XSPEC Tutorial and Statistics

Basic steps for X-ray spectral analysis

Step 1: setup commands

xspec

xspec> data pn_25.grp
xspec> ignore bad
xspec> ignore **-0.3 7.2-**
xspec> cpd /xw
xspec> plot ldata
xspec> show all

Energy: with "."
Otherwise: channels

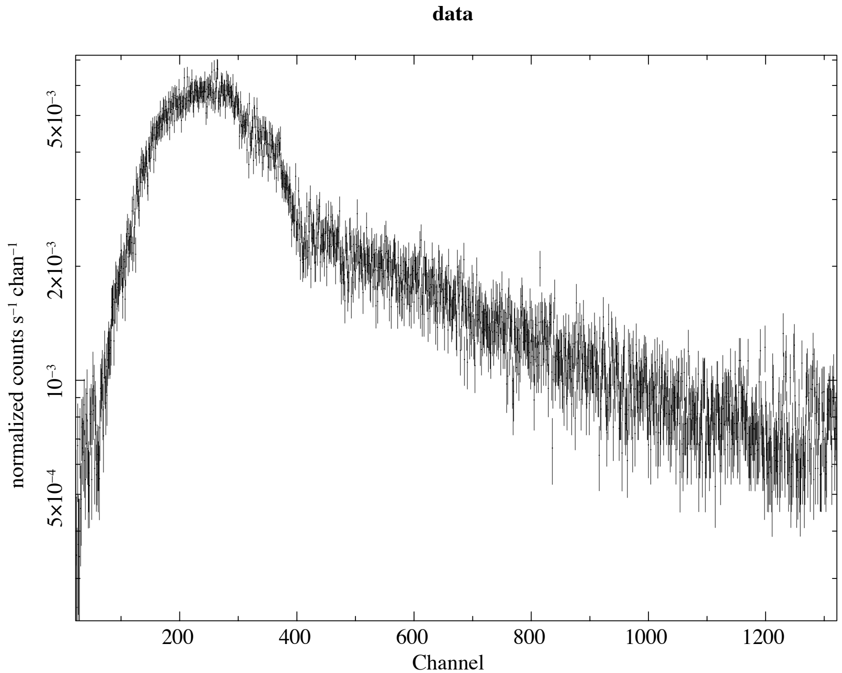
```
         XSPEC version: 12.8.2
  Build Date/Time: Thu Jul 10 09:26:57 2014
```

```
XSPEC12>data pn_25.grp

1 file 1 spectrum
Spectrum 1  Spectral Data File: pn_25.grp
Net count rate (cts/s) for Spectrum:1  2.652e+00 +/- 7.965e-03 (96.8 % total)
 Assigned to Data Group 1 and Plot Group 1
  Noticed Channels:  23-1321
  Telescope: XMM Instrument: EPN  Channel Type: PI
  Exposure Time: 4.441e+04 sec
 Using fit statistic: chi
 Using test statistic: chi
 Using Background File                back_spectrum.fits
  Background Exposure Time: 4.441e+04 sec
 Using Response (RMF) File            pn.rmf for Source 1
 Using Auxiliary Response (ARF) File  pn.arf

 Spectral data counts: 121704
```

