ORCA 2500 Foundations of Data Science

HW 1

Yihang Ding, yd2459

Yd2459@columbia.edu

Problem 1

A study at UCLA investigated factors that might result in greater attention to the health and education of girls in rural India. One such factor is information about job opportunities for women. The idea is that if people know that educated women can get good jobs, they might take more care of the health and education of girls in their families, as an investment in the girls future potential as earners. The study focused on 160 villages outside the capital of India, all with little access to information about call centers and similar organizations that offer job opportunities to women. In 80 of the villages chosen at random, recruiters visited the village, described the opportunities, recruited women who had some English language proficiency and experience with computers, and provided ongoing support free of charge for three years. In the other 80 villages, no recruiters visited and no other intervention was made. At the end of the study period, the researchers recorded data about the school attendance and health of the children in the villages.

1. Did this analysis have a treatment group and a control group? If so, describe the two groups.

   Yes, in the first 80 villages, recruiters visited the village and recruited relatively educated women, which may make villagers realize the importance of education. Hence it is a treatment group. Whereas in the rest of the villages nothing was done, hence it is a control group.

2. Was this an observational study or a randomized controlled experiment?

   It was a controlled experiment.

3. The study reported (Jensen, Department of Public Policy at UCLA, 2010), Girls aged 5-15 in villages that received the recruiting services were 3 to 5 percentage points more likely to be in school and experienced an increase in Body Mass Index, reflecting greater nutrition and/or medical care. However, there was no net gain in height. For boys, there was no change in any of these measures. Why do you think the author points out the lack of change in the boys?

   Because during the experiment, in the first 80 villages only girls were recruited, there are no direct measure taken on boys, thus by pointing out the boys' situation, it shows the attitude of villagers.

Problem 2

An ad for ADT Security Systems says, when you go on vacation, burglars go to work ... According to FBI statistics, over 25% of home burglaries occur between Memorial Day and Labor Day. Do the data in the ad support the claim that burglars are more likely to go to work during the summer vacation period than at other times?

The number of days between Memorial Day to Labor Day is 99, which is 99/365*100%=27.1% of the year. As the value >25%, there's no increase of home burglaries during the time, thus the data does not support the ad.

Problem 3

Myopia, or nearsightedness, results from a number of genetic and environmental factors. In 1999, Quinn et al studied the relation between myopia and ambient lighting at night (for example, from nightlights or room lights) during childhood.

1. The data were gathered by the following procedure, reported in the study. Between January and June 1998, parents of children aged 2-16 years ... that were seen as outpatients in a university pediatric ophthalmology clinic completed a questionnaire on the childs light exposure both at present and before the age of 2 years. Was this study observational, or was it a controlled experiment?

   It's observational.

2. The study found that of the children who slept with a room light on before the age of 2, 55% were myopic. Of the children who slept with a night light on before the age of 2, 34% were myopic. Of the children who slept in the dark before the age of 2, 10% were myopic. The study concluded that, The prevalence of myopia ... during childhood was strongly associated with ambient light exposure during sleep at night in the first two years after birth. Do the data support this statement? You may interpret strongly in any reasonable qualitative way.

   Yes, as the myopic rate changed significantly as the light ambient condition changes. Though there might be more factors that affect the myopic rate, but at least there is a association between light ambient and myopic

3. On May 13, 1999, CNN reported the results of this study under the head- line, Night light may lead to nearsightedness. Does the conclusion of the study claim that night light causes nearsightedness?

   No, the conclusion may only claim there's a association between night light and nearsightedness. Light ambient may not necessarily be the cause.

4. The final paragraph of the CNN report said that several eye specialists had pointed out that the study should have accounted for heredity. Myopia is passed down from parents to children. In what way do you think this fact might have affected the data?

   For families that have hereditary myopia, there might be a preference in the light ambient. For instance, if these families are more likely to keep lights open during the night, there will be a significant increase in the proportion of myopia, and thus the result is affected.