SNU AO Seminar Notes

Astronomical Observation and Lab at Seoul National University

Yoonsoo P. Bach

This book is prepared since 2019 Spring,

Seminars are given since 2016 Fall.

Contents

Introduction

In math, people prefer to use "$\log_e$" or just "log" rather than "ln", which is the opposite to many scientific fields. Due to its simplicity (especially when typing), I will also stick to the usage of "ln" which is identical to "$\log_e$". But to avoid any confusion, I will always specify the base 10 ("$\log_{10}$").

At many places in physics and astronomy, people use $\equiv$ as "is defined as" or any kind of "trivial by the definition" cases. But I will use $:=$ for the definition and use $\equiv$ to mean "equality holds trivially by the definition". For example, under 1-D linear homogeneous isotropic matter, linear electric polarization is written as $P \equiv \epsilon_0 \chi E$ or $\chi := P/\epsilon_0 E$. The latter is the *definition* of electric susceptability, while the former is *trivially true by the definition*.

For distributions, $\sim$ means "follows", i.e., $X \sim \mathcal{N}(\mu, \sigma^2)$ means "the random variable X follows the Gaussian (normal) distribution with mean μ and standard deviation σ (thus variance σ^2)." When this is an approximation, the symbol with dot "$\dot{\sim}$" is used, such as $X \dot{\sim} \mathcal{N}(\mu, \sigma^2)$. In this case, it is interpreted as "approximately follows."

In this note, I used some math-like notations like theorem (Thm). However, the theorem here is not necessarily the same as that of mathematics. I tried to put some important assumptions and summary notes to theorem.

The references are

- "Walpole" is Walpole et al. 2013, "Essentials of Probabilities and Statistics".
- "Sivia" is Sivia and Skilling 2006, "Data Analysis A Bayesian Tutorial", 2/e.

Chapter 1

Statistics - Basic

In any chapter in this book, especially in this chapter, I will assume (1) you have really basic knowledge on statistics, e.g., what "Gaussian", "probability", "standard normal distribution", "random variable" means and/or (2) you know how to use Google so that you can find concepts which you don't know. The reasons we cannot touch much of the mathematical details are because what we have to deal with in this class is rather a practical usages of statistics in astronomy, and because I have not been a good math student.

Let's start with a question.

Ex 1

For the **same star**, consider the following three scenarios:

- Case A: Researcher 1 says $m_1 = 14.0^{\rm m} \pm 0.1^{\rm m}$; Researcher 2 says $m_2 = 15.0^{\rm m} \pm 1.0^{\rm m}$.
- Case B: Researcher 1 says $m_1 = 14.0^{\rm m} \pm 0.1^{\rm m}$; Researcher 2 says $m_2 = 15.0^{\rm m} \pm 0.3^{\rm m}$.
- Case C: Researcher 1 says $m_1 = 14.0^{\rm m} \pm 0.1^{\rm m}$; Researcher 2 says $m_2 = 15.0^{\rm m} \pm 0.1^{\rm m}$.

For each of the three cases: Are the two studies coincide? If so, with how much confidence would you say so? Wait, but what does that "±" sign means in rigorous mathematical sense?

Some of the previous course takers answered that they coincide, because the "3-sigma rule" says, e.g., for case A, $m_1 = 14.0^{\rm m} \pm 0.3^{\rm m}$ & $m_2 = 15.0^{\rm m} \pm 3.0^{\rm m}$ overlaps with each other. Crudely speaking, this makes sense, but it is of course not the "publication level" reasoning.

To give you the answer: It is an ill-defined question. The answer can change based on many factors; to pick one, the number of observations each reasercher used for the determination of the magnitude. If we assume that number is 5 for both researchers, for example, we can give some meaningful answers: We *cannot reject* $m_1 = m_2$ for case A & B with 90 % confidence, and we *can reject* $m_1 = m_2$ for case C with 90 % confidence. Depending on the confidence level, the answer may also change. Note that the expression that *can/cannot reject* is very important! You should **NOT** say you *accept* $m_1 = m_2$, never ever. Let me explain why throughout this chapter.

1.1 The n-σ

Maybe you are familiar of the 1-σ, 2-σ, and 3-σ words (cf. fig. 1.1). For example, mean $\pm$1-σ contains $68.27\cdots$ % of the total area of a normal distribution. Similarly 3-σ conatains $99.73\cdots$ % of it, so it is very unlikely to get a sample outside of the mean $\pm$3-σ.

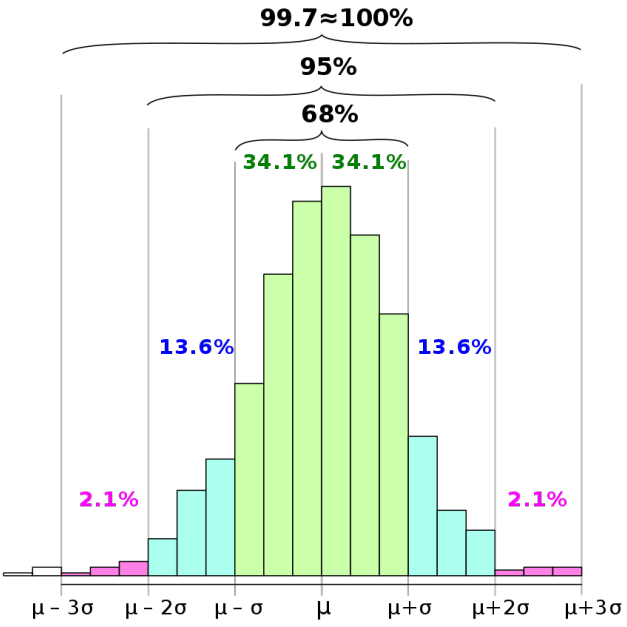

Figure 1.1: A test sampling from Gaussian (normal) distribution (from Wikimedia).

In this sense, the "n-σ" is defined as the x-axis value which gives the integrated area of certain value. More precisely, we have to call this confidence interval (CI):

Def 1 (n-σ *Confidence Interval; CI*)

n-σ **confidence interval** is defined as *an* interval which contains $\Phi(n) - \Phi(-n)$ of the total area, usually centered at the mean or median, where $\Phi := (2\pi)^{-1/2} \int_{-\infty}^{x} e^{t^2/2} dt$ is the cumulative distrubution function of standard normal distribution.

Following this definition, 1-σ CI is the interval such that the integrated area is $0.6827\cdots$ of the total area. Similarly, 2-σ CI is that with $0.9545\cdots$, 3-σ CI is that with $0.9973\cdots$, etc. The term CI is not necessarily limited to Gaussian (normal) distribution: You can apply it to any distribution, and give an interval such as $0.11^{+0.03}_{-0.01}$ as the 1-σ CI, for instance.

Sometimes it is more convenient to call, e.g., 1-σ CI as the $68.27\cdots\%$ CI. We also define the *significance level* α as 100% minus this percentage, i.e., the 1-σ CI is of significance level $\alpha = 1 - 0.6827\cdots = 0.3173\cdots$:

$$\text{1-}\sigma \text{ CI} \quad \equiv \quad 68.27\cdots\% \text{ CI} \quad \equiv \quad \text{CI of significance level } 0.3173\cdots \ . \tag{1.1}$$

In physical and astronomical sciences we often use the terminology of "n-σ CI", but in mathematics and other branches of sciences and engineering, the "k % CI" or the "significance level α" are used dominantly. Among them, 90 %, 95 %, and 99 % CIs, i.e., CIs with significance level $\alpha = 0.10$, 0.05, and 0.01 are found frequently.

Because when we say "n-σ", we are just omitting the term "CI" at the end of it, this is not n times the standard deviation (σ): What we mean by "n-σ" is actually "n-σ CI" and is *not necessarily* $n \times \sigma$. They are the same for, e.g., Gaussian distribution, because the n-σ terminology is *defined* using it. For Gaussian (normal) distribution, which is a distribution that appears very often in natural sciences, such as 2-σ is nothing but 2 times the standard deviation (σ). For almost all of the distributions other than Gaussian, this is not the case.

1.2 The Meaning of ± Sign

Simple answer: The ± sign means the 1-σ confidence interval (CI). It is also called the "**error-bar**".

Mathematically speaking, the ± sign, i.e., the 1-σ CI, does *not* fully describe the uncertainty. The best way to show or describe the uncertainty is to show a graph of **probability distribution function** (**pdf** or p.d.f.) as in fig. 1.2. In many publications or books, people do not show this, because (1) the distribution is very similar to Gaussian or a widely known distribution and it is clearly written in the text or assumed as all the readers know that, (2) detailed calculation is not of interest and/or does not affect the final result, (3) the writer is lazy.

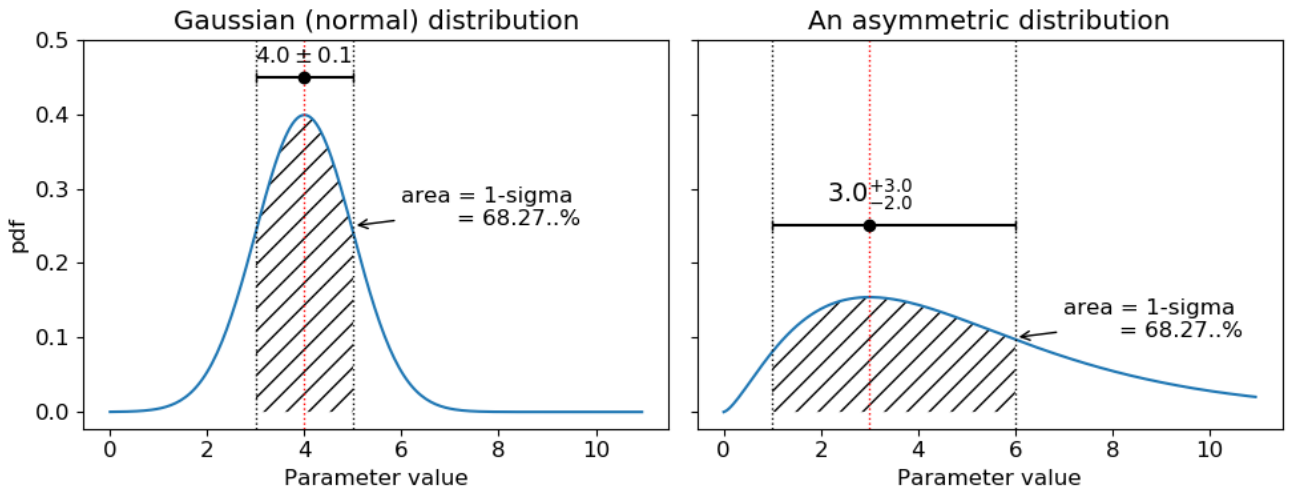

Figure 1.2: The probability distribution function (pdf) of two examples of distributions, namely, Gaussian and an asymmetric distributions. I used a chi-square distribution for the latter for plotting purpose.

The uncertainty ranges shown in fig. 1.2 (4.0 ± 0.1 and $3.0^{+3.0}_{-2.0}$) are determined such that the lower/upper bounds include $68.27\cdots\%$ of the total area. It is trivial for a Gaussian distribution: mean ± 1-σ. But if the distribution is asymmetric (right panel of the figure), there are few choices to set such bounds. First and the most widely used one is to set the bounds such that the distance between upper/lower bounds are minimized (but include $68.27\cdots\%$ of the total area). Second is to make it symmetric while include $68.27\cdots\%$ of the total area, so that you can use ± sign for simplicity (3.0 ± 2.3 for example). Third choice would be something like FWHM: find x-axis values such that the pdf value is $0.5 \times \text{pdf}_{\text{max}}$. The last two bounds are simple to calculate but not as accurate as the first choice.

Very assymetric probability distributions appear, e.g., in cosmological sciences*, and exhaustive statiscal analyses are conducted on such models to *reject* (I repeat; *reject, not accept*) cosmological models. The reason is that, in cosmology, we have only one single sample (our universe), and the statistical analyses given the observational data is of utmost importance. Of course similar exhaustive statistical analyses should be conducted on any research field if the data/object/target is of such importance. We don't do that in many observational studies purely because (1) we don't have human power to do all that and (2) the results won't change much (but no one in history thoroughly checked this).

*Google "posterior distribution cosmological constants".

1.3 Caveat on the "n-σ CI" Notation

Let me emphasize again: What we mean by "n-σ" is actually "n-σ CI" and is *not necessarily* $n \times \sigma$.

It maybe is tempting to convert the 1-σ CI to the 3-σ CI. For instance, from fig. 1.2, 4.0 ± 0.1 and $3.0^{+3.0}_{-2.0}$ to 4.0 ± 0.3 and $3.0^{+9.0}_{-6.0}$. This is true for the former (Gaussian) but wrong for the latter (non-Gaussian). The lower bound of the second parameter now become negative ($3.0 - 6.0 = -3.0$), which is not even physically correct if this parameter is defined to be positive. What you have to do is, get the 3-σ CI which should contain $0.9973\cdots$ of the total area from the distribution given in fig. 1.2. There can be at least three choices to do it, which I mentioned in the previous section.

1.4 Central Limit Theorem (CLT)

The Central Limit Theorem (CLT) is at the heart of all the observational or experimental sciences. It is stated as*

Thm 1 (*Central Limit Theorem; CLT*)

Consider a random sample (observation, measurement, etc) with size n and the mean value of this sample is $\bar{X}$. Then $(\bar{X} - \mu)/(\sigma/\sqrt{n})$ approaches the standard normal distribution ($\mathcal{Z}$) as $n \to \infty$, where μ and σ^2 are the (finite) mean and variance of the population:

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{Z} \quad (\text{as } n \to \infty) \ . \tag{1.2}$$

The standard deviation (second moment) is sometimes undefined. An example is the Cauchy distribution, and I will show you later with an example when this distribution can appear. There is one more important theorem, which we usually skip to mention:

Thm 2 (*Sample Variance Distribution*)

Consider a random sample (observation, measurement, etc) with size n and the sample variance is S^2. Then $(n-1)S^2/\sigma^2$ follows a chi-squared distribution with degrees of freedom $(n-1)$, where σ^2 is the variance of the population:

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} \tag{1.3}$$

Also the definition of the Student t-distribution†:

*For strict definitions of random sample, population, independence, the statement of CLT in mathematical senses, etc, I recommend mathematical statistics textbooks.

†This definition is directly copied from Walpole et al. p.177

Def 2 (*Student t-distribution*)

Let Z be a standard normal random variable and V a chi-squared random variable with v degrees of freedom. If Z and V are independent, then the distribution of the random variable T, where

$$T = \frac{Z}{\sqrt{V/v}} \tag{1.4}$$

is given by the density function

$$h(t) = \frac{\Gamma\left[\frac{v+1}{2}\right]}{\Gamma\left[\frac{v}{2}\right]\sqrt{\pi v}}\left(1+\frac{t^2}{v}\right)^{-\frac{(v+1)}{2}} \quad , \quad -\infty < t < \infty \ . \tag{1.5}$$

This is known as the Student t-distribution with v degrees of freedom.

We can see that the t-distribution approaches the standard normal distribution as the degrees of freedom $v \to \infty$, because the power term takes the $e^{-t^2/2}$ form.

Thm 3 (*Practical Usage of the CLT*)

Consider a random sample (observation, measurement, etc) with size n large enough and the mean value of this sample is $\bar{X}$. If μ, σ^2, and S^2 are the (finite) mean and variance of the population, and the sample variance, respectively:

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim T_{(n-1)} \ . \tag{1.6}$$

Sometimes people empirically say $n \geq 30$ is enough. This really depends on the situation and I will skip this issue here. The sketch of the proof of the theorem is simple. By the definition of the t-distribution and Thm 2,

$$T = \frac{Z}{\sqrt{V/v}} = \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \tag{1.7}$$

so the theorem is plausible. For your information, the sample variance is defined as

$$S^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X-\bar{X})^2 \tag{1.8}$$

and the sample standard deviation S is the square root of this.

Note here the difference between Thm 1 and Thm 3: σ is changed to S, and the standard normal distribution is changed to the t-distribution of $(n-1)$ degrees of freedom. Since the true variance σ^2 is unknown, it is difficult to use the CLT (Thm 1) directly. But thanks to Thm 2, we can use the variance of the sample, S^2, which is measurable, and utilize the t-distribution, which is slightly bothersome than the standard normal distribution but still useful.

Also be aware that CLT says the *expectation value of the mean* is normally distributed, not the *sample* is so.

Ex 2 (*Coin Tossing and the CLT*)

Consider a coin-tossing experiment and assign ± 1 to heads and tails. Each experiment will give you ± 1, but never 0. Meanwhile, we know a fair coin should have $\mu = 0$ (same probability of heads and tails). After $n = 100$ experiments, say you obtained $\bar{X} = 0.01$ and $S = 0.1$, so

$$T = \frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{0.01-\mu}{0.1/\sqrt{100}} = 1 - 100\mu \ .$$

Since $T \sim t_{99}$, the significance level $\alpha = 0.05$ confidence interval, i.e., the confidence interval containing $100(1-\alpha)\% = 95\%$, can be

$$t_{99,0.025} < T < t_{99,0.975} \quad \rightarrow \quad -1.9842 < 1 - 100\mu < 1.9842 \quad \rightarrow \quad \mu \in [-0.0098,\, 0.0298] \ .$$

The notation $t_{\nu,x}$ means the input argument of the t-distribution with the degrees of freedom ν such that $\int_{-\infty}^{t_{99,x}} h(t)dt = x$. Thus, to get the two-tail CI of significance level α, we need to calculate $t_{\nu,\alpha/2}$ and $t_{\nu,1-\alpha/2}$. For a symmetric distribution like t here, $t_{\nu,1-\alpha/2} = -t_{\nu,\alpha/2}$. This is very widely used standard notation.

If you just calculated without pondering about the meaning of the caculation, you may be surprised that your next experiment gives either $+1$ or -1, while the expectation is $\mu \in [-0.0098, 0.0298]$. This is because the result from CLT is, as described, about the *mean* of the samples, not the *single sample*. Thus, the error-bar from the CLT is not necessarily predicting the possible range of *future experiments*, but it just confines the *position of the mean*.

The range which predicts the future experiments is called the prediction interval. You may learn the prediction interval (PI) to clarify this difference.

The two-tail significance level α interval of the t-distribution calculable in python by

```
from scipy.stats import t
alpha = 0.95
nu = 99

# Note that the alpha below is NOT the significance level,
# but the area under the curve!
lo, hi = t.interval(alpha=alpha, df=nu)
print(f"Confidence Interval: [{lo:.4f}, {hi:.4f}]")
```

Note: Uncertainty of Median

The 1-σ error-bar of the median is more difficult to handle than that of the mean (CLT: Thm 1). But we can get a result from simplifying assumptions, which are not necessarily true for real observations:

Thm 4 (*Uncertainty of Median*)

Consider $n(\gg 1)$ samples are independently drawn from $X \sim \mathcal{G}(\mu, \sigma^2)$, a general continuous distribution with finite mean and variance μ and σ^2 with pdf $p(x)$. The uncertainty of the median estimator is

$$\Delta\text{med} \approx \frac{1}{2p(\nu_0)\sqrt{n-1}} \tag{1.9}$$

and if $\mathcal{G} = \mathcal{N}$, i.e., a normal distribution,

$$\Delta\text{med} \approx \sqrt{\frac{\pi/2}{n-1}}s = \sqrt{\frac{\pi}{2}\frac{n}{n-1}}\Delta\text{mean} \approx 1.25\Delta\text{mean} \ , \tag{1.10}$$

where s is the sample standard deviation and $\Delta\text{mean} = s/\sqrt{n}$ is the error-bar of the mean from Thm 3.

Proof of general distribution : To prove it, say the true median is ν_0 and samples $\{x_1, \cdots, x_n\}$ are sorted as increasing order. For an integer m, set $n = 2m+1$, because $n \gg 1$. The sample median is then $\nu = x_{m+1}$. The probability of having sample median ν is calculable by noting that we have to sample m of sample with $x_i \leq \nu$ and the other m with $x_i > \nu$:

$$f(\nu) = \frac{n!}{m!m!}q^m(1-q)^m \ ,$$

where $q = q(\nu) = \mathbb{P}\{X_i \leq \nu\} = \int_{-\infty}^{\nu} p(x)dx$. $p(x)$ is the pdf of the normal distribution.

The Taylor expansionof q at $\nu = \nu_0$:

$$\begin{aligned} q(\nu) &= q(\nu_0) + \frac{1}{1!}q'(\nu_0)(\nu-\nu_0) + O\big((\nu-\nu_0)^2\big) \\ &\approx \frac{1}{2} + q'(\nu_0)(\nu-\nu_0) \\ &\equiv \frac{1}{2} + p(\nu_0)(\nu-\nu_0) \ . \end{aligned}$$

Then

$$\begin{aligned} f(\nu) &\approx \frac{n!}{m!m!}\left[\frac{1}{2} + p(\nu_0)(\nu-\nu_0)\right]^m \left[\frac{1}{2} - p(\nu_0)(\nu-\nu_0)\right]^m \\ &= \frac{n!}{m!m!}\left[\frac{1}{4}\big(1 - 4p(\nu_0)^2(\nu-\nu_0)^2\big)\right]^m \\ &= \frac{n!}{m!m!4^m}\left[1 - \frac{4mp(\nu_0)^2(\nu-\nu_0)^2}{m}\right]^m \\ &\approx \frac{n!}{m!m!4^m} e^{-4mp(\nu_0)^2(\nu-\nu_0)^2} \ . \end{aligned}$$

This has the form of Gaussian distribution ($\propto e^{-(x-\mu)^2/2\sigma^2}$) with mean ν_0 and variance $\frac{1}{4(n-1)p(\nu_0)^2}$ since $m = (n-1)/2$. Thus,

$$\nu \sim \mathcal{N}\left(\nu_0, \frac{1}{4(n-1)p(\nu_0)^2}\right) \quad \rightarrow \quad \Delta\text{med} \approx \frac{1}{2p(\nu_0)\sqrt{n-1}} \ . \tag{1.11}$$

Q.E.D.

Proof of Gaussian distribution : Now if $p(x)$ is a Gaussian, median is equal to mean, so the pdf becomes $p(\nu_0) = p(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$, and

$$\nu \sim \mathcal{N}\left(\nu_0, \frac{\pi/2}{n-1}\sigma^2\right) \ .$$

If we denote $A = \sqrt{(n-1)/(\pi/2)}$, the statistic $Z = A\frac{\nu-\nu_0}{\sigma}$ will follow the standard normal distribution. On the other hand, $V = \frac{(n-1)s^2}{\sigma^2}$ follows the chi-squared distribution of degrees of freedom $(n-1)$ by Thm 2. Then

$$T = \frac{Z}{\sqrt{V/(n-1)}} = A\frac{\nu-\nu_0}{s}$$

will follow a t-distribution with degrees of freedom $(n-1)$ by Def 2. T is nearly a Gaussian when n is large (rule-of-thumb: when $n \gtrsim 30$), so

$$\nu \dot{\sim} \mathcal{N}\left(\nu_0, \frac{\pi/2}{n-1}s^2\right) \quad \rightarrow \quad \Delta\text{med} \approx \sqrt{\frac{\pi/2}{n-1}}s \ . \tag{1.12}$$

Q.E.D.

1.5 Meaning of the Confidence Interval

I circunvented the definition of the CI so far. It is defined as

Def 3 (*Confidence Interval; CI*)

> A confidence interval of significance level α of a parameter X is the interval of X value such that if we conduct the identical parameter estimation process in many *parallel universes* (i.e., ensemble) which have the identical true value X_{true}, fraction of $(1-\alpha)$ of such universes could have calculated a CI such that the X_{true} is included in that CI.

This is not a practically meaningful definition, but philosophically important. Let me give an example to elaborate the meaning of this.

Ex 3 (*Meaning of the Confidence Interval*)

Imagine there are many *parallel universes* and the same observer observes the same star in each universe. Because of random errors (which we will learn later but that includes Poisson photon noise, readout noise, etc), the observer at each universe will obtain slightly different observational results. The observers will observe the star for n times, and calculate their own mean and error-bar as $\bar{X}$ and $S/\sqrt{n}$ from their observations. Note again that $S/\sqrt{n}$ is the uncertainty of the *mean* value.

For visualization, the true magnitude of the star is $\mu = 14.0^{\mathrm{m}}$ and true standard deviation of the observation $\sigma = 0.2^{\mathrm{m}}$ with $n = 9$ are used for the generation of fig. 1.3. The universe ID 0 of both the upper and lower panels show roughly $m = 14.0^{\mathrm{m}} \pm 0.1^{\mathrm{m}}$. That is, when the true magnitude and true standard deviation are given as 14.0^{m} and 0.2^{m}, an observer at one of the universes (in our case universe ID 0) got $m = 14.0^{\mathrm{m}} \pm 0.1^{\mathrm{m}}$, and that is our universe.

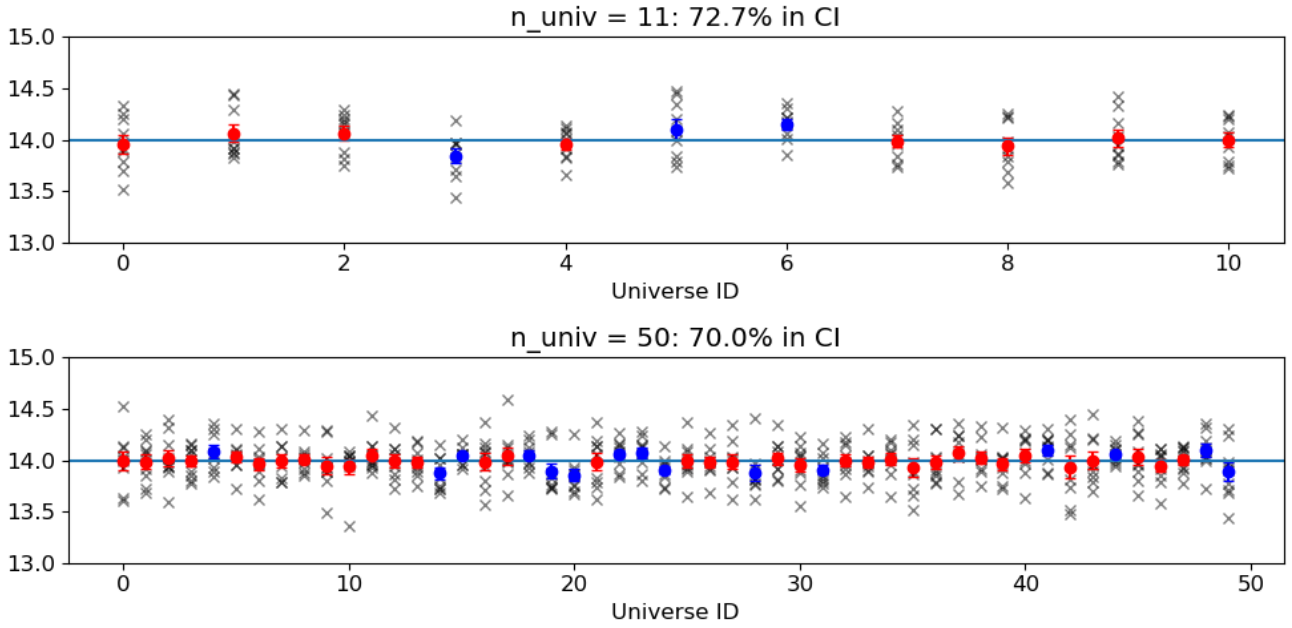

Figure 1.3: Simulation to show the concept of CI and CLT. Black crosses are the observation in each universe, and the circles are the mean and its uncertainty from the CLT $S/\sqrt{n}$. Red circles are the univserses where the confidence interval contains the true mean $\mu = 14.0^{\mathrm{m}}$ and blue circles are the others. In the title, the fraction of the universes which contains the true mean within their own confidence intervals out of `n_univ` simulated universes is shown. This fraction approaches $0.6827\cdots$ as the number of universe gets infinity.