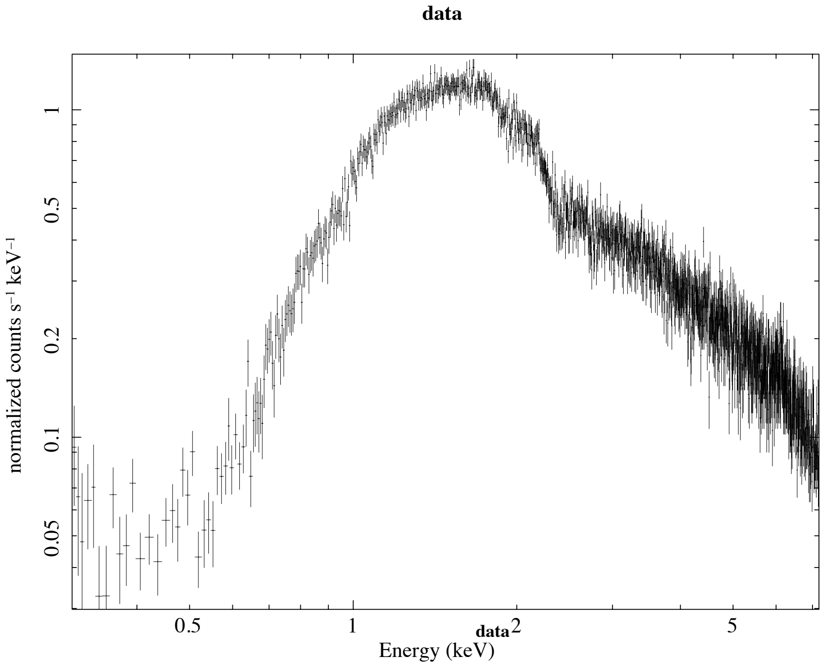
Fraction of src counts/total

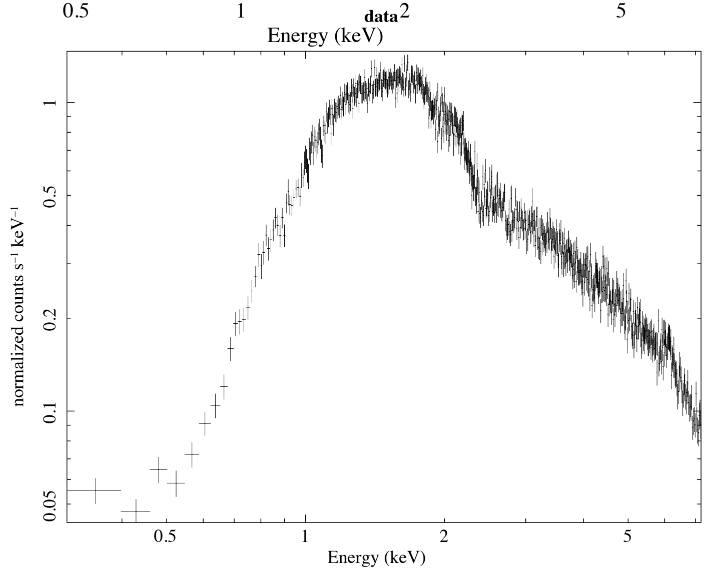
loaded bkg and response files

Channel scale



Energy scale





xspec> setplot rebin 10 30

minimum significance — max number of bins

(just for plotting purposes)

They are linked via the RMF (redistribution matrix file)

Step 2: choice of the models

Additive models

agauss	apec	bapec	bbody	bbodyrad	bexrav
bexriv	bkn2pow	bknpower	bmc	bremss	bvapec
bvvapec	c6mekl	c6pmekl	c6pvmkl	c6vmekl	cemekl
cevmkl	cflow	compLS	compPS	compST	compTT
compbb	compmag	comptb	compth	cplinear	cutoffpl
disk	diskbb	diskir	diskline	diskm	disko
diskpbb	diskpn	eplogpar	eqpair	eqtherm	equil
expdec	ezdiskbb	gadem	gaussian	gnei	grad
grbm	kerrbb	kerrd	kerrdisk	laor	laor2
logpar	lorentz	meka	mekal	mkcflow	nei
npshock	nsa	nsagrav	nsatmos	nsmax	nsmaxg
nsx	nteea	nthComp	optxagn	optxagnf	pegpwrlw
pexmon	pexrav	pexriv	plcabs	posm	powerlaw
pshock	raymond	redge	refsch	rnei	sedov
sirf	smaug	srcut	sresc	step	vapec
vbremss	vequil	vgadem	vgnei	vmcflow	vmeka
vmekal	vnei	vnpshock	vpshock	vraymond	vrnei
vsedov	vvapec	vvgnei	vvnei	vvnpshock	vvpshock
vvrnei	vvsedov	zagauss	zbbody	zbremss	zgauss
zpowerlw					

Moltiplicative models

SSS_ice	TBabs	TBgrain	TBvarabs	absori	acisabs
cabs	constant	cyclabs	dust	edge	expabs
expfac	gabs	heilin	highecut	hrefl	lyman
notch	pcfabs	phabs	plabs	pwab	recorn
redden	smedge	spexpcut	spline	swind1	uvred
varabs	vphabs	wabs	wndabs	xion	zTBabs
zbabs	zdust	zedge	zhighect	zigm	zpcfabs
zphabs	zredden	zsmdust	zvarabs	zvfeabs	zvphabs
zwabs	zwndabs	zxipcf			

Syntax:

M1*M2*(A1+A2+M3*A3)

M=multiplicative model: modifies incident flux

A=additive model: source of emission

Other models

Convolution Models:

cflux	cpflux	gsmooth	ireflect	kdblur	kdblur2
kerrconv	lsmooth	partcov	rdblur	reflect	rgsxsrc
simpl	zashift	zmshift			

Mixing Models:

ascac	projct	suzpsf	xmmpsf

Pile-up Models:

pileup

Example:

model wabs*(powerlaw+gaussian)

$$M(E) = \exp\left[-n_H \sigma(E)\right]$$

$$A(E) = KE^{-\alpha}$$

$$A(E) = K\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(E-E_l)^2}{2\sigma^2}\right)$$

Step 3: model fit and best-fitting solution

Absorption due to our Galaxy: you need to include it in **all** spectral fits

xspec> nh

```
XSPEC12>nh
Equinox (d/f 2000)[2000]
RA in hh mm ss.s or degrees[159.386] 04 18 21.3
DEC in dd mm ss.s or degrees[56.171] 38 01 36
  >> Leiden/Argentine/Bonn (LAB) Survey of Galactic HI
  LII , BII 161.675682   -8.819546
  Requested position at X and Y pixel    22.78    103.39
  Search nH in  4  X  4 box
  Each pixel is  0.675 deg 0.675 deg
  nH calculated using all points within
   1.0000 deg from input position
     RA        DEC        Dist         nH
   64.1051  37.3970   0.7360     2.65E+21
   65.1324  37.3425   0.8071     2.80E+21
   64.3226  37.9446   0.2251     2.93E+21
   63.5005  38.5376   0.9979     3.00E+21
   65.3701  37.8774   0.6333     2.89E+21
   64.5573  38.4821   0.4561     3.09E+21
   65.6250  38.4019   0.8984     2.84E+21
   64.8094  39.0092   0.9978     2.99E+21
  LAB >> Average nH (cm**-2)  2.90E+21
  LAB >> Weighted average nH (cm**-2)  2.91E+21
/usr/local/heasoft-6.16/x86_64-apple-darwin10.8.0/bin/nh
```

