

Mathematics, Biology, and Computer Science

[1]

Brief explanations of the four choices are provided below.

(1) **Continuity of Functions**

For a real-valued function $f(x)$, if the limit of $f(x)$ at any real number a exists and equals $f(a)$, then $f(x)$ is said to be continuous at a .

(4) **Correlation Coefficient**

An index indicating the strength of the linear relationship between two random variables X and Y , which takes values between -1 and 1 . It is obtained by dividing the covariance $\text{Cov}[X, Y]$ by the standard deviations σ_X and σ_Y .

(9) **Breadth-First Search (BFS)**

A graph search algorithm that explores nodes starting from the shallowest node (the node with the shortest distance from the start).

(14) **Nucleotide**

A fundamental unit composing DNA and RNA, consisting of a phosphate group, a sugar (ribose or deoxyribose), and a base.

[2]

Problem 1

(1) 1

(2) $(\sinh kx)' = k \cosh kx$, $(\cosh kx)' = k \sinh kx$

(3) $\sinh 2x = 2 \sinh x \cosh x$, $\cosh 2x = 2 \cosh^2 x - 1$

(4) Performing integration by substitution using $x = \sinh u$, we obtain:

$$\int \frac{1}{\sqrt{x^2 + 1}} dx = \log(x + \sqrt{x^2 + 1}) + C.$$

Problem 2

(1) When $(x, y) = (\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$.

(2) Let the three sides of the rectangular prism be x , y and z . Then we obtain (1) - (4) below using Lagrange's method of undetermined multipliers.

1. $yz - \lambda(y + z) = 0$
2. $xz - \lambda(x + z) = 0$
3. $xy - \lambda(x + y) = 0$
4. $xy + yz + zx - 3 = 0$

From (1) and (2), we obtain $x = y$. From (2) and (3), we get $y = z$. Thus, $x = y = z$, that is a cube with a side length of $6^{\frac{1}{3}}$.

Problem 3

(1) From the conditional expression in D , we obtain the following four inequalities:

$$\begin{aligned} y &\leq -x, \\ y &\geq -x + 1, \\ y &\leq x, \\ y &\geq x - 1. \end{aligned}$$

The region defined by these inequalities is a quadrilateral with four points: $(0, 0)$, $(\frac{1}{2}, \frac{1}{2})$, $(\frac{1}{2}, -\frac{1}{2})$, and $(1, 0)$.

(2) By converting D into the integration bounds, the range of x is $0 \leq x \leq 1$, and for each x , the range of y is $-(x - 1) \leq y \leq x$. Thus, the integral can be expressed as:

$$I = \int_0^1 \int_{-(x-1)}^x \frac{x-y}{1+x+y} dy dx.$$

The inner integral is

$$I_1 = \int_{-(x-1)}^x \frac{x-y}{1+x+y} dy.$$

By introducing the substitution $u = 1 + x + y$, the calculation leads to the final result:

$$I = \frac{1}{2} \ln 2 - \frac{1}{4}.$$

[3]

Mathematics, Biology, and Computer Science

Problem 1

(1)

$$\frac{(36-40)^2}{40} + \frac{(33-30)^2}{30} + \frac{(22-20)^2}{20} + \frac{(9-10)^2}{10} = \frac{4+3+2+1}{10} = 1$$

(2) $4 - 1 = 3$

Problem 2

(1) Body weight does not strongly provides a good predictor of body fat percentage.

(2) $1.33/218.42 = 0.006$

(3) ② (The p-value for the gradient means the confidence interval includes 0.)

Problem 3

1. Did the analysis predict landslides at each location?
2. Landslides can be predicted using rainfall, maximum slope, distance from rivers, or land use.
3. (i) (iii) (v) (vi)
4. Perform a multivariate logistic regression analysis using (6) as the dependent variable.
5. The maximum slope of a given mesh can be estimated from data of (3) in adjacent meshes. The distance from rivers can be calculated from (1).

[4]

1.1 By the multiplication theorem, $p(A \cap B) = p(B|A)p(A) = p(A|B)p(B)$. From this, we obtain Bayes' theorem:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)}.$$

1.2 We solve using Bayes' theorem. The probability of interest, $p(\text{non-infected} | \text{positive})$, can be expressed by Bayes' theorem as follows:

$$p(\text{non-infected} | \text{positive}) = \frac{p(\text{positive} | \text{non-infected}) \cdot p(\text{non-infected})}{p(\text{positive})}.$$

The calculations for each term are as follows:

$$p(\text{non-infected}) = 0.9998 \quad (\text{proportion of non-infected individuals}),$$

$$p(\text{infected}) = 0.0002 \quad (\text{proportion of infected individuals}),$$

$$p(\text{positive} | \text{infected}) = 0.86 \quad (\text{positive rate among infected individuals}),$$

$$p(\text{negative} | \text{infected}) = 0.14 \quad (\text{negative rate among infected individuals}),$$

$$p(\text{positive} | \text{non-infected}) = 0.070 \quad (\text{positive rate among non-infected individuals}),$$

$$p(\text{negative} | \text{non-infected}) = 0.930 \quad (\text{negative rate among non-infected individuals}).$$

