Astronomical Observing Techniques 2019

Lecture 5: Your Noise is My Signal

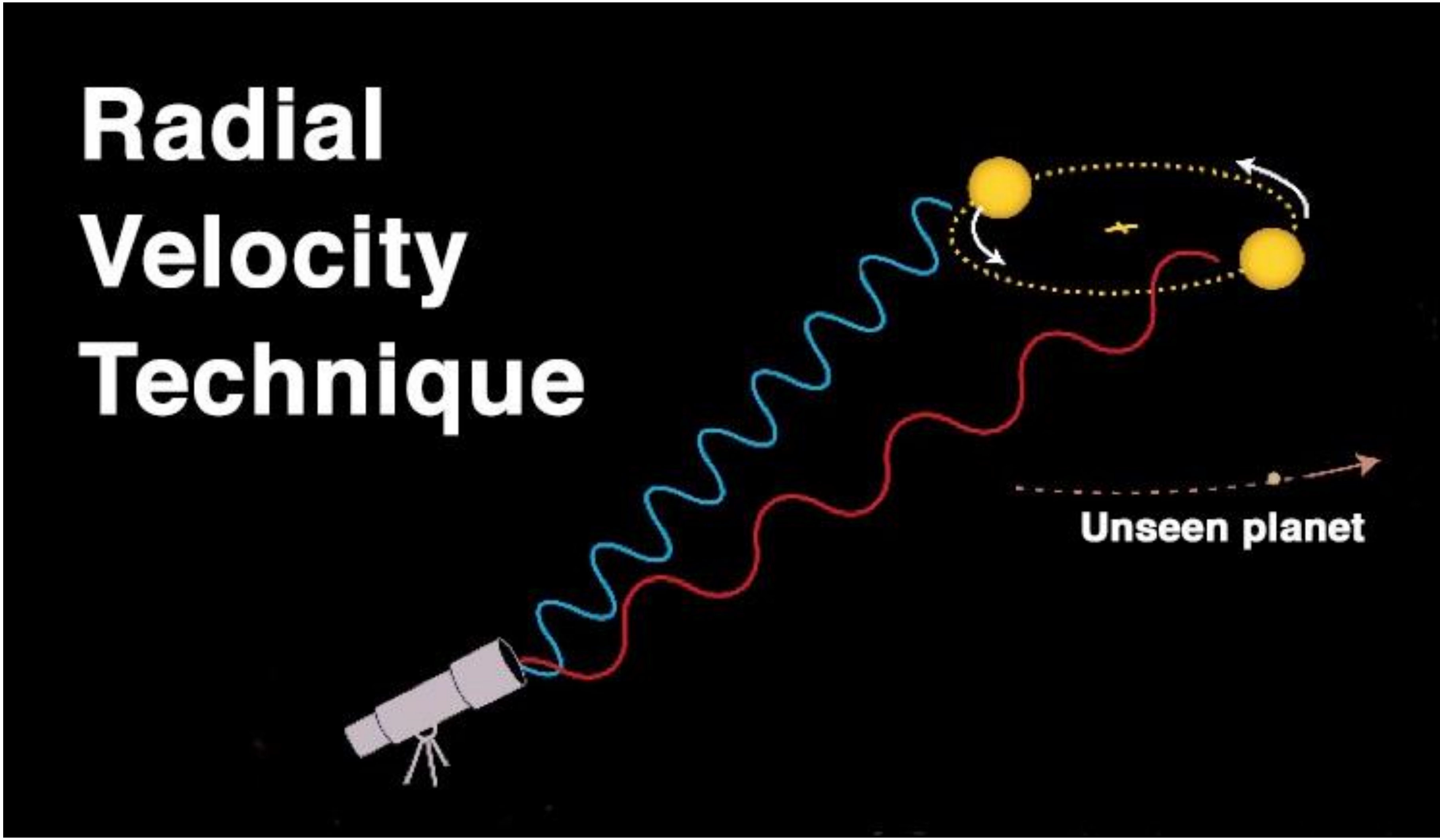
Huub Rottgering
rottgering@strw.leidenuniv.nl

Content

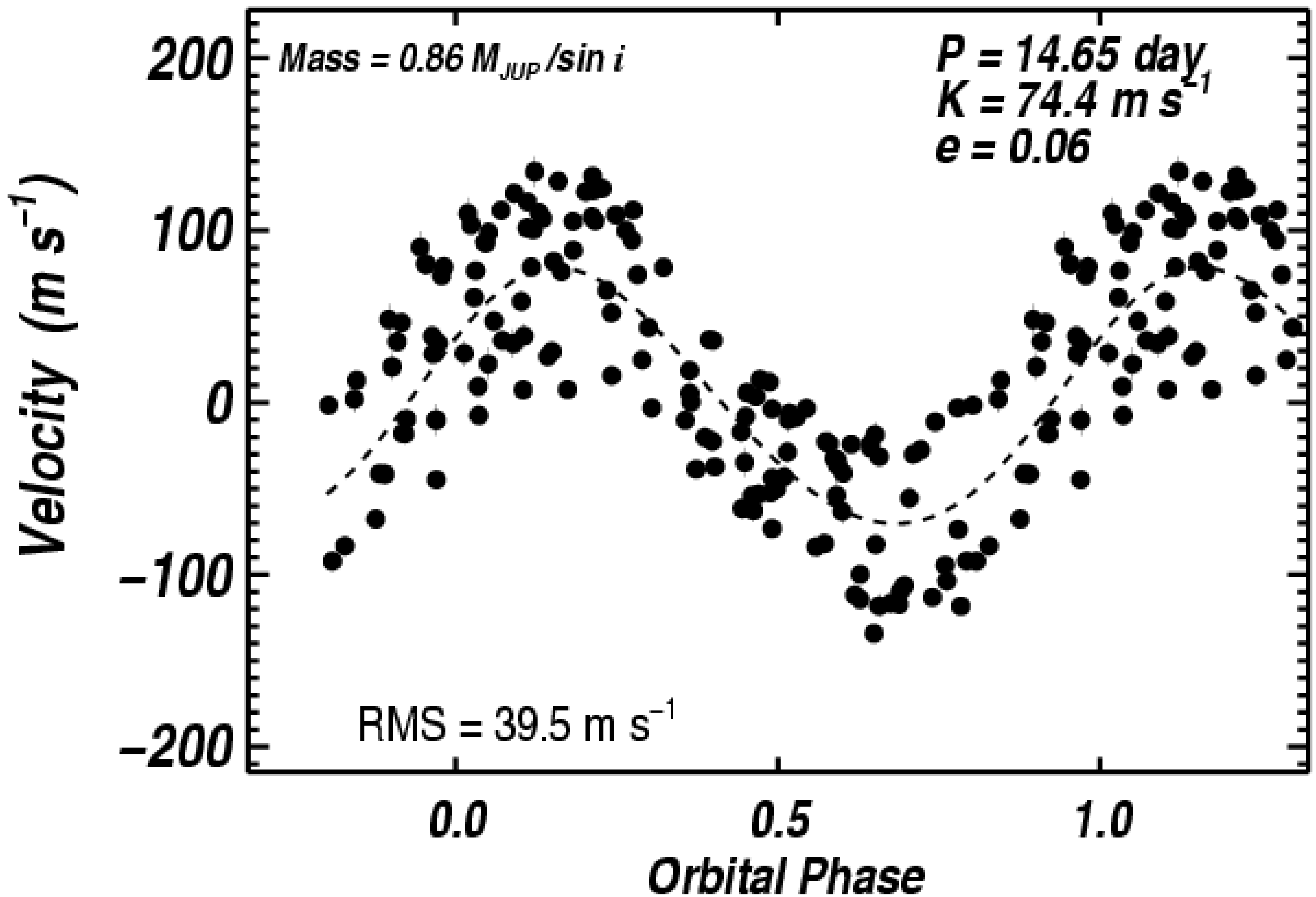
Signal and Noise

- Signal: data that is relevant for our science
- Noise: data that is irrelevant for our science
- Signal-to-noise ratio (SNR): ratio of relevant to irrelevant information

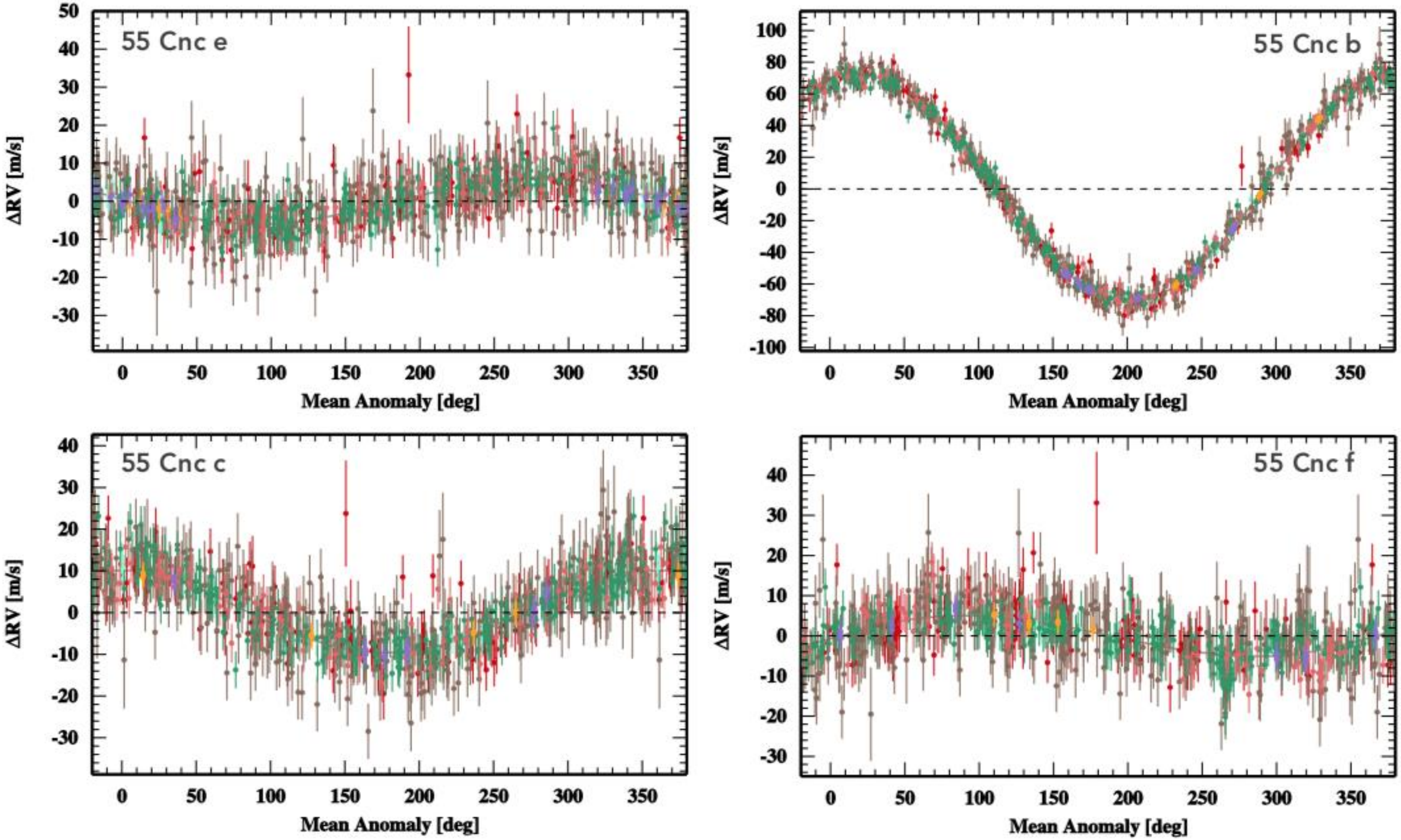
- Definition of signal and noise inherently depends on science objectives


Radial
Velocity
Technique
Unseen planet

Exoplanet in Radial Velocity



exoplanets.org/

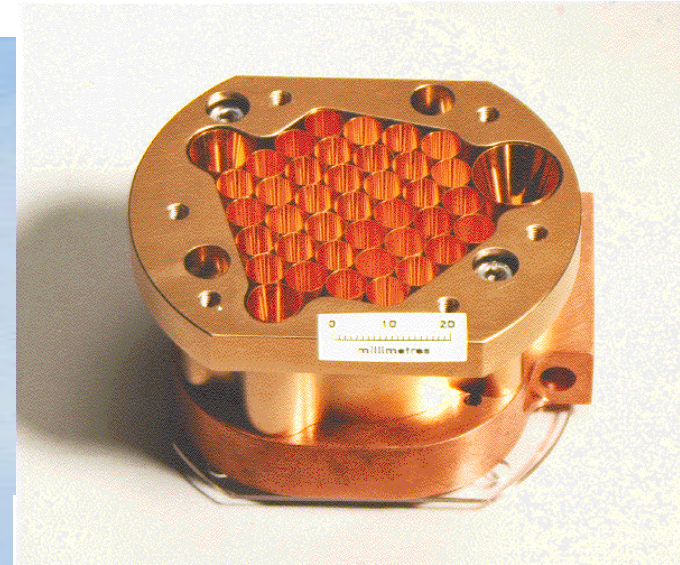


[Bourrier etal. 2018A&A...619A...1B]

