1) Given that a double stranded DNA molecule contains 21% of Adenosine, find the corresponding percentage of Guanosine it contains
   A) 21%
   B) 42%
   C) 29%
   D) 25%

The DNA is double stranded: there are as many As as Ts, and as many Gs and Cs. Since Adenosine represents 21% of the content, T represents another 21%; therefore G+C represents 58%, and since G and C are in equal amount, Guanosine represents 29% of the content.

2) The codon for leucine is CUC. How many different amino acids could possibly result from single-base substitutions?
   A) 1
   B) 3
   C) 5
   D) 7

There are 9 options for single-base substitutions, depending on which base is mutated:
- First base. The mutated codons are AUC, GUC, and UUC, which code for Ile, Val, and Phe, respectively.
- Second base. The mutated codons are CAC, CCC, and CGC, which code for His, Pro, and Arg, respectively.
- Third base: The mutated codons are CUA, CUU, and CUG. All three code for Leu.
There are therefore 7 possible amino acids.

3) A bacterial population increases from 100 to 104,857,600 in 15 hours. What is the generation time of this culture?
   A) 1 hour
   B) 2 hours
   C) 45 minutes
   D) 30 minutes

a. Compute number of generations: \( P_n = P_0 \cdot 2^n \) \( \Rightarrow n = \log_2(\frac{P_n}{P}) = \log_2(104857600/100) = 20 \)
b. Generation time : \( t = \frac{15}{20} = 0.75 \) hours = 45 minutes

4) The Lenski experiment is designed to mimic evolution in the laboratory: bacteria are grown over a large number of generations, they experience spontaneous mutations and some of these mutations are selected for better fitness with the environment. How many spontaneous mutations do occur in a population of bacteria? Is this number large enough that we can expect to see results from such an experiment? To get an estimate, let us consider that there are on average 3x10^7 bacteria, that the genome of one bacterium is 5 10^6 base pair long, that the
intrinsic mutation rate is $5 \times 10^{-10}$ mutation per base pair, and that we have observed 50,000 generations. What is your estimate of the total number of point mutations that occurred:

A) $3.75 \times 10^8$
B) $3.75 \times 10^9$
C) $1.875 \times 10^8$
D) $5 \times 10^6$

$$N_{mutation} = (\text{Population size}) \times (\text{Number of base pair}) \times (# \text{ mutation / base pair}) \times (# \text{ of generation})$$

$$= 3 \times 10^7 \times 5 \times 10^6 \times 5 \times 10^{-10} \times 50000 = 3.75 \times 10^9$$

5) A professor decides to illustrate in class the Luria and Delbruck experiment. To start, she prepares a batch of E. Coli. This batch is then exposed to bacteriophage T1, and then directly divided into 5 samples and spread on petri dishes. The professor then divides the students into 4 groups and asks each group to predict what they expect to see in terms of number of resistant colonies on each petri dish the day after. Here are their answers:

<table>
<thead>
<tr>
<th>Petri dish</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>10</td>
<td>14</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>20</td>
<td>15</td>
<td>50</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>100</td>
<td>12</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>30</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Mean</td>
<td>0</td>
<td>32</td>
<td>14.4</td>
<td>50</td>
</tr>
<tr>
<td>Variance</td>
<td>0</td>
<td>1570</td>
<td>12.3</td>
<td>0</td>
</tr>
</tbody>
</table>

Which of these four groups is most likely to have predicted the correct behavior?

A) Group 1
B) Group 2
C) Group 3
D) Group 4

The same culture is exposed to the phage: all 5 petri dishes will show the same behavior: Poisson distribution with mean similar to variance.

6) Assume that the whole DNA corresponding to the human genome can be represented as a ribbon of length $L=1$ meter and width $W = 2$nm. Assume also that this DNA contains $3 \times 10^9$ bases. From these numbers, compute the corresponding density of letters per nm$^2$. If we could create a material that would allow storage with the same density, how many books could we store on 1 cm$^2$ of such material, assuming that a book contains on average 200,000 letters.

A) $7.5 \times 10^8$
B) $7.5 \times 10^9$
C) $1.5 \times 10^{10}$
D) $3.0 \times 10^8$
Density of letter = (Number of letter) / (area of DNA ribbon) = 3 \times 10^9 / (2 \times 10^9) = 1.5 \text{ letter} / \text{nm}^2
Number of letters / 1\text{cm}^2 = 1.5 \text{ letter} / \text{nm}^2 \times 10^{14} \text{ nm}^2 / \text{cm}^2 = 1.5 \times 10^{14}
Number of books: $1.5 \times 10^{14} / 200000 = 0.75 \times 10^9 = 7.5 \times 10^8$

7) You have been given the task to predict the structure of a small protein with 21 amino acids. You design a very simple method for this. Each amino acid defines a basic block. In the process of building a model for the protein, you have 5 possibilities when attaching a new amino acid to the current chain. For example, you can build 25 different conformations for a protein of 3 amino acids. Once a conformation for the full protein is built, you have at your disposition a tool that can check if it is correct. Assuming that it takes no time to build all possible conformations for the protein of 21 amino acids, and that it takes 1 ns (nanosecond) to check one of these conformations, how long (in seconds) would it take approximately to find the correct structure for the protein (assume that you will have to check all conformations).

A) 20 seconds
B) 100,000 seconds
C) 3,600 seconds
D) 1,000,000 seconds

Total number of conformations: $5^{20}$
Total time to check them: $5^{20} \times 10^{-9} = \text{(approx)} 100,000$

8) Which of these statements about information and noise is (are) likely to be wrong?
A) A processor is always able to distinguish information from noise; therefore noise is never a problem.
B) What is information and noise is defined by the context.
C) Noise is an unwanted perturbation to a wanted signal.
D) Noise level can be quantified as the variability of a measurement around its mean.

9) You have designed *E. coli* such that it can react to light. In the presence of light it generates a white dot, while in the absence of light it generates a black dot. You want to use a bio-film of dimension 1 cm x 1 cm as a synthetic camera; assuming that the bacteria cover uniformly your bio-film, and that each bacterium has the shape of a square, with a side of 1 µm, what would be the resolution of your camera, in megapixels (1 Megapixel = $10^6$ pixels):
A) 1 megapixel
B) 10 megapixels
C) 50 megapixels
D) 100 megapixels

# of pixels = #bacteria = Surface area/(size of a bacterium) = 1 cm x 1 cm / ($10^{-4}$ cm x $10^{-4}$ cm) = $10^8$