Next, the overall positive rate $p(\text{positive})$ is computed as:

$$p(\text{positive}) = p(\text{positive} | \text{infected}) \cdot p(\text{infected}) + p(\text{positive} | \text{non-infected}) \cdot p(\text{non-infected}).$$

Substituting the values gives:

$$p(\text{positive}) = 0.86 \cdot 0.0002 + 0.070 \cdot 0.9998 = 0.070158.$$

Finally, substituting into Bayes' theorem yields:

$$p(\text{non-infected} | \text{positive}) = \frac{0.070 \cdot 0.9998}{0.070158} \approx \frac{0.069986}{0.070158} \approx 0.9975.$$

Therefore, when the test result is positive, the probability that the person is non-infected is approximately 0.9975.

Mathematics, Biology, and Computer Science

- 2 The probability density function of a normal distribution $f(x | \mu, \sigma^2)$ is defined as:

$$f(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right).$$

In general, for N data points $\{x_1, x_2, \dots, x_N\}$ that are assumed to follow a normal distribution, the likelihood function $L(\mu, \sigma^2)$ is the product of the probabilities of those data points:

$$L(\mu, \sigma^2) = \prod_{i=1}^N f(x_i | \mu, \sigma^2).$$

Taking the logarithm (the log-likelihood function) and expanding gives:

$$\ell(\mu, \sigma^2) = \sum_{i=1}^N \left[-\frac{1}{2} \log(2\pi\sigma^2) - \frac{(x_i - \mu)^2}{2\sigma^2} \right] = -\frac{N}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^N (x_i - \mu)^2.$$

Hence, the maximum likelihood estimate $\hat{\mu}$ of μ is obtained by solving:

$$\frac{\partial \ell(\mu, \sigma^2)}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^N (x_i - \mu) = 0,$$

which gives

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N x_i.$$

Thus, for the given data points $\{-2, -1, 0, 1, 2\}$, the maximum likelihood estimate of the mean parameter is:

$$\hat{\mu} = \frac{-2 + (-1) + 0 + 1 + 2}{5} = \frac{0}{5} = 0.$$

Similarly, the maximum likelihood estimate $\hat{\sigma}^2$ of σ^2 is obtained by solving:

$$\frac{\partial \ell(\hat{\mu}, \sigma^2)}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^N (x_i - \hat{\mu})^2 = 0,$$

which leads to

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{\mu})^2.$$

Therefore, for the data points $\{-2, -1, 0, 1, 2\}$, the maximum likelihood estimate of the variance parameter is:

$$\hat{\sigma}^2 = \frac{1}{5} [(-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2] = \frac{10}{5} = 2.$$

- 3 (1) Assuming the data points are independent, the likelihood function is given by:

$$L(\mu | \{x_i\}) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^N \exp\left(-\sum_{i=1}^N \frac{(x_i - \mu)^2}{2\sigma^2}\right).$$

(2) We derive the posterior distribution

$$p(\mu | \{x_i\}, \mu_0, \sigma_0^2).$$

Using Bayes' theorem, the posterior distribution is:

$$p(\mu | \{x_i\}, \mu_0, \sigma_0^2) \propto L(\mu; \{x_i\}) \cdot p(\mu).$$

Here, $p(\mu)$ is the prior distribution, which is assumed to follow a normal distribution $N(\mu_0, \sigma_0^2)$:

$$p(\mu) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{(\mu - \mu_0)^2}{2\sigma_0^2}\right).$$

Substituting this into the posterior distribution:

$$p(\mu | \{x_i\}, \mu_0, \sigma_0^2) \propto \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^N (x_i - \mu)^2 - \frac{(\mu - \mu_0)^2}{2\sigma_0^2}\right].$$

By completing the square in the exponent, we obtain:

$$p(\mu | \{x_i\}, \mu_0, \sigma_0^2) \propto \exp\left(-\frac{(\mu - \mu_{\text{post}})^2}{2\sigma_{\text{post}}^2}\right),$$

Mathematics, Biology, and Computer Science

where the posterior mean μ_{post} and variance σ_{post}^2 of the posterior distribution are given by:

$$\mu_{\text{post}} = \frac{\frac{1}{\sigma^2} \sum_{i=1}^N x_i + \frac{\mu_0}{\sigma_0^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_0^2}},$$

$$\sigma_{\text{post}}^2 = \left(\frac{N}{\sigma^2} + \frac{1}{\sigma_0^2} \right)^{-1}.$$

4.1 Calculating yields $a_2 = 1$, $a_3 = -4$, and $a_4 = -11$.

4.2

```
def a(n) { if (n == 0) return 1; if (n == 1) return 2; return 2 * a(n-1) - 3 * a(n-2); }
```

[5]

Problem 1

(1) Fixed points $x^* = 0, K$.

