

例題 2.2

表 2.1 の観測値/期待値は、Saxonov 論文の Table 1 では Normalized CpG fraction に相当する。Saxonov 論文の Table 1 では、1 kb upstream regions の GC content は 0.53、Observed CpG fraction は 0.042 であった。この GC content の値を用いて、(1) CC および(2) GA の期待値を小数第 4 位までそれぞれ示せ。(3) 上記 1 kb upstream regions の情報を用いて、CG の観測値/期待値を小数第 3 位まで示せ。

解答

- (1) C の出現確率は $0.53 / 2 = 0.265$ なので、CC の期待値は $0.265 \times 0.265 = 0.0702$ 。
(2) G の出現確率は $0.53 / 2 = 0.265$ 。A の出現確率は $(1 - 0.53) / 2 = 0.235$ 。よって GA の期待値は $0.265 \times 0.235 = 0.0623$ 。(3) CG の出現確率（期待値）は $0.265 \times 0.265 = 0.0702$ 、そして Observed CpG fraction は 0.042 なので、 $0.042 / 0.0702 = 0.598$ 。