

BIOE 489D Bioinformatics Homework #3

1. If the cdf is a probability and hence is dimensionless, what are the units of the pdf?
2. Draw and annotate a normal pdf and show how to illustrate the probability that a random variable will be less than some specific value.
3. Draw and annotate an extreme value pdf and show how to illustrate the probability that a random sequence alignment score will have a score greater than some specific value.
4. Write down the definitions of the mean and standard deviation of a sample.
5. For DNA with match/mismatch 5/-4, and gap costs 12/4, determine the alignment score for:

```
2      CCAACTACAAAGGTTGGCCATTG
      | | | | | | | | | | | | | |
23     CTA ACTAC - GAGTCTGTCTATTG
```

6. For protein with PAM250 and gap costs 12/1, determine the alignment score for

```
Query  4      LTPEEKSAVTALWGKVNVD--EVGGE
        L+   E   V  +WGKV  D    G E
Sbjct  3      LSDGEWQLVLNVWGKVEADIPGHGQE
```