Hello!

Statistics

Ajit Kembhavi

Kaustubh Vaghmare Sheelu Abraham

Data
Analytics:

Find patterns, correlations, classes, outliers and meaning in data

Domain Knowledge
Mathematics
Statistics

"Of the 76 astronomers indexed in Abell (1982) who flourished from ancient times up to 1850, 49 appear in one or more of the statistical histories of Stiegler, Hald, Pearson and Franklin."

-Virginia Trimble

Gaussian and Poisson Distributions, Linear and Rank Correlations, Maximum Likelihood, χ^2, Kolmogorov-Smirnov Test…

Mean, Variance, Covariance

$$\langle x \rangle = \frac{1}{n}\sum_i x_i$$

Mean

$$\sigma_x^2 = \frac{1}{n}\sum_i (x_i - \langle x \rangle)^2)$$

Variance

$$\mathrm{cov}(x, y) = \frac{1}{n}\sum_i (x_i - \langle x \rangle)(y_i - \langle y \rangle)$$

Covariance

Continuous Distributions

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

Continuous Distribution

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

Normalisation

$$< x >= \int_{-\infty}^{\infty} x\, p(x)\, dx$$

Mean

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - \langle x \rangle)^2\, p(x)\, dx$$

Variance

The Gaussian Distribution

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

$$\langle x \rangle = \mu$$

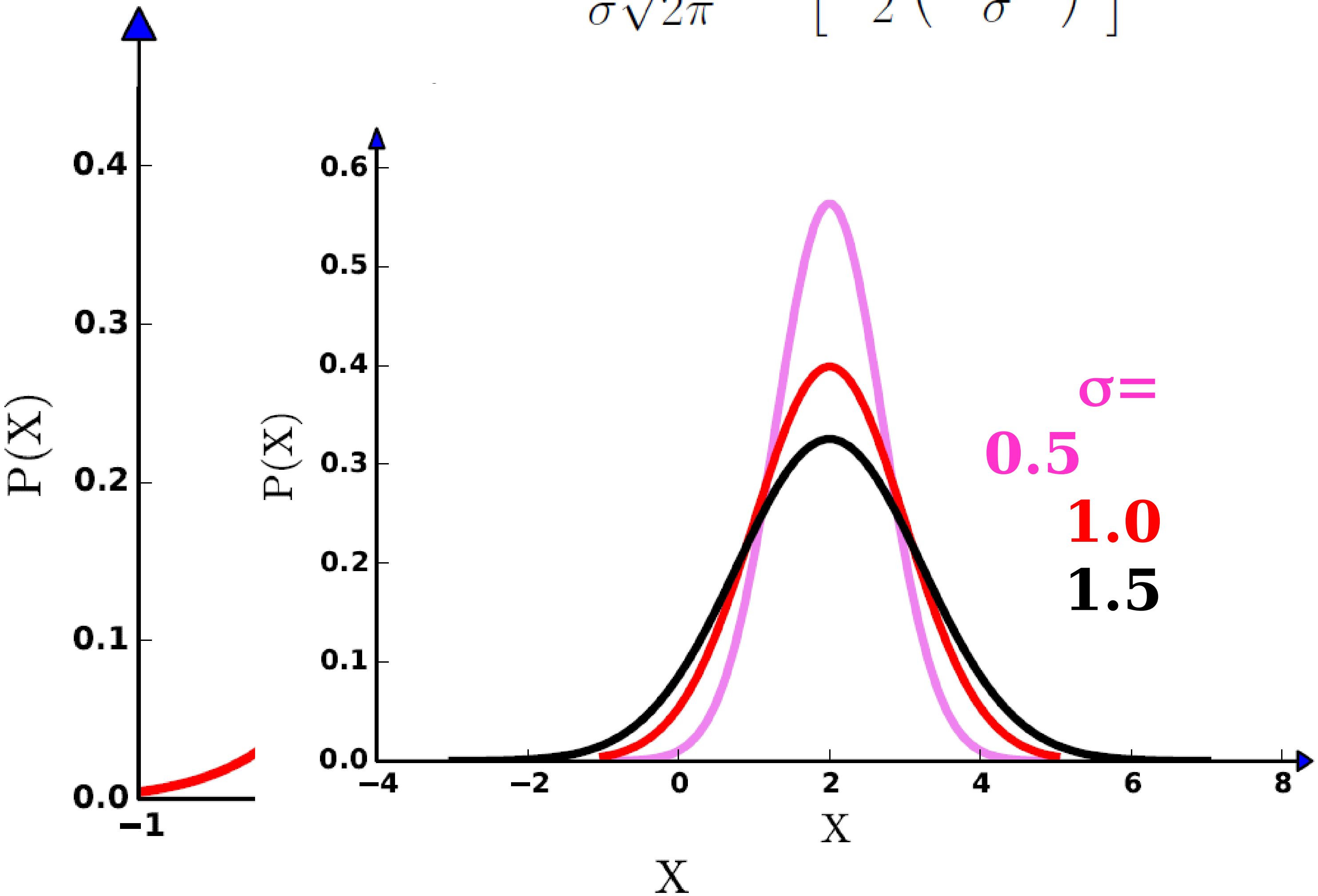
$$\sigma_x^2 = \sigma^2$$

$$\mathrm{FWHM} = 2(2\ln 2)^{1/2}\sigma = 2.354\sigma$$

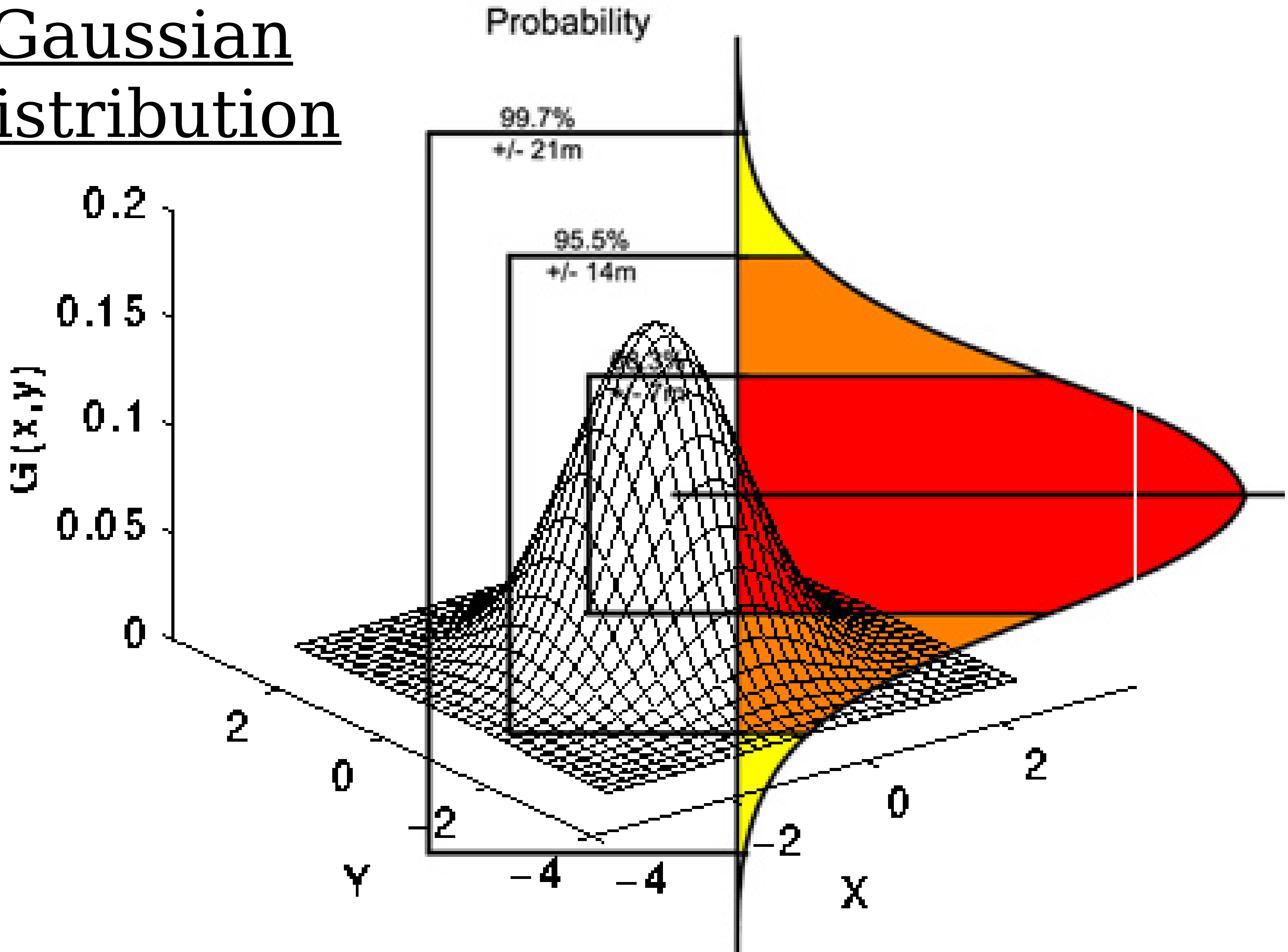
$$\int_{-\infty}^{\infty} G(x)\, dx = 1$$

