You can use Mathematica as an aid for many of the computations, however make sure to do hand calculations where suitable as well.
Note: The Mathematica based solution are in `https://github.com/yoninazarathy/MATH7501-2021/blob/master/Assignment2Sol/Sol2Mathematica.pdf`

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.

   Solution:
   (i) The limit is 0 when $\alpha < 0$. If $\alpha = 0$ $n^\alpha \equiv 1$ and hence the limit is 1. If $\alpha > 0$, the limit is ∞.

   (ii) The series converges when $\alpha < -1$. If $\alpha \geq 0$ Then the limit of x_n is not 0 and hence the series certainly doesn't converge. If $\alpha \in [-1, 0)$ then even though $\lim_{n\to\infty} x_n = 0$ the series still does not converge. This follows from the p-test theorem for convergence of a series.

(iii) In agreement with the previous answer for $\alpha = -2$ the series convergences. In fact it converges to $\pi^2/6$ as you can calculate via Mathematica. The explicit calculation to obtain this value is more involved.See Mathematica file for solution.

(iv) In agreement with (ii), when $\alpha = -1$ the series does not converge. See Mathematica file for solution of the numerical approximation of γ and for plotting e_n.

(v) In the ith iteration of step 2a, we may consider $x_1, x_2, \ldots, x_i$. The chance of x_i (the last one) being greater than each of $x_1, x_2, \ldots, x_{i-1}$ is exactly $1/i$. Hence for the first iteration ($i = 1$) step 2a will happen always (expectation = 1). For the second iteration ($i = 2$), step 2a will happen with probability 1/2 and hence the expectation is 1/2. For the third iteration it will happen with probability 1/3 and the expectation is 1/3. Hence after n iterations the expectation is,

$$L = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}.$$

(vi) See Mathematica file.

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,

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

Solution-Comment:
This is a standard proof in analysis (note that it is beyond the level of material needed for the exam). See for example `https://en.wikibooks.org/wiki/Calculus/Proofs_of_Some_Basic_Limit_Rules`.

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

Solution-Comment:
This is a standard explanation. See for example the video `https://www.youtube.com/watch?v=R4o7sraVMZg`.

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.

Solution:
The Taylor series of $f(u) = e^u$ around $u = 0$ is,

$$e^u = \sum_{k=0}^{\infty} f^{(k)}(0)\frac{(u-0)^k}{k!} = \sum_{k=0}^{\infty} e^0 \frac{u^k}{k!} = \sum_{k=0}^{\infty} \frac{u^k}{k!}.$$

Hence setting $u = \lambda$ we get,

$$e^\lambda = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}.$$

Multiplying this by $e^{-\lambda}$ we get,

$$1 = \sum_{x=0}^{\infty} e^{-\lambda}\frac{\lambda^x}{x!}.$$

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.

   Solution:
   (i)
   $$\frac{df(x)}{dx} = -xe^{-\frac{x^2}{2}}.$$
   Solving
   $$\frac{df(x)}{dx} = 0$$
   we get $x = 0$ as the unique solution. Further,
   $$f''(x) = \frac{df(x)}{dx^2} = \big(-xe^{-\frac{x^2}{2}}\big)' = -e^{-\frac{x^2}{2}} + x^2e^{-\frac{x^2}{2}} = (x^2 - 1)e^{-\frac{x^2}{2}}.$$
   Hence $f''(0) < 0$ showing that the critical point $x = 0$ is a maximum.

   (ii) We solve for $f''(x) = 0$ or,
   $$(x^2 - 1)e^{-\frac{x^2}{2}} = 0.$$
   The Solution is $x = \pm 1$.

   (iii) See Mathematica solution.

   (iv) See Mathematica solution.

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).

Solution:

(i) $x^4 = x^2x^2$. Thus,

$$(x^4)' = (x^2)'x^2 + x^2(x^2)' = 2xx^2 + x^2(2x) = 4x^3.$$

(ii) $x^4 = (x^2)^2$. Thus,

$$(x^4)' = 2(x^2)2x = 4x^3.$$

(iii) $\frac{f}{g} = f\frac{1}{g}$. We also need that $(u^{-1})' = -u^{-2}$ (this was not stated in the question). Thus,

$$\Big(\frac{f}{g}\Big)' = f'\frac{1}{g} + f\Big(\frac{1}{g}\Big)' = f'\frac{1}{g} + f(-\frac{1}{g^2})g' = \frac{f'g - fg'}{g^2}.$$

(iv) Use the fact that $x^{-7} = \frac{1}{xx^2x^2x^2}$. Then just using the rules above it follows.

(v) Use the fact that $(x+5)^2 = x^2 + 10x + 25$.

(vi) Follows directly from 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

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

Solution:
This could be done with calculus. The gradient is,

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

If we equate it to 0 and solve for η, we obtain $\eta = \overline{x}$. Further, the second derivative is $2n > 0$ (positive) indicating that $\overline{x}$ is a minimum.

Alternatively, without using calculus, we may represent the sum of squares as an upward facing parabola in η via,

$$\underbrace{n}_{a}\eta^2 + \underbrace{\Big(-2\sum_{i=1}^n x_i\Big)}_{b}\eta + \underbrace{\sum_{i=1}^n x_i^2}_{c}.$$

This then allows us to read off the minimum value of the parabola at $\eta = -\frac{b}{2a} = \overline{x}$. In any case we see that the sample mean has the interpretation of minimizing the sum of squared deviations in the data.

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 (2), as well as the fact that there is exactly one pair of functions (c, s) satisfying (2).

(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 $x \in \mathbb{R}$.

Solution:
As we are given that there is only a single pair of functions (c, s) that satisfy (2) we may realize (and verify) that $c(x) = \cos(x)$ and $s(x) = \sin(x)$:

$$\begin{aligned} &\sin'(x) = \cos(x), \qquad \cos'(x) = -\sin(x) \qquad \text{for all } x \in \mathbb{R} \\ &\sin(0) = 0, \qquad \cos(0) = 1. \end{aligned} \tag{2}$$

(i) Now $\cos(x)^2 + \sin(x)^2 = 1$.

(ii)

$$\begin{aligned} \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \dots \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}. \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \dots \\ &= \sum_{n=1}^{\infty} (-1)^{(n-1)} \frac{x^{2n-1}}{(2n-1)!} \overset{\text{or}}{=} \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}. \end{aligned}$$

(iii) See Mathematica based solution.

(iv) It is the case that $p = \pi/2$ and it is known that $\sin(\pi/2 - x) = \cos(x)$ and $\cos(\pi/2 - x) = \sin(x)$.