

Announcements

Make sure to sign in on both Google forms: [this one](#) and [this one](#).

Pset 2 due Friday 2/9.



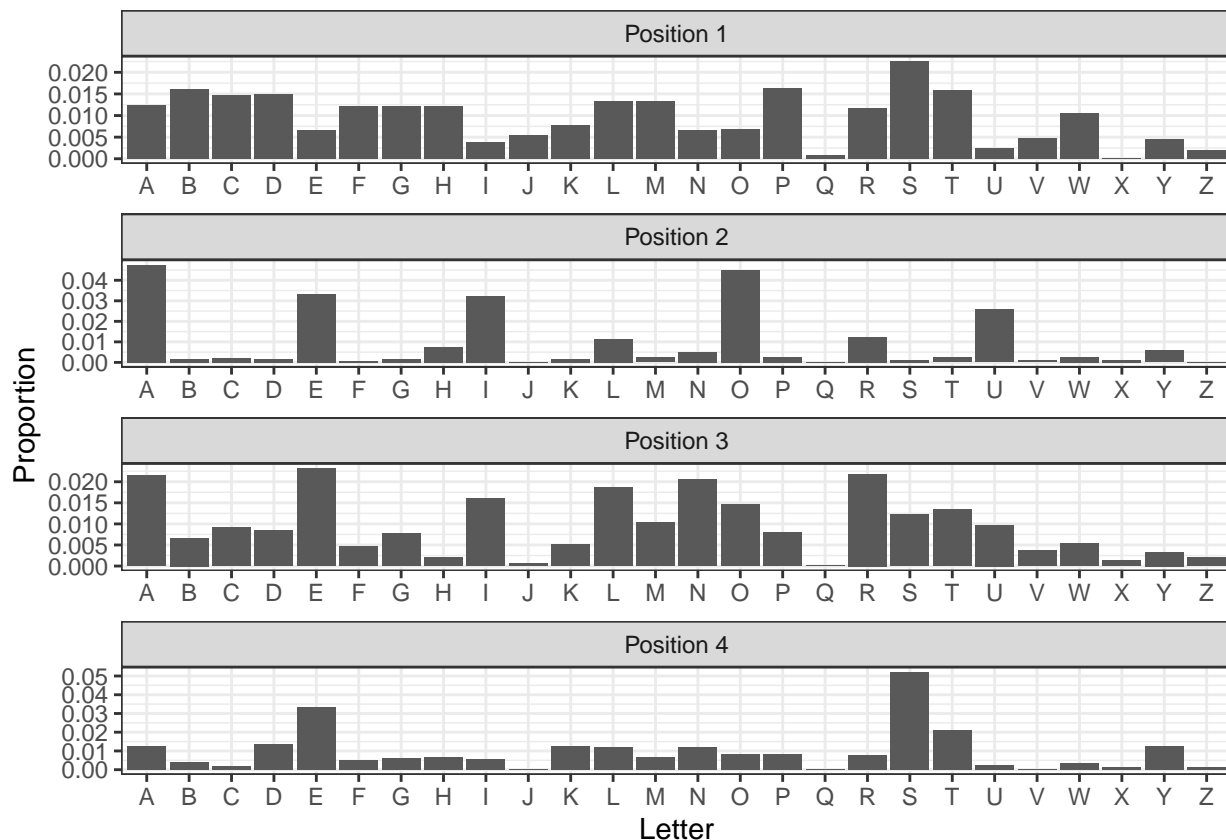
Alphabet Soup

The following question deals with the set of four letter words in the Scrabble dictionary [available here](#).

In this question, we will be modeling the letters in four letter words as draws from a multinomial distribution: $\vec{Y} \sim \text{Mult}_k(4, \vec{p})$. Recall that the Multinomial PMF is

$$P(Y_1 = n_1, \dots, Y_k = n_k) = \frac{n!}{\prod_{i=1}^k n_i!} \prod_{i=1}^k p_i^{n_i}$$