Alternatively (web tool): http://heasarc.nasa.gov/cgi-bin/Tools/w3nh/w3nh.pl

based on the LAB survey (Kalberla+05): http://www.astro.uni-bonn.de/~webaiub/english/tools_labsurvey.php xspec> mo pha*po

pha: accounts for the Galactic N_H (moltiplicative model)
po: powerlaw model (additive model)

It is possible to provide values to the parameters at every step of the fitting process

-1 means frozen parameter (the same as the Command *freeze* # of the parameter; opposite: *thaw*)

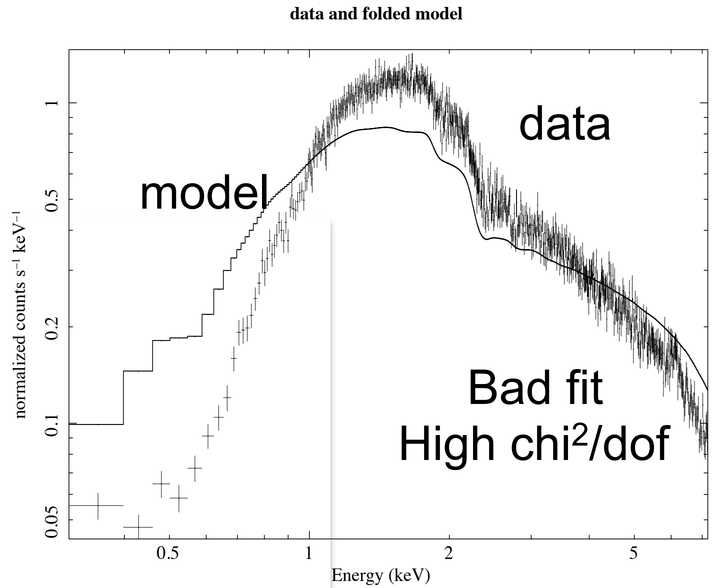
```
XSPEC12>mo pha*po

Input parameter value, delta, min, bot, top, and max values for ...
              1      0.001(      0.01)          0          0     100000      1e+06
1:phabs:nH>2.91e-1 -1
              1       0.01(      0.01)         -3         -2          9         10
2:powerlaw:PhoIndex>1.8
              1       0.01(      0.01)          0          0      1e+20      1e+24
3:powerlaw:norm>1e-6

========================================================================
Model phabs<1>*powerlaw<2> Source No.: 1   Active/On
Model Model Component  Parameter  Unit     Value
 par  comp
   1    1   phabs      nH         10^22    0.291000     frozen
   2    2   powerlaw   PhoIndex            1.80000      +/-  0.0
   3    2   powerlaw   norm                1.00000E-06  +/-  0.0
________________________________________________________________________
```

parameter number

number of the component

xspec> query yes
xspec> renorm
xspec> fit 100

```
========================================================================
Model phabs<1>*powerlaw<2> Source No.: 1   Active/On
Model Model Component  Parameter  Unit     Value
 par  comp
   1    1   phabs      nH         10^22    0.291000
   2    2   powerlaw   PhoIndex            0.895250
   3    2   powerlaw   norm                3.65973E-03
________________________________________________________________________
```




```
Fit statistic : Chi-Squared =       10561.52 using 1299 PHA bins.

Test statistic : Chi-Squared =       10561.52 using 1299 PHA bins.
 Reduced chi-squared =       8.143041 for   1297 degrees of freedom
 Null hypothesis probability =   0.000000e+00
```

dof=degrees of freedom=*number of datapoints – number of free parameters*=
=1299-2=1297

X^2/dof close to unity means that it is a good fit (not in this case!)

Null hypothesis probability=probability that the model is correct for those datapoints (if close to 1)

All the adopted models should be physically motivated according to the known source properties (or classification, or from other wavelengths)

Step 3a: adding components

xspec> addcomp 2 zpha   adding zpha as # component (#=order in the model)
xspec> fit 100

