2018 AO 1
TA Lecture Note

(Statistics + Ground Based Aperture Photometry)

Yoonsoo P. Bach

`https://ysbach.github.io/AO_LectureNotes`

Statistics

What does m = 14 +- 0.1 mean?

① $m_1 = 14 \pm 0.1$
$m_2 = 15 \pm 1.0$

② $m_1 = 14 \pm 0.1$
$m_2 = 15 \pm 0.8$

$m_1 \stackrel{?}{=} m_2$

Cannot $m_1 \neq m_2$ (90%)

③ $m_1 = 14 \pm 0.1$
$m_2 = 15 \pm 0.3$

Can $m_1 \neq m_2$

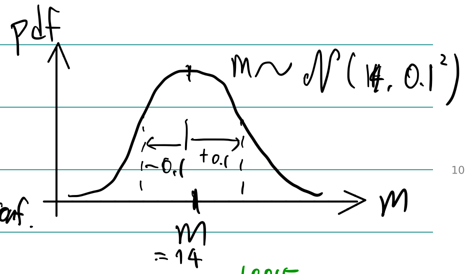
Basic Knowledge

※ $m \pm \Delta m$: Gaussian "normally"

Gaussian mean, Gaussian stddev.

pdf = probability distribution function
확률 분포 함수
確率分布函數

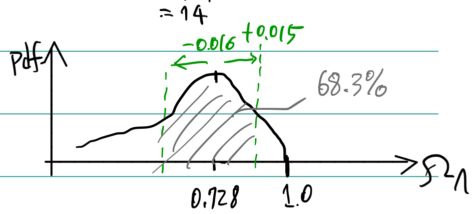
$14 \pm 0.1 \xrightarrow{3\sigma} 14 \pm 0.3$: 13.7 ~ 14.3 w/ 99.8% conf.

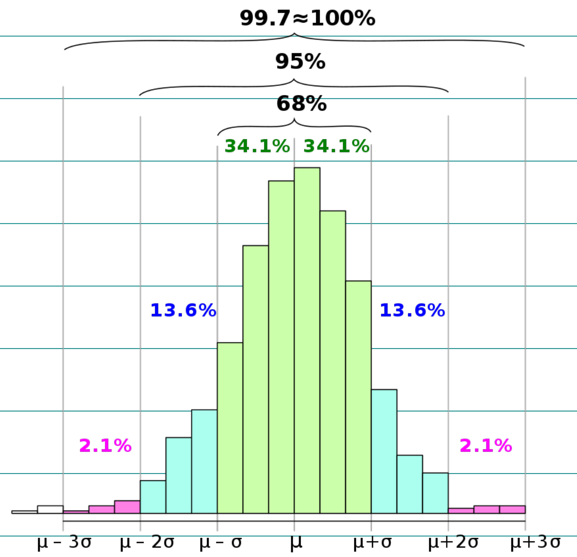


- Non-Gaussian Error $\Omega_\Lambda = 0.728^{+0.015}_{-0.016}$: 0.031 ✓

$0.728^{+0}_{-0.035}$ 0.035

cdf = cumulative d f
누적 분포 함수
累積分布函數



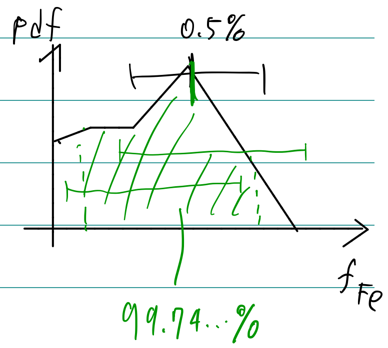


$$\Pr(\mu - \sigma \le X \le \mu + \sigma) \approx 0.6827$$
$$\Pr(\mu - 2\sigma \le X \le \mu + 2\sigma) \approx 0.9545$$
$$\Pr(\mu - 3\sigma \le X \le \mu + 3\sigma) \approx 0.9974$$

$f_{Fe} \simeq 0.5 \pm 0.2\%$

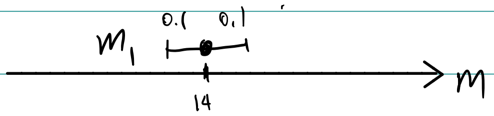
$3\sigma \rightarrow$ −0.1 — 1.1 %

↳ no range ≠ n × "1-σ"

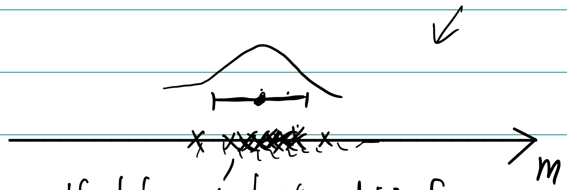


• What does it mean by $m = 14 \pm 0.1$ mag?

... 1st obs



means



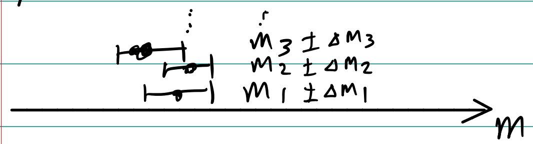
if obs'n was made at Δt [s] of exposure, we may have obtained these values as m.

central limit thm

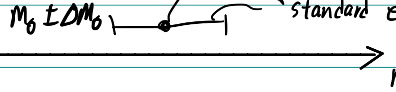
$$\text{CLT} : X \sim G(\mu, \sigma^2) \longrightarrow \bar{X}_N \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{N}\right) \Rightarrow$$ We can use Gaussian

↳ symmetric error

If we've made 100 obs'ns,



$\Longrightarrow$



weighted mean: Maximal Likelihood Estimator (MLE) of the true mean

$$m_0 = \frac{\sum_{i=1}^{100} m_i / \Delta m_i^2}{\sum_{i=1}^{100} (1/\Delta m_i^2)} \qquad \Delta m_0 = \sqrt{\frac{1}{\sum \Delta m_i^2 / (\Delta m_i^2)^2}} = \sqrt{\frac{1}{\sum_{i=1}^{100} \Delta m_i^{-2}}}$$

If $\Delta m_1 = \cdots = \Delta m_{100} \equiv \Delta m$ $\sim 0.5^{\text{mag}}$

$$m_0 = \frac{\sum m_i}{100} \qquad \Delta m_0 = \frac{\Delta m}{\sqrt{100}} = \underbrace{\frac{\Delta m}{10}}_{0.05^{\text{mag}}}$$