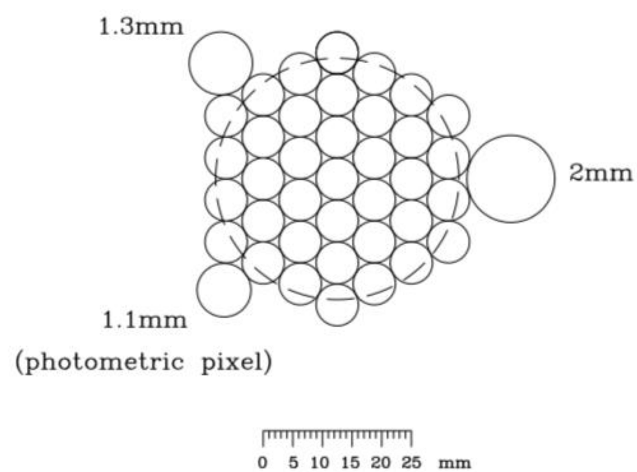
Table 3 Best-fitted solution for the planetary system orbiting 55 Cnc. For each parameter, the median of the posterior is considered, with error bars computed from the MCMC marginalized posteriors using a 68.3% confidence interval. $\sigma_{(O-C)X}$ corresponds to the standard deviation of the residuals around this best solutions for instrument X, and $\sigma_{(O-C)\,\mathrm{all}}$ the weighted standard deviation for all the data. All the parameters probed by the MCMC can be found in the Appendix, in Tables .2 and .3

Param.	Units	55 Cnc e	55 Cnc b	55 Cnc c	55 Cnc f	magnetic cycle	55 Cnc d
P	[d]	$0.73654737^{+1.30\,10^{-6}}_{-1.44\,10^{-6}}$	$14.6516^{+0.0001}_{-0.0001}$	$44.3989^{+0.0042}_{-0.0043}$	$259.88^{+0.29}_{-0.29}$	$3822.4^{+76.4}_{-77.4}$	$5574.2^{+93.8}_{-88.6}$
K	[m s^{-1}]	$6.02^{+0.24}_{-0.23}$	$71.37^{+0.21}_{-0.21}$	$9.89^{+0.22}_{-0.22}$	$5.14^{+0.26}_{-0.25}$	$15.2^{+1.6}_{-1.8}$	$38.6^{+1.3}_{-1.4}$
e		$0.05^{+0.03}_{-0.03}$	$0.00^{+0.01}_{-0.01}$	$0.03^{+0.02}_{-0.02}$	$0.08^{+0.05}_{-0.04}$	$0.17^{+0.04}_{-0.04}$	$0.13^{+0.02}_{-0.02}$
ω	[deg]	$86.0^{+30.7}_{-33.4}$	$-21.5^{+56.9}_{-89.8}$	$2.4^{+43.1}_{-49.2}$	$-97.6^{+37.0}_{-51.3}$	$174.7^{+16.6}_{-14.1}$	$-69.1^{+9.1}_{-7.9}$
T_C	[d]	$55733.0060^{+0.0014}_{-0.0014}$	$55495.587^{+0.013}_{-0.016}$	$55492.02^{+0.34}_{-0.42}$	$55491.5^{+4.8}_{-4.8}$	$55336.9^{+45.5}_{-50.6}$	$56669.3^{+83.6}_{-76.5}$
a	[AU]	$0.0154^{+0.0001}_{-0.0001}$	$0.1134^{+0.0006}_{-0.0006}$	$0.2373^{+0.0013}_{-0.0013}$	$0.7708^{+0.0043}_{-0.0044}$	–	$5.957^{+0.074}_{-0.071}$
M	[M_{Jup}]	$0.0251^{+0.0010}_{-0.0010}$	–	–	–	–	–
M	[M_{Earth}]	$7.99^{+0.32}_{-0.33}$	–	–	–	–	–
M.sin i	[M_{Jup}]	–	$0.8036^{+0.0092}_{-0.0091}$	$0.1611^{+0.0040}_{-0.0040}$	$0.1503^{+0.0076}_{-0.0076}$	–	$3.12^{+0.10}_{-0.10}$
M.sin i	[M_{Earth}]	–	$255.4^{+2.9}_{-2.9}$	$51.2^{+1.3}_{-1.3}$	$47.8^{+2.4}_{-2.4}$	–	$991.6^{+30.7}_{-33.1}$

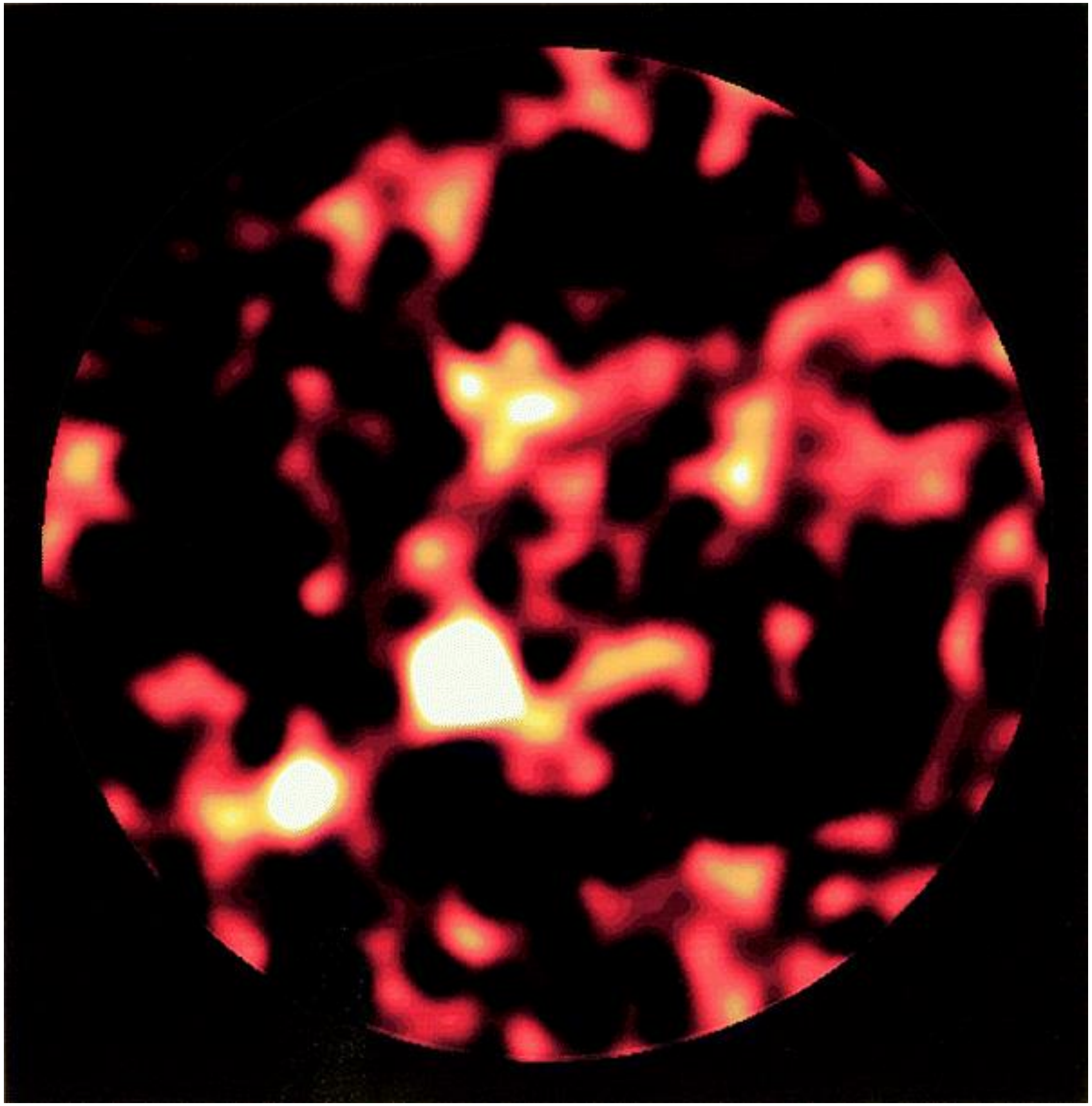
JCMT+SCUBA



(37 detectors)

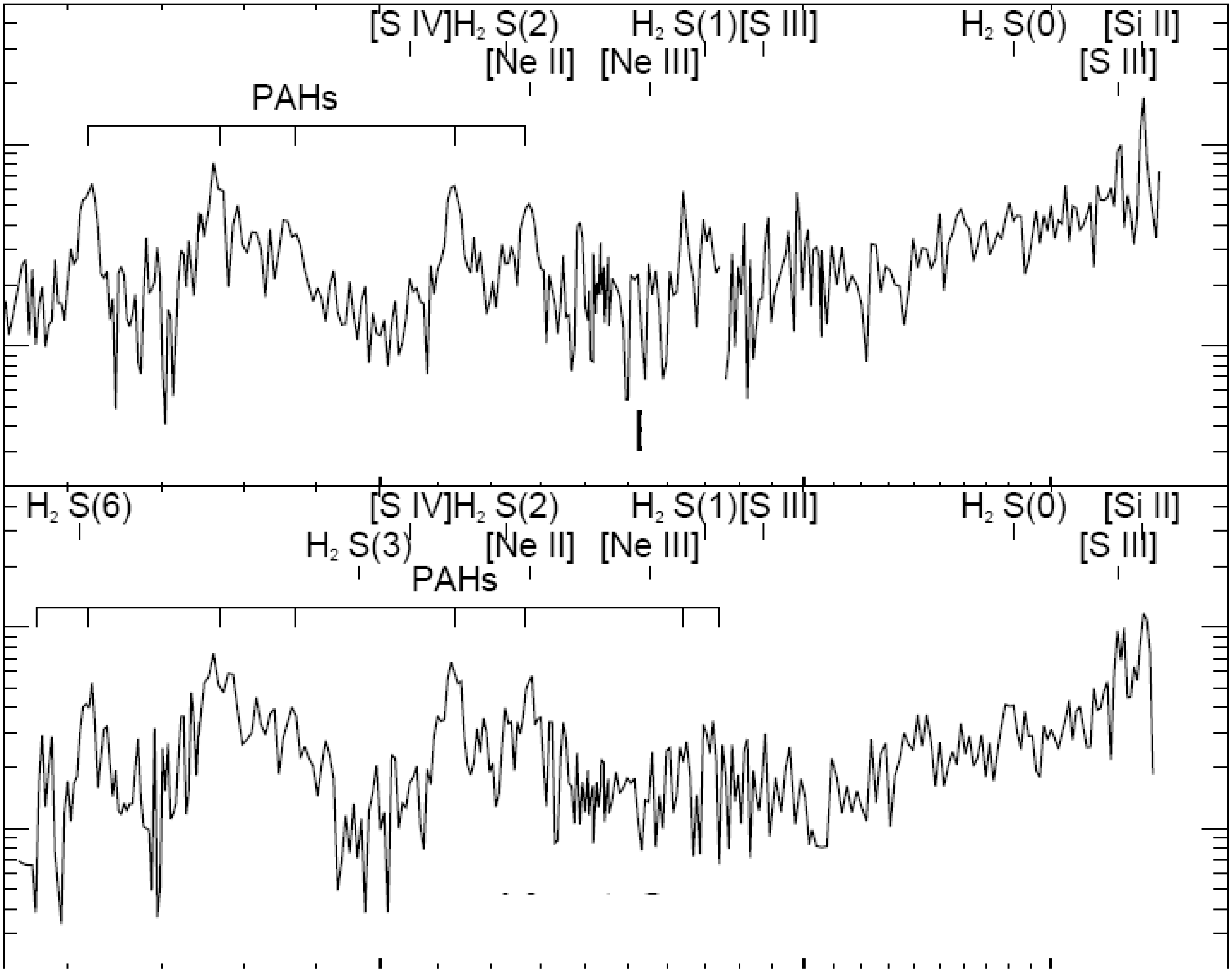


Signal or Noise?



