DRAFT 0.1

Glen Cowan
11 August 2020

Note on Convolution of Gaussians

This note records some properties of models of Gaussian data with a mean that includes an additive bias that is constrained by a Gaussian control measurement.

1 Maximum Likelihood (ML) method

Consider a measurement that consists of two independent quantities: the primary measurement y and a control measurement u. They are both modeled with Gaussian pdfs,

$$p(y|\mu,\theta) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-(y-\mu-\theta)^2/2\sigma_y^2} \,, \tag{1}$$

$$p(u|\theta) = \frac{1}{\sqrt{2\pi}\sigma_u} e^{-(u-\theta)^2/2\sigma_u^2} \,. \tag{2}$$

with means $E[y] = \mu+\theta$, $E[u] = \theta$, and variances σ_y^2 and σ_u^2 as shown. Here μ is the parameter of interest and θ is a nuisance parameter that represents an additive bias. Suppose that σ_y and σ_u are known.

The likelihood function is the product of the two Gaussian terms,

$$L(\mu,\theta) = p(y,u|\mu,\theta) = p(y|\mu,\theta)p(u|\theta) \,, \tag{3}$$

which gives the log-likelihood (up to an additive constant) of

$$\ln L(\mu,\theta) = -\frac{1}{2}\left[\frac{(y-\mu-\theta)^2}{\sigma_y^2} + \frac{(u-\theta)^2}{\sigma_u^2}\right] \,. \tag{4}$$

Setting the derivatives of $\ln L$ with respect to μ and θ gives the maximimum-likelihood estimators (MLEs)

$$\hat{\mu} = y - u \,, \tag{5}$$

$$\hat{\theta} = u \,. \tag{6}$$

Since the variances $V[y] = \sigma_y^2$ and $V[u] = \sigma_u^2$ are taken as known, the variance of $\hat{\mu}$ is

$$V[\hat{\mu}] = V[y-u] = V[y] + (-1)^2 V[u] = \sigma_y^2 + \sigma_u^2 \,, \tag{7}$$

or equivalently we can write the standard deviation of $\hat{\mu}$ as $\sigma_{\hat{\mu}} = \sqrt{\sigma_y^2 + \sigma_u^2}$. We refer to this as the ML result: the two "error" sources σ_y and σ_u add in quadrature.

2 Marginal likelihood method

As an alternative to the ML procedure of Sec. 1, we can treat the nuisance parameter θ in a Bayesian sense and assign to it a prior pdf $\pi(\theta)$. Suppose we take the original prior (the "ur-prior"), before the control measurement, to be a constant,

$$\pi_0(\theta) = \text{const.} \tag{8}$$

Then given the Gaussian distributed measurement u, the knowledge about θ is updated by Bayes' theorem to become

$$\pi(\theta) \propto p(u|\theta)\pi_0(\theta) \ . \tag{9}$$

Using a Gaussian distribution for u and normalizing as a pdf for θ gives again a Gaussian,

$$\pi(\theta) = \frac{1}{\sqrt{2\pi}\sigma_u} e^{-(\theta-u)^2/\sigma_u^2} = \frac{1}{\sqrt{2\pi}\sigma_\theta} e^{-\theta^2/\sigma_\theta^2} \tag{10}$$

Here in the fianl equality we substituted $u = 0$, i.e. the the best estimate for the bias is zero, otherwise the measurement y would be corrected to account for it. In addition the standard deviation σ_u was relabled σ_θ as this quantity now reflects the (Bayesian) uncertainty in θ.

Next, the prior of Eq. (10) can be used with the Gaussian distributed y to find the marginal likelihood

$$\begin{aligned} L_{\rm m}(\mu) &= p(y|\mu) = \int p(y|\mu,\theta)\pi(\theta)\, d\theta \\ &= \int \frac{1}{\sqrt{2\pi}\sigma_y} e^{-(y-\mu-\theta)^2/2\sigma_y^2} \frac{1}{\sqrt{2\pi}\sigma_u} e^{-\theta^2/\sigma_u^2}\, d\theta \ . \end{aligned} \tag{11}$$

The integral in Eq. (11) can be carried out by combining the arguments of the exponential terms and completing the square. One obtains

$$L_{\rm m}(\mu) = \frac{1}{\sqrt{2\pi(\sigma_y^2+\sigma_u^2)}} \exp\left[-\frac{1}{2}\frac{(y-\mu)^2}{\sigma_y^2+\sigma_\theta^2}\right] \ . \tag{12}$$

The marginal model defined by Eq. (12) corresponds to an average of those defined by the $p(y|\mu,\theta)$ in Eq. (1) with the average taken with respect to the prior $\pi(\theta)$ from Eq. (10). The ML estimator from the marginal model and its variance are found as before to be

$$\hat{\mu}_{\rm m} = y \ , \tag{13}$$

$$V[\hat{\mu}_{\rm m}] = \sigma_y^2 + \sigma_\theta^2 \ . \tag{14}$$

As in the original ML case one sees that the effect of the uncertan nuisance parameter θ is to increase the variance by $\sigma_y^2 \to \sigma_y^2 + \sigma_\theta^2$.