Normal times normal is normal!

One of the remarkable properties of a normal distribution is that (as a family) they are invariant under a bunch of operations. One of these operations is multiplying them together (in these of pointwise multiples of normal densities). This is the source of a standard formula for applying meta-analysis, but also has other applications such as to Bayesian analysis of summary statistics, as we explain below.

Pointwise multiples of normal densities The following lemma explain how to pointwise multiply two normal densities, i.e. to work out

$$N(x|\mu_1, \sigma_1^2) \times N(x|\mu_2, \sigma_2^2)$$

for any given point x. It says that the result is proportional to another normal density, and it computes the mean and variance of this density. Moreover, it also calculates the constant of proportionality, which turns out to be computable as yet another normal density - but this time not computed at the point of evaluation, but at the difference of the original distribution means.

The lemma holds in both univariate and multivariate cases; both versions are given below.

Lemma 1. *(Univariate version) If μ_1,μ_2 are two means and σ_1^2, σ_2^2 are variances then:*

$$N(x|\mu_1, \sigma_1^2) \times N(x|\mu_2, \sigma_2^2) = const \times N(x|a, A)$$

where

$$A = \frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} \qquad \textit{and} \qquad a = A \cdot \left(\frac{\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2} \right)$$

Moreover the constant is also computable as a multivariate normal density evaluated at $\mu_1 - \mu_2$,

$$const = N\left(\mu_1 - \mu_2 | 0, \sigma_1^2 + \sigma_2^2\right)$$

NOTE: the lemma can be recursively applied to extend to multiple distributions. Specifically:

$$\prod_{i=1}^{n} N(x|\mu_i, \sigma_i^2) = \text{const } \times N(x|a_n, A_n)$$

where

$$A_n = \frac{1}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}} \quad \text{and} \quad a_n = A_n \cdot \sum_{i=1}^{n} \frac{\mu_i}{\sigma_i^2} \tag{1}$$

Proof. The lemma says this is true for $n = 2$. Moreover if it is true for n, then by the lemma

$$A_{n+1} = \frac{1}{\frac{1}{A_n} + \frac{1}{\sigma_{n+1}^2}} = \frac{1}{\sum_{i=1}^{n+1} \frac{1}{\sigma_i^2}}$$

and

$$a_{n+1} = A_{n+1} \cdot \left(\frac{a_n}{A_n} + \frac{\mu_{n+1}}{\sigma_{n+1}^2} \right) = A_{n+1} \cdot \left(\sum_{i=1}^{n+1} \frac{\mu_i}{\sigma_i^2} \right)$$

which is what we claimed. $\square$

Here is the multivariate version:

Lemma 2. *(Multivariate version) If μ_1,μ_2 are two means and Σ_1, Σ_2 are covariance matrices then:*

$$MVN(x|\mu_1, \Sigma_1) \times MVN(x|\mu_2, \Sigma_2) = const \times MVN(x|a, A)$$

where

$$A = \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1} \qquad \textit{and} \qquad a = A\left(\Sigma_1^{-1}\mu_1 + \Sigma_2^{-1}\mu_2\right)$$

(if the matrices are invertible). Moreover the constant is also computable as a multivariate normal density evaluated at $\mu_1 - \mu_2$,

$$const = MVN\left(\mu_1 - \mu_2 | 0, \Sigma_1 + \Sigma_2\right)$$

Application to fixed-effect meta-analysis In a meta-analysis all we are typically given is a point estimate ($\hat{\beta}_i$) and a standard error (v_i) for a collection of studies $i = 1, \cdots, N$. And the question is how to combine these to make a join estimate.

One way to think of this is that $\hat{\beta}_i, v_i$ form an approximate representation of the likelihood from study i as a normal distribution. (This approximation often is good as long as there is lots of data). If the studies are independent, the combined likelihood should be obtained by multiplying across studies, giving

$$\text{combined likelihood}(x) = \prod_i N(x|\hat{\beta}_i, v_i^2)$$

Now we apply (1) to find that the combined likelihood is again normal with variance and mean:

$$v_{\text{meta}}^2 = \frac{1}{\sum_i \frac{1}{v_i^2}} \quad \text{and} \quad \hat{\beta}_{\text{meta}} = v_{\text{meta}}^2 \cdot \sum_i \frac{\hat{\beta}_i}{v_i^2}$$

This is the most standard way to do meta-analysis. Its full name is *inverse variance-weighted fixed-effect meta-analysis* because:

1. The meta-analysis estimate $\hat{\beta}_{\text{meta}}$ is an average among all the estimates passed in, with each term weighted by the inverse of its variance. (This is the right thing, since a high variance would mean an uncertain estimate that should have low weight).

2. The underlying assumption is that there is one true parameter value β being estimated that is being estimated. This is fixed (assumed not to vary) across studies.

The above derivation makes clear that this meta-analysis is really just the process of forming the combined likelihood function (up to the approximation by normal likelihoods) and estimating from it in the usual way.

Application to Bayes factors We can also apply the lemma to Bayesian analysis, by using a normal prior $MVN(\mu, \Sigma)$ and a normal likelihood $MVN(\hat{\beta}, V)$, where V is the variance-covariance matrix of the loglikelihood. (As discussed in lectures, this is approximately applicable in fairly general settings when data quantities are large). In particular suppose we have a vector of true effects β and want to compare the model $\beta \sim MVN(\mu_1, \Sigma_1)$ to a model where all the effects are zero. Having seen some data, we want to know how probable M_1 is relative to M_0. This can be figured out using Bayes theorem written:

$$\begin{aligned}\frac{P(M_1|\text{data})}{P(M_0|\text{data})} &= \frac{P(\text{data}|M_1)}{P(\text{data}|M_0)} \cdot \frac{P(M_1)}{P(M_0)} \\ &= \frac{\int_\beta P(\text{data}|\beta, M_1)P(\beta|M_1)}{P(\text{data}|\beta = 0)} \cdot \frac{P(M_1)}{P(M_0)}\end{aligned}$$

The first term is the Bayes factor, given the change odds in the data for model 1 relative to the null model. The last term is the prior odds of model 1 relative to the null model. (We must choose this).

The term being integrated is a product of normals so we can use the lemma. The integral works out to be the constant term in the lemma = i.e. $MVN\left(\hat{\beta} - \mu; 0, \Sigma + V\right)$. Hence the BF can be computed as:

$$\text{Bayes factor} = \frac{MVN(\hat{\beta} - \mu; 0, \Sigma + V)}{MVN(\hat{\beta}; 0, V)}$$

This is one Bayesian analysis that is easy to implement !