Centroid Statistics

revised 3 September 2018, DESI-DOC 4006v7, M.Lampton UCB SSL

One Dimension: Given a set of data bins or pixels numbered with an index i, the chi-squared statistic measures the discrepancy between an adjustable mathematical model f_i and a set of given fixed data d_i:

$$\chi^2 = \sum_i (\frac{f_i - d_i}{\sigma_i})^2$$

Here, σ represents the RMS statistical uncertainty in the contents of each bin, assumed uniform across all bins, so that each term is expected to be the order of 1. Generally the model will have adjustable parameters that can be explored to best fit the data. The 1-dimensional problem addresses a single spectral line whose shape is known but whose intensity a and location b are to be determined along with their errors. I use a two parameter model linear in both a and b: $f_i = a(g_i - bg'_i)$ where g_i is a fixed nominal image profile that specifies the fraction of signal that falls into each bin i, so $\Sigma g_i = 1$, and g'_i is its derivative with respect to i, units of fractional change per unit i, with $\Sigma g'_i = 0$. The total signal is then a, and a lateral displacement of its centroid is described by b. I expand χ^2 into adjustable sums and fixed sums:

$$\begin{aligned}\chi^2 &= \sum_i \Big(\frac{a(g_i - bg'_i) - d_i}{\sigma_i}\Big)^2 \\ &= \frac{a^2}{\sigma^2}\sum_i g_i^2 + \frac{a^2b^2}{\sigma^2}\sum_i g_i'^2 + \frac{1}{\sigma^2}\sum_i d_i^2 - \frac{2a^2b}{\sigma^2}\sum_i g_i g'_i - \frac{2a}{\sigma^2}\sum_i g_i d_i + \frac{2ab}{\sigma^2}\sum_i g'_i d_i\end{aligned}$$

Two of these terms do not matter: Σd^2 has no derivatives, and $\Sigma g_i g'_i = 0$ by orthogonality. The best fit parameter estimates are their values for which χ^2 is minimum, i.e. their χ^2 derivatives are zero. The variances of a and b depend only on their second derivatives (quadratic terms) in this sum, since a change that gives $\Delta\chi^2 = 1$ sets the range of $\pm 1\sigma_{param}$ uncertainty in that parameter [1]. That is, $var(p) = 2/(d^2\chi^2/dp^2)$:

$$\begin{aligned}\frac{d\chi^2}{da} &= \frac{2a}{\sigma^2}\sum_i g_i^2 - \frac{2}{\sigma^2}\sum_i g_i d_i = 0 \quad \text{so} \quad \hat{a} = \frac{\Sigma g_i d_i}{\Sigma g_i^2} \quad \text{and} \quad var(a) = \frac{\sigma^2}{\Sigma g_i^2} \\ \frac{d\chi^2}{db} &= \frac{2a^2b}{\sigma^2}\sum_i g_i'^2 + \frac{2a}{\sigma^2}\sum_i g'_i d_i = 0 \quad \text{so} \quad \hat{b} = \frac{-\Sigma g'_i d_i}{a\,\Sigma g_i'^2} \quad \text{and} \quad var(b) = \frac{\sigma^2}{a^2\,\Sigma g_i'^2}\end{aligned}$$

The weighted noise in the profile is $\sqrt{var(a)} = \sigma/\sqrt{\Sigma g_i^2}$. Define the spot breadth $B = \sqrt{\Sigma g_i^2}/\sqrt{\Sigma g_i'^2}$:

$$SNR = \frac{\hat{a}}{\sigma}\cdot\sqrt{\Sigma g_i^2} \quad \text{and} \quad \sigma_{centroid} = \frac{B}{SNR}$$

Here, B depends on the shape of the spot profile. For a Gaussian spatial profile with standard deviation s or $FWHM = 2.35482\,s$, $\Sigma g_i^2 = 1/(2\sqrt{\pi}s)$ and $\Sigma g_i'^2 = 1/(4\sqrt{\pi}s^3)$, so that $B = \sqrt{2}s = 0.60056\cdot FWHM$:

$$\sigma_{centroid} = \frac{0.6\,FWHM}{SNR}$$

Two Dimensions: This follows the one dimensional case but with $f_{i,j} = a(g_{i,j} - b\ dg/di - c\ dg/dj)$. The χ^2 statistic now has ten terms, six of which are useful, giving the three estimators $\hat{a}, \hat{b}, \hat{c}$ and their variances as in the one dimensional case. The breadth B is again $\sqrt{2}s$ and the rms centroid error is the same as in the one dimensional case.

[1] Press et al., *Numerical Recipes in C*, 2nd ed., p.696, eqn 16.6.4, 1992.