The Gaussian Distribution

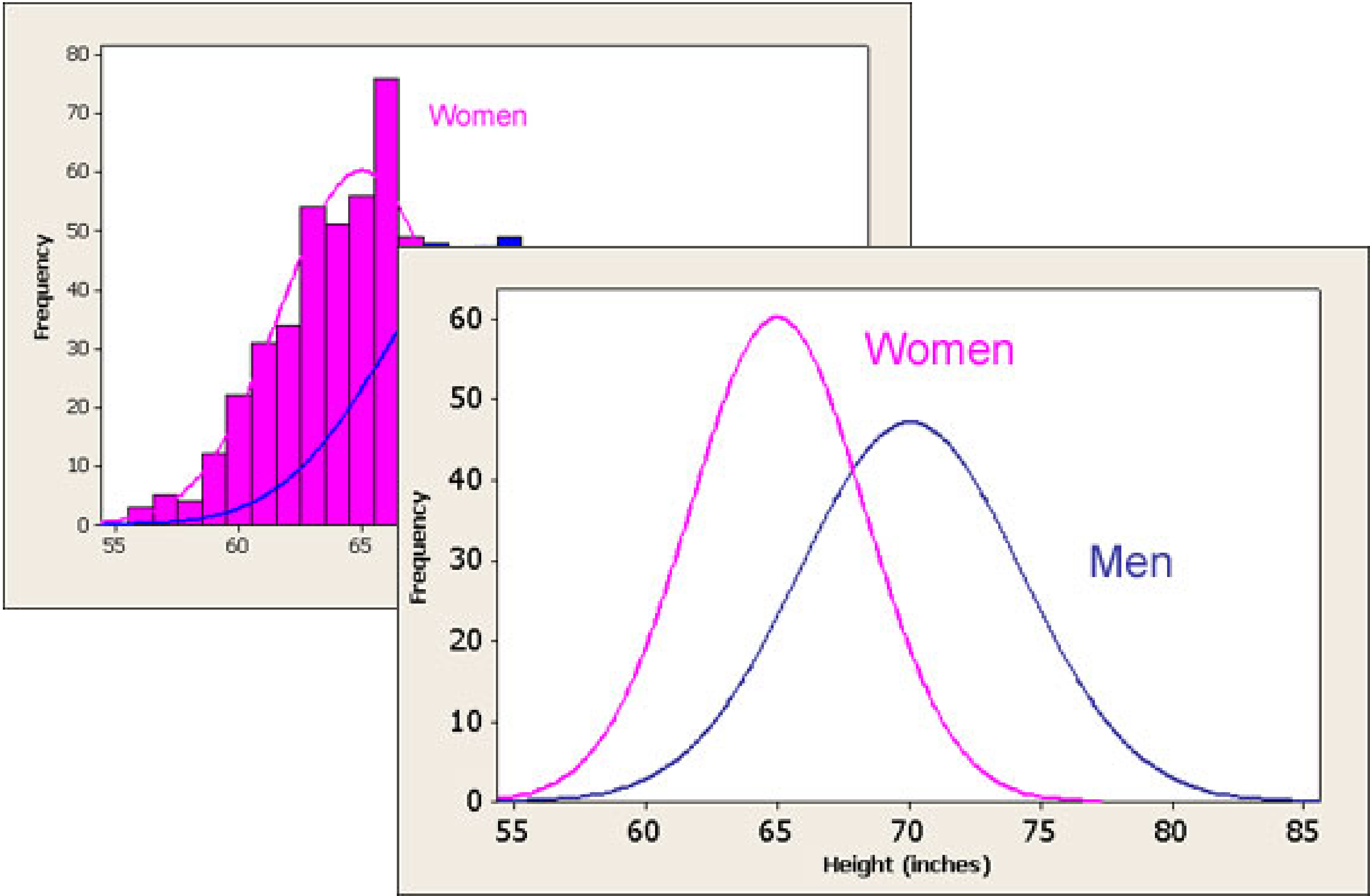
$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$



