

情報生物学 1/1

- [1] 確率変数  $X$  と  $Y$  は独立で、それぞれ  $N(0, 1)$ ,  $N(0, 3)$  の正規分布に従うものとする。(計 40 点)
- (1)  $aX + bY$  の平均を答えなさい。(  $a, b$  は任意の実数) (5 点)
  - (2)  $aX + bY$  の分散を答えなさい。(5 点)
  - (3)  $X$  と  $X^2$  の相関係数を答えなさい。(10 点)
  - (4)  $X$  と  $Y$  の相関係数を答えなさい。(10 点)
  - (5)  $X$  と  $X + Y$  の相関係数を答えなさい。(10 点)
- [2] ゲノム比較では、BBH (bidirectional best hit) の関係にある遺伝子を仮のオーソログ (ortholog) として扱うことがある。ゲノム  $X$  上の遺伝子  $A$  とゲノム  $Y$  上の遺伝子  $B$  が BBH の関係にあるとは、ゲノム  $X$  上の遺伝子  $A$  を問い合わせとして、ゲノム  $Y$  上の遺伝子の類似性検索を行うと遺伝子  $B$  が最高スコアを示すものとして検出され、逆に遺伝子  $B$  を問い合わせとしてゲノム  $X$  上の遺伝子の類似性検索によって、遺伝子  $A$  が最高スコアを示すことをさす。BBH によるオーソログを定義することの問題点を述べなさい。(30 点)
- [3] 種は異なるが同じタンパク質をコードしている遺伝子（例えば、ヒトとマウスのヘモグロビン  $\alpha$  の遺伝子）の塩基配列を比較して得られる同義置換率 (= 同義置換数/同義サイト数) と非同義置換率 (= 非同義置換数/非同義サイト数) の大小関係、またそのような関係が生じる原因について、以下のキーワードを含めて説明しなさい。(30 点)
- キーワード：偽遺伝子、正の淘汰