1. Suppose 5 different hypothesis tests have been conducted, with p-values of: Test 1 ($p = 0.07$), Test 2 ($p = 0.001$), Test 3 ($p = 0.015$), Test 4 ($p = 0.005$), Test 5 ($p = 0.05$).

   (a) Using the Bonferroni correction, which tests should be accepted or rejected when the family-wise error rate is $\alpha = 0.05$.

   (b) Using the HolmBonferroni method, which tests should be accepted or rejected when the family-wise error rate is $\alpha = 0.05$.

Solutions:

(a) Since there are five tests, the Bonferroni correction states that a hypothesis should be rejected if $p < \alpha/5 = 0.01$. Thus, Tests 2 and 4 should be rejected and Tests 1, 3, and 5 should be accepted.

(b) We begin by arranging the p-values in increasing order: 0.001, 0.005, 0.015, 0.05, 0.07. We need to determine the smallest k such that the k-th p-value in the arranged list is greater than $Q_k = \alpha/(5 + 1k)$. For $k = 1, \ldots, 5$, the rounded values of Q_k are 0.01, 0.0125, 0.017, 0.025, and 0.05. In this case, $k = 4$ is that smallest k. As a result, we reject hypothesis corresponding to the first three p-values in the ordered list and accept the remaining. Thus, Tests 2, 3, and 4 should be rejected and Tests 1 and 5 should be accepted

2. Suppose would like to model sunlight using a discrete time Markov chain. In particular, suppose we have the following data

	sunny tomorrow (1)	**cloudy tomorrow (2)**
sunny today (1)	86	23
cloudy today (2)	18	88

The number in parenthesis are the labeling of the vertices in the corresponding discrete time Markov chain. Find the maximum likelihood estimates of the transition probabilities.

Solutions:

$$\begin{aligned}
p_{1|1} &= 86/(86+23) = 0.79 \\
p_{2|1} &= 23/(86+23) = 0.21 \\
p_{1|2} &= 18/(18+88) = 0.17 \\
p_{2|2} &= 88/(18+88) = 0.83
\end{aligned}$$

3. Assume that items produced are supposed to be normally distributed with mean 3 and standard deviation 2. To monitor this process, subgroups of size 10 are sampled. If the following represents the averages of the first 10 subgroups, does it appear that the process was in control with significance level 0.05?

Solutions:

Subgroup No.	$\bar{X}$
1	3.4
2	4.1
3	1.9
4	3.7
5	3.7
6	3.1
7	2.9
8	1.8
9	4.3
10	4.3

Solutions:

When in control, $X \sim N(3, 2)$. n = 10.

$$LCL = \mu - \sigma \cdot z_{(1-\alpha/2)}/\sqrt{n} = 3 - 2 \cdot 1.96/\sqrt{10} = 1.76$$
$$UCL = \mu + \sigma \cdot z_{(1-\alpha/2)}/\sqrt{n} = 3 + 2 \cdot 1.96/\sqrt{10} = 4.24$$

$\bar{X}^9 > UCL$, so the process is out of control starting with the 9th subgroup.

4. Suppose $X_1 \sim \mathcal{N}(\mu, 1)$ and $X_2 \sim \mathcal{N}(\mu, 2)$ are independent. Using X_1, X_2, construct a tight level 0.95 confidence interval for μ.

Solutions:

Using the formula for the unbiased estimate of μ that minimizes the expected squared loss gives $\hat{\mu} = 2X_1/3 + X_2/3$. From the properties of Gaussian distributions, we have $\hat{\mu} \sim \mathcal{N}(\mu, \frac{2}{3})$. Thus, we have

$$\underline{\mu} = \hat{\mu} - \sigma \cdot z_{(1-\alpha/2)} = 2X_1/3 + X_2/3 - \frac{2}{3} \cdot 1.96 = 2X_1/3 + X_2/3 - 1.31$$

$$\overline{\mu} = \hat{\mu} + \sigma \cdot z_{(1-\alpha/2)} = 2X_1/3 + X_2/3 + \frac{2}{3} \cdot 1.96 = 2X_1/3 + X_2/3 + 1.31$$

5. Suppose $X_1, \ldots, X_n$ are iid samples from a distribution with pdf $f(u) = \frac{1}{2}\exp(-|x-\theta|)$. Formulate the MLE for θ as a minimization problem with a summation.

Solutions:

The negative log-likelihood is proportional to $\sum_{i=1}^{n} |x - \theta|$, and so the MLE is

$$\hat{\theta} = \arg\min \sum_{i=1}^{n} |x - \theta|. \tag{1}$$

As an aside, the solution of this optimization problem is $\hat{\theta} = \text{median}(X_i)$, though deriving this solution requires using optimization theory beyond the scope of IEOR 165.