Problem 1

Answer the following questions:

a. Consider the random vector \( Y \) with mean vector \( \mu \) and variance covariance matrix \( \Sigma \). If \( a'Y \) follows univariate normal for every vector \( a \) show that \( Y \sim N(\mu, \Sigma) \).

b. Let \((X_i, Y_i), i = 1, 2, \ldots, n\) be a random sample from a bivariate normal distribution (the \( n \) pairs are independent). Consider the vector \( W = \begin{pmatrix} X \\ Y \end{pmatrix} \), where \( X = (X_1, X_2, \ldots, X_n) \) and \( Y = (Y_1, Y_2, \ldots, Y_n) \). Find the distribution of \( W \).

c. Refer to question (b). Find the conditional distribution of \( X \) given \( Y \).

d. Suppose \( \epsilon_0, \epsilon_1, \ldots, \epsilon_n \) are independent \( N(0, \sigma) \) and let \( Y_i = \epsilon_i + c\epsilon_{i-1} \), where \( c \) is a known constant. Show that \( Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \) follows multivariate normal. What is the mean and variance covariance matrix?

e. Let \( X_1, \ldots, X_n \) be i.i.d. random variables with \( X_i \sim N(\mu_1, 1) \) and Let \( Y_1, \ldots, Y_m \) be i.i.d. random variables with \( Y_i \sim N(\mu_2, 1) \). All the random variables are independent. What is the distribution of the expression \( \sum_{i=1}^n (X_i - \mu_1)^2 + \sum_{i=1}^m (Y_i - \mu_2)^2 \). What is the mean and variance of this expression?
Problem 2
Answer the following questions:

a. The multinomial distribution is defined as follows: A sequence of $n$ independent experiments is performed and each experiment can result in one of $r$ possible outcomes with probabilities $p_1, p_2, \ldots, p_r$ with $\sum_{i=1}^{r} p_i = 1$. Let $X_i$ be the number of the $n$ experiments that result in outcome $i$, $i = 1, 2, \ldots, r$. Then, $P(X_1 = x_1, X_2 = x_2, \ldots, X_r = x_r) = \frac{n!}{n_1!n_2!\cdots n_r!} p_1^{x_1} p_2^{x_2} \cdots p_r^{x_r}$. Find the mean and variance covariance matrix of the vector $\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_r \end{pmatrix}$

b. Refer to question (a). Find the variance of $\sum_{i=1}^{r} X_i$.

c. Let $M_{X_i,X_j}(t_i, t_j)$ be the joint moment generating function of $X_i$ and $X_j$. Show that $\frac{\partial^2 M(0,0)}{\partial t_i \partial t_j} - \begin{bmatrix} \frac{\partial M(0,0)}{\partial t_i} \\ \frac{\partial M(0,0)}{\partial t_j} \end{bmatrix} = \text{cov}(X_i, X_j)$.

d. Suppose the random variable $X$ has following pdf: $f(x) = \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $0 < x < \infty$. Find the mean and variance of $X$ without integration (by using a suitable transformation).

e. Suppose $X_1, X_2, X_3$ are independent with $X_1 \sim N(2, 2)$, $X_2 \sim N(5, 5)$, and $X_3 \sim N(4, 4)$. Let $Q = X_1 + 2X_2 - 3X_3 - 4$. Find the distribution of $Q$. 
Problem 3
Answer the following questions:

a. Find the moment generating function of a Bernoulli random variable and use it to find the moment generating function of binomial. (This is a different method for finding the moment generating function of binomial that was discussed in class.)

b. Let $X$ be a random variable with $E[X^m] = (m + 1)!2^m$, with $m = 1, 2, 3, \ldots$. Find the moment generating function of $X$ and its distributions.

c. Let $X \sim \Gamma(\frac{n}{2}, \beta)$. Find the distribution of $Y = \frac{2X}{\beta}$ using the method of cdf and the method of moment generating functions.

d. Let $U$ be random variable with moment-generating function $M_U(t) = e^{500t+5000t^2}$. Find $P(27100 < (U - 500)^2 < 50200)$.

e. Let $X_1, X_2, \cdots, X_n$ be i.i.d. random variables and each one follows the exponential distribution with parameter $\lambda$. Show that $Q = 2\lambda \sum_{i=1}^{n} X_i$ follows the $\chi^2$ distribution. What are the degrees of freedom.
Problem 4
Answer the following questions:

a. Suppose the number of pine trees in a certain forest follows the Poisson distribution with parameter $\lambda$ per meter$^2$. Suppose we randomly select a point (say $A$) in this forest (not a pine tree, just a point). Let $X$ be the distance from this point to the nearest pine tree and let $Y$ be the distance from this point to the second nearest pine tree (see graph below). Find the probability density function of $X$ and then show that the random variable $\lambda \pi X^2$ follows the exponential distribution with mean 1. Note: The parameter $\lambda$ here is given per meter$^2$. The parameter $\lambda$ of a circle with radius $r$ is $\lambda \pi r^2$.

![Diagram of a forest with a closest and second closest pine tree to point A](image)

b. Explain how to compute the probability that the distance until we observe the first pine tree from point $A$ is larger than $c$ meters.

c. Refer to question (a). Find the probability density function of $Y$. ($x$ is fixed when we are considering the pdf of $Y$.) Show that the random variable $\lambda \pi (Y^2 - X^2)$ follows the exponential distribution with mean 1.

d. Suppose now we randomly select $m$ points in this forest. Show that $2 \lambda \pi \sum_{i=1}^{m} X_i^2$ and $2 \lambda \pi \sum_{i=1}^{m} (Y_i^2 - X_i^2)$ follow a gamma distribution. What are the parameters of these distributions?

e. Let $s = \lambda \pi \sum_{i=1}^{m} X_i^2$ and $t = \lambda \pi \sum_{i=1}^{m} (Y_i^2 - X_i^2)$. If $s$ and $t$ are independent show that $\frac{\sum_{i=1}^{m} X_i^2}{\sum_{i=1}^{m} Y_i^2} \sim$ beta($m$, $m$).