I will close this section with an excerpt from Walpole p.234:

The interpretation of a CI is often misunderstood. It is tempting to conclude that the parameter falls inside the CI with probability of(, e.g.,) 0.95. (But this is not ture.) A CI merely suggests that if the experiment is conducted and data are observed again and again, about 95% of such intervals will contain the true parameter.

1.6 Answer to the Question

In statistics, the *null hypothesis*, denoted H_0, is the one that is the simplest, and that has value if it is *rejected* (not *accepted*). The *alternative hypothesis*, H_1, is another possibility if H_0 is not true. In the

simplest case, H_1 is the complementary of H_0. For example, $H_0 : \mu = 0$ and $H_1 : \mu \neq 0$.

Let's consider the opening question: $m = 14.0^{\mathrm{m}} \pm 0.1^{\mathrm{m}}$. Now you understand that 14.0^{m} means that the mean value* from the observation, and $\pm 0.1^{\mathrm{m}}$ means the uncertainty or the error-bar of the Gaussian probability distribution of that mean value, $S/\sqrt{n}$. If it were not Gaussian, the author should have given more information. Note that the sample standard deviation is $\sqrt{n}$ times the error-bar.

To answer the question, we set the hypotheses†:

$$H_0 : m_1 = m_2 \quad ; \quad H_1 : m_1 \neq m_2 \ , \tag{1.13}$$

where $m_1 = 14.0^{\mathrm{m}} \pm 0.1^{\mathrm{m}}$ and m_2 is the other (from researcher 2). m_1 and m_2 here means the "true" magnitude of the star, not the measured value! If it were the measured value, they are just different, since one is 14.0 and the other is 15.0. The two true means may be different because (1) researchers may have mistakenly observed different star, (2) the star is maybe a variable (including binary), (3) explanet occulted some part of the star, or any other scenario is possible.

Now we have to choose which formula to use. When two measurements with $\bar{x}_1 \pm s_1$ and $\bar{x}_2 \pm s_2$ from the number of observations n_1 and n_2 are given, and μ_1 and μ_2 are the true means of two sampling distributions that will be tested under the hypothesis testing, there are two different formulae:

- The true variances are unknown and different (Satterthwhite approximation‡):

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \tag{1.14}$$

  follows the t-distribution with degrees of freedom

$$\nu \approx \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}} \ . \tag{1.15}$$

- The true variances are unknown but the same:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{1/n_1 + 1/n_2}} \tag{1.16}$$

  follows the t-distribution with degrees of freedom $\nu = n_1 + n_2 - 2$ and

$$s_p = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{\nu} \ . \tag{1.17}$$

Which one will you choose? To answer our original question, it is more reasonable to assume the true variances are not necessarily identical, because they may have used different instruments at different sky conditions. So I will use the first choice.

To apply the formula, you need the number of observations from each researcher, i.e., the n_1 and n_2 values. As I mentioned, it is an ill-defined question since this numbers are not given. So let me assume $n_1 = n_2 = n = 5$. The question can be reformulated as $\bar{x}_1 = 14.0$, $\bar{x}_2 = 15.0$, $s_1 = 0.1\sqrt{n}$, and

*It can actually be the median value, most probable value, or whatever the representative value.

†There are some "standard" ways to set the hypotheses, and it may seem very sudden if you are not familiar with these. Please refer to basic applied statistics textbooks for more examples, e.g., p.246 and p.264 of Walpole et al.

‡SatterthwhiteFE (1946, Biometrics Bulletin, 2, 110), "An Approximate Distribution of Estimates of Variance Components."

$s_2 = 1.0\sqrt{n}$, $0.3\sqrt{n}$, $0.1\sqrt{n}$, where I dropped the magnitude sign for brevity. Under the null hypothesis ($H_0 : \mu_1 = \mu_2$), $\mu_1 - \mu_2 = 0$. For the three s_2 values,

$$\begin{aligned} t = \frac{(14.0 - 15.0) - 0}{\sqrt{(s_1^2 + s_2^2)/n}} \quad &\rightarrow \quad t \;=\; -0.995, \quad -3.162, \quad -7.07 \\ \nu = (s_1^2 + s_2^2)(n-1) \quad &\rightarrow \quad \nu \;=\; 4.08, \quad 4.32, \quad 8.0 \\ &\qquad\quad \approx \quad 4, \quad 4, \quad 8 \end{aligned} \tag{1.18}$$

The values are in the order for $s_2 = (1.0, 0.3, 0.1)\sqrt{n}$ cases. The (two-tail) significance level α test is listed in table 1.1.

Table 1.1: The confidence interval calculation for the example question.

CI %	significance level α and $\frac{\alpha}{2}$	$\nu = 4$ case	$\nu = 8$ case
90 %	0.10 & 0.05	$[-2.1318, 2.1318]$	$[-1.8595, 1.8595]$
95 %	0.05 & 0.025	$[-2.7764, 2.7764]$	$[-2.3060, 2.3060]$
99 %	0.01 & 0.005	$[-4.6041, 4.6041]$	$[-3.3554, 3.3554]$

The answers to the questions are, assuming $n_1 = n_2 = 5$,

- Case A ($t = -0.995$ with $\nu = 4$): The t value is inside of all three CIs. Thus, we cannot reject H_0 under 90 %, 95 %, and 99 % CI criterion (cannot reject under the significance level $\alpha = 0.10, 0.05, 0.01$).
- Case B ($t = -3.162$ with $\nu = 4$): The t value is inside of 90 % and 95 % CI but outside of the 99 % CI. Thus, we reject H_0 with $\alpha = 0.10$ and 0.05, but cannot reject it with $\alpha = 0.01$.
- Case C ($t = -7.07$ with $\nu = 8$): We reject H_0 with all the three $\alpha = 0.10, 0.05, 0.01$.

By *rejecting the null hypothesis*, $H_0 : m_1 = m_2$, we mean that it is unlikely that the true magnitude value of the two studies are identical. By *failing in rejecting the null hypothesis*, we mean that it is impossible to conclude whether m_1 and m_2 are the same under the given significance level.

The results are as expected: For Case A, the error-bar of m_2 is too large, which means it is almost impossible to reject the claim that $m_1 = m_2$, so the we always fail to reject the null hypothesis. For Case C, the error-bars are too small and thus m_1 is of course vastly different from m_2, so the null hypothesis is always rejected.

1.7 Probability Distributions and Poisson Noise

There are a myriad of probability distributions (PDs) in math, due to hard works of mathematicians lived in the era of frequentist statistics. In this section, I will just give some definitions and properties of few of them, which are frequently visited by astronomical literatures.

1.7.1 Binomial Distribution

First, let me define Bernoulli process:

Def 4 (*Bernoulli Process*)

A Bernoulli process is a process (experiments) which has the following properties (Walpole p.102):

1. Repeated trials where each trial (outcome) is classified as either a success (1) or failure (0).
2. The success probability (p) should remain constant throughout the process.
3. Each trial (outcome) must be independent to each other.

A coin-tossing experiment can be a Bernoulli process: The probability of success (heads; H) $p = 0.5$ and failure (tails; T) is $q = 1 - p = 0.5$. Each coin-tossing is independent of each other. The probability to have the outcome "HHT" in this order is $p \cdot p \cdot q = 1/8$.

A binomial distribution arises when we describe the probability distribution of the number X which is the number of success in n Bernoulli trials (Walpole p.102).

Def 5

A random variable X follows a binomial distribution ($X \sim \text{Binom}(n, p)$) has the probability distribution

$$b(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x} \tag{1.19}$$

for the success probability p of n Bernoulli trials.

For the reason why we use a word "binomial", see math textbooks such as Walpole p.103. This distribution has mean, variance, and standard deviation of:

Thm 5 (*Binomial Distribution Mean and Variance*)

A binomial distribution with parameters n and p, it has

$$\text{mean} = np \quad ; \quad \text{var} = npq \quad ; \quad \text{std} = \sqrt{npq}\ . \tag{1.20}$$

Proof of Binomial Dsitribution Mean and Variance: Consider O_i is the outcome of each Bernoulli trial, i.e., 1 if success and 0 if failure, of the i-th trial. Then $X = \sum_{i=1}^n O_i$. The expected value of X is $E(X) = E(\sum_{i=1}^n O_i) = E(O_1) + \cdots + E(O_n)$ since they are independent, and this becomes $E(X) = np$.

For variance, we use a relationship $\text{Var}(O_i) = E(O_i^2) - (E(O_i))^2$. But since the value O_i^2 is identical to O_i, $E(O_i^2) = E(O_i)$, so $\text{Var}(O_i) = p(1-p) \equiv pq$ for $q := 1-p$. $X = \sum_{i=1}^n O_i$ gives $\text{Var}(X) = \sum_{i=1}^n \text{Var}(O_i) = npq$.

1.7.2 Poisson Distribution

The Poisson process is a fancy naming for some special "counting". In astronomy, we count photons (well, actually the CCD counts the photoelectrons) over time. This is why Poisson process is dealt with so much weight in astronomy. In social sciences we can count the death toll over time, in experiments we could count the lattice of a randomly shaped crystal over the distance. A formal definition of Poisson process is*

Def 6 (*Poisson Process 1*)

The Poisson process $N(t)$ for $t \geq 0$ (t can be time, distance, or similar things), with rate λ is defined by

1. $N(0) = 0$
2. $N(t)$ has independent increment
3. $N(t_2) - N(t_1)$ follows Poisson distribution of rate $\lambda(t_2 - t_1)$ for $t_1 < t_2$.

The definition of Poisson distribution is given below. It can be proven that it is identical to

*See, e.g., `http://dept.stat.lsa.umich.edu/~ionides/620/notes/poisson_processes.pdf` and `https://www.probabilitycourse.com/chapter11/11_1_2_basic_concepts_of_the_poisson_process.php`

Def 7 (*Poisson Process 2*)

The Poisson process $N(t)$ for $t \geq 0$ (t can be time, distance, or similar things), with rate λ is defined by

1. $N(0) = 0$
2. $N(t)$ has independent increment
3. $\mathbb{P}(N(h) = 1) = \lambda h + o(h)$ for small h.
4. $\mathbb{P}(N(h) \geq 2) = o(h)$ for small h.

Here $o(h)$ is any function such that $\lim_{h\to 0} o(h)/h = 0$. The last two items in this second definition can be re-phrased like this:

3. For a very small interval h, the probability of one single "counting" (Poisson process) occurs in that interval is proportional to the length of h.
4. For a very small interval h, the probability of two or more "counting" (Poisson process) occurs in that interval is nearly 0.

These must be independent of whether the counting happened outside of the interval, because the second item of the definition states the increment is independent.

Def 8 (*Poisson Distribution*)

A random variable X which describes a Poisson process of rate λ follows the Poisson distribution ($X \sim \text{Pois}(\lambda t)$) and has the following pdf

$$p(x; \lambda t) = \frac{(\lambda t)^x}{x!} e^{-\lambda t} \tag{1.21}$$

for non-negative integer x.

Thm 6 (*Poisson Distribution Mean and Variance*)

The mean, variance, and the standard deviation of the Poisson distribution $p(x; \lambda t)$ are

$$\text{mean} = \lambda t \quad ; \quad \text{var} = \lambda t \quad ; \quad \text{std} = \sqrt{\lambda t}\ . \tag{1.22}$$

The reason we use λt not just λ is clear from Thm 6: the mean changes as t changes, and we want to express $p(x; \text{mean})$ rather than $p(x; \text{mean}/t)$. For instance, if the photon influx is $\lambda = 10\,\text{photons/s}$, the mean of photon count is a function of time as λt, so for $t = 10\,\text{s}$, the pdf is $p(x; 100\,\text{photons})$. If the particles are distributed with mean density along the line of sight $\lambda = 0.1\,\text{particle/m}$, the mean of particles along the line of sight is a function of the reaching distance λd, so the pdf for $d = 5\,\text{m}$ is $p(x; 0.5\,\text{particles})$. We, therefore, **frequently denote a Poisson distribution of *mean*** m, instead of *rate* λ.

1.7.3 Relationships Between Distributions

There are some relationships I want to emphasize because they will be used to understand statistical processes in astronomy.

Thm 7 (*Binomial to Poisson*)

For $X \sim \text{Binom}(n, p)$, if (1) $n \to \infty$, (2) $p \to 0$ and (3) $np \xrightarrow{n\to\infty} \mu < \infty$,

$$X \dot{\sim} \text{Pois}(x; \mu = np) \tag{1.23}$$

Although I did not describe the Gaussian (normal) distribution, I am assuming you are familiar with it. One thing to note is that, both binomial and Poisson distributions are discrete ones, while normal

distribution is continuous one. Therefore, the approximations given below, which approximate binomial and Poisson to normal distribution, should be used with care when the discreteness is important.

Thm 8 (*Binomial to Normal*)

For $X \sim \mathrm{Binom}(n, p)$, it has mean np and variance npq (see Thm 5). Then, if $n \to \infty$,

$$X \dot{\sim} \mathcal{N}(np, npq) \tag{1.24}$$

Thm 9 (*Poisson to Normal*)

For $X \sim \mathrm{Pois}(\lambda t)$, it has mean λt and variance λt (see Thm 6). Then, if $\lambda t \to \infty$,

$$X \dot{\sim} \mathcal{N}(\lambda t, \lambda t) \tag{1.25}$$

Although I will omit here, the proof uses Stirling's formula $x! \approx \sqrt{2\pi x} x^x e^{-x}$ and $\ln[(1+\varepsilon)^{\lambda t(1+\varepsilon)+1/2}] \approx \lambda t\varepsilon + \lambda t\varepsilon^2/2$, where $x = \lambda t(1+\varepsilon)$ with $\lambda t \gg 1$ and $\varepsilon \ll 1$. Stirling's formula has relative error of less than 1 % when $x = 1$ and is almost negligible for any x of interest in many cases. Considering all the approximations in the proof, we can safely assume that Poisson distribution is quite Gaussian (normal) in many astronomical contexts, unless x is very small. In practical sense, $\lambda t > 100$ is large enough, and in astronomical uses, an error, i.e., error-bar, arise from small mean ($\lambda t \lesssim 50$), is not very large compared to other approximations we use in date reduction.

Ex 4 (*Babies in Hospital 1*)

(Problem provided by prof. Jae-Kwang Kim at the Dept. of Math at KAIST in 2018)

There are two hospitals A/B, and new babies are born every day. At A/B, $N_A = 45$ and $N_B = 15$ babies are born everyday, respectively. They both recorded the number of days when $\geq 60\,\%$ of the babies have XX chromosome, say N'_A and N'_B. Which will have higher value? For simplicity, assume a baby can only have either XX or XY chromosome, and the probability of being either of these is the same ($P_X = P_Y = 0.5$).

For a hospital, assuming the events are independent, the probability of having $\geq 60\,\%$ of XX babies is

$$p_{A,B} = \sum_{k=\lceil 0.6N_{A,B}\rceil}^{N_{A,B}} \binom{N_{A,B}}{k} P_X^k P_Y^{N_{A,B}-k} = 0.5^{N_{A,B}} \sum_{k=\lceil 0.6N_{A,B}\rceil}^{N_{A,B}} \binom{N_{A,B}}{k},$$

for the case of hospitals A/B, respectively. A simple calculation gives this probability are $p_A = 0.0899$ ($\lceil 0.6N_A\rceil = 27$ for A) and $p_B = 0.304$ ($\lceil 0.6N_B\rceil = 9$ for B) for hospitals A and B, respectively. For 1 year, 365 days, we can calculate $N'_A = 32.8$ and $N'_B = 111.0$. Therefore, $N'_B > N'_A$.

In the problem above, we found hospital B will have more XX babies than A. But with what confidence? In principle, you can just calculate using binomial distribution, but there's a simpler workaround.

Define a random variable N_i, such that it is 1 when $\geq 60\,\%$ of the babies have XX chromosome on the i-th date; otherwise, 0. Then the expectation of N_i is $E(N_i) = p_{A,B}$, for the case of hospitals A/B, respectively. Since N_i^2 must have identical value as N_i, $E(N_i^2) = p_{A,B}$. Then the variance becomes $V(N_i) = E(N_i^2) - E^2(N_i) = p_{A,B} - p_{A,B}^2 = p_{A,B}(1 - p_{A,B}) \equiv p_{A,B}q_{A,B}$.

The total number of days is $N_{A,B} = \sum_{i=1}^{365} N_i$, and if N_i are all independent,

$$\mu_{A,B} = E(N_{A,B}) \quad = E\left(\sum_i N_i\right) = np_{A,B}$$
$$\sigma_{A,B} = V(N_{A,B}) \quad = V\left(\sum_i N_i\right) = n \times V(N_i) = np_{A,B}q_{A,B}$$

where $n = 365$. Note that N follows binomial distribution, which is approximated as Gaussian distribution in this case (roughly speaking when p is not close to 0 or 1 and np is large enough). Therefore, $N_{A,B} \dot{\sim} \mathcal{N}\left(\mu_{A,B}, \sigma^2_{A,B}\right)$, or

$$N_A \dot{\sim} \mathcal{N}(32.8,\, 29.9) = \mathcal{N}(32.8,\, 5.46^2) \quad ; \quad N_B \dot{\sim} \mathcal{N}(111.0,\, 77.2) = \mathcal{N}(111.0,\, 8.79^2)$$

Since this is a case when μ and σ are known, we can use standard normal z statistic for confidence calculation. Take null hypothesis as $H_0 : d = N_A - N_B = 0$ and then $z = \frac{\mu_A - \mu_B}{\sqrt{\sigma_A^2 + \sigma_B^2}} = 12.4$. That means, we are confident with 12.4-σ, which is much larger than 99.999 %. More realistic null hypothesis is $H_0 : d = N_A - N_B < 0$, but this will have only small effect to the result.

1.7.4 Usage in Astronomy

As mentioned, photon counting is a key process in observational astronomy, and it is tightly bound to Poisson process. I will try to clarify this concept by the examples below.

Ex 5 (*Photon Counting and Poisson Process*)

The photon counting process of astronomy resembles the Poisson Process. If $N(t)$ is the number of photons during the exposure time t, it is trivial that the first two items of Def 7 is satisfied. The number of photons arriving in an infinitesimal time interval $h = dt$ must be 0 or 1, but not larger than 1. To illustrate, if a total 10,000 photon should come during 10 sec of exposure, we can take $h = 1\,\text{ns}$ ($\ll 10\,\text{s}/10,000 = 1\,\text{ms}$) to meet this condition. Thus, photon counting is a Poisson process.

In reality, however, photon counting is done by electronic devices, such as CCD. Before going deeper into the subtlety, let me give another example dealing with photon noise:

Ex 6 (*Photon Counting and Poisson Noise*)

We saw the photon counting is a Poisson process; let me further assume the photoelectron counting process on CCD is, too (photoelectrons are generated by these photons when they excite bound electrons). Thus, we can use eqs. (1.21) and (1.22).

For example, if we collected $\lambda t = 10,000$ electrons during the exposure time t, the 1-σ uncertainty (1-σ CI) or the standard deviation of this counting is $\sqrt{10,000} = 100$, so we denote $N = 10,000 \pm 100$. Since the λt is large enough, Thm 9 says this is just a Gaussian distribution of mean 10,000 and standard deviation 100 (1 % of the mean). Since magnitude is (const) $- 2.5 \log_{10}$(count), differentiation gives a first-order estimation of the uncertainty in the magnitude $\Delta m = \left| -\frac{2.5}{\ln 10} \frac{\Delta(\text{count})}{\text{count}} \right| = 0.011^{\text{m}}$ (error-bar), which is small enough for some scientific purposes.

Note that the term $\frac{\Delta(\text{count})}{\text{count}} = \frac{1}{\sqrt{\text{count}}}$ is inversely proportional to the square root of count itself. The more we collect the count, i.e., the longer the exposure time, we will get smaller error-bar.

Now a tricky part comes in:

Ex 7 (*Photon Counting and Poisson Noise in ADU*)

In CCD, the number of photon (N_γ) will generate certain number of photoelectrons (N_e), and this is measured by the electric potential (roughly speaking, potential $= V_e \propto N_e e$ where e is the electron charge). This voltage will be translated into integer numbers by electric circuit, and this integer is said to have a unit ADU (analog-to-digital unit) or DN (data number). In summary, a conversion of this order happens: $N_\gamma \to N_e \to V_e \to N_{\rm ADU}$. We define the (electron) gain g [electrons/ADU] such that $N_e = gN_{\rm ADU}$.

If $g = 2\,\mathrm{e/ADU}$ and $N_e = 10,000$ as in the previous example, $N_{\rm ADU} = 5,000$. If $N_{\rm ADU}$, i.e., the ADU counting is a Poisson process, we must hace $N_{\rm ADU} = 5000 \pm 71$, which gives magnitude error $\Delta m_{\rm ADU} = \left|-\frac{2.5}{\ln 10}\frac{71}{5000}\right| = 0.015^{\rm m}$. The error-bar has changed! What has just happened?

The key here is that we only checked the photon counting process is a Poisson process. Rigorously speaking, photoelectron counting and ADU counting are **not Poisson processes**.

Then what process it is for photoelectron counting?

Ex 8 (*Photoelectron Follows a Binomial Distribution*)

Any device which collect photon does not actually collect photon, but they output *something* which is proportional to the number of photons incident to it. In CCD, the output is the number of electrons which is measured by electric potential. The key here is the *conversion of photon into* something *must be a probabilistic process*. In CCD, the conversion happens with a paramter called the quantum efficiency (QE). Therefore, for a fixed number of photon, N_γ, which follows a Poisson distribution $N_\gamma \sim \mathrm{Pois}(N_\gamma)$, the outcome, N_e, will follow a binomial distribution $N_e \sim \mathrm{Binom}(N_\gamma, QE)$. Verbally put, "photoelectron will follow a binomial distribution with the paramter n which follows a Poisson distribution and the parameter p which is believed to be fixed thanks to engineering process".

Although I will not give a proof, $N_e \sim \mathrm{Binom}(n = \mathrm{Pois}(N_\gamma), p = QE)$ is approximated as $N_e \dot\sim \mathrm{Pois}(N_\gamma QE)$. In observation, since $N_e \approx N_\gamma QE \gg 1$, $N_e \dot\sim \mathcal{N}(N_e, N_e)$ (Thm 9).

How about the ADU count?

Ex 9 (*ADU is Approximately Gaussian*)

The ADU counting is a very strange thing: For electron gain g, ADU increase by 1 for every g electrons, i.e., a step function which is very difficult to deal with in analytical sense. Regardless of its final distribution, it is clear that this conversion to the original N_e will destroy the original distribution. Fortunately, the count N_e or $N_{\rm ADU}$ is large enough in most observations, so that the discreteness of conversion is ignorable. Hence, we can treat them as if a smooth continuous convertable variables by $N_e = gN_{\rm ADU}$. Then

$$N_{\rm ADU} \dot\sim \mathcal{N}(N_e/g, N_e/g^2) \quad \text{or} \quad N_{\rm ADU} \dot\sim \mathcal{N}(N_{\rm ADU}, N_{\rm ADU}/g) \tag{1.26}$$

That means $N_{\rm ADU}$ follows a Gaussian, but not like usual Poisson distribution (because ADU counting is not a Poisson process). Going back to Ex 7, we have to divide the error-bar of the count by the gain, so $\Delta m_{\rm ADU} = \left|-\frac{2.5}{\ln 10}\frac{\sqrt{5000/g}}{5000}\right| = 0.011^{\rm m}$, which is identical result as Ex 6.

For a mathematical treatment of the ADU conversion (step function), you may find some research papers useful*.

*See, e.g., Merline W. J. & Howell S. B. 1995, Exp. Astron. **6**, *163*, "A REALISTIC MODEL FOR POINT-SOURCES IMAGED ON ARRAY DETECTORS: THE MODEL AND INITIAL RESULTS".

1.8 Error Propagation

When we have to calculate the uncertainty of a parameter derived from other parameters that have their own uncertainties, what we usually use is the *error propagation*. For example, $g' = 14.0 \pm 0.1$ and $r' = 15.0 \pm 0.1$, what is the uncertainty of the $g' - r'$ color? Many astronomers will say it is -1.0 ± 0.14, where the error here is $\sqrt{\sigma_{g'}^2 + \sigma_{r'}^2}$ (σ is the error-bar). If the temperature is $T = 100 \pm 1$K, what is the fractional uncertainty of bolometric luminosity ($L = \sigma_{\rm SB} T^4$)? Now people will say $\Delta L/L \approx \frac{4\sigma_{\rm SB} T^3 \Delta T}{\sigma_{\rm SB} T^4} = 4\Delta T/T = 4\%$.

The way how we do this is that, for a function $f = f(x|a, b, c)$ where (a, b, c) is a set of parameters:

$$(\Delta f)^2 \approx \left(\frac{\partial f}{\partial a}\Delta a\right)^2 + \left(\frac{\partial f}{\partial b}\Delta b\right)^2 + \left(\frac{\partial f}{\partial c}\Delta c\right)^2 . \tag{1.27}$$

Here, $(\Delta f)^2$ is the variance of the function f, ΔX is the standard error (such that $X \pm \Delta X$ is the 1-σ CI for X), and we frequently use Δf, the standard error of f, as *the error-bar*. If any parameters are dependent, we need a covariance term σ_{ab} for instance, and this is one of the reasons why the above error-propagation is only an approximation. This formula itself is similar to getting the gradient. In different notation, e.g., sum of pixel values, $I := \sum_i I_i$, which is reasonable to assume I_i's are independent, the identical formula becomes

$$(\Delta I)^2 \equiv \left(\Delta\left(\sum I_i\right)\right)^2 = \sum_i (\Delta I_i)^2 . \tag{1.28}$$

Note that this is only an approximation, and in reality, we need more complicated calculation. Also it is impossible to know how accurate this approximation is, before you conduct detailed calculation. The reason astronomers use it in spite of the pitfall is because most cases we are only interested in the *rough* estimation of the error-bar.

1.9 The Chi-Square Minimization

You may have heard of it, or simply the *least square* something. Least square or the chi-square minimization is a process to find the set(s) of model parameters which has the biggest (or bigger than threshold) *likelihood* in Bayesian statistics. The set of "best" model parameters is the one with the largest likelihood, so we call it the maximum likelihood estimator, **MLE**.

Consider independent and identically distributed (*i.i.d.*) random variables*, e.g., flux as a function of λ. It is independent because the measurement at each wavelength itself is independent from that of any other wavelength. It is identically distributed because and we will assume each error-bar is Gaussian. The second assumption is not necessarily correct, and actually that happens many times (e.g., flux is Gaussian distributed but magnitude is not, because it is log of Gaussian distribution).

