by taking the difference between the mean response at a given level and *the average of the response at all levels of the variable*.

The test on the effect of *time* labeled Time [1-2] compares the first level of *time* to the *mean of the 2 levels of time*. We obtain an estimate for *time* of - 6.1, because level 1 (change at 0 day) is contrasted to the mean of the responses obtained at the 2 light levels (so it is ½ smaller than 12.2, the difference between the response at time=24 and time =0).

**Analysis of the flower VS light and timing problem: are the lines parallel?**

In multiple regression, an explanatory variable for an *interaction* is constructed as the product of the 2 explanatory variables that are thought to interact.

The model with interaction can be written as:

$\mu\{\text{flowers} \mid \text{light}, \textit{day24}\} = \beta_0 + \beta_1 \text{ light} + \beta_2 \textit{ day24} + \beta_3 (\text{light x } \textit{day24})$

This model yields the following 2 models, which are used to estimate the mean response at the 2 values of *day24*:

day24= 0: $\mu\{\text{flowers} \mid \text{light, time=0}\} = \beta_0 + \beta_1 \text{ light}$ (reference level)
day24=1: $\mu\{\text{flowers} \mid \text{light, time=1}\} = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) \text{ light}$

The single multiple regression model states that both the intercept and slope in the regression of flowers on light depend on timing (i.e. *day24*):

At time = 0, the intercept is $\beta_0$ and the slope $\beta_1$.
At time =24, the intercept is $\beta_0 + \beta_2$ and the slope $\beta_1 + \beta_3$. This model allows ***both*** for different slopes and intercepts.

<<Display 9.8>>

**Analysis of the interaction model flower VS light and timing in JMP**

**Analysis of Variance**

| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Model | 3 | 3467.2765 | 1155.76 | 26.5490 |
| Error | 20 | 870.6598 | 43.53 | Prob > F |
| C. Total | 23 | 4337.9362 | | <.0001 |

**Effect Tests**

| Source | Nparm | DF | Sum of Squares | F Ratio | Prob > F |
|---|---|---|---|---|---|
| Time | 1 | 1 | 886.9504 | 20.3742 | 0.0002 |
| Intensity | 1 | 1 | 2579.7500 | 59.2597 | <.0001 |
| time*intensity | 1 | 1 | 0.5760 | 0.0132 | 0.9096 |

**Parameter Estimates**

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | 77.385 | 3.07118 | 25.20 | <.0001 |
| time[1] | -6.079167 | 1.346802 | -4.51 | 0.0002 |
| Intensity | -0.040471 | 0.005257 | -7.70 | <.0001 |
| time[1]*(intensity-525) | -0.000605 | 0.005257 | -0.12 | 0.9096 |

Because there is no evidence that the coefficient for the interaction ($\beta_3$) is different from zero, we conclude that the effect of light intensity *does not* depend on the timing of change in light intensity (two-sided p-value = 0.91, from a t-test for interaction, 20 degrees of freedom).

We conclude that the parallel line model describes well the relationship between number of flowers and light intensity and timing.

**Fitting your own model with Dummy variables**

Letting JMP choose the reference level is not always practical; also, comparing the effect of each level of a categorical variable is not easy when there are more than 2 levels (JMP uses the difference between each level and the mean of all levels).

You can code your own Dummy variables directly in the data table, which allows you to choose your *reference level*.

<Display 9.7>

For example, a separate-means model (ANOVA) could be fitted to describe the effect of light by using 5 Dummy variables for the 6 levels of light. (Alternatively you could specify *light* as a *nominal* variable in JMP.)

**Example:** In the flower VS light and timing problem, you can create <u>one</u> new column (time has 2 levels) for a Dummy variable *day24*, that takes a value of 0 if time = 0 and 1 if time =24. Using that column to estimate the effect of timing, you get:

Parameter Estimates

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | 71.305834 | 3.273772 | 21.78 | <.0001 |
| Intensity | -0.040471 | 0.005132 | -7.89 | <.0001 |
| Day24 | 12.158333 | 2.629557 | 4.62 | 0.0001 |

Here the effect of the categorical variable time is estimated by taking the difference between the level 24 days and the *reference level* 0 day.

