Statistical inference on SELEX sequencing data

Anthony Aylward

2018-07-28

The `selex` package includes an example of SELEX count data.

```
library(selex)
example_counts
#>    a  c  g  t
#> 1 25 25 25 25
#> 2 50 20 15 15
#> 3 75  8  7  7
#> 4 95  2  1  1
#> 5 95  2  1  1
```

First, fit a multinomial logit regression model to the counts.

```
fit <- selex_multinom(example_counts, weights = c(8, 1, 2, 4, 8), ref = "c")
summary(fit)
#> Call:
#> multinom(formula = counts ~ cycle, weights = weights)
#>
#> Coefficients:
#>   (Intercept)      cycle
#> a 0.011978665  1.0665507
#> g 0.002133017 -0.1765759
#> t 0.002638545 -0.1766294
#>
#> Std. Errors:
#>   (Intercept)      cycle
#> a  0.09531140 0.06128190
#> g  0.09787694 0.09098062
#> t  0.09786460 0.09097057
#>
#> Residual Deviance: 3288.078
#> AIC: 3300.078
```

Then, numerically compute p-values for the coefficients.

```
selex_pvals(fit)
#>         a         g         t
```

```
#> 0.0000000 0.6271237 0.5765662
```