So $m = 14 \pm 0.1$ means

- 1 obs'n & $m_1 = 14$, $\delta m_1 = 0.1$

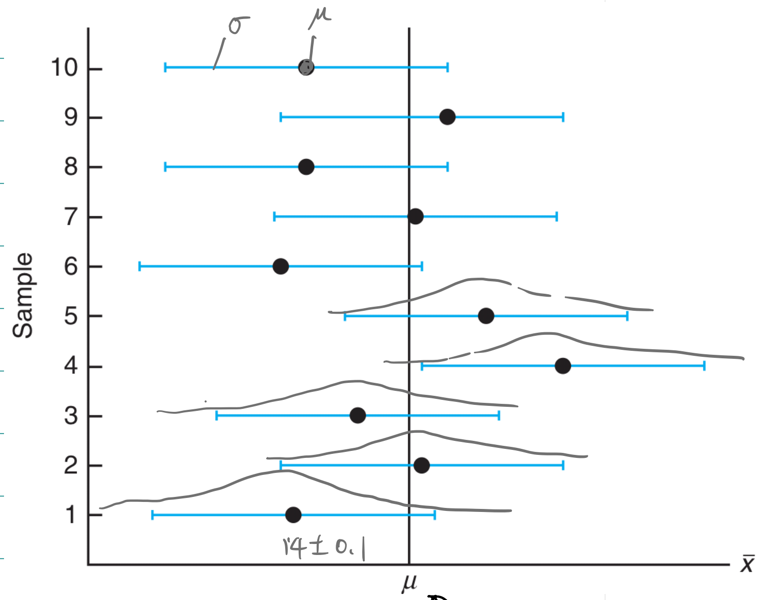
or

- 100 obs'n & $m_1 = 14.xx$, $\Delta m_1 = 1.0$
  - $m_2 = 15.xx$, $\Delta m_2 = 1.0$
  - $m_3 = 14.xx$, $\Delta m_3 = 1.0$
  - $\vdots$ $\vdots$

or • Many possibilities

Summary

$\Delta m_i = \Delta m$ (const)



$m_{100} \pm \Delta m$

$m_3 \pm \Delta m$

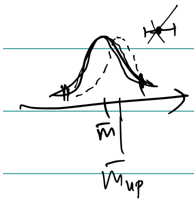
$m_2 \pm \Delta m$

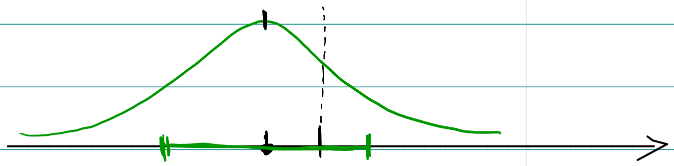
$m_1 \pm \Delta m$

obs mean → $\bar{m}$

true mean → μ

※ Assume no systematic error





$\mu \pm \sigma$

$\uparrow \quad \uparrow$

$\hat{\mu} \quad \hat{\sigma}$

$$\bar{m} \pm \underset{\hat{\sigma}\,?}{s.e} = \frac{1}{N}\sum m_i \pm \frac{\sigma}{\sqrt{N}}$$

Interpretation : μ w/i $\bar{m} \pm s.e$ w/ 68.27% probability NO!

⇒ Rather : ~~If more obs's are made ($m_{101} \pm \Delta m$, $m_{102} \pm \Delta m$, ……)~~
~~the updated $\bar{m}_{up}$ is w/i $\bar{m} \pm s.e$.~~

If the obs'ns are made under identical circumstances, (ensemble)
68.27% of the $\bar{m} \pm s.e$ intervals would contain μ.

※ Thanks to Sanghyuk Moon

In each universe in such ensemble, you will estimate the parameter p (e.g., mean) and its error k * δp (k is a constant which is a function of the significance level, and δp is a function of "sample" standard deviation or other sampling statistic). This p & δp can differ from universe to universe, although k is constant once the significance level is fixed.
Since the true value of p, say μ, is a fixed constant (but unknown to us), we cannot say that our universe which gave p_0 & δp_0 will have μ within p_0 - δp_0 ~ p_0 + δp_0.
But rather, we can say that (for the above example case)

"if the same strategy is applied to all the universes,
68.27% of such ensemble of universes will contain μ in I_i,
but saying that 'our universe will have μ within I_0 with 68.27% probability' is wrong."

Here, I_i is the interval of [p_i - δp_i , p_i + δp_i] for the i-th universe,
and i=0 is our universe.

5.13 Potential Misconceptions and Hazards; Relationship to Material in Other Chapters

The concept of a *large-sample confidence interval* on a population parameter is often confusing to the beginning student. It is based on the notion that even when σ is unknown and one is not convinced that the distribution being sampled is normal, a confidence interval on μ can be computed from

$$\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}.$$

In practice, this formula is often used when the sample is too small. The genesis of this large sample interval is, of course, the Central Limit Theorem (CLT), under which normality is not necessary in practice. Here the CLT requires a known σ, of which s is only an estimate. Thus, n must be at least as large as 30 and the underlying distribution must be close to symmetric, in which case the interval remains an approximation.

There are instances in which the appropriateness of the practical application of material in this chapter depends very much on the specific context. One very important illustration is the use of the t-distribution for the confidence interval on μ when σ is unknown. Strictly speaking, the use of the t-distribution requires that the distribution sampled from be normal. However, it is well known that any application of the t-distribution is reasonably insensitive (i.e., **robust**) to the normality assumption. This represents one of those fortunate situations which occur often in the field of statistics in which a basic assumption does not hold and yet "everything turns out all right!"

It is our experience that one of the most serious misuses of statistics in practice evolves from confusion about distinctions in the interpretation of the types of statistical intervals. Thus, the subsection in this chapter where differences among the three types of intervals are discussed is important. It is very likely that in practice the **confidence interval is heavily overused**. That is, it is used when there is really no interest in the mean; rather, the question is "Where is the next observation going to fall?" or often, more importantly, "Where is the large bulk of the distribution?" These are crucial questions that are not answered by computing an interval on the mean.

The interpretation of a confidence interval is often misunderstood. It is tempting to conclude that the parameter falls inside the interval with probability 0.95. A confidence interval 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. Any beginning student of practical statistics should be very clear on the difference among these statistical intervals.

Another potential serious misuse of statistics centers around the use of the χ^2-distribution for a confidence interval on a single variance. Again, normality of the distribution from which the sample is drawn is assumed. Unlike the use of the t-distribution, the use of the χ^2 test for this application is **not robust to the normality assumption** (i.e., the sampling distribution of $\frac{(n-1)S^2}{\sigma^2}$ deviates far from χ^2 if the underlying distribution is not normal).

