Re: Computation of AIC and AIC with weights

Source: http://sci.tech−archive.net/Archive/sci.stat.math/2007−09/msg00447.html

- *From*: joseph.yarmus@xxxxxxxxxx
- *Date*: Thu, 27 Sep 2007 06:20:04 −0700

On Sep 26, 11:22 am, joseph.yar...@xxxxxxxxxx wrote:

> This message was also posted to R−help.
> In accordance with Venables and Ripley, SAS documentation and other sources AIC with sigma^2 unknown is calculated as:
> AIC = −2LL + 2* #parameters = n log(RSS/n) + 2p
> For the fitness data: (http://support.sas.com/ctx/samples/index.jsp?sid=927), SAS gets an AIC of 64.534 with model oxygen = runtime. (SAS STAT User's Guide. Chapter 61. pp 3956, the REG Procedure). This value of AIC accords with n log(RSS/n) + 2p and p = 2.
>
> When I run the same problem in R ver 2.5.1, I get
>
>> rt.glm =glm(oxy ~ runtime, data=fitness)
>> rt.glm
>
> Call: glm(formula = oxy ~ runtime, data = fitness)
>
> Coefficients:
> (Intercept) runtime
> 82.422 −3.311
>
> Degrees of Freedom: 30 Total (i.e. Null); 29 Residual
> Null Deviance: 851.4
> Residual Deviance: 218.5 AIC: 154.5
>
> I get very close to what R gets if the constant term is included in −2LL, (31*Log(2*pi)+n−1), divide RSS by n−1 and the number of parameters is 3 (the predictor, the intercept and the error term)> 31 * (log(2*pi)+log(sum(rt.glm$res^2)/30)) + 30 + 2 * 3
> [1] 154.5248
>
>> AIC(rt.glm)
>
> [1] 154.5083

> 3 questions:
> 1) Why the discrepancy between SAS and R?
> 2) Why the slight difference between my calculation in R and R's AIC?
> 3) How should AIC be computed if row weights are used in the linear model?
>
> Thanks!
>
> −joe yarmus

Digging into the R−code behind AIC for linear models, I see:
AIC = nobs * (log(dev/nobs * 2 * pi) + 1) + 2 − sum(log(wt)) + 2 * p
dev = sum(wt * (y − mean(y))^2
For the unweighted case, this translates directly to −2LL with the number of parameters including both intercept and error term (as represented by the constant + 2) and the unknown
sigma−squared = sum((y − mean(y)^2)/ nobs (rather than nobs−1).
However, with weights, I am at a loss to understand the expression, because, given
−2LL = nobs * (log(2 * pi * sigma^2) + sum(wt * (y − mean(y))^2/ sigma^2
if sigma^2 = sum(wt * (y − mean(y))/sum(wt)
then
−2LL = nobs * (log(2 * pi *dev/nobs) + log(nobs) − log(sum(wt)) + sum(wt)
so if wt = 1 all is fine because
−2LL = nobs * (log(2 * pi * dev/nobs) + 1)
What am I missing? Thanks!

−joe yarmus

.