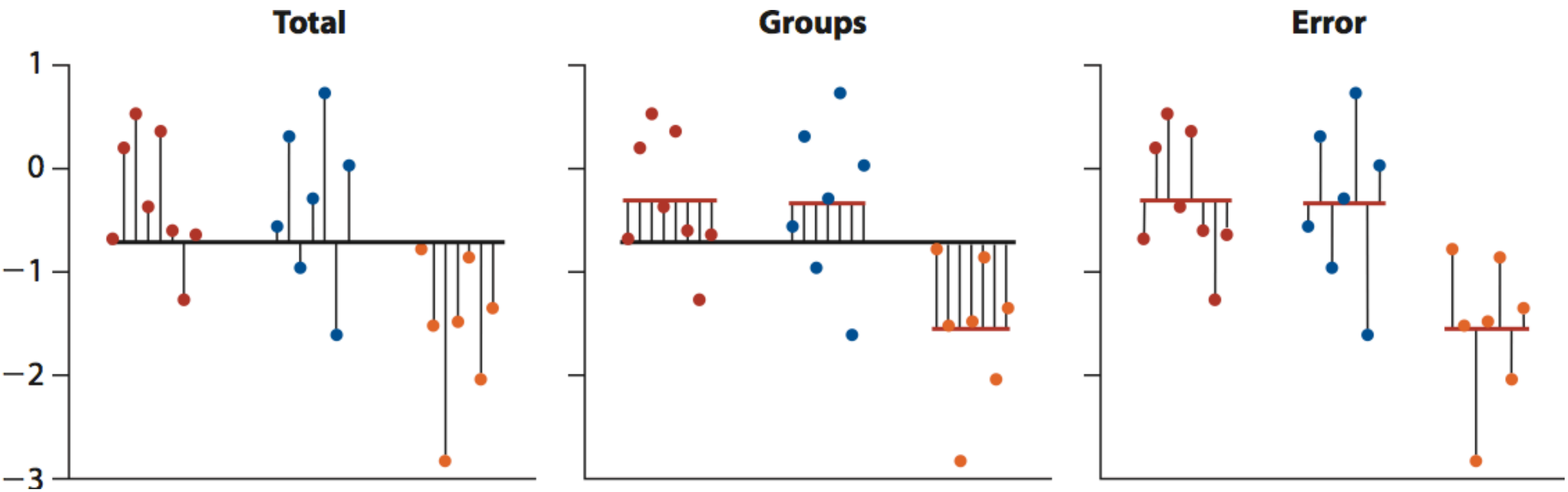
# General Linear Models

# General linear model

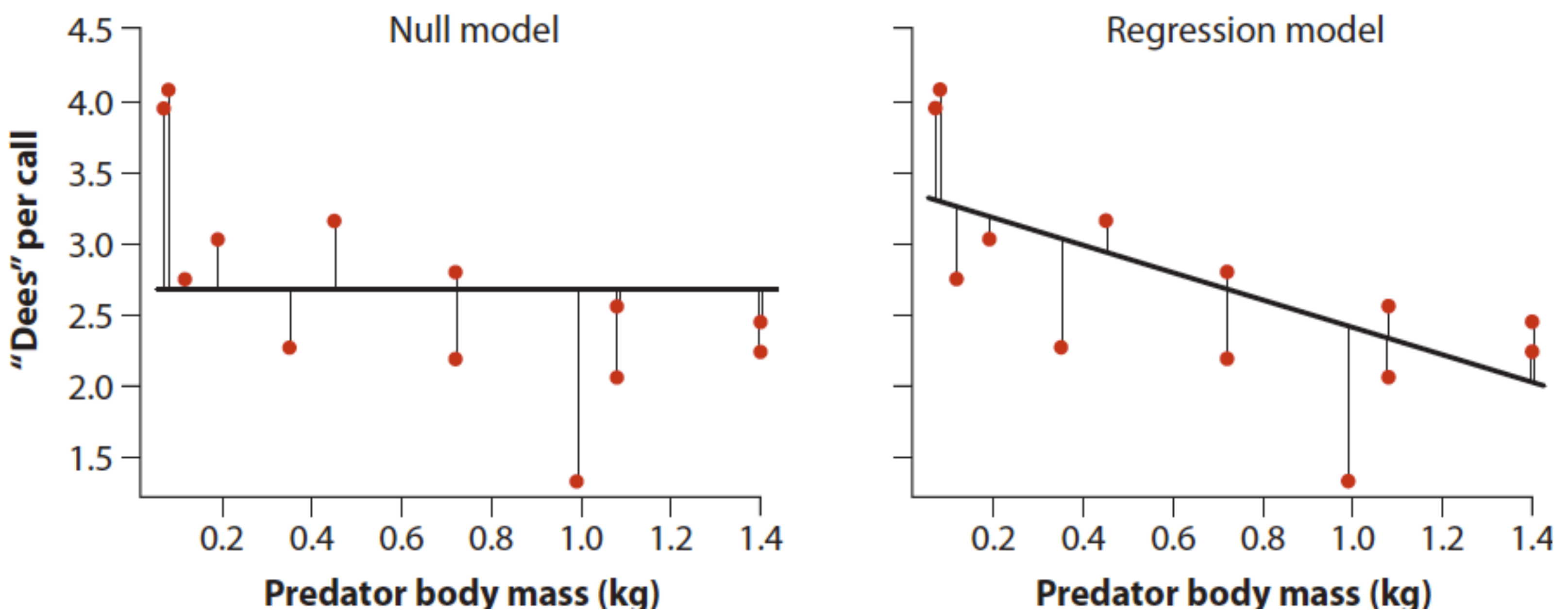
- Response variable, **Y**, can be represented by a linear model plus random error
  - Scatter of Y measurements around the model is random error
- So far, we have looked at (univariate) ANOVA



# General linear model

– We have also looked at the linear regression

$$\boxed{Y = \alpha + \beta X + \varepsilon}$$



# General linear model

– Extends the linear regression in two ways

- More explanatory variables (>1)
- Allows use of **categorical** explanatory variables

- Example:
  – Linear model for single-factor ANOVA

$$\boxed{Y = \mu + A}$$

# General linear model

– Extends the linear regression in two ways

- More explanatory variables
- Allows use of **categorical** explanatory variables
- Example:
  – Linear model for single-factor ANOVA

$$Y = \mu + \boxed{A}$$

Grand Mean

Treatment Effect

# General linear model

- Linear Model for single-factor ANOVA
- Linear Regression

$Y = \mu + A_i$ $\qquad A_i$ = group mean - μ

$$Y = \alpha + \beta X$$

You are fundamentally fitting two models in both of these cases

RESPONSE = CONSTANT + VARIABLE

# General linear model

We have already encountered some methods which are actually specific examples of the general linear model:

- Linear regression
- One factor ANOVA
- t-test

Additionally, we will encounter these:

- Analysis of covariance (video 20)
- Multiple regression

# General linear model

| Linear Model | Other Name | Example-study Design |
|---|---|---|
| $Y = \mu + X$ | Linear Regression | Dose-Response |
| $Y = \mu + A$ | One-way ANOVA | Completely randomized |
| $Y = \mu + A + b$ | Two-way ANOVA, no replication | Randomized block |
| $Y = \mu + A + B + A{*}B$ | Two-way, fixed effects ANOVA | Factorial Experiment |
| $Y = \mu + A + b + A{*}b$ | Two-way, mixed effects ANOVA | Factorial Experiment |
| $Y = \mu + X + A(+A{*}X)$ | Analysis of Covariance (ANCOVA) | Observational Study |
| $Y = \mu + X_1 + X_2 + X_1{*}X_2$ | Multiple Regression | Dose-Response |

*Match the linear model with the "name of the test" and a given (example study design).*

Variable Descriptions:

mu = Constant X = Numerical explanatory variable Y = Numerical response variable

A, B = Fixed, categorical variables b = Random-effect categorical variable (ex. blocking)

1) Y = mu + X
2) Y = mu + A
3) Y = mu + X + A
4) Y = mu + A + B + A*B

I) "Linear Regression" (Dose-response)
II) "One-way, single factor, ANOVA" (Completely randomized)
III) "Two-way, fixed-effect ANOVA" (Factorial experiment)
IV) "Analysis of covariance, ANCOVA" (Observational study)