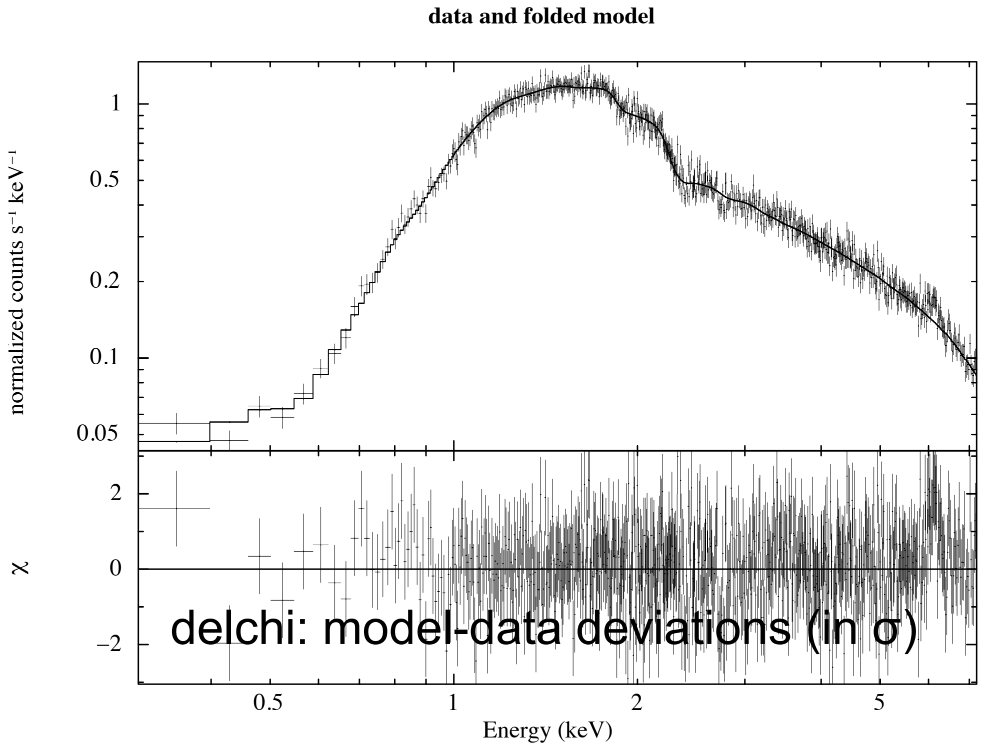
```
XSPEC12>addcomp 2 zpha

Input parameter value, delta, min, bot, top, and max values for ...
              1      0.001(      0.01)         0          0     100000      1e+06
2:zphabs:nH>1
              0      -0.01(      0.01)    -0.999     -0.999         10         10
3:zphabs:Redshift>.048

Fit statistic : Chi-Squared =          31014.39 using 1299 PHA bins.

Test statistic : Chi-Squared =         31014.39 using 1299 PHA bins.
 Reduced chi-squared =         23.93086 for   1296 degrees of freedom
 Null hypothesis probability =    0.000000e+00
 Current data and model not fit yet.

========================================================================
Model phabs<1>*zphabs<2>*powerlaw<3> Source No.: 1   Active/On
Model Model Component  Parameter  Unit     Value
 par  comp
   1    1   phabs      nH         10^22    0.291000     frozen
   2    2   zphabs     nH         10^22    1.00000      +/-  0.0
   3    2   zphabs     Redshift            4.80000E-02  frozen
   4    3   powerlaw   PhoIndex            0.895250     +/-  4.39599E-03
   5    3   powerlaw   norm                3.65973E-03  +/-  1.85394E-05
```

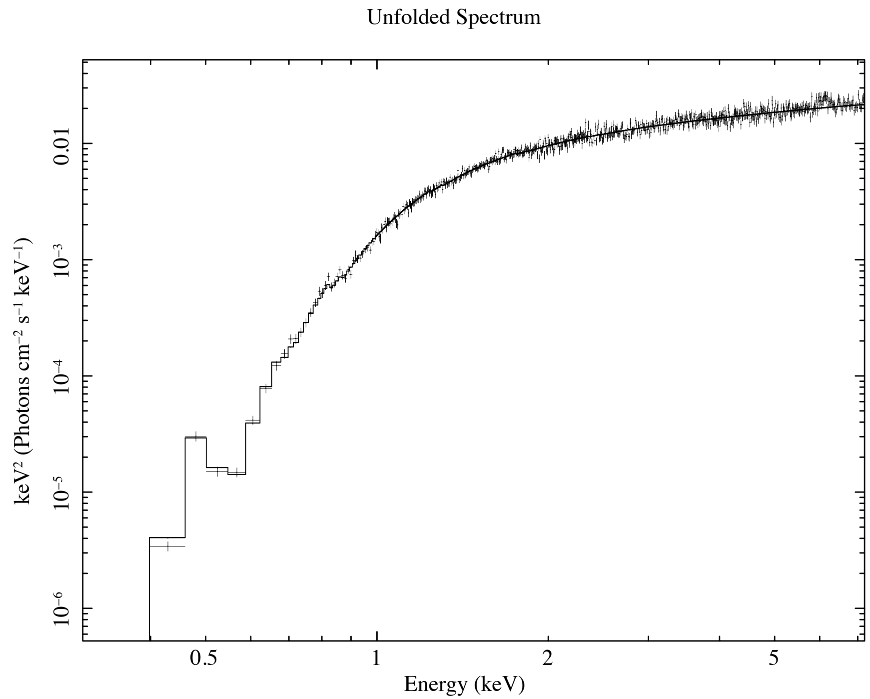
xspec> plot ldata delchi
xspec> plot model   to plot the input model

xspec> plot eeufspec delchi

Convolved with the response matrix

eeufspec: unfolded spectrum in E^2 F(E)





Step 4: statistical test: χ^2

Test to compare the observed distribution of the results with that expected

$$\chi^2 = \sum_{k=1}^{n} \frac{(O_k - E_k)^2}{\sigma_k^2}$$

O_k=observed values (spectral datapoints)
E_k=expected values (model)
σ_k=error on the measured values
k=number of datapoints (bins after rebinning)

$$\chi^2 / dof \approx 1$$

→ the observed and expected distributions are similar

Applicability of χ^2 statistics

χ^2 statistic

$$S = \sum_i (\underbrace{S_i - B_i t_s / t_b}_{O_K} - \underbrace{m_i t_s}_{E_K})^2 / (\underbrace{(\sigma_S)_i^2 + (\sigma_B)_i^2}_{\sigma_K})$$

where S_i = src counts in the I={1,…,N} data bins with exposure t_S, B_i = background counts with exposure t_B and m_i = model predicted count rate; $(\sigma_S)^2$ and $(\sigma_B)^2$ = variance on the src and background counts, typically estimated by S_i and B_i

BUT

the χ^2 statistic fails in low-counting regime
(few counts in each data bin)