Gaussian Distribution



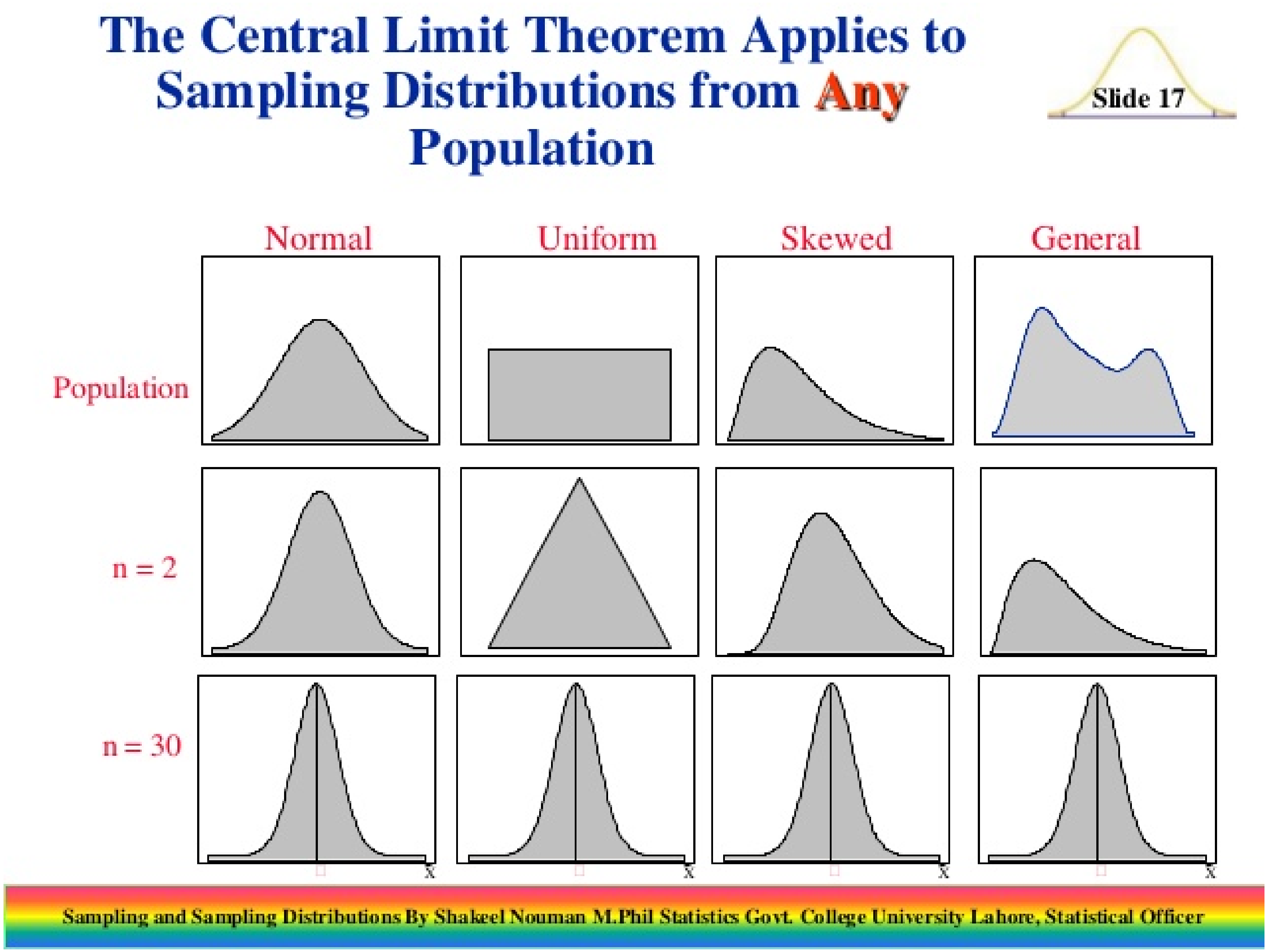
Distribution of Heights



The Central Limit Theorem

When independent random variables are added, their sum tends towards the Gaussian distribution, regardless of the distribution of the original variable.