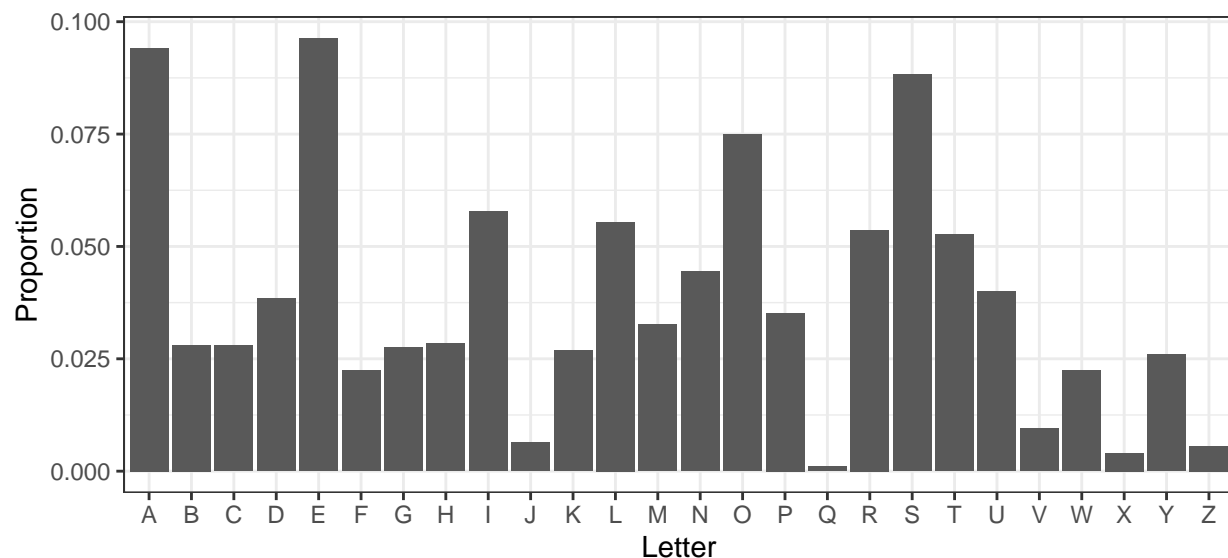
1. What should k be?
2. Suppose we generated a draw from the distribution. What additional step would we have to perform to construct a word from the draw?
3. What assumption of the Multinomial distribution is obviously violated here?



4. Find the likelihood function $L(\vec{p}; \mathbf{Y})$ where \mathbf{Y} is a $n \times k$ matrix with one row for each word and one column for each letter. (E.g. “face” would become a row 1, 0, 1, 0, 1, 1, 0, ..., 0.) Then, find the log likelihood function. What constants can we drop?

5. If we maximized the log likelihood as it stands now, we would end up with $p_j = \infty$ for all p_j . To prevent this, we’ll restrict the sum of the p_j with a [Lagrangian constraint](#). Specifically, find the gradient of $\ell(\vec{p}; \vec{y}) + \lambda(1 - \sum_{j=1}^k p_j)$ (the derivative with respect to each p_j), set it equal to 0, solve for p_j , and then use the fact that $\sum_{j=1}^k p_j = 1$ to solve for λ . Explain in words what this MLE is.

6. The following shows the \hat{p}_j from the data.



7. Suppose you generate a set of 4 letters from the multinomial distribution above and put the 4 letters in a random order. What is the probability of producing the word “stat”? Find this in two ways: Find this in two ways: first, condition on the multinomial. Second, use counting.
9. Find a maximum likelihood estimator for $\text{Cov}(\vec{Y}_{[1]}, \vec{Y}_{[2]})$, the covariance between the number of As and the number of Bs.
10. Find a method of moments estimator for the covariance between the number of As and the number of Bs.
11. Compare the MSEs of these two estimators through simulation with $n = 30$ “words” in each draw.
MLE MSE MOM MSE
4.218867e-05 9.550612e-04

Logistic Logic

The Logistic(s) distribution is defined to be the distribution of $s \log \left(\frac{U}{1-U} \right)$ where $U \sim \text{Unif}(0, 1)$.

1. Find the CDF $F(y)$ of the Logistic distribution. (Hint: Let Y have the Logistic distribution and let U have the Uniform distribution. Then, write Y in terms of U , isolate U , and use the Uniform PDF.)

2. Find the PDF of the Logistic distribution.

3. Let Y have the Logistic distribution. Find $E(Y)$ and $\text{Var}(Y)$. You may use the facts that $\int_0^1 \ln \left(\frac{x}{1-x} \right) dx = 0$ and $\int_0^1 \ln^2 \left(\frac{x}{1-x} \right) dx = \pi^2/3$.

4. Suppose we are measuring vehicle velocities on a congested highway which are distributed $\text{Logistic}(s)$. (It makes sense for the distribution to be symmetric around 0 since the vehicles are equally likely to be going either direction). However, our instrument can only measure velocities in the range of $[-c, c]$ (for any physicists in the room, c does not represent the speed of light). We want to estimate s despite this limitation. Find the likelihood function for s given n_1 observed velocities Y_1, \dots, Y_{n_1} , n_2 velocities less than $-c$ and n_3 velocities more than c .

5. A closed form solution for the maximum likelihood estimator of s does not exist (or at least I wasn't able to find it after an hour of number pushing, and Google doesn't seem to have it either). Instead, consider the following method of moments estimator for s :

$$\hat{s} = \sqrt{\frac{3}{n_1 \pi^2} \sum_{i=1}^{n_1} Y_i^2}$$

Describe the logic behind this estimator.

6. Find the sign of the bias of \hat{s} for s with an argument about how $E\left(\frac{3}{n_1 \pi^2} \sum_{i=1}^{n_1} Y_i^2\right)$ compares to s^2 and Jensen's inequality. (Since the square root function is concave, Jensen's inequality says $E(\sqrt{X}) < \sqrt{E(X)}$.)