estimates stats — Model-selection statistics

Description

`estimates stats` reports model-selection statistics, including the Akaike (AIC), consistent Akaike's (CAIC), corrected Akaike's (AICc), and Schwarz's Bayesian (BIC) information criteria. These measures are appropriate for maximum likelihood models.

If `estimates stats` is used for a non–likelihood-based model, such as `qreg`, missing values are reported.

Quick start

Display AIC and BIC for last estimation command

```
estimates stats
```

Display AIC and BIC for stored estimates `m1` and `m2`

```
estimates stats m1 m2
```

Same as above, but display CAIC instead of AIC

```
estimates stats m1 m2, aicconsistent
```

Same as above, but display AICc instead of AIC

```
estimates stats m1 m2, aiccorrected
```

Same as above, but display all AIC, BIC, AICc, and CAIC

```
estimates stats m1 m2, all
```

Specify $N = 1{,}000$ for calculation of BIC

```
estimates stats, n(1000)
```

Same as above, but use 10 degrees of freedom

```
estimates stats, n(1000) df(10)
```

Menu

Statistics > Postestimation

Syntax

```
estimates stats [namelist] [ , options ]
```

where *namelist* is a name, a list of names, `_all`, or `*`. A name may be `.`, meaning the current (active) estimates. `_all` and `*` mean the same thing.

options	Description
`aiccorrected`	report AICc instead of AIC
`aicconsistent`	report CAIC instead of AIC
`all`	report all four information criteria: AIC, BIC, AICc, and CAIC
`n(#)`	specify N to be used in calculating BIC, AICc, and CAIC; see [R] **IC note**
`df(#)`	specify degrees of freedom k to be used in calculating AIC, BIC, AICc, and CAIC
`icdetail`	produce a table showing the type of N used in BIC, AICc, and CAIC calculation

`collect` is allowed; see **[U] 11.1.10 Prefix commands**.

Options

`aiccorrected` specifies that AICc be computed instead of AIC. This information criterion is a second-order approximation and is recommended for small sample sizes.

Only one of `aiccorrected`, `aicconsistent`, or `all` is allowed.

`aicconsistent` specifies that CAIC be computed instead of AIC. This information criterion is a consistent version of AIC; that is, the probability of selecting the "true model" approaches 1 as sample size increases.

Only one of `aicconsistent`, `aiccorrected`, or `all` is allowed.

`all` produces a table showing all four information criteria: AIC, BIC, AICc, and CAIC.

Only one of `all`, `aiccorrected`, or `aicconsistent` is allowed.

`n(#)` specifies N to be used in calculating BIC, AICc, and CAIC; see [R] **IC note**.

`df(#)` specifies degrees of freedom k to be used in calculating AIC, BIC, AICc, and CAIC. By default, k is the number of estimated parameters.

`icdetail` produces a table showing the type of N used in BIC, AICc, and CAIC calculations. Most estimation commands use the number of observations in the estimation sample for the information criteria. For some models, however, other types of N, such as the number of cases in choice models, should be used. When the default table of `estimates stats` contains more than one type of N, specifying `icdetail` allows you to see the different types of N used for BIC, AICc, and CAIC.

Remarks and examples

If you type `estimates stats` without arguments, a table for the most recent estimation results will be shown:

```
. use https://www.stata-press.com/data/r19/auto
(1978 automobile data)
. logistic foreign mpg weight displ
  (output omitted)
. estimates stats
Akaike's information criterion and Bayesian information criterion
```

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	74	-45.03321	-20.59083	4	49.18167	58.39793

```
Note: BIC uses N = number of observations. See [R] IC note.
```

Regarding the note at the bottom of the table, N is an ingredient in the calculation of BIC, AICc, and CAIC; see [R] **IC note**. The note changes if you specify the `n()` option, which tells `estimates stats` what N to use. By default, N is the number of observations used in fitting the model.

Regarding the table itself, `ll(null)` is the log likelihood for the constant-only model, `ll(model)` is the log likelihood for the model, `df` is the number of degrees of freedom, and `AIC` and `BIC` are the Akaike and Bayesian information criteria, respectively.

Models with smaller values of an information criterion are considered preferable.

`estimates stats` can compare estimation results:

```
. use https://www.stata-press.com/data/r19/auto
(1978 automobile data)
. logistic foreign mpg weight displ
  (output omitted)
. estimates store full
. logistic foreign mpg weight
  (output omitted)
. estimates store sub
. estimates stats full sub
Akaike's information criterion and Bayesian information criterion
```

Model	N	ll(null)	ll(model)	df	AIC	BIC
full	74	-45.03321	-20.59083	4	49.18167	58.39793
sub	74	-45.03321	-27.17516	3	60.35031	67.26251

```
Note: BIC uses N = number of observations. See [R] IC note.
```

You can use the `all` option to compare the models based on all four information criteria

```
. estimates stats full sub, all
Information criteria
```

Model	N	ll(null)	ll(model)	df
full	74	-45.03321	-20.59083	4
sub	74	-45.03321	-27.17516	3

```
Note: BIC, AICc, and CAIC use N = number of observations.
      See [R] IC note.
```

Model	AIC	BIC	AICc	CAIC
full	49.18167	58.39793	49.76138	62.39793
sub	60.35031	67.26251	60.69317	70.26251

```
Legend: AIC is Akaike's information criterion.
        BIC is Bayesian information criterion.
        AICc is corrected Akaike's information criterion.
        CAIC is consistent Akaike's information criterion.
```

All four information criteria suggest that the `full` model is preferable.

Stored results

`estimates stats` stores the following in `r()`:

Matrices

`r(S)` matrix with columns (`N`, `ll(null)`, `ll(model)`, `df`, and information criteria) and rows corresponding to models in the table

Methods and formulas

See [R] **IC note**.

Also see

[R] **estimates** — Save and manipulate estimation results