Problem 1. You are building a classifier for the sentiment of Russian adjectives. The following 100 adjectives have been sampled from the class of positive adjectives, to use as training data. The adjectives have been analyzed into a stem and suffix.

Adjective	Stem + suffix	Count
красивый	красив + ый	10
красивая	красив + ая	18
красивую	красив + ую	12
приятный	приятн + ый	10
приятная	приятн + ая	32
приятную	приятн + ую	18

a. (5 points) Based on the training data, give estimates for the probabilities of the individual stems and suffixes below.

P(красив- | positive) = P(приятн- | positive) =

P(-ый | positive) = P(-ая | positive) = P(-ую | positive) =

b. (6 points) Suppose that the stem and suffix are conditionally independent, given the class (that is, a naive Bayes model). If the probability estimates you just calculated exactly describe the class of positive adjectives, how many instances of each word would you expect to find in a sample of 100 words drawn from the class of positive adjectives?

красивый

красивая

красивую

приятный

приятная

приятную