H_0: null hypothesis
영 가설
零假說
H_1: AN alternative hypothesis
대립 가설
對立假說

- Returning to the question: $m_1 = 14 \pm 0.1$, $m_2 = 15 \pm 1.0$, 15 ± 0.8, 15 ± 0.3 ?

$H_0: m_1 = m_2 \Leftrightarrow m_1 - m_2 = 0$
$H_1: m_1 \neq m_2 \Leftrightarrow m_1 - m_2 \neq 0$

Derivations for each case is given in statistics textbooks.

notation
σ = population std dev
$s := \hat{\sigma}$ = sample std dev
t: Student's t distribution
v: degree of freedom for the t.

H_0	Value of Test Statistic	H_1	Critical Region
$\mu = \mu_0$	$z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$; σ known	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$z < -z_\alpha$ $z > z_\alpha$ $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
$\mu = \mu_0$	$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$; $v = n - 1$, σ unknown	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$z = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$; σ_1 and σ_2 known	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$z < -z_\alpha$ $z > z_\alpha$ $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$t = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p\sqrt{1/n_1 + 1/n_2}}$; $v = n_1 + n_2 - 2$, $\sigma_1 = \sigma_2$ but unknown, $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$t' = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$; $v = \dfrac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$, $\sigma_1 \neq \sigma_2$ and unknown	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$t' < -t_\alpha$ $t' > t_\alpha$ $t' < -t_{\alpha/2}$ or $t' > t_{\alpha/2}$
$\mu_D = d_0$ paired observations	$t = \dfrac{\bar{d} - d_0}{s_d/\sqrt{n}}$; $v = n - 1$	$\mu_D < d_0$ $\mu_D > d_0$ $\mu_D \neq d_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$

To answer the question, we need #(obs'ns) for each m_1 & m_2.

Say both are obtained by 5 obs'ns. $\bar{x}_1 = m_1$, $\bar{x}_2 = m_2$, $s_1 = 0.1$, $s_2 = 1.0$ or 0.8 or 0.3
$n_1 = 5$, $n_2 = 5$, $d_0 = 0$

$$t' = \frac{(m_1 - m_2) - 0}{\sqrt{(s_1^2 + s_2^2)/n}} = -2.225 \text{ or } -2.774 \text{ or } -7.071$$

s_2 1.0 $\quad s_2 = 0.8 \quad s_2 = 0.3$

$$v = \frac{\left(\frac{s_1^2 + s_2^2}{n}\right)^2}{\frac{s_1^4 + s_2^4}{n^2(n-1)}} = \frac{(s_1^2 + s_2^2)^2 (n-1)}{s_1^4 + s_2^4} = 4.09 \text{ or } 4.12 \text{ or } 4.88$$

$\xrightarrow{\text{integer}}$ 4 $\quad$ 4 $\quad$ 5

two-tail significance level α-test

$t^{[v]}_{\alpha/2}$

$\alpha = 0.20$	$\frac{\alpha}{2} = 0.10$	0.941	0.920
$\alpha = 0.10$	$\frac{\alpha}{2} = 0.05$	1.533	1.476
$\alpha = 0.05$	$\frac{\alpha}{2} = 0.025$	2.776	2.571
$\alpha = 0.02$	$\frac{\alpha}{2} = 0.01$	2.999	2.757
$\alpha = 0.01$	$\frac{\alpha}{2} = 0.005$	4.604	4.032
$\alpha = 0.001$	$\frac{\alpha}{2} = 0.0005$	8.610	6.869

Failed in rejection of H_0 at $\alpha \leq 0.05$: Cannot say $m_1 \neq m_2$ w/ 95%
You could reject H_0 at $\alpha > 0.05$: But you can // // 90%

Summary

- You could say

$$m_1 = 14 \pm 0.1 \neq m_2 = 15 \pm 0.3 \qquad (N_{obs,1} = N_{obs,2} = 5)$$

w/ 99+% confidence.

- You could say

$$m_1 = 14 \pm 0.1 \neq m_2 = 15 \pm 1.0 \qquad (N_{obs,1} = N_{obs,2} = 5)$$
$$\pm 0.8$$

but only w/ < 90% confidence.

- Although ≠ golden rule, people normally use $\alpha < 0.05$

※ P-value : The likelihood that your conclusion is wrong.
(ex 99.9 % confidence ⟷ P-value = 0.1 %)

- $\underbrace{H_0}_{m_1 = m_2}$ is rejected w/ significance level $\alpha = 0.05$: P-value < 0.05

↳ $m_1 \neq m_2$

$\boxed{\text{QUIZ}}$

Courtesy
prof. Jae-Kwang Kim
(KAIST Dept. Math)

There are two hospitals A/B. New babies are born (~50% girl/boy) every day:
A = 45 babies, B = 15 babies.
they recorded the number of days when 60+ % babies are girl. Which hospital will have more such days?
(+) With how much confidence you can say that?

Bin. A : $\sum_{k=27}^{N} \binom{N}{k} p^k q^{N-k} = \left(\frac{1}{2}\right)^N \sum \binom{N}{k} \sim 0.0899$ $(\times 365 = 32.8)$

$(60^{+}\% \geq 27)$

B : $\sum_{k=9}^{N} \binom{N}{k}$ " $= \left(\frac{1}{2}\right)^N \sum_{k=9}^{N} \binom{N}{k} \sim 0.304$ $(\times 365 = 111.0)$

$(60^{+}\% \geq 9)$

Formula $X_i \sim Bin(p)$ $X_i = \begin{cases} 1 & \text{when } 60+\% \text{ are girls} \\ 0 & \text{o.w} \end{cases}$ for the day i

$(i = 1, \cdots, 365)$

$E(X_i) = p$

$V(X_i) = E(X_i^2) - [E(X_i)]^2 = p - p^2 = p(1-p) \equiv pq$

$X = \Sigma X_i$

$E(X) = E(\Sigma X_i) = np$ $(n = 365)$

$V(X) = V(\Sigma X_i) = n V(X_i)$

$= npq$

A : $np = 32.8$ $\sqrt{V} = \sqrt{365 \cdot 0.0899 \cdot (1 - 0.0899)} \sim 5$

B : $np = 111$ $\sqrt{V} = \ldots$ $\sim$ few

- Poisson noise

3.5 Poisson Distribution and the Poisson Process

Experiments yielding numerical values of a random variable X, the number of outcomes occurring during a given time interval or in a specified region, are called **Poisson experiments**. The given time interval may be of any length, such as a minute, a day, a week, a month, or even a year. For example, a Poisson experiment can generate observations for the random variable X representing the number of telephone calls received per hour by an office, the number of days school is closed due to snow during the winter, or the number of games postponed due to rain during a baseball season. The specified region could be a line segment, an area, a volume, or perhaps a piece of material. In such instances, X might represent the number of field mice per acre, the number of bacteria in a given culture, or the number of typing errors per page. A Poisson experiment is derived from the **Poisson process** and possesses the following properties.