The Central Limit Theorem



Poisson Distribution

$$P(N) = \frac{e^{-\mu}\mu^N}{N!}$$

Distribution

$$\sum_0^\infty P(N) = 1$$

Normalised to Unity

$$\langle N \rangle = \sum_0^\infty N P(N) = \mu$$

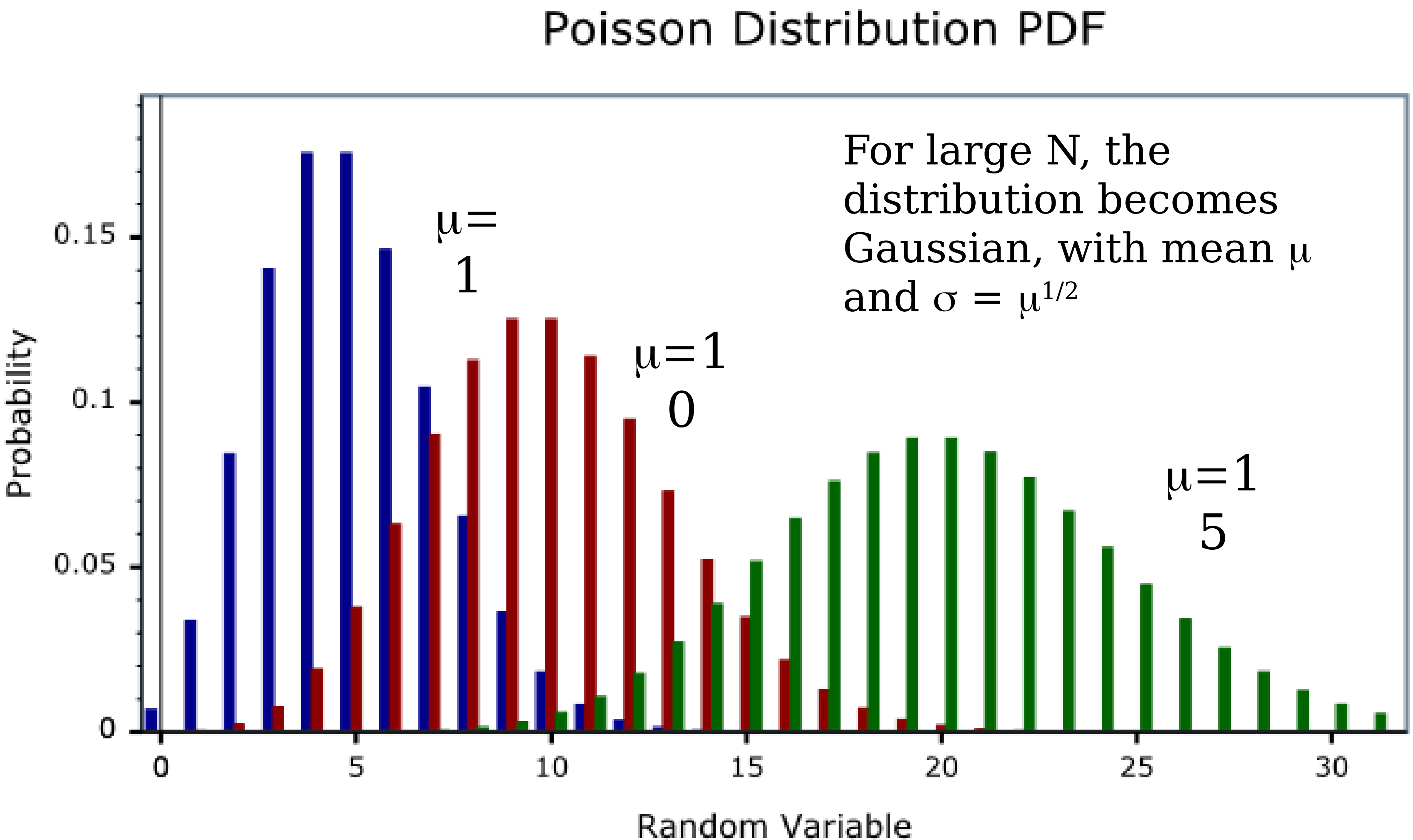
Mean

$$\sigma^2(N) = \sum_0^\infty (N - \langle N \rangle)^2 P(N) = \mu$$

Variance

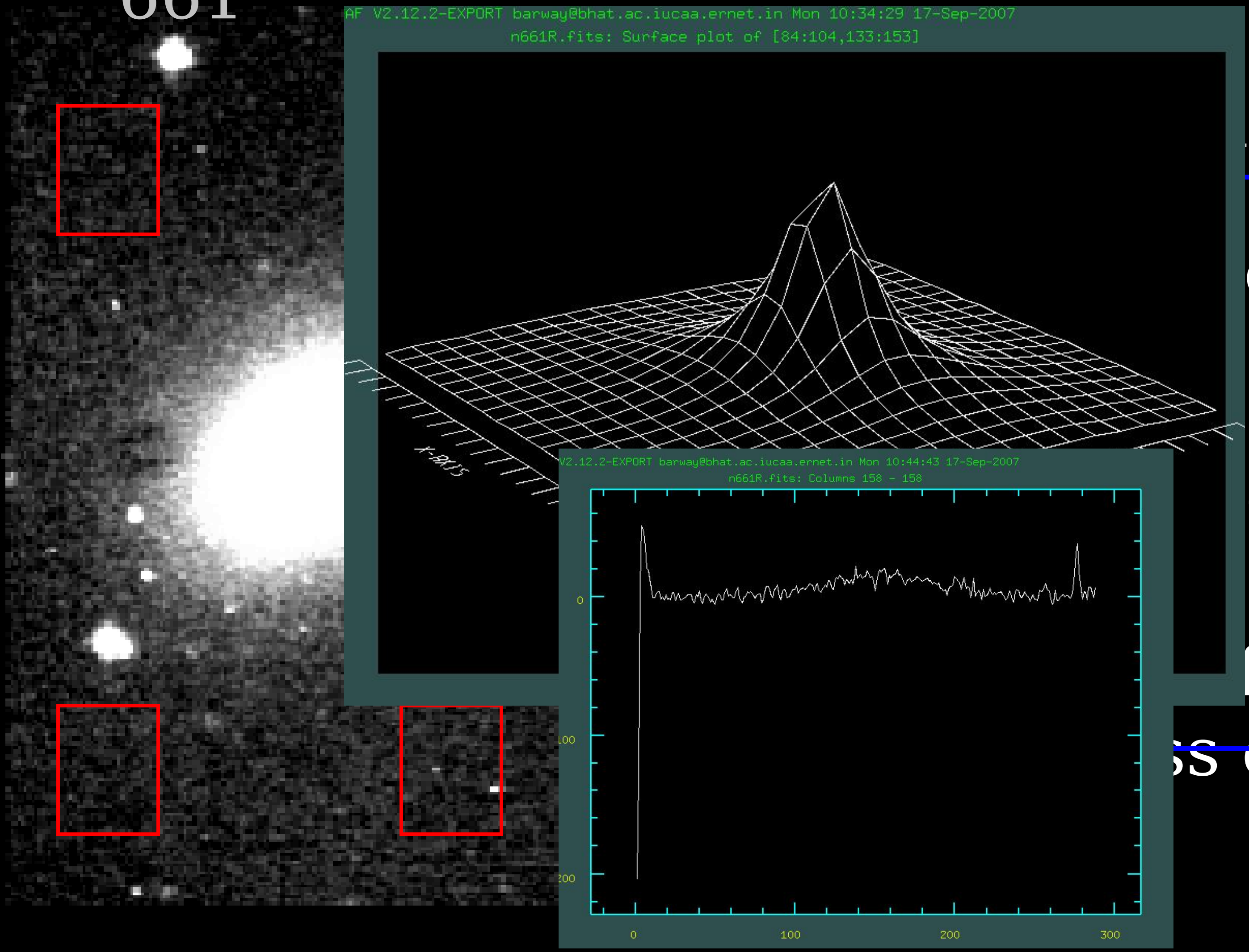
$$\sigma = \sqrt{\mu} \simeq \sqrt{N}$$

rms Deviation



Poisson Distribution PDF
For large N, the distribution becomes Gaussian, with mean μ and $\sigma = \mu^{1/2}$
$\mu=$
1
$\mu=1$
0
$\mu=1$
5
Probability
0.15
0.1
0.05
0
0
5
10
15
20
25
30
Random Variable