We will study Bayes' theorem later, but let me introduce it first to grap the meaning of chi-square statistic:

$$P(\theta|D, I) = \frac{P(D|\theta, I)P(\theta|I)}{P(D, I)}$$
$$P(\theta|D, I) \propto P(D|\theta, I)P(\theta|I)$$

*Random variables X_i for $i = 0, \cdots, N$ are called i.i.d. if they (1) follow identical probability distribution, say $f(x)$, and (2) $f(X = x) = f(X_1 = x_1) \times \cdots \times f(X_N = x_N)$, where $x = (x_1, \cdots, x_N)$ (same for X).

When we have data points, (x_i, y_i), which are all independent, and $y \sim N(y_i, \sigma_i^2)$. Then one of the most naïve goodness-of-fit statistic variables can be the square-sumed error:

$$\text{SSE} = \sum_i (y_i - f(x_i|\theta))^2 \tag{1.29}$$

Since y_i are all Gaussian, the probability of obtaining such data is obtained from the multiplication law:

$$P(D = \{y_i\}) = \prod_i A e^{-(f(x_i|\theta)-y_i)^2/2\sigma_i^2}$$

or taking log:

$$\ln P(D = \{y_i\}) = \ln A - \sum_i \frac{(f(x_i|\theta) - y_i)^2}{2\sigma_i^2} \tag{1.30}$$

Now getting parameters which maximize P is identical to (1) getting them which maximize $\ln P$, and this is also identical to (2) getting them which mininize the summation in the exponent. Therefore, the exponent, which is chi-square statistic:

$$\chi^2 = \sum_i \frac{(y_i - f(x_i|\theta))^2}{\sigma_i^2} \ , \tag{1.31}$$

should be minimized to get the most probable answer to the fitting problem. If the error-bars of all data points are identical, σ_i is nothing but a constant, so minimizing chi-square is identical to minimizing SSE.

1.10 Interpolation

When we fit an analytic function to the data, we usually use a statistic (e.g., the chi-square value or the BIC) that indicates the goodness-of-fit (also called figure of merit). But what if we don't know or not interested in the true analytic function behind the scene? For example, for aperture trace or sense function in spectroscopy, we have absolutely no idea what is the true analytic function. In this situation, **interpolation** comes in.

Def 9 (*Interpolation*)

An interpolation function is a function that smoothly connects *all* the observed data points.

For instance, a linear interpolation is a function which connects all the data points with linear segments as in fig. 1.4. In cubic spline, you fit a 3rd order polynomial function (total 4 unknowns). For points (x_i, y_i) where $i = 0 \ldots n$, we want to obtain n spline curves $S_i(x)$ for $i = 0 \ldots n-1$. All such interpolation functions should be *smoothly connected* to each other by

- $S_i(x_i) = y_i \quad (i = 0, \ldots, n-1)$.
- $S_{i-1}(x_i) = y_i \quad (i = 1, \ldots, n)$.
- $S_i'(x_i) = S_{i-1}'(x_i) \quad (i = 1, \ldots, n-1)$.
- $S_i''(x_i) = S_{i-1}''(x_i) \quad (i = 1, \ldots, n-1)$.
- $S_0''(x_0) = S_{n-1}''(x_n) = 0$.

In python, you can use

```
from scipy.interpolate import UnivariateSpline
# linear spline interpolation
interp_1 = UnivariateSpline(x, y, s=0, k=1)
# cubic spline interpolation
```

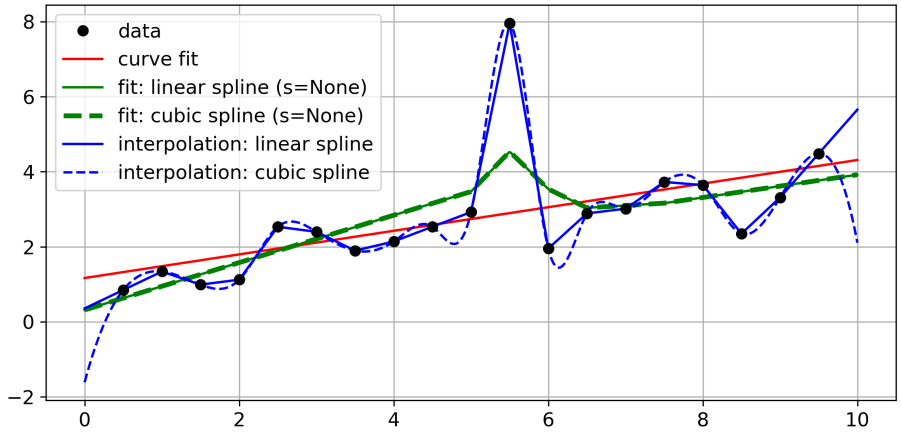

Figure 1.4: A display of linear/cubic spline interpolations and spline fittings.

```
interp_3 = UnivariateSpline(x, y, s=0, k=3)

# use it as y_interp_1 = interp_1(x_values)
```

The parameter `s` is the smoothing factor. That is,

$$\sum_{i=1}^{n} w_i^2 (y_i - S(x_i))^2 \leq s \ . \tag{1.32}$$

If no error-bar is present, we can just ignore w_i (weights), and if there exists σ_i, we can use $w_i = 1/\sigma_i$, and the above equation becomes nothing but a chi-square. In that case, the χ^2 smaller than the total degrees of freedom (roughly the number of data points) is a good measure, so `s = len(x)` is a good guess for spline fitting. If you put $s = 0$, it means $S(x_i)$ must be the same as y_i, i.e., the fitted function must go through all the data points: this is the interpolation.

From the figure, cubic spline interpolation is more smooth and the linear one looks too discrete, so you may always want to use cubic version. A caveat is that cubic spline has large fluctuation when there is any large scatter, as seen from $x = 5$ and $x = 6$. This "shooting" effect is severe when you fit a sense function (spectroscopic flux calibration) to a spectroscopic standard star: if the star has an absorption line, the cubic fit will be problematic near that wavelength region.

1.11 Change of Variables

In a formal error-propagation calculation, the change of variables is very important, because sometimes the physical paramters which describe the function you are interested in may be different from the observables which you obtain from expeirments or observations. We actually do the change of variables without much mathematical instroduction. An example is when $m = -2.5 \log_{10} I$, we differentiate it to get $\Delta m = \frac{2.5}{\ln 10} \frac{\Delta I}{I} \approx 1.086 \frac{\Delta I}{I}$ (as in Ex 6). This way, we changed the error of the variable I to that of m. Formally, if a conversion happens from parameters (or random variables) $X_1, \cdots, X_M$ to $Y_1, \cdots, Y_M$, we use Jacobian

$$\begin{aligned} f(Y_1, \cdots, Y_M) \delta Y_1 \cdots \delta Y_M &= g(X_1, \cdots, X_M) \delta X_1 \cdots \delta X_M \\ f(Y_1, \cdots, Y_M) &= \left| \frac{\partial (Y_1, \cdots, Y_M)}{\partial (X_1, \cdots, X_M)} \right| g(X_1, \cdots, X_M) \end{aligned} \tag{1.33}$$

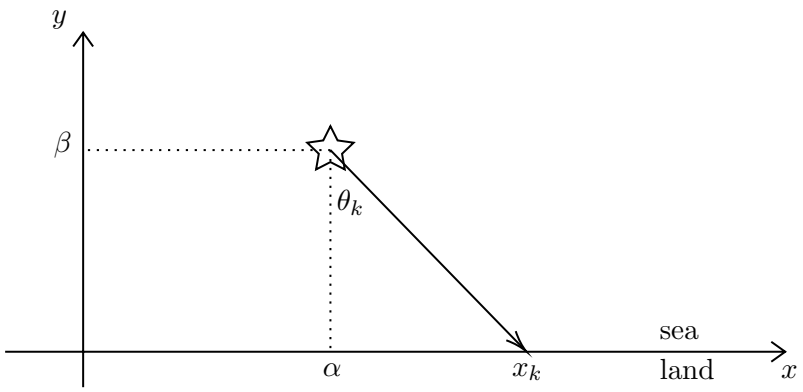

Figure 1.5: A schematic diagram for the lighthouse problem.

where, e.g.,

$$\left|\frac{\partial(Y_1,\, Y_2)}{\partial(X_1,\, X_2)}\right| := \begin{vmatrix} \frac{\partial Y_1}{\partial X_1} & \frac{\partial Y_1}{\partial X_2} \\ \frac{\partial Y_2}{\partial X_1} & \frac{\partial Y_2}{\partial X_2} \end{vmatrix} . \tag{1.34}$$

Ex 10 (*Cauchy distribution: Lighthouse Problem*)

(This is taken from Sivia (§2.4), which is recited from GullSF 1988, which also adopted it from "Cambridge Part 1A examples sheet".)

Consider a lighthouse at $(x, y) = (\alpha, \beta)$. The lighthous emits an infinitely thin lightray to a uniformly random direction θ (angle counterclockwise from $-y$ direction), as in section 1.11. What is the probability distribution of x_k for the k-th position on the shore which receives the light?

Sicne θ is uniformly distributed, the pdf is $f(\theta_k|\alpha, \beta) = 1/\pi$ (uniform over $\pm\pi/2$; You may use $1/2\pi$ considering uniform over $\pm\pi$, but that will change the result only by up to a multiplication of a constant, 2). Since $\tan\theta_k = (x_k - \alpha)/\beta$, the change of variable gives

$$g(x_k|\alpha, \beta) = \frac{\partial\theta_k}{\partial x_k} f(\theta_k|\alpha, \beta) = \frac{1}{\pi}\frac{\beta}{\beta^2 + (x_k - \alpha)^2} .$$

This is the Cauchy distribution with the "γ" parameter equals β.

Mathematically, Cauchy distribution does not have a mean (simple proof can be $\lim_{a\to\infty}\int_{-a}^{+a} g(x)dx = 0$ is not equal to $\lim_{a\to\infty}\int_{-2a}^{+a} g(x)dx = \infty$). Also the standard deviation cannot be defined since the mean of x^2 also diverges. If μ and σ are undefined, the CLT in the form of Thm 1 cannot be used.

Chapter 2

Idea of Photometry

2.1 The Point-Spread Function (PSF)

Stars must be point sources. But from our experience, we know that they are never a point, but a usually nearly circular extended sources on the CCD. The image on the CCD when a single perfect point source is imaged, is called the **point spread function (PSF)**. There are basically two mechanisms responsible for this: diffraction and seeing. Although diffraction can also be a part of seeing, it is somewhat special than others: It is just impossible to eliminate the diffraction, while other sources of seeing can be removed (see below).

2.1.1 Diffraction

When a light ray passes through a finite-sized aperture, diffraction must occur. Due to this, a point source is blurred to a finite size after it passed through an aperture (e.g., telescope or DSLR camera aperture).

Ex 11 (*Diffraction*)

For $D = 1\,\mathrm{m}$ in visible (550 nm wavelength), the diffraction limit is $1.22\frac{\lambda}{D} = 0.14\,\mathrm{arcsec}$. You may be familiar with this formula. Using the proportionality, you can memorize it as $\theta_{\mathrm{min}}^{(550\,\mathrm{nm})} = \frac{0.14}{D[\mathrm{m}]}\mathrm{arcsec}$.

It is more complicated in reality: the mirrors and other optical instruments located in the optical path affect the final shape of a point source by diffractions. See fig. 2.1 for the examples of simulated PSFs from optics, i.e., when the seeing effects are neglected.

As you can see in fig. 2.1, although the detailed PSF are different (mid row), the long-range features are weaker compared to the central nearly circular features (bottom row) in cases except for the last column, which is not of our interest.

2.1.2 Seeing

Even if we ignore the diffraction, a point source is blurred due to two main reasons: (1) air turbulence and (2) discontinuous movement of the telescope since the telescope gear is not a perfectly smooth continuous ideal machine. The blurring caused by these are called *natural (atmospheric)* and *telescopic seeing*, respectively. The atmospheric seeing makes a point source an extended circular source (for details, search

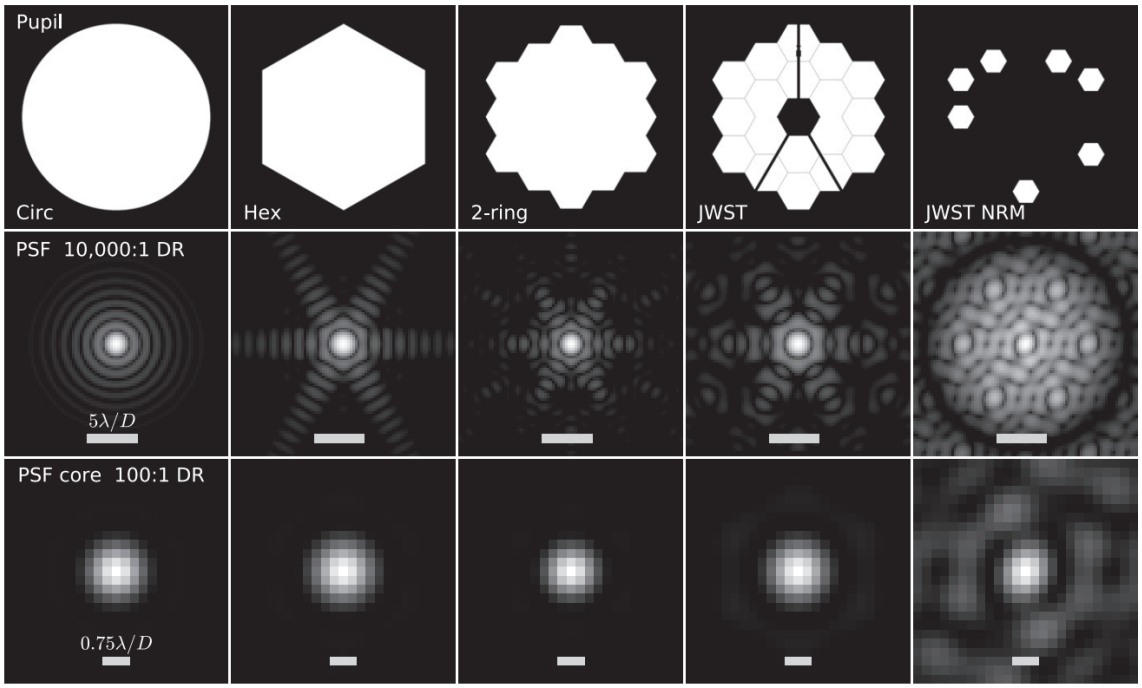

Figure 2. Comparison of the point spread functions (PSFs) of different shaped apertures. The top row shows apertures as labeled. The middle row shows the corresponding PSFs with 10,000:1 dynamic range at consistent stretch and pixel scale. The bottom row shows the cores of the same PSFs with 100:1 dynamic range. Note the relatively small core and deep null surrounding the *JWST* NRM PSF.

Figure 2.1: Example of PSFs. For circular aperture, the PSF shows concentric diffraction pattern as we learned in general physics course. For JWST case, you can see the secondary mirror and the mirror holding pipe (black things) makes the PSF different compared to the 2-ring case. The bottom row is not just a zoom-in of the PSF, but also the stretch (contrast) is different from the mid row (DR means the dynamic range). Direct excerpt from Figure 2 of FordKES+, 2014, ApJ, 783, 73.

for the theory of, e.g., Andrei Kolmogorov), while the telescope seeing *may* make it very elongated or irregular depending on the quality of the tracking system. If the tracking system is well established on a telescope, the final telescope seeing effect is quite circular.

All these effects can be removed if we move our telescope in space. The atmospheric turbulence is now gone, and the telescope can be very stable such that the telescopic seeing is ignorable. This is what is called the diffraction-limited case.

2.1.3 Summary of PSF

To summarize:

$$\text{PSF} = \text{diffraction} + \text{seeing}$$
$$\text{seeing} = \text{atmospheric seeing} + \text{telescopic seeing}$$

and practically speaking,

$$\text{PSF shape} \approx \text{circular} \ .$$

This is why we can talk about the circular aperture photometry. Also

$$\text{diffraction} \ll \text{other sources of seeing}$$
$$\text{(unless adaptive optics or space telescope used)} \ .$$

Actually, for many ground-based observations, it is not necessarily important to distinguish diffraction from other seeing effects. I feel like people tend to use the word "PSF" when the PSF is measured rather accurately (e.g., PSF photometry), and use the word "seeing" when we assumed circular PSF (e.g., aperture photometry).

2.1.4 Seeing Disc Size

As mentioned above, when we assume circular PSF (as we will do frequently in aperture photometry), we use the word "seeing" for the PSF. *Seeing disc* is the word for the circular shape of the PSF, and the *seeing disc size* is a measure of the size of that circle. The definition of the "size" is not trivial, but many people use the FWHM of the stellar radial profile*. This seeing disc size measurement is used to find the focal position and give a sense to the data quality.

The seeing disc FWHM is dependent on wavelength, weather, airmass, etc. At Seoul National University observatories in Seoul, it is roughly 1 to 3 arcsec at Bldg. 46 and 3 to 6 arcsec at Bldg. 45. For Subaru telescope at Hawai'i, for example, it is less than 1 arcsec†. Compared to the diffraction size $\theta_{\min}$, these are much larger. It can go even below if we use adaptive optics (then diffraction $\approx$ total seeing and non-circular PSF may be important).

The fact that the seeing disc size (i.e., size of the circular PSF) on the CCD is normally much larger than the diffraction scale hint that atmospheric and telescope seeing are very important. Since the telescope seeing is roughly constant over time for continuous tracking of a good telescope, the change in seeing can also strongly dependent on the weather condition. So astronomers use seeing disk size (usually rough estimate of FWHM) as a proxy of the weather condition.

It can also be used as a data-quality-indicator.

Ex 12 (*Contamination by Other Object*)

If you have your faint target near a bright stellar source with 1 arcsec separation and if the seeing disc FWHM is roughly 1 arcsec, your faint target is contaminated by the nearby bright star. A possibility is to do photometry of "your target + bright star" and subtract the flux of the "bright star" from catalog or future/previous observations. In this case, although you can get the flux of the target, its uncertainty may increase significantly.

Ex 13 (*Slit Width Determination*)

For spectroscopy, seeing disc size may determine the slit width you have to use to collect the target's light into the slit. It will be dealt in the spectroscopy chapter.

Ex 14 (*Finding Focal Position from Seeing FWHM*)

First find a random bright star that do not vary over short period of time (e.g., about 1 hour). You then expose and check the FWHM of the PSF (assuming it is circular). Now tune the focal position, e.g., by moving the position of the secondary mirror, and expose again, and check the FWHM again. Doing this repeatedly, you can get the FWHM as a function of the focal position. When the FWHM becomes the minimum, that is *the* focal position (fig. 2.2). Normally this curve is roughly second order polynomial.

*Radial profile means the pixel values of the star image as a function of the distance from the center of the star: $I(r)$. From many free/commercial products, it is rather easily drawn, such as `ginga`, `SAO ds9`, or `Maxim DL 6`. They also automatically calculates FWHM for you. The details on how to find the "center" will be studied later. Here, just think about something like Gaussian curve.

†`https://www.naoj.org/Observing/Telescope/ImageQuality/Seeing/`

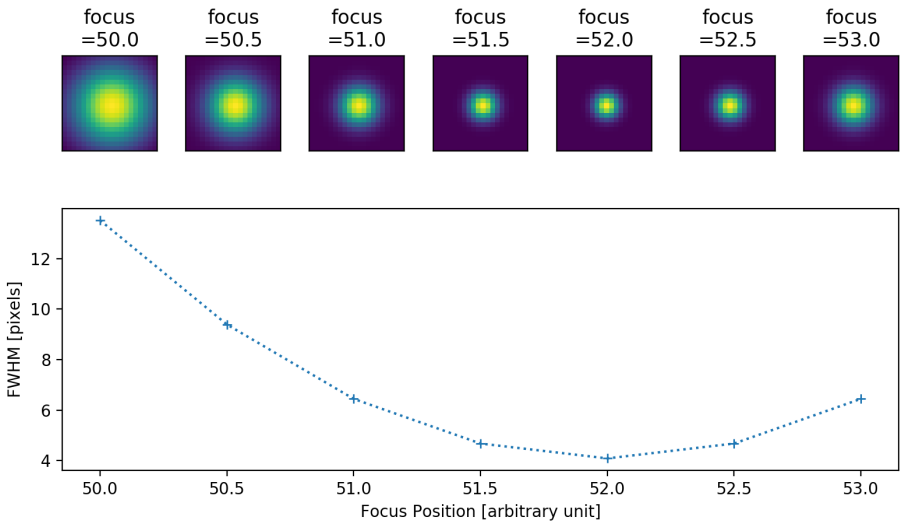

Figure 2.2: Finding focal position. While tuning the focal position value (x-axis), check the FWHM of the stellar profile. In this fake simulation, it is found that the focal position of 52.0 is *the* focal position with FWHM ~ 4 pixels. Multiplying with the pixel scale will give the FWHM in angular units (e.g., if the pixel scale is 0.8 arcsec/pixel, FWHM after the focusing is $4\,\text{pixel} \times 0.8\,\text{arcsec/pixel} = 3.2\,\text{arcsec}$)

Our goal in observation is to "gather all the photons from the target". For a faint source, you have to tune the focal position very accurately, so that the signal from the source gathers to as small number of pixels as possible. Otherwise, the signal will be buried in sky values or other noise sources (dark current or readnoise) and our goal may not be achieved. The non-circular PSFs for high-quality data shares this same philosophy: the "sprays" of PSF should be analyzed. But for a very bright source, you may even intentionally defocus before the exposure. This is because we can avoid saturation and blooming from the central pixels while the signal is so strong that the noise is overwhelmed by the source signal. Then our goal is achieved even if we forget about the accurate focal position or accurate PSF.

2.2 Centroid

Even though we know the accurate PSF, we cannot use it for photometry if we cannot determine the center or the origin. There are various ways to determine this "origin", but the most widely used one is centroiding (center of mass).

Let I_i for $i = 1, 2, \cdots, N$ means the i-th pixel with $(x, y) = (x_i, y_i)$ of the original $N_x \times N_y$ 2-D pixel array so that $N = N_x N_y$. Say we set a threshold I_0 (e.g., background value, signal-to-noise ratio of 3, etc). Then change $I_i = 0$ if $I_i < I_0$. For $\xi = x$ or y, the centroid is by definition

$$\xi_c = \frac{\sum_i I_i \xi_i}{\sum_i I_i} \ . \tag{2.1}$$

By using the error-propagation, the uncertainty of ξ_c is, considering ξ_i's are fixed with no uncertainty,

$$(\Delta\xi_c)^2 = \left(\frac{\partial \xi_c}{\partial I_1}\right)^2 (\Delta I_1)^2 + \cdots + \left(\frac{\partial \xi_c}{\partial I_N}\right)^2 (\Delta I_N)^2 \ . \tag{2.2}$$

But

$$\frac{\partial \xi_c}{\partial I_k} = \frac{\partial}{\partial I_k}\left(\frac{I_1\xi_1 + \cdots + I_N\xi_N}{I_1 + \cdots + I_N}\right) = \frac{\xi_k}{\sum_i I_i} - \xi_c \ , \tag{2.3}$$

so

$$(\Delta\xi_c)^2 = \sum_k \left(\frac{\xi_k}{\sum_i I_i} - \xi_c\right)^2 (\Delta I_k)^2 = \frac{\sum_k (\xi_i - \xi_c)^2 (\Delta I_k)^2}{\sum_i I_i} \ . \tag{2.4}$$

Pixels near the object are likely to have high pixel values, so we approximate the uncertainty ΔI_i is just the Poisson noise of the pixel value (which is source + sky + read noise), i.e., eq. (1.22), $\Delta I_i = \sqrt{I_i}$. Then expanding the square and using the definition of ξ_c, we get

$$\Delta\xi_c = \sqrt{\frac{\sum_k (\xi_i - \xi_c)^2 I_k}{\sum_i I_i}} = \sqrt{\frac{s_c^2}{\sum_i I_i}} \tag{2.5}$$

where

$$s_c^2 = \frac{\sum_k \xi_k^2 I_k}{\sum_i I_i} - \xi_c^2 \ . \tag{2.6}$$

There are some other techniques which involves Gaussian fitting: marginalize the image onto the x- and y-axes, and do the 1-D Gaussian fitting to each of them. Also the sigma-clipping inside the centroiding box (cbox) can be used (e.g., Ma+ 2009, Opt.Exp., 17, 8525).

2.3 Aperture Sum

We need first to find the total flux inside the aperture. The sky subtraction will be done later. Consider a very simple case in fig. 2.3: The centroid is calculated following eq. (2.1), and it is trivial that the centroid is at the center of the shown image. If we put circular aperture of radius 1 pixel (middle panel), the pixel contribution from the pixel with value 10 (right panel) is

$$\text{pixel value} \times \frac{\text{area in aperture}}{\text{total area}} = 10 \times \frac{\pi/4}{1} = 7.854 \ . \tag{2.7}$$

Doing the same calculation for all the 4 pixels inside the aperture, you will get $N_{\text{apsum}} = 47.12$, where apsum means the "aperture sum".

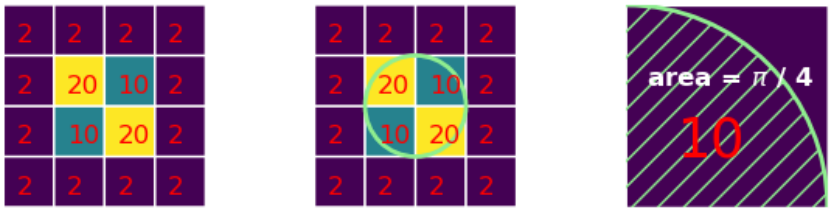

Figure 2.3: Left: An example pixel values near the source. Middle: An aperture of radius 1 pixel shown as green circle. Right: A zoom in of one pixel with pixel value 10.

One thing to note is the uncertainty in position. In this 4 by 4 pixel example, $\sum_i I_i = 84$ and $\xi_c = (x_c, y_c) = (1.5, 1.5)$ under the 0-indexing. For Δx_c, we consider only the x-coordinate for eq. (2.5) caluclation. Note that in that case, $\sum_k (x_k^2 I_k)$ for all the $4 \times 4 = 16$ pixels is nothing but $\sum_{k=1}^{4} (x_k^2 I_k')$

where $I' = [8, 34, 34, 8]$ is the marginalized intensity along the x direction from fig. 2.3. Then $\Delta x_c = 0^2 \times 8 + 1^2 \times 34 + 2^2 \times 34 + 3^2 \times 8 = 242$. This must be the same for the y-axis, and thus $s_c^2 = 242/84 - 1.5^2 = 0.63$, so $\Delta x_c = \Delta y_c = \sqrt{0.63/84} = 0.087$ pixel.

Here, I did not touch the topics such as (1) how to select the aperture radius, $r_{\rm ap}$ and (2) in reality, we need to fit some function to the radial profile, but which function should we use? These will be dealt in the next chapter.

Since this aperture sum includes not only the source but also the sky values (sky value is always non-zero for ground-based observations), we need to subtract the sky.