Alternative solutions in case of low photon statistics

i. To rebin the data so that each bin contains a large enough number of counts

BUT: loss of information and dependence on the rebinning method adopted

ii. To modify S so the it performs better in low-count regime (e.g., by estimating the variance for a given data bin using the average counts from the surrounding bins; Churazov+96)

BUT: it would need Montecarlo simulations to properly support the result

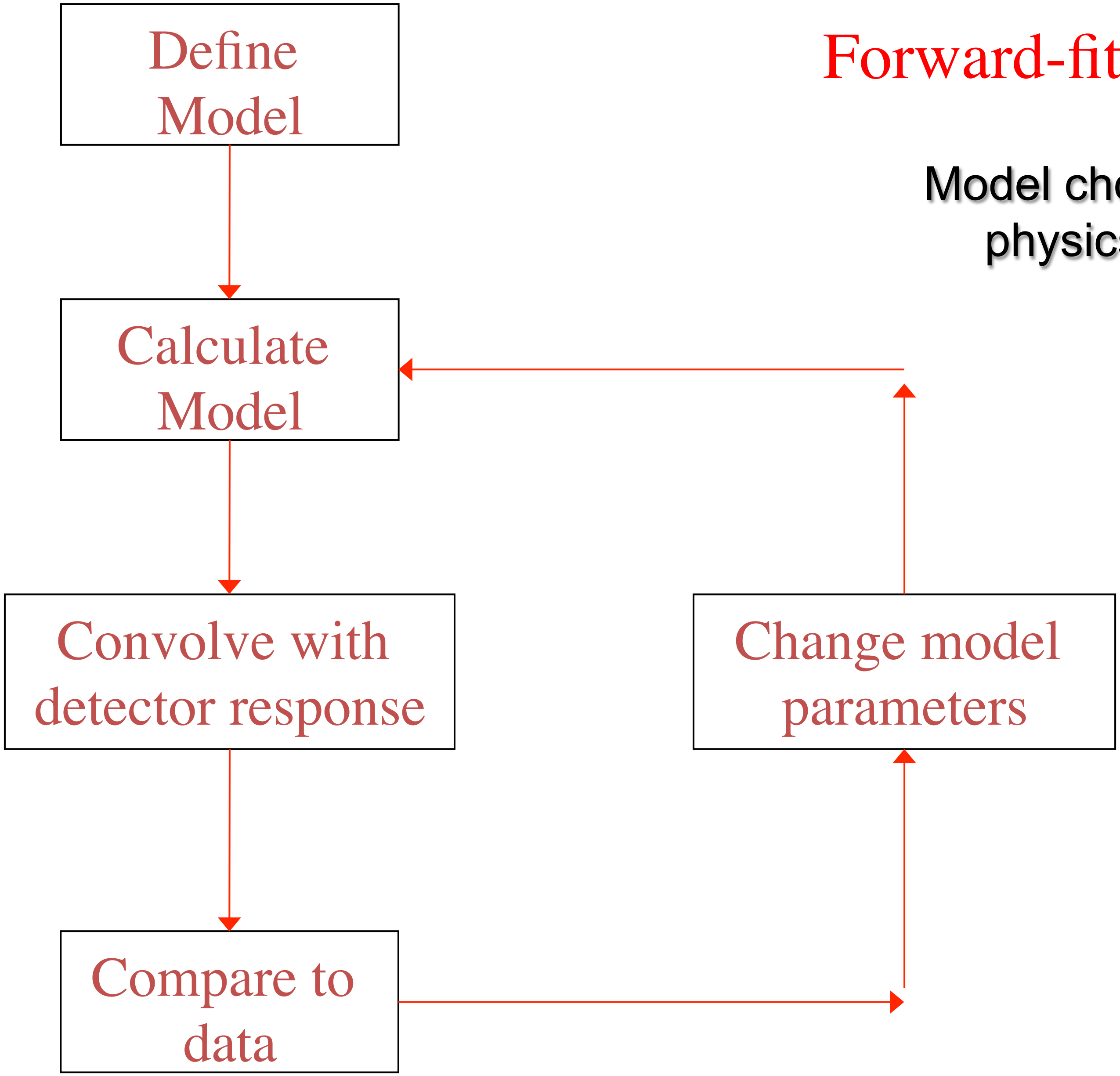
iii. To construct a maximum-likelihood estimator based on the Poisson distribution of the detected counts (Cash79; Wachter+79). ML means finding the best fit of parameters that maximizes the Poisson likelihood