We conclude that increasing light intensity decreased the mean number of flowers by an estimated 4.0 flowers per plant per 100 μmol / $m^2$ / sec (t = - 7.89, df = 21, P < 0.0001), after accounting for the effect of timing. Beginning the light treatment 24 days prior to PFI increased mean number of flowers by 12.2 flower per plant (t = 4.62, df = 21, P = 0.0001), after accounting for variation in light intensity.

**Performing the equivalent of a t-test with a dummy variable and regression**

**Example:** Compare height of 2 populations of plants, one which received only water (n = 10) and the other a fertilizer (n=8).

Recall we compared heights between populations with a 2-sample t-test, which yielded: $t_{16} = 2.99$, P= 0.0087.

Identically, we could have created an *indicator variable* to identify the 2 treatment levels and used linear regression.

Say we created an indicator variable DUMMY coded as:

0 for the control group
1 for the fertilizer group

We then perform the simple linear regression as:

$$\text{Height} = \beta_0 + \beta_1 \text{ DUMMY}$$

For DUMMY = 0, the predicted mean height is : height = $\beta_0$
($\beta_0$ is the average of the control group, i.e. the response when fertilizer = 0)

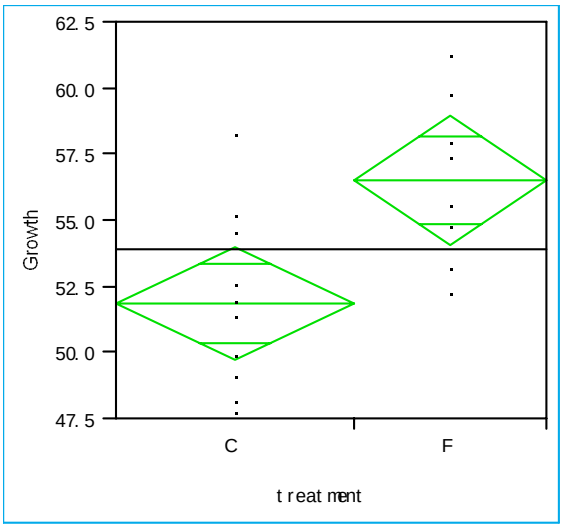
For DUMMY = 1, the predicted mean height is: height = $\beta_0 + \beta_1$
($\beta_1$ is the difference between the average of the treatment and control group)

The test for $\beta_1$ (the coefficient of DUMMY) yields t = 2.99, P = 0.0087, with $\beta_0 = 51.91$ and $\beta_1 = 4.64$.

The difference in *mean response* due to fertilizer is 4.64 (control = 51.91; fertilizer = 56.55), which is the difference between the mean response for each level of the factor (i.e. $\bar{y}_1 \text{ and } \bar{y}_2$).

Thus, the coefficient of the *indicator variable* ($\beta_1$) is the *difference* between the mean response for the *indicated category* (=1) and the mean for the *reference category* (=0).

**t-test** **Regression with indicator**




t-Test

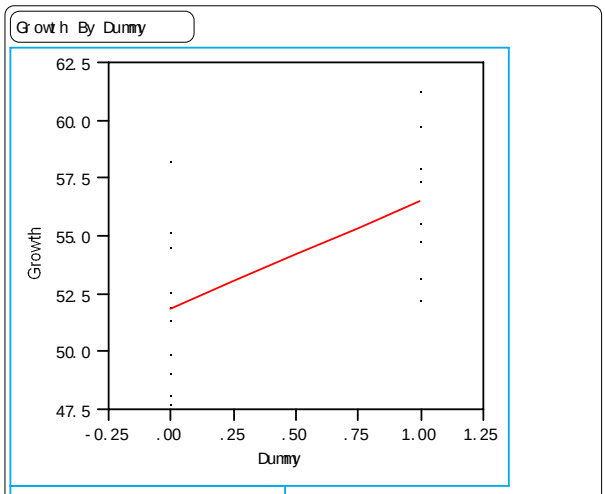
| | Difference | t-Test | DF | Prob>\|t\| |
|---|---|---|---|---|
| Estimate | -4.64000 | -2.988 | 16 | 0.0087 |
| Std Error | 1.55265 | | | |
| Lower 95% | -7.93145 | | | |
| Upper 95% | -1.34855 | | | |

