1. Compare the data structures "array" and "linked list" (e.g., their advantages and disadvantages). (6%)

2. Write a recursive program to calculate \( n! \). (15%)

3. (1) What is a stack? What is a queue? (8%)
(2) Write the pseudo code of inserting/deleting an item to/from a circular queue. (10%)

4. (15%)
Write an algorithm to count the number of internal nodes in a binary tree \( T \).

5. (1) Use adjacency matrix and adjacency list to represent the graph in Figure 1. (6%)
(2) Perform DFS (Depth First Search) and BFS (Breadth First Search) to search the graph in Figure 1 (8%)

![Graph Diagram]

6. (1) What are the average time and worst time to sort \( n \) objects using the following algorithms? (a) bubble sort; (b) quick sort; (c) heap sort; (d) merge sort (8%)
(2) Quicksort is an unstable sorting method. Use quicksort to determine the order of the records with keys 36, 21, 15, 38, 52, 15, 6, 48, and explain why this sorting method is unstable. (8%)
(3) What kind of data structure you would use for the heap in the heapsort program? Explain why. (4%)

7. (1) What is a hash function? (3%)
(2) Use two different ways to design a hash function. (3%)
(3) What is hash collision? (3%)
(4) How to solve the collision? (3%)
1. Each question is worth 5 points, totaling 30 points.

The owner of a restaurant serving Continental-style entrées was interested in studying ordering patterns of patrons for the Friday-to-Sunday weekend time period. Records of demand for dessert during the time period were maintained, along with two other variables: the gender of patrons and whether a beef entrée was ordered. The results are as follows.

<table>
<thead>
<tr>
<th>Dessert Ordered</th>
<th>Gender</th>
<th>Beef Entrée</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Yes</td>
<td>96</td>
<td>40</td>
</tr>
<tr>
<td>No</td>
<td>224</td>
<td>240</td>
</tr>
<tr>
<td>Total</td>
<td>320</td>
<td>280</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Beef Entrée</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>65</td>
<td>71</td>
<td>136</td>
</tr>
<tr>
<td>No</td>
<td>348</td>
<td>116</td>
<td>464</td>
</tr>
</tbody>
</table>

(1) What is the probability that a patron who does not order a beef entrée orders a dessert?

(2) Is ordering a beef entrée independent of whether the patron orders dessert?

(3) Consider a random experiment based on the observed data of three characteristics: dessert ordering, gender and beef entrée ordering. How does the outcome space look like? How many elements are included in the outcome space?

(4) Define a random variable D by converting the elements in the outcome space with 1 being dessert ordered and 0 for not ordered. What is the probability distribution of D? What are the mean and variance of D?

(5) Define another random variable G by converting the elements in the outcome space with 1 being male and 0 for female. Find the covariance between D and G. What can you suggest to the restaurant owner?

Define another random variable B by converting the elements in the outcome space with 1 being beef entrée ordered and 0 for not ordered. Find the covariance between D and B. What can you suggest to the restaurant owner?
2. 每小題 10 分，共 40 分

For each of the following hypothesis tests, identify the hypotheses, the type of test, the test statistic in use, the rejection region, and the assumptions behind each test.

(1) Nine experts rated four brands of Colombian coffee in a taste-testing experiment. A rating on a 7-point scale (1 = extremely unpleasing, 7 = extremely pleasing) is given for each of four characteristics: taste, aroma, richness, and acidity, and the summated rating, accumulated over all four characteristics represent the quality of a certain brand of Colombian coffee.

(2) A pet food company is desired to know which product line, kidney- or shrimp-based cat foods, is of better quality. An experiment is conducted to compare the two products with a sample of 20 cats selected from the population at a local animal shelter. Ten cats were randomly assigned to each of the products being tested. Each of the cats was presented with 3 ounces of the selected food in a dish at feeding time. The researchers defined the variable to be measured as the number of ounces of food consumed within a 10-minute time interval right after the filled dish was presented.

(3) Shipments of meat, meat by-products, and other ingredients are mixed together in several filling lines at a pet food canning factory. Operations manager suspect that although the mean amount filled per can of pet food is usually stable, the variability of the cans filled in line A is greater than that of line B. A random sample of 10 8-ounce cans from line A and 12 8-ounce cans from line B is collected accordingly.

(4) A bank branch located in a commercial district of a city had the business objective of improving the process for serving customers during the noon-to-1:00 P.M. lunch period. To do so, the waiting time (defined as the time the customer enters the line until he or she reaches the teller window) needs to be shortened to increase customer satisfaction. The previous standard waiting time is 5 minutes. A random sample of 15 customers during this hour is recorded over a period of a week.
3. 每小題 6 分，共 30 分

An investigation of ethnic differences in reports of pain perception was presented. A sample of 55 blacks and 159 whites participated in the study. Subjects rate (on a 13-point scale) the intensity and unpleasantness of pain felt when a bag of ice was placed on their foreheads for two minutes. The mean pain intensity for blacks was 8.0 and for whites, 6.8.

(1) Why is it dangerous to draw a statistical inference given the above information only?

(2) Give possible values of the missing sample standard deviations that would lead you to conclude (at $\alpha = 0.05$) that blacks, on average, can stand for a pain intensity rating more than 7.5.

(3) What assumption(s) do you need to perform test in (2)?

(4) Give possible values of the missing sample standard deviations that would lead you to conclude (at $\alpha = 0.05$) that blacks, on average, have a higher mean intensity rating than whites.

(5) Give possible values of the missing sample standard deviations that would lead you to an inconclusive decision (at $\alpha = 0.05$) regarding whether blacks or whites have a higher mean intensity rating.