NGC
661



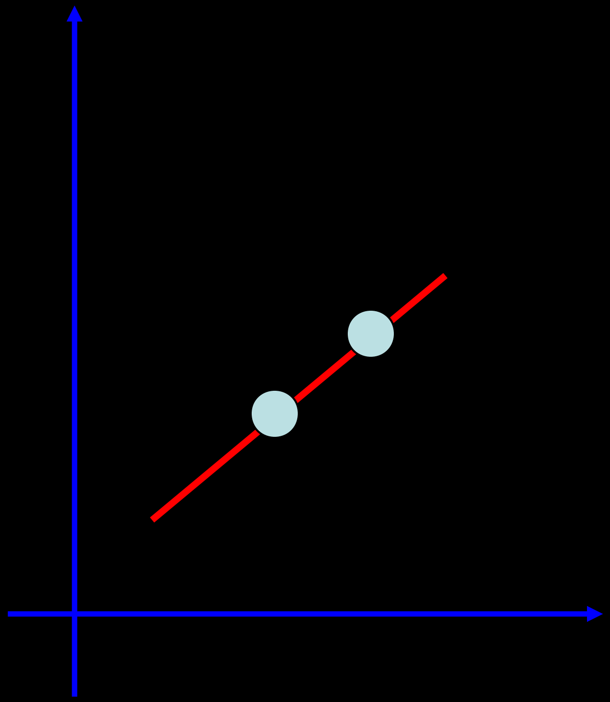
on

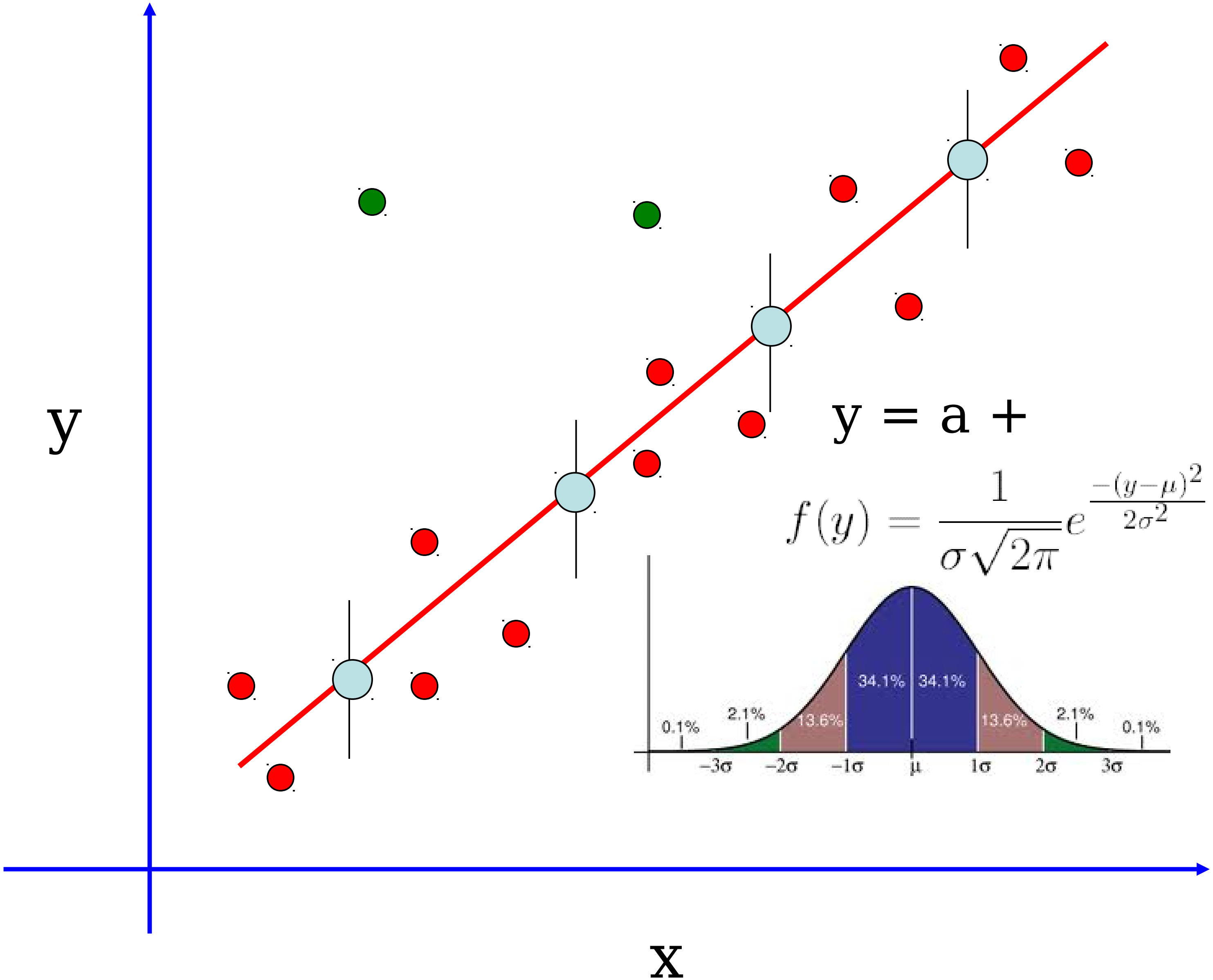
ound
sky

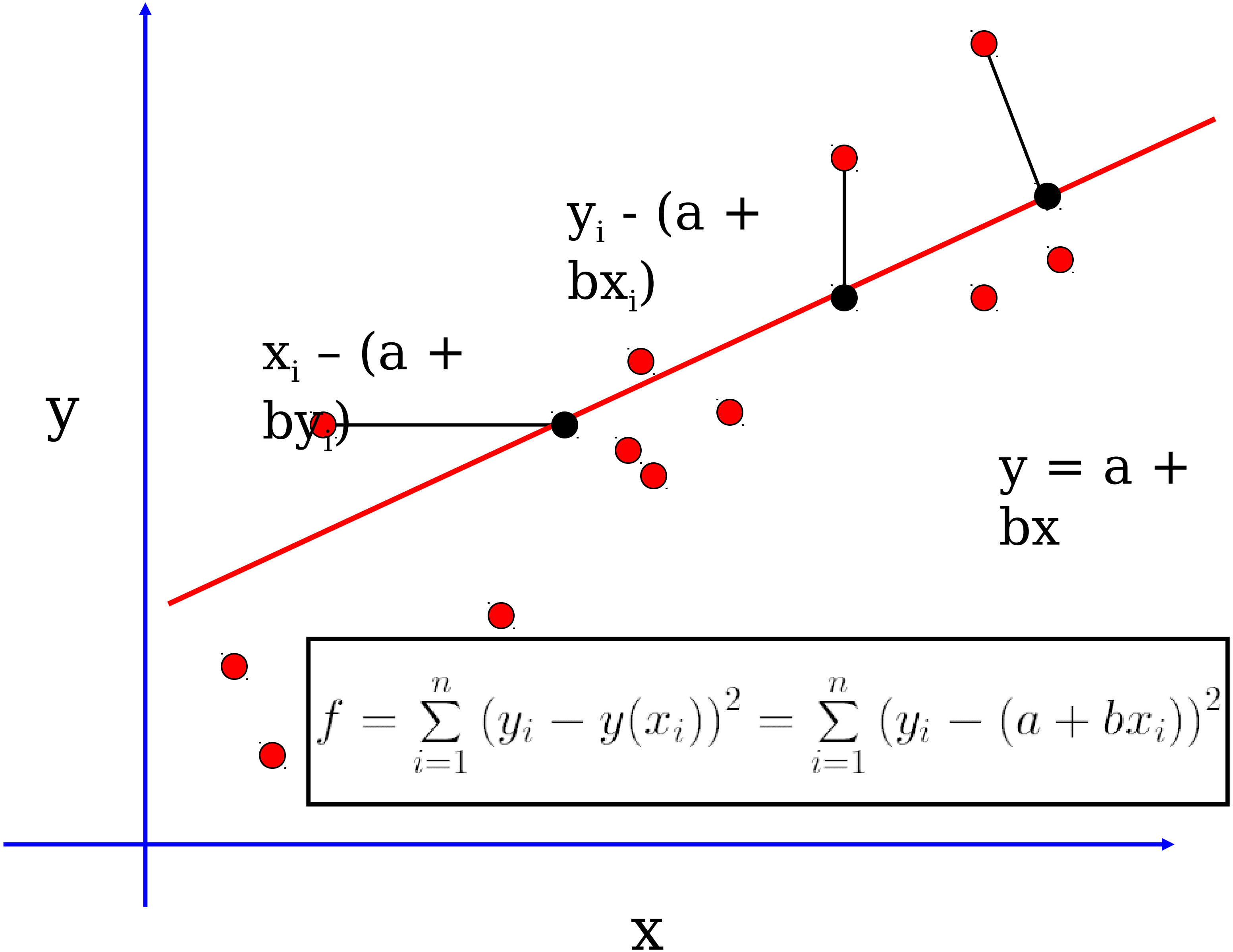
to

face
ss of the

Fitting a Straight Line




y
x
y = a +
$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(y-\mu)^2}{2\sigma^2}}$
34.1% 34.1%
13.6% 13.6%
2.1% 2.1%
0.1% 0.1%
−3σ −2σ −1σ μ 1σ 2σ 3σ


y
x
y_i - (a + bx_i)
x_i – (a + by_i)
y = a + bx
$f = \sum_{i=1}^{n} (y_i - y(x_i))^2 = \sum_{i=1}^{n} (y_i - (a + bx_i))^2$

Least Square Minimisation

$$f = \sum_{i=1}^{n} (y_i - y(x_i))^2 = \sum_{i=1}^{n} (y_i - (a + bx_i))^2$$

$$\frac{\partial f(a,b)}{\partial a} = 0, \quad \frac{\partial f(a,b)}{\partial b} = 0$$

