Takehome Exam 8 Answers

Question 1.

True or False - MrBayes is extremely reliable in predicting the correct tree. As a result, the support values that the program produce are conservative and should be considered more reliable than bootstral support values.

True
False

Question 2.

True or False - the car trunk analogy illustrates that genes can be under purifying selection without increasing the fitness of the individuals carrying the gene.

True
False

Question 3.

True/False Phylogenetic reconstruction using Markov chain Metropolis Coupled Monte Carlo sampling aims to find the phylogenetic tree that is most probable given the data by walking around in tree (and other parameters) space with a biased walk and sampling the trees.

True
False

Question 4.

What allowed Walter Fitch to beat of the CDC in picking the strain of flu to vaccinate, year after year, until the CDC finally started doing things his way?

A. He had a huge team of researchers on the project, while the CDC just had one retired professor on the project
B. He had enormous computing power at his disposal, while the CDC was using a pocket calculator
C. He had new, modern laboratory equipment, allowing him more to obtain more accurate sequences than the CDC
D. He had decades of knowledge of evolution and with it was able to develop the novel idea of using an evolutionary pattern he spotted in flu phylogenies to predict next years' epidemic
E. A, B, and C

Question 5.

When considering data obtained from flipping one coin four times and obtaining all tails, what will the maximum likelihood approach calculate?
(Consider that there are three models possible for this coin toss: 1. A fair coin model. 2. A coin with both sides heads. And 3. A coin with both sides tails. Priors are 1. 99.8%, 2. 0.1%, 3. 0.1%)

A. The probability of obtaining all tails, averaged over all possible models (i.e. $((.5)^4 * 0.998) + (0 * 0.001) + (1.0 * 0.001))$

B. The probability of obtaining all tails, given the model that maximizes this probability (i.e. 100% and it will always choose the third model)
C. The probability of obtaining all tails when using a fair coin (i.e. $(.5)^3 * 0.998)$)
D. The probability of obtaining all tails, without considering possible models. This is possible because a robot is used to explore probability space.
E. Maximum likelihood is not applicable to coin toss data, only nucleotide or amino acid sequence data can be used.

Question 6.

Which computing programs use the MCMCMC approach?

A. MrBayes and Beast
B. Seaview
C. PSI-BLAST
D. Njplot
E. Swiss pdb

Question 7.

Which of the following is a factor when calculating the bootstrap support of a clan in a phylogenetic tree?

A. Branch lengths
B. The number of times the clan is recovered in the analysis of a set of bootstrap samples
C. The number of times the clan is recovered in the trees sampled in a random biased walk through tree space
D. Lineage sorting
E. The Gamma parameter and among site rate variation

Question 8.

Which of the following is a factor when calculating the posterior probability of a clan in a phylogenetic tree using an MCMCMC approach?

A. Branch lengths
B. The number of times the clan is recovered in the analysis of a set of bootstrap samples
C. The number of times the clan is recovered in the trees sampled in a random biased walk through tree space
D. Lineage sorting
E. The Gamma parameter and among site rate variation

Question 9.
What does ASRV stand for, and what does this mean?

Among site rate variation.
Sites in a DNA or protein sequence undergo substitution events with different frequency.

Question 10.
Which processes allow favorable genetic changes to be combined into the same individual, speeding up the rate of evolution?

A. Gene duplication and neofunctionalization
B. Genetic drift
C. Punctuated equilibrium
D. Sex and HGT
E. None of the above.

Question 11.
Which of the following programs can align nucleotide sequences based on a protein alignment?

A. MrBayes
B. Seaview
C. psiBLAST
D. Clustalo
E. Cluster

Question 12.
Which value of non-synonymous/synonymous rate ratio (+omega or dN/dS) would you expect for a protein-coding gene that encodes an enzyme vital for photosynthesis?

A. 0.01 or smaller
B. about 1.0
C. 1.2 or larger

Question 13.
Why is a Bayesian consideration advantageous, as opposed to the Maximum Likelihood estimate, for the coin toss example used in previous question?

It takes prior knowledge into account. Thus a really strange model (such as Gremlins) needs a lot of support through the data to overcome the prior.

Question 14.
Why might amino acids on the outside of virus capsids be under positive selection?

A. They interact with the immune system and need to change to evade recognition and capture
B. These positions are under strong selection to maintain function, because they are important to the virus
C. They interact with the host DNA and need to change as the host evolves
D. Binding of host antibodies triggers mutations in the virus
E. All the above

Question 15.
A dN/dS ratio smaller than 1 implies positive or Darwinian selection

True
False

Question 16.
The shape parameter of the Gamma distribution is often used to describe ASRV. Order the values below from the smallest to the largest amount of ASRV in the distribution.

1. 5
2. .4
3. 1
4. 20

Correct Order
4. 20
1. 5
3. 1
2. .4

Question 17.
The Gamma distribution can approximate (or be identical) to several other distributions. Which shape parameter results in a Gamma distribution that is similar to a Normal Distribution?

10
2
1
.5

Question 18.
The Gamma distribution can approximate (or be identical) to several other distributions. Which shape parameter results in a Gamma distribution that is similar to a Poison Distribution?

10
2
1
.5

Question 19.
The Gamma distribution can approximate (or be identical) to several other distributions. Which shape parameter results in a Gamma distribution that is similar to an Exponential Distribution?

10
2
1
.5

Question 20.
The estimation of dN/dS ratios has been used to (this does not imply that this is a recommended or uncontroversial practice)

A. identify sites in a protein experiencing diversifying selection
B. identify proteins that have a function that contributes to the molecules or the organisms fitness
C. Identify a branch in a molecular phylogeny along which a protein experienced positive selection
D. Identify a branch in a molecular phylogeny along which a single position in protein experienced positive selection
E. All of the above

Question 21.
You do a phylogenetic reconstruction of protein encoding nucleotide sequences. You observe exactly the same number of synonymous and non-synonymous substitutions. What does this suggest for the type of selection acting on the protein?

A. the sequences evolve neutrally.
B. purifying selection
C. diversifying selection
D. Darwinian selection