Assuming equal variances

Means for Oneway Anova

| Level | Number | Mean | Std Error |
|---|---|---|---|
| C | 10 | 51.9100 | 1.0351 |
| F | 8 | 56.5500 | 1.1573 |

Std Error uses a pooled estimate of error variance




Linear Fit

Growth = 51.91 + 4.64 Dummy

Summary of Fit

| | |
|---|---|
| RSquare | 0.358223 |
| RSquare Adj | 0.318112 |
| Root Mean Square Error | 3.273272 |
| Mean of Response | 53.97222 |
| Observations (or Sum Wgts) | 18 |

Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Model | 1 | 95.68711 | 95.6871 | 8.9308 |
| Error | 16 | 171.42900 | 10.7143 | Prob>F |
| C Total | 17 | 267.11611 | | 0.0087 |

Parameter Estimates

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | 51.91 | 1.0351 | 50.15 | <.0001 |
| Dummy | 4.64 | 1.552649 | 2.99 | 0.0087 |

## Polynomial Regressions

When the change in the mean of the response variable as a function of the explanatory variable is not monotonic, the straight-line model is not adequate. By including polynomial terms based on the original explanatory variables, regression can be developed for relationships that exhibit curvature.

**Example:** Develop a model to describe the relationship between corn yield (bu/ac) and amount of rainfall (inches) in 6 U.S. states between 1890 and 1927.

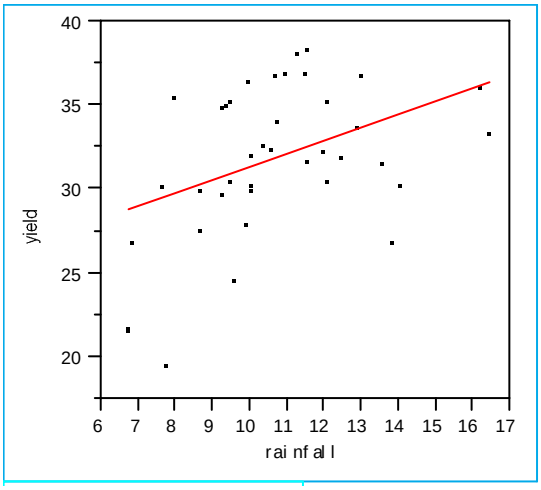
**Linear Fit:** YIELD = 23.552 + 0.775 RAINFALL

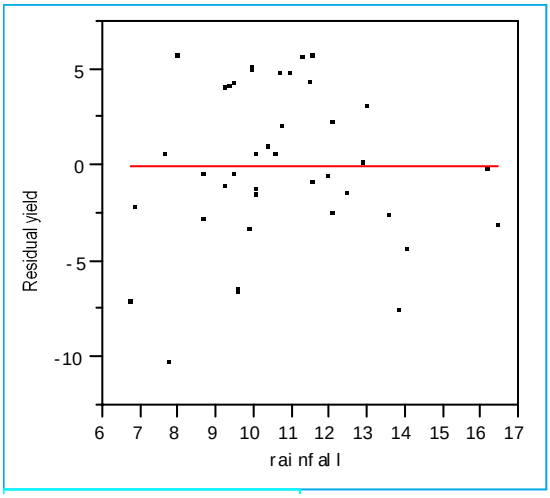
**Parameter estimates:**

Parameter Estimates

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | 23.552102 | 3.236462 | 7.28 | <.0001 |
| rainfall | 0.7755493 | 0.293864 | 2.64 | 0.0122 |

$R^2$= 0.16




Add quadratic term (i.e. rainfall $^2$) to account for curvature (produce a new column RAINFALL X RAINFALL with CALCULATOR in JMP; or use FIT POLYNOMIAL in the Fit Y by X platform):

**Polynomial Fit degree =2:**

YIELD = -5.014 + 6.004 RAINFALL – 0.229 RAINFALL $^2$ (**from Fit Model Platform**)

