Instructions: Answer the following questions. Make and state your own assumptions for questions where the information is not sufficient for you to solve them. For example, if you need the corresponding p-value of a normally distributed random variable evaluated at 2.5, you may indicate the value as, for example, $Pr(x \geq 2.5)$, where $x \sim \mathcal{N}(0, 1)$.

1. (15 points) Comment the statement: "Two mutually exclusive events may not be independent." You may draw some diagrams to explain your answer.

2. (20 points) A bus starts with 6 people and stops at 10 different stops. Assuming that passengers are equally likely to depart (that is, get off) at any stop, find the probability that no two passengers leave at the same bus stop. That is, find the probability that at each stop at most only one passenger gets off the bus.

3. (25 points) A random variable x has an exponential probability probability distribution with density function given as:

$$f(x) = ce^{-25x} \qquad x \leq 0$$

where c is an unknown constant.

   (a) (10 points) Find the value of c so that $f(x)$ is a pdf.

   (b) (5 points) Find the mean μ and standard deviation σ of x.

   (c) (10 points) Find the probability that x will fall within $(\mu - 2\sigma, \mu + 2\sigma)$.

4. (40 points) From Statistics course(s), we learned the following:
Let $x_1, ..., x_N$ be the outcomes of a random variable x with probabilities $p(x_1), ..., p(x_N)$, respectively. Then the mean and variance of x are:

$$\mu = \sum_{i=1}^{N} x_i p(x_i) \qquad (1)$$

$$\sigma^2 = \sum_{i=1}^{N} (x_i - \mu)^2 p(x_i) \qquad (2)$$

When $p(x_i) = \frac{1}{N}, \forall \ i = 1, ..., N$, (1) and (2) are:

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i \qquad (3)$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2 \qquad (4)$$

We also learned that if $(x_1, \ldots, x_N)$ represents an iid random sample of x, then the estimates of mean and variance are:

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{5}$$

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \hat{\mu})^2 \tag{6}$$

(a) (5 points) Why do $p(x_i)$'s appear in (1) and (2), but not in (5) and (6)?

(b) (10 points) Prove that $\hat{\mu}$ is unbiased.

(c) (10 points) When $p(x_i) = \frac{1}{N}$, $\forall \ i = 1, \ldots, N$, then (3) and (5) look the same. So, do we have to assume in (5) and (6) that the sample $(x_1, \ldots, x_N)$ is drawn from a distribution that assigns equal probability to each of the possible outcomes?

(d) (15 points) In (6), the sum of squares is divided by $N-1$, but in (4) the sum of squares is divided by N. A statistician says that, in (4), we know the population, so the variance is unbiased by dividing the sum of squares by N. But in (6), the population is unknown, so one has to divide the sum of squares by $N-1$ to get an unbiased estimate. Also, he says that "as long as the sample size is larger than the number of possible outcomes (that is, the population size), then one can make inferences about the paramters of the random variable." Is he correct? What do you think about his statement? Anything wrong? Please comment and explain.