Properties of the Poisson Process

interval I, small interval dI

X = # of outcomes

1. $\#(\text{outcomes} \in I)$ is indep of $\#(\text{outcomes} \in I')$
2. $\text{Prob}[X=1 \in dI] \propto \text{length}(dI)$
3. $\text{Prob}[X \geq 2 \in dI \to 0] = 0$

1. The number of outcomes occurring in one time interval or specified region of space is independent of the number that occur in any other disjoint time interval or region. In this sense we say that the Poisson process has no memory.
2. The probability that a single outcome will occur during a very short time interval or in a small region is proportional to the length of the time interval or the size of the region and does not depend on the number of outcomes occurring outside this time interval or region.
3. The probability that more than one outcome will occur in such a short time interval or fall in such a small region is negligible.

The number X of outcomes occurring during a Poisson experiment is called a **Poisson random variable**, and its probability distribution is called the **Poisson distribution**. The mean number of outcomes is computed from $\mu = \lambda t$, where t is the specific "time," "distance," "area," or "volume" of interest. Since the probabilities depend on λ, the rate of occurrence of outcomes, we shall denote them by $p(x; \lambda t)$. The derivation of the formula for $p(x; \lambda t)$, based on the three properties of a Poisson process listed above, is beyond the scope of this book. The following formula is used for computing Poisson probabilities.

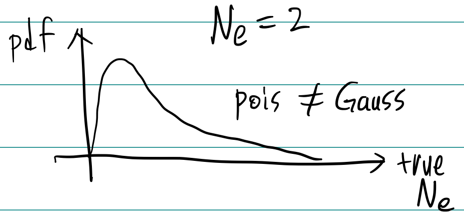
$$X \sim \text{Pois}(\lambda) \; : \; \mu = \sigma^2 \equiv \lambda$$

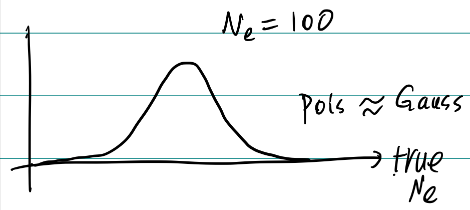
- Measurement of photoe^- (photon) ~ Poisson distribution.

- Ex) N_e e^-'s from a ✡

$\Rightarrow \mu = N_e \;\;\&\;\; \sigma = \sqrt{N_e}$

$\Rightarrow$ # of e^-'s from the ✡ $= N_e \pm \sqrt{N_e}$





Date 2018-04-02

Aperture Photometry

Consider you finished preprocessing. gain $= g$ [e/ADU]

ronoise $= N_{ro}^{(e)}$ [e]

[1] ● : centroid

unit = ADU

2	2	2	2
2	10	20	2
2	20	10	2
2	2	2	2

①

$$I_i = \sum_{j=-L}^{j=L} I_{i,j}$$

and

$$J_j = \sum_{i=-L}^{i=L} I_{i,j},$$

where $I_{i,j}$ is the intensity (in ADU) at each x, y pixel; the mean intensities are determined from

$$\bar{I} = \frac{1}{2L+1} \sum_{i=-L}^{i=L} I_i$$

and

$$\bar{J} = \frac{1}{2L+1} \sum_{j=-L}^{j=L} J_j;$$

and finally the intensity weighted centroid is determined using

$$x_c = \frac{\sum_{i=-L}^{i=L} (I_i - \bar{I}) x_i}{\sum_{i=-L}^{i=L} (I_i - \bar{I})}$$

for all $I_i - \bar{I} > 0$ and

$$y_c = \frac{\sum_{j=-L}^{j=L} (J_j - \bar{J}) y_j}{\sum_{j=-L}^{j=L} (J_j - \bar{J})}$$

for all $J_j - \bar{J} > 0$.

2	2	2	2
2	10	20	2
2	20	10	2
2	2	2	2

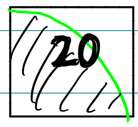
② Fit Gaussian 1D or 2D

③ Optionally you can do 3-σ clip.

(Ma+ '99, Optics Express, 17, 8525)

[2] Aperture sum?

$r_{ap} = 1$



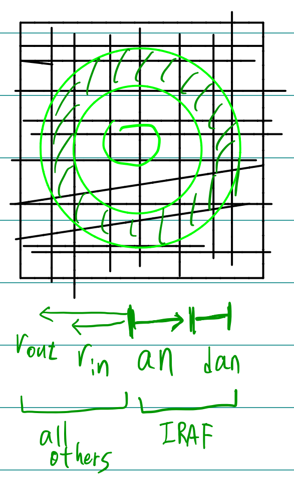
$20 \times \frac{◯}{□} = \frac{\pi/4}{1}$

$\Rightarrow N_{apsum} = \left(\frac{\pi/4}{1} \cdot 20 + \frac{\pi/4}{1} \cdot 10 \right) \times 2$

$= 15\pi$ [ADU]

$n_{pix} = \pi$ [pixel]

[3] sky?



$n_{sky}^{tot} = 105$

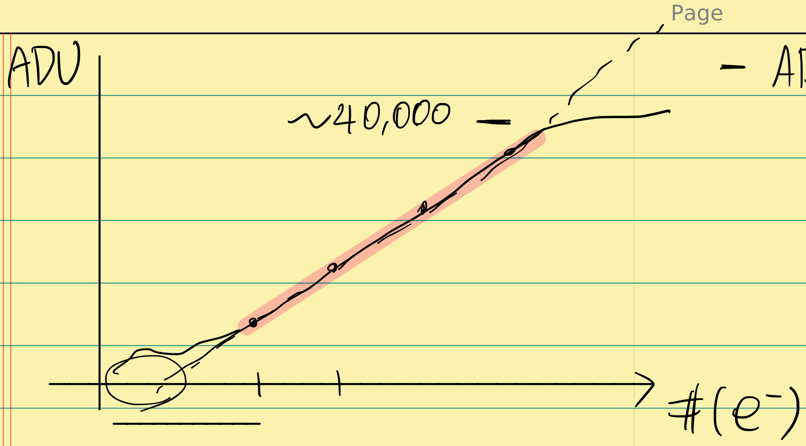
after 3-σ clipping, $n_{rej} = 5$ (rejected)