Least Square Fit

y = a + bx

$$b = \frac{1}{\Delta}\left(\sum_i x_i y_i - \sum_i x_i \sum_i y_i\right)$$

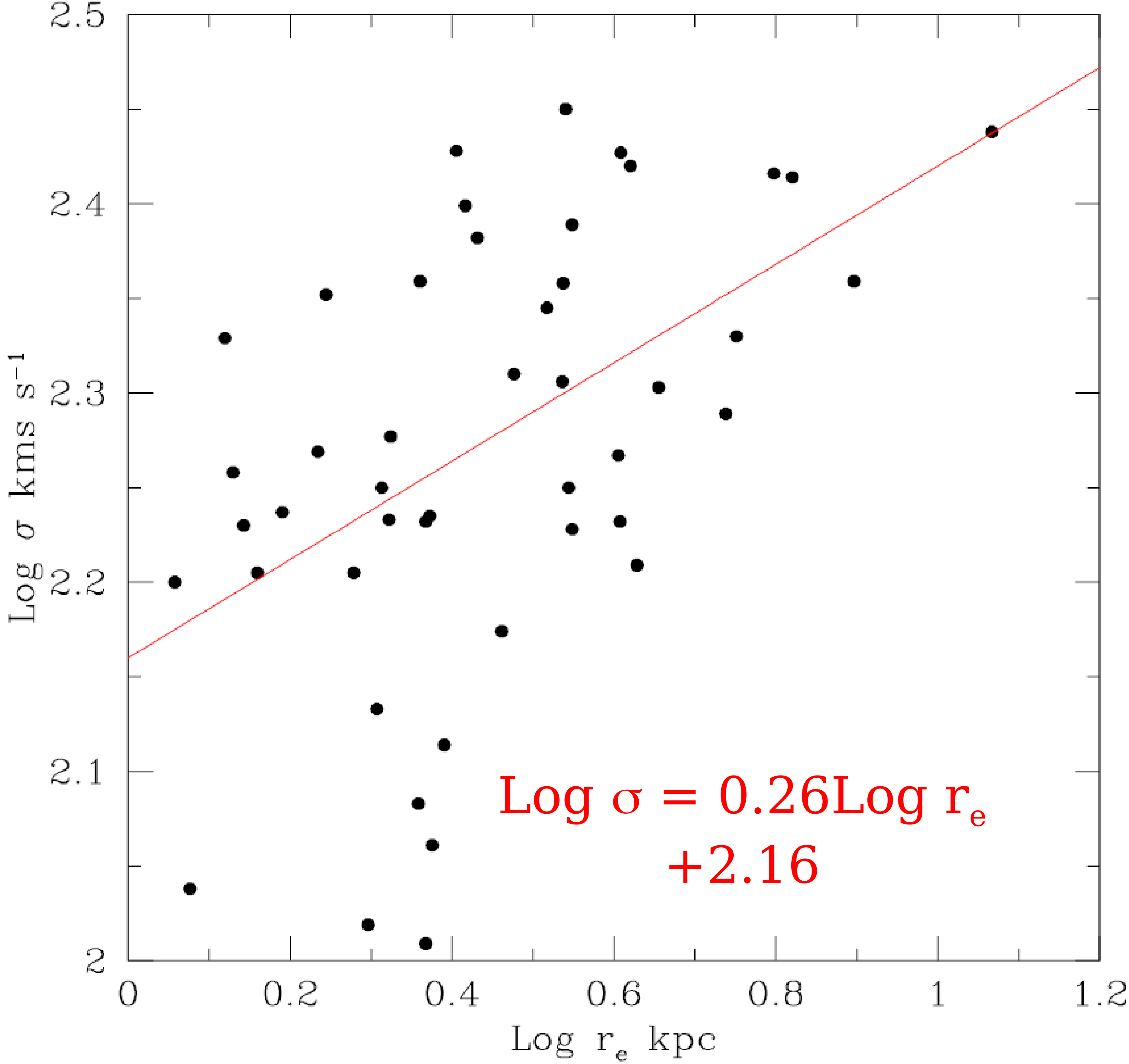
Slope

$$a = \frac{1}{\Delta}\left(\sum_i x_i^2 \sum_i y_i - \sum_i x_i \sum_i x_i y_i\right)$$

Intercept pt

$$\Delta = N \sum_i x_i^2 - \left(\sum_i x_i\right)^2$$

Coma Cluster Galaxies Size-Velocity Dispersion



$$\text{Log}\,\sigma = 0.26\,\text{Log}\,r_e + 2.16$$

Straight Line Fit – Exchanged Variables

$$x = a' + b'y$$

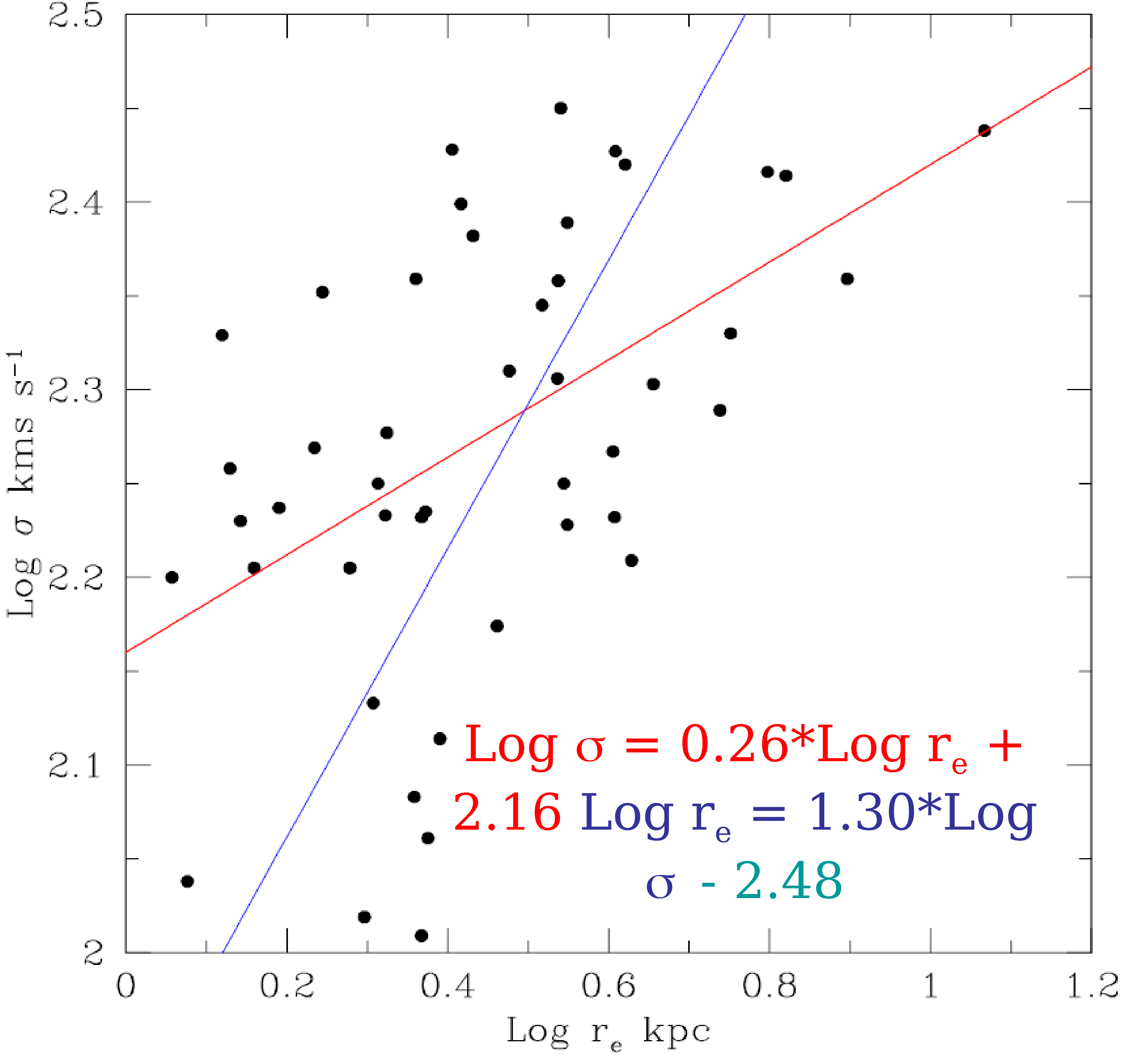
Minimization of the χ^2 function leads to the best fit values

$$b' = \frac{1}{\Delta'}\left(\sum_i x_i y_i - \sum_i x_i \sum_i y_i\right)$$

$$a' = \frac{1}{\Delta'}\left(\sum_i y_i^2 \sum_i x_i - \sum_i y_i \sum_i x_i y_i\right)$$

$$\Delta' = N\sum_i y_i^2 - \left(\sum_i y_i\right)^2$$

Coma Cluster Galaxies Size-Velocity Dispersion



Types of Straight Line Fit

Babu & Feigelson
1991

OLS(y|x)