2.4 Sky Estimation

Estimation of the sky is one of the most trickest part in observational astronomy. There are two most widely used methods: annulus and 2D interpolation. I will first explain few formulae, and then explain these two methods.

2.4.1 Simple Parametric Sky Estimation

There have been some parametric methods to estimate correct sky value. The most widely used ones include the "MMM" relation, or the "mean-median-mode" relation, and regard the mode (most frequent value) as the constant sky value. This has been used by virtually all astronomical softwares, including IRAF, IDL, SExtractor, and python packages. It is the simplest and does not require much computational speed, while gives reasonable results in many cases.

Why mode? The mission in determining the *constant* sky value is to find a *robust*, i.e., trustful constant value. Average (mean) is very vulnerable to few outliers even though we did sigma-clipping, and the median is less robust than mode, because if faint object with S/N ratio ~ 1 is in the sky region, the median can also be overestimate the sky. Thus, mode is the most widely used simple statistic for sky estimation.

The determination of the modal value is very difficult, because the it is not simple to define it for the *sampled data*. If the distribution is not a mathematically continuous one, i.e., if it is discrete as we always encounter after sampling process, the mode is the largest value when we draw a *histogram*. The histograms are basically sensitive to the subjective selection of bins, and you can change the modal value to an unexpected value depending on the bin you choose (see fig. 2.4).

Since 1800s, people empirically knew that the mode has the following relationship which holds for a moderetely asymmetric distribution:

$$\text{mode} \approx 3 \times \text{median} - 2 \times \text{mean} \ , \tag{2.8}$$

or equivalently, *the distance between the mode and median is two-thirds of that between mode and mean.* DoodsonAT (1917, Biometrika Trust, 11, 425) first gave mathematical proof of it using some mathematical tricks including Laplace's and De Morgan's (1836), and it is recognized by many statisticians including KendallMG (1943; or the second edition of it in 1945, The Advanced Theory of Statistics 2e, Vol 1, p.35). People now refer to more recent references like KendallMG+StuartK (1977, The Advanced Theory of Statistics, Vol. 1. Charles Griffin & Co., London).

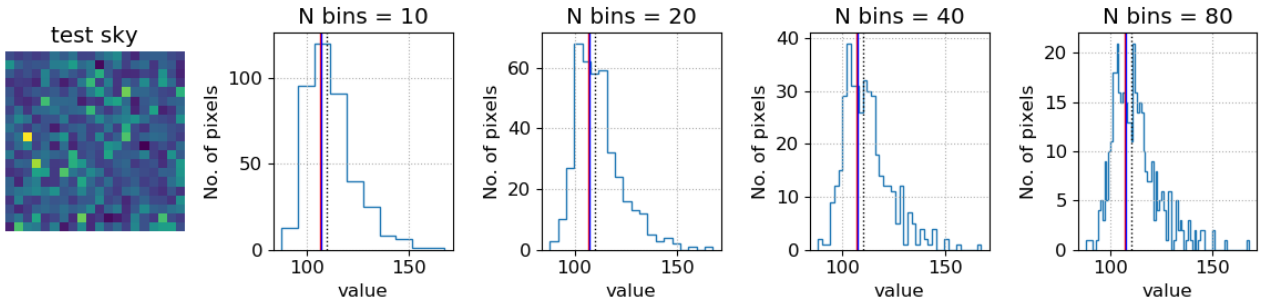

Figure 2.4: Left: A fake sky of 20 by 20 pixels generated by $100 + \mathcal{N}(0, 5^2) + 5\chi^2_2 + 3\mathcal{U}(0, 1)$ (normal distribution, chi-squared distribution, and uniform distribution, respectively). Others: The histogram of the pixel values with different bin numbers with fixed bin width. The black dotted lines is the median, and the red and blue solid lines are the mode estimation from eqs. (2.8) and (2.9), respectively (almost indistinguishable). As can be seen, the mode value does not lie at the same position, and it even become bimodal when the number of bins is 80.

SExtractor (BertinE+ArnoutsS, 1996, A&AS, 117, 393) uses a variant of this:

$$\text{mode} \approx 2.5 \times \text{median} - 1.5 \times \text{mean} \ . \tag{2.9}$$

There is no clear reason provided in the original paper nor the SEx user manual, but it is just descibed that it was found to be more accurate when they tested. The estimation process works like this: Get the mean, median, and sample standard deviation (mean, med, std, respectively) from the sky pixel values after sigma-clipping. If $\frac{\text{mean}-\text{med}}{\text{std}} > 0.3$, use mode $\approx$ med*. Otherwise, use eq. (2.9).

IDL MMM uses eq. (2.8). IRAF also uses eq. (2.8) internally, but it gives the mean value if mean $<$ med. In astropy-related packages (e.g., `photutils`), it is the user's choice. In all cases, you can set the sigma-clipping factors, e.g., a such that data outside med $\pm\, a \times$ std can be rejected, and n, the maximum iterations this process should be done until no more data are rejected. Whether to use med $\pm\, a \times$ std or mean $\pm\, a \times$ std, whether to give asymmetric a values (upper and lower sigma-clipping), etc are up to the user.

I recommend to follow what SEx does, with sigma-clipping of $a \sim 3$ and $n \geq 5$, because that became a virtual standard in parametric sky estimation.

2.4.2 Other Ways for Sky Estimation

As can be seen from fig. 2.4, the sky histogram is mostly skewed towards right. Because of this, BijaouiA (1980, A&A, 84, 81) first introduced to fit the sky histogram with Gaussian multiplied by Laplace:

$$p(I) = \frac{e^{\sigma^2/2a^2}}{a} e^{-(I-s)/a} \text{erf}_c\left(\frac{\sigma}{a} - \frac{I-s}{\sigma}\right) \tag{2.10}$$

where

$$\text{erf}_c(x) := \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-t^2/2} dt \ . \tag{2.11}$$

*Although the publication and SEx manual says they use mean if $\frac{\text{mean}-\text{med}}{\text{std}} > 0.3$, but the SEx software actually returns med, not mean (described in photutils v0.6, Background Estimation (photutils.background))

IrwinMJ (1985, MNRAS, 214, 575) argued that this method for modal estimation (Bayesian maximum likelihood is used) is better compared to mean, median, and Gaussian fitting. But because of the complexity of the calculation, Irwin used simpler method, i.e., get the smoothed histogram near the estimated mode and fit a Gaussian. BeardSM+ (1990, MNRAS, 247, 311) used slightly different method: smooth the (sky pixel) frequency histogram with moving box filter, find the mode estimate with it, sample the pixels nearby it which has pixel values larger than half of it. Then it does a cubic polynomial fit to the unsmoothed original histogram and finds the peak from this polynomial as the mode estimate.

For more, you may refer to Appendix A of AkhlaghiM and IchikawaT (2015, ApJS, 220, 1) and section 3.1 of MasiasM+ (2012, MNRAS, 422, 1674).

2.4.3 Annulus Sky

The simplest way is to use circular annulus. This method assumes the sky is nearly constant near the terget of interest, PSF is circular, and there is no nearby celestial object. Under these simplifying assumptions, the sky must be azimuthally symmetric and a function of radius centered on the centroid of the target. If we set the inner and outer radii $r_{\rm in}$ and $r_{\rm out}$ much larger than the seeing disc size to define the circular annulus, the pixel values within that annulus are now the "sample" of sky values, and should not have clear tendency along the radial and azimuthal direction. fig. 2.5 shows the result of centroiding and sky estimation.

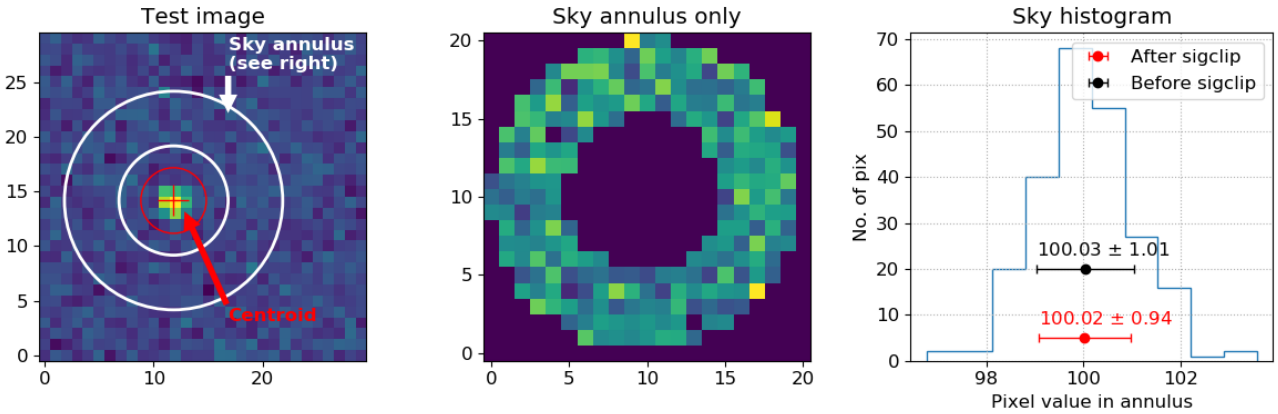

Figure 2.5: Left: The test image. Centroid is determined by eq. (2.1) and annulus is defined as $r_{\rm in} = 5$ and $r_{\rm out} = 10$ pixels, respectively. Middle: Plotting the sky annulus region only, to just visualize the random fluctuations in the sky region. Clearly, there is no radial trend and it seems quite symmetric along the azimuth. Right: The histogram of the sky pixel values in the annulus. The red/black markers are the median and sample standard deviation of the sky values after and before sigma-clipping, respectively. The sigma clip was done with 3-sigma 50-iterations.

Consider we obtained the sky value $m_{\rm sky}$ somehow, e.g., median of sky, eq. (2.9) after sigma-clipping, etc. There can be many ways to estimate the uncertainties of this $m_{\rm sky}$. Let me call this $\Delta m_{\rm sky}$. The fluctuation of sky value can critically affect the final result of the photometry. But once a robust way to estimate the sky is set and if the target is bright enough, the uncertainty due to the sky value is not so large. Thus, people do not care too much about the accurate pdf of the $m_{\rm sky}$ value. The most widely used way is to use the practical CLT (Thm 3) to $m_{\rm sky}$, treating as if it is the mean of the sky samples

(although we almost never use the mean for $m_{\rm sky}$). Then

$$\Delta m_{\rm sky} = \frac{s_{\rm sky}}{\sqrt{n_{\rm sky}}} \ . \tag{2.12}$$

Here, $s_{\rm sky}$ and $n_{\rm sky}$ are the sample standard deviation and the number of the sky pixels survived after the sigma-clipping, respectively. So we are assuming here that the mode has similar uncertainty as the mean.

2.4.4 Source Extractor Sky

We will not cover Source Extractor (SE or SEx or SExtractor) deeply. But simply put, it works like the description below. Say we have the image of 1000 by 1000 pixel.

1. **Meshing**: Chop the image by given box size. For example, if you set the box size as $(50, 25)$, 20 by 40 (in total 800) "pads" or "meshes" will be made. Each pad will of course have 50 by 25 (in total 1250) pixels.
2. **Filtering**: Determine the size of the median filter. If the size is $(3, 3)$, for instance, a 3 by 3 median filter will be used to reduce sharp noised pixels in each pad (similar to Gaussian convolution). The filter size should be determined as a function of seeing disc size.
3. **Rejecting Meshes**: In this median-filtered pad, the sigma-clipping is done. If too many pixels inside a pad is rejected, that pad is rejected for the next background estimation step. Such a threshold, say exclusion percentile, must be given by the user. For example, if that percentile is 10 %, any of the 800 mesh which rejected more than 10 % of its pixels from sigma-clipping (10 % out of 1250 pixels, i.e., 125 pixels) will be regarded as "bad region" for the sky estimation.
4. **Sky at each Mesh** (bottom middle panel of fig. 2.6): Now come back to the original un-filtered meshes. From the pixel values, estimate the sky value and its uncertainty following eq. (2.9) and the description below the equation. For the meshes which are flagged as "bad" from previous step are not used for this process. As a result, you will have 20 by 40 array of sky values, while some elements may be empty or NaN, if some meshes are flagged ad "bad".
5. **Interpolation** (top middle panel of fig. 2.6): From the estimated single sky value at each mesh, i.e., the 20 by 40 array, we do interpolation to estimate the sky at all the 1000 by 1000 pixels. The estimation uncertainties should also be calculated for each pixel.

After the sky subtraction, I plotted the Box plot of the image with x-axis 70 to 90 and y-axis 50 to 150 (python indexing) in fig. 2.7.

As can be noticed, SEx background estimation*, unlike the annulus sky estimation technique, is estimating the sky values at all the pixels, not for each source. Thus, it takes much longer time. Also you can see that lots of parameters (rejection percentile, box size, filter size, interpolation method and related parameters, etc) should be set for SEx to work! In some cases, astronomers run SExtractor again and again to do get expected results, and find best combination of such parameters. But these parameters of course depend on the instruments, sky conditions, etc, so if the results are suspicious, some fine tuning must be made or you have to find different ways to do photometry than SExtractor.

*We tend to call "background" rather than "sky" when we are dealing with SExtractor.

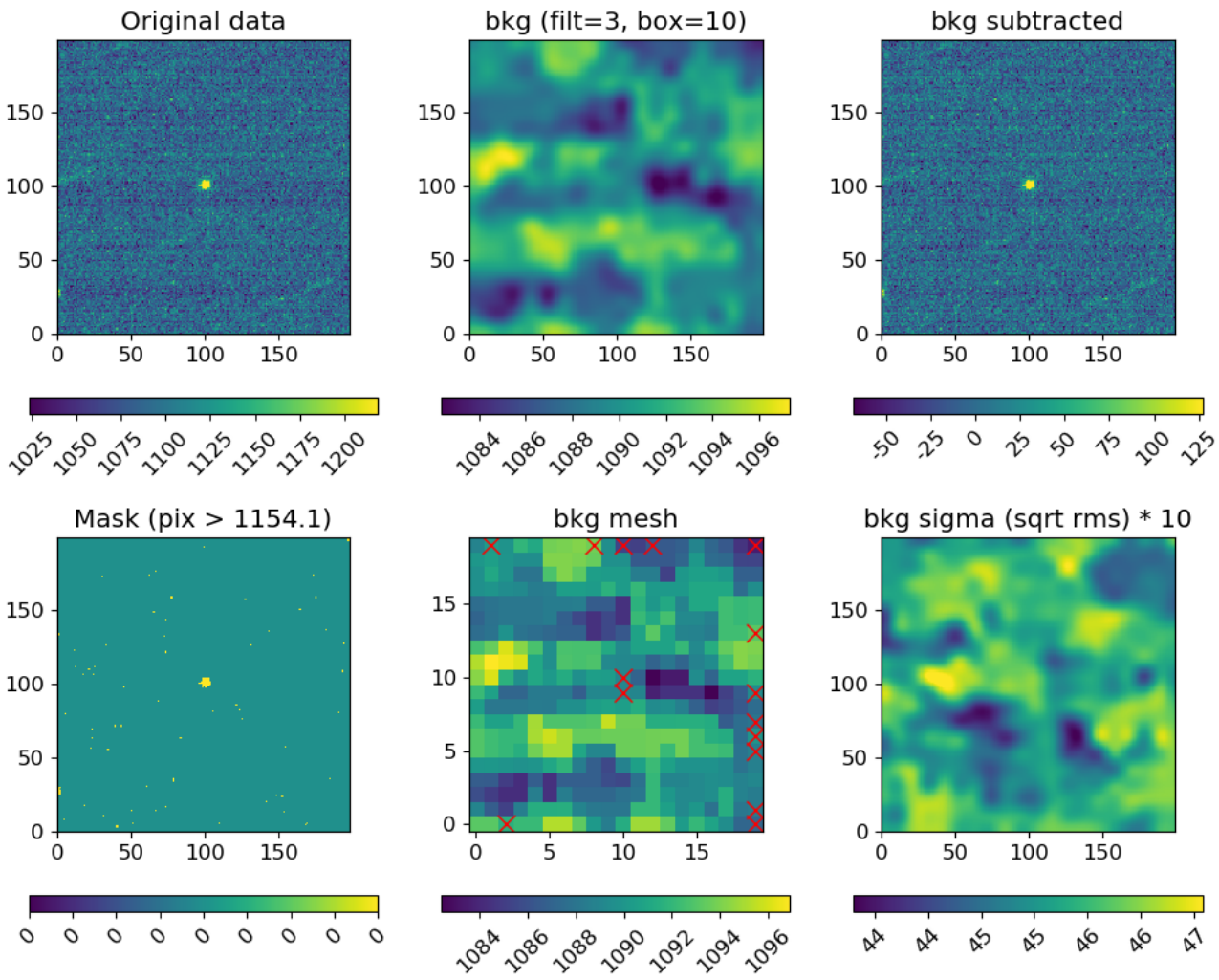

Figure 2.6: Example of the SExtractor background estimation usage. The original data is the targeto-centric combined image from 13 KMTNet images from SAAO observatory. The pixels with value 1154.1, which is the med + 3std after the sigma-clipping (3-sigma, 5-iter clipping) are masked because those can be stars (bottom left). The background meshes with the estimated sky values (eq. (2.9)) are shown in bottom center, and the interpolated one is given in top middle. Note that the sky level changes at maximum around 12 ADU out of the pixel value 1000+. Normally this should be smaller like 5 ADU or 2 ADU within this small region (200 by 200 pixel). The background sigma shown in bottom right is around 4 ADU (note the figure is 10 times the sigma), but many times the sky sigma is smaller than this.

2.5 Aperture Shape in Real Reduction

We most frequently use circular aperture photometry. To determine how large the aperture should be, we need information about the radial profile of the PSF. Historically, there are three important radial profiles: Gaussian, Moffat, and Penny:

1. Gaussian. The Gaussian function.
2. Moffat. A purely empirical function.
3. Penny. A purely empirical function.

(links to Wikipedia)

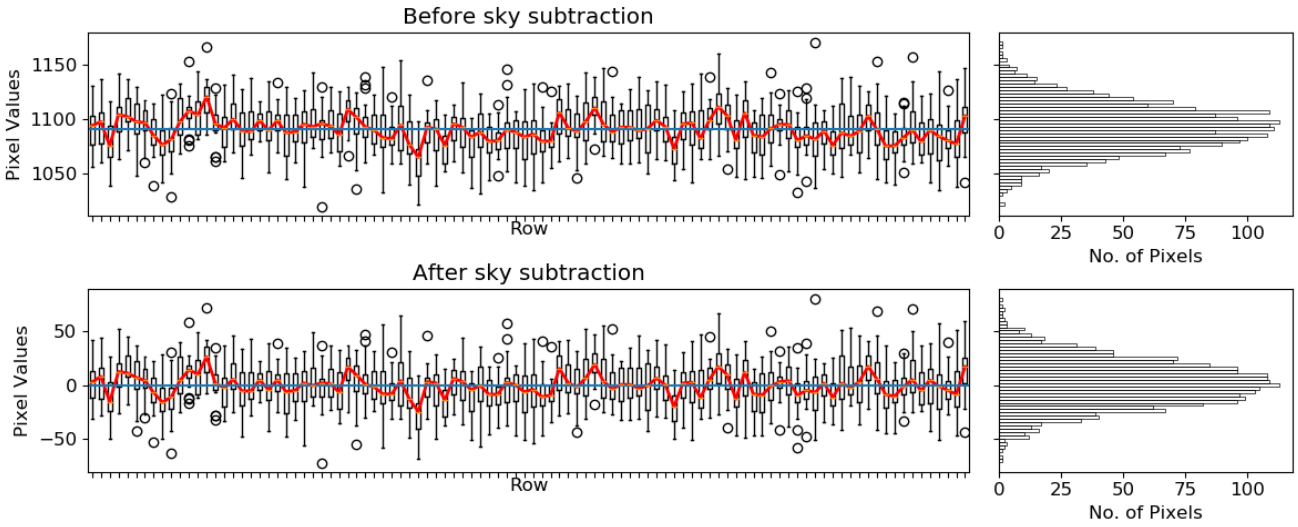

Figure 2.7: The Box plot and pixel value histogram of before/after the sky subtraction, of a region defined by `data[50:150, 70:90]` (python indexing). The horizontal axes of Box plots correspond to the "row" or the y index. Blue horizontal lines are the median of all the pixels in that region, and red solid lines are the medians

Among parameters for PSFs, one that is very widely used in astronomy is the FWHM (full width at half-maximum), as described before. FWHM is defined as the distance between the two points which have the value $\frac{1}{2}f_{\max}$, i.e., $f(r = \text{HWHM}) = \frac{1}{2}f_{\max}$ for HWHM = FWHM/2. FWHM is also used for Lorenzian profile, which has infinite standard deviation (the second centered moment diverges, so the standard deviation is not defined, but FWHM can). The profiles of PSFs and their FWHMs are described below.

In the descriptions below, I will assume the center $\mathbf{r}_0$ is clearly fixed and the PSF is azimuthally symmetric (circular) and thus $r = |\mathbf{r} - \mathbf{r}_0|$ is the only free parameter except for other parameters described in the model.

2.5.1 Gaussian

The circular, i.e., azimuthally symmetric Gaussian PSF is like this:

$$f_{\text{Gauss}}(r|A, \sigma) = Ae^{-r^2/2\sigma^2} \tag{2.13}$$

Here, r is the distance from the center (r_0), and σ is the standard deviation of the profile. This is the most famous form when it comes to numerical calculation, because all the normalization constants are not so important in PSF fitting.

The normalization constant $A = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N$ for an N-D Gaussian, such that the integration is normalized, i.e., $\int_{-\infty}^{+\infty} f_{\text{Gauss}}(r)2\pi r dr = 1$. The 2-D Gaussian of integrated flux I is, therefore,

$$f_{\text{Gauss}}(r|I, \sigma) = \frac{I}{2\pi\sigma^2}e^{-r^2/2\sigma^2} \tag{2.14}$$

This is useful because the total integrated flux is an input parameter. Conversion between this and eq. (2.13) is very easy.

Table 2.1: Moffat profile comparison. `astropy`'s notation is strikingly confusing, because of its `alpha` notation, which completely goes against the original Moffat paper and other software implementations (e.g., IRAF). Note: `AsPyLib` is developed by J. Caron in Python 2 and the development is halted since 2013.

Symbol	astropy	IRAF, AsPyLib	Description
r			radial distance from the center ($r_0 = (x_0, y_0)$)
R	`gamma`	`alpha`	core width
β	`alpha`	`beta`	power

The FWHM is calculated by setting $f_{\rm Gauss}(r = \mathrm{HWHM}) = \frac{A}{2}$. First we obtain $(\mathrm{HWHM})^2 = (2\ln 2)\sigma^2$ and thus,

$$\mathrm{FWHM(circGauss)} := 2\sqrt{2\ln 2}\sigma \ . \tag{2.15}$$

Re-formulating eq. (2.14) with FWHM (F for brevity):

$$f_{\rm Gauss}(r|I, F) = I\frac{4\ln 2}{\pi F^2}e^{-(4\ln 2)r^2/F^2} \ . \tag{2.16}$$

This is not widely used, although both I and F are what we're most interested in. Rather, people tend to convert F to σ and use eq. (2.13) or eq. (2.14).

2.5.2 Moffat

The first suggestion of using Moffat profile was made in 1969 by A. F. J. Moffat in Moffat AFJ 1969, A&A, 3, 455. It is purely an empirical function, and is one of the most widely accepted form, because of its simple form yet powerful in fitting many circular profiles. Comparison of notations is given in section 2.5.2.

The circular, i.e., azimuthally symmetric Moffat PSF is:

$$f_{\rm Moffat}(r|A, R, \beta) = A\left[1 + \left(\frac{r}{R}\right)^2\right]^{-\beta} \ . \tag{2.17}$$

The parameter R is called the core width and β is called the power. As in Gaussian case, this is the most widely used form in numerical calculation.

Normalizing as we did for Gaussian, $\int_{-\infty}^{+\infty} f_{\rm Moffat}(r)2\pi r dr = 1$, we have the constant $A = \frac{\beta-1}{\pi R^2}$. Therefore, Moffat profile has a restriction that $\beta > 1$. The Moffat profile for the integrated flux I is, similar to Gaussian,

$$f_{\rm Moffat}(r|I, R, \beta) = I\frac{\beta-1}{\pi R^2}\left[1 + \left(\frac{r}{R}\right)^2\right]^{-\beta} \tag{2.18}$$

Similar to Gaussian, this is also widely used.

The FWHM is calculated by setting $f_{\rm Moffat}(\mathrm{HWHM}) = \frac{A}{2}$. As we did in Gauss example, we obtain $\mathrm{HWHM}^2 = R^2(2^{1/\beta} - 1)$,

$$\mathrm{FWHM(circMoffat)} = 2R\sqrt{2^{1/\beta} - 1} \ . \tag{2.19}$$

Using FWHM (F for brevity), we re-formulate Moffat profile eq. (2.18)*:

$$f_{\rm Moffat}(r|I, F, \beta) = I\frac{4(2^{1/\beta} - 1)(\beta - 1)}{\pi F^2}\left[1 + 4(2^{1/\beta} - 1)\left(\frac{r}{F}\right)^2\right]^{-\beta} \ . \tag{2.20}$$

*Trujillo+2001, MNRAS, 328, 977

Thm 10 (*High power Moffat is Gaussian*)

For $\beta \to \infty$,

$$f_{\rm Moffat}(r|I,F,\infty) = f_{\rm Gauss}(r|I,F) \ . \tag{2.21}$$

Practically $\beta \gtrsim 100$ is enough for this approximation.

Proof of Thm : For $\beta \to \infty$, we approximate $2^{1/\beta} - 1 \approx \frac{\ln 2}{\beta}$, so

$$f_{\rm Moffat}(r|I,F,\infty) \approx I\frac{\tilde{A}}{\pi} \lim_{\beta\to\infty} \frac{\beta-1}{\beta}\left[1+\frac{\tilde{A}r^2}{\beta}\right]^{-\beta} \ ,$$

for $\tilde{A} = \frac{4\ln 2}{F^2}$. Note that $\lim_{\beta\to\infty}\left[1+\frac{\tilde{A}}{\beta}\right]^{-\beta} \equiv e^{-\tilde{A}}$ and $\frac{\beta-1}{\beta}$ becomes unity. Substituting these, we get the approximated formula identical to eq. (2.16), or to eq. (2.13) with eq. (2.15). Q.E.D.

Also, if you look at the pdf of the Students' t-distribution, Moffat is a special case of the t-distribution, but with a simpler re-parameterization (original t-distribution has a clear mathematical reason why it appeared, but it has normalization constants including the Gamma functions, etc, so computationally a bit more complicated).

2.5.3 Penny

It was Franz*, who first used Lorenzian-like profile for stars. Then Penny† used it with linear interpolation in sky level to do the UBV photometry for visual binaries close to each other. After that, Penny and Dickens‡, using their own package (`STARMAN`), introduced Gaussian and Lorentzian mixed model§.

