Model Selection Cheat Sheet

Akaike Information Criteria

$$2 \times \text{log-likelihood} + 2p$$

BIC is consistent in selecting the true model. AIC is minimax-rate optimal for estimating the regression function.

You cannot have both consistency and minimax-rate optimality (Yang 2005)

References

[1] Yuhong Yang. Can the Strengths of AIC and BIC Be Shared? A Conflict between Model Indentification and Regression Estimation. *Biometrika*, 92(4):937–950, 2005.