Appendix: Useful Gaussian Expressions
For $g(x)$ a 1D Gaussian with unit area and rms width s,

$$g(x) = \frac{1}{\sqrt{2\pi}\ s}\ exp(-\frac{x^2}{2s^2})$$
$$g^2(x) = \frac{1}{2\pi\ s^2}\ exp(-\frac{x^2}{s^2})$$
$$\int_{-\infty}^{+\infty} g^2(x)dx = \frac{1}{2\sqrt{\pi}\ s}$$
$$g'(x) = \frac{-x}{\sqrt{2\pi}\ s^3}\ exp(-\frac{x^2}{2s^2})$$
$$g'^2(x) = \frac{x^2}{2\pi\ s^6}\ exp(-\frac{x^2}{s^2})$$
$$\int_{-\infty}^{+\infty} g'^2(x)dx = \frac{1}{4\sqrt{\pi}\ s^3}$$
$$B = \frac{\sqrt{\int g^2 dx}}{\sqrt{\int g'^2 dx}} = \sqrt{2}\ s$$

The 2D Cartesian coordinate Gaussian with unit volume and rms core radius s gives

$$g(x,y) = \frac{1}{2\pi s^2} exp(\frac{-x^2}{2s^2}) exp(\frac{-y^2}{2s^2})$$
$$g^2 = \frac{1}{4\pi^2 s^4} exp(\frac{-x^2}{s^2}) exp(\frac{-y^2}{s^2})$$
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g^2 dxdy = \frac{1}{4\pi s^2}$$
$$g' \equiv \frac{dg}{dx} = \frac{-x}{2\pi s^4}\ exp(\frac{-x^2}{2s^2}) exp(\frac{-y^2}{2s^2})$$
$$g'^2 = \frac{x^2}{4\pi^2 s^8} exp(\frac{-x^2}{s^2}) exp(\frac{-y^2}{s^2})$$
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g'^2 dxdy = \frac{1}{8\pi s^4}$$
$$B_x = B_y = \frac{\sqrt{\int\int g^2 dxdy}}{\sqrt{\int\int g'^2 dxdy}} = \sqrt{2}\ s$$

The 2D polar coordinate Gaussian with unit volume and rms core radius s gives the EE_{50} radius:

$$g(r) = \frac{1}{2\pi s^2}\ exp(-\frac{r^2}{2s^2})$$
$$frac(R) = \frac{1}{2\pi s^2}\int_0^R 2\pi\ r\ dr\ exp(\frac{-r^2}{2s^2}) = 1 - exp(\frac{-R^2}{2s^2})$$
$$EE50\ radius = \sqrt{2\ ln(2)}\ s$$
$$FWHM = 2\sqrt{2\ ln(2)}\ s = 2.35482\ s$$

Therefore the ratio of $B_x/FWHM$ or $B_y/FWHM$ is 0.60056.

Useful Moffat Expressions

The 2D Moffat distribution[2] is another commonly used model for the point spread function of astronomical images. Its core radius is represented by the parameter α. Its skirt steepness is represented by the parameter β. The King distribution[3] is a Moffat whose $\beta = 3/2$; it is commonly used as an smooth approximation to a circular diffraction pattern. (But note: although its EE50 is well defined, the rms radius moment diverges for $\beta \le 2$.) Astronomical seeing is commonly modeled with $\beta = 7/2$; see for example Dey and Valdes. [4] Its radial moments are most easily obtained in polar coordinates:

$$g(r) = \frac{\beta - 1}{\pi\alpha^2} \cdot \frac{1}{(1 + \frac{r^2}{\alpha^2})^\beta}$$

$$\int_0^R 2\pi r g(r) dr = 1 - \left(1 + R^2/\alpha^2\right)^{1-\beta}$$

$$FWHM = 2\sqrt{2^{1/\beta} - 1}\ \alpha$$

$$RMS\ radius = \frac{\alpha}{\sqrt{\beta - 2}}$$

$$EE50\ radius = \sqrt{2^{1/(\beta-1)} - 1}\ \alpha$$

Cartesian coordinates are needed for obtaining the centroid errors in the (x,y) plane:

$$g(x, y) = \frac{\beta - 1}{\pi\alpha^2} \cdot \frac{1}{(1 + \frac{x^2}{\alpha^2} + \frac{y^2}{\alpha^2})^\beta}$$

$$g' \equiv \frac{dg}{dx} = \frac{-2\beta(\beta - 1)}{\pi\alpha^4} \cdot \frac{x}{(1 + \frac{x^2}{\alpha^2} + \frac{y^2}{\alpha^2})^{\beta+1}}$$