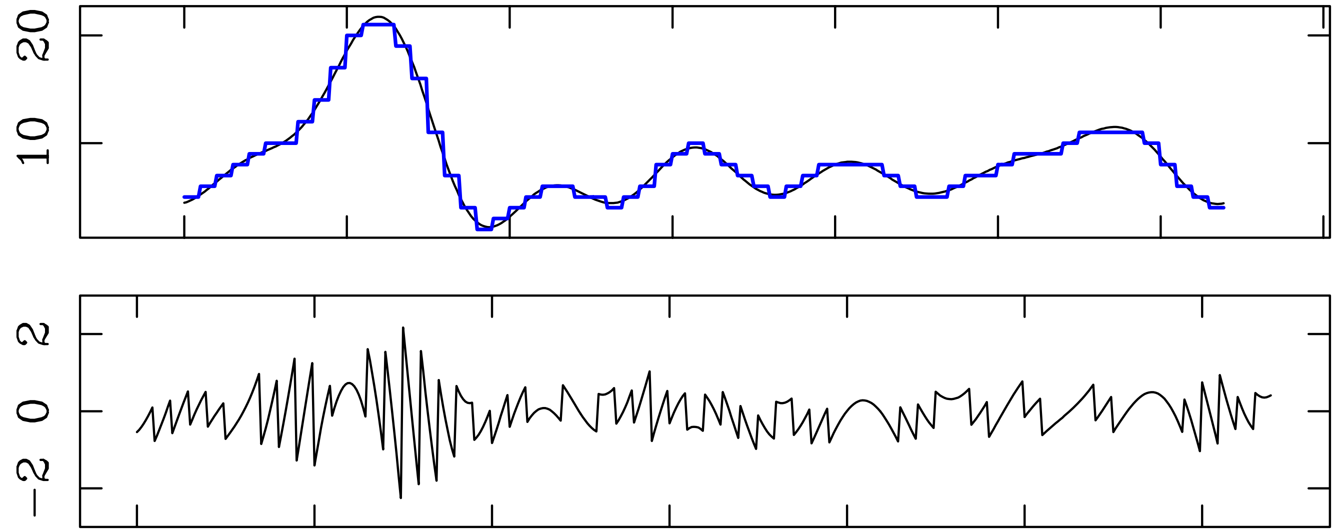
SCUBA 850μm map of the Hubble deep field (Hughes et al. 1998, Nature volume 394, pages 241–247

What is Noise? And what is real Signal?

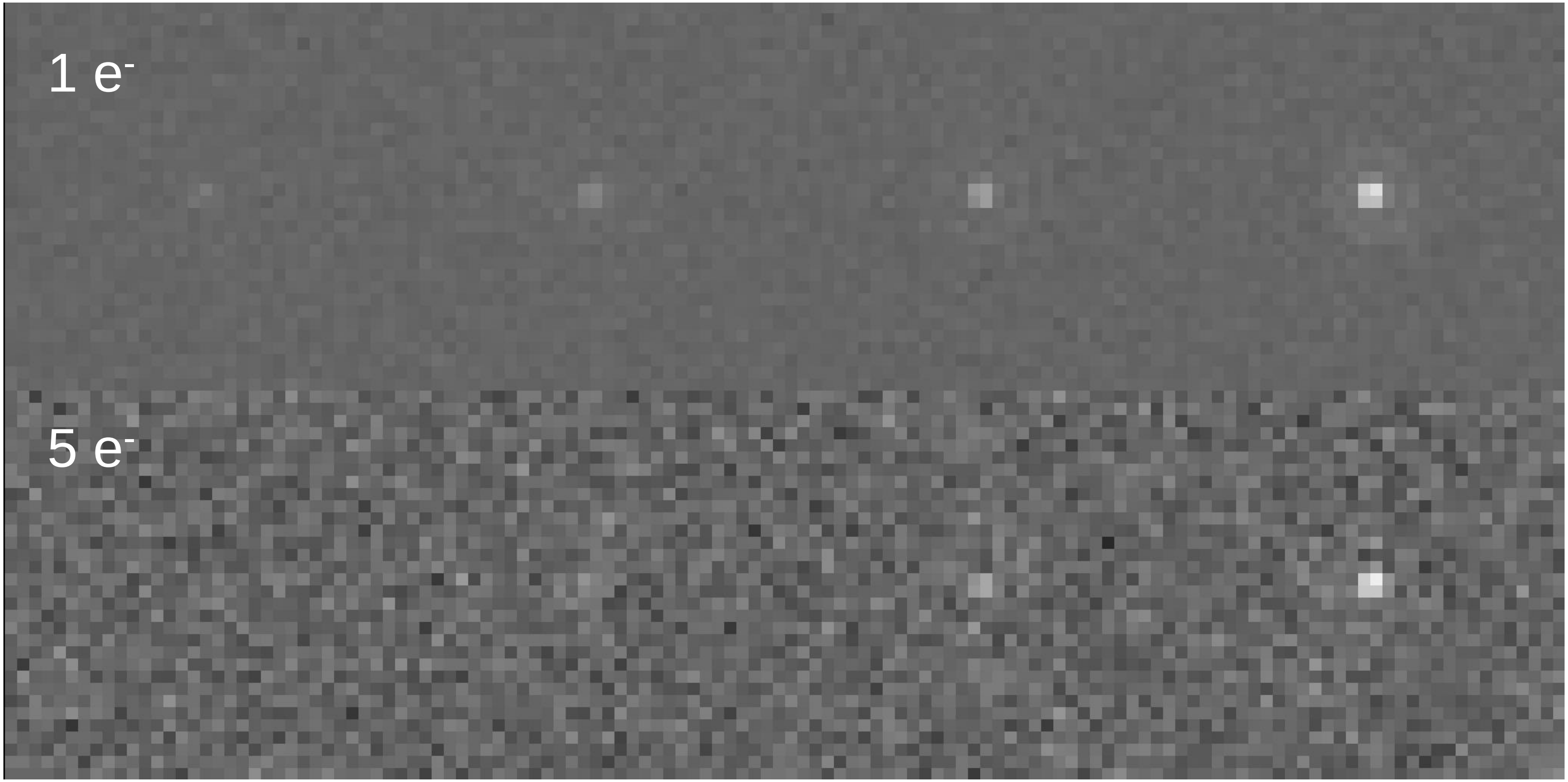


Example 1: Digitization/Quantization Noise

- Analog-to-Digital Signal Converter (ADC).
- Number of bits determines dynamic range of ADC
- Resolution: 12 bit 2^{12} = 4096 quantization levels
  16 bit 2^{16} = 65636 quantization levels
- Discrete, "artificial" steps in signal levels → noise



Example 2: Read-out noise in CCDs



7e- 12e- 20e- 38e-

Some Sources of Noise in Astronomical Data

Noise type	Signal	Background
Photon shot noise	X	X
Scintillation	X	
Cosmic rays		X
Image stability	X	
Read noise	X	X
Dark current noise	X	X
Charge transfer efficiencies (CCDs)	X	X
Flat fielding (non-linearity)	X	X
Digitization noise	X	X
Other calibration errors	X	X
Image subtraction	X	X

Probability

The **Calculus of Probabilities** is the mathematical theory which allows us to predict the statistical behaviour of complex systems from a basic set of fundamental axioms. Probabilities come in two distinct forms: discrete, where P_i is the probability of the i^{th} event occurring, and continuous, where $P(x)$ is the probability that the even, or **random variable**, x, occurs.

1. **The Range of Probabilities:** The probability of an event is measurable on a continuous scale, such that $P(x)$ is a real number in the range $0 \leq P(x) \leq 1$.

2. **The Sum Rule:** The sum of all discrete possibilities is

$$\sum_i P_i = 1. \tag{1}$$

For a continuous range of random variables, x, this becomes

$$\int_{-\infty}^{\infty} dx\, p(x) = 1, \tag{2}$$

where $p(x)$ is the **probability density**. The probability density clearly must have units of $1/x$.

Credits: https://www.roe.ac.uk/~ant/Teaching/Astronomical%20Stats/

Probability distributions can be characterized by their **moments**.

Definition:

$$m_n \equiv \langle x^n \rangle = \int_{-\infty}^{\infty} dx\, x^n p(x), \tag{6}$$

is the n^{th} moment of a distribution. The angled brackets $\langle \cdots \rangle$ denote the **expectation value**. Probability distributions are normalized so that

$$m_0 = \int_{-\infty}^{\infty} dx\, p(x) = 1 \tag{7}$$

(Axiom 2, The Sum Rule).

The first moment,

$$m_1 = \langle x \rangle, \tag{8}$$

gives the **expectation value** of x, called the **mean**: the average or typical expected value of the random variable x if we make random drawings from the probability distribution.

Centred moments are obtained by shifting the origin of x to the mean;

$$\mu_n \equiv \langle (x - \langle x \rangle)^n \rangle. \tag{9}$$

The second centred moment,

$$\mu_2 = \langle (x - \langle x \rangle)^2 \rangle, \tag{10}$$

is a measure of the **spread** of the distribution about the mean. This is such an important quantity it is often called the **variance**, and denoted

$$\sigma^2 \equiv \mu_2. \tag{11}$$

We will need the following, useful result later:

$$\sigma^2 = \langle (x - \langle x \rangle)^2 \rangle = \langle (x^2 - 2x\langle x \rangle + \langle x \rangle^2) \rangle = \langle x^2 \rangle - \langle x \rangle^2. \tag{12}$$

The variance is obtained from the mean of the square minus the square of the mean. Another commonly defined quantity is the square-root of the variance, called the **standard deviation**; σ. This quantity is sometimes also called the **root mean squared (rms) deviation**[3], or **error**[4].

Three Probability Density Functions

- Binomial distribution
- Poisson distribution
- Gaussian/normal distribution

Binomial distribution

The Binomial distribution allows us to calculate the probability, P_n, of n successes arising after N independent trials.

Suppose we have a sample of objects of which a probability, p_1, of have some attribute (such as a coin being heads-up) and a probability, $p_2 = 1 - p_1$ of not having this attribute (e.g. tails-up). Suppose we sample these objects twice, e.g. toss a coin 2 times, or toss 2 coins at once. The possible out comes are hh, ht, th, and tt. As these are independent events we see that the probability of each distinguishable outcome is

$$\begin{aligned} P(hh) &= P(h)P(h), \\ P(ht+th) &= P(ht)+P(th) = 2P(h)P(t), \\ P(tt) &= P(t)P(t). \end{aligned} \tag{15}$$

These combinations are simply the coefficients of the binomial expansion of the quantity $(P(h) + P(t))^2$. In general, if we draw N objects, then the number of possible permutations which can result in n of them having some attribute is the n^{th} coefficient in the expansion of $(p_1 + p_2)^N$, the probability of each of these permutations is $p_1^n p_2^{N-n}$, and the probability of n objects having some attribute is the binomial expansion

$$P_n = C_n^N p_1^n p_2^{N-n}, \quad (0 \leq n \leq N), \tag{16}$$

where

$$C_n^N = \frac{N!}{n!(N-n)!} \tag{17}$$

are the **Binomial coefficients**. The binomial coefficients can here be viewed as statistical weights which allow for the number of possible indistinguishable permutations which lead to the same outcome. This distribution is called the general **Binomial**, or **Bernoulli distribution**. We can plot out the values of P_n for all the possible n (Figure 1) and in

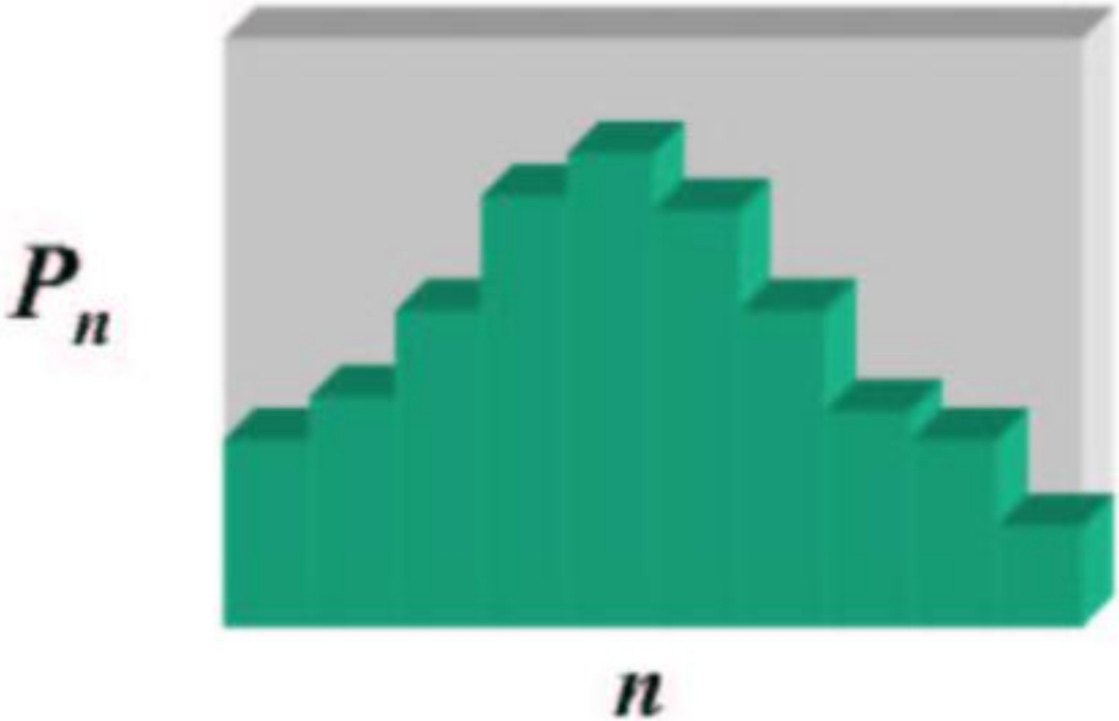


Figure 1: *Histogram of a binomial distribution*

doing so have generated the predicted probability distribution which in this case is the binomial distribution whose form is determined by N, n, p_1. If we have only two possible outcomes $p_2 = 1 - p_1$. The Binomial distribution can be generalised to a multinomial distribution function.

