

GammaLib - Action #2209

Add WSTAT as option to GCTAOnOffObservation

10/09/2017 03:31 PM - Knödlseeder Jürgen

Status:	Closed	Start date:	10/09/2017
Priority:	Normal	Due date:	
Assigned To:	Tibaldo Luigi	% Done:	100%
Category:		Estimated time:	0.00 hour
Target version:	1.5.0		
Description			
Allow for an optional WSTAT statistics in GCTAOnOffObservation			

History

#1 - 10/17/2017 03:58 PM - Tibaldo Luigi

- Status changed from New to In Progress
- Assigned To set to Tibaldo Luigi

WStat corresponds to the case of Poisson signal with measured Poisson background, as opposed to the case already implemented of Poisson signal with modeled Poisson background. See for example the [XSpec description](#) (see Poisson data with Poisson background).

#2 - 10/20/2017 12:13 PM - Tibaldo Luigi

- % Done changed from 0 to 90

WSTAT is implemented and tested in branch 2209-add-WSTAT.

Only issue: I can compile the ctools on this gammalib version but there is a test failure related to csspec

```
1 errors occurred in module "csspec":
Error in Test csspec from Python:
**** ERROR in GPythonTestSuite::test: <class 'TypeError'>in method 'GARf__getitem__', argument 2 of type 'int'
(N16GPythonException10test_errorE)
```

Jürgen, could you please take a look at it and advise? I guess we need to implement anyway in ctlike and csspec the possibility to toggle between different statistics for the On/Off observations.

#3 - 10/21/2017 02:06 PM - Knödlseeder Jürgen

- Target version set to 1.5.0

I integrated the code into the devel branch.

I do not have a GARf__getitem__ error in csspec, I guess this is due to an older version of ctools that you may have on your side.

However, I get the following error running the ctools tests:

```
**** ERROR in GPythonTestSuite::test: <type 'exceptions.ValueError'>**** ERROR in GCTAOnOffObservation::likelihood(GModels&,
GOptimizerPars&, GMatrixSparse&, GVector&, double&, double&): Invalid value. Invalid statistics "POISSON" encountered. Either specify
"POISSON-MODELEDBKG" or "POISSON-MEASUREDBKG". (N16GPythonException10test_errorE)
```

The reason for this is simple: by default the statistic in the ctlike tool is POISSON, and since you now expect POISSON-MODELEDBKG or POISSON-MEASUREDBKG an error occurs. This cannot be changed in the unit tests since csspec has so far no statistic parameter.

For the time being I changed the POISSON-MODELEDBKG to POISSON (and also accept CSTAT) and POISSON-MEASUREDBKG to WSTAT. I'm wondering whether we should change the naming of the statistic, since POISSON and GAUSSIAN are in fact not very accurate descriptions. We could use:

- CSTAT instead of POISSON
- WSTAT
- CHISQ instead of GAUSSIAN

Except of ctlike, none of the other likelihood tools allows actually to set the statistic. We need to add parameters to the tools for that. And also: we need to add the capability to GObservations::read to extract the statistic from an XML file so that one can mix observations with different statistics.

I created #2237 for the addition of the parameters (should by the way rename the parameter to statistic) and #2238 for the change of the GObservations::read method.

#4 - 10/23/2017 12:43 PM - Tibaldo Luigi

Thanks Jürgen.

For the naming scheme, I would avoid CSTAT and WSTAT. They are opaque names, and, even more importantly, ambiguous. For example, the statistics called CSTAT in XSPEC is the same called WSTAT in Sherpa.

My preferred names would be: Poisson, Gaussian (or chisq), Poisson-measuredbkg (or Poisson-profilebkg).

If this is already integrated in devel shall we close the action item?

#5 - 10/23/2017 03:01 PM - Knödseder Jürgen

We can use other names, but they should be as short as possible to avoid typos and to be also easily citable in a publication. Ideally it would be good to reuse names that are already used by others.

As far as I can see, Xspec is using cstat for both statistics, which is unfortunate (see <https://heasarc.gsfc.nasa.gov/xanadu/xspec/manual/XSappendixStatistics.html>). Sherpa is making the distinction between both, so I would tend to follow this path. In addition, wstat is not known to Xspec, hence one can argue that there is no formal conflict there.

