

Mathematics, Biology, and Computer Science

[1]

We give an example solution for only (1).

(1) Taylor Expansion

A Taylor expansion is a method for approximating a function as an infinite series centered at a specific point. If the function is sufficiently smooth, one can express its behavior as a polynomial using the function's value and its derivatives at that point. The qualitative behavior of a differential equation near a fixed point can be understood by analyzing a low-order polynomial approximation. Specifically, the Taylor series of a function is expressed as:

$$f(a+x) = f(a) + f'(a)x + \frac{f''(a)}{2!}x^2 + \dots$$

[2]

Question 1

(1) $f'(x) = \tan x, f''(x) = \frac{1}{\cos^2 x}$

(2) Omitted.

(3) $x^2 + (y-1)^2 = 1$

Question 2

(1) $P_0(x) = 1, P_1(x) = x, P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}, P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x$

(2) Omitted.

(3) $a_0 = \frac{\int_{-1}^1 f(x)P_0(x)dx}{\int_{-1}^1 P_0(x)P_0(x)dx} = \frac{\int_{-1}^1 (x^3+x^2+x+1)dx}{2} = \frac{4}{3}$

Question 3

(1) $\tan \frac{\theta}{2} = \frac{\sin \frac{\theta}{2}}{\cos \frac{\theta}{2}} \therefore \sin \frac{\theta}{2} = \tan \frac{\theta}{2} \cdot \cos \frac{\theta}{2}$

$\sin^2 \frac{\theta}{2} + \cos^2 \frac{\theta}{2} = 1$

$\therefore \tan^2 \frac{\theta}{2} + 1 = \frac{1}{\cos^2 \frac{\theta}{2}}$

$\therefore \cos^2 \frac{\theta}{2} = \frac{1}{1+\tan^2 \frac{\theta}{2}}$

$\sin \theta = 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} = 2 \tan \frac{\theta}{2} \cdot \cos^2 \frac{\theta}{2} = \frac{2x}{1+x^2}$

$\cos \frac{\theta}{2} = \sqrt{1 - \sin^2 \frac{\theta}{2}} = \frac{1-x^2}{1+x^2}$

(2) $\int_0^1 \frac{1}{1+x^2} dx - \int_0^1 \frac{x}{1+x^2} dx = \frac{\pi}{4} - \frac{1}{2} \ln 2$

[3]

Question 1

(1)

	female	male
Master's program	131.25	189.75
Doctoral program	43.75	63.25

(2) $(2-1)^2 = 1$

Question 2

(1) Density can explain seed weight well.

(2) $10554/12259 = 0.86$

(3) ③ (The regression coefficient and correlation coefficient have the same positive or negative value.)

Question 3 One sample answer is shown below:

1. Are there any scenes that they thought were more fair or unfair compared to others?
2. Questions 1 to 5 and the ratio of fair/unfair responses are independent.
3. Perform a chi-square independence test with 4 degrees of freedom on a cross tabulation table with 'Questions 1/2/3/4/5' as rows and 'Fair responses/Unfair responses' as columns.

[4]

Mathematics, Biology, and Computer Science

Question 1

A Hidden Markov Model (HMM) probabilistically models unobservable hidden states S_1, S_2, \dots, S_N and the observed data O_1, O_2, \dots, O_T generated from them. An HMM is specified by the transition probabilities between hidden states $P(S_t | S_{t-1})$ and the emission (observation) distributions $P(O_t | S_t)$.

An HMM consists of the following components:

- **Hidden states** $S = \{S_1, S_2, \dots, S_N\}$
- **State transition probabilities** $A = \{a_{ij}\}$, which denote the probability of transitioning from state i to state j .
- **Emission probabilities** $B = \{b_i(o)\}$, which give the probability of observing o when the system is in state i .

HMMs are widely used in speech recognition, natural language processing, and bioinformatics, for tasks such as state inference and optimal state-sequence prediction.

Question 2

- (1) Choose x_1 and x_3 so that the total satisfaction is 130.
- (2) $D[i][j] = \max(D[i-1][j], D[i-1][j - p_i/100] + s_i)$.
- (3) Pseudo-code in Python style:

```

i = 0
for j = 0, ..., U:
    D[i][j] = 0
for i = 1, ..., n:
    for j = 0, ..., U:
        if  $p_i/100 > j$ :
            D[i][j] = D[i-1][j]
        else:
            D[i][j] = max(D[i-1][j], D[i-1][j -  $p_i/100$ ] +  $s_i$ )
return D[n][U]
```

- (4) First, in the program of (3), prepare an auxiliary array that records, for each $D[i][j]$, which entry achieved the maximum: either $(i-1, j)$ or $(i-1, j - p_i/100)$. Then, trace back from $D[n][U]$ following these recorded predecessors to recover the selected items.