xspec> statistic chi (default)
xspec> statistic cstat

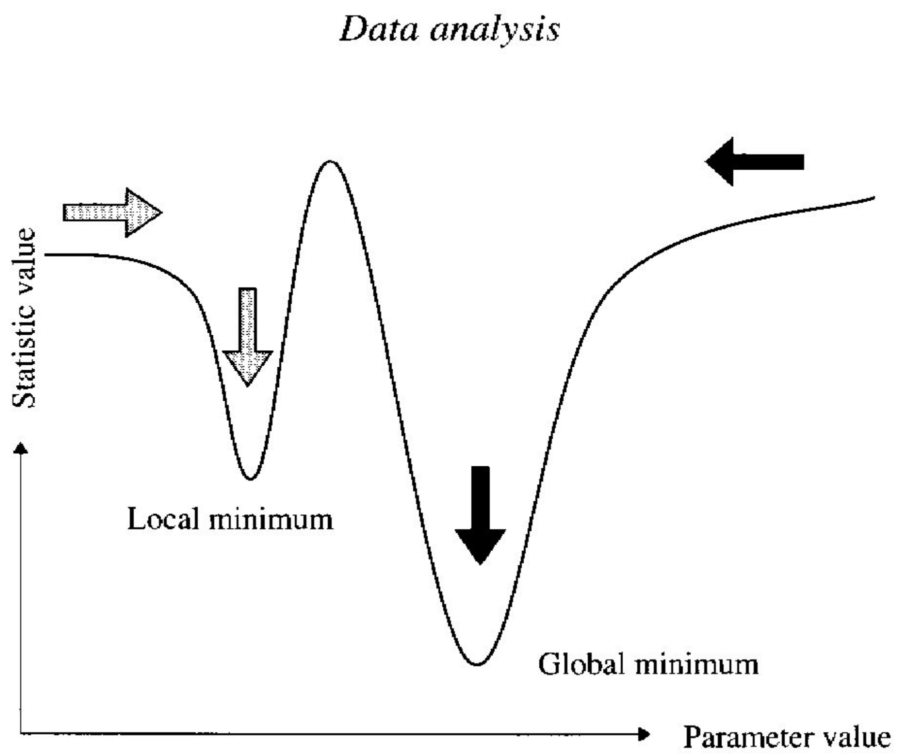
Binned data, χ^2 statistics ⇔ Gaussian statistics
Unbinned data, C-statistics ⇔ Poisson statistics

Forward-fitting algorithm

Model choice based on the physics of the source



Global vs. local minimum



If the fit process is started at the "right place", then it will converge to the true minimum
The more complicated the model and the more highly correlated the parameters, then the more likely that the algorithm will hardly find the true minimum

xspec> newpar 2 1.8 — to assign a new value to a parameter (2 here) (useful also to move from a local minimum…)

χ^2 in a nutshell

Reduced χ^2 large ⇔ $P(\chi^2)$ small

a. Errors are under-estimated
b. The model does not describe the data well

Reduced χ^2 small ⇔ $P(\chi^2)$ large

c. Errors are over-estimated
d. Data selected in a particular way?

Step 5: error estimate

xspec> error #
(#=number of the parameter)

$\Delta\chi^2$ as a Function of Confidence Level and Degrees of Freedom

p	ν = 1	2	3	4	5	6
68.3%	1.00	2.30	3.53	4.72	5.89	7.04
90%	2.71	4.61	6.25	7.78	9.24	10.6
95.4%	4.00	6.17	8.02	9.70	11.3	12.8
99%	6.63	9.21	11.3	13.3	15.1	16.8
99.73%	9.00	11.8	14.2	16.3	18.2	20.1
99.99%	15.1	18.4	21.1	23.5	25.7	27.8

Avni76

```
========================================================================
Model phabs<1>*zphabs<2>*powerlaw<3> Source No.: 1   Active/(
Model Model Component  Parameter  Unit     Value
 par  comp
   1    1   phabs      nH         10^22    0.291000     frozen
   2    2   zphabs     nH         10^22    0.465420     +/-  6.99525E-03
   3    2   zphabs     Redshift            4.80000E-02  frozen
   4    3   powerlaw   PhoIndex            1.59881      +/-  9.35898E-03
   5    3   powerlaw   norm                9.98424E-03  +/-  1.19425E-04
________________________________________________________________________

   Using energies from responses.

Fit statistic : Chi-Squared =         1286.80 using 1299 PHA bins.

Test statistic : Chi-Squared =         1286.80 using 1299 PHA bins.
 Reduced chi-squared =        0.992903 for   1296 degrees of freedom
 Null hypothesis probability =   5.667071e-01
 Weighting method: standard
XSPEC12>error 4
 Parameter    Confidence Range (2.706)
     4      1.58343       1.6143     (-0.0153884,0.0154818)
```

These are **NOT** the errors

2.706: 90% confidence level for one parameter of interest

Confidence	sigma	delta_chi-square
68.3%	1.0	1.00
90.0%	1.6	2.71
95.5%	2.0	4.00
99.0%	2.6	6.63
99.7%	3.0	9.00

1 parameter of interest

Ex.1: Error at 90% confidence level for one parameter of interest:
xspec> error #param 2.71

Ex. 2: Error at 90% confidence level for two parameters of interest:
xspec> error #param 4.61

Ex. 3: Error at 99% confidence level for one parameter of interest:
xspec> error #param 6.63

Table entry for p is the critical value $(\chi^2)^*$ with probability p lying to its right.

