1. The Erlang density is given for \( \lambda > 0 \) and positive integer \( n \):

\[
f_X(x) = \begin{cases} 
\frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} & x \geq 0 \\
0 & \text{otherwise}
\end{cases}
\]

When \( n = 1 \) we have the special case of the Exponential density. Let \( Y_n = \sum_{i=1}^{n} X_i \) where the \( X_i \) are independent and Exponentially distributed with parameter \( \lambda \).

(a) Using moment generating functions, show that \( Y_n \) has the Erlang distribution with parameter \( \lambda \) and \( n \).

(b) Use the relationship you showed above to argue that the mean and variance of the Erlang are \( n/\lambda \) and \( n/\lambda^2 \). (Hint: you do not need to do any significant calculations.)

2. Let \( R = X_1 + X_2 + \cdots + X_N \) where the \( X_i \) are i.i.d. Gaussian with \( \mu_X = \sigma_X^2 \) and \( N \) is a Poisson variable with \( \mu_N = \sigma_N^2 \).

(a) Compute the MGF of \( R \), \( \theta_R(t) \). Does this correspond to any of the common MGF (Gaussian, Exponential, Erlang, Poisson, Bernoulli, or Binomial)?

(b) Compute \( \mu_R \) and \( \sigma_R \) for \( \mu_X = 100, \mu_N = 1 \).

(c) Compute \( \mu_R \) and \( \sigma_R \) for \( \mu_X = 1, \mu_N = 100 \). Explain in words the similarity or difference between this result and the previous results in (b).

NB: I want you to compute \( \sigma_R \) instead of \( \sigma_R^2 \) to see how the variability compares with the mean in each case.

3. Let \( S = (X_1 + X_2 + \cdots + X_N)/\sigma \) where the \( X_i \) are independent zero-mean random variables (which may be from different distributions); with variance \( \sigma_i^2 \) and \( \sigma = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_N^2} \). Note that \( S \) is zero mean with variance 1. To be concrete, for all \( i \), let \( X_i = +1 \) or \( -1 \) with equal probability.

(a) Generate at least \( k = 100 \) samples of \( S \) with \( N = 10 \). Plot a histogram using bin sizes that are 0.25 wide. Repeat with \( N = 100 \). Do either of these look normal?

(b) An alternative approach is to use a Q-Q plot (Q stands for Quantile). Let \( S_1, S_2, \ldots, S_k \) represent the \( k \) outcomes. Further assume that they are sorted so that \( S_1 \leq S_2 \leq \cdots \leq S_k \). To see how well we fit a distribution, \( D \), we plot \( S_i \) vs. \( F^{-1}(i/(k+1)) \), where \( F^{-1}(x) \) is the inverse CDF of \( D \). If and only if the data and distribution match, the plot should be well fit by a straight line.

Generate the Q-Q plots for the data you generated in part (a) versus the normal distribution. Is it well fit by the normal distribution?

4. Repeat the previous problem for the following distribution \( X_i = \pm \frac{1}{\sqrt{2}} \) with equal probability. What is going on here in relation to the central limit theorem?