OLS(x|y)

Bisector

Orthogonal Minimization

$$b_{bis} = \frac{b_1 b_2 - 1 + \sqrt{(1+b_1^2)(1+b_2^2)}}{b_1 + b_2}$$

$$a_{bis} = \langle y \rangle - b_{bis} \langle x \rangle$$

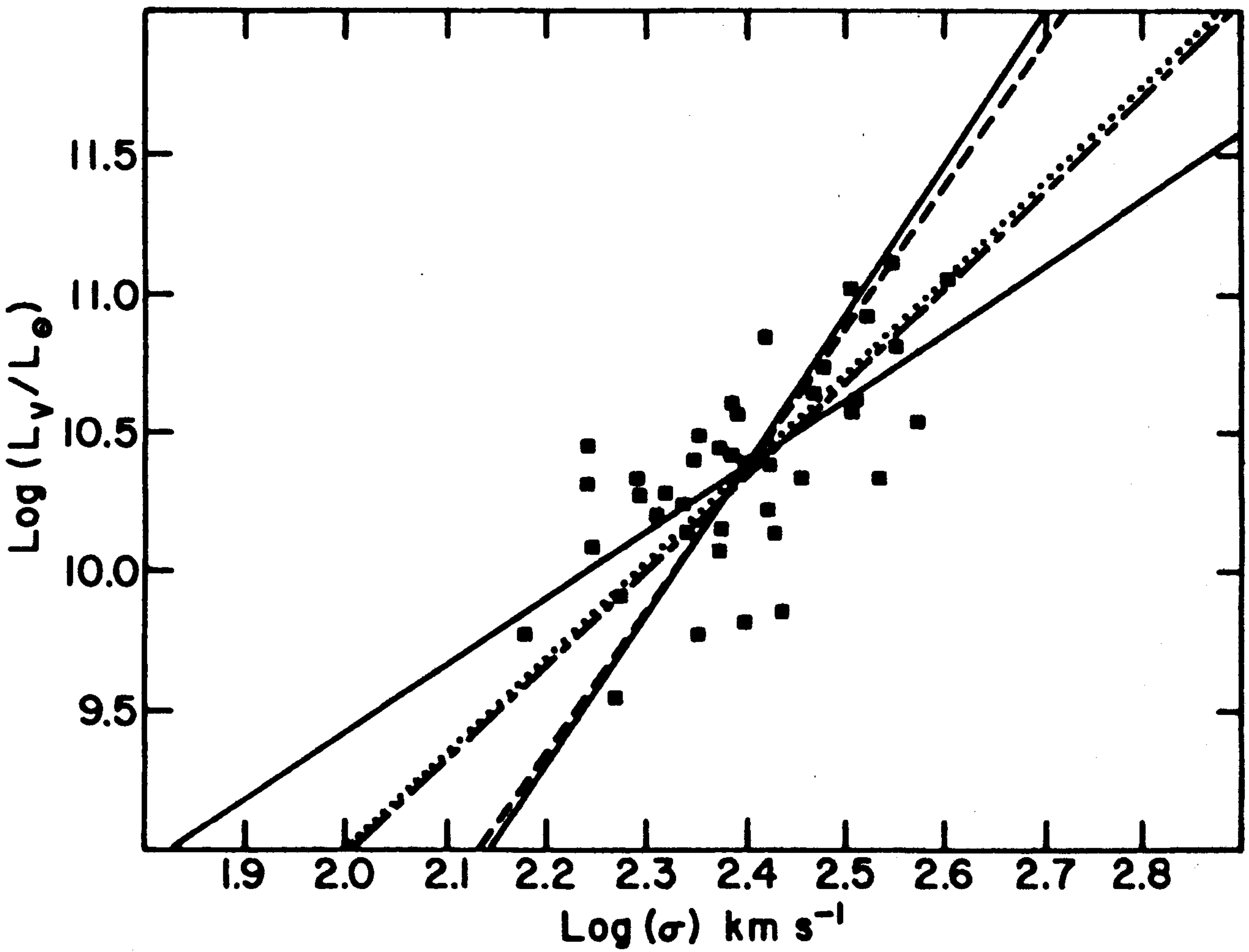


FIG. 2.—Example of a data set with large scatter obtained from Schechter's (1980) measurements of the Faber-Jackson relation in elliptical galaxies. The luminosity is in solar luminosity units. The two solid lines present OLS($Y \mid X$) (*shallowest line*) and OLS($X \mid Y$) (*steepest line*). The dot-dashed line, dashed line, and dotted line represent the OLS bisector, OR, and RMA, respectively.

Correlation Coefficient

The correlation coefficient r is defined as

$$r = \sqrt{bb'} = \frac{1}{\sqrt{\Delta\Delta'}}\left(N\sum_i x_i y_i - \sum_i x_i \sum_i y_i\right)$$

For points along a perfect straight line,

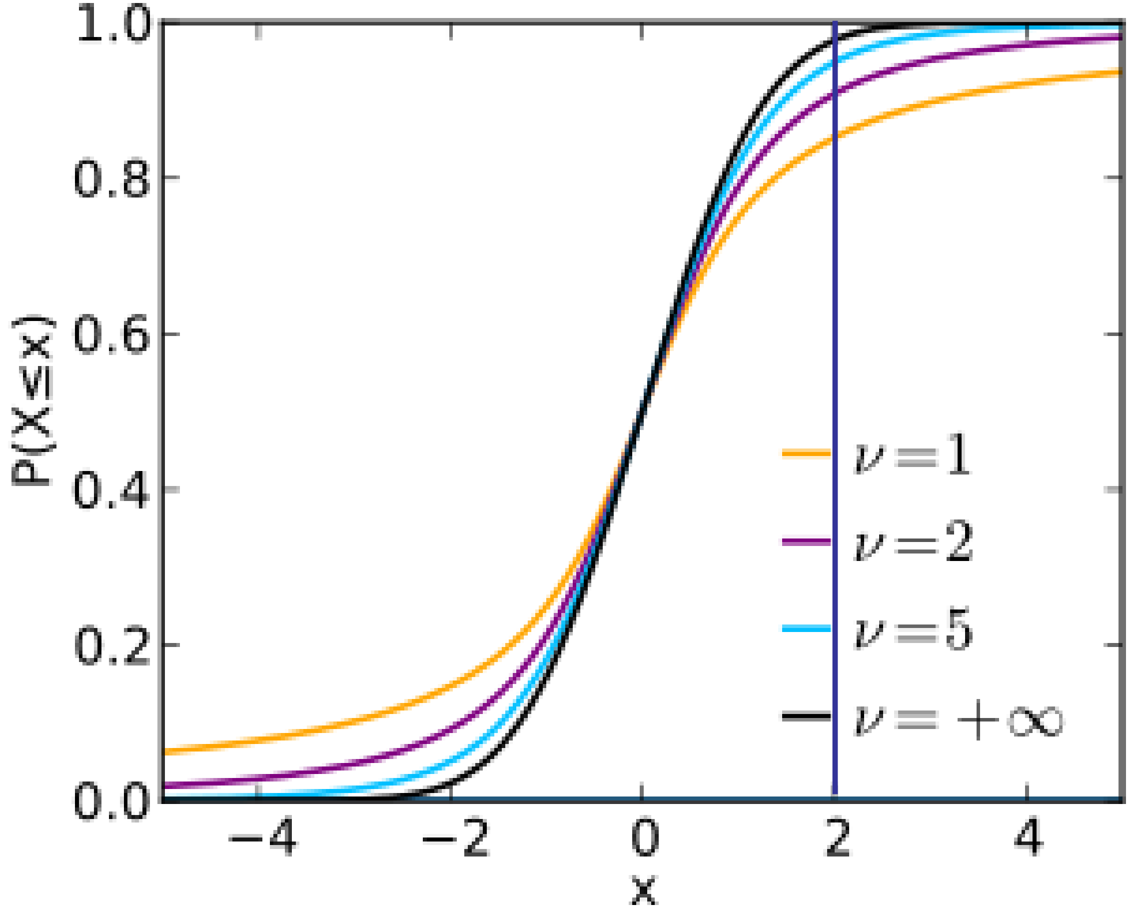
$$r = 1, \quad b' = \frac{1}{b}, \quad a' = -\frac{a}{b}$$

