

2015 LSA Summer Institute: Computational Psycholinguistics

Homework #1

Klinton Bicknell (Northwestern) and Roger Levy (UCSD)

Tuesday 7 July 2015

Due: Friday 10 July 2015

The following questions relate to ideal probabilistic categorization of instances of the sound categories /b/ and /p/. Assume that a single informative cue (VOT) distinguishes between these categories, and that the distributions of VOT values for these categories can be approximated by Gaussian distributions with means of $\mu_b = 0$ and $\mu_p = 50$. Imagine a context in which the prior probabilities of the two categories differ, $p(/b/) = 0.75$ and $p(/p/) = 0.25$.

For a given VOT value x , we can calculate the posterior distribution on the category c that token came from $p(c|x)$ using Bayes rule:

$$p(c|x) = \frac{p(x|c)p(c)}{p(x)} \quad (1)$$

$$= \frac{p(x|c)p(c)}{\sum_{c'} p(x|c')p(c')} \quad (2)$$

where the prior $p(c)$ is as given above, the likelihood $p(x|c)$ is given by the Gaussian probability density function

$$p(x|c) = \frac{1}{\sigma_c \sqrt{2\pi}} \exp \left[-\frac{(x - \mu_c)^2}{2\sigma_c^2} \right] \quad (3)$$

and the normalizing constant in the denominator is evaluated by summing across all possible hypotheses $c' \in \{/b/, /p/\}$:

$$p(c|x) = \frac{p(x|c)p(c)}{p(x|/b/)p(/b/) + p(x|/p/)p(/p/)} \quad (4)$$

1. Imagine that both categories had equal variances $\sigma_b^2 = \sigma_p^2 = 144$. Under this assumption, calculate the posterior $p(/p/)$ for a VOT value of 25 ms. *Bonus: also calculate the posterior for VOT values of -25, 0, 50, and 75 ms, and plot these values against VOT.*
2. In fact, VOTs for voiceless stops such as /p/ are more variable than those for voiced stops such as /b/. This means that the Gaussian approximations of these categories should have different variances, such as $\sigma_b^2 = 64$ and $\sigma_p^2 = 144$. Assuming these values, calculate the posterior for a VOT value of 25 ms again. How does the categorization curve change? Why? *Bonus: again, calculate the posterior for all VOT values mentioned in problem (1) bonus and make another plot.*
3. Continuing to assume the unequal-variance parameters as in (2), calculate the posterior for the very low VOT of -200 ms. There is some counter-intuitive behavior: what is it? What does this counter-intuitive behavior tell us about the limitations of the model we've been using? *Bonus: extend your bonus graph from (2) with VOT values of -50, -100, -150, and -200 ms to see how this counter-intuitive effect develops.*