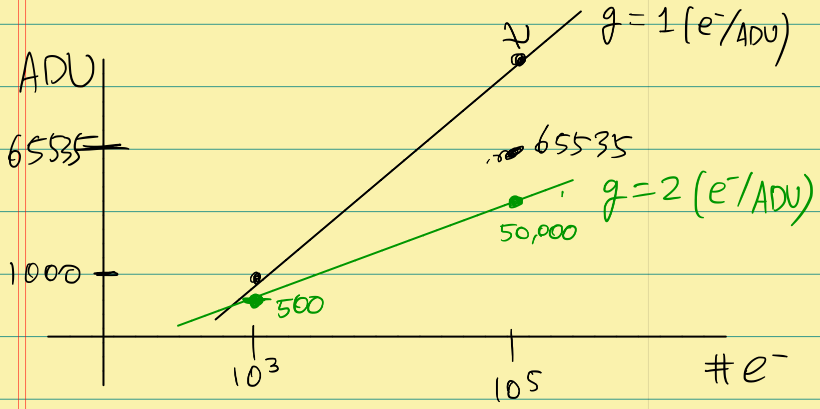
$n_{sky} = 100$

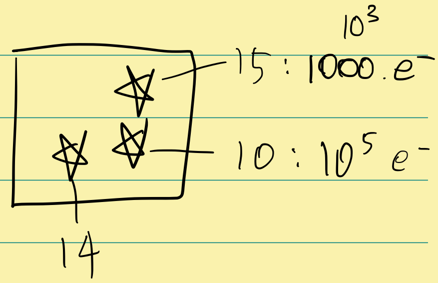
$m_{sky} = 2$ [ADU]

$s_{sky} = 0.2$ [ADU]

- $ADU_{max} = 2^{16} - 1 = 65535$







4 Source sum

$$N_{source}^{(ADU)} = N_{apsum}^{(ADU)} - m_{sky}^{(ADU)} \times n_{ap}$$
$$= 15\pi - 2 \times \pi$$
$$= 13\pi \approx 40 \ [ADU]$$

5 Source error

Error propagation of (*) : $\Delta N_{source}^{(ADU)} = \sqrt{(\delta N_{apsum}^{(ADU)})^2 + (\Delta m_{sky}^{(ADU)} \times n_{ap})^2}$

But

* Poisson noise must be calculated in e^- unit, NOT ADU.

$$- \ \Delta N_{apsum}^{(ADU)} = \sqrt{N_{apsum}^{(ADU)} \times g + N_{ro}^{(e)2} n_{ap}} \ /g$$
$$= \sqrt{N_{apsum}^{(ADU)}/g + N_{ro}^{(e)2}/g^2 \cdot n_{ap}}$$

Poissonian & gaussian

"ADU" × g = "e"

$$- \ \delta m_{sky}^{(ADU)} = s_{sky}^{(ADU)}/\sqrt{n_{sky}}$$

so

$$\Delta N_{source}^{(ADU)} = \sqrt{\frac{N_{apsum}^{(ADU)}}{g} + N_{ro}^{(e)2}/g^2 \cdot n_{ap} + \frac{s_{sky}^{(ADU)2} n_{ap}^2}{n_{sky}}}$$
$$= \sqrt{\frac{N_{source}^{(ADU)}}{g} + \left[\frac{m_{sky}^{(ADU)}}{g} + \left(\frac{N_{ro}^{(e)}}{g}\right)^2\right] n_{ap} + \frac{s_{sky}^{(ADU)2} n_{ap}^2}{n_{sky}}}$$

Ex $g=2$, $N_{apsum}^{(ADU)}=5$

Poisson noise "$\sqrt{N}$"

X $\sqrt{5} = 2.xx$

O $\frac{\sqrt{5 \times g}}{g} = \sqrt{\frac{5}{2}} = 1.xx$

If $g = 2 \ e^-/ADU$, $N_{ro}^{(e)} = 4$,

$$\Delta N_{source}^{(ADU)} \simeq \sqrt{\underbrace{\frac{40}{2}}_{pois} + [\underbrace{\frac{2}{2}}_{pois} + \underbrace{(\frac{4}{2})^2}_{ro}]\pi + \underbrace{\frac{0.2^2 \cdot \pi^2}{100}}_{\text{sky estimation}}} \approx 6$$

negligible

So $N_{source} \approx 40 \pm 6$ ADU

* Error Propagation

Function	Variance	Standard Deviation
$f = aA$	$\sigma_f^2 = a^2\sigma_A^2$	$\sigma_f = \lvert a\rvert\sigma_A$
$f = aA + bB$	$\sigma_f^2 = a^2\sigma_A^2 + b^2\sigma_B^2 + 2ab\,\sigma_{AB}$	$\sigma_f = \sqrt{a^2\sigma_A^2 + b^2\sigma_B^2 + 2ab\,\sigma_{AB}}$
$f = aA - bB$	$\sigma_f^2 = a^2\sigma_A^2 + b^2\sigma_B^2 - 2ab\,\sigma_{AB}$	$\sigma_f = \sqrt{a^2\sigma_A^2 + b^2\sigma_B^2 - 2ab\,\sigma_{AB}}$
$f = AB$	$\sigma_f^2 \approx f^2\left[\left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 + 2\frac{\sigma_{AB}}{AB}\right]$ [11][12]	$\sigma_f \approx \lvert f\rvert\sqrt{\left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 + 2\frac{\sigma_{AB}}{AB}}$
$f = \frac{A}{B}$	$\sigma_f^2 \approx f^2\left[\left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 - 2\frac{\sigma_{AB}}{AB}\right]$ [13]	$\sigma_f \approx \lvert f\rvert\sqrt{\left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 - 2\frac{\sigma_{AB}}{AB}}$
$f = aA^b$	$\sigma_f^2 \approx \left(abA^{b-1}\sigma_A\right)^2 = \left(\frac{fb\sigma_A}{A}\right)^2$	$\sigma_f \approx \lvert abA^{b-1}\sigma_A\rvert = \left\lvert\frac{fb\sigma_A}{A}\right\rvert$
$f = a\ln(bA)$	$\sigma_f^2 \approx \left(a\frac{\sigma_A}{A}\right)^2$ [14]	$\sigma_f \approx \left\lvert a\frac{\sigma_A}{A}\right\rvert$
$f = a\log_{10}(bA)$	$\sigma_f^2 \approx \left(a\frac{\sigma_A}{A\ln(10)}\right)^2$ [14]	$\sigma_f \approx \left\lvert a\frac{\sigma_A}{A\ln(10)}\right\rvert$
$f = ae^{bA}$	$\sigma_f^2 \approx f^2(b\sigma_A)^2$ [15]	$\sigma_f \approx \lvert f(b\sigma_A)\rvert$
$f = a^{bA}$	$\sigma_f^2 \approx f^2(b\ln(a)\sigma_A)^2$	$\sigma_f \approx \lvert f(b\ln(a)\sigma_A)\rvert$
$f = a\sin(bA)$	$\sigma_f^2 \approx [ab\cos(bA)\sigma_A]^2$	$\sigma_f \approx \lvert ab\cos(bA)\sigma_A\rvert$
$f = a\cos(bA)$	$\sigma_f^2 \approx [ab\sin(bA)\sigma_A]^2$	$\sigma_f \approx \lvert ab\sin(bA)\sigma_A\rvert$
$f = A^B$	$\sigma_f^2 \approx f^2\left[\left(\frac{B}{A}\sigma_A\right)^2 + (\ln(A)\sigma_B)^2 + 2\frac{B\ln(A)}{A}\sigma_{AB}\right]$	$\sigma_f \approx \lvert f\rvert\sqrt{\left(\frac{B}{A}\sigma_A\right)^2 + (\ln(A)\sigma_B)^2 + 2\frac{B\ln(A)}{A}\sigma_{AB}}$

