1. Suppose that \( f(x) \) is differentiable on \((0, \infty)\) and that \( f'(x) \to 0 \) as \( x \to \infty \). Let \( g(x) = f(x + 1) - f(x) \). Show that \( g(x) \to 0 \) as \( x \to \infty \).

2. Let \( f \) be a convex function on a closed interval \([a, b]\).
   (a) Show that \( f \) is bounded on \([a, b]\).
   (b) Show that \( f \) is continuous on \((a, b)\).

3. The moment generating function of a random variable \( X \) is the function of \( t \) defined by \( M_X(t) = E[e^{tX}] \), provided the expectation exists for all \( t \) in a neighbourhood of zero. Assuming that it does, and that \( X \) has a mean \( \mu_X \), show that \( M_X(t) \geq e^{\mu_X} \), so that
   \[
   \mu_X \leq \frac{\log M_X(t)}{t} \quad \text{for } t > 0.
   \]
   Show that this inequality becomes an equality as \( t \to 0 \). Point out where you need to perform a certain interchange of operations in your derivation.

4. Let \( A \) be a \( p \times p \) positive definite matrix. Show that, if \( a \) is any vector with \( ||a|| = 1 \), then
   \[
   (a'Aa) (a'A^{-1}a) \geq 1.
   \]
   (Hints: 1. A symmetric matrix is ‘almost’ diagonal. 2. Suppose first that \( A \) is diagonal. Write out the inequality as a statement about expectations, interpreting the squares of the elements of \( a \) as probabilities. (Are they? Why?) 3. Think about Jensen’s Inequality.)

5. Prove the following special case of Slutsky’s Theorem: If \( X_n \xrightarrow{pr} c \) as \( n \to \infty \), and that \( a_n \to a \) (finite) as \( n \to \infty \), where \( \{a_n\} \) is a sequence of constants, then \( a_nX_n \xrightarrow{pr} ac \).

6. Let \( X \) be a r.v. denoting the age of failure of an electrical component, and assume that \( X \) has a d.f. \( F \) with density \( f \). The failure rate is defined as the probability of failure in a finite interval of time, given the age of the component, say \( x \). This is therefore given by
   \[
   P(x \leq X \leq x + h | X \geq x).
   \]
   The hazard rate is defined as the instantaneous failure rate:
   \[
   h(x) = \lim_{h \to 0} \frac{P(x \leq X \leq x + h | X \geq x)}{h}.
   \]
(a) Show that
\[ h(x) = \frac{f(x)}{1 - F(x)}. \]

(b) Show that \( X \) has a constant hazard rate iff it has an exponential distribution (i.e. \( f(x) = \lambda e^{-\lambda x} \) for some \( \lambda \)).

7. Prove: If \( f(x) \) is strictly increasing, twice differentiable, and convex on \([a, \infty)\) then it is unbounded.

8. In the theory of robust estimation in Statistics, one encounters the differential equation
\[ 2\psi'(x) - \psi^2(x) = -\lambda^2, \]
where \( \lambda \) is a positive constant. There are several solutions to this equation – \( \psi(x) = \lambda \) is an obvious one. Without actually solving the equation, show that if \( \psi(x) \) is any bounded solution, then \( \psi(x) \leq \lambda \) for all \( x \).

9. Show that, if an angle \( \theta \) is uniformly distributed over \((-\pi/2, \pi/2)\), then \( Y = \tan \theta \) has the Cauchy distribution, with density \( f_Y(y) = 1/[\pi (1 + y^2)] \), \(-\infty < y < \infty\). (This is also known as “Student’s” t-distribution on 1 degree of freedom.)

10. (a) Prove: If \( h(x) > 0 \) for all \( x \), then \( h(x) \) is maximized at \( x_0 \) if and only if \( \log h(x) \) is maximized at \( x_0 \).

(b) Suppose that a random sample \( x = (x_1, ..., x_n) \) has a density, depending on an unknown parameter \( \sigma^2 \), of the form
\[ f(x; \sigma^2) = h(t; \sigma^2) g(x), \quad \text{for} \]
\[ h(t; \sigma^2) = \frac{n^{-1} (\frac{(n-1)\sigma^2}{2\sigma^2})^{\frac{n-1}{2}} \exp\{-\frac{(n-1)t}{2\sigma^2}\}}{\Gamma\left(\frac{n-1}{2}\right)}, \quad 0 < t < \infty. \]

Here \( t \) is the value of a random variable \( T = T(x) \) computed from the sample, and \( g(\cdot) \) does not depend upon the parameter. If we observe the value of \( T \) then we can compute an estimate of \( \sigma^2 \). This estimate, called the Maximum Likelihood Estimate (MLE), is the value of \( \sigma^2 \) which maximizes \( h(t; \sigma^2) \), hence \( f(x; \sigma^2) \). Obtain the MLE. That is, if \( t \) is observed, what is the estimate of \( \sigma^2 \)? [Note: In the sampling situation in which this problem arises, \( T \) is a ‘sufficient statistic’ for \( \sigma^2 \) and contains all the information in the sample about that parameter, through its density \( h(t; \sigma^2) \).]