$R^2$ = 0.26

**Parameter Estimates**

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | -5.014664 | 11.44158 | -0.44 | 0.6639 |
| RAINFALL | 6.004283 | 2.03895 | 2.94 | 0.0057 |
| Rain2 | -0.229364 | 0.088635 | -2.59 | 0.0140 |

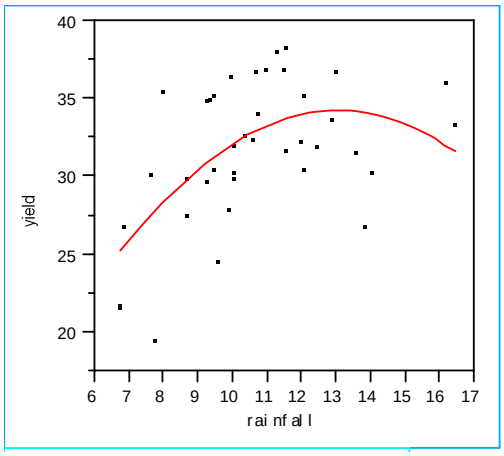
**Polynomial Fit degree =2:**

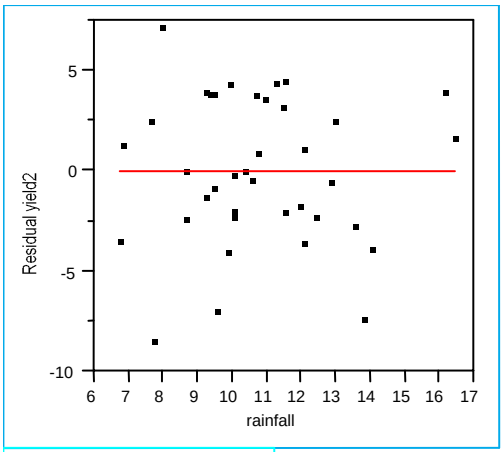
YIELD = 21.66 + 1.057 RAINFALL – 0.229 (RAINFALL – 10.784)$^2$ (**from Fit Y by X Platform**)

$R^2$ = 0.26

**Parameter Estimates**

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | 21.660175 | 3.094868 | 7.00 | <.0001 |
| RAINFALL | 1.0572654 | 0.293956 | 3.60 | 0.0010 |
| (RAINFALL-10.7842)^2 | -0.229364 | 0.088635 | -2.59 | 0.0140 |



**Interpretation:** The linear model seems inadequate**.** The quadratic model describes the curvature in the relationship between rainfall and corn yield that was evident is the scatterplot and patterns of the residuals.

With a significant quadratic term in the model, *corn yield depends on the specific value taken by rainfall*. The mean yield increases up to about 14 inches of rain per summer, then corn yield declines with additional amount of rain. It seems that there is an *optimum* amount of rain.

In fact, because $\beta_2$ is *negative* (i.e. the relationship is concave-down), the value of X that maximizes $\mu\{Y \mid X\} = \beta_0 + \beta_1 X + \beta_2 X^2$ is:

$X_{\max} = -\beta_1 / 2\beta_2$ (use $\beta_1$ and $\beta_2$ from the Fit ***Model platform*** to solve for the optimum).

For a *positive* $\beta_2$ (relationship concave-up), the value above would minimize the mean response (see Sleuth p. 290).

**Interpretation depends on other explanatory variables included in the model**

Example: Determine the association between weight (Y), height ($X_1$) and age ($X_2$) in 12 children.

Multiple Regression: weight = $\beta_0 + \beta_1$ height + $\beta_2$ age

| Term | Estimate | Std Err | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | 6.55 | 10.94 | 0.60 | 0.564 |
| Height | 0.72 | 0.26 | 2.77 | 0.021 |
| Age | 2.05 | 0.94 | 2.19 | 0.056 |

A one-unit increase in height resulted in an estimated 0.72 unit increase in mean weight (t = 2.77, P= 0.021), independently of the effect of age. There is also evidence that a one-unit increase in age was associated with an increase in weight of 2.05 units (t = 2.19, P = 0.056), after accounting for the effect of height.