example proof

pf) $df = a \cdot \frac{1}{\ln(10)} \cdot \frac{1}{bA} \cdot d(bA) = \frac{1}{a \ln 10} \frac{dA}{A}$.

$\boxed{6}$ ADU to instrumental magnitude

Say exposure time $t_{exp} = 0.1\,s$

$$m_{inst} := -2.5 \log_{10} N^{(e)}_{source}/t_{exp} = -2.5 \log_{10} g\, N^{(ADU)}_{source}/t_{exp}$$

$$= -2.5 \log_{10} (2 \cdot 40 / 0.1)$$

$$= -7.26$$

$$\Delta m_{inst} = \left| -2.5 \cdot \frac{1}{\ln(10)} \frac{\Delta N^{(ADU)}_{source}}{N^{(ADU)}_{source}} \right|$$

$$= \left| -1.086 \frac{\Delta N}{N} \right|$$

$$\approx \left| -1.086 \frac{6}{40} \right|$$

$$= 0.163$$

← Even though you use $\Delta N^{(e)}/N^{(e)}$, not ADU, the ratio will be the same:

$$\frac{\Delta N^{(e)}}{N^{(e)}} = \frac{g \cdot \Delta N^{(ADU)}}{g \cdot N^{(ADU)}} = \frac{\Delta N^{(ADU)}}{N^{(ADU)}}$$

So

$$m_{inst} = -7.26 \pm 0.163.$$

NOTE

Δm is independent of exposure time on the definition of the instrumental magnitude above.
Also it is independent of whether you used ADU or electrons.
The coefficients 1.086 should be familiar to you (2.5/ln(10)).
Since 1.086 ~ 1, 10% error (ΔN/N = 0.1) is directly converted to Δm = 0.1.
But if the error gets larger, e.g., 80% error, this error calculation doesn't hold anymore.
This is because we assumed infinitesimal change in N (remember we did "differentiation", which basically takes the first order change of Taylor expansion).

NOTE

The magnitude we obtained, -7 mag, is extremely bright value as you may have already realized.
This bright magnitude comes from our definition of m_inst. Pogson's formula states

$$m_1 - m_2 = -2.5 \log_{10} \frac{N_1}{N_2} \quad (N \text{ can be } \# \text{ of } e^-, \text{ ADU, flux density, etc})$$

But what we did in the definition of m_inst is to set m_1 = m_inst and m_2 = 0 with N_2 = 1. The 0-magnitude star MUST have N_2 >> 1 !! This is why instrumental magnitude must be "transformed" into the "standard" magnitude system, which we normally use.

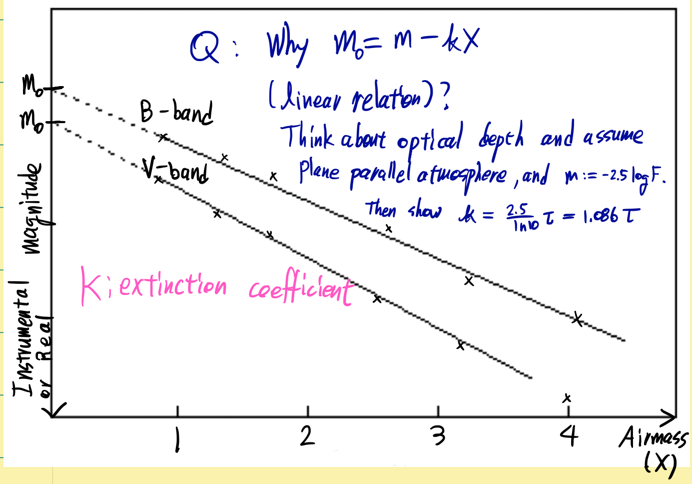
The next step, [7], is where we will learn this "standardization". But before doing so, we need to understand atmospheric extinction.

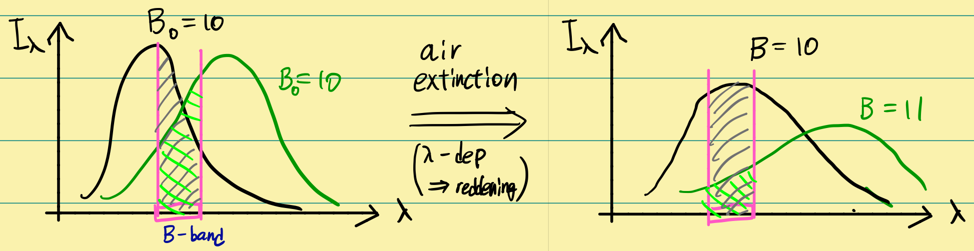
Atmospheric Extinctions

Before going further to obtain real magnitude m from the instrumental magnitude m_inst, we must understand about the atmospheric extinction.

Basically the atmospheric extinction is wavelength dependent. Thus in spectroscopy, the determination of extinction must be done at every single wavelength bin. This can be eased when we assume atmospheric extinction is constant over the night: Then we can use the standard star data to normalize the flux of the given spectrum (called flux calibration). This is reliable when we made observation at high altitude and when the standard star is observed right after/before the object observation.

$m_0 = m - k X$





So in general, $k = k(\lambda)$, or in photometry, $k(\text{filter})$, i.e., different at each filter: $k_B \neq k_V \neq k_{g'} \neq \cdots$

Normally we denote $k_A = k_A' + k_A'' C$ where A: filter & C: C.I. (B-V, g-r, etc).

※ k_A may differ from target to target but k_A' & k_A'' are assumed to be constants.