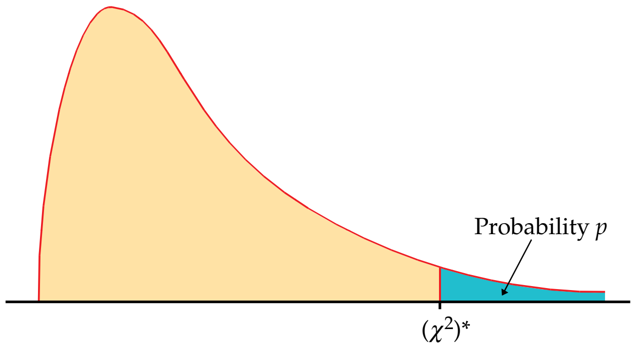


TABLE F

χ^2 distribution critical values

Parameter of interest → df 1, 2

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11

Step 5a: contour plots

xspec> steppar par1 min_value max_value #steps par2 min max #steps

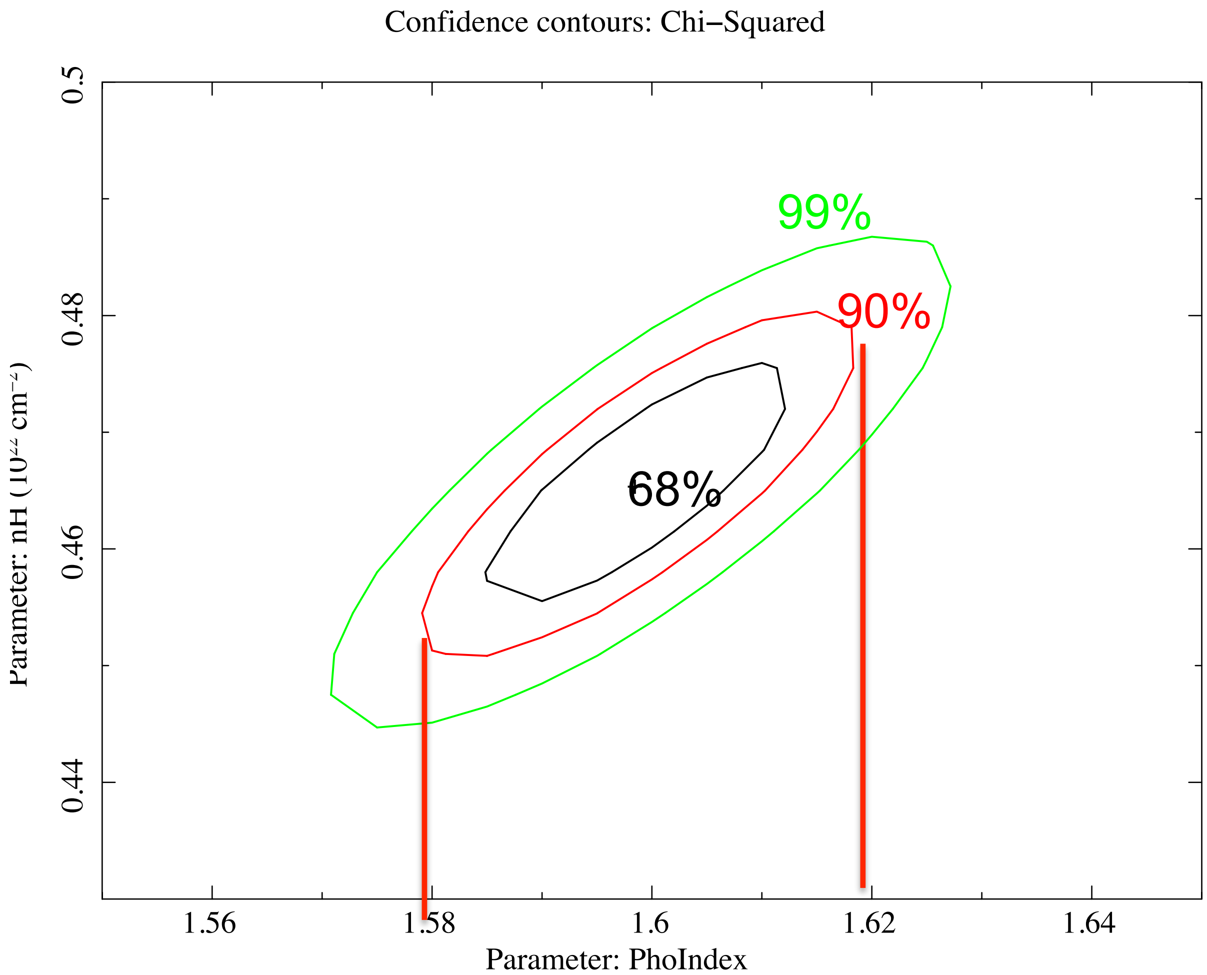
Perform a fit while stepping the value of a parameter through a given range

```
XSPEC12>stepp 4 1.55 1.65 10 2 0.43 0.50 10
```

Parameters involved in the fit

Chi-Squared	Delta Chi-Squared		PhoIndex 4		nH 2
1317	30.234	0	1.55	0	0.43
1313.8	26.991	1	1.56	0	0.43
1317.3	30.457	2	1.57	0	0.43
1327.4	40.613	3	1.58	0	0.43
1344.2	57.438	4	1.59	0	0.43
1367.7	80.908	5	1.6	0	0.43
1397.8	111	6	1.61	0	0.43
1434.5	147.68	7	1.62	0	0.43
1477.7	190.92	8	1.63	0	0.43
1527.5	240.68	9	1.64	0	0.43
1583.7	296.94	10	1.65	0	0.43
1527.3	240.45	10	1.65	1	0.437
1476.6	189.75	9	1.64	1	0.437
1432.3	145.51	8	1.63	1	0.437
1394.6	107.76	7	1.62	1	0.437
1363.3	76.53	6	1.61	1	0.437
1338.7	51.852	5	1.6	1	0.437
1320.6	33.752	4	1.59	1	0.437
1309.1	22.255	3	1.58	1	0.437
1304.2	17.383	2	1.57	1	0.437
1306	19.155	1	1.56	1	0.437

xspec> plot contours



90% c.l.: the photon index varies in the range 1.58−1.62 (vs. 1.58−1.61 using the *error* command). Slight differences are explained because in the case of the *error* command, the uncertainty was computed for one parameter of interest.