The original profile suggested by Penny is a sum of a circular Gaussian and elliptical modified-Lorentzian:

$$\begin{aligned} &f(r|A,Q,\sigma_1,\sigma_2,\sigma_3,\sigma_4,P,\theta,P_Q,R_Q) \\ &= A\left\{ \frac{1}{1+\left[\left(\frac{x'}{\sigma_1}\right)^2+\left(\frac{y'}{\sigma_2}\right)^2\right]^{\frac{P}{2}\left(1+\sqrt{\left(\frac{x'}{\sigma_3}\right)^2+\left(\frac{y'}{\sigma_4}\right)^2}\right)}} + Q\exp\left[-\left(\frac{r}{R_Q}\right)^{P_Q}\right]\right\} , \end{aligned} \tag{2.22}$$

for

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} \cos\theta & \sin\theta \\ -\cos\theta & \sin\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

where $\mathbf{r}_0 = (x_0, y_0)$ is the center, and θ is the tilting angle of the Lorentzian. The notation $\sigma_{1,2,3,4}$ are $R_{\rm maj}$ (`RX`), $R_{\rm min}$ (`RY`), $RP_{\rm maj}$ (`PRX`), and $RP_{\rm min}$ (`PRY`) in the `POORMAN`'s notations. It is no more widely used, because of the large number of free parameters.

*FranzOG 1973, JRASC, 67, 81

†PennyAJ 1979, MNRAS, 187, 829

‡PennyAJ & DickensRJ, 1986, MNRAS, 220, 845, which is cited for more than 110 times as of Oct 2019

§Unfortunately, however, both the original Penny & Dickens paper and Penny (1995, StaUN 141) pp.44–45 have so many typos in the star profile formulae...

In IRAF, it is implemented* in two ways: `penny1` and `penny2`.

†

Simple way of using it is, assuming $Q = 0$, $P = 2$, and $\sigma_3, \sigma_4 \to \infty$, so that

$$f(r|A, \theta, \sigma_1, \sigma_2, \tag{2.23}$$

2.5.4 IRAF

Here I describe how IRAF does centroiding in its IRAF `imexamine` task. It is crude but always works decently, so I thought it's a good idea to summarize it for benchmark purpose‡.

First, set r, the circular aperture radius (default `radius=5`). Then the sky level is determined from an annulus with the inner radius `radius+buffer` (default `buffer=5`) and the outer radius `radius+buffer+width` (default `width=5`), centered at the initial center position. By default (`xorder=0`, `yorder=0`), the sky value is set as the median value in this annulus. If both of `xorder` and `yorder` are larger than 0, then the 2-D polynomial of these orders is fitted to the pixel values from the annulus. Mathematically speaking, `xorder=yorder=1` means a constant value (the weighted average value).

Then the centering is done by fitting 1-D Gaussian or Moffat to the marginalized pixel values (default is Moffat by `fittype="moffat"`), for the square box with width = height = $2r$. In pythonic way, you are fitting 1-D profile to `np.sum(data, axis=0)` and `np.sum(data, axis=1)` to get the y and x centers, respectively. The pixels below the average value in this box are masked in the centering process§.

It is repeated for certain times (default `iterations=3`), until the updated center position resides in the same pixel as the previous step¶. In every iteration, the aperture radius r is updated to 3 FWHM using the FWHM obtained from the last iteration (from this definition, r remains as the initial value if `iterations=1`). Any pixel outside the image will be assumed to have a constant value (default `constant=0`).

The weighting for each pixel is calculated by

$$w(r_{\mathtt{i,j}}) = \begin{cases} r_{\mathtt{i,j}}^{-2} & r_{\mathtt{i,j}} \le \mathrm{FWHM}/2 \\ e^{-\left(\frac{r_{\mathtt{i,j}}}{\mathrm{FWHM}/2}-1\right)^2} & r_{\mathtt{i,j}} > \mathrm{FWHM}/2 \end{cases}, \tag{2.24}$$

*The source code is at `noao.digiphot.daophot.dailib.profile.x`.

†Although there is no reference, "Photometry using IRAF v. 2.5.1" (2012-03-25) by Keunhong Park, Jinhyuk Ryu, Insung Jang, and Ho Seong Hwang notes that (p. 8): "`penny1` or `penny2` of IRAF fits the analytical stellar profiles well, while `moffat15` or `moffat25` fits the stellar profiles well if observed on cloudy nights." I cannot even understand the meaning of it...

‡The description below is the description of IRAF `imexamine` when you hit `a` key on IRAF. Since `imexamine` is not only for the 2-D analysis, it consists of parameters for the 1-D extraction (when you hit `j` or `k`), for instance. You may want to fit a Gaussian to the pixels along a vertical line (i.e., $N \times 1$ array). For those purposes, `imexamine` has more parameters like `sigma`, etc. Please don't be confused when you read the doc.

§Not clear from the doc: How are the weighting scheme applied in this 1-D profile fittings? As in the 2-D profile fit? Or no error-bars?

¶Not clear from the doc: I worried what if the center position is at the edge of the pixel. I guess, maybe, even though the center position is updated from the right edge of the i-th to the left edge of the $(i+1)$-th pixel, the centering box will remain the same in the upcoming iterations, so the centering process will be halted (converged) at the left edge of the $(i+1)$-th pixel.

where FWHM/2 is also called HWHM or $r_{\rm half}$ and $r_{\rm i,j}$ is the distance in pixel unit to the `(i, j)`-th pixel from the center*. This weighting scheme is different from usual least square statistic (which uses $w = 1/\sigma^2$).

Once this fitting is done, put a circular aperture with radius r. Because the purpose of `imexamine` is not an accurate photometry, it crudely estimates the flux by summing the pixel values for pixels which has its center within the aperture (same as putting `method="center"` to photutils aperture).

Two more parameters are calculated, namely, ellipticity e and angle pa:

$$e = \sqrt{\frac{(M_{xx} - M_{yy})^2 + 4M_{xy}^2}{M_{xx} + M_{yy}}} \quad ; \quad \mathrm{pa} = \frac{1}{2}\,\mathrm{atan}\left(\frac{2M_{xy}}{M_{xx} - M_{yy}}\right), \tag{2.25}$$

where the moments

$$M_{\rm ij} = \frac{\sum I \times i \times j}{\sum I} \tag{2.26}$$

for pixel value I and $i, j \in [x, y]$.

In realistic photometry, such as IRAF `psfmeasure` task, we must use $I - \mathrm{sky}$ instead of I.

Some old packages including `ROMAFOT`, `STARMAN`, `DAOPHOT`, `DoPHOT`, `WOLF`, `LUND`, `CAPELLA` are summarized in this lecture note and StetsonPB 1992, ASPC, 25, 297

Note

The shape of circular apetures were briefly discussed. In reality, you may encounter apertures of ellipse, box, pill-box, or any other scientifically meaningful ones. Now what you have to understand more is the aperture *size*.

2.6 Aperture Size in Real Reduction

We will stick to the circular apertures here, too. Any other compilcated apertures may have to be used depending on the scientific reasons. Here I will describe three *ideas* of finding the aperture size. When you have to apply aperture size in professional research, these should not be used blindly, but you have to find your own way which is both mathematically robust and scientifically meaningful for your specific purpose.

2.6.1 Maximum SNR (not recommeneded)

The singal-to-noise ratio is a

*Not clear from the doc: Center of the last iteration?

Chapter 3

Statistics in Photometry

So far we've seen how the photometry should be done. But we haven't seriously discussed how to analize the uncertainties, i.e., the error-bars of each measurement such as magnitude.

3.1 Pixel-wise Uncertainty

An object frame consist of at least bias, dark, flat, cosmic-ray, target's signal, and sky, as well as readout noise. For the j-th pixel, they are denoted as:

- $\tilde{N}_j$: the raw pixel value of the object frame in ADU.
- B_j: the bias in ADU.
- D_j: the dark in ADU.
- F_j: the flat in ADU (*not necessarily* normalized; see **Note** at the end).
- C_j: the cosmic-ray in ADU.
- I_j: the object signal count in ADU.
- S_j: the sky signal in ADU.

In some cases, such as infrared detectors, the gain and readout noise may differ for each pixel, so I will denote gain and readout noise of the j-th pixel as g_j [e/ADU] and R_j [e]. Then we define the bias-subtracted pixel value

$$N_j = \tilde{N}_j - B_j = D_j + (I_j + S_j + C_j)F_j \tag{3.1}$$

Note that the dark is not affected by the flat value. Since this pixel value should follow a Poisson distribution, which is approximated as a Gaussian distribution,

$$N_j \mathrel{\dot{\sim}} \mathcal{N}\left(N_j,\ \frac{N_j}{g_j} + \left(\frac{R_j}{g_j}\right)^2\right) \ . \tag{3.2}$$

3.1.1 Dark Estimation

To estimate dark D_j, you should have taken nearly tens of dark frames, and combined it. From the median value of the frames, you may have obtained the estimation of the dark $\hat{D}_j$ (for brevity, I will just use D_j without the caret $\hat{\cdot}$). During the combination, you can obtain the uncertainty of the dark at the j-th pixel, ΔD_j, as

$$\Delta D_j = \mathrm{sstd}(D_j^i) \ , \tag{3.3}$$

where D_j^i is the dark current (ADU) of the j-th pixel at the i-th dark frame, sstd is the sample standard deviation function. Here it is assumed D_j roughly follows a Gaussian distribution so that sstd becomes an unbiased estimator* of the true standard deviation. Note that you should *not* divide it by the number of dark frames (such as Thm 1), because **what you will need is not the uncertainty of the *mean* value, but the uncertainty of one pixel value that will have been added to the object frame**.

Therefore, to the first order,

$$\begin{aligned} \tilde{O}_j = N_j - D_j &\dot{\sim} \mathcal{N}\left(\tilde{O}_j,\ \frac{N_j}{g_j} + \left(\frac{R_j}{g_j}\right)^2 + (\Delta D_j)^2\right) \\ \text{or } &\dot{\sim} \mathcal{N}\left(\tilde{O}_j,\ \frac{\tilde{O}_j}{g_j} + \frac{D_j}{g_j} + \left(\frac{R_j}{g_j}\right)^2 + (\Delta D_j)^2\right) \ . \end{aligned} \tag{3.4}$$

3.1.2 Flat Estimation

Unlike the dark current, which is a probabilistic electron generation, the flat, i.e., the pixel sensitivity, is a characterestic value of the detector and optics, so it should not change at each exposure. Therefore, following the CLT (Thm 1):

$$\Delta F_j = \frac{\text{sstd}(F_j^i)}{\sqrt{N_{\text{Flat}}}} \ , \tag{3.5}$$

where F_j^i is the flat frame pixel value (ADU) of the j-th pixel at the i-th flat frame, sstd is the sample standard deviation function and N_{Flat} is the number of the used flat frames.

Generally speaking, F_j must be about 10^4 ADU or more, so that the signal-to-noise ratio from Poisson statistics is larger than about 100. Moreover, you should have obtained, say, $N_{\text{flat}} = 9$, this will increase by a factor of 3. Therefore, the uncertainty in the flat ΔF_j, is around 0.1 % order. In some observatories, people take hundreds of flats at one time to get flat as good as roughly 0.01 % order, assuming it won't change over certain period of time.

Although there are mathematically known ratio distribution, i.e., the pdf of $Z := X/Y$ where $X, Y \sim \mathcal{N}(\mu_{X,Y}, \sigma^2_{X,Y})$ and the covariance is zero, what we are interested in is only a rough estimation of the uncertainties†. From the propagation of error,

$$O_j = \frac{N_j - D_j}{F_j} = I_j + S_j + C_j \dot{\sim} \mathcal{N}\left(O_j,\ O_j^2\left[\left(\frac{\Delta\tilde{O}_j}{\tilde{O}_j}\right)^2 + \left(\frac{\Delta F_j}{F_j}\right)^2\right]\right) \ . \tag{3.6}$$

Substituting previously obtained distribution of $\tilde{O}_j$,

$$O_j \dot{\sim} \mathcal{N}\left(O_j,\ \frac{O_j}{g_j F_j} + \frac{D_j}{g_j F_j^2} + \left(\frac{R_j}{g_j F_j}\right)^2 + \left(\frac{\Delta D_j}{F_j}\right)^2 + O_j^2\left(\frac{\Delta F_j}{F_j}\right)^2\right) \ . \tag{3.7}$$

*An unbiased estimator of a random variable X, $\hat{X}$ is defined such that the expectation value of $\hat{X}$ is the same as the true value X_{true}.

†If you are interested in, see FiellerEC (1932) Biometrika, 24, 428; HinkleyDV (1969) Biometrika, 56, 635; Díaz-Francés+RubioFJ (2013) Statistical Papers, 54, 309, as well as Wikipedia. For the ratio of N_j/F_j, i.e., when the denominator has so high signal-to-noise ratio, the distribution is roughly a Gaussian with $\mathcal{N}(\mu_X, \sigma_X^2)$ as we assume in this section.

3.1.3 Final Pixel-wise Uncertainty

Many times the C_j is removed by the so-called *cosmic-ray rejection* algorithms.

Thm 11 (*Pixel-wise Error*)

The final, cosmic-ray removed pixel value will follow

$$\begin{aligned} O_j^{\mathrm{cr}} = \frac{N_j - D_j}{F_j} - C_j = I_j + S_j &\dot\sim \mathcal{N}\left(O_j^{\mathrm{cr}},\ \frac{O_j}{g_j F_j} + \frac{D_j}{g_j F_j^2} + \left(\frac{R_j}{g_j F_j}\right)^2 + \left(\frac{\Delta D_j}{F_j}\right)^2 + O_j^2\left(\frac{\Delta F_j}{F_j}\right)^2\right) \\ &\dot\sim \mathcal{N}\left(O_j^{\mathrm{cr}},\ \frac{N_j}{g_j F_j^2} + \left(\frac{R_j}{g_j F_j}\right)^2 + \left(\frac{\Delta D_j}{F_j}\right)^2 + O_j^2\left(\frac{\Delta F_j}{F_j}\right)^2\right) \end{aligned} \tag{3.8}$$

Here, O_j is the bias, dark, and flat corrected pixel value before the cosmic-ray removal, while N_j is only bias-subtracted pixel value. Usually the uncertainty from the estimation for C_j is ignored as it is too difficult to estimate.

In practical application, the following approximations may be used:

1. Calculating the ΔD_j term all the time is annoying, and moreover, it will be much smaller than the O_j term. Therefore, people just ignore it and set it to 0. Theoretically, however, $\Delta D_j > (R_j/g_j)$, because $\Delta D_j = \mathrm{sstd}(D_j^i)$ and $D_j^i \dot\sim \mathcal{N}\left(D_j^i,\ D_j^i/g_j + (R_j/g_j)^2\right)$. Therefore, more reasonable approximation, or the lower limit of the uncertainty, would be R_j/g_j.
2. Frequently, $\Delta F_j/F_j$ is much smaller than $\Delta\tilde{O}_j/\tilde{O}_j$ (or we can increase N_{Flat} to force that this holds), so the flat-error term is negligible. StetsonPB, for instance, asked the user to give a constant $\Delta F_j/F_j \equiv \sigma_F$ for all pixels, such as 0, 0.01, or 0.0075, in `DAOPHOT`.
3. Many times we use the normalized flat (for more discussion, see below) such that its mean or median is 1, so $F_j \sim 1$ for all pixel. Then all F_j in eq. (3.8) can be just ignored.
4. The D_j term is negligible for many times. It is not negligible only for hot pixels where $D_j \gg 1$, but likely the observer did not put the target of interest at where hot pixels present. Moreover, state-of-the-art CCDs, such as Subaru FOCAS for instance, has too small dark current $D_j < 0.1\,\mathrm{e/s}$. Therefore, most important parts in the object frames will have negligible D_j compared to O_j^{cr}. Therefore, this term is also ignored many times.

Summarizing, eq. (3.8) is approximated as

$$\begin{aligned} O_j^{\mathrm{cr}} &\dot\sim \mathcal{N}\left(O_j^{\mathrm{cr}},\ \frac{O_j}{g_j F_j} + \frac{D_j}{g_j F_j^2} + \left(\frac{R_j}{g_j F_j}\right)^2 + \left(\frac{\Delta D_j}{F_j}\right)^2 + O_j^2\left(\frac{\Delta F_j}{F_j}\right)^2\right) \\ &\dot\sim \mathcal{N}\left(O_j^{\mathrm{cr}},\ \frac{O_j}{g_j F_j} + \frac{D_j}{g_j F_j^2} + 2\left(\frac{R_j}{g_j F_j}\right)^2 + \sigma_F O_j^2\right) \\ &\dot\sim \mathcal{N}\left(O_j^{\mathrm{cr}},\ \frac{O_j}{g_j} + \frac{D_j}{g_j} + 2\left(\frac{R_j}{g_j}\right)^2 + \sigma_F O_j^2\right) \\ &\dot\sim \mathcal{N}\left(O_j^{\mathrm{cr}},\ \frac{O_j}{g_j} + 2\left(\frac{R_j}{g_j}\right)^2 + \sigma_F O_j^2\right) \end{aligned} \tag{3.9}$$

Most frequently people even drop the factor 2 and σ_F, and write $O_j^{\mathrm{cr}} \dot\sim \mathcal{N}\left(O_j^{\mathrm{cr}},\ O_j/g_j + (R_j/g_j)^2\right)$.

Note

The flat frames are frequently assumed to be normalized, i.e., the mean or median of F_j values of the frame is around 1. In the generalized error estimation given in this book, this normalization is not important. However, astronomers (including all tasks of IRAF) have conventionally used a very simplified version of error estimation, that is, the Poisson component of the pixel error is approximated as $\sqrt{O_j/g_j}$: $O_j \dot{\sim} \mathcal{N}(O_j, O_j/g_j)$. Note the factor F_j is missing! The true signal-to-noise ratio is $O_j/\Delta O_j \approx \sqrt{O_j/(g_j F_j)}$, while in this approximation, it is $\sqrt{O_j/g_j}$. If the flat was normalized, $F_j \sim 1$, so these two are similar. If it were not normalized, however, the signal-to-noise ratio in this classical approximation will be underestimated by a factor of $\sqrt{F_j} \sim \sqrt{10^4\,(\mathrm{ADU})} = 100$!! You will be calculating wrong uncertainty in that case.

From STSDAS package, `stsdas/pkg/hst_calib/wfpc/noise/fitnoise.x`: noise = R if pixel value not positive, otherwise, $\sqrt{R^2 + N/g + (N \times \texttt{scalen}/100)^2}$ where `scalen` is maybe an uncertainty of pixel (flat uncertainty?) in percentage.

Chapter 4

Standardization

4.1 Problem Statement

The standardization process is a process to convert the instrumental values to physical values. In CCD, what is being read is the potential ($\propto$ number of electrons $\propto$ flux). But what does a CCD pixel value mean? 1 ADU can mean 1 Jy at one CCD but at different one it can mean 5 Jy because it is designed to be insensitive to photons for some reason. Thus, what astronomers do is

1. Make a list of objects which have known flux (e.g., star A has spectrum of blahblah, and it has $\mathrm{V}_{\mathrm{std}}$ magnitude or flux I_{std} in the V-band). These stars are called **standard stars**.
2. Observe the target and the standard stars simultaneously. If they cannot be in the same field of view, observe them at the same night when airmasses are not too different and weather is not changed largely.

Now the power of CCD comes in: It's highly linear, i.e., the pixel counts of N (of the target of interest) and N_{std} are very much proportional to the original flux, I and I_0. Thus, you can use Pogson's formula, because what it requires is only the ratio of flux:

$$\mathrm{V} - \mathrm{V}_{\mathrm{std}} = -2.5 \log_{10} \frac{I}{I_{\mathrm{std}}} = -2.5 \log_{10} \frac{N}{N_{\mathrm{std}}} \ . \tag{4.1}$$

Ex 15 (*Simplest Standardization*)

If the aperture photometry gave pixel count of 1000 for a standard star of $\mathrm{V}_{\mathrm{std}} = 10.00^{\mathrm{m}}$ and the object had pixel count of 500, the above formula will give $\mathrm{V} = 10.75^{\mathrm{m}}$.

In practice*, especially when stars with catalogued magnitude are not in the same field of view as the target of interest, it is very difficult to use these formulae. In such a case, we have two FITS images to compare: one for target and one for standard star (usually far away from the target). A direct comparison of N and N_0 is difficult because

1. The atmosphre exists. The magnitude we observe on ground is different from the one we would have observed outside of the atmosphere (space). This gives the k' and k'' terms in eq. (4.2) below.
2. The CCD is not ideally simple. For example, if it is more sensitive to redder wavelength, making red stars brighter than they should be. This gives the k term in eq. (4.2) below.

*From here, I extensively refered to Ch. 6 of "A Practical Guide to Lightcurve Photometry and Analysis" by Brian D. Warner, 2e.

If all these are considered with proper approximations (derived later in this chapter), we can obtain the following second-order approximation of the standard magnitude of an object seen on CCD:

$$\begin{aligned} M_f &= m_f + (\text{effect of atmosphere}) + (\text{effect of CCD}) \\ &\approx m_f - k'_f X - k''_f XC + z_f + k_f C \\ &\equiv m_{0f} + z_f + k_f C \ , \end{aligned} \tag{4.2}$$

where

$$m_f \equiv m_{0f} + k'_f X + k''_f XC \tag{4.3}$$

and

- f: The filter (V, B, g', etc).
- X: airmass (the simplest approximation is the secant of zenith angle, $\sec Z$).
- M_f: The *standard* apparent magnitude (or the *true* apparent magnitude) at filter f.
- m_f: The *instrumental* magnitude ($m_f = -2.5 \log_{10} N$).
- m_{0f}: The extra-atmospheric magnitude (m_f value we would have obtained if we were in space $X = 0$).
- C: The *true* color index*, e.g., B − V or r′ − i′.
- k'_f: The first order extinction coefficient at filter f.
- k''_f: The second order extinction coefficient at filter f.
- z_f: The zero point at filter f.
- k_f: The system transform coefficient at filter f.
- Note: lower- and upper-cased letters are used for the *instrumental* and *true* magnitudes, respectively†. From eq. (4.2):

$$\mathrm{V} - \mathrm{V}_{\mathrm{std}} = (\mathrm{v} - \mathrm{v}_{\mathrm{std}}) + k'_{\mathrm{V}}(X - X_{\mathrm{std}}) + k''_{\mathrm{V}}(XC - X_{\mathrm{std}}C_{\mathrm{std}}) + k_{\mathrm{V}}(C - C_{\mathrm{std}}) + \Delta z_{\mathrm{V}} \neq \mathrm{v} - \mathrm{v}_{\mathrm{std}} \ . \tag{4.4}$$

So the calculation given in the example is true only if the airmass of the object and standard star are identical AND the true color indices of them are identical. Otherwise, we cannot simply equate the right hand side of eq. (4.1) (= v − $\mathrm{v}_{\mathrm{std}}$) to the left hand side (= V − $\mathrm{V}_{\mathrm{std}}$). In space, we can remove all the atmosphere related terms since $X = 0$. Also $\Delta z \approx 0$ is assumed (discussed in section 4.2.4), so only the k_f term remains. This is why space observation is powerful.

4.2 Understanding the Standardization Formula

In this section, I will discuss about the terms in eq. (4.2) with some realistic data and plots. At the same time, I will give a derivation of the equation. Many textbooks only give the former; I do not want to go against that trend, but I wanted more quantitative explanations and justifications of those at the same time.

4.2.1 Atmospheric Extinction

The atmospheric extinction is dependent on the wavelength as in fig. 4.1. The extinction is expressed as mag/airmass, i.e., the extincted magnitude when airmass $X = 1$, i.e., "$m_f - m_{0f}$ at $X = 1$" or $k'_f + k''_f C$

*Not necessarily include filter f, but it is better that the wavelength ranges of the selected two filters "contain" the range of f for interpolation purpose.

†For example, v, b, $m_{g'}$ are instrumental magnitudes of an object and V, B, and $M_{g'}$ are true appparent magnitudes of it.

in the language of eq. (4.2). The extinction is severe at shorter wavelengths, and that is why the Sun looks redder when it rises or sets (i.e., when airmass is larger).

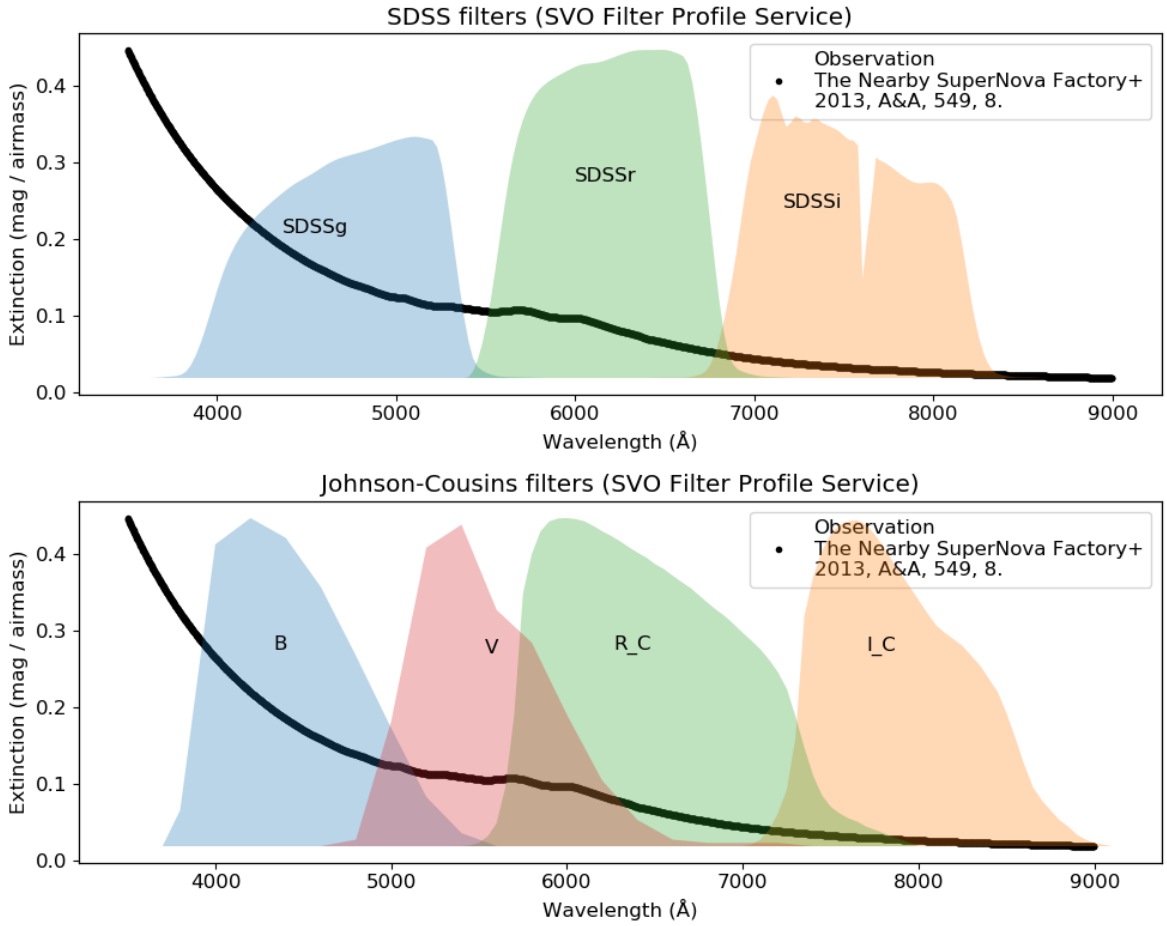

Figure 4.1: The atmospheric extinction as a function of wavelength at Mauna Kea, *based on some 4285 standard star spectra obtained on 478 nights spread over a period of 7 years obtained by the Nearby SuperNova Factory using the SuperNova Integral Field Spectrograph.* (excerpt from The Nearby SuperNova Factory+ 2013, A&A, 549, 8). The SDSS and Johnson-Cousins filters' filter profiles are overplotted.