A general analytic expression for the single integrals of g^2 and g'^2 is provided by Gradshteyn and Ryzhhik[5]:

$$\int_0^{+\infty} \frac{x^{\mu-1} dx}{(p + qx^\nu)^{n+1}} = \frac{1}{\nu p^{n+1}} \left(\frac{p}{q}\right)^{\mu/\nu} \frac{\Gamma(\mu/\nu)\Gamma(1 + n - \mu/\nu)}{\Gamma(1 + n)}$$

In general the g^2 and g'^2 double integrals are complicated, but thanks to the properties of the Gamma function, they simplify for β equal to an integer or half integer. The Python SymPy package[6] evaluates these double integrals, as does the Wolfram Alpha web tool[7]. In the table below I list my results. As expected, their relative breadths $B/FWHM$ are larger than the Gaussian ratio 0.60065 but they approach it when the skirts are steep.

Double Integrals of the Moffat Distribution

β	$\int\int g^2 dx\, dy$	$\int\int g'^2 dx\, dy$	B/α	$FWHM/\alpha$	$B/FWHM$
3/2	$1/(8\pi\alpha^2)$	$(3 \times 1)^2/(16 \times 6\ \pi\alpha^4)$	1.15470	1.53284	0.75331
2	$4/(12\pi\alpha^2)$	$(4 \times 2)^2/(16 \times 10\ \pi\alpha^4)$	0.91287	1.28719	0.70920
5/2	$9/(16\pi\alpha^2)$	$(5 \times 3)^2/(16 \times 15\ \pi\alpha^4)$	0.77460	1.13050	0.68158
3	$16/(20\pi\alpha^2)$	$(6 \times 4)^2/(16 \times 21\ \pi\alpha^4)$	0.68313	1.01965	0.66997
7/2	$25/(24\pi\alpha^2)$	$(7 \times 5)^2/(16 \times 28\ \pi\alpha^4)$	0.61721	0.93598	0.65943
4	$36/(28\pi\alpha^2)$	$(8 \times 6)^2/(16 \times 36\ \pi\alpha^4)$	0.56695	0.86996	0.64576

[2] Moffat, A.F.J., Astron & Astrophys. vol 3, 455-464, 1969.
[3] King, I.R., PASP vol 83, 199-201, 1971 and PASP vol 95, 163-168, 1983.
[4] Dey, A., and Valdes, F., PASP vol126, 296-311, 2014.
[5] Gradshteyn and Ryzhik, Tables of Integrals Series and Products, 7th ed., sec. 3.241 No.4 p.322, 2007.
[6] www.sympy.org/en/index.html
[7] www.wolframalpha.com/widgets

Comparing Moffats with the Gaussian

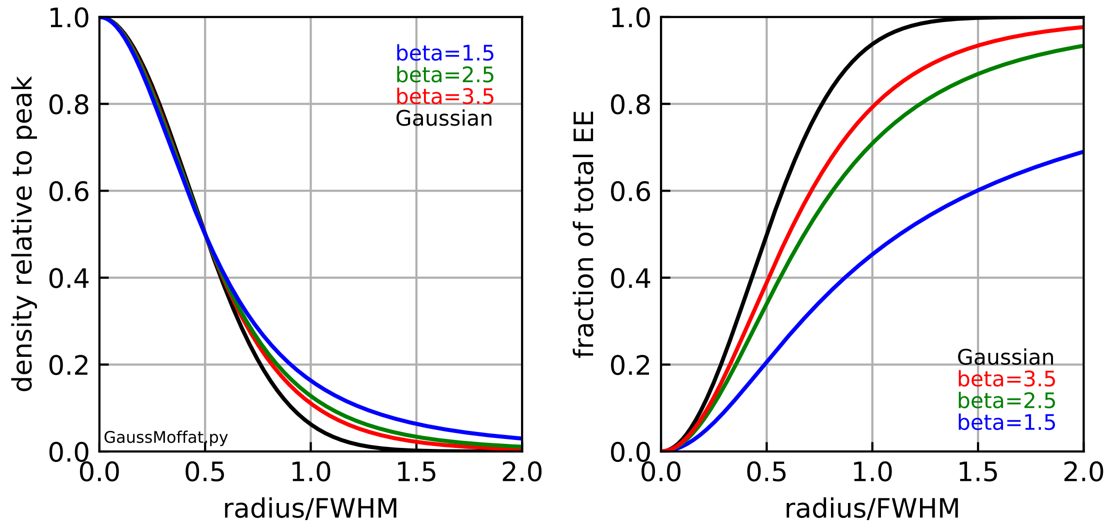

Figure 1: Left: densities of 2D Gaussian and three 2D Moffats; the extended wings of the Moffats are barely visible. Right: the encircled energies of these distributions, now clearly showing the effect of the extended Moffat wings.