**Clicker Choices:**

A) 1=I, 2=II, 3=III, 4=IV
B) 1=I, 2=II, 3=IV, 4=III
C) 1=II, 2=I, 3=IV, 4=III
D) 1=III, 2=IV, 3=I, 4=II
E) 1=II, 2=I, 3=III, 4=IV

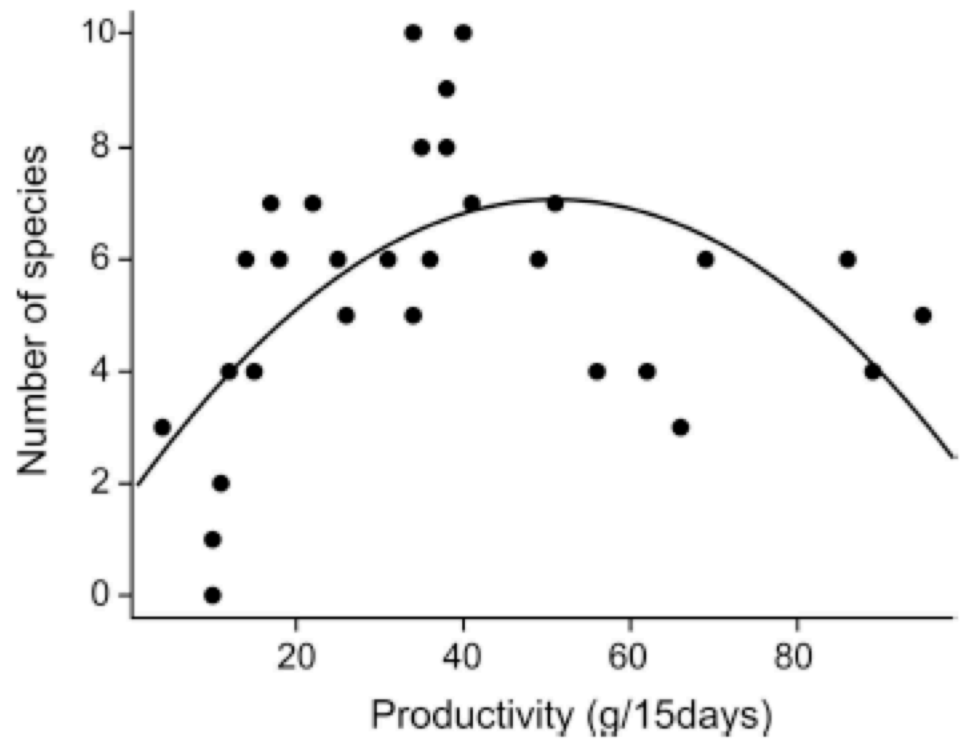
# Note: General linear model

In the form:

$$Y = \beta_0+\beta_1X_1+\beta_2X_2+\ldots+\text{error}$$

Doesn't have to be LINEAR relationship:

$$Y = \beta_0+\beta_1X+\beta_2X^2$$ (Quadratic)



# General linear models:

$H_0$: Treatment means are same

$H_A$: Treatment means are not all the same

Significance of a treatment variable is tested by comparing the fit of two models, $H_0$ and $H_A$, to the data by using **F-test**

$$\text{F-test} = \frac{H_A}{H_0} = \frac{\text{Constant + Variable}}{\text{Constant}}$$

*Does the additional parameter, the variable, improve the fit of the data **significantly**?*

– ANOVA table

– P-value leads to rejection or FTR $H_0$

# Often appropriate/useful to investigate >1 explanatory variable simultaneously

- Efficiency
- Interactions

## Three major approaches:

- Blocking
  - Improve detection of treatment effects
  - If nuisance variable is known and controllable
- Factorial experiment
  - Investigate effects of ≥ 2 treatment variables
  - Interactions
- Covariates
  - Confounding variables
  - Nuisance variable is known but uncontrollable

# Assumptions of general linear models:

- Test these by using **residual plots**
- Same as for regression and ANOVA

  - Random sample from the population of possible measurements

  - Normal distribution

  - Variance of response variable is the same for all combinations of the explanatory variables