Consider an object with spectrum $S_0(\lambda)$, measured at space, is observed at an airmass of X. The intensity at filter with profile $f_f(\lambda)$ without atmosphere is $\int_0^\infty S_0(\lambda)f_f(\lambda)d\lambda$. Because $f_f(\lambda)$ can be set as non-zero for $\lambda \in (\lambda_1, \lambda_2)$ and 0 otherwise, it can also be written as $\int_{\lambda_1}^{\lambda_2} S_0(\lambda)f_f(\lambda)d\lambda$. If the spectrum undergoes atmospheric extinction described by optical depth of $\tau(\lambda)$, the intensity after the filter throughput is $\int_{\lambda_1}^{\lambda_2} S_0(\lambda)f_f(\lambda)e^{-\int \tau(\lambda)dX}d\lambda \approx \int_{\lambda_1}^{\lambda_2} S_0(\lambda)f_f(\lambda)e^{-\tau(\lambda)X}d\lambda$, where $e^{-\tau(\lambda)X}$ is an approximation of $e^{\int -\tau(\lambda)dX}$, where the integration is along the optical path along the atmosphere. When extinction is not severe, i.e., $\tau(\lambda)X \ll 1$ for all λ of interest, $e^{-\tau(\lambda)X} \approx 1 - \tau(\lambda)X$ (approx 1). Also when $A \ll 1$,

$\log_{10}(1-A) \approx -A/\ln 10$ (approx 2). Combining these information, Pogson's formula states

$$\begin{aligned} m_f - m_{0f} &= -2.5\log_{10}\left(\frac{\int_{\lambda_1}^{\lambda_2} S_0(\lambda)f(\lambda)e^{-\tau(\lambda)X}d\lambda}{\int_{\lambda_1}^{\lambda_2} S_0(\lambda)f(\lambda)d\lambda}\right) \\ \text{(using approx 1)} &\approx -2.5\log_{10}\left(1 - \frac{\int_{\lambda_1}^{\lambda_2} S_0(\lambda)f(\lambda)\tau(\lambda)d\lambda}{\int_{\lambda_1}^{\lambda_2} S_0(\lambda)f(\lambda)d\lambda}X\right) \\ \text{(using approx 2)} &\approx \frac{2.5}{\ln 10}\frac{\int_{\lambda_1}^{\lambda_2} S_0(\lambda)f(\lambda)\tau(\lambda)d\lambda}{\int_{\lambda_1}^{\lambda_2} S_0(\lambda)f(\lambda)d\lambda}X \ . \end{aligned} \tag{4.5}$$

Remembering $I(\lambda) = I_0(\lambda)e^{-\tau(\lambda)X}$, we have extinction (magnitude)

$$\Delta m(\lambda) = -2.5\log_{10}\left(\frac{I(\lambda)}{I_0(\lambda)}\right) = 1.086\tau(\lambda)X \ .$$

Then the y-axis of fig. 4.1 is $1.086\tau(\lambda)$, so you can roughly understand that the y-axis represents $\tau(\lambda)$. Hence, $\tau(\lambda)X \ll 1$ (approx 1) is reasonable. In cases such as short wavelength (shorter than B/g) and high airmass observation, this assumption may break down. This is why classical photometric observers dislike observations at airmass $X \gtrsim 1.5-2$ which corresponds to elevation smaller than $48°-30°$. For polarimetry, however, only the *ratio* of two electric field vector directions are important, so airmass does not matter* The error due to the approximation, however, may not be severe compared to other error sources (e.g., changing weather).

Now we want to further assume that, $\tau(\lambda) \approx \tilde{c}_1 + \tilde{c}_2\lambda$ within the wavelength range of (λ_1, λ_2). This is similar to approximating the black markers in fig. 4.1 within each filter as a line because its y-axis is nothing but 1.086τ. Then

$$m_f - m_{0f} \approx 2.5\left(c_1 + c_2\frac{\int_{\lambda_1}^{\lambda_2} S_0(\lambda)f(\lambda)\lambda d\lambda}{\int_{\lambda_1}^{\lambda_2} S_0(\lambda)f(\lambda)d\lambda}\right)X \ . \tag{4.6}$$

Here, c_1 and c_2 are also constants. If the filter is fixed (e.g., V-band or SDSS g′ filter, etc), the only unknown thing in the second term in the parentheses is $S_0(\lambda)$, i.e., the spectral shape. If it is a black body spectrum, the shape of S_0 is determined uniquely once the color index C is known. Even if it is not a perfect black body, it is reasonable to assume the spectral shape, $S_0(\lambda)$, and color index, C, have *nearly* one-to-one relationship†. Fortunately most wildely used color indices (such as $B-V$ or gri colors) are more like one-to-one for *many* (but not all) cases. Thus, C is an indicator of $S_0(\lambda)$, so the second term is roughly a function of C, say $c_2\tilde{S}(C)$. The final assumption we make here is that the second term is $c_2\tilde{S}(C) \approx c_{3f} + c_{4f}C$ as the first-order approximation. Here c_{3f} and c_{4f} have subscript f because they depend on the *filter* profile, but not on the spectral shape under our simplifying assumptions, because all the dependency from the spectral shape is absorbed into C. Then

$$m_f - m_{0f} \approx 2.5(c_1 + c_{3f} + c_{4f}C)X \equiv k'_f X + k''_f CX \tag{4.7}$$

*For example, `ItoT+2018, NatCo, 9, 2486` demonstrated the polarization degree is not seriously affected by airmass even up to 7 compared to that of 1.03, at least less than 0.05 %p. This is because atmospheric scattering is basically a *forward* scattering, which should not induce any additional polarization degree although the total intensity should decrease.

†For example, if you look at the color-color diagram of stars, A0, F0, and G0 stars all share similar $U-B$ colors, i.e., color index and spectral shape are not one-to-one. This happens because (1) star spectra are not perfect black bodies and (2) filter profile is not "flat" as a function of λ. But to the first-order approximation, it is acceptable.

These are the origins of k'_f and k''_f in eq. (4.2).

To illustrate the result, I used the SDSS filter system as shown in fig. 4.2 and calculated how much magnitude extinction happens depending on the black body temperature at airmass $X = 1$ in the following table. Note the color C can be any gri color, such as $\mathrm{r}' - \mathrm{i}'$, and it is determined once the blackbody temperature is given.

Blackbody Temperature	Extinction magnitude $m_{\mathrm{g}'} - m_{0\mathrm{g}'} = k'_{g'} + k''_{g'}C$	$m_{\mathrm{r}'} - m_{0\mathrm{r}'} = k'_{r'} + k''_{r'}C$	$m_{\mathrm{i}'} - m_{0\mathrm{i}'} = k'_{i'} + k''_{i'}C$
3000 K	0.142^{m}	0.081^{m}	0.034^{m}
6000 K	0.158^{m}	0.084^{m}	0.035^{m}
20 000 K	0.171^{m}	0.087^{m}	0.036^{m}

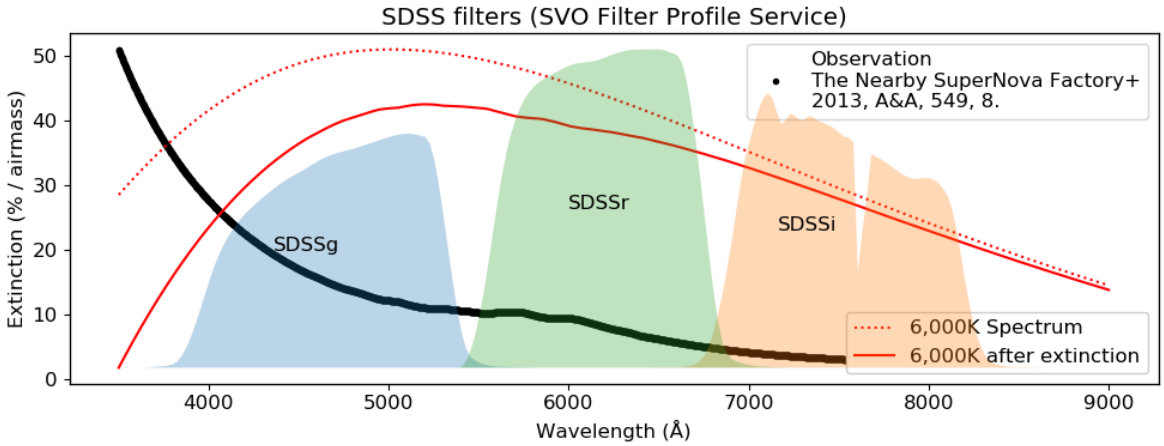

Figure 4.2: A black body radiation spectrum ($T = 6000\,\mathrm{K}$) before and after the extinction at SDSS bands. Note that the y-axis is changed to % per airmass (cf. fig. 4.1) by $10^{-0.4\Delta m}$.

As can be seen, *the extinction is stronger for higher temperature (lower color index)*, so likely $k''_f < 0$. However, the difference of extinction between the objects gets smaller as we look at longer wavelength. Thus, $k''_f \sim 0$ except for B or u filters.

These facts can also be understood qualitatively. The higher the temperature the spectrum will have more fraction of energe at shorter wavelength. Considering that the atmospheric extinction is stronger at shorter wavelength (see black markers in fig. 4.2), high termperature object will get more "penalty" when it comes into the atmosphere. That is why high temperature object is more strongly extincted, i.e., k''_f is negative. At the same time, since the amount of extinction drops significantly at wavelengths of r and i bands, and making k''_f itself very small.

SmithJA+ (2002, AJ, 123, 2121) determined the coefficients for SDSS filters at 1.0 m Ritchey–Chrétien telescope at the USNO Flagstaff Station, from the observations in 1998–2000 in table 4.1. The k'_f value fluctuates much for each night (see Fig 6 of the original publication), so I just took representative values from visual inspection. The k''_f values were obtained by two independent pipelines, `superExcal` (method 1) and `solve_network` (method 2). Both are quite consistent except for the u filter. The color C used for each filter is the nearby filter color: $\mathrm{u}' - \mathrm{g}'$ for u′, $\mathrm{g}' - \mathrm{r}'$ for g′, etc, and $\mathrm{i}' - \mathrm{z}'$ for both i′ and z′. The atmospheric extinction coefficients (k'_f, k''_f) all change as a function of time. We just hope they are

reasonably constant during our observation of our targets and standard stars. From experience, we know k''_f is always very small for filters longer than u or B-band, but is not necessarily ignorable because $k''_f XC$ might be larger than the accuracy you want to obtain.

Table 4.1: The extinction coefficients of SDSS from SmithJA+ (2002, AJ, 123, 2121).

Parameter	u′	g′	r′	i′	z′
k'_f	$> +0.5$	$+0.20 \pm 0.05$	$+0.10 \pm 0.05$	$+0.05 \pm 0.05$	$+0.05 \pm 0.05$
k''_f method 1	-0.021 ± 0.003	-0.016 ± 0.003	-0.004 ± 0.003	$+0.006 \pm 0.003$	$+0.003 \pm 0.003$
k''_f method 2	-0.032	-0.015	0.000	$+0.005$	$+0.006$

Ex 16 (*Ignoring the Second Order Extinction Term*)

Consider an observation at $X = 2$ of $C = 0.2$ star. If you ignore the second order extinction term, you are making $k''_f XC = 0.4k''_f$ of uncertainty. According to table 4.1, this is most likely smaller than 0.01 magnitude.

The Sun has $C = \mathrm{g}' - \mathrm{r}' \sim 0.5$, and then the error for Sun-like star becomes $1.0k''_f$: It is $< 0.01^{\mathrm{m}}$ for r, i, and z bands. It is about 0.02−0.03 mag for u′.

The red M0 stars have $C = \mathrm{g}' - \mathrm{r}' \lesssim 1.5$ and the error is now up to $3.0k''_f$. The accuracy of 0.01^{m} can be achieved in riz bands, but risky.

The calculation above is only for SDSS observatory at altitude of 2.3 km. But fortunately, BuchheimB (2005, SASS, 24, 111) found that the $k''_f \lesssim 0.005$ for V band even at many low-altitude observatories (including Bochum observatory at altitude 200 m, Vainu Bappu Observatory at altitude 700 m), so likely we expect the correction from the second-order extinction term is small enough.

4.2.2 Transformation Coefficient

The sensitivity of the optics, such as filter, lens, mirror, and/or CCD cover glass, is also a function of λ. The argument is identical to atmospheric extinction, but there is no X (similar parameter will be something like the optical depth of materials blocking CCD pixel, but that should be a device-dependent constant). Then the same logic leads us to the conclusion that there should be a color term which tunes the final output of the CCD count, and that is the $\tilde{k}_f c$ ("transformation") term. Here, c is the color index of the object after all the atmospheric extinction, and the true color before it enters telescopic optics. Once we assume there exists a function f_c such that $f_c(C) = c$ is a one-to-one function and $f_c(C) \approx \tilde{c}_5 + \tilde{c}_6 C$ for constants $\tilde{c}_5$ and $\tilde{c}_6$, the transformation term becomes $k_f C + \tilde{c}_{7f}$ for the filter-dependent constant $\tilde{c}_{7f}$. This constant is finally absorbed in to another filter-dependent constant, called the zero point, z_f. Therefore we reach eq. (4.2).

The transformation coefficient, which I denoted k_f, is fortunately nearly constant for the given device. Warner argues that it is enough to update k_f (Warner uses notation of T_f) value only about 2–4 times a year, unless you physically changed the device elements (e.g., filter, CCD, lens, etc). Moreover, from experience, we know that this is nearly zero: $|k_f| \lesssim 0.1$. Many cases $|k_f| \lesssim 0.01$. Since the range of color indices are $\max(\Delta C) \lesssim 1^{\mathrm{m}}$, we have $|k_f C| \lesssim 0.1$, and in many cases, $|k_f C| \lesssim 0.01$.

4.2.3 A Note on Linearity

As we noted at the beginning part of this chapter, CCD is highly linear. That means, $N = gN_e = \alpha N_\gamma$ where N is the pixel count (after *bias* and *dark* subtraction), N_e is the total number of photo-electrons,

g is the electronic gain of the CCD (a constant; unit of counts per electrons), and N_γ is the photon incident to the CCD from the true photon number $N_{\gamma 0}$. No higher-order terms, no other constants. The α value may differ from pixel to pixel due to the inhomogeneity of optics or CCD pixels, but they are homogeneized by the so-called *flat fielding*, so here I can safely say it is strictly constant over all the pixels.

Any kind of extinction (atmosphere or optics in front of the CCD) is multiplicative only to the linear term, i.e., $N_\gamma = N_{\gamma 0} \times \text{something}_1$ (e.g., $e^{-\tau}$). There is neither higher-order terms like $N_{\gamma 0}^2$ nor addition of constant. Therefore, $N = \text{something}_2 \times N_{\gamma 0}$ and the instrumental magnitude

$$m_f := -2.5 \log_{10} N = \text{something}_3 - 2.5 \log_{10} N_{\gamma 0} \equiv \text{something}_3 + M_f \ . \tag{4.8}$$

Thus, thanks to the linearity of CCD, we have **no additional coefficient *multiplied*** in front of m_f or M_f. If, for example, N were $\alpha N_\gamma + \alpha'$, or there were other terms in the extinction (proportional to N^2 or a constant radiation from the optics), this simple relation wouldn't hold.

To emphasize, *you should not worry about whether to* multiply *something in front of M_f or m_f to satisfy eq.* (4.2). Their coefficients *must be unity.* From our experiences, most observational experts and electrical engineers would say that you should only care about this if you are sure that some parts of the optics have serious problems (e.g., your CCD underwent serious problem and shows non-linearity).

4.2.4 A Note on Zero Point

The zero point z_f is a constant to convert the instrumental magnitude (which is nothing but a -2.5 multiplied by $\log_{10}(\text{count})$) to a realistic standard magnitude system astronomers have been using. In intensity sence, this "addition of a constant" in magnitude represents a "multiplication of a constant" to the instrumental count to measure the intensity.

What is a typical value of zero point? From telescopes with diameter $\gtrsim 1\,\text{m}$ in Seoul, let's say we are observing objects of around 15^m with 100 seconds exposure. Say the sky-subtracted peak value of our target is* $\sim 10^4$ and the integral of the profile gives intensity of $\sim 10^5$, and the intensity per unit time is 10^3 (divided by exposure time). The instrumental magnitude is $m \sim -2.5 \log_{10} 10^3 = -7.5$. This means $z_f \sim 22.5$. If we do realistic calculation, z_f is mostly within a range of 20 to 25. This is why, in IRAF, the default initial guess of z_f is 25 mag.

Zero point must be a constant unless the device is affected by external disturbance in our simple model dealt in this chapter. In reality it is true that this zero point fluctuate at each exposure, and that is because of the imperfect readout process of CCD electronics. $\Delta z_\text{V} \approx 0$ is assumed in this chapter. Although you may be unconfortable, but this is what is assumed even in professional space telescope data reduction processes, if this fact makes you more comfortable.

4.3 Standardization Applied to Photometry

Now that we justified the usage of eq. (4.2), let's find the cases of application. The simplest case is the *differential photometry.*

*Because CCDs are usually operated in 16-bit unsigned integer mode which represents 0 to 65,535, and non-linearity appears around 40,000, the sky subtracted peak pixel value of an intermediate-brightness object in the FOV is order of $\sim 10^4$ after sky subtraction

If there are many celestial objects (of course including your target) in the field of view with known standard magnitudes, we can use them as standard stars. Although there can be variable stars and galaxies*, if most of the objects with known magnitudes are non-variable stars, those *outliers* will be smoothed out. Thus, we just assume all the celestial objects in the field of view with known standard magnitudes as "standard stars".

This technique is very widely used in variable star and asteroidal light curve observations. This is widely used than absolute photometry, because it is annoying and difficult to observe standard stars at different airmasses while observing your target, which requires telescope time and human labor. In asteroidal science, even single-filter differential photometry is frequently used.

4.3.1 Differential Photometry: Single-Filter

Consider there are many stars in the FOV with known (catalogued) magnitudes at multiple filters, so that the true apparent magnitude M_f and color C are known. Say, from the photometry, we could determine the instrumental magnitude of stars m_f. Rearranging eq. (4.2):

$$M_f - m_f = (z_f - k'_f X) + (k_f - k''_f X)C \equiv a_f(X) + b_f(X)C \tag{4.9}$$

The first term in the RHS, $a_f(X)$, is a constant for all the stars in the same FITS frame, as they will have nearly identical airmass† X and the identical zero point z_f. $M_f - m_f = a_f(X)$ for $C = 0$, i.e., a perfectly flat spectrum. The second term is color-dependent, but as we discussed before, it is likely to be very small. Sometimes people just call this value (LHS) as "zero point", although I will stick to use this term for z_f for clarity.

Ex 17 (*How small is the second term?*)

Consider observatinos made in wavelength ranges around V or gri bands. From table 4.1, $|k''_f| \sim 0.000-0.020$ and we expect that $k_f \lesssim 0.02$ for most optics. Since k''_f is mostly negative, $k_f - k''_f X$ is likely be positive, but not always (see table 4.1). Also we have $1 < X \lesssim 2$ in most cases. Then if you play with many combinations of numbers, $b_f(X)C = (k_f - k''_f X)C \lesssim 0.05C$ and most likely much smaller than that. Note that the color index is mostly $-1 \lesssim C \lesssim +1$.

What we do now is to draw several test plots as in fig. 4.3. It is better if the color of the stars span wider than about 0.5 ($\max(C) - \min(C) \gtrsim 0.5$). On the left side of the figure, I plotted M_f VS m_f for $f = \mathrm{R}$ (Johnson–Cousins $\mathrm{R_C}$ filter). The fitted slope is ~ 1.015, which is near the unity as we expect from linearity. Sometimes it becomes as high as 1.05 throughout the night. Below is a residual plot. Normally, since the uncertainty in the magnitude measurement increases for larger magnitude (fainter star), the scatter of the residual must increase at large magnitude. The right panel of the figure shows the residual as a function of catalogued stars. The color-dependent slope is 0.02, and it was $\lesssim 0.05$ throughout the night; this is small indeed as we expected above. As we do not know our target's color, assuming the

*Galaxies can have spectra significantly different from those of black bodies. Therefore, the coefficients k_f and k''_f, which are approximations of spectral shape (i.e., not k'_f), should be different from those derived from standard *stars*, which are black bodies to the first order. But mostly this effect is not serious because, as we discussed before, both $k_f C$ and $k''_f XC$ terms are anyway very small. For this reason, people use k_f and k''_f derived from standard stars for their target galaxies (or any non-black body like spectra). If you really worry about this, you must conduct spectroscopic observation, not broad-band photometry.

†If you observed near the horizon, ΔX across the FOV may not be negligible. For accurate photometry, this should also be taken into account.

color uncertainty of ±0.5 will give around ±0.03 ***systematic offset to the magnitude*** of our target. Note that this is a systematic parallel shift in magnitude, not a random error. Therefore, the *shape* of the light curve will not be affected by this, althought the magnitude *value* may have been affected by a constant.

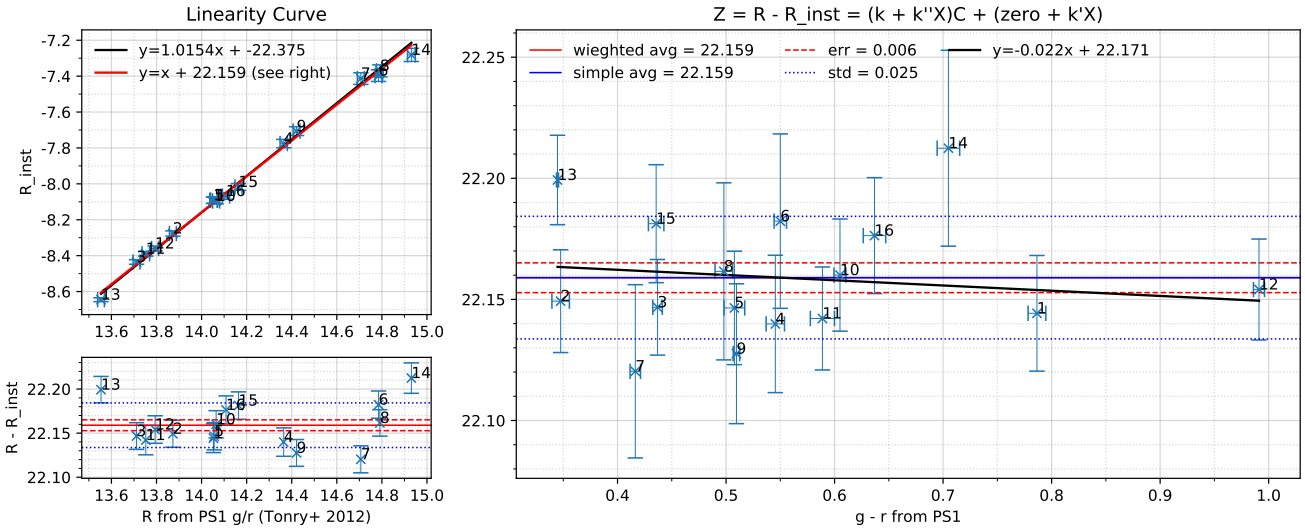

Figure 4.3: A typical linearity curve with residual (left) and color-zero plot (right) on 2018-10-12 (UT 19:08:29) observation at SNUO 1-m telescope for the field stars near an asteroid (155140) 2005 UD. I used PS1 DR1 catalog, within 13.5 to 15 mag, removed any object with flag of quasar, variable, galaxy, etc, and dropped any pairs of stars if there's any nearby object in PS1 DR1 catalog.

A summary of the three plots:

Table 4.2: Differential Photometry Diagnostic Plots

Name	x-axis	y-axis	Appearance	Comments
Linearity Curve	M_f	m_f	straight line slope of unity	$\lvert\text{slope} - 1\rvert \gtrsim 0.01$ means you need to check your photometry and/or device linerity
Linearity Residual	M_f	$M_f - m_f$	constant value regardless of M_f	Scatter is larger for large M_f objects ($\because$ faint object's mag is less accurate)
Color Dependency	C	$M_f - m_f$	$\sim$ constant regardless of M_f	If you can see a clear trend, the color-terms in eq. (4.9) not negligible.

If graphs do not look as expected, that means (1) many field objects are variable or highly non-blackbody-like galaxies so that they are not suitable as "standard stars" and the basic assumptions of eq. (4.2) break down, (2) the airmass is too high that the 2nd order approximations do not hold, (3) catalog M_f suffer from unknown uncertainties or systematic errors so the catalog value is not reliable, (4) m_f was measured incorrectly, (5) many other possibilities. Check these before going further, because your final results may be affected.

Consider you determined the (a_f, b_f) for all the FITS files you obtained through a night, where each FITS file has different airmass. Then, if you

- plot a_f as X, the intercept and slope are z_f and $-k'_f$,
- plot b_f as X, the intercept and slope are k_f and $-k''_f$.

Note, however, the following assumptions should be met:

1. Atmospheric conditions (extinction coefficients k' and k''), remain constant over the night
2. Zero point z_f remain constant over the night
3. The approximations used for eq. (4.2) (2nd-order approximation) must hold

Because not all these are met in Seoul sky, the plots you made will not look like to have a linear trend. It can be largely scattered so that a linear fit does not seem to make sense, clear non-linear trend appears (mostly due to the third assumption breaks down at large airmass), etc.

Therefore, as a simple yet reasonable approximation, I assumed $M_f - m_f$ for all the objects in a single FITS file should be a constant, and found this value. Then added it to the instrumental magnitude m_f of the target of interest.

4.3.2 Differential Photometry: Multi-Filter

When we have more than one filter observation, we can even eliminate the systematic offset due to the color uncertainty of the object. Just write the equation for two filters, x and y

$$\begin{cases} M_x - m_x &= (z_x - k'_x X_x) + (k_x - k''_x X_x) C_{xy} = a_x(X_x) + b_x(X_x) C_{xy} \\ M_y - m_y &= (z_y - k'_y X_y) + (k_y - k''_y X_y) C_{xy} = a_y(X_y) + b_y(X_y) C_{xy} \end{cases} \tag{4.10}$$

Note here that, for field stars with known standard magnitudes, M_x, M_y, and thus $C_{xy} = M_x - M_y$ should all be known. The m_x and m_y are known from a photometry to the image. Although z values are assumed to be constant throughout the night for the same detector, I explicitly put z_x and z_y for generality. Following the logic of single-filter case, if we plot $M_x - m_x$ as ordinate and C_{xy} as abscissa for N field stars, the y-intercept is a_x and the slope is b_x. Same goes for the y filter. So (a_x, b_x) and (a_y, b_y) are determined with the uncertainties.

