This homework can be completely hand-written, except for 1(c) and 4(f)-4(g). Your answers for the R parts can be completely separate from your hand-written answers for the other problems, if you would like.

1. Let $y_i \sim \text{Binom}(n_i, p)$ for $i = 1, 2, \ldots, m$ be $m$ independent binomial random variables.
   (a) Find the maximum likelihood estimate of $p$ using calculus.
   (b) Given the data in the following table, what is the MLE of $p$?

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $n_i$ | 29 | 53 | 61 | 62 | 51 | 62 | 53 | 49 | 71 | 26 |
| $y_i$ | 18 | 31 | 34 | 33 | 27 | 33 | 28 | 23 | 33 | 12 |

   (c) Use R to plot the likelihood function for a range of $p$ values between zero and one. The "dbinom" function will be useful here. The example code in "MLE.ex.rain.r" shows a similar example with Bernoulli-distributed data.

2. Let $y_i \sim N(\mu, \sigma^2)$ for $i = 1, 2, ..., n$ be $n$ independent and identically-distributed Gaussian random variables. Assume that the variance is known: $\sigma^2 = 4$.
   (a) Find the maximum likeihood estimate of $\mu$. This will be $\hat{\mu}_{ML} = g(y_1, y_1, \ldots, y_n)$ for some function $g(\cdot)$.
   (b) Find $E(\hat{\mu}_{ML})$ and $Var(\hat{\mu}_{ML})$.
   (c) Show that $\hat{\mu}_{ML} = \hat{\mu}_{OLS}$ where $\hat{\mu}_{OLS}$ is the value of $\mu$ that minimizes the sum of squared errors:

$$LS(\mu; \mathbf{y}) = \sum_{i=1}^{n} (y_i - \mu)^2$$

3. Let $\mathbf{y}$ be a $n \times 1$ random vector distributed as

$$\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$$

   where $\mathbf{X}$ is a $n \times p$ matrix of real numbers and $\boldsymbol{\beta}$ is a $p \times 1$ vector of real numbers.
   (a) What is the distribution of $\mathbf{z} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$?
   (b) What is the distribution of $\bar{\mathbf{y}} = \sum_{i=1}^{n} y_i/n$?

4. Let $\mathbf{z} \sim N(\mu\mathbf{1}, \sigma^2\mathbf{I})$ and $\mathbf{w} \sim N(\gamma\mathbf{1}, \mathbf{R})$ be independent multivariate normal distributions of dimension $n$.
   (a) Let $\mathbf{b} = [\mathbf{z}'\mathbf{w}']'$ be a $2n \times 1$ column-vector formed by stacking $\mathbf{z}$ and $\mathbf{w}$ together. What is the distribution of $\mathbf{b}$?
   (b) What is the distribution of $(\mathbf{z} - \mu\mathbf{1})'(\mathbf{z} - \mu\mathbf{1})/\sigma^2$?
   (c) What is the distribution of $\mathbf{y} = \mathbf{z} + \tau \cdot \mathbf{w}$?
   (d) What is the distribution of $(\mathbf{w} - \gamma\mathbf{1})'\mathbf{R}^{-1}(\mathbf{w} - \gamma\mathbf{1})$?
   (e) Let the $i, j$-th element of $\mathbf{R}$ be given by $R_{ij} = \exp(-|i - j|/10)$. What is $Cov(y_i, y_j)$? What is $cor(y_i, y_j)$?
   (f) Do this question in R. Set $\mu = \gamma = 0$ and let $n = 20$. Use "rmvnorm" in the "mvtnorm" package to simulate 9 realizations of $\mathbf{y}$ when $\sigma = \tau = 1$. The following code may be used to make the $\mathbf{R}$ and $\mathbf{I}$ matrices:

```
R=matrix(0,nrow=n,ncol=n)
for(i in 1:n){
  for(j in 1:n){
    R[i,j]=exp(-abs(i-j)/10)
  }
}

I=diag(n)
```

(g) Repeat (e) for $\sigma = 10$ and $\tau = 1$.

(h) Repeat (e) for $\sigma = 1$ and $\tau = 10$.

5. Let $x_1$ and $x_2$ be independent, standard normal random variables. Show that the random variables $y = x_1 + x_2$ and $z = x_1 - x_2$ are independent by showing that their covariance is zero. Note, zero covariance implies independence ONLY for normal random variables.

6. Suppose the value of a stock on Day $i$ is $x_i$. On day $i$ the value of the stock is distributed as follows:

$$x_i = \begin{cases} e_0 & , \quad i = 0 \\ x_{i-1} + e_i & , \quad i > 0 \end{cases} , \quad e_t \overset{iid}{\sim} N(\mu, \sigma^2).$$

(a) What is the expected value of the stock on Day 10?

(b) What is the joint distribution of $(x_0, x_1, x_2, x_3, x_4)'$? Hint: Write $\mathbf{x} = \mathbf{A}\mathbf{e}$ for some matrix $\mathbf{A}$.

(c) What is $Cov(x_1, x_4)$?