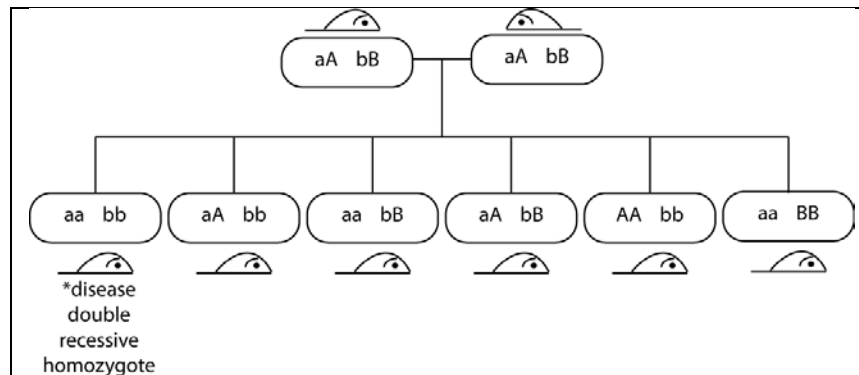


Name: _____

Question 1

You are studying a genetic disease using a mouse model. One hypothesized model for the disease is that it involves two genes that independently exhibit classical Mendelian inheritance patterns, and the disease only occurs when both loci are homozygous for the recessive allele. This model predicts that $1/16^{\text{th}}$ of the offspring from a pair of doubly heterozygous parents (i.e., both parents are heterozygous in both genes) will have the disease.



You have two breeding pairs that are known from genotyping to be such double heterozygotes. In one breeding pair, 2 out of 42 offspring (from several litters) have the disease. In the second breeding pair, 10 out of 50 offspring have the disease.

- (A) What test would you use to determine if the rates of disease incidence of these two pairs are significantly different from each other? Explain your reasoning, and what conclusions you might be able to draw from this test, but do not perform the test.

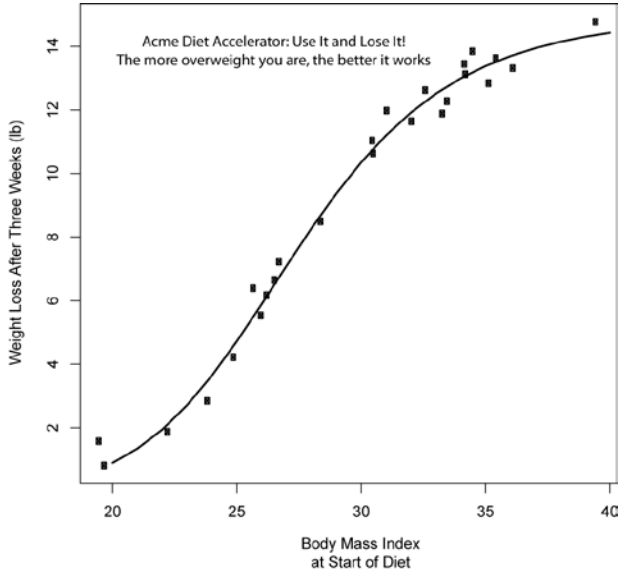
Name:

- (B) What analysis would you perform to determine if the data you have is consistent with the prediction that $1/16^{\text{th}}$ of the offspring will have the disease? Explain your reasoning, and what conclusions you might be able to draw from this test, but do not perform the analysis.

Name: _____

Question 2

The Acme Vitamin Company sells a dietary supplement that they claim accelerates weight loss for overweight and obese dieters. Amazingly, the supplement work regardless of the kind of diet you use, and it is organic! Their literature cites a scientific study, and includes the following figure:



	Body Mass Index
	$BMI = \frac{weight (lb)}{[height (ft)]^2} \times 4.88$
Normal	18.5 – 23.0
Overweight	23.0 – 27.5
Obese	27.5 – 40.0

(A) Comment critically on the analysis implied in the figure above.