The mean of the binomial distribution is

$$\begin{aligned}
\langle n \rangle &= \sum_{n=0}^{N} nP \\
&= \sum_{n=0}^{N} n \frac{N!}{n!(N-n)!} p_1^n p_2^{N-n} \\
&= \sum_{n=1}^{N} \frac{N!}{(n-1)!(N-n)!} p_1^n p_2^{N-n} \\
&= \sum_{n=1}^{N} \frac{N(N-1)!}{(n-1)!(N-n)!} p_1 p_1^{n-1} p_2^{N-n} \\
&= Np_1
\end{aligned}$$

For $p_1 \neq p_2$ the distribution is asymmetric, with mean $\langle n \rangle = Np_1$, but if is large the shape of the envelope around the maximum looks more and more symmetrical and tends towards a Gaussian distribution - an example of the **Central Limit Theorem** at work. More of this later!

1.5.2 The Poisson distribution

The **Poisson distribution** occupies a special place in probability and statistics, and hence in observational astronomy. It is the archetypical distribution for point processes. It is of particular importance in the **detection** of astronomical objects since it describes **photon noise**. It essentially models the distribution of randomly distributed, independent, point-like events, and is commonly taken as the null hypothesis.

It can be derived as a limiting form of the binomial distribution:

The probability of n "successes" of an event of probability p is

$$P_n = C_n^N p^n (1-p)^{N-n} \tag{19}$$

after N trials.

Let us suppose that the probability p is very small, but that in our experiment we allow N to become large, while keeping the mean finite, so that we have a reasonable chance of finding a finite number of successes n. That is we define $\lambda = \langle n \rangle = Np$ and let $p \rightarrow 0$, $N \rightarrow \infty$, while λ =constant. Then,

$$P_n = \frac{N!}{n!(N-n)!} \left(\frac{\lambda}{N}\right)^n \left(1 - \frac{\lambda}{N}\right)^{N-n} \tag{20}$$

(skipping some maths)

Using **Stirling's approximation**, where $x! \rightarrow \sqrt{2\pi}e^{-x}x^{x+1/2}$ when $x \rightarrow \infty$, and letting $N \rightarrow \infty$, we find[5]

$$P_n = \frac{\lambda^n e^{-\lambda}}{n!} \tag{24}$$

This is **Poisson's distribution** for random point processes, discovered by him in 1837.

1.5.3 Moments of the Poisson distribution:

Let's look at the moments of the Poisson distribution:

$$m_i = \langle n^i \rangle = \sum_{n=0}^{\infty} n^i P_n \tag{25}$$

The mean of Poisson's distribution is

$$\begin{aligned}
\langle n \rangle &= \sum_{n=0}^{\infty} n \frac{\lambda^n e^{-\lambda}}{n!} \\
&= \sum_{n=1}^{\infty} \frac{\lambda^n e^{-\lambda}}{(n-1)!} \\
&= \sum_{n=1}^{\infty} \frac{\lambda \lambda^{n-1} e^{-\lambda}}{(n-1)!} \\
&= \lambda
\end{aligned} \tag{26}$$

i.e. the expectation value of Poisson's distribution is the factor λ. This makes sense since λ was defined as the mean of the underlying Binomial distribution, and kept constant when we took the limit. Now lets look at the second centred moment (i.e. the variance):

when we took the limit. Now lets look at the second centred moment (i.e. the variance):

$$\begin{aligned}
\mu_2 &= \sum_{n=0}^{\infty}(n-\langle n\rangle)^2\frac{\lambda^n e^{-\lambda}}{n!} \\
&= \sum_{n=0}^{\infty} n^2\frac{\lambda^n e^{-\lambda}}{n!} - \lambda^2 \\
&= \sum_{n=1}^{\infty} \frac{n\lambda^n e^{-\lambda}}{(n-1)!} - \lambda^2 \\
&= \sum_{n=1}^{\infty} \frac{n\lambda\lambda^{n-1} e^{-\lambda}}{(n-1)!} - \lambda^2 \\
&= \sum_{n=0}^{\infty} \frac{(n+1)\lambda\lambda^n e^{-\lambda}}{n!} - \lambda^2 \\
&= (\lambda+1)\lambda - \lambda^2 \\
&= \lambda.
\end{aligned} \tag{27}$$

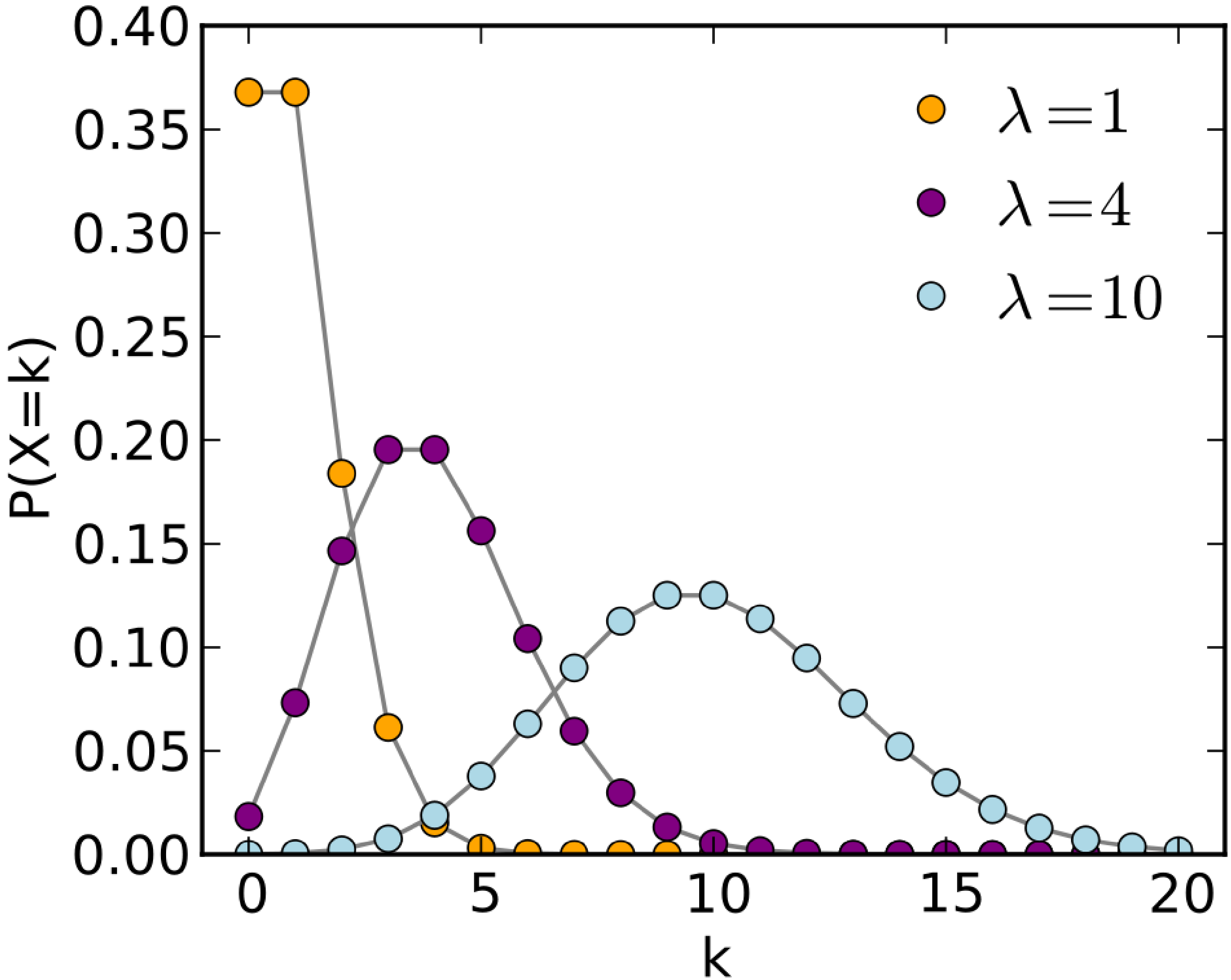
So the variance of Poisson's distribution is also λ. **This means that the variance of the Poisson distribution is equal to its mean.** This is a very useful result.

Poisson distribution

$$P_n = \frac{\lambda^n e^{-\lambda}}{n!}$$

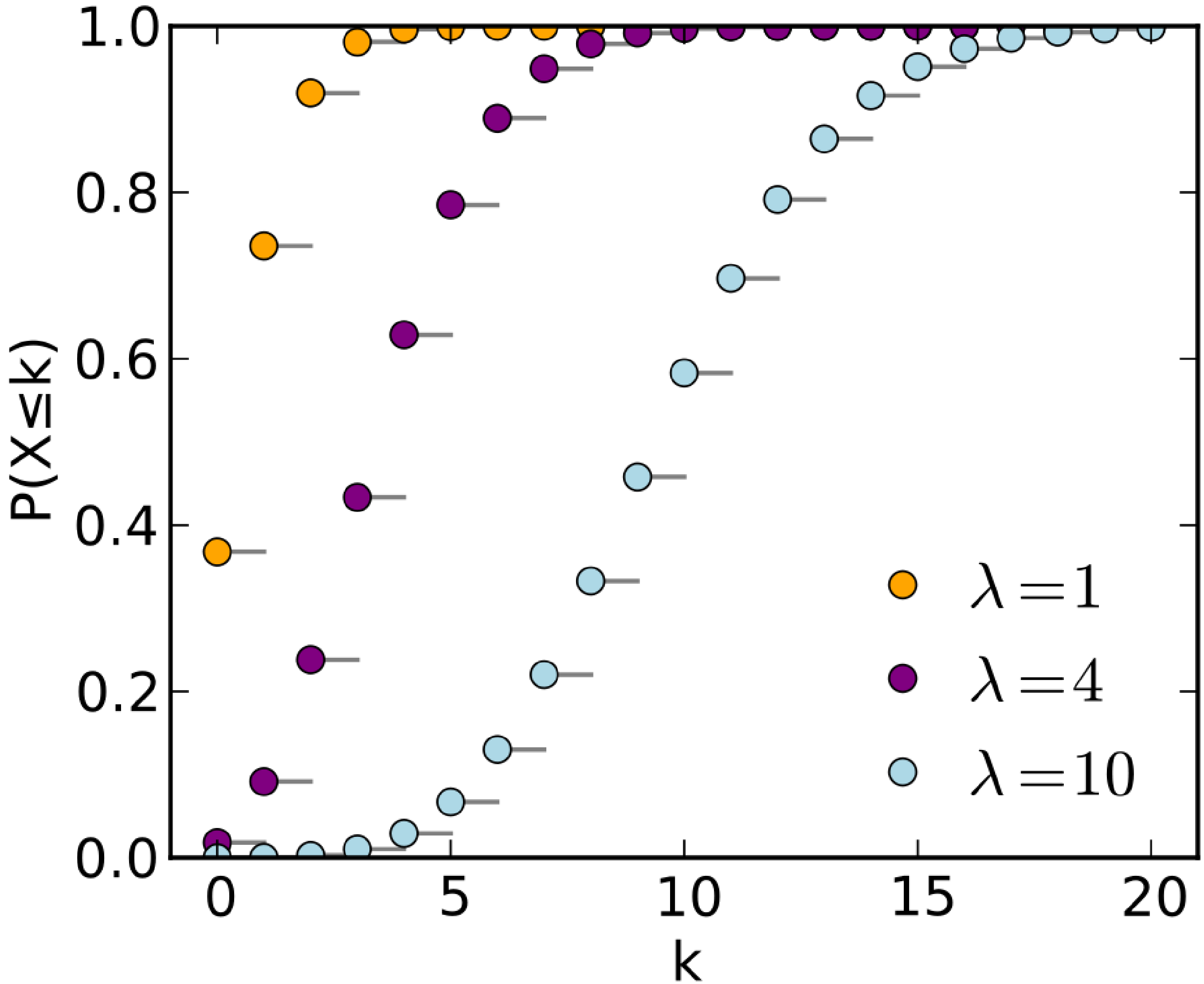
- *n:* number of occurrences of an event
- *λ: expected* (average) number of occurrences, hence the mean of P_n is λ
- variance of P_n is λ
- standard deviation of P_n is square root of λ

Poisson PDF



0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
P(X=k)
0
5
10
15
20
k
$\lambda = 1$
$\lambda = 4$
$\lambda = 10$

Poisson Cumulative Distribution Function



Poisson Noise

- Poisson noise has Poisson distribution
- probability of number of events occurring in constant interval of time/space if events occur with known *average rate* and *independently* of each other
- example: fluctuations in photon flux in finite time intervals Δt: chance to detect k photons with average flux of λ photons

1.5.4 A rule of thumb

Lets see how useful this result is. When counting photons, if the expected number detected is n, the variance of the detected number is n: i.e. we expect typically to detect

$$n \pm \sqrt{n} \tag{28}$$

photons. Hence, just by detecting n counts, we can immediately say that the uncertainty on that measurement is $\pm\sqrt{n}$, without knowing anything else about the problem, and only assuming the counts are random. This is a very useful result, but beware: when stating an uncertainty like this we are assuming the underlying distribution is Gaussian (see later). Only for large n does the Poisson distribution look Gaussian (the Central Limit Theorem at work again), and we can assume the uncertainty $\sigma = \sqrt{n}$.