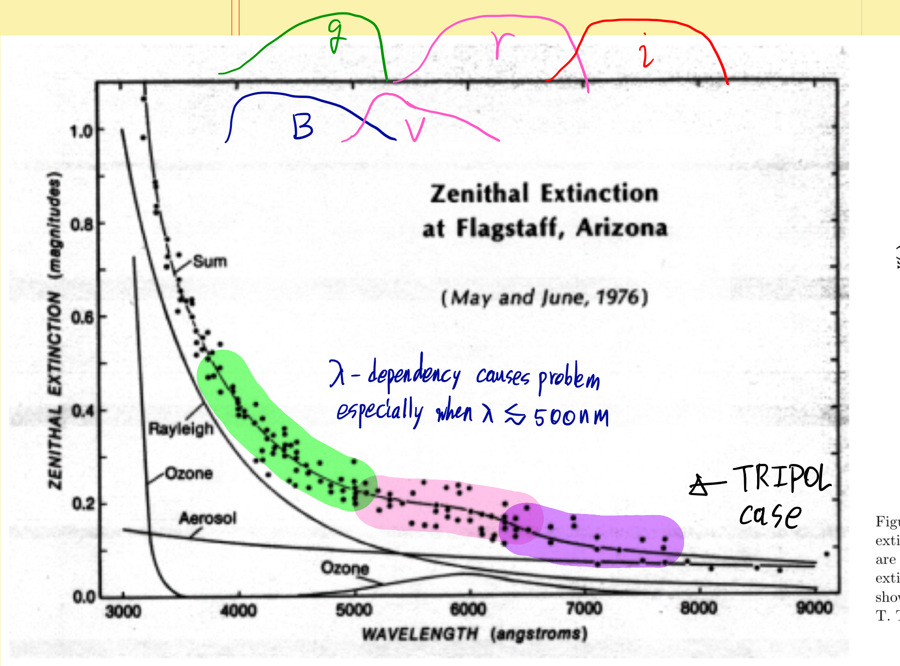


Figure 19.2: Extinction looking "straight up" (towards zenith) from Flagstaff, showing components of extinction. (from "A New Absolute Calibration of Vega" Sky and Telescope Oct 1978)

Seasonal effect

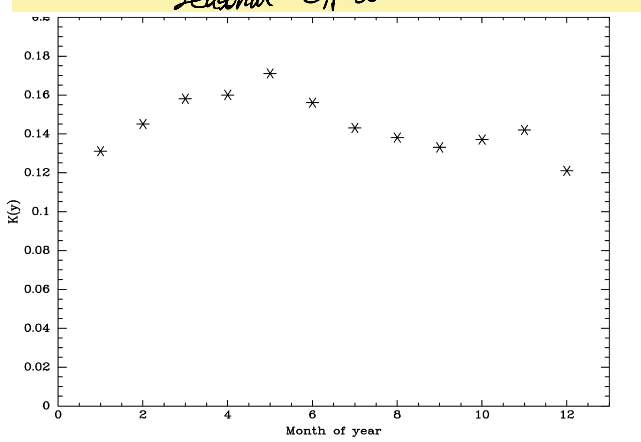


Figure 19.3: Seasonal variation of extinction over Flagstaff. Plotted are the monthly median extinction values (in mag) in the y band, an intermediate band filter centered at 5500 Å. These are for the years 1976-1980, when there was no significant volcanic contribution to the atmospheric extinction. Note that while there is a definite seasonal pattern, the individual nightly values (not shown here) show considerable night-to- night scatter. (data taken from G. W. Lockwood and D. T. Thompson AJ 92 p. 976 , 1986)

remember: This page describes about a simple example, not the real case.
In reality, we cannot determine each parameters separately as in this case. See next page.

So $m_A = m_A(X, C) = m_A(0) + k_A X = m_A(0) + (k_A' + k_A'' C) X$

Strategy to find k_A, k_A', and k_A'' at the filter A :

① Get k at filters of interest : observe std ☆ at many X
$\rightarrow$ get m_{inst}
$\rightarrow$ Plot m_{inst} VS X and do linear regression

Here we are assuming that k and z (see next page) is constant over the night

② observe 'red-blue' pairs : Select 2 std ☆'s, e.g., $g-r = -0.3$ & $+0.3$.

$$
\begin{array}{r}
k_1 = k' + k'' C_1 \\
- \quad k_2 = k' + k'' C_2 \\
\hline
\Delta k = k'' \Delta C
\end{array}
\quad \Rightarrow \quad k'' = \frac{\Delta k}{\Delta C}
$$

$$k' = k_1 - k'' C_1 = k_2 - k'' C_2$$

Here we are assuming that k1, k2, k', k'', and z (see next page) are constant over the night

$\rightarrow$ get k'' at this filter
get k' from $k = k' + k'' C$

Example

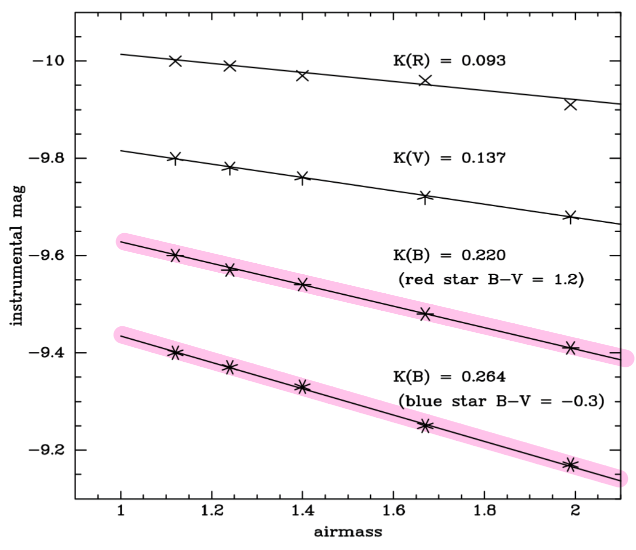


Figure 19.1: Example instrumental mag vs. airmass for stars. This shows that the K value is the slope of the line in the mag vs. airmass plot. The instrumental magnitudes have been arbitrarily shifted so they would not overlap. An exercise for the student: from the information on the B extinction for two stars as shown on the graph, calculate K'_B and K''_B.

