Lecture Notes VI – Model selection by AIC and BIC

Marina Meilă
mmp@stat.washington.edu

Department of Statistics
University of Washington

April, 2021

AIC and BIC

AIC and BIC

Hold for

- parametric $\mathcal{F} = \{f_\theta\}$ with θ a vector of parameters, e.g. (μ, σ^2), γ, a, b (for logistic)
- log-likelihood $I(\theta) = -\ln P(x^{1:n}|f_\theta)$
- $f^{ML} \in \mathcal{F}$ estimated by Maximum Likelihood
- (for BIC: $\frac{\partial^2 I}{\partial \text{parameters}}$ non-singular at f^{ML})

Akaike's Information Criterion (AIC)

$$AIC(f^{ML}) = I(f^{ML}) - d, \tag{1}$$

where $d = \#\text{parameters}(f)$, and $n =$ the size of $\mathcal{D}$.

The Bayesian Information Criterion (BIC)

$$BIC(f^{ML}) = -n\hat{L}(f^{ML}) - \frac{d}{2}\ln n, \tag{2}$$

with $d = \#\text{parameters}(f)$