There is no evidence that $\beta_0$ (intercept) is different from zero (as expected, the mean weight of children is zero when both height and age are equal to zero).

Let's re-analyze the data but this time only considering the relationship between height and weight.

Simple linear regression: weight = $\beta_0 + \beta_1$ height

| Term | Estimate | Std Err | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | 6.19 | 12.84 | 0.48 | 0.64 |
| Height | 1.07 | 0.24 | 4.44 | 0.0013 |

The coefficient associated with height is *larger* than before. This is because taller children also tend to be older. The coefficient of height in this *reduced model* contains the effect of height on weight and in addition reflects the association between age and weight.

In the model Weight = $\beta_0 + \beta_1$ height + $\beta_2$ age, the coefficient of *height* reflects the association between height and weight *after accounting for the effect of age*. In other words, the coefficient of *height* describes the association between height and weight *for children of the same age*.

**A strategy for data analysis (with few explanatory variables)**

Because values of the regression coefficients change when different effects are included in the regression model, one should not think that a single model will be better than all others in describing the data (especially with complex models).

There is a general strategy that can help developing an *inferential model.*

After considering the questions of interests and choosing a design:

1. *Explore* the data with graphical tools: consider transformations; check outliers.
2. *Formulate an inferential model* by wording questions in term of model parameters.
3. *Check* the model fit. a) if appropriate, fit a richer model, i.e. one with interactions or curvature
   b) examine residuals and check for nonconstant variance and outliers
   c) drop unnecessary extra terms
4. *Infer* answers to the questions of interest using appropriate tools, such as confidence intervals and hypothesis tests.
5. *Communicate results* in language of ecology, not statistics.

**Example**. What lifestyles allow evolution of large brains in mammals?

<Display 9.4>

a) Assuming that brains are energetically costly to produce, one hypothesis is that brain size in mammals is limited by the rate at which mothers can provide nutrition to their offspring. If this is true, we would expect a *negative* association between brain weight and litter size.

b) Another reasonable assumption is that producing larger brains require more time than smaller ones. Thus one would expect brain weight to be associated *positively* with gestation period.

c) Because body size is related to brain weight, metabolic rate, and gestation length through allometric relationships, the two hypotheses above should be investigated after accounting for variation in body weight.

Finding the expected relationships would provide support for the hypotheses that production of large brains:

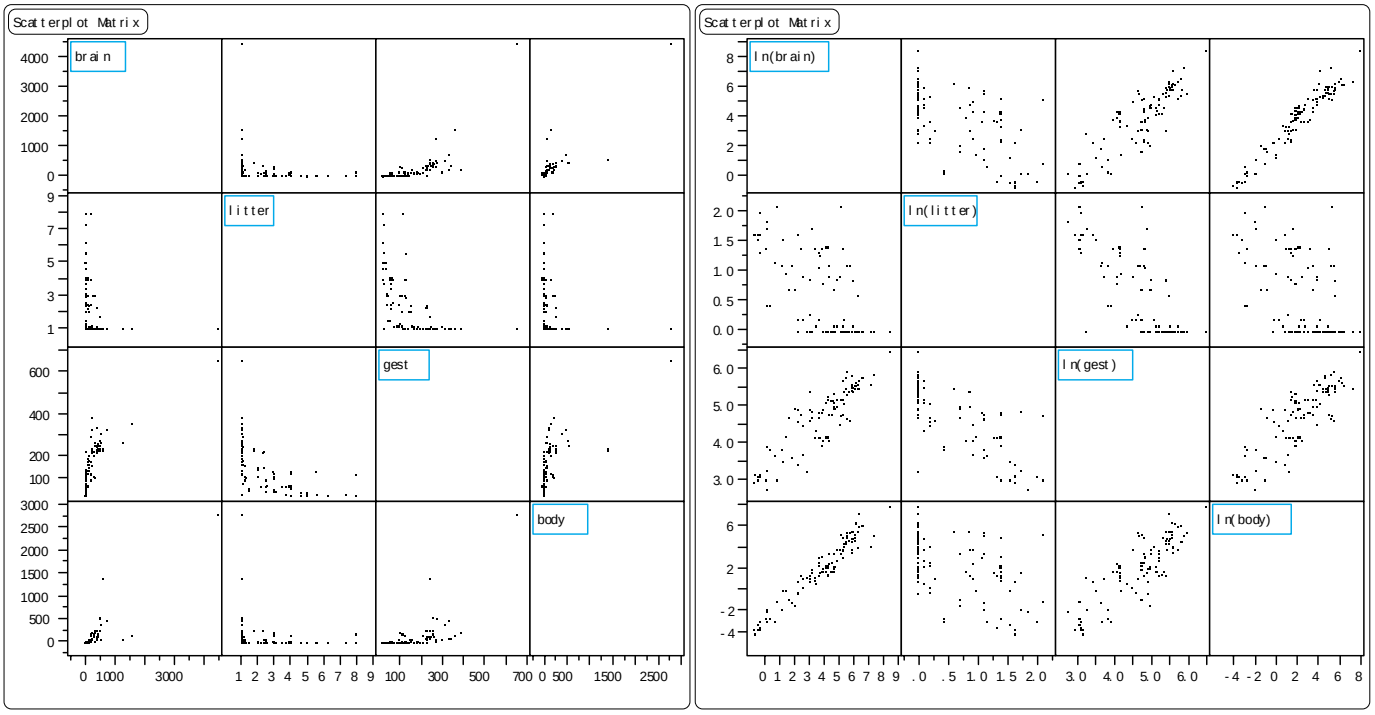
1) requires energy efficiency from mothers (could it be related to diet quality?)