Then for the target of interest, which M_x and M_y are unknown,

$$\begin{cases} M_x^{\text{target}} - m_x^{\text{target}} &= a_x(X_x) + b_x(X_x) C_{xy}^{\text{target}} \\ M_y^{\text{target}} - m_y^{\text{target}} &= a_y(X_y) + b_y(X_y) C_{xy}^{\text{target}} \end{cases} \tag{4.11}$$

or since $C_{xy}^{\text{target}} = M_x^{\text{target}} - M_y^{\text{target}}$, (dropping X_x and X_y for brevity)

$$C_{xy}^{\text{target}} = \frac{(m_x^{\text{target}} - m_y^{\text{target}}) + (a_x - a_y)}{1 - (b_x - b_y)} \ . \tag{4.12}$$

Putting this back to the original equation,

$$\begin{cases} M_x^{\text{target}} = m_x + a_x + b_x \frac{(m_x^{\text{target}} - m_y^{\text{target}}) + (a_x - a_y)}{1 - (b_x - b_y)} \\ M_y^{\text{target}} = m_y + a_y + b_y \frac{(m_x^{\text{target}} - m_y^{\text{target}}) + (a_x - a_y)}{1 - (b_x - b_y)} \end{cases} \tag{4.13}$$

Now we have the standard magnitude of the target in both x and y filters. Since we have taken the color of the target into account, ***there is no systematic offset*** as in single-filter case.

What if we have more than 2 filters, say x, y, and z? Solve for x and y as above by using (a_x, b_x). Then solve for y and z, but this time using color as C_{yz}, not C_{xy}, using (a'_y, b'_y). Theoretically $a_y = a'_y$, because there is no color term. However, $b_x \neq b'_x$, because k''_y is defined for C_{xy} for the first case, but it is defined for C_{yz} for the second case. In reality, because what we get is only the best-fit values with uncertainties, a_x and a'_x can also be different (likely within certain amount of error-bar).

4.4 Photometry Using Standard Stars

When we need accurate absolute photometry, photometric standard star observation is essential. Unlike "field stars with known magnitude," standard stars are confirmed as non-variable stars with very accurate magnitudes. By observing them, we determine the coefficients (k_f, k'_f, and k''_f) and zero point (z_f), mostly for at least two filters, and get the magnitude of our target of interest with high accuracy. Hence, I will only talk about multi-filter observation of standard stars.

A different thing for standard star frames is that there is only one star with known magnitude in the FOV, and there will be only one point in the right panel of fig. 4.3. Therefore, we select at least two standard stars with different colors, sometimes called a ***blue–red pair***. Observe one standard star at an airmass. When the other standard star reaches similar airmass, observe it, so that you have at least two stars at the same airmass. Then you have two points in the right panel of fig. 4.3 at the given airmass. Repeat this for many airmasses, so that you obtain the a_f and b_f for many airmasses for each filter. Following the logic of multi-filter case, you can determine all the coefficients.

If you are not interested in zero point and the transformation coefficient k_f, you can follow this calculation: Consider a standard star with ID $= i$ is observed for N_i times, and denote each observation as j ($j = 1, \cdots, N_i$), and write the parameters of eq. (4.2) as $M_{f,i}$, C_i (M and C are fixed values for a standard star, so no j is needed), $m_{f,i}^{(j)}$, $X_{f,i}^{(j)}$, etc. We here assume the zero point, z_f, and the coefficients k_f, k'_f, and k''_f are almost constant over the night*. Then

$$\begin{cases} M_{f,i} - m_{f,i}^{(1)} &= z_f + k'_f X_{f,i}^{(1)} + k''_f X_{f,i}^{(1)} C_i + k_f C_i \ , \\ M_{f,i} - m_{f,i}^{(2)} &= z_f + k'_f X_{f,i}^{(2)} + k''_f X_{f,i}^{(2)} C_i + k_f C_i \ , \\ \quad\vdots \qquad\qquad & \quad\vdots \\ M_{f,i} - m_{f,i}^{(N_i)} &= z_f + k'_f X_{f,i}^{(N_i)} + k''_f X_{f,i}^{(N_i)} C_i + k_f C_i \ . \end{cases} \tag{4.14}$$

For the first two observations of the same standard star (ID $= i$) at filter f, subtracting two,

$$\Delta m_{f,i}^{(1,2)} = (k'_f + k''_f C_i) \Delta X_{f,i}^{(1,2)} \quad \rightarrow \quad \left(\frac{\Delta m}{\Delta X}\right)_{f,i}^{(1,2)} = k'_f + k''_f C_i \ . \tag{4.15}$$

Of course $\Delta m_{f,i}^{(1,2)} = m_{f,i}^{(1)} - m_{f,i}^{(2)}$ and $\Delta X_{f,i}^{(1,2)} = X_{f,i}^{(1)} - X_{f,i}^{(2)}$.

Therefore, if we plot $\left(\frac{\Delta m}{\Delta X}\right)_{f,i}$ as a function of color C_i of the standard stars with different color indices (colors of them are all known), the linear fit will give intercept of k'_f and slope of k''_f. For star i, we have N_i observations, so we have $\binom{N_i}{2} = N_i(N_i-1)/2$ points at the single C_i value. If we have N standard stars of wide color range, we can have $N \sum_i \binom{N_i}{2}$ points to fit the linear line. For a simple blue–red pair, $N = 2$, so you have two x-axis values (color index), but many points at each x-axis value.

*Even SDSS standard stars were also observed and analyzed under this assumption. See SmithJA (2002, AJ, 123, 2121)

Chapter 5

Statistics - Bayesian

Spectroscopy is a great starting point for Bayesian statistics to those who're interested in astronomy. This is because it has plenty of data points along the x-axis (e.g., wavelength) with error-bars, and we usually fit a simple analytic function to the dataset.

Consider you have the 1-D spectrum as in fig. 5.1. The big question is this: **How likely is that the peak is an actual line(s), not due to random noise?**

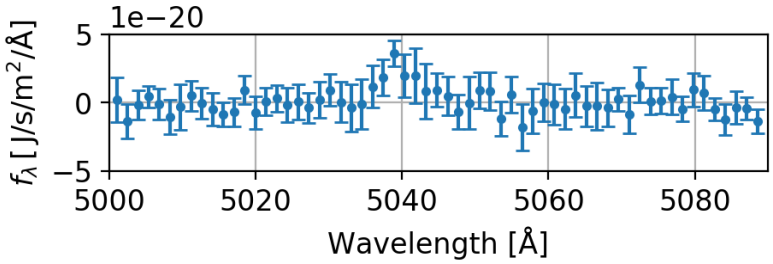

Figure 5.1: The data extracted from Greco+2018ApJ, without Gaussian fit. Note that $1\,\mathrm{erg/s/cm^2/Å} = 1\,\mathrm{mW/m^2/Å}$. See the notebook `Spectroscopy_Simulation` for the codes I used in this chapter.

To test this, let's set up hypotheses: The null hypothesis (H_0) is that there is no line, while the alternative hypothesis (H_1) is that there *is* a line*. Note that we haven't specified the properties of the line (amplitude, width, etc) yet. Then our strategy must be

1. We have to check which hypothesis is more likely (assuming there're only two possibilities, H_0 and H_1).
2. If H_0 is more likely, we accept the null hypothesis, i.e., "we argue the non-existence of line in this data with certainty of xxx".
3. If H_1 is more likely, we have to find the line properties (amplitude, width, etc) in the form of $x \pm dx$.

The first is called the **Model Selection** and the third is called the **Parameter Estimation**. If the second is the case, things are simple: no line! That's all. Model selection is *discrete* (finite number of different hypotheses), while parameter estimation is usually *continuous*.

*Reminder: *the null hypothesis is what we want to reject.*

5.1 Bayes Theorem

Thm 12 (*Bayes' Theorem*)

If H, D, and I are the hypothesis, data, and prior information, respectively, Bayes' theorem states

$$P(H|D,I) = \frac{P(H|I) \times P(D|H,I)}{P(D|I)} \ . \tag{5.1}$$

Usually it is simply put as

$$P(H|D,I) \propto P(H|I) \times P(D|H,I) \ . \tag{5.2}$$

The **posterior** $P(H|D,I)$ is understood as "the probability that the hypothesis is true, given the data equals to D and we have prior information I." The **prior** $P(H|I)$ can be quite subjective, since it means "the probability that the hypothesis is true." In most cases, we take either uniform prior or Jeffrey's prior (see below). The denominator, $P(D|I) := \int P(D|H,I)P(H|I)dH$ is understood as the normalization constant for the posterior.

The term $P(D|H,I)$ is called the **likelihood**, and this is what we have calculated during our high school course works:

Ex 18 (*High school exam problem: calculating liklihood*)

Given the hypothesis that a person has possibility 0.5 to win a game, what is the probability for her to win 3 games in a row?

In this case, the hypothesis is $H = (p = 0.5)$, and the data $D =$ (win 3 in a row). The prior knowledge is that "the game rule does not change", "the person doesn't use faul measures to make the possibility to change over time", etc, which are just trivial assumptions. Most importantly, we have to assume that the results of each game should be *independent*, so that we can use the multiplication law. Therefore,

$$P(D|H,I) = (1/2)^3 = 1/8 \ .$$

In real scientific problems, I can be something like "the fundamental physical laws do not change over time", "this galaxy is spiral for sure (if you're working on rotational curves of S galaxies, you may assume your samples are definitely S, not E or Irr)", etc.

When you are finding the mean of the data, it is a single parameter case, and therefore, H just means the mean value, so you can write something like this: $P(H|D,I) = P(m|D,I)$. When you have multiple parameters, usually people denote the set of paramers as $\vec{\theta}$, so $P(H|D,I) = P(\vec{\theta}|D,I)$. If we're fitting the data with, say, a 3rd order polynomial with 4 parameters, we can understand that I includes "the data is described by a 3rd order polynomial with 4 parameters".

5.2 Towards the Model Selection

5.2.1 Model Selection Concept

The term "odd" is used for $\frac{\text{probability}}{1-\text{probability}}$. In this case the $P(H_1|I)/P(H_0|I)$ is called the prior odds because $P(H_1|I) + P(H_0|I) = 1$, as there is no other possibility other than H_0 and H_1. Similarly, since $P(D|H_0,I) + P(D|H_1,I) = 1$, the LHS is also called the posterior odds.

Now come back to the original question: model selection. The model selection is done based on the **odds ratio**:

$$R := \frac{P(H_1|D,I)}{P(H_0|D,I)} = \frac{P(H_1|I)}{P(H_0|I)} \times \frac{P(D|H_1,I)}{P(D|H_0,I)} = \text{odds(prior)} \times B_{10} \ . \tag{5.3}$$

Here B_{10} is called the **Bayes' factor**.

If $R > 1$, we select H_1 over H_0 and vice versa. If $R = 1$, we can't give any conclusion. Equivalently, we can use $\ln R$ and compare it with 0. This is because we frequently get extremely large or small R values (e.g., 10^{-40}), which is inappropriate for some computer programming.

5.2.2 Prior Selection

There are two major priors: uniform and Jeffreys' prior.

- **Uniform prior** assumes uniform probability for the model within reasonable range.
- **Jeffreys' prior** assumes the probability proportional to the determinant of the Fisher information matrix. Simply put, for the parameters like standard deviation, if the value is small (has more information), we give more weight to it.

Another possibility is that iteratively updated prior based on the accumulated data, as sometimes used in artificial intelligence.

Ex 19 (*Uniform prior*)

Consider an emission line fitting to the 1-D spectrum. If you are sure that the line must be Gaussian with center $\lambda_c \in [650, 655]$ nm, the uniform prior of the line center will be $\lambda_c \sim \mathcal{U}(650, 655)$, which has the pdf of $p(\lambda_c) = 1/5$ for $\lambda_c \in [650, 655]$ nm and 0 otherwise. That is, uniform prior regards it is equally likely to have any value within the bound specified by the user.

5.2.3 Likelihood Calculation

Assume all data values, e.g., the f_λ values, are independent*. Denoting the i-th pixel's value as D_i,

$$P(D|H, I) = P(D_1, \cdots, D_N|H, I) = P(D_1|H, I) \times \cdots \times P(D_N|H, I) \ .$$

Under H_0 (no emission line), $P(D_i|H_0, I)$ is the probability to measure D_i electrons when there is only the pixel noise σ_i (no actual line). Each pixel we have Poissonian noise which is nearly Gaussian, and sky estimation noise term, which is difficult to quantify but we usually assume it is Gaussian, and the Gaussian readout noise in standard CCD. Therefore, each pixel is assumed to have a Gaussian noise. Thus,

$$P(D_i|H_0, I) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left\{\left(-\frac{(D_i - 0)^2}{2\sigma_r^2}\right)\right\} \quad \rightarrow \quad P(D|H_0, I) = \prod_{i=1}^{N} P(D_i|H_0, I)$$

Then the **log-liklihood** is

$$\ln P(D|H_0, I) = C_\sigma - \sum_{i=1}^{N} \frac{D_i^2}{2\sigma_i^2} \ , \tag{5.4}$$

where $C_\sigma := \sum \ln\left(\sqrt{2\pi}\sigma_i\right)^{-1}$ is a constant. Note that even the second term is a calculable constant once the data is given.

For H_1, we can just change the zero mean to the Gaussian line profile. Consider the best-fit line profile is described as $g(x|\vec{\theta_0})$ for $\vec{\theta_0} = (A, w, \lambda_c)$ (amplitude, width sigma, central wavelength). Then

$$\ln P(D|H_1, I) = C_\sigma - \sum_{i=1}^{N} \frac{(D_i - g(x_i|\vec{\theta_0}))^2}{2\sigma_i^2} \equiv C_\sigma - \frac{1}{2}\chi^2 \ . \tag{5.5}$$

*A strange thing happens on CCD sometimes maybe? If this is true, then nothing is independent. See BooneK+18 PASP (`2018PASP..130f4504B`) "A Binary Offset Effect in CCD Readout and Its Impact on Astronomical Data ".

Here, χ^2 is the usual chi-square statistic from the data and model, because in our case, the error-bars are independent and normally distributed. Finding the best-fit parameter set, $\vec{\theta}_0$, is done by the least-square fitting (also called the χ^2-minimization).

Although it is mathematically too trivial, let me put another theorem to emphasize.

Thm 13

Maximizing (log-)likelihood is identical to minimizing χ^2, when the error-bars are independent and follows Gaussian.

5.2.4 Model Selection Calculation

Now let's do the real calculation to compare H_0 and H_1. Because I want to use amplitude rather than flux as a free paramter, the gaussian function will be $g(x|\vec{\theta}_0) = Ae^{-(\lambda-\lambda_c)^2/2w^2}$. The best fit Gaussian function to the data shown in fig. 5.1 is found to have the following parameters:

$$\text{amplitude} = 3.213 \times 10^{-20} \quad ; \quad \lambda_c = 5039.4\,\text{Å} \quad ; \quad w = 2.33\,\text{Å} \tag{5.6}$$

with the integrated flux $\log_{10}(F_{\text{O III}}/\text{mW m}^{-2}) = -15.73$, because $F = A\sqrt{2\pi w^2}$. This matches well with the original publication -15.7 ± 0.1.

From the data, $C_\sigma = 2739.234$. The log-likelihood of H_0 and H_1 are

$$\begin{aligned} \ln P(D|H_0, I) &= C_\sigma - \sum_{i=1}^{N} \frac{D_i^2}{2\sigma_i^2} && = 2720.810 \\ \ln P(D|H_1, I) &= C_\sigma - \sum_{i=1}^{N} \frac{(D_i - g(x_i|\text{amplitude}, \lambda_c, w))^2}{2\sigma_i^2} && = 2730.136\ . \end{aligned}$$

Assume the probability of H_0 being true is the same as that of H_1 being true. The Bayes' ratio in eq. (5.3) therefore becomes just a Bayes' factor:

$$R = B_{10} = \frac{e^{2730.136}}{e^{2720.810}} = 1.12 \times 10^4 \gg 1\ , \tag{5.7}$$

which means H_1 is extremely more likely. Thus, we now believe there must be an emission line, and have to find the CI of the parameters (e.g., flux value).

5.2.5 Model Selection with AIC, BIC

But wait, is this all? No.

The pitfall of this naïve approach using Bayes' ratio is that, you can minimize χ^2 to 0, by increasing the number of fitting parameters. When the number of free parameters is equal to the number of data points, you must be able to make a function f such that $D_i - f(\lambda_i)$ is always 0. Does the N-parameter model a better choice than 1-parameter case, for example, if just the ratio is large? It can't be.

The "number of paramter" problem is not an easy thing to solve, but we have simple rule-of-thumbs: The Akaike Information Criteria, AIC, and the Bayesian Information Criteria, BIC. The derivations are not shown here, because it can become a bit lengthy while that derivation itself is not at the heart of the understanding.

Thm 14 (*Bayesian Information Criterion; BIC*)

For the $N(\to\infty)$ data points, the BIC for the model with n free parameters denoted as $\vec{\theta}$ is given as

$$\mathrm{BIC} := n\ln N - 2\ln P(D|\vec{\theta}_0, I) = n\ln N + \chi^2_{\min} - 2C_\sigma \ , \tag{5.8}$$

where $\vec{\theta}_0$ is the best-fit parameter which results in the minimum χ^2, or maximum likelihood ($P(D|H,I)$). For the same data, comparing with two different models, the difference in the BICs is used:

$$\Delta\mathrm{BIC} = (n_1 - n_2)\ln N - 2\ln\frac{P(D|\vec{\theta}_{0,1}, I)}{P(D|\vec{\theta}_{0,2}, I)} = (n_1 - n_2)\ln N + (\chi^2_{\min,1} - \chi^2_{\min,2}) \ . \tag{5.9}$$

The second equalities above including χ^2 hold only if the error-bars are independent and normally distributed.

Some people prefer to define in the opposite sign and/or half of this value to remove the factor 2 in front of the log-likelihood.

Note that BIC is usable only if $N \gg n$. Also the prior distribution only affects when you find $\vec{\theta}$, not when calculating the BIC.

Since RafteryAE's work*, the following criteria for choosing models are widely used:

$\lvert\Delta\mathrm{BIC}\rvert \in$	[0, 2]	[2, 6]	[6, 10]	10+
Evidence	Not worth more than a bare mention	Positive	Strong	Very Strong

Thm 15 (*Akaike Information Criterion; AIC*)

For the $N(\to\infty)$ data points, the BIC for the model with n free parameters denoted as $\vec{\theta}$ is given as

$$\mathrm{AIC} := 2n - 2\ln P(D|\vec{\theta}_0, I) = 2n + \chi^2_{\min} - 2C_\sigma \ , \tag{5.10}$$

where $\vec{\theta}_0$ is the best-fit parameter which results in the minimum χ^2, or maximum likelihood ($P(D|H,I)$). For the same data, comparing with two different models, the difference in AICs is used:

$$\Delta\mathrm{AIC} = 2(n_1 - n_2) - 2\ln\frac{P(D|\vec{\theta}_{0,1}, I)}{P(D|\vec{\theta}_{0,2}, I)} = 2(n_1 - n_2) + (\chi^2_{\min,1} - \chi^2_{\min,2}) \ . \tag{5.11}$$

The second equalities above including χ^2 hold only if the error-bars are independent and normally distributed.

Although I am not so familiar with hard-core statistics, it seems like there are debates about which should be preferred (BIC or AIC). BIC is known to prefer the "true model", if it is in our set of alternative hypotheses, with probability 1 when $N \to \infty$, while AIC doesn't. On the other hand, if the data is too few (note that the difference in BIC and AIC is the $\ln N$ term), BIC tend to seek for the model which explains that small dataset, so it can prefer worse model which tries to explain the bad data points than AIC.

Now let's calculate ΔBIC and ΔAIC for our sample data to test H_0 VS H_1. As before, $C_\sigma = 2739.234$, and log-likelihoods are 2720.810 and 2730.136. Then because H_0 has no parameter and H_1 has three paramters, and $N = 61$,

$$\begin{aligned} \Delta\mathrm{BIC} &= (0-3)\ln 61 - 2(2720.810 - 2730.136) &&= 6.32 \\ \Delta\mathrm{AIC} &= 2(0-3) - 2(2720.810 - 2730.136) &&= 12.65 \end{aligned}$$

*RafteryAE (1995) "Bayesian Model Selection in Social Research", Sociological Methodology, 25, 111

The fact that these are positive means we should prefer the H_1, but not as strong as what we've seen from the odd's ratio.

5.3 Towards the Parameter Estimation

5.3.1 Brute-Force

From the last section, we learned we have a clear evidence that there is a line. Then how can we estimate the line properties with uncertainties? This is the same as we did in the statistics chapter. First is the **brute-force** fixed-grid search scheme

In the chi-square sense, what we have to do are

1. Calculate the chi-square statistic at each parameter space position.
2. Keep only those with $\chi^2 < \chi^2_{\rm min} + \Delta(n_\theta, \alpha)$
   - Δ: inverse cdf (cumulative distribution function) of χ^2 distribution.
   - α: significance level ($\alpha = 0.6827$ for 1-sigma)
   - n_θ: number of free parameters.
   - In python, you can do `delta = scipy.stats.chi2.ppf(0.6827, n_param)`
3. These are the models "within 1-sigma level confidence interval."
4. Get the min/max of each of the parameters and set these as lower/upper limit of the parameters.
5. The *center* of the parameters can be obtained by simple maximum likelihood estimation, such as least-square fitting.

For the 1-sigma contour of 2 parameters, $\chi^2 < \chi^2_{\rm min} + 2.30$. The marginalized pdfs is usually drawn together to grasp the distribution of the parameters. If you have used chi-square statistic, you can use the fact that $P \propto e^{-\chi^2/2}$. Define $\bar{P} = e^{-\chi^2/2}$. Then the normalization constant will be $A = \sum_{parameters} \bar{P}$, i.e., you can use $P = \bar{P}/A$ as the normalized probability values.

In this example, because the authors mentioned that the w parameter is obtained from the Hα fitting, I just fixed the w value as the best fit value. From zooming in the oringinal paper, I couldn't find any difference from theirs to ours.

In the figure, the green vertical lines:

- dashed = our best fit central wavelength
- dotted = the uncertainty range from the paper's redshift uncertainty measured from Hα line ($1742 \pm 19\,$km/s)
- Because their uncertainty is from Hα, which has signal much better than O^{2+}, the error-bar is much smaller than ours.

and the blue horizontal lines:

- dashed = paper's flux ($10^{-15.7}\,$mW/m^2) converted to amplitude
- dotted = paper's uncertainty around the paper's value.
- The difference between our fit and paper's value is only 0.03 dex (maybe the authors obtained the identical value but just dropped the significance numbers)
- Error bars seem slightly understimated, but in reality if we use error bar of 0.14 dex in log scale, it's similar to ours. Maybe the authors just did not care about such detailed numbers, which is understandable.

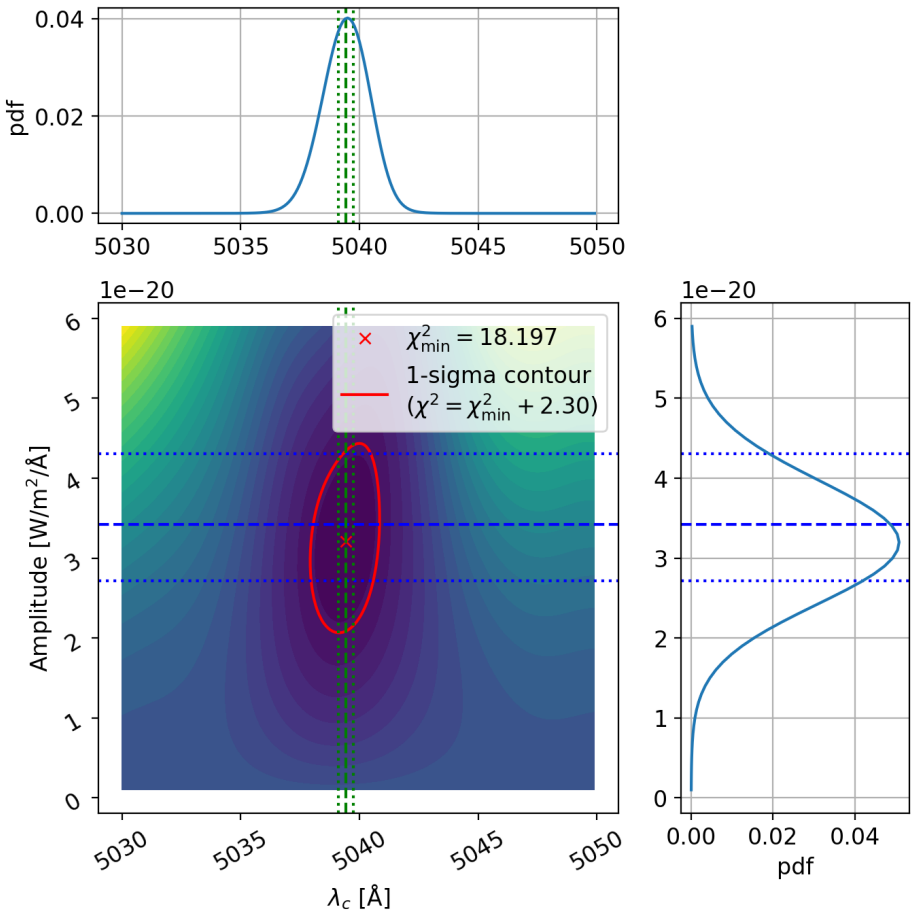

Figure 5.2: Our fitting of Gaussian curve to the emission line. The 2-D contour shows the χ^2 map in the 2-D parameter space. The 1-D plots show marginalized pdf from $P = e^{-\chi^2/2}/\bar{P}$.

Notes

The original paper argue that they used gaussian with flat spectrum*. When I tried this, however, the best fit constant is `-1.36e-21` , which must be a marginally visible negative shift to the fitted Gaussian, but I cannot find this by zooming in the paper's plot. Thus, I guess they may have just ignored this flat offset or set a bound that this constant must be positive, so that the best fit value is just 0.

5.3.2 Markov Chain Monte Carlo (MCMC)

Markov chain is a fancy name of "memoryless":

$$p(x_5|x_4, x_3, x_2, x_1) = p(x_5|x_4) \tag{5.12}$$

The Monte Carlo (MC)† came from the casino "Monte Carlo" in Monaco. Stanislaw Ulam and John von Neumann, in 1940, wanted to find best-fit parameters to make a better nuclear weapon at the Los Alamos

*From the paper: "When fitting the [O III] λ5007 line, we assume a flat continuum plus a single Gaussian line profile with standard deviation given by the Hα fit."

†Historical background adopted from KalinovaV's lecture note 2017 Feb.

National Laboratory. They arrived at the idea of (now-called) Monte Carlo, but wanted it to be secret to enemies, so they chose the secret name MC, where Ulam's uncle used to play gamble.

When you have, e.g., 7 parameters to fit 100 data points, and if the model is too complicated, the usual grid searching in the previous section is computationally impossible. Therefore, the "hopping" in the N-dimensional parameter space is suggested and that's MC. Currently I don't have complete plan to cover MCMC in AO class, but you may refer to many available packages from websites*. I recommend you to try `emcee` (although I used `pymc3`, but I feel `emcee` is more standard and safe as it's classical. `pymc3` is too much a black box to my eyes, and my friends kind of agreed).