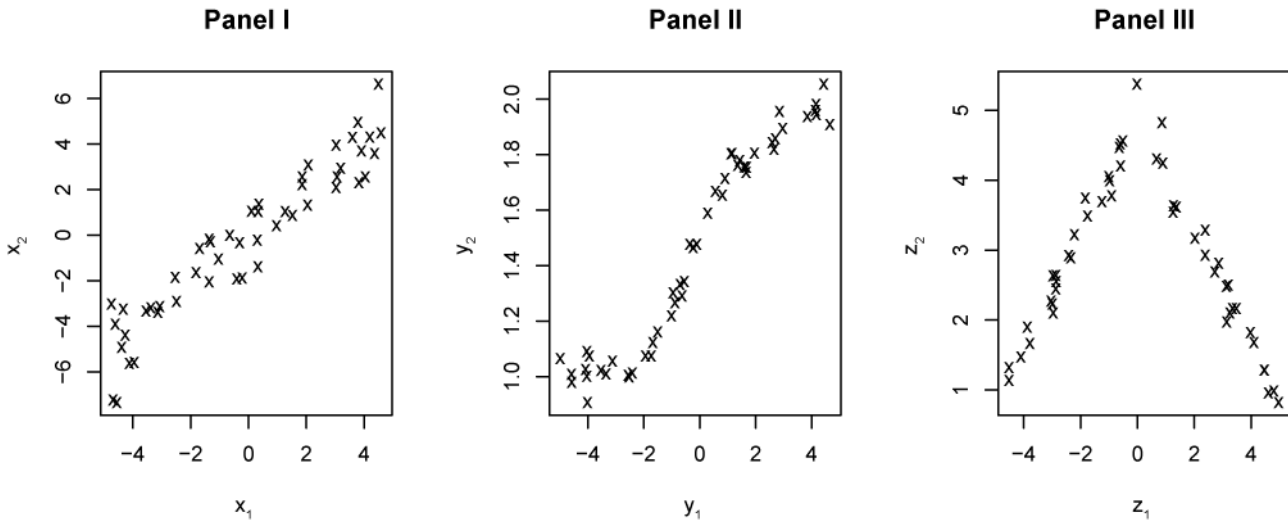
Name:

(B) What study design and data analysis protocol would you suggest to test the manufacturer's claim?

Name: _____

Question 3

Each of the panels below shows a plot of two independently measured variables. In each case, some form of correlation is clearly apparent.



(A) For Panel I, what test and/or procedure would you use to characterize and quantify the observed correlation? Explain why.

(B) For Panel II, what test and/or procedure would you use to characterize and quantify the observed correlation? Explain why.

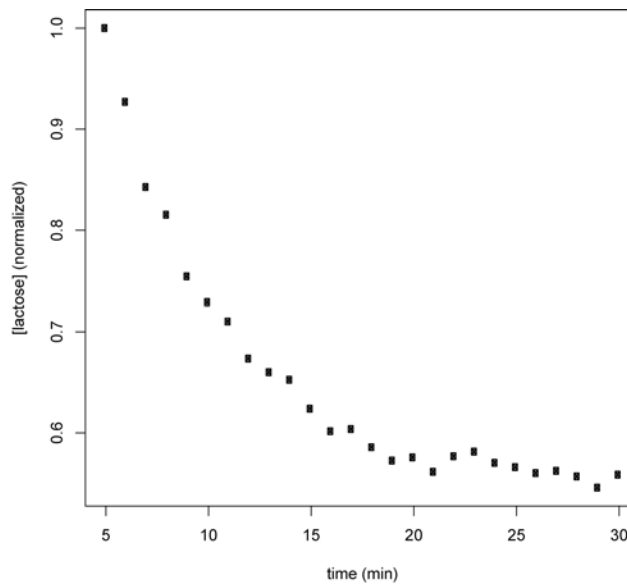
Name:

- (C) For Panel III, what test and/or procedure would you use to characterize and quantify the observed correlation? Explain why.

Name: _____

Question 4

A plot of lactose concentration in a yeast growth medium is plotted as a function of time.



Three possible mathematical models are proposed:

Model I	Model II	Model III
$[lactose] = \beta_I \cdot e^{-t/\lambda_I}$	$[lactose] = \beta_{II} \cdot e^{-(t/\lambda_{II})^2}$	$[lactose] = \beta_{III} \cdot e^{-(t/\lambda_{III})^n}$

Assume you have fit the data to each model using a non-linear least squares regression algorithm, and determined the best fit β s and λ s.

(A) Which model do you think will have the lowest sum of squares of residuals? Is this necessarily the best model? Why or why not?

(B) Which models can be compared using the F-test? Why or why not? If any of these comparisons are not appropriate for the F-test, what can you do to compare the models?

Name:

(C) Which of these models are amenable to variable transformations that would allow for a linear least square fitting algorithm to be used? Show how you would do this.

(D) Can you compare the results of these transformed models using the F-test? Why or why not? If any of these comparisons are not appropriate for the F-test, what can you do to compare the models?

Name:

[no question on this page – just extra space]

Name:

[no question on this page – just extra space]