

1. In a study to assess the effects of the ingestion of methylmercury in fish, Skerfving et al (1974, *Environmental Research*) compared a group of 23 subjects who had eaten large quantities (at least 3 meals per week for at least 3 years) of fish contaminated with methylmercury, with a group of 16 subjects who did not regularly consume contaminated fish and who ate far less fish of all kinds. They found that mercury levels in the blood of the first group were on average much higher than those for the second group.
  - (a) What kind of study is this? Be as specific as you can and support your answer.  
**Solution:** *It is a cohort study, because,*
    - i. *Its goal is to demonstrate cause-effect.*
    - ii. *No treatments are imposed on experimental units (hence it is not a controlled experiment).*
    - iii. *The groups for comparison (cohorts) are formed on the basis of the presumed causal effect.*
  - (b) From this study, can we deduce that ingestion of methylmercury causes elevated blood levels of mercury?  
**Solution:** *No, since this is not a controlled experiment.*
2. Suppose the study in question 1 had been conducted as follows: two groups were formed based on blood mercury levels (low, high), each subject was questioned concerning his/her diet over the past three years, and methylmercury intake estimated and compared for the two groups.
  - (a) What kind of study would it be?  
**Solution:** *Only item iii in the solution to 1(a) would change, with the two groups being constructed based on levels of the response (blood mercury level). This means it is a case-referent study, with cases being the high blood mercury group and referents the low blood mercury group.*
  - (b) From this study, can we deduce that ingestion of methylmercury causes elevated blood levels of mercury?  
**Solution:** *No, since this is not a controlled experiment.*
3. Suppose the study in question 1 had been conducted as follows: Volunteers were randomly divided into two groups. Each group was assigned a diet and was carefully monitored for adherence. The diets were identical except that the diet for one group included a substantial amount of types of fish contaminated with methylmercury while the diet for the other excluded these sources of methylmercury. The subjects were followed for three years, and their blood mercury levels compared at the end of that time.
  - (a) What kind of study would it be?  
**Solution:** *Since treatments (methylmercury in fish or not) were imposed on experimental units (subjects) to observe a response (blood mercury levels), this is a controlled experiment.*
  - (b) From this study, can we deduce that ingestion of methylmercury causes elevated blood levels of mercury?  
**Solution:** *Yes, if the experiment is conducted properly.*
4. Figure 1 shows a frequency histogram and summary statistics for a set of data.
  - (a) Verbally summarize the pattern of variation.  
**Solution:** *The data are unimodal and skewed right.*
  - (b) Choose an appropriate measure of location and scale for these data and give the value and an interpretation of each.  
**Solution:** *For location, the median, 0.6726 (preferred) or the mode, 0.25. For scale, the IQR, 1.0. Interpretations: half the data at or below the median, 0.6726, half at or above it; greatest concentration of data are near the mode, 0.25; IQR is the range of the middle half of the data.*
5. The table below displays a set of data taken over time.

|      |     |     |     |     |     |     |     |     |     |     |     |     |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Time | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  |
| Y    | 1.1 | 3.5 | 2.5 | 2.4 | 4.8 | 3.9 | 4.1 | 6.7 | 5.6 | 5.5 | 7.9 | 7.0 |

(a) Construct a plot appropriate to check for stationarity. What pattern(s) does it reveal?

**Solution:** See the line plot in Figure 2. This shows a linear trend and a three-term cycle about the trend.

(b) What is the appropriate moving average to estimate the trend? Use it to compute estimates of the trend for time values 4, 5 and 6.

**Solution:** A three term moving average is appropriate. Its value at time 4 is  $(2.4+2.5+3.5)/3=2.80$ , and the values at times 5 and 6 are 3.23 and 3.70, respectively.

| Moments  |          |          |          |
|----------|----------|----------|----------|
| N        | 105.0000 | Sum Wgts | 105.0000 |
| Mean     | 0.9183   | Sum      | 96.4175  |
| Std Dev  | 0.9591   | Variance | 0.9198   |
| Skewness | 1.9167   | Kurtosis | 4.3221   |
| USS      | 184.1987 | CSS      | 95.6623  |
| CU       | 104.4449 | Std Mean | 0.0936   |

| Quantiles |        |       |        |
|-----------|--------|-------|--------|
| 100% Max  | 5.0985 | 99.0% | 4.0354 |
| 75% Q3    | 1.2068 | 97.5% | 3.7144 |
| 50% Med   | 0.6726 | 95.0% | 3.2815 |
| 25% Q1    | 0.2067 | 90.0% | 2.0610 |
| 0% Min    | 0.0010 | 10.0% | 0.0869 |
| Range     | 5.0975 | 5.0%  | 0.0478 |
| Q3-Q1     | 1.0000 | 2.5%  | 0.0156 |
| Mode      | .      | 1.0%  | 0.0087 |

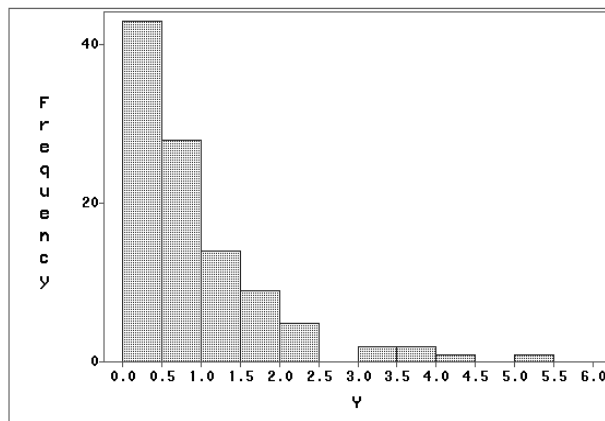


Figure 1: Frequency histogram and summary statistics for question 4.

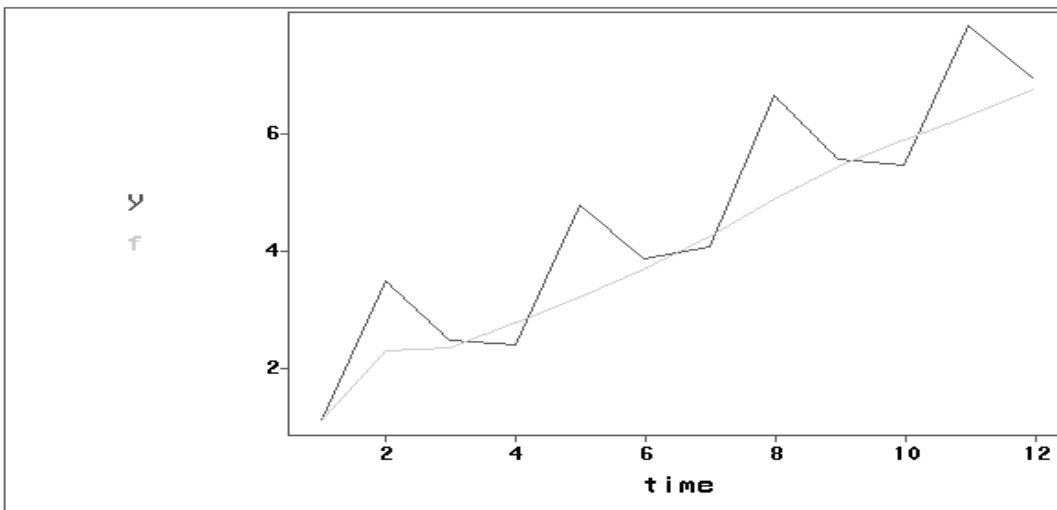


Figure 2: *Line plot of data and three term moving average, question 5.*