2) may require time (could it be disfavored in “r-selected” species and favored in “K-selected” species?).

response variable: brain weight (g)
explanatory variables: body weight (kg), gestation length (days), litter size

*Which, if any, variables are associated with brain size, after accounting for variation in body size?*

**Not transformed** **Log transformed**



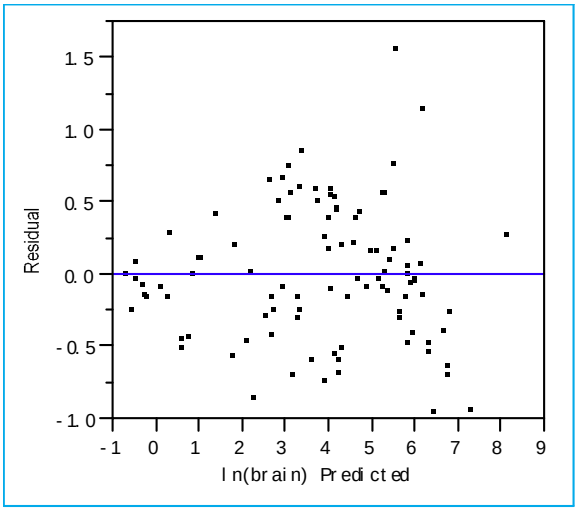

Response: ln(brain)

Summary of Fit

| | |
|---|---|
| RSquare | 0.953695 |
| RSquare Adj | 0.952185 |
| Root Mean Square Error | 0.474755 |
| Mean of Response | 3.864575 |
| Observations (or Sum Wgts) | 96 |

Parameter Estimates

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | 0.8548219 | 0.661672 | 1.29 | 0.1996 |
| ln(litter) | -0.310071 | 0.115927 | -2.67 | 0.0089 |
| ln(gest) | 0.4179421 | 0.140782 | 2.97 | 0.0038 |
| ln(body) | 0.5750714 | 0.032588 | 17.65 | <.0001 |



There is evidence that brain weight is positively associated with gestation length (two-sided p-value = 0.0038 for a test that the slope is zero) after accounting for the effect of body weight and litter size. Brain weight is also negatively associated with litter size (two-sided p-value = 0.0089) after accounting for body weight and gestation period.

Although there are likely *cluster effects* due to taxonomic relatedness that may strongly influence statistical inferences, this analysis provides some evidence in favor of the ideas that rate at which mothers provide nutrition to their offspring or gestation length could limit the evolution of large brains in mammals.