Question 3

- (1) The likelihood function is

$$L(p_1, p_2, \dots, p_6 | c_1, c_2, \dots, c_6) = \prod_{i=1}^6 \binom{n}{c_i} p_i^{c_i} (1 - p_i)^{n - c_i}.$$

- (2) Apply the method of Lagrange multipliers to the likelihood in (1). The solution is $\hat{p}_i = \frac{c_i}{n}$.
- (3) The count c_i of face i follows a binomial distribution, so the estimator of its probability p_i is the binomial mean $\hat{p} = c/n$ (subscript i omitted below).

$$Z = \frac{c - np}{\sqrt{np(1-p)}} = \frac{c/n - p}{\sqrt{p(1-p)/n}} \tag{1}$$

For sufficiently large n , Z is approximately standard normal, and $\hat{p} = c/n$ is approximately

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right). \tag{2}$$

Let $z_{1-\alpha/2}$ denote the upper $\alpha/2$ quantile of the standard normal (e.g., when $\alpha = 0.05$, $z_{0.975} = 1.96$). Then

$$-z_{1-\alpha/2} \leq \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \leq z_{1-\alpha/2}. \tag{3}$$

Since \hat{p} is a consistent estimator of p , when n is large it is reasonable to replace p in the denominator by \hat{p} . As a result, the two-sided $100(1-\alpha)\%$ confidence interval is

$$\hat{p} - z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}. \tag{4}$$

Mathematics, Biology, and Computer Science

Question 1

- (1) $\dot{x} = x(x-1)(x+1)$
(2) $\dot{x} = x^2(x-1)^2$
(3) $\dot{x} = \sin x$
(4) $\dot{x} = r - x^2, \quad r = 0$
(5) $\dot{x} = (x^2 + r)(x+1)(x-1), \quad r = 0$

Question 2 (1) $a + b, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad a - b, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

- (2) (x, y) satisfying $y = x$
(3) (x, y) satisfying $y > x$

Question 3 (1) $h = H/(rK), \quad a = A/K$

(2) $x^* = 0, 1$

(3) (summary)

$a \leq 1$ and $h \leq \left(\frac{1+a}{2}\right)^2$ or $a > 1$ and $h < a$.

[6]

Question 1

Ribosomes play a role in the synthesis of protein by reading the genetic information of mRNA and binding amino acids.

Question 2

The 16S rRNA gene is commonly found in bacteria and archaea and is essential for protein synthesis. Since the gene contains alternating conserved and variable regions, PCR amplification can target multiple bacteria and archaea using a universal primer set that binds to the conserved regions flanking the variable ones. By analyzing the nucleotide sequences of these variable regions and referencing a database, it is possible to identify the species present. For these reasons, the 16S rRNA gene is used in studying microbiota diversity.

Question 3

The following reagents and materials are required:

- Template DNA (PCR products, genome, or reverse transcription products are all acceptable): used as the template for PCR.
- Forward primer: specifies the start or end point of the PCR.
- Reverse primer: specifies the end or start point of the PCR.
- Polymerase: an enzyme that catalyzes the polymerization of nucleotides.
- dNTPs: contain four types of nucleotides—dATP, dGTP, dTTP, and dCTP—which are the basic building blocks of a DNA strand. Each nucleotide consists of a sugar, a base, and a phosphate group.
- Mg^{2+} (optional): enhances binding between polymerase and nucleotides and increases annealing stability.
- PCR buffer (optional): maintains stable PCR amplification by preserving the pH of the PCR solution and optimizing primer binding to the template DNA.

Question 4

Bacteria or archaea whose DNA sequences don't match the primer used cannot be detected. Additionally, species with similar sequences cannot be distinguished. Furthermore, species not listed in the database cannot be identified.

Question 5

(Answer of example)

Aim: To investigate whether electromagnetic waves and cool environments influence microbiota by analyzing the metagenomes in microwaves and refrigerators.

- Swab the interior of microwaves and refrigerators, then extract metagenomes.
- Conduct PCR amplifications targeting the 16S rRNA gene using the metagenomes as templates.
- Sequence the PCR products, classify species based on the 16S rRNA database for bacteria and archaea, and compare the microbiomes of microwaves and refrigerators.

[7]

Mathematics, Biology, and Computer Science

Question 1

1. B
2. C
3. Rough endoplasmic reticulum (RER) has ribosomes attached and synthesizes proteins, whereas smooth endoplasmic reticulum (SER) primarily synthesizes lipids and metabolizes drugs.
4. A, C, E

Question 2

1. Elongation rate = (on-rate \times [G-actin]) – off-rate
 $= 11 \mu M^{-1} s^{-1} \times 5 \mu M - 1.4 s^{-1}$
 $= 55 - 1.4 = 53.6$ subunits/s
2. Membrane protrusion speed = 53.6 subunits/s \times 2.7 nm/subunit
 $= 144.72$ nm/s

Question 3

1. To prevent unwanted degradation in the cytoplasm and avoid damage to cellular components.
2. It has a safety mechanism in which enzyme activity is significantly reduced around neutral pH.
3. Fusing enzyme X with a fluorescent protein (e.g., GFP) and observing with a fluorescence microscope.
4. From the Michaelis – Menten equation,
 $K_m = \frac{V_{max}[S]}{V} - [S] = \frac{10 \times 1}{4} - 1 = 1.5$ mM

Question 4

1. Phase-contrast microscope, because it allows observation of internal dynamics in live cells without staining.
2. Electron microscope, as it provides higher resolution than optical microscopes, enabling detailed observation of fine structures.
3. Fluorescence intensity decreases due to photobleaching caused by repeated exposure to excitation light.