1. Let X and Y be two random variables with the joint probability density function

$$f(x,y) = \begin{cases} 2(x+y), & 0 \le x \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

   (a) Let $Z = X + Y$. Show that the joint probability density function $f_{Y,Z}(y,z)$ is

$$f_{Y,Z}(y,z) = 2z, \quad 0 \le -y + z \le y \le 1.$$

   (b) Derive the conditional probability density function of Y given $Z = z$.

2. Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$ population and let

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2.$$

   (a) If μ is unknown and σ^2 is known, show that $\bar{X}_n$ is a complete and sufficient statistic and S_n^2 is an ancillary statistic for μ. Hence, $\bar{X}_n$ and S_n^2 are independent by Basu's Theorem.

   (b) Again, if μ is unknown and σ^2 is known, find the constant c such that

$$E\left(c\bar{X}_n \sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right) = \mu.$$

   (c) Now, if both μ and σ^2 are unknown and X_{n+1} is a new observation, using the fact the $\bar{X}_n$ and S_n^2 are still independent in this case to find the constant k such that

$$\frac{k(\bar{X}_n - X_{n+1})}{S_n}$$

   follows a t distribution. Identify the degree of freedom of the t distribution specifically.

3. Let Let $X_1, \ldots, X_n$ be a random sample from an exponential distribution with pdf

$$f_X(x) = \theta e^{-\theta x}, \quad x > 0, \quad \theta > 0,$$

   and cdf

$$F_X(x) = 1 - e^{-\theta x}.$$

(a) Let $X_{(n)} = \max\{X_1, \ldots, X_n\}$ be the maximum order statistic. Show that a new random variable $Z_{(n)} = F_X(X_{(n)})$ has pdf

$$f_{Z_{(n)}}(z) = nz^{n-1}, \quad 0 < z < 1,$$

and $E(Z_{(n)}) = n/(n+1)$.

(b) Find the limiting distribution of $Y_n = \theta X_{(n)} - \log(n)$, using the fact that $\lim_{n\to\infty}(1 - x/n)^n = e^{-x}$ for a constant $x > 0$.

4. In statistics, homogeneity means equal variance between different groups. Therefore, estimation of variance can be of great interest. Say, a random sample of size n, $X_1, \ldots, X_n$, is collected from $N(0, \theta^2)$. One may use $T_n = n^{-1}\sum_{i=1}^n X_i^2$ to estimate the variance θ^2.

   (a) Show that T_n converges in probability to θ^2 and that the limiting distribution of $\sqrt{n}(T_n - \theta^2)$ is $N(0, 2\theta^4)$. Use the result to construct an approximate 95% confidence interval for θ^2. That is, find an interval (L, U) such that $P(L \le \theta^2 \le U) \approx 0.95$.

   (b) One way to stabilize the variance estimation is to find a transformation function $g(\cdot)$ such that the limiting variance of $g(T_n)$ is free of θ, or even better, $\sqrt{n}\{g(T_n) - g(\theta^2)\}$ converges in distribution to a random variable whose distribution is free of θ. Provide one such transformation function.

   (c) Use the $g(T_n)$ you found in (b) to construct another approximate 95% confidence interval for θ^2. Compare it to the one in (a) and comment on which one you would prefer.