Course Preview: Designing and Running Randomized Evaluations

This course preview is meant to give prospective learners the opportunity to get a taste of the content and exercises that will be covered in the course. While there are no prerequisites for this online course, it is recommended that learners have some familiarity with economics or statistics. Each question below is tied to concepts that will appear in this course, all of which it would be good to feel comfortable with. If you are new to these subjects, or eager to refresh your memory, please do consult the available resources below, and be prepared to refer to these resources over the course of the class. Try to first answer these questions without consulting the resources, but fear not if you do consult them - being an agile user of outside resources will help you succeed in this course.

A score of 60% or above in this course preview indicates that you are ready to take this course, while a score below 60% indicates that you should further review some concepts in the attached materials before commencing the course.

Useful Resources:

- *Review of key statistical concepts*: Khan Academy: Statistics and probability
- *Introduction to econometrics*: Mastering 'Metrics: The path from cause to effect
- *Overview of RCTs*: Running Randomized Evaluations: The book and the blog

1. **Probability**: A fair coin is tossed three times. What is the probability of getting at least two heads? (1 point)

2. **Exponential Functions**: As an epidemiologist, you realize that you can use an exponential function to map the spread of malaria, as follows:

$$Malaria_{\text{today}} = 10 + 1.5 * (Malaria_{\text{yesterday}})^2$$

where $\text{Malaria}_{\text{today}}$ is the number of individuals in the population with malaria today, while $\text{Malaria}_{\text{yesterday}}$ is the number of individuals in the population with malaria yesterday.

If you know there are 16 cases of malaria today, how many were there yesterday? (1 point)

3. **Measurement**: In your role as a clinic manager, you conduct monthly tests on your nurses' sanitation practices. In July, only 50% of nurses passed the sanitation test. (1 point)

Thinking ahead to the test in August, would you rather see a:

(A) a 50% increase in the number of nurses who passed

(B) a 50 percentage point increase in the number of nurses who passed

(C) A and B are equivalent, so both responses are correct

4. **Interpreting Regressions**: The administration of a school district in Rajasthan, India wishes to understand the effect of class size on learning outcomes, measured as:

$$score = \alpha + \beta_1 * ratio + \beta_2 * income + \beta_3 * attendance$$

where the variables are defined as:

Dependent Variable (Y) :

- *score* = student's score on last year's final exam (on a 100 point scale)

Independent Variables (X) :

- *ratio* = number of students per teacher in the student's classroom
- *income* = annual household income of student (in thousands of Rupees)
- *attendance* = number of class days missed by the student each month

Data was collected and analyzed, producing the following suite of regression models. Please read the regression table below and answer the four associated questions.

4.1. (a) In the first regression, what is the effect of adding one more student to the classroom (e.g. increasing the student-teacher ratio by one degree) on student test scores? Is this result statistically significant? (2 points)

Table 1: Regressions of Student-Teacher Ratio on Student Test Scores

Variable Label	(1)	(2)	(3)
Student-teacher ratio	-2.070**	-0.649	0.047
	(0.719)	(0.552)	(0.495)
Parental income		1.769**	1.436**
		(0.821)	(0.075)
Days absent			-1.41*
			(0.583)
Constant	48.36**	42.75**	41.12**
	(1.036)	(0.830)	(0.5956)
n	525	525	525

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

(b) How does the effect of the student-teacher ratio on test scores differ across the regression models? What does this imply about the true impact of the student-teacher ratio on test scores? (2 points)

4.2. What is the standard error for parental income in the second regression model? (1 point)

(a) 525
(b) 0.821
(c) 0.01
(d) 1.769

4.3. Consider the third regression model. If a student misses three days of class, what is the predicted change in test scores? (1 point)

(a) A decrease of 4.23 points
(b) A decrease of 1.41 points
(c) An increase of 0.583 points
(d) An increase of 4.23 points

4.4. What does n represent in the above regression table? (1 point)

(a) alpha level
(b) sample size
(c) standard error
(d) error term