1.5.5 Detection of a Source

A star produces a large number, $N \gg 1$, of photons during its life. If we observe it with a telescope on Earth we can only intercept a tiny fraction, $p \ll 1$, of the photons which are emitted in all directions by the star, and if we collect those photons for a few minutes or hours we will collect only a tiny fraction of those emitted throughout the life of the star.

So if the star emits N photons in total and we collect a fraction, p, of those, then

$$\begin{aligned} \lambda &= Np \\ N &\to \infty \\ p &\to 0. \end{aligned} \tag{29}$$

So if we make many identical observations of the star and plot out the frequency distribution of the numbers of photons collected each time, we expect to see a Poisson distribution.

Conversely, if we make one observation and detect n photons, we can use the Poisson distribution to derive the probability of all the possible values of λ: we can set confidence limits on the value of λ from this observation. And if we can show that one piece of sky has only a small probability of having a value of λ as low as the surrounding sky, then we can say that we have detected a star, quasar, galaxy or whatever at a particular **significance level** (i.e. at a given probability that we have made a mistake due to the random fluctuations in the arrival rate of photons). A useful rule of thumb here is

$$\lambda_S \geq \lambda_B + \nu\sqrt{\lambda_B} \tag{30}$$

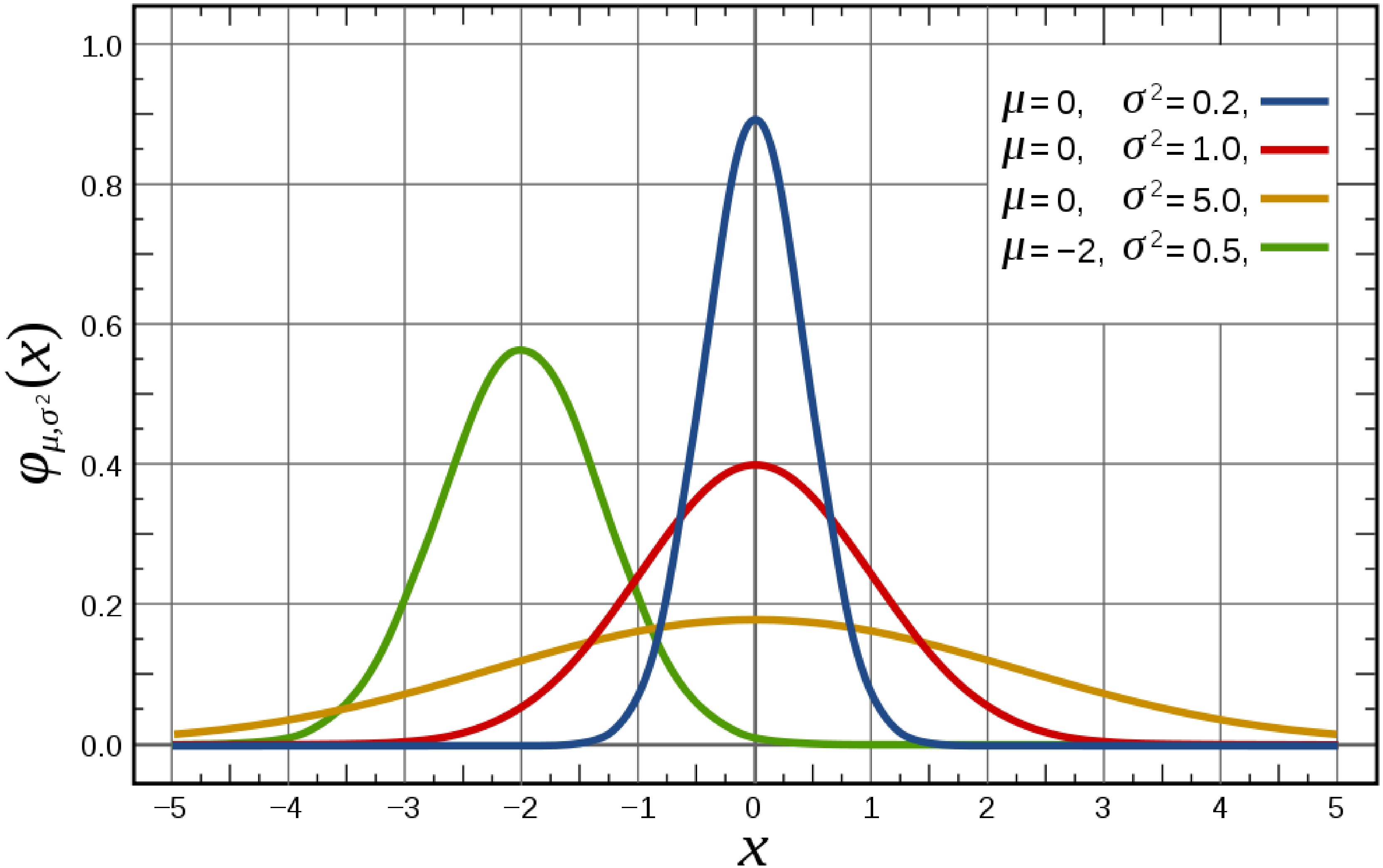
where λ_S is the mean counts from the source, λ_B, is the mean background count, and ν is the detection level. Usually we take $\nu = 3$ to be a detection, but sometimes it can be $\nu = 5$ or even $\nu = 15$ for high confidence in a detection.

1.6.1 The Gaussian Distribution

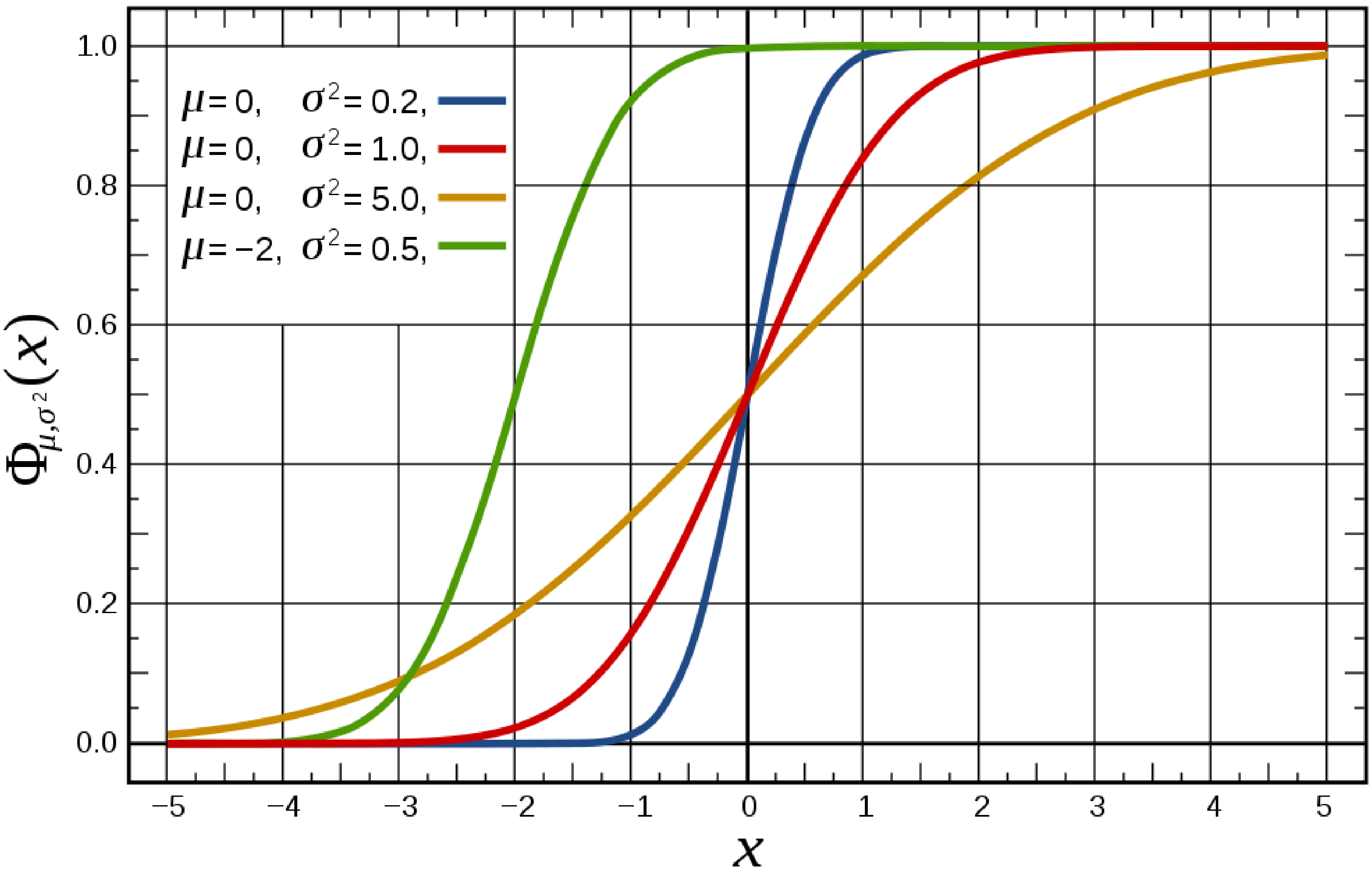
A limiting form of the Poisson distribution is the Gaussian distribution.

