

[Re] Local alignment statistics - Table 1 of Altschul and Gish 1996

Nathan Brouwer

11/6/2019

The original table

The code reproduces the original table and some calculations in the original paper.

Original data

```
#sequence length (this is NOT m*n)
mn      <- c(191,245,314,403,518,665,854,1097,1408,1808,2322,2981)

# mean length of alignments
l.orig  <- c(22.6 ,25.8 ,29.4, 32.4,36.3,40.3,43.9,48.1,51.6,55.1,59.1,63.5)

# log(m*n)
## this is the log of the search space
## eg log(search space)
ln.n.m. <- c(10.25, 10.78,11.30,11.83,12.35,12.87,13.39,13.91,14.43,14.94,15.45,15.96)

# u (mu)
u.orig  <- c(26.45,28.31,30.21,32.04,33.92,35.94,37.84,39.75,41.71,43.54,45.53,47.32)

# lambda
lambda.orig <- c(0.298,0.286,0.282,0.275,0.279,0.273,0.272,0.275,0.268,0.271,0.267,0.270)

# K
K.orig   <- c(0.073,0.055,0.051,0.041,0.048,0.041,0.040,0.046,0.036,0.041,0.035,0.040)
```

Create table

```
table1 <- data.frame(mn = mn,
                     l = l.orig,
                     ln.nm = round(log(mn*mn),3),
                     ln.n.m.,
                     u = u.orig,
                     lambda = lambda.orig,
                     K = K.orig)
```

Look at the table

table1

##	mn	l	ln.nm	ln.n.m.	u	lambda	K
## 1	191	22.6	10.505	10.25	26.45	0.298	0.073

```
## 2 245 25.8 11.003 10.78 28.31 0.286 0.055
## 3 314 29.4 11.499 11.30 30.21 0.282 0.051
## 4 403 32.4 11.998 11.83 32.04 0.275 0.041
## 5 518 36.3 12.500 12.35 33.92 0.279 0.048
## 6 665 40.3 13.000 12.87 35.94 0.273 0.041
## 7 854 43.9 13.500 13.39 37.84 0.272 0.040
## 8 1097 48.1 14.001 13.91 39.75 0.275 0.046
## 9 1408 51.6 14.500 14.43 41.71 0.268 0.036
## 10 1808 55.1 15.000 14.94 43.54 0.271 0.041
## 11 2322 59.1 15.500 15.45 45.53 0.267 0.035
## 12 2981 63.5 16.000 15.96 47.32 0.270 0.040
```

Save the table

```
table1 <- write.csv(table1,file = "table1.csv")
```