Step 6: source flux and luminosity

xspec> flux 2 10   (flux band in keV)
xspec> lum 2 10 0.048 (lum band redshift)

```
XSPEC12>flux 2 10
 Model Flux 0.0062186 photons (4.5024e-11 ergs/cm^2/s) range (2.0000 - 10.000 keV)
XSPEC12>ne 2 0

Fit statistic : Chi-Squared =       144391.0 using 1299 PHA bins.

Test statistic : Chi-Squared =       144391.0 using 1299 PHA bins.
 Reduced chi-squared =       111.4128 for   1296 degrees of freedom
 Null hypothesis probability =   0.000000e+00
 Current data and model not fit yet.
XSPEC12>lum 2 10 0.048
Model Luminosity 2.4791e+44 ergs/s  (2.0000 - 10.000 keV rest frame)
     (z = 0.0480 H0 =  70.0 q0 =  0.00 Lambda0 = 0.730)
```

Flux is *observed* (no correction for absorption) and in the observed-frame band

Luminosity needs to be intrinsic (so, put N_H = 0) and is reported in the source rest frame

Step 7: the F-test

Model 1: absorbed powerlaw
Model 2: absorbed powerlaw + iron emission line
xspec> addcomp 3 zgauss
xspec> fit 100

```
XSPEC12>addcomp 3 zgauss

Input parameter value, delta, min, bot, top, and max values for ...
            6.5       0.05(     0.065)         0          0      1e+06      1e+06
4:zgauss:LineE>6.4
            0.1       0.05(     0.001)         0          0         10         20
5:zgauss:Sigma>.01 -1
              0      -0.01(      0.01)    -0.999     -0.999         10         10
6:zgauss:Redshift>.048
              1       0.01(      0.01)         0          0      1e+20      1e+24
7:zgauss:norm>1e-6

Fit statistic : Chi-Squared =        1284.34 using 1299 PHA bins.

========================================================================
Model phabs<1>*zphabs<2>(zgauss<3> + powerlaw<4>) Source No.: 1   Active/On
Model Model Component  Parameter  Unit     Value
 par  comp
   1    1   phabs      nH         10^22    0.291000     frozen
   2    2   zphabs     nH         10^22    0.470750     +/-  7.09342E-03
   3    2   zphabs     Redshift            4.80000E-02  frozen
   4    3   zgauss     LineE      keV      6.40830      +/-  2.18809E-02
   5    3   zgauss     Sigma      keV      1.00000E-02  frozen
   6    3   zgauss     Redshift            4.80000E-02  frozen
   7    3   zgauss     norm                2.65689E-05  +/-  4.58946E-06
   8    4   powerlaw   PhoIndex            1.61154      +/-  9.66835E-03
   9    4   powerlaw   norm                1.01037E-02  +/-  1.23074E-04
________________________________________________________________________
```

```
Fit statistic : Chi-Squared =        1253.29 using 1299 PHA bins.

Test statistic : Chi-Squared =        1253.29 using 1299 PHA bins.
 Reduced chi-squared =        0.968541 for   1294 degrees of freedom
 Null hypothesis probability =    7.868667e-01
```

$\Delta\chi^2/\Delta dof=33.5/2$

$\chi^2/dof=1253.3/1294$
vs. 1286.8/1296 (no line)

xspec> ftest chi2_mod2 dof_mod2 chi2_mod1 dof_mod1

➔ Low F value: low statistical significance of the added component

xspec> ftest 1253.3 1294 1286.8 1296

```
XSPEC12>ftest 1253.3 1294 1286.8 1296
 F statistic value = 17.2939 and probability 3.87222e-08
```

Large F value=low probability = significant improvement due to the additional component

Use the F-test to evaluate the improvement to a spectral fit due to the assumption of a different model, with additional terms
Conditions: (a) the simpler model is nested within the more complex model; (b) the extra parameters have Gaussian distribution (not truncated by the parameter space boundaries) − BUT see also Protassov+02 on caveats

$$P_f(f;\upsilon_1,\upsilon_2) = \frac{\chi_1^2/\upsilon_1}{\chi_2^2/\upsilon_2}$$

$$\propto \Delta\chi^2/k$$

The larger this ratio is, the larger the improvement is in the spectral fitting
k=number of additional parameters

Other useful commands

in XSPEC

> setplot rebin #1 #2 (to rebin the data; #1 indicates the number of σ)
> show all
> show files
> show notice
> save all bestfit.xcm (save the best fit model with the data)
> save model bestmodel.xcm (save only the best fit model, without the data)

in PLOT

> time off
> csize 2 (character size)
> msize (marker size)
> label top (title of the plot)
> label filename (title of the file)
> hardcopy nomefile.ps/cps (save a figure)
> plot
> wen nomefile (writes two files, one with data and the other with plot settings)

Step 7: the F-test

Model 1: pow
Model 2: pow+line

xspec> ftest chi2_mod2 dof_mod2 chi2_mod1 dof_mod1
➔ Low F value: low statistical significance of the added component

Use the F-test to evaluate the improvement to a spectral fit due to the assumption of a different model, with additional terms
Conditions: (a) the simpler model is nested within the more complex model; (b) the extra parameters have Gaussian distribution (not truncated by the parameter space boundaries)

$$P_f(f;\upsilon_1,\upsilon_2) = \frac{\chi_1^2/\upsilon_1}{\chi_2^2/\upsilon_2}$$

$$\propto \Delta\chi^2/k$$

The larger this ratio is, the larger the improvement is in the spectral fitting
k=number of additional parameters