**There are 6 questions (80 points).**

* ***No collaboration is allowed; giving or receiving any unauthorized assistance will be treated as cheating and will automatically get 0 for this test.***
* ***You may not disclose the content of the exam to anyone else.***
* Your solutions should be compiled in a word file, and be submitted to Canvas by the due time.
* For multiple choice questions, there is only one correct answer.
* You may receive partial credits by showing your work even if your final answer is incorrect. You may want to submit your Excel file as well.

The data file **ED\_patients.xlsx** includes the patient visit information to a local emergency department (ED) in early 2015. Each row contains the visit information of a patient. For each patient, we have the following information:

--- Visit ID, age, arrival time to the ED;

--- Severity group: a number from 1 to 5, where 1 indicates most urgent and 5 indicates least urgent;

--- Disposition: Admitted – the patient is admitted to the hospital; Discharged – the patient is discharged home;

--- Waiting time (in minutes): how long did the patient wait in the ED before being seen by a doctor.

**Q1. Analyze patient waiting time.**

(1) Make a histogram for the patient waiting time. Starting from the minimum waiting time, use every 20 minutes to generate bins, until the maximum waiting time is covered. (10 pts)

(2) The hospital management team is interested to know if there is any evidence that the average patient waiting time is greater than 98 minutes. Set the alternative hypothesis to be *H*a: ** 98. Report the value of the test statistics (keep 3 decimal places) and the p-value of the test (keep 3 decimal places). (10 pts)

**Q2. Compare patient waiting time.** We are interested to learn if the waiting times of admitted patients are shorter than that of discharged patients.

(1) Which of the following is the appropriate procedure to compare the performance between the two sales forces? (5 pts)

A. One sample z-test

B. One sample t-test

C. Independent two-sample t-test

D. Matched-pairs t-test

E. Chi-squared test of independence

F. Chi-squared test of goodness of fit

G. One-way ANOVA

(2) What is the p-value for the hypothesis test ? Here, is the average waiting time of admitted patients, and is the average waiting time of discharged patients. (10 pts)

**Q3.** We are interested to know if there is any difference in the mean waiting times of patients of severity groups 2, 3, and 4.

(1) (5 pts) To study whether patients waiting times depend on their severity groups, what is the appropriate model to be used? \_\_\_\_\_\_\_\_\_\_\_

A. One sample z-test

B. One sample t-test

C. Independent two-sample t-test

D. Matched-pairs t-test

E. Chi-squared test of independence

F. Chi-squared test of goodness of fit

G. One-way ANOVA

(2) State the null hypothesis of the test. (5 pts)

(3) What is the value of the test statistic (*hint: the value to left of the p-value in your Excel output*)? What is the p-value of the test (report your Excel output)? (5 pts)

**Q4.** Patients arrive to the ED during different hours of the day (the Excel function **hour()** can help us obtain the hour of day of a patient’s ED visit time). Because patients of different severity groups bring different levels of workload to the ED (e.g., a patient of severity group 1 is much harder to treat than a patient of severity group 5), we want to learn if there is any association between severity groups and hour of day. Such information can help ED management to make better staffing decisions.

(1) What is the null hypothesis for testing if there is any dependence between patient severity groups and hour of day? (5 pts)

(2) What is the p-value of the test (report your Excel output)? At a significance level of 0.01, what is the conclusion of the test? (15 pts)

**Q5.** Do you think that patients visit the ED with equal chance among different days of the week (from Monday to Sunday)? Use a hypothesis test to answer the question. State the null and alternative hypotheses, and calculate the test statistic and the p-value. The Excel function **weekday()** can help us obtain the day of week of a patient’s ED visit time. (10 pts)