① K_B is determined from the fittings for star red & blue.
$k_{1B} = 0.220$; $k_{2B} = 0.264$

② $k''_B = \frac{\Delta k}{\Delta C} = \frac{-0.044}{1.5} = -0.03$

③ $k'_B = k_1 - k'' C_1 = k_2 - k'' C_2 = 0.256$

7 Standardization

From the standard star observation, we obtained the k values (k' and k") at a given filter.
Recall Pogson's formula:

$$m_1 - m_2 = -2.5 \log_{10} N_1 / N_2$$

we put m2=0 and N2=1 to define the instrumental magnitude. If N2 was truly the N of a real 0-magnitude star, m1 would have been larger than m_inst:

$$m_{real} = m_{inst} + z \qquad z: \text{zero point}$$

z is dependent on the electronics of the time when the image was taken. So this value may differ from CCD to CCD and every time you read out the CCD, it may change. Plus, a stronger effect is the color effect: The filter you have may be slightly differ from the ideal A-band filter (A = V, B, g, r, ...) and the CCD may have QE = QE(λ). Thus, z = z(C), i.e., a function of color of the star.

$$z \approx z_0 + a_A C \qquad (a_A: \text{const})$$

Finally

$$m_{A,inst} \approx (m_{A,real} + k_A X) + (z_0 + a_A C)$$

$$\approx m_{A,real} + k_A' X + k_A'' C X + z_0 + a_A C \quad \cdots (*)$$

※ I didn't use 'C_real', but just C. For $m_{A,real}$, I put 'real' to avoid confusion with $m_{A,inst}(X=0)$, etc.

A	The filter of interest (B, V, g, r, i, ...)
$m_{A,real}$	The standardized magnitude in A-band
k'_A	The 1st order extinction coefficient in A-band
k''_A	The 2nd order extinction coefficient in A-band
z_0	(the constant) zero point
C	The real color index (B-V, g-r, or whatever appropriate)
a_A	The color term coefficient in A-band

Nominal values

Filter	U	B	V
k'	0.4	0.2	0.1
k"	nearly 0		
a	[-0.1, +0.1]		

z_0 can be any value but mostly ~25.
Q (why this can be so similar for many CCDs?)

Due to the fact that C is the "real" color index, we must observe at least two bands and compare this C and the instrumental color index.

Write (*) for A & B bands and subtract to get

$$C_{inst} = (1 + a_C) C + k_C' X + k_C'' C X \quad \cdots (\dagger)$$

where k'_C, k''_C, and a_C are $\square_A - \square_B$.

Nominal values

Color	U-B	B-V
k"	nearly 0	
a	nearly 0	

Normally we are not so much interested in very accurate measurements, and even we do, our instruments are not perfect enough to detect k" value precisely. It is thus usual to assume k" = 0 unless you have to do the performance evaluation.

$$(*): \quad m_{A,inst} = m_{A,real} + k_A' X + z_0 + a_A C$$

$$(\dagger): \quad C_{inst} = k_C' X + (1 + a_C) C$$

Page Date

Strategy:

(1) Observe 2 standard stars (red-blue pair) as well as your target.
--- using, e.g., g and r band.
--- at many different airmasses on ONE single night.
--- Be sure that there must be a pair of g and r data at given airmass X.
--- It is better to observe standard stars such that their X range includes the X of your object.

(2) Do photometry and get instrumental magnitudes at each X, each star.

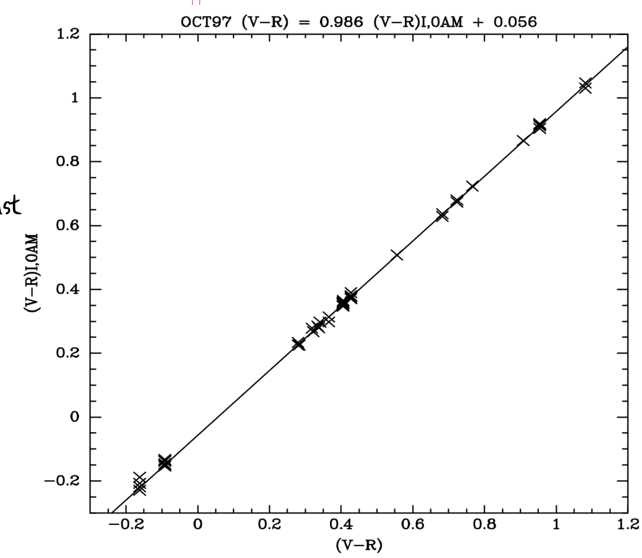
(3) Solve the previous two equations using, e.g., python's ``scipy.optimize.curve_fit``.
--- Determine k', k", a, and z values and their errors.

(4) Use these parameters to determine the m_real and C of the object.
--- Knowing magnitude in g and g-r color means you know r magnitude too.

(5) If you want to use i band too, do the same thing with, e.g., r and i bands.
--- You must have been made the observation of the two standard stars as in step 1.
--- Then you'll get r and r-i.
--- Compare this r with step (4) to validate your data reduction.

Simple example case

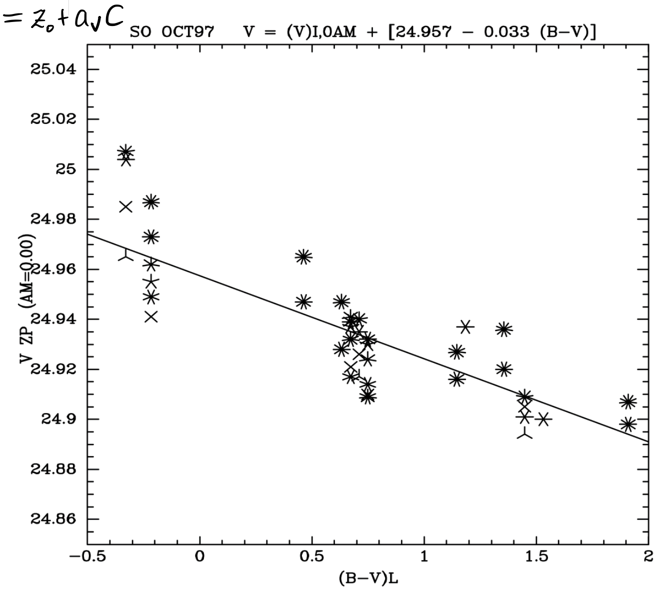
C_{inst}



C (real)

Figure 20.1: Steward 2.3m V−R transformation. Plotted on the x axis are the standard (Landolt) V−R colors of standard stars. Plotted on the y axis are the instrumental colors, corrected to zero airmass. The title gives the transformation equation. The sharp- eyed among you may see that the fit does not seem to be quite right for the two bluest sets of points. This is because the fit was done to de-emphasize these points, as none of the objects we were interested in were anywhere near this blue.

$Z_V(C)$
$= Z_0 + a_V C$



C (real)

Figure 20.2: Steward Observatory 2.3m telescope V transformation. Different symbols are for different nights during the run. There is a slight color dependence to the V mag zero point. The transformation equation is shown in the title.

Page Date