NAME: _____

*Instructions: Write your answers clearly on the exam paper. Show work to receive partial credit. Use the back of the page if you need more space but please note that you have done so. The exam is worth a total of **80 points**.*

1. [*20 points*] Shown below is a histogram of the carbon dioxide (CO_2) emissions per person for 48 countries with populations of at least 20 million. The questions that follow refer to this histogram.
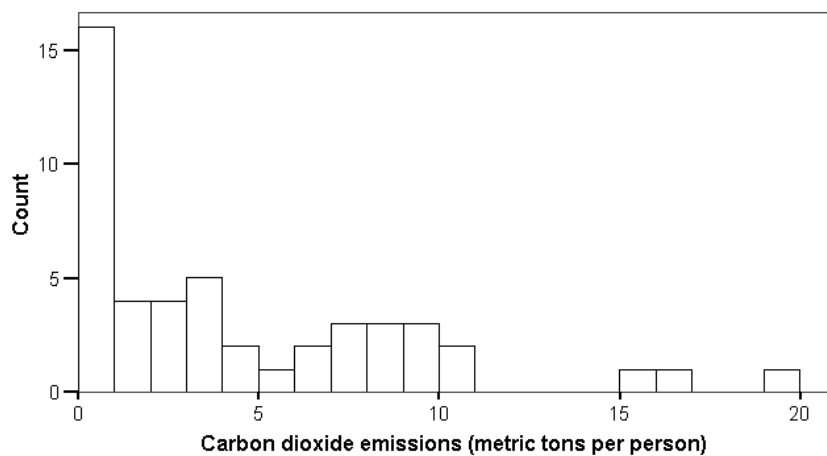
(a) Is the variable CO_2 emissions continuous or discrete? Explain with reference to the histogram.

(b) How many countries have CO_2 emissions greater than 12 metric tons per person?

(Problem 1. continued)

(c) Describe the shape of the histogram and interpret it in context.

(d) Because you do not have the detailed data, you cannot determine the 5-number summary exactly. You should, however, be able to estimate the Q_1 and Q_3 to within 1 metric ton. Give your estimates and a brief justification.

(e) In the space below, provide a rough sketch of the normal probability plot (QQ plot) for this data set. Your plot need not be very detailed but you should explain why the shape of your plot is consistent with the shape of the histogram.

2. [*15 points*] The number of mutants in a random sample of 8 organisms is a random variable X that takes values 0, 1, 2, 4, 8 with probabilities as shown in the table below.

x	0	1	2	4	8
Probability	p	0.5	0.25	0.125	0.0625

(a) What must be the value of p for this to be a valid probability distribution?

(b) Find the mean μ_X and variance σ_X^2.

(c) The number out of n samples of 8 organisms that have no mutants can be approximated by a Normal distribution with mean np and variance $np(1-p)$. Suppose 200 samples of 8 organisms are drawn. Find the approximate probability that no more than 10 of these 200 samples have zero mutants.

3. [*20 points*] Too much cholesterol in the blood increases the risk of heart disease. Young women are generally less afflicted with high cholesterol than other groups. The cholesterol levels for women aged 20 to 34 follow an approximately Normal distribution with mean 185 milligrams per deciliter (mg/dl) and standard deviation 39 mg/dl.

(a) Cholesterol levels above 240 mg/dl demand medical attention. What percent of young women have levels above 240 mg/dl?

(b) In a random sample of 10 women aged 20 to 34, what is the probability that exactly one has a cholesterol level that demands medical attention, i.e., is above 240 mg/dl?

(c) Fill in the following statement: 90% of young women aged 20 to 34 have cholesterol levels between _____ and _____ mg/dl. (Consider the middle 90% of the distribution.)

(d) Middle-aged men 55 to 64 have cholesterol levels that follow an approximately Normal distribution with mean 222 mg/dl and standard deviation 37 mg/dl. What proportion of middle-aged men have cholesterol levels above the mean level for women 20 to 34?

4. [*10 points*] Intense or repetitive sun exposure can lead to skin cancer. Patterns of sun exposure, however, differ between men and women. A study of cutaneous malignant myeloma in the Italian population found that 15% of skin cancers are located on the head and neck, another 41% on the trunk, and the remaining 44% on the limbs. Moreover, 56% of those with a skin cancer on the head are women, as are 37% of those with a skin cancer on the trunk but 80% of those with a skin cancer on the limbs.

(a) What percent of all individuals with skin cancer are women? Draw a probability tree to support your answer to this question.

(b) A woman is found with skin cancer. What is the chance that the cancer is located on her limbs?

5. [*15 points*] Some researchers believe autism may be marked by different brain growth patterns in early life, even before the diagnosis is made. Studies have linked brain size in infants and toddlers to a number of future ailments, including autism. One study looked at the brain sizes of 30 autistic boys and 12 nonautistic boys (control) who had received an MRI scan as toddlers. The table below provides their whole-brain volumes in cubic milliliters:

Autistic									
1311	1250	1292	1419	1401	1297	1202	1336	1308	1353
1515	1461	1365	1364	1362	1303	1278	1247	1333	1340
1319	1286	1223	1241	1229	1209	1171	1154	1128	1230
Control									
1040	1180	1207	1179	1115	1133	1298	1263	1194	1198
1230	1114								

Put yourself in the role of the statistical analyst for this study. Use the supporting statistics and plots on the following page to write a paragraph or two addressing the following questions. You many continue on the back of this page.

- Do autistic boys tend to have larger brains on average than nonautistic boys during the toddler years?

- Can an MRI scan be used with great confidence to determine whether a toddler will develop autism?

- Do the MRI scans need to be transformed (say with square root or log transformation) before they can be used in statistical tests that require measurements have a Normal distribution?

Supporting Information for Question 5

Volume in cubic ml

Group	Mean	N	Std. Deviation
Autistic	1297.57	30	88.42
Control	1179.25	12	70.68
Total	1263.76	42	98.97

```
      Control     |    | Autistic
                4 | 10 |
                  | 10 |
          3 1 1   | 11 | 2
        9 9 8 7   | 11 | 5 7
            3 0   | 12 | 0 0 2 2 3 4 4
            9 6   | 12 | 5 7 8 9 9
                  | 13 | 0 0 1 1 3 3 4
                  | 13 | 5 6 6 6
                  | 14 | 0 1
                  | 14 | 6
                  | 15 | 1
```


Normal Q-Q Plot of Volume in cubic ml — Group: Autistic

Normal Q-Q Plot of Volume in cubic ml — Group: Control