Obs & Stats HW 6

Ryan Hofmann

15 March 2017

PDF for a Difference of Gaussians

1. The elements $\sigma_{\delta_i \delta_j}$, with $\delta = x - y$ are defined as

$$\begin{aligned} \sigma^2_{\delta_i \delta_j} &= V[\delta] = E[\delta_i \delta_j] - E[\delta_i]E[\delta_j] \\ &= E[(x_i - y_i)(x_j - y_j)] - E[x_i - y_i]E[x_j - y_j] \end{aligned} \tag{1}$$

   Using the properties of the expectation value, and assuming that x and y are independent,

$$\begin{aligned} \sigma^2_{\delta_i \delta_j} &= E[x_i x_j - x_i y_j - y_i x_j + y_i y_j] - E[x_i - y_i]E[x_j - y_j] \\ &= E[x_i x_j] - E[x_i y_j] - E[y_i x_j] + E[y_i y_j] - (E[x_i] - E[y_i])(E[x_j] - E[y_j]) \\ &= E[x_i x_j] + E[y_i y_j] - E[x_i]E[x_j] - E[y_i]E[y_j] \\ &= (E[x_i x_j] - E[x_i]E[x_j]) + (E[y_i y_j] - E[y_i]E[y_j]) \\ &= \sigma^2_{x_i x_j} + \sigma^2_{y_i y_j} \end{aligned} \tag{2}$$

   In other words, the variance of an arithmetic sum or difference of uncorrelated variables is the arithmetic sum of the individual variances. This is the vector generalization of the simple error propagation formula introduced in undergrad physics.

2. I am not sure how to proceed here. I assume I need to write down the PDF for δ, then substitute the x and y stuff, but I am unclear as to how that works with vectors, and I have failed to find anything of significant help online. I know we did something similar in homework 5, but that was just a single vector with two scalar components. How do I extend that to N-dimensional vectors?

Gaussian Error Estimation

1. For 10 random samples, the characteristics are:

$$\hat{\mu}_x = 0.130 \qquad \hat{\sigma}^2_{x,\mu_x=0} = 0.506 \qquad \hat{\sigma}^2_x = 0.543 \tag{3}$$

2. Using the standard formulas for variance of the mean and variance,

$$\hat{\sigma}^2_{\hat{\mu}_x} = 0.051 \qquad \hat{\sigma}^2_{\hat{\sigma}^2_x} = 0.101 \tag{4}$$

$$\sigma^2_{\mu_x} = 0.100 \qquad \sigma^2_{\sigma^2_x} = 0.200 \tag{5}$$

   The numerical values are a factor of two smaller than the true values, which is still within a couple sigma.

3. Over 1000 trials, the confidence interval contains the true mean ($\mu_x = 0$) 65.1% of the time. Increasing the number of experiments, the frequency converges to $\sim 65\%$. This is slightly less than the expected value of 68%. This is because the variance of the mean is not distributed as a Gaussian, but rather has a χ^2 distribution.