I think the fundamental difference between CSTAT and WSTAT is that CSTAT assumes an arbitrary background model while WSTAT assumes that the background rate per steradian is identical to the On and the Off region (only by making this assumption one can solve the equation). I noted that one could in principle extent the statistics to take into account a difference in the background rate per steradian as derived from a background template.

#6 - 10/24/2017 01:29 PM - Tbaldo Luigi

I see the point of having shorter names. We could stay with Sherpa names, that is also the standard in HESS.

I don't think that to use WSTAT one needs to assume that the background rate per sr is the same in the On and Off regions. This is true only for the reflected regions background estimation method, but would not be true, e.g., for the ring background estimation method (in the latter case the alpha parameter takes into account a different background rate with offset from the center of the camera).

The main difference between CSTAT and WSTAT as far as I understand is that for CSTAT you need to assume a spectral shape for the background for both the On and Off regions, while for WSTAT you just assume the spectral shape is the same, and then the background rate in each spectral bin as a nuisance parameter that you profile over in the likelihood computation.

#7 - 10/24/2017 02:58 PM - Knödlseher Jürgen

user#266 wrote:

I see the point of having shorter names. We could stay with Sherpa names, that is also the standard in HESS.

I don't think that to use WSTAT one needs to assume that the background rate per sr is the same in the On and Off regions. This is true only for the reflected regions background estimation method, but would not be true, e.g., for the ring background estimation method (in the latter case the alpha parameter takes into account a different background rate with offset from the center of the camera).

Indeed, you can take differences in rate into account in the alpha parameter, but for the time being I think we only use the solid angle to compute alpha, hence we assume that the rate per solid angle is the same.

The main difference between CSTAT and WSTAT as far as I understand is that for CSTAT you need to assume a spectral shape for the background for both the On and Off regions, while for WSTAT you just assume the spectral shape is the same, and then the background rate in each spectral bin as a nuisance parameter that you profile over in the likelihood computation.

I think you can only do this by making an implicit assumption of the background ratio between On and Off regions. In the WSTAT formula you have the same mu-term (corrected by alpha) in the On and Off region.

#8 - 10/24/2017 03:00 PM - Tbaldo Luigi

The main difference between CSTAT and WSTAT as far as I understand is that for CSTAT you need to assume a spectral shape for the background for both the On and Off regions, while for WSTAT you just assume the spectral shape is the same, and then the background rate in each spectral bin as a nuisance parameter that you profile over in the likelihood computation.

I think you can only do this by making an implicit assumption of the background ratio between On and Off regions. In the WSTAT formula you have the same mu-term (corrected by alpha) in the On and Off region.

Isn't this the case also in CSTAT?

#9 - 10/24/2017 03:02 PM - Knödseder Jürgen

user#266 wrote:

The main difference between CSTAT and WSTAT as far as I understand is that for CSTAT you need to assume a spectral shape for the background for both the On and Off regions, while for WSTAT you just assume the spectral shape is the same, and then the background rate in each spectral bin as a nuisance parameter that you profile over in the likelihood computation.

I think you can only do this by making an implicit assumption of the background ratio between On and Off regions. In the WSTAT formula you have the same mu-term (corrected by alpha) in the On and Off region.

Isn't this the case also in CSTAT?

Indeed. I just meant that there is no black magic behind WSTAT that avoids making an assumption about the background shape.

#10 - 10/24/2017 03:06 PM - Tibaldo Luigi

The main difference between CSTAT and WSTAT as far as I understand is that for CSTAT you need to assume a spectral shape for the background for both the On and Off regions, while for WSTAT you just assume the spectral shape is the same, and then the background rate in each spectral bin as a nuisance parameter that you profile over in the likelihood computation.

I think you can only do this by making an implicit assumption of the background ratio between On and Off regions. In the WSTAT formula you have the same mu-term (corrected by alpha) in the On and Off region.

Isn't this the case also in CSTAT?

Indeed. I just meant that there is no black magic behind WSTAT that avoids making an assumption about the background shape.

No black magic indeed. WSTAT requires fewer assumptions/free parameters than CSTAT, but still does require some sort of assumption.

#11 - 10/30/2017 10:34 PM - Knödseder Jürgen

- *Status changed from In Progress to Closed*

- *% Done changed from 90 to 100*