(2) Stability is determined from the slope of the phase portrait $x-\dot{x}$. When $r > 0$, $x=0$ is unstable and $x=K$ is stable. When $r < 0$, $x=0$ is stable and $x=K$ is unstable.

(3) Since $\dot{x} = 0$ always holds, all x are fixed points, and therefore maintains the value of $x(0)$, exhibiting neutral stability.

Problem 2

(1) $\dot{y} = -\frac{k}{m}x - \frac{b}{m}y$

(2) $A = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix}$

(3) The eigenvalues of A are complex numbers when $b^2 - 4mk < 0$.

Problem 3

(1) $x(t) = x_0 e^{-ky_c t}$ (2) $z(t) = x_0(1 - e^{-ky_c t})$ (3) $\dot{x} = k_1 a x - k_2 x^2 - k_3 b x$ (4) $k_1 a - k_3 b > 0$ (5) Transcritical bifurcation

[6]

Problem 1

a. Deoxyribose (sugar), b. Thymine, c. Guanine, d. Histone, e. Chromatin, f. Genome, g. 46, h. 23, i. Ribosome, j. Promoter, k. Transcription factor, l. RNA polymerase, m. Transfer RNA (tRNA), n. Peptide bond, o. Folding.

Problem 2

(1) Heat to around $94 - 98^\circ\text{C}$. Hydrogen bonds between bases break, separating DNA strands.

(2) For example, Fw: TACATTCTACAACCTACAGCC

Rv: AATTCGTTTGGGCTACTCTA

(3) DNA polymerase

Problem 3

a. Scaffold DNA, b. Staple strand, c. Sticky end

Problem 4

5' -TACGAGCCAGGGGAACCTCGTA-3' Loop-releasing DNA: 5' -TTCCCCTGGCTCGTA-3'

Problem 5

(1) First, calculate the molar concentration according to the Lambert-Beer law. $1.0 \times 10^{-6} \text{ mol/L}$

Next, convert the molar concentration to ng/ μL . 3.3 ng/ μL

(2) 25 μL : DNA, 75 μL H_2O

[7]

Problem 1

(1)

a. Nucleus: Holds the cell's genetic information and serves as the center for DNA replication and transcription.

b. Mitochondria: Generates ATP, the energy source of cells.

c. Golgi apparatus: Storage and secretion of substances such as proteins.

d. Lysosome: Breaks down waste materials and foreign substances within the cell.

Mathematics, Biology, and Computer Science

(2)

a. Structural features

Microtubules: Hollow fibers with a diameter of 25 nm, primarily composed of 13 protofilaments.

Actin filaments: Fibers with a double-helix structure and a diameter of approximately 7 – 9 nm.

b. Nucleotides used for polymerization

Microtubules: GTP

Actin filaments: ATP

c. Names of the protein constituent units

Microtubules: $\alpha\beta$ -tubulin

Actin filaments: G-actin

d. Two main functions within the cell

Microtubules: Chromosome segregation, intracellular material transport, ciliary movement, etc.

Actin filaments: Cell motility, support of cell shape, cell division, etc.

e. General names of complementary motor proteins (for microtubules, provide two for a complete answer)

Microtubules: Kinesin, Dynein

Actin filaments: Myosin

(3)

Lamellipodia , Filopodia

(4)

Nucleation: The initial stage of polymerization where a small number of monomers aggregate to form a stable nucleus (oligomer). ※ If the answer indicates the starting point of fiber formation to some extent, it is considered correct.

Critical concentration: The monomer concentration at which the rates of polymer elongation and dissociation are equal. Above the critical concentration, the polymer grows; below it, the polymer dissociates.

(5)

Microtubules: γ -tubulin complex (centrosome may also be accepted).

Actin filaments: Arp2/3 complex

Problem 2

a. 3D fluorescence imaging of cells: CLSM

b. Fluorescently labeled cytoskeleton on the cover glass surface in a system with a large amount of fluorescent dye: TIRFM

c. Internal molecular structure of a virus: TEM

d. Observation of cell surface structure in living cells at nanometer-scale resolution : HS-AFM

e. Observation of bacterial surface structure at nanometer-scale resolution: SEM

Problem 3

a. Purification of His-tag fusion protein: C b. Separation and analysis of proteins of different molecular sizes (with denaturation): B c.

Fractionation of proteins of different molecular sizes: E d. Purification of GST-tag protein: A e. Separation of polymerized microtubules

from cell extract: D

Problem 4

The dissociation constant, K_d is calculated using the following equation:

$$K_d = k_{\text{off}}/k_{\text{on}} = 1.0 \times 10^{-14}\text{M}$$