

Homework 1 (convert scores to 100)  
Suggested Solutions

Problem 1

- For each study, describe what treatment is being compared and what response is being measured to compare the treatments.

Study 1: Treatment: hair color (1 point)  
Response: pain tolerance level (1 point)

Study 2: Treatment: age (1 point)  
Response: finger-print quality (1 point)

Study 3: Treatment: use of new diagnostic machine (1 point)  
Response: service time (1 point)

Study 4: Treatment: gender (1 point)  
Response: exam result (1 point)

- Which of the studies would be described as experiments and which would be described as observational studies?

Study 1: Observational study (1 point)  
Study 2: Observational study (1 point)  
Study 3: Experiment (1 point)  
Study 4: Observational study (1 point)

- For the studies that are observational, could an experiment have been carried out instead? If not, briefly explain why not.

Study 1: No. We have no control over which color we assign to each subject (assuming we are talking about natural hair color). ( 2 points, 1 for No and 1 for reason)

Study 2: No. Can't control people's age. But a longitudinal study of the same individuals tracked over time may provide a designed experiment for effects of aging on fingerprint quality.

Study 4: No. We have no control over which gender we assign to each student. ( 2 points, 1 for No and 1 for reason)

- For the studies that are experiments, briefly discuss what forms of blinding would be possible to be used.

Study 3: possible answers (1 point)

a)The person analyzing the data is given a list of service times for each car, but does not know which list corresponds to each service method ( i.e. using new diagnostic machine or not using new diagnostic machine). OR

b) The mechanic does not know whether they are using the new diagnostic machine or the old machine. The machines could have been disguised so that they both look the same.

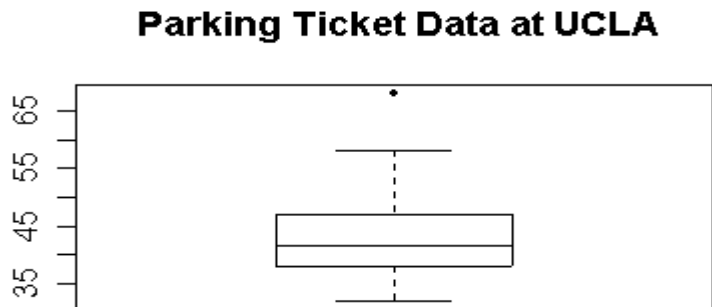
• In which of the studies has blocking been used? Briefly describe what was blocked and why it was blocked.

Study 3: Blocking by mechanic has been used in this study. This was done because the difference in work experience between mechanics could also have an effect on the time for service. ( 2 points, 1 point for what was blocked and 1 point for why)

• Calculate the five number summary for the data (min, lower quartile, median, upper quartile, and max).

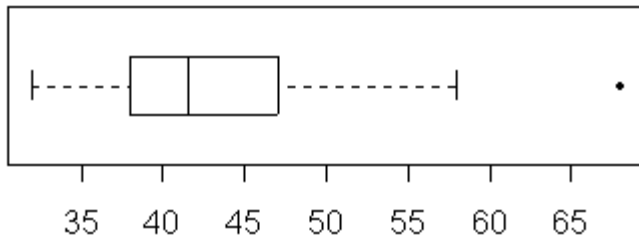
- |                    |            |
|--------------------|------------|
| Minimum: 32        | ( 1 point) |
| Lower Quartile: 38 | ( 1 point) |
| Median: 41.5       | ( 1 point) |
| Upper Quartile: 47 | ( 1 point) |
| Maximum: 68        | ( 1 point) |

• Draw a box plot for this set of data. Show your working. ( 8 points, 1 point for labeling axis correctly( either vertical or horizontal and the numbers could be labeled differently), 1 point for minimum, 1 point for Q1, 1 point for median, 1 point for Q3, 1 point for  $Q3+1.5(IQR)$  position, 1 point for the outlier, 1 point for drawing the plot correctly)



**OR:**

## Parking Ticket Data at UCLA



- Using your plots, in plain English, briefly comment on the data and the story the above summary/plots convey. ( 6 points, students may have slightly different answers)

From the analysis, it can be seen that the number of parking tickets given out at UCLA over a period of 30 days is about 42 per day, since the median is 41.5. The highest number of parking tickets given out on a single day during the period was 68, and the least number was 32. Most days had between 38 and 47 tickets given out.

1.  $P(H_B) = 11\%$
2.  $P(H_A \text{ and } W_A) = P(H_A) * P(W_A) = 0.3 * 0.3 = 0.09$
3.  $P(H_{AB} \cup W_{AB}) = P(H_{AB}) + P(W_{AB}) - P(H_{AB} \text{ and } W_{AB}) = 0.1 + 0.1 - 0.1^2 = 0.19$
4.  $P(H_A \text{ and } W_A) + P(H_B \text{ and } W_B) + P(H_o \text{ and } W_o) + (H_{AB} \text{ and } W_{AB}) = 0.3 * 0.3 + 0.11 * 0.11 + 0.49 * 0.49 + 0.01 = 0.3522$

- Suppose that  $X \sim \text{Normal}(\mu=3, \sigma^2=16)$  compute the Z-scores for the following numbers and state how many SD's each of these numbers is away from  $\mu$ :
  - $-5 \quad Z = (-5-3)/4 = -2$
  - $11 \quad Z = (11-3)/4 = 2$
  - $5 \quad Z = (5-3)/4 = 0.5$
  - $1.4 \quad Z = (1.4-3)/4 = -0.4$
  - What is the probability  $P(-3 \leq X \leq -1)$ ? Is it different from  $P(-3 < X < -1)$ ? Why? Same Probabilities!  $Z_1 = (-3-3)/4 = -1.5$ ;  $Z_2 = (-1-3)/4 = -1$

$$P(-3 < X < -1) = (0.4332 - 0.3413) = 0.09$$


---