*A great compilation is available here

Chapter 6

CCD and Detector Parameters

In this chapter, I will deal with some miscellaneous but important topics about CCD or other types of detectors (e.g., there is no available CCD in infrared wavelength, so technically they're not CCD).

6.1 Calibration Frames

There are mostly three types of calibration frames in observational astronomy: bias, dark, and flat. Others inlcude comparison arc lamp images, various types of flat (dome, sky, etc). In this section, I referred HowellSB (2006) "Handbook of CCD Astronomy" 2e, chapter 4.

6.1.1 Detector Readout

To understand why we need the bias, which is a technical than scientific reason, we need to know how the detector is being read out, so I separated this part from the section 6.1.3. When a CCD is being read out*, the electronics need to convert the number of (photo-)electrons into a digital unit (so-called ADU or DN), where the number of ADU is, by the definition, number of electrons divided by the electron gain.

But how can it count the number of electrons? Consider we have a voltmeter, attached to the 'measuring position' (the output gate). By the clock pattern of CCD, the **correlated double sampling** (CDS) process is done as follows:

1. Reset the output gate's voltage: $V_{\mathrm{out},1} = V_{\mathrm{reset}}$.
2. Measure V_{out}.
3. Dump the electrons to the output gate, so the voltage will change by V_{e}.
4. Measure the output gate's voltage again $V_{\mathrm{out},2}$ (which is now $V_{\mathrm{reset}} + V_{\mathrm{e}}$)
5. Get $\Delta V = V_{\mathrm{out},2} - V_{\mathrm{out},1} = V_{\mathrm{e}}$ (trivially the number of electrons is linearly related with this)
6. Iterate this again and again until all the pixels are read out.

The V_{e} is then converted to ADU or DN by the Analog-to-Digital Converter (ADC).

What I want to emphasize here is that the processes number 1 (resetting) and 4 (dumping) are not perfect, so the voltage over time ($V(t)$) graph is never a combination of linear lines as in fig. 6.1. The resetting and dumping involves movement of electrons, which results in a damped oscillation in the $V(t)$

*It is assumed you know how the CCD or photodetectors transfer electrons; if not, search for, e.g., CCD clock pattern.

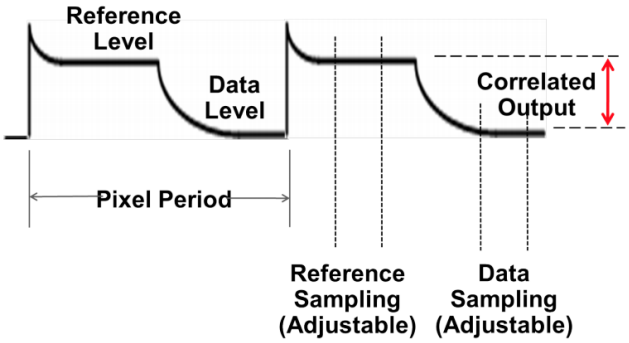

Figure 6.1: CCD output and correlated doubling sampling. The vertical axis is the voltage and horizontal axis is time. You can see the voltage is not a combination of linear line segments, but noisy curves. (Image from Texas Instruments, 2018: ti.com/lit/an/snaa322/snaa322.pdf)

graph. If we had plenty of time, we can just wait until it stablized; however, we need to read out pixels as soon as possible (consider reading out 1000 by 1000 pixel CCD: to read out it in 1 sec, we need to read each pixel in 1 μs!). If it takes long time, you may lose observation time, and the dark current at the first-read pixel and the last-read pixel may differ, etc. Therefore, it is impossible to avoid the *readout noise*. For some CCDs, there are options you can choose to compromise these readout time and readout noise.

6.1.2 Signed and Unsigned `int`

In computer science, there are two ways to express integers: signed and unsigned. For instance, an 8-bit integer of number 3 can be expressed as `0000 0011`. How about negative 3? `1000 0011`. That means, the first part of the 8-bit is used as *sign*. From this sense, an 8-bit integer can express from $(-2^7) = -128$ to $(+2^7 - 1) = +127$. But in astronomical science, we know there should be no negative pixel value, and therefore there is absolutely no need to use the first bit as sign bit. Then we can store the numbers from 0 to $2^8 - 1 = 255$, called the *unsigned* integer format. Thus we could express almost double amount of data (twice the resolution) without losing any storage.

Although astronomers use `BZERO` and `BSCALE` technique in the FITS file format in reality, the idea is similar to unsigned integer. The output images are frequently **assumed to have no negative pixel values**. This is one of the reasons why we need the bias frame, explained a bit more in section 6.1.3.

6.1.3 Bias

Def 10 (*Bias*)

A **bias frame** is ideally a frame with neither source signal nor thermal electrons. In practice, it is obtained with 0-sec (or the shortest possible) exposure time with shutter closed.

Now you know we have to go with readout noise; we cannot remove it. As technology improves, the readout noise R decreased down to only few electrons per readout. Early days when CCD was first used in astronomy, it was at least tens of electrons (e.g., $R \sim 50\,\mathrm{e}$, and it's still true in infra-red detectors), or up to hundreds of electrons. Since gain is $\sim 1{-}10\,\mathrm{e/ADU}$, that means each pixel might have at least $\pm$ tens of ADU uncertainty. Thus, even if you take a bias frame, which must have nearly a constant value

b, you may have a Gaussian distribution of, e.g., $\mathcal{N}(b, 10^2)$.

Consider a 1K by 1K CCD with bias follows Gaussian of $\mathcal{N}(b, 10^2)$. You have 10^6 pixels, and if you observe for 1 year, you may have taken 10^5 to 10^6 frames, which translates into the number of pixels read out per year of 10^{10} to 10^{11} pixels. If the readout was a perfect Gaussian, about one pixel of pixel value $v = -7\sigma = b - 70$ should have been detected at least once. Probabilistically speaking, even $b - 100$ may happen once in our lives. Thus, for all pixels to have non-negative value, we need to have b value of at least around 100 or larger. Therefore, most detectors have bias level of around 100 to 1000 or even more, depending on the readout noise.

You may ask that, given the full capacity of the usual detectors (16-bit `int`) is $2^{16} - 1 = 65535$ ADU, so we need to take as small bias value as possible to achieve as large dynamic range (DR) as possible. You may have seen a detector such as $R \sim 3\,\text{ADU}$ while $b \sim 1000\,\text{ADU}$. It seems $b \sim 100\,\text{ADU}$ is more than enough and about 900 ADU capacity (DR) is wasted. In practice, however, we are not losing the DR at all. The CCD anyway loses linearity or saturate at around 40k (40,000 ADU), i.e., anyway you should carefully use any CCD pixels which have pixel values higher than around 40k after bias subtraction. In principle, therefore, it's even possible to have a bias of 10,000 ADU without losing any scientific performance.

A short note: bias is also temperature dependent due to the electronics performance, as well as dark.

6.1.4 Dark

Def 11 (*Dark*)

A **dark frame** is ideally a frame with no source signal but with thermal electrons. In practice, it is obtained with preset exposure time with shutter closed.

In the potential well made by the detector electronics, some electrons are not *photo*electrons, but `thermal`electrons. The former is the electrons pumped up from the valence to conduction band due to the signal photons, while the latter is those which obtained enough kinetic energy from random thermal energy distribution of electrons according to quantum mechanics. Simply put, probabilistically, few electrons may have obtained enough thermal energy to be detached from any parent atom/molecule of the electronics, and captured in the potential well after the detachment. The thermal energy is a sensitive (exponential) function of temperature, so this probability is dependent of the electronics temperature.

In reality, there are at least 5 differnt sources of dark current which have different functional forms*. Among them, the most widely cited source of the dark is ther *backside dark current*, which is the dominant one for backside-illuminated CCDs:

$$d(T) \propto P_S T^{1.5} \exp\left[-\frac{E_g(T)}{2k_B}\left(\frac{1}{T} - \frac{1}{300\,\text{K}}\right)\right] \quad [\text{electrons/sec/pixel}] \tag{6.1}$$

where P_S is the pixel size, T is the tempareture, and

$$E_g(T) = a - \frac{bT^2}{c + T} \tag{6.2}$$

for $a = 1.1557$, $b = 7.021 \times 10^{-4}$, and $c = 1108$ are the experimentally determined constants. The dark current is very sensitive to temperature, and gets enormously large as temperature increases.

*See JanesickJR (2001) "Scientific Charge-Coupled Devices" chapter 7 and PankoveJ (1971) "Optical Processes in Semiconductors" p. 27.

In real observations, therefore, it is highly recommended to (1) reduce the temperature at least much lower than the room temperature and (2) try to maintain the temperature constant throughout the observation (otherwise, dark current may significantly vary).

6.1.5 Flat

Def 12 (*Flat*)

A **flat frame** is ideally a frame with both the flat signal with thermal electrons. The flat signal means a spatially (and spectrally/polarimetrically for spectroscopic/polarimetric observation) homogeneous, high signal-to-noise ratio signal. In practice, it is obtained with flat lamp and dome flat, twilight/dawn flat, and/or dark night flat.

Flat is not an easy topic: "To CCD experts, the term "flat field" can cause shivers to run up and down their spine*". The pixel-to-pixel non-uniformity of quantum efficiency at the given wavelength is not necessarily identical, so flat frame is obtained to remove this variation.

The flat field *in photometry* is composed of mainly the following two components:

1. **large-scale variation**: Even though the telescope is uniformly illuminated, vignetting can occur because of the unavoidable pixel-scale variation, i.e., distortion, throughout the field of view.
2. **small-scale variation**: The vignetting effect within, e.g., 10 by 10 pixels may be negligible. But each pixel can have different sensitivity, so local variation may appear.

The large scale variation could be extracted even from the object frame by using large median or boxcar kernel (it has been used since IRAF: see `ILLUMCOR`). This is sometimes needed because the variation pattern may differ from flat to object, caused by a slight movement of loosely fixed optics during the motion of the telescope.

Especially for the small-scale variation, we need to think the difference in the spectral response of each pixel. Consider two nearby pixels, where vignetting is ignored. Say the filter profile is $f(\lambda)$, i-th pixel has the spectral response $r_i(\lambda)$, flat and object have the spectral energy distribution $s_f(\lambda)$ and $s_o(\lambda)$, respectively. Then what we have from flat and the object frame at the i-th pixel are

$$F_i = \int_0^\infty s_f(\lambda) r_i(\lambda) f(\lambda) d\lambda \quad ; \quad I_i = \int_0^\infty s_o(\lambda) r_i(\lambda) f(\lambda) d\lambda \ . \tag{6.3}$$

By the flat field correction, I_i/F_i, what we are *hoping* is that the integrand function $\tilde{f}(\lambda)$ is identical for all the pixels:

$$I_i/F_i = \int_0^\infty s_o(\lambda) \tilde{f}(\lambda) d\lambda \quad ; \quad \tilde{f}(\lambda) = \frac{r_i(\lambda) f(\lambda)}{F_i} = \frac{r_i(\lambda) f(\lambda)}{\int_0^\infty s_f(\lambda) r_i(\lambda) f(\lambda) d\lambda} \tag{6.4}$$

This is not always true (because the $r_i(\lambda)$ do not simply cancel out), and it is a reason from the technological point of view why the instrument-dependent correction term, the transformation coefficient (k_f; see eq. (4.2)) appears, and making the spectroscopic flat very difficult to achieve.

One lesson from this is that, it *might be* risky to use a monochromatic flat (e.g., an LED), but even a spectrally flat lamp may not be a perfect flat, unless our flat has the spectrum identical to the (unknown) target object. In spectroscopy, however, a spectrally flat lamp is important to obtain good signal-to-noise ratio for all wavelength pixels.

*HowellSB (2006) "Handbook of CCD Astronomy" 2e, p.67.

6.2 Gain and Readout Noise

Now that you are familiar with preprocessing and data reduction. In the error-analysis, you may have used the gain and readout noise to estimate the pixel noise. But how can we determine the gain and readout noise? Frequently both of them are *provided from the CCD manufacturer*, but sometimes the user has to determine them. There are few ways to determine those two.

6.2.1 Gain and Readout Noise in FITS Header

The two parameters usually appear in the FITS header. Gain appears as the keyword `GAIN`, but many times people use the keyword `EGAIN`, which is not preferred. The readout noise is also called the read noise in short, and appear as `RDNOISE`. Sometimes `RONOISE` is used, but not preferred.

6.2.2 Janesick's Method

Janesick's method is the most classical way of deriving the gain and readout noise value. Although it's the most widely used in many textbooks, they mostly don't provide even simple ideas of proof, I here provide the full proof as well as the formulae.

Thm 16 (*Janesick's Method*)

If the two flat images have pixel values of F_1 and F_2 and two biases have B_1 and B_2 (all in ADU), the gain and readout noise are

$$g = \frac{(\bar{F}_1 + \bar{F}_2) - (\bar{B}_1 + \bar{B}_2)}{\sigma^2_{F_1-F_2} - \sigma^2_{B_1-B_2}} \text{ [e/ADU]} \quad ; \quad R = g\frac{\sigma_{B_1-B_2}}{\sqrt{2}} \text{ [e]} \tag{6.5}$$

Here $\bar{X}$ means the average of all the pixels in the frame X, and σ_X is the true standard deviation of the frame X, estimated from the sample standard deviation $\sigma_X \approx \sqrt{(\sum_i (X_i - \bar{X})^2)/(N-1)}$.

Proof of Janesick's Method : Note that in this theorem, by saying *flat*, we are implicitly assuming that those frames should share the identical expected values. Also the pixel-wise sensitivity variation is ignored, as well as cosmic-ray events, bad pixels, etc.

The bias frame is nothing but the offset voltage added with readout noise. Therefore, if the true bias level is b in ADU, any bias image will follow a normal distribution:

$$B \sim \mathcal{N}\left(b, \left(\frac{R}{g}\right)^2\right) \text{ [ADU]} \quad \rightarrow \quad B_1 - B_2 \sim \mathcal{N}\left(0, 2\left(\frac{R}{g}\right)^2\right) \text{ [ADU]}$$

where g is introduced in the denominator to convert R, in [e], to [ADU]. Hence,

$$\sigma^2_{B_1-B_2} = 2\frac{R^2}{g^2} \ ,$$

so the second equation is proven. Since the LHS is approximated by the sample standard deviation, you can write $\sigma^2_{B_1-B_2} \approx$ `np.std(B1-B2, ddof=1)**2` in python.

The (raw) flat image consist of photons with dark plus bias level. Therefore, if f is the true flat level *plus dark* in ADU, any flat will roughly follow a normal distribution, similar to the bias case:

$$F \dot{\sim} \mathcal{N}\left(f + b, \frac{f}{g} + \left(\frac{R}{g}\right)^2\right) \text{ [ADU]} \quad \rightarrow \quad \begin{cases} F_1 - F_2 \dot{\sim} \mathcal{N}\left(0, 2\frac{f}{g} + 2\left(\frac{R}{g}\right)^2\right) & \text{[ADU]} \\ (F_1 + F_2) - (B_1 + B_2) \dot{\sim} \mathcal{N}\left(2f, 2\frac{f}{g} + 4\left(\frac{R}{g}\right)^2\right) & \text{[ADU]} \end{cases}$$

The first term in the variance is the Poisson noise term* and the second term is the readoise term, respectively. From these, you can extract

$$\sigma^2_{F_1-F_2} - \sigma^2_{B_1-B_2} \approx 2\frac{f}{g}$$
$$(\bar{F}_1 + \bar{F}_2) - (\bar{B}_1 + \bar{B}_2) \approx 2f$$

This proves the first equation. Since the σ values are approximated by the sample standard deviation, $\sigma^2_{F_1-F_2} - \sigma^2_{B_1-B_2} \approx$ `np.std(F1-F2, ddof=1)**2 - np.std(B1-B2, ddof=1)**2` in python, and from simple mathematics, $(\bar{F}_1 + \bar{F}_2) - (\bar{B}_1 + \bar{B}_2) =$ `np.mean((F1+F2) - (B1+B2))` in python.

Although we can use any one flat and one bias out of two of each to obtain $F_i - B_i \dot{\sim} \mathcal{N}(f, f/g + 2(R/g)^2)$, Janesick's method uses all the information from all the four frames at once. In real application, we may take a lot of bias and flat frames, say N_b and N_f frames, respectively. Then select 2 from each, making $\binom{N_b}{2} \times \binom{N_f}{2}$ possible gain and readout noise estimations. We can use the mean of those results to estimate the gain and readout noise and their uncertainties by sample standard deviation of the estimates. Because there always are vignetting in flat frames, you **must extract only the smooth part of flat** (and thus the corresponding region in bias) for this analysis to meet the assumptions given in the beginning of the proof, and should not naïvely use all the pixels in the image.

Also note that, in real physical unit system, both e/ADU and e are unitless, so don't be confused if you see something like $\sqrt{R^2/g}$ has the unit of ADU, not $\sqrt{\text{e} \cdot \text{ADU}}$.

6.2.3 Graphical Method (Line Fitting)

There is another method, which has no name as far as I know. It fits a linear line to some values. **This method is preferred over Janesick's method**, because (1) it's a linear regression so you can simply estimate the uncertainty of g and R from simple statistics, (2) using the identical data obtained from this method, you can check the linearity of the detector.

Similar to Janesick's method, consider you have a master bias B, and assume the flat frame has no sensitivity variation in the pixels we select. Take $N_f^{(i)} (\geq 2)$ flat images at each exposure time t_i, $t_i \in \{t_1, t_2, \cdots, t_N\}$ $(t_1 < t_2 < \cdots < t_N)$, so that the expected value of each exposure time will be $I_0 t_i$.

Thm 17 (*Gain and Readnoise Determination - Graphical (not recommended)*)

For each exposure, select 2 flats among $N_f^{(i)}$ flat frames, and calculate the variance of the difference between two frames, $\sigma^2\left(F_j^{(i)} - F_k^{(i)}\right)$ $(j \neq k)$, so that you have $\binom{N_f^{(i)}}{2}$ values. Then

$$\sigma^2\left(F_j^{(i)} - F_k^{(i)}\right) = 2\left[\frac{1}{g}f^{(i)} + \left(\frac{R}{g}\right)^2\right] \quad \rightarrow \quad \begin{cases} g = \dfrac{2}{\text{slope}} & [\text{e/ADU}] \\ R = g\sqrt{\text{intercept}}/2 & [\text{e}] \end{cases} \tag{6.6}$$

where $f^{(i)} \approx \text{mean}\left(\sum_j (F_j^{(i)} - B)\right)$ is the true flat *plus dark* in ADU. There will be $N_f^{(i)}$ points for the fixed $f^{(i)}$ value on the x-axis, and the number of x values is N.

*Reminder: the Poisson noise term is also called photon noise or shot noise in astronomy. Both photoelectron from true flat and dark current follow Poisson distribution. The Poisson distribution is usually assumed to be Gaussian, because $f/g \gg 1$ (Poisson distribution asymptotically approaches to Gaussian when the mean value gets larger, as we saw in Thm **??**).

There is actually no need to prove this: from Thm 16, we already derived $F_1 - F_2 \dot{\sim} \mathcal{N}(f, 2f/g + 2(R/g)^2)$, thus the above equation is proven. The assumptions used are that all flats share same bias (true bias b is estimated from the master bias B), all flats share identical gain and readout noise, and that mean$\left(\sum_j (F_j^{(i)} - B)\right)$ can be used as the estimator for the true signal in ADU. One may think some extended version of assumptions such as the variation of the detector's temperature is ignorable, etc. All these assumptions are too trivial for well-established instruments, so manytimes we just do not explicitly state these.

Usually the error-bar of $f^{(i)}$ will be very small because you will use tens of thousands of pixels to estimate this, and the CLT (Thm 1) decreases the error-bar. There is no need to spend long time to accurately derive the error-bar in the y-direction either, because this kind of *performance evaluation* is a one-time task per year or so, and hence you can just plot a huge number of y values rather than thinking about error-bars of each point. Let's just ignore error-bars and proceed the line fitting.

As mentioned before, the data used in this analysis can be **recycled for the linearity check**. We can simply plot $f^{(i)}$ as a function of t_i, and you should see a linear line. Depending on the dynamic range of the detector, the linearity will break at some exposure $t_i > t_{i_{\text{crit}}}$. The $F^{(i_{\text{crit}})} = f^{(i_{\text{crit}})} + b$ will then be the maximum ADU value you should trust. Any pixel value higher than that must be cautiously dealt when you do the data analysis.

To increase the number of data points in linear regression, you need to take too many flats. A better way is to chop each frame into mesh (e.g., 10 by 10 pixel meshes), and use each such region as one data point. If you have 1 flat at 5 exposures ($N_f^{(i)} = 1$ for $i = 1, .., 5$) and flat is 1,000 by 1,000 pixels, you can make 10,000 meshes per each exposure! Therefore, a simpler method is :

Thm 18 (*Gain and Readnoise Determination - Graphical (recommended)*)

For each flat, subtract bias and chop it into many meshes. For each mesh, calculate pixel mean and variances. If the incident flux is assumed to be homogeneous in this small mesh area, variance = mean$/g + (R/g)^2$.

Fit a line for (x, y) where x: mean$(= f =$ flat $-$ bias$)$ and y: variance for all meshes and exposures. Then

$$\begin{cases} g = \dfrac{1}{\text{slope}} & \text{[e/ADU]} \\ R = g\sqrt{\text{intercept}} & \text{[e]} \end{cases} \tag{6.7}$$

Note the factor 2 is not appearing in this method.

6.2.4 Note

A good thing for the above two methods are that they are independent of the spectral energy distribution (SED) of the flat. This is because both the gain and readout noise are device's output-gate-related, not related to the pixel's sensitivity or whatever. Thus, important things are the pixel count and mathematics (statistics), which means you can calculate gain and readout noise even with a monochromatic laser, for example. In real observation, it's better not to use monochromatic light source for the flat (see section 6.1.5).

Chapter 7

Growth Curve

NOTE: This part is heavily based on `StetsonPB 1990 PASP 102 932`.

Selecting an appropriate aperture radius for circular PSF photometry is not an easy task. Since 1960s, it has been extensively studied including Moffat, and it was blooming in 1990s thanks to many pioneers who tried to combine the analytical description of stellar profile and empirical counterparts. Here I will summarize the most powerful technique called the **growth curve (GC) analysis**, which maybe out of scope for undergraduate courses, but conceptually easy.

7.1 Introduction

In photometry, we need to set the aperture as large as possible if we want to collect all the flux from the star. But wait, how large is large? It should be "sufficiently large that seeing, tracking, and focus errors do not affect the fraction of the star's flux which falls *outside* the aperture. $\cdots$ The computer-defined "aperture" within which the pixel values are summed in the two-dimensional data array is a direct analog to the physical aperture placed at the telescope focus in a standard photometer." (from Stetson 1990).

Ststson further describes "$\cdots$ to measure as many stars as possible through a series of *several* concentric apertures of increasing radius and to calculate the observed magnitude differences between successive apertures for each star. These are plotted as a function of radius, and a smooth curve is sketched through them to yield the average "growth curve" of the frame. The average magnitude differences between successive apertures are then read from this curve and summed from the outside inward to yield cumulative corrections from each of the smaller apertures to the system of the largest. This multiple-aperture technique offers several advantages over the simpler two-aperture method." (Stetson 1990) The earliest description about the growth curve analysis is, Stetson notes, `Rich+ 1984, ApJ, 286, 517`.

In short, what we do in GC analysis is like this. First, assume the PSF is circular. If i is the frame number (i-th CCD frame) and j is the star id (j-th star):

1. Find the center of the star and fix it ($\mathbf{r}_{0,i,j}$)
2. Set the sky annulus for the frame (say inner/outer radii $r_{\mathrm{in},i}$ and $r_{\mathrm{out},i}$, respectively, and you may fix it identical for all stars if you want).
3. Do aperture photometry for many r_k, e.g., $r_k = 1, 2, \cdots, r_{\mathrm{in}}$ pixels.
4. Plot the sky-subtracted delta-magnitude $\delta_{i,j,k} = m_{i,j}(r_k) - m_{i,j}(r_{k-1})$ as a function of $(r_k + r_{k-1})/2$ for many bright stars.

5. Find a smooth function which describes this curve (discussed below).
6. Find, e.g., Δ value such that True mag $= m(r=\infty) = m(r=r_k) - \Delta(r_k)$.
7. Do photometry for j-th star with $r_{k'}$ (where the SNR gets maximum), and get the true magnitude by $\Delta(r_{k'})$.

7.2 Stellar Profile

We have discussed the stellar profiles like Gaussian and Moffat. Stetson 1990 describes a more general stellar profile*:

$$I(r, X_i|R_i) = [a + bX_i]M(r|R,\beta) + [1 - (a + bX_i)]\left\{cG(r|\sigma_i) + [1 - c]H(r|s\sigma_i)\right\} \ .$$

Here, using the notations in eqs. (2.14) and (2.18),

$$\begin{aligned} M(r|R,\beta) = f_{\rm Moffat}(r|I=1,R,\beta) &= \frac{\beta-1}{\pi R^2}\left[1+\left(\frac{r}{R}\right)^2\right]^{-A} \ , \\ G(r|\sigma_i) = f_{\rm Gauss}(r|I=1,\sigma_i) &= \frac{1}{2\pi\sigma_i^2}e^{-r^2/2\sigma_i^2} \ , \\ H(r|r_0) = \frac{1}{2\pi r_0^2}e^{-r/r_0} \ . & \end{aligned}$$

Actually there is no reason to introduce $s\sigma_i$ instead of a totally new parameter, say $\tilde{s}$, but Stetson used $s\sigma_i$. These functions have simple integral form:

$$\begin{aligned} M_I(r_k|R,\beta) = \int_0^{r_k} M(r|I=1,R,\beta)(2\pi r)dr &= 1 - \frac{1}{(1+(r_k/R)^2)^{\beta-1}} \ , \\ G_I(r_k|\sigma) = \int_0^{r_k} G(r|I=1,\sigma_i)(2\pi r)dr &= 1 - e^{-r_k^2/2\sigma_i^2} \ , \\ H_I(r_k|r_0) = \int_0^{r_k} H(r|I=1,r_0)(2\pi r)dr &= 1 - \left[1+\frac{r_k}{r_0}\right]e^{-r_k/r_0} \ , \end{aligned} \tag{7.1}$$

and all have integral of unity if $r_k \to \infty$ (As noted with eq. (2.18), $A > 1$ is required). The GC is basically an azimuthally averaged profile, so the integrated version of I is required:

$$I_I(r, X_i) = [a + bX_i]M_I(r|R,\beta) + [1 - (a + bX_i)]\left\{cG_I(r|\sigma_i) + [1 - c]H_I(r|s\sigma_i)\right\} \ . \tag{7.2}$$

*The original notation was $I(r, X_i; R_i, A, B, C, D, E) = (B+EX_i)M(r;A)+(1-B-EX_i)[CG(r;R_i)+(1-C)H(r;DR_i)]$, i.e., $R_i = \sigma_i$, $A = \beta$, $B = a$, $C = c$, $D = s$, $E = b$. I made a more generalization by adding the core width R, which was set to 1 in Stetson's paper.