

### **Module 4A Questions:**

1. Researchers measure **Activity** and **Glucose12** levels in a group of mice. They calculate a **correlation coefficient of  $r = -0.72$**  between Activity and Glucose12.

**a)** Explain what this correlation tells you *and what it does NOT tell you* about the relationship between Activity and Glucose12.

**b)** If the researchers fit a **regression model** with Glucose12 as the outcome and Activity as the predictor, what additional information does regression provide that correlation alone cannot?

2. A study of 12 mice finds:

- A **correlation** of  **$r = 0.55$**  between Weight12 and Il6 expression.
- A **regression model**:

$$\widehat{\text{Il6}} = 1.2 + 0.18 \times \text{Weight12}$$

**a)** Why does correlation treat the relationship between Weight12 and Il6 as *symmetric* (or not a dependent/independent relationship), while regression does not?

**b)** In the regression model above, what does the slope (0.18) tell you that the correlation coefficient cannot?

3. Researchers model the relationship between **Activity** and **Glucose12** in mice and obtain the following estimated regression line:

$$\widehat{\text{Glucose12}} = 180 - 6.5 \times \text{Activity}$$

**a)** Interpret the slope: what does “-6.5” mean in the context of mouse physiology?

**b)** Predict the Glucose12 value for a mouse with an Activity score of **5.0**.

**c)** A mouse has Activity = 7.0 and an actual Glucose12 of 142 mg/dL. Compute the **residual** (observed – predicted) and explain what it means.

4. Researchers want to estimate a simple linear regression predicting **Glucose12** from **Activity** using the 12-mouse dataset. You are given the following summary statistics:

- Mean Activity:

$$\bar{X} = 5.0$$

- Mean Glucose12:

$$\bar{Y} = 140$$

- Sum of cross-products:

$$\sum(X_i - \bar{X})(Y_i - \bar{Y}) = -336$$

- Sum of squares for Activity:

$$\sum(X_i - \bar{X})^2 = 21.5$$

Use these to answer the following:

- Compute the slope of the regression line
- Compute the y-intercept using means of the two variables
- Write the final regression equation
- Using your regression equation, predict the Glucose12 value for a mouse with activity 6.0.
- If an actual mouse had Activity=6.0 and Glucose12=150 mg/dL, compute the residual.