You can use Mathematica as an aid for many of the computations, however make sure to do hand calculations where suitable as well.

1. Consider sequences of numbers of the form $x_n = n^\alpha$ for $n = 1, 2, 3, \ldots$ with $\alpha \in \mathbb{R}$.

   (i) For what values of α does $\lim_{n\to\infty} x_n = 0$?

   (ii) Set now $S = \sum_{k=1}^{\infty} x_k$. For what values of α does the series, S, converge?

   (iii) Consider $\alpha = -2$. Use Mathematica to analytically evaluate S. Use this result to suggest an algorithm for numerically approximating the constant π and implement it in Mathematica.

   (iv) Consider $\alpha = -1$. In this case S is called the harmonic series. It holds that,

   $$\sum_{k=1}^{n} x_k = \log(n) + \gamma + e_n,$$

   where γ is Euler's gamma constant and e_n is an $o(1)$ sequence. Numerically approximate γ and plot the sequence e_n.

   (v) Assume you have a data set of numbers $(x_1, x_2, \ldots, x_n)$ where each number is randomly generated (uniform) on the interval $[0, 1]$ and all numbers are statistically independent. You now run an algorithm to find,

   $$m = \max_{i=1,\ldots,n} x_i,$$

   as follows:

   Step 1: Set $m = -\infty$ and $L = 0$

   Step 2: Loop on $i = 1, \ldots, n$:

   Step 2a: If $x_i > m$ then set $m = x_i$ and set $L = L + 1$ (increment L).

   Step 3: Return m as the maximum and L as the number of times a new maximum was found.

   Determine the expected value of L as a function of n. Hint: consider the harmonic series.

   (vi) Carry out experimentation by repeating the algorithm above for $10,000$ data sets of lengths $n = 1, 2, 3, 10$. See if the approximation $\log(n) + \gamma$ is useful and comment on your results.

2. Use the formal (ϵ, N), definition of the limit of a sequence to show that if $\lim_{n\to\infty} x_n = L_1$ and $\lim_{n\to\infty} y_n = L_2$ then,

   $$\lim_{n\to\infty} x_n y_n = L_1 L_2.$$

3. Provide a detailed geometric explanation of why $\frac{d}{dx}\cos(x) = -\sin(x)$.

4. The Poisson distribution is a discrete probability distribution defined by the probabilities $p(x) = e^{-\lambda}\frac{\lambda^x}{x!}$ for $x = 0, 1, 2, 3, \ldots.$ Prove that $\sum_{x=0}^{\infty} p(x) = 1$ and thus since $p(x) \geq 0$, the sequence $\{p(x)\}$ is probability distribution.

5. The function $f(x) = e^{-\frac{x^2}{2}}$ is one of the most important functions in statistics. In this problem you will investigate several properties of this function.

   (i) Show that the function has a unique maximum at $x = 0$. Use derivatives.

   (ii) Find the points of $f(x)$ where the second derivative is 0, these are called inflection points.

   (iii) Say you wish to approximate $f(x)$ on the interval $[-a, a]$ and use a Taylor series centered around $x = 0$. You use K terms to obtain a Taylor polynomial, $\tilde{f}(x)$ (a polynomial of order $K - 1$). Find the minimal K such,

   $$\max_{x \in [-a,a]} |f(x) - \tilde{f}(x)| < 10^{-4},$$

   for $a = 0.5, 1.0, 1.5$. You may use Mathematica.

   (iv) The Wikipedia page `https://en.wikipedia.org/wiki/Hermite_polynomials` presents the first eleven "probabilist's Hermite polynomials". Use Mathematica to reproduce this result and find the 12'th probabilist's Hermite polynomial with the same definition.

6. Assume that you only know: (1) That derivatives are linear. (2) That the derivative of a constant is 0. (3) That the derivative of x is 1. (4) $\frac{d}{dx}x^2 = 2x$. (5) The product rule. (6) The chain rule.
   Use (1)-(6), or a subset to obtain each of the following:

   (i) $\frac{d}{dx}x^4$ (do it using the product rule).
   (ii) $\frac{d}{dx}x^4$ (do it using the chain rule).
   (iii) The quotient rule for derivatives.
   (iv) $\frac{d}{dx}x^{-7}$.
   (v) $\frac{d}{dx}(x+5)^2$ (do it using based on (1)-(4)).
   (vi) $\frac{d}{dx}(x+5)^2$ (do it using the chain rule).

7. Given data points, $x_1, \ldots, x_n$ show that the sample mean, $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is also the minimizer, η of

   $$\sum_{i=1}^{n}(x_i - \eta)^2.$$

8. Let $s : \mathbb{R} \to \mathbb{R}$ and $c : \mathbb{R} \to \mathbb{R}$ be two real-valued functions satisfying the following:

   $$\begin{aligned} &s'(x) = c(x), \qquad c'(x) = -s(x) \qquad \text{for all } x \in \mathbb{R} \\ &s(0) = 0, \qquad c(0) = 1. \end{aligned} \tag{1}$$

   This question concerns properties of the functions c and s. You may only use (1), as well as the fact that there is exactly one pair of functions (c, s) satisfying (1).

   (i) Show that $c(x)^2 + s(x)^2 = 1$ for all $x \in \mathbb{R}$.
   (ii) Find the Taylor series for both $c(x)$ and $s(x)$ about $x = 0$.
   (iii) Let $p > 0$ be the first positive root of c. Set up a bisection method in Mathematica to approximate p, starting with the inteval $[0, 2]$, and using at least 10 bisections. You should use the Taylor series approximation for c found in part (ii) up to the x^{10} term.
   (iv) Show that $s(p - x) = c(x)$ and $c(p - x) = s(x)$ for all $x \in \mathbb{R}$.