DRAFT 0.0

Glen Cowan
14 May 2014

Comment on "related parameters"

It can happen that the model for two independent measurements, say, x and y, contain the same parameter of interest μ and a common nuisance parameter θ, such as the jet-energy scale. To combine the measurements one constructs the full likelihood

$$L(\mu, \theta) = P(x|\mu, \theta)P(y|\mu, \theta) \ . \tag{1}$$

Although the parameter θ is thought of initially as being common to the two measurements, this may not be a good approximation. For example, the two analyses may use jets with different angles and energies, so a single energy-scale parameter θ may not represent an accurate model.

One way of extending the model is to assume that the appropriate nuisance parameter for one of the measurements, say, x, is θ and for the other, y, there is a different value θ'. One can propose a relation between the two, such as

$$\theta' = \theta + \varepsilon \ , \tag{2}$$

where ε is an additional nuisance parameter, which we expect to be small.

The frequentist approach to the problem is to treat the best estimates for θ and ε as if they were measured quantities (they may or may not actually result from real measurements). Suppose these values are $\tilde{\theta}$ and $\tilde{\varepsilon}$. Here tildes are used instead of hats for the estimates because the hats will be used later with a different meaning in the profile likelihood ratio.

One might model $\tilde{\theta}$ and $\tilde{\varepsilon}$ as, for example, independent and Gaussian distributed, i.e.,

$$p(\tilde{\theta}, \tilde{\varepsilon}|\theta, \varepsilon) = \text{Gauss}(\tilde{\theta}|\theta, \sigma_{\tilde{\theta}})\text{Gauss}(\tilde{\varepsilon}|\varepsilon, \sigma_{\tilde{\varepsilon}}) \ , \tag{3}$$

where $\sigma_{\tilde{\theta}}$ and $\sigma_{\tilde{\varepsilon}}$ are the standard deviations (or "nominal errors") for the estimates of θ and ε. Since the starting point was that the two parameters θ and θ' represent the same thing, one would usually take $\tilde{\varepsilon} = 0$.

For this model it is easy to work out the covariance between $\tilde{\theta}$ and $\tilde{\theta}'$. We are treating $\tilde{\theta}$ and $\tilde{\varepsilon}$ as independent, so therefore we have

$$\text{cov}[\tilde{\theta}, \tilde{\theta}'] = \text{cov}[\tilde{\theta}, \tilde{\theta} + \tilde{\varepsilon}] = \text{cov}[\tilde{\theta}, \tilde{\theta}] + \text{cov}[\tilde{\theta}, \tilde{\varepsilon}] = \sigma^2_{\tilde{\theta}} \ . \tag{4}$$

Furthermore we have the variance of $\tilde{\theta}'$, $\sigma^2_{\tilde{\theta}'} = \sigma^2_{\tilde{\theta}} + \sigma^2_{\tilde{\varepsilon}}$, so the correlation coefficient for $\tilde{\theta}$ and $\tilde{\theta}'$ is

$$\rho[\tilde{\theta}, \tilde{\theta}'] = \frac{\text{cov}[\tilde{\theta}, \tilde{\theta}']}{\sigma_{\tilde{\theta}}\sigma_{\tilde{\theta}'}} = \frac{1}{\sqrt{1 + \frac{\sigma^2_{\tilde{\varepsilon}}}{\sigma^2_{\tilde{\theta}}}}} \ . \tag{5}$$

This latter relation is not in fact necessary; the entire analysis can proceed using the model

$$L(\mu, \theta, \varepsilon) = P(x|\mu, \theta)P(y|\mu, \theta, \varepsilon)p(\tilde{\theta}, \tilde{\varepsilon}|\theta, \varepsilon) \; , \tag{6}$$

where here the joint distribution for $\tilde{\theta}$ and $\tilde{\varepsilon}$ has been written symbollically as p; this can be a product of Gaussians as above or some other model as appropriate.

Values of μ are then tested using the usual profile likelihood ratio,

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}}, \hat{\hat{\varepsilon}})}{L(\hat{\mu}, \hat{\theta}, \hat{\varepsilon})} \; . \tag{7}$$