Here are some BIC citations of interest:

Wasserman (2000) in J of Math Psych gives a formula (27) which implies that a BIC-related difference between two models is logBij where B is the Bayes factor for choosing between model i and j. Wasserman's (27) says that logBij is approximately what Mplus calls minus 1/2 BIC. This means that 2log Bij is in the Mplus BIC scale apart from the ignorable sign difference.

Kass and Raftery (1995) in J of the Am Stat Assoc gives rules of evidence on page 777 for 2log_e Bij which say that >10 is very strong evidence in favor of the model with largest value.

So, to conclude, this says that an Mplus BIC difference > 10 is strong evidence against the model with the highest Mplus BIC value (I hope I got that right).

Raftery has a Soc Meth chapter:

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. Sociological Methodology, 25, 111-163.

that talks about Bij from a SEM perspective.

There's also a good discussion about this here:

http://www.statmodel.com/discussion/messages/23/2232.html?1209409498

Aaron M. Thompson posted on Monday, August 30, 2010 - 6:14 pm
Dr. Muthen,

Thank you for the great resources. I hate to belabor this point, but I am a stickler for accuracy and I am an intervention researcher - not a mathematician. Last summer, I took an ICPSR course and learned about the Raferty citation for calculating a more interpretable BIC using the Mplus chi2 in the formula "chi2-df (ln(N))". This calculation produces a BIC that is comparable across nonnested models following the Raferty rule >10.

However, as Mplus LTA output does not give a chi2, but only a LgLkd chi2, I am assuming that I can not use this statistic in this calculation, am I correct in my understanding?

Therefore, following your suggestions using the results from my models, 2ln of the BIC (19355.681) for model i = 19.741, 2ln of the BIC (18956.107) for model j = 19.699. The difference between Bij is less than 10. Thus, according to your explanation, this is is "strong" statistical evidence for retaining the more parsimonious model with the larger BIC (i.e. keep model i over model j). Is my interpretation of this accurate?

Thanks again for your time and consideration.

Bengt O. Muthen posted on Tuesday, August 31, 2010 - 2:36 pm
Comparing models using the formula "chi2-df (ln(N))" is the same as using the Mplus BIC = -2logL + p*ln(N), where p is the number of parameters. Note that

chi2 = -2(logL_a - logL_b),

where a is a model nested within b. In the usual SEM case b is the totally unrestricted model called H1. Note also that

df = p_b - p_a,

where p is the number of parameters.

So when you look at the difference between the BIC of two models using the formula chi2-df (ln(N)) there is a canceling out of the terms -2logL_b and of the terms p_b*ln(N). This means that BIC differences are the same for both formulas. And this means that we should view a BIC difference > 10 as strong evidence that the model with lower BIC is better.