$$p(x) = \frac{1}{\sqrt{2p}} e^{-\frac{(x-m)^2}{2s^2}}$$

Normal/Gaussian PDF



Normal Cumulative Distribution



Gaussian Noise

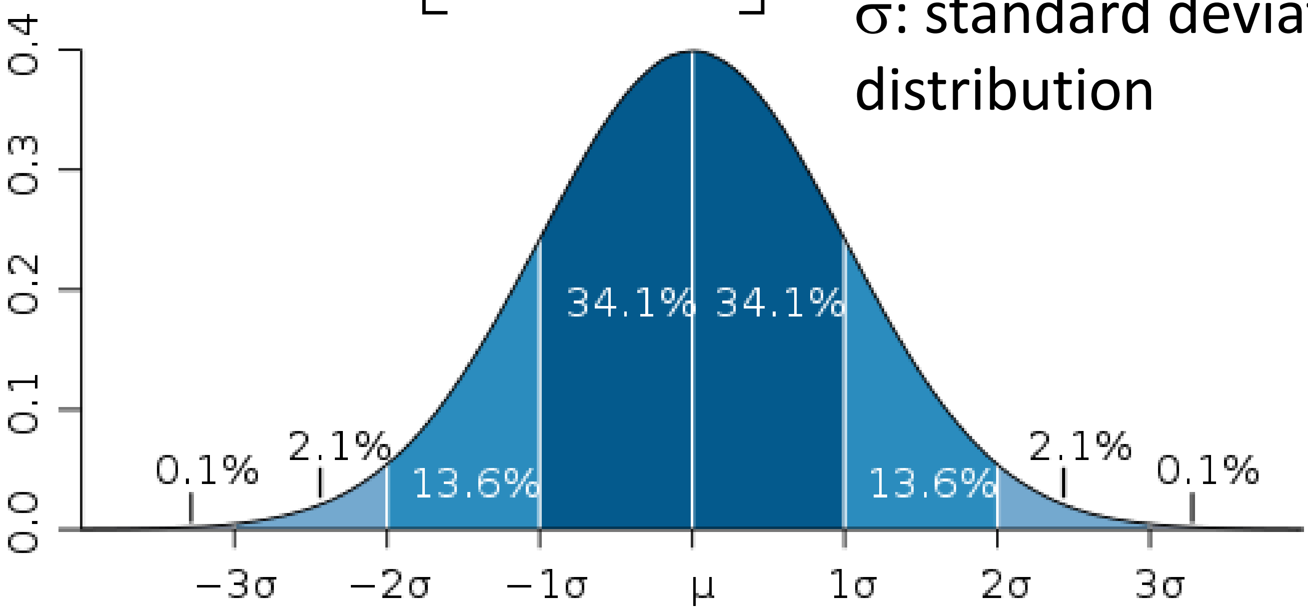
- Gaussian noise has Gaussian (normal) distribution
- Sometimes (incorrectly) called white noise (uncorrelated noise)

$$S = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

x: actual value

μ: mean of distribution

σ: standard deviation of distribution



1-σ ~ 68%

2-σ ~ 95%

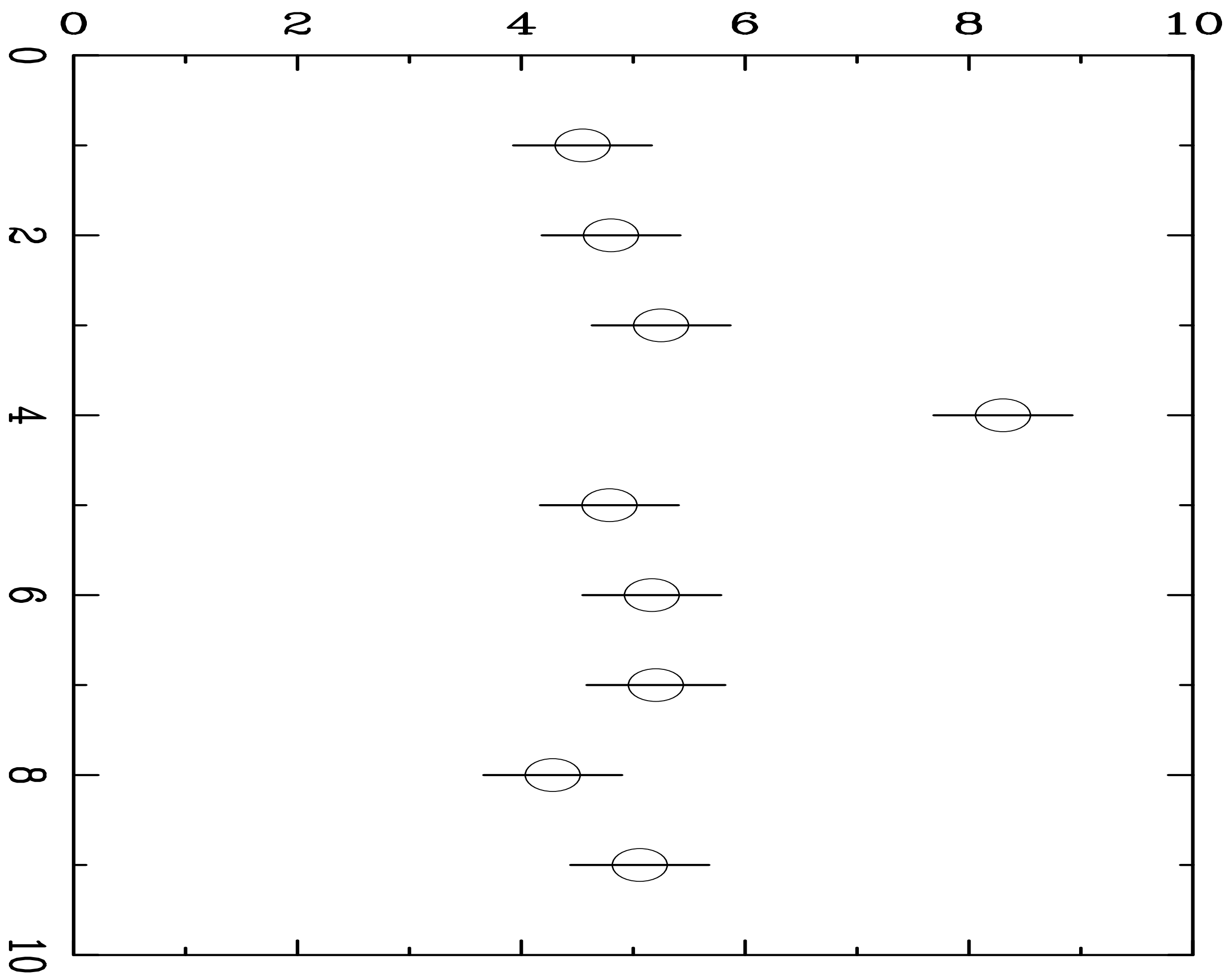
3-σ ~ 99.7%

Astronomers often consider S/N > 3σ or > 5σ as significant

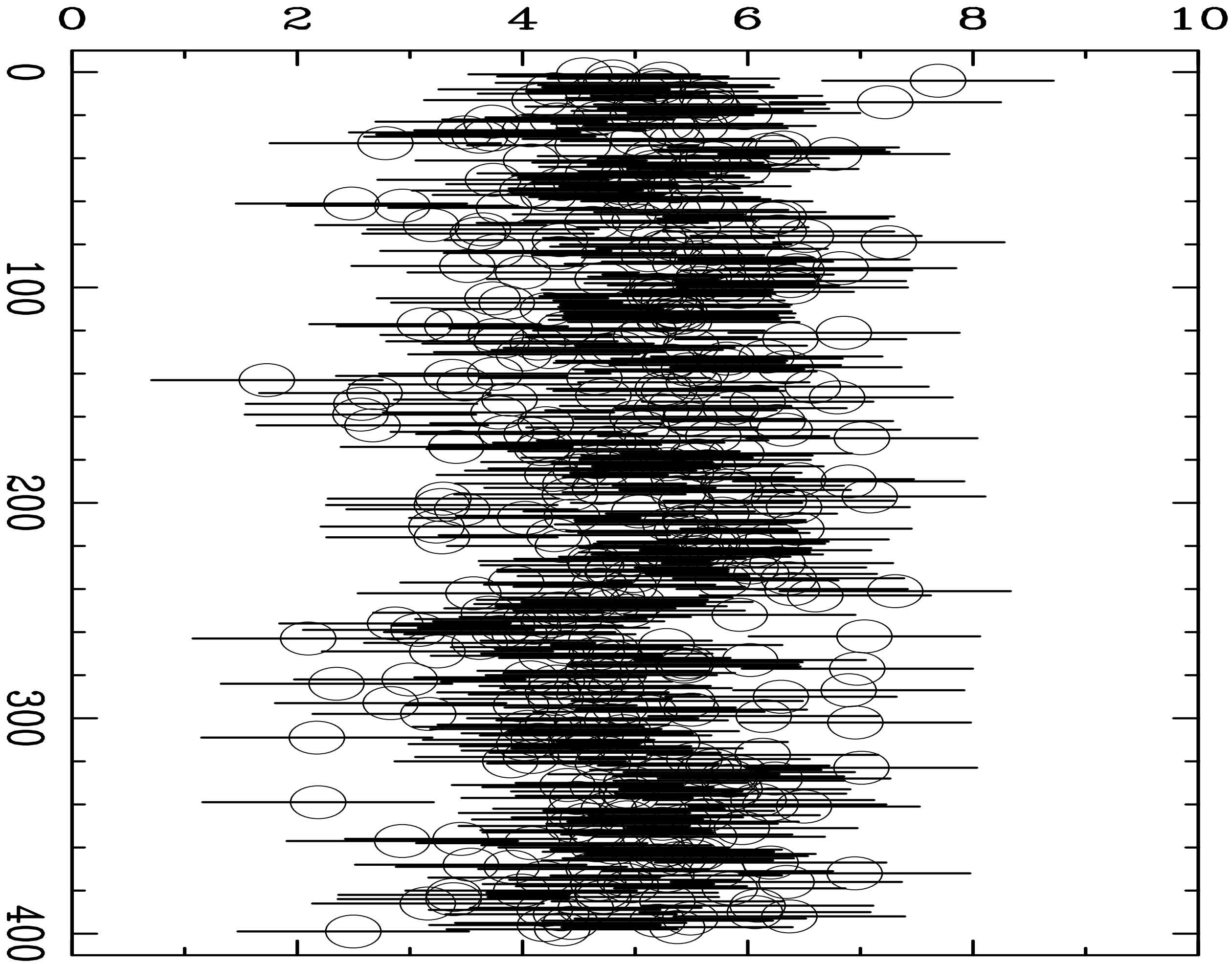
Scientists Like Gauss

- Reasons to like Gaussian distribution:
  - variances of independent Gaussian distributions add
  - Poisson distribution approximates Gaussian for large numbers
  - combinations of different distributions tend to lead to Gaussian distribution
- Always check whether the distribution is Gaussian

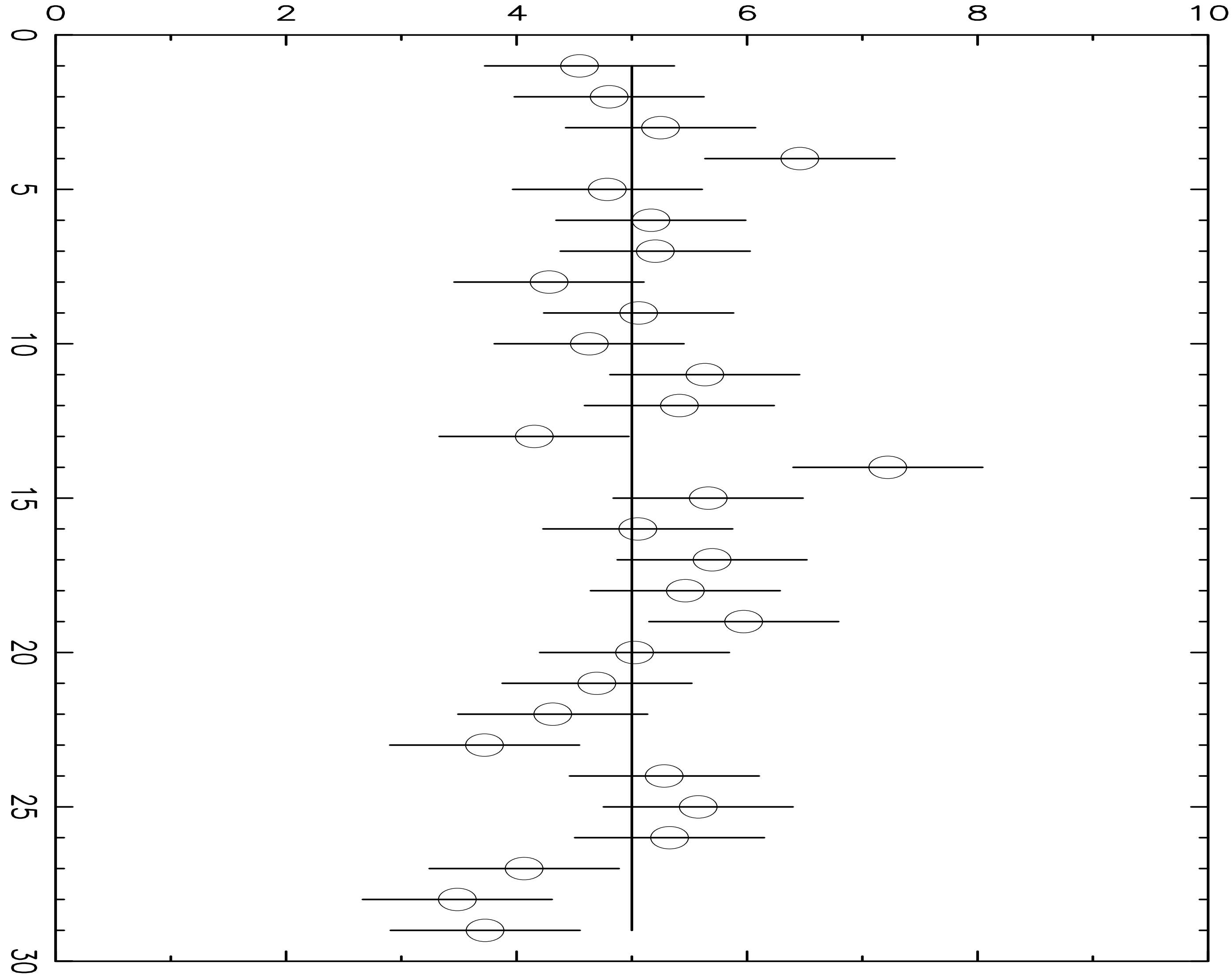
Signal or Noise?



