**Fahri Ulkat**

**090220756**

**Problem Set 1 Solutions**

**QUESTION 1**

**1.Econometrics**: Ecometrics is a way of analyzing economic data by using statistical and mathematical methods to test the theories and forecasting the future values.

**2.Experimental and Observational Data:** Experimental data is acquired by an experiment which is made to collect data to check a treatment or policy or to review a causal effect. Observational data is acquired from existing behavior, not from an experiment.

**3. Cross-sectional Data:** Cross-sectional data refers that data collected at a single point in time with various variables (age, income, gender, educatin level etc.).

**4. Panel Data:** Panel data ,also known as longitudinal data, is a data type which contains various entities and observations of these entities over multiple periods.

**5.Time Series Data:** Time Series data are data for a specific entity and its obsevrations during the multiple time periods.

**6. An ideal randomized controlled experiment (IRCE):** IRCE is an experiment with at least 2 groups randomly chosen to eliminate other effects and invesitage the causal effect perfectly.

**7. The Causal Effect:** The causal effect is the impact that one variable has on another.

**8. Confounding variable (omitted variable):** When analyzing a causal effect there might be some other variables that affects the causality. Those variable are called confounders.

**QUESTION 2**

1. $\sum_{i=1}^{n} ax_i = a\sum_{i=1}^{n} x_i$

$\sum_{i=1}^{n} ax_i = ax_1 + ax_2 + ax_3 + \cdots + ax_n$

$= a(x_1 + x_2 + x_3 + \cdots + x_n)$

$= a\sum_{i=1}^{n} x_i$ //

2. $\sum_{i=1}^{n} (x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$

$\sum_{i=1}^{n} (x_i + y_i) = x_1 + y_1 + x_2 + y_2 + x_3 + y_3 \cdots + x_n + y_n$

$= (x_1 + x_2 + x_3 + \cdots + x_n) + (y_1 + y_2 + y_3 + \cdots + y_n)$

$= \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$ //

3. $\sum_{i=1}^{n} b = n \times b$

$\sum_{i=1}^{n} b = b + b + b + b + \cdots b$ (n times)

$= n \times b$ //

4. $\sum_{i=1}^{n} (a + bx_i + cy_i)^2 = n.a^2 + b^2\sum_{i=1}^{n} x_i^2 + c^2\sum_{i=1}^{n} y_i^2 + 2ab\sum_{i=1}^{n} x_i + 2ac\sum_{i=1}^{n} y_i + 2bc\sum_{i=1}^{n} x_i y_i$

$\sum_{i=1}^{n} (a + bx_i + cy_i)^2 = (a + bx_1 + cy_1)^2 + (a + bx_2 + cy_2)^2 + \cdots + (a + bx_n + cy_n)^2$

$= a^2 + 2abx_1 + 2acy_1 + b^2x_1^2 + 2bcx_1y_1 + c^2y_1^2$

$+ a^2 + 2abx_2 + 2acy_2 + b^2x_2^2 + 2bcx_2y_2 + c^2y_2^2$

$+ \cdots$

$+ a^2 + 2abx_n + 2acy_n + b^2x_n^2 + 2bcx_ny_n + c^2y_n^2$

$= na^2 + \sum_{i=1}^{n} 2abx_i + \sum_{i=1}^{n} 2acy_i + \sum_{i=1}^{n} b^2x_i^2 + \sum_{i=1}^{n} 2bcx_iy_i + \sum_{i=1}^{n} c^2y_i^2$

$= na^2 + b^2\sum_{i=1}^{n} x_i^2 + c^2\sum_{i=1}^{n} y^2 + 2ab\sum_{i=1}^{n} x_i + 2ac\sum_{i=1}^{n} y_i + 2bc\sum x_iy_i$

//

**QUESTION 3**

After drawing a number between 1-10 from the bag, let's form two groups of 5 people each according to whether it is odd or even. Let the first group study three hours more per person than the other group. After both groups take the same midterm exam, the results can be analyzed and the effect of extra studying on the score can be examined. However, there are some obstacles to actually implementing this experiment, such as: students' current state of anxiety, lack of background knowledge, differences in the sources studied, intelligence levels, etc.