Signal or Noise?



Correct Statistics?

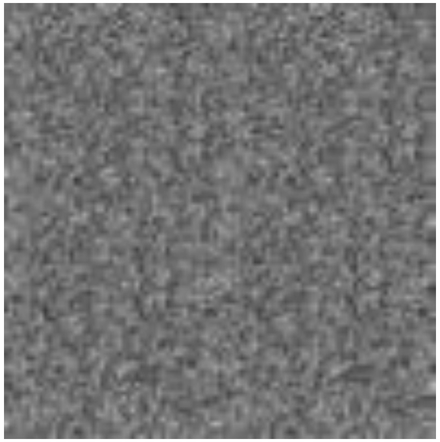


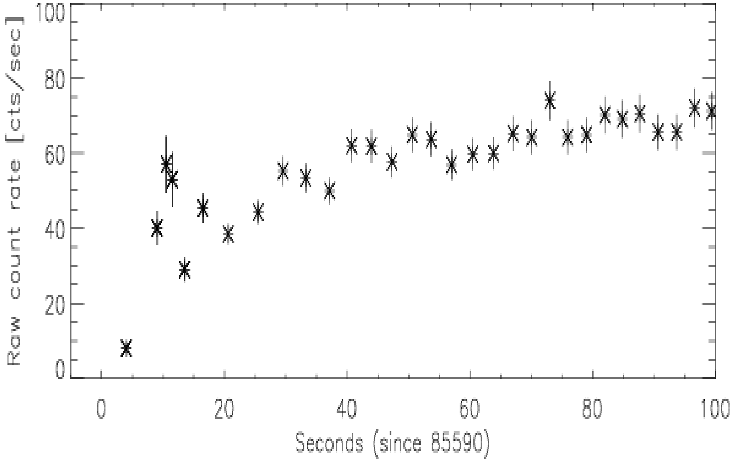
Noise Measurement

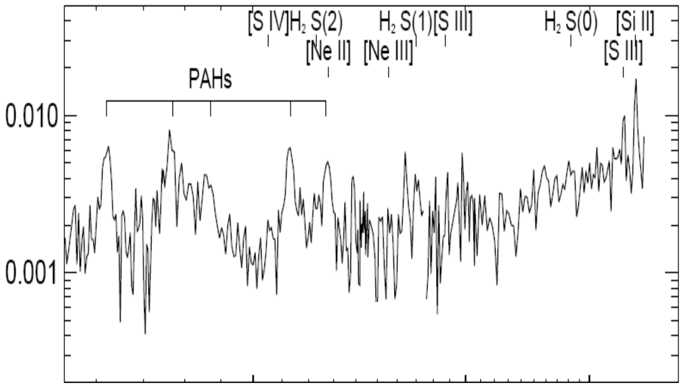
If purely Gauss or Poisson noise distribution, <u>no other systematic</u> noise and no correlations,

then the spatial distribution (neighbouring pixels) of the noise is equivalent to the temporal distribution (successive measurements with one pixel)

This is analogous to throwing 5 dices once versus throwing one dice







Case 1: Spatial noise (detector pixels)

Case 2: Repeated measurements in time (time series)

Case 3: Spectrum (dispersed information)

Poisson Noise and Integration Time

- Integrate light from uniform, extended source on detector
- In finite time interval *Δt*, expect average of λ photons
- Statistical nature of photon arrival rate ➔ some pixels will detect more, some less than λ photons.
- Noise of average signal λ (i.e., between pixels) is $\sqrt{\lambda}$
- Integrate for *2×Δt* → expect average of *2×λ* photons
- Noise of that signal is now $\sqrt{2\lambda}$, i.e., increased by $\sqrt{2}$
- With respect to integration time *t*, noise will only increase $\sim\sqrt{t}$ while signal increases $\sim t$

Signal-to-Noise Ratio: Infrared Imaging

- Measurement:
  - signal (in #photons) from source plus background (in #photons) towards source
- <B>: average sky background (# photons) from source area
- SNR = $S/N = (S + B - \langle B \rangle)/\sqrt{\langle B \rangle}$

SNR and Integration Time

Assuming the signal suffers from Poisson shot noise. Let's calculate the dependence on integration time t_{int}:

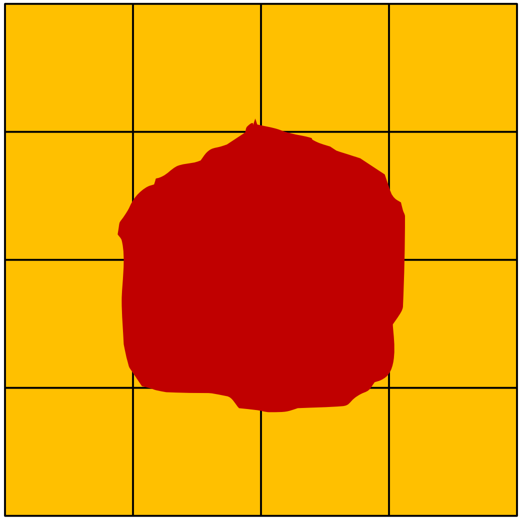
Integrating t_{int}: $$\sigma = \frac{S}{N}$$

Integrating $n \times t_{int}$: $$\sigma = \frac{n \cdot S}{\sqrt{n \cdot B}} \overset{N=\sqrt{B}}{=} \sqrt{n}\,\frac{S}{N} \quad \Rightarrow \quad \frac{S}{N} \propto \sqrt{t_{\text{int}}}$$

Need to integrate four times as long to double the SNR

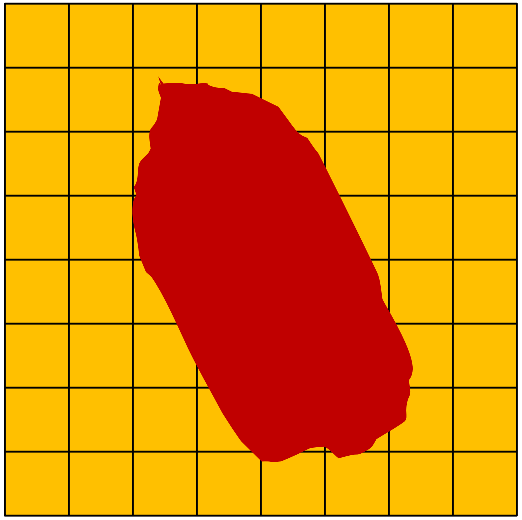
SNR Dependence on Source

Background (=noise)   Target



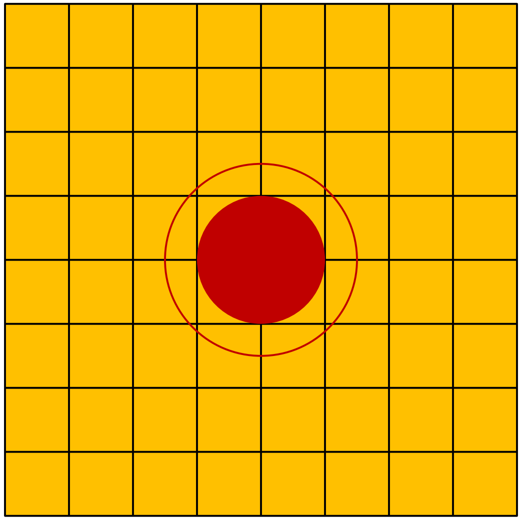
Seeing-limited point source

- pixel size ~ seeing
- PSF ≠ f(D)



Diffraction-limited, extended source

- pixel size ~ diff.lim
- PSF = f(D)
- target >> PSF



Diffraction-limited, point source

- pixel size ~ diff.lim
- PSF = f(D)
- target << PSF

(D: diameter telescope)

Case 1: Seeing-limited "Point Source"

Signal* = S; Background* = B; Noise = N; Telescope diameter = D

θ_{seeing} ~ const

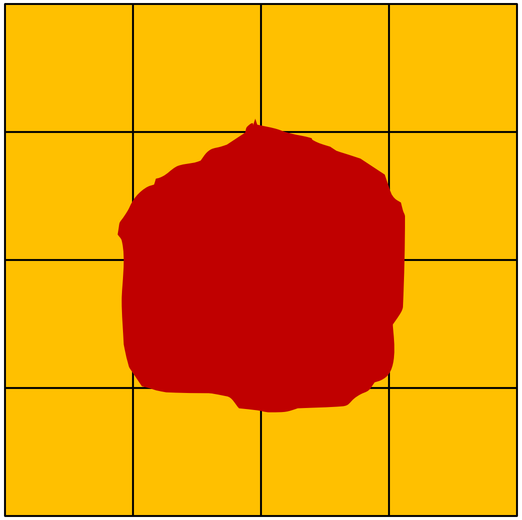
If detector is Nyquist-sampled to θ_{seeing} :

$S \sim D^2$ (area)

$B \sim D^2 \rightarrow N \sim D$ (Poisson std.dev)

$\rightarrow S/N \sim D$

$\rightarrow t_{int} \sim D^{-2}$



*per pixel

Case 2: Diffraction-limited extended Source

Signal* = S;   Background* = B;   Noise = N;   Telescope diameter = D

*per pixel

"Diameter" of PSF ~ const

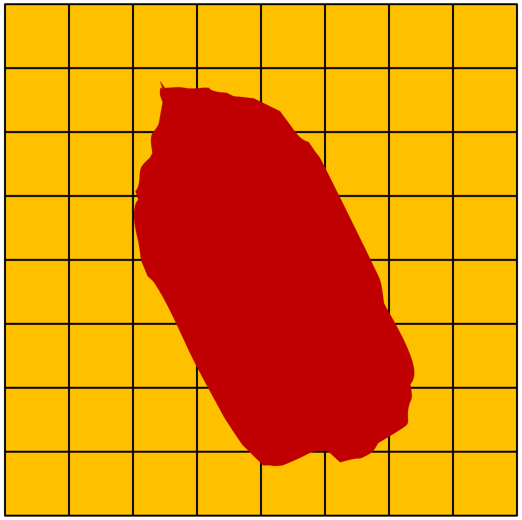
If detector Nyquist sampled to θ_{diff}: pixel ~ D^{-2} but $S \sim D^2$

D^2 (telescope size) and D^{-2} (pixel FOV) cancel each other → no change in signal

same for the background flux

➔ S/N ~ const → t_{int} ~ const

→ *no gain for larger telescopes!*

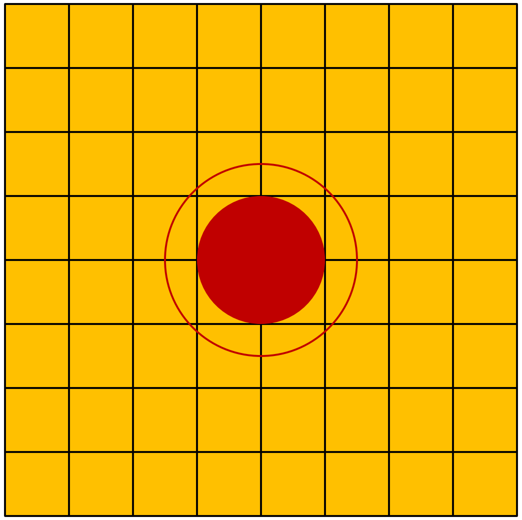


Note:

Total signal / Noise = $\sim D^2 / \sim \sqrt{(D^2)} \sim D$

Case 3: Diffraction-limited "Point Source"

Signal* = S; Background* = B; Noise = N; Telescope diameter = D



"$S/N = (S/N)_{\text{light bucket}} \cdot (S/N)_{\text{pixel scale}}$"

(i) Effect of telescope aperture: → $S/N \sim D$

- Signal $S \sim D^2$
- Background $B \sim D^2$ → $N \sim D$

(ii) Effect of pixel FOV (if Nyquist sampled to θ_{diff}): → $S/N \sim D$

- $S \sim$ const (pixel samples PSF = all source flux)
- $B \sim D^{-2}$ → $N \sim D^{-1}$

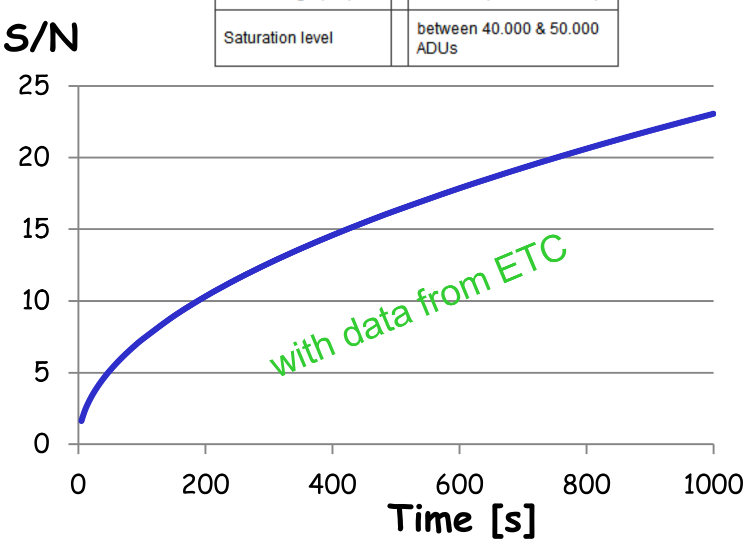
(i) and (ii) combined $S/N \sim D^2$ → $t_{int} \sim D^{-4}$

→ huge gain: 1hr ELT = 3 months VLT

*per pixel

Instrument Sensitivity Example: HAWK-I

Operating temperature	75K, controlled to 1mK
Dark current [e-/s] (at 75K)	between 0.10 & 0.15
Read noise* (DCR)	~ 12 e-
Read noise* (NDR)	~ 5 e-
Linear range (1%)	60.000e- (~30.000 ADUs)
Saturation level	between 40.000 & 50.000 ADUs



http://www.eso.org/observing/etc/bin/ut4/hawki/script/hawkisimu

Input Flux Distribution

Uniform (constant with wavelength)
NOTE: Please use the "Uniform" template spectrum instead of this option.

Template Spectrum: A0V (Pickles) (9480 K) Redshift z = 0.00

Blackbody: Temperature : 15000.00 K

Single Line : Lambda: 1250.000 nm Flux: 50.000 10^{-16} ergs/s/cm^2 (per arcsec^2 for extended sources) FWHM: 1.000 nm

Target Magnitude and Mag.System: K = 20.00 Vega AB

Magnitudes are given per arcsec^2 for extended sources.

Spatial Distribution:
- Point Source
- Extended Source diameter: 1.00 arcsec
- Extended Source (per pixel)

The Magnitude (or flux) is given per arcsec^2 for extended sources.

Sky Conditions

Airmass: 1.20

Seeing: 0.80 arcsec (FWHM in V band)

Instrument Setup

Filter: K

Detector mode: Non-destructive Read-out (NDR)

Results

S/N ratio: S/N = 100.000

Exposure Time: NDIT = 100

DIT = 60.000 sec

Instrument Sensitivity: Example

$$\sigma = \frac{S_{el}}{N_{tot}}$$

detected signal

total noise

number of pixels

background noise

background flux

read noise

dark current

intensity

pixel FOV

...

Detected Signal

Detected signal S_{el} depends on:

- source flux density S_{src} [photons s^{-1} cm^{-2} μm^{-1}]
- Strehl ratio SR (ratio of actual to theoretical maximum intensity)
- spectral bandwidth $\Delta\lambda$ [μm]
- telescope aperture A_{tel} [m^2]
- detector responsivity $\eta_D G$
- transmission of the atmosphere η_{atm}
- total throughput of the system η_{tot}, which includes:
    - reflectivity of all telescope mirrors
    - reflectivity (or transmission) of all instrument components, such as mirrors, lenses, filters, beam splitters, grating efficiencies, slit losses, etc.
- integration time t_{int} [s]

$$S_{el} = S_{src} \cdot SR \cdot \Delta\lambda \cdot A_{tel} \cdot \eta_D G \cdot \eta_{atm} \cdot \eta_{tot} \cdot t_{\text{int}}$$

Total Noise

Total noise N_{tot} depends on:

- number of pixels n_{pix} of one resolution element
- background noise per pixel N_{back}

$$N_{tot} = N_{back}\sqrt{n_{pix}}$$

total background noise N_{back} depends on:

- background flux density S_{back}
- integration time t_{int}
- detector dark current I_d
- pixel read-out noise (N_{read}) and detector frames (n)

$$N_{back} = \sqrt{S_{back} \cdot t_{\mathrm{int}} + I_d \cdot t_{\mathrm{int}} + N_{read}^2 \cdot n}$$

Background Flux Density S_{back}

depends on:

- total background intensity $B_{tot} = \left(B_T + B_A\right) \cdot \eta_{tot}$
  B_T, B_A are thermal emissions from telescope and atmosphere (~black body)
- pixel size and field of view: A x Ω
- detector responsivity: $\eta_D G$
- spectral bandwidth: Δλ

$$S_{back} = B_{tot} \cdot A \times \Omega \cdot \eta_d G \cdot \Delta\lambda$$

Instrument Sensitivity

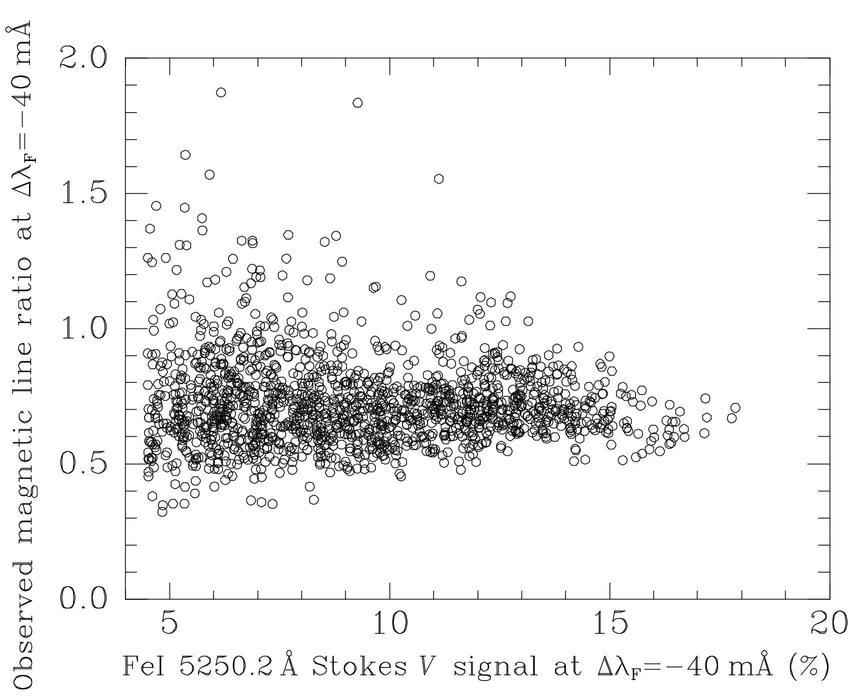
Putting it all together, the total signal to noise ratio for a given experiment is:

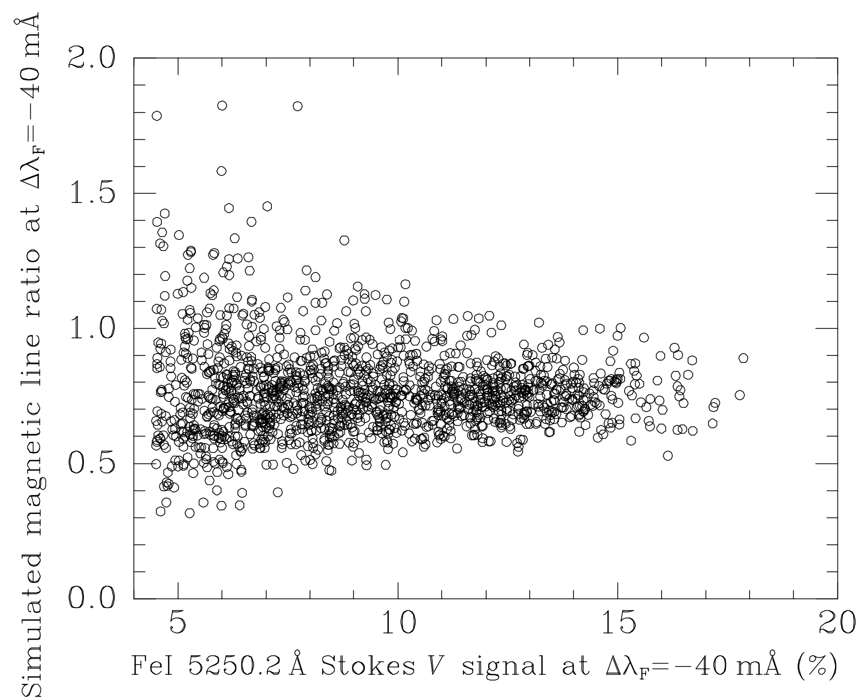
$$\sigma = \frac{S_{el}}{N_{tot}} = \frac{S_{src} \cdot SR \cdot \Delta\lambda \cdot A_{tel} \cdot \eta_D G \cdot \eta_{atm} \cdot \eta_{tot} \cdot t_{\text{int}}}{N_{back}\sqrt{n_{pix}}}$$

$$\Rightarrow \; S_{src} = \frac{\sigma \cdot \sqrt{S_{back} \cdot t_{\text{int}} + I_d \cdot t_{\text{int}} + N_{read}^2 \cdot n} \cdot \sqrt{n_{pix}}}{SR \cdot \Delta\lambda \cdot A_{tel} \cdot \eta_D G \cdot \eta_{atm} \cdot \eta_{tot} \cdot t_{\text{int}}}$$

Noise Propagation: Example

- ratio of two spectral lines as a function of depth of one of the spectral lines
- noise of ratio strongly changes with line depth
- noise affects average ratio





Noise Propagation

- same as error propagation
- function *f(u,v,...)* depends on variables *u,v, ...*
- estimate variance of *f* knowing variances $\sigma_u{}^2$, $\sigma_v{}^2$,... of variables *u,v, ...*

$$\sigma_f^2 = \frac{1}{N-1} \lim_{N\to\infty} \sum_{i=1}^{N} \left(f_i - \bar{f}\right)^2$$

- make assumption / approximation that average of *f* is well approximated by value of *f* for averages of variables: $\bar{f} = f(\bar{u},\bar{v},\text{¼})$

Noise Propagation (cont.)

- Taylor expansion of *f* around average:

$$f_i - \bar{f} \approx (u_i - \bar{u})\frac{\partial f}{\partial u} + (v_i - \bar{v})\frac{\partial f}{\partial v} + \cdots$$

- variance in *f*:

$$\sigma_f^{\,2} \approx \lim_{N\to\infty}\sum_{i=1}^{N}\left[(u_i - \bar{u})\frac{\partial f}{\partial u} + (v_i - \bar{v})\frac{\partial f}{\partial v} + \cdots\right]^2$$

$$= \lim_{N\to\infty}\sum_{i=1}^{N}\left[(u_i - \bar{u})^2\left(\frac{\partial f}{\partial u}\right)^2 + (v_i - \bar{v})^2\left(\frac{\partial f}{\partial v}\right)^2 + 2(u_i - \bar{u})(v_i - \bar{v})\frac{\partial f}{\partial u}\frac{\partial f}{\partial v} + \cdots\right]$$

Noise Propagation (cont.)

- variances of *u* and *v*

$$s_u^2 \equiv \lim_{N\to\infty} \sum_{i=1}^{N} (u_i - \overline{u})^2; \qquad s_v^2 \equiv \lim_{N\to\infty} \sum_{i=1}^{N} (v_i - \overline{v})^2$$

- covariance of *u* and *v* *(can be negative!)*

$$s_{uv}^2 \equiv \lim_{N\to\infty} \sum_{i=1}^{N} (u_i - \overline{u})(v_i - \overline{v})$$

- combine Taylor expansion and these definitions

$$s_f^2 = s_u^2 \left(\frac{\partial f}{\partial u} \right)^2 + s_v^2 \left(\frac{\partial f}{\partial v} \right)^2 + 2 s_{uv}^2 \frac{\partial f}{\partial u} \frac{\partial f}{\partial v} + ¼$$

Noise Propagation (cont.)

- from before

$$s_f^{\ 2} = s_u^{\ 2}\left(\frac{\partial f}{\partial u}\right)^2 + s_v^{\ 2}\left(\frac{\partial f}{\partial v}\right)^2 + 2s_{uv}^{\ 2}\frac{\partial f}{\partial u}\frac{\partial f}{\partial v} + ¼$$

- if differences $(u_i - \bar{u})$ and $(v_i - \bar{v})$ not correlated ➔ sign of product as often positive as negative ➔ covariance small compared to other terms
- if differences are correlated ➔ most products $(u_i - \bar{u})(v_i - \bar{v})$ have the same sign ➔ cross-correlation term can be large