The statistical significance of the correlation can be assessed using the parameter

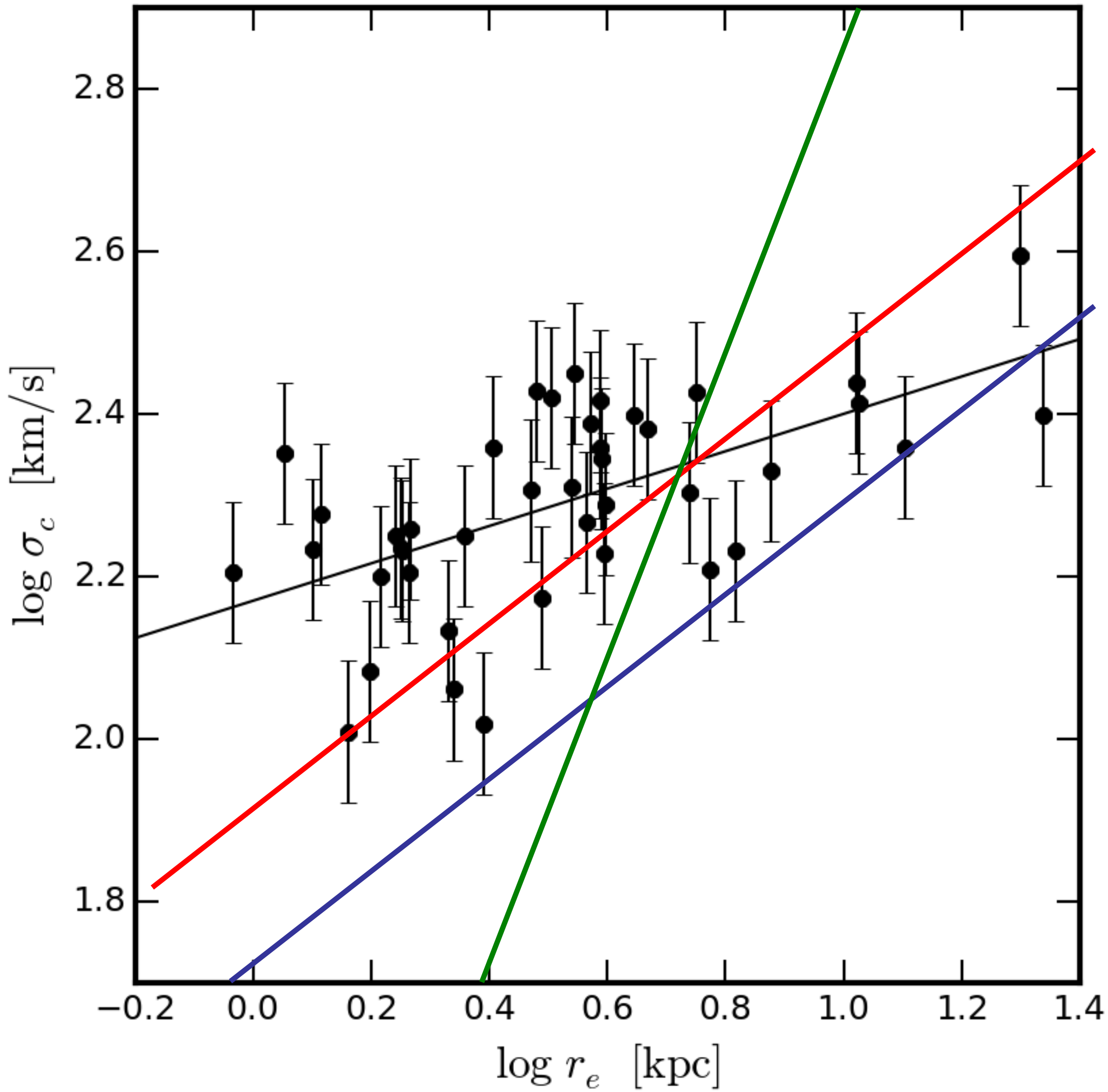
$$t = r\sqrt{\frac{N-2}{1-r^2}}.$$

For an uncorrelated parent population, the statistic t has Student's distribution.

Student’s t - Distribution



Maximum Likelihood Method



$y = a + bx$

$y_{im} = a + bx_i$

Maximum Likelihood Method

Suppose we are fitting N data points (x_i, y_i), $i = 1, \ldots, N$ to a model with M adjustable parameters.

The model predicts a functional relationship between measured independent and dependent variables:

$$y(x_i) = y(x_i; a_1, a_2, \ldots, a_N)$$

Given a particular set of parameters, what is the probability that the data could have occurred?

Maximum Likelihood Method

For Gaussian deviations the probability of obtaining deviations y_i from model values y_{im} = $y(x_i ; a_k)$ is

$$\boxed{P(y_i;a_k) = \prod_{i=1} \frac{1}{\left(\sigma_i\sqrt{2\pi}\right)^N} \exp\left[\frac{-(y_i - y(x_i;a_k))^2}{2\sigma_i^2}\right]}$$

$$\ln P(y_i;a_k) = -\boxed{\sum_i \frac{(y_i - y(x_i;a_k))^2}{2\sigma_i^2}} - N \ln \sum_i \sigma_i \sqrt{2\pi}$$

$$\ln P(y_i;a_k) = -\frac{\chi^2}{2} - N \ln \sum_i \sigma_i \sqrt{2\pi}$$

Minimising χ^2 therefore maximises the probability of obtaining the observed deviations, given the model

χ2 Distribution

$$\boxed{\chi^2 = \sum_{i=1}^{n} \frac{\left(y_i - y(x_i)\right)^2}{\sigma_i^2}}$$

$$y(x_i) = y(x_i; a_1, a_2, \ldots, a_m)$$

$$\boxed{f(\chi^2; \nu) = \frac{(\chi^2)^{(\nu-2)/2} \exp(-\chi^2/2)}{2^{\nu/2}\Gamma(\nu/2)}}$$

ν = n-m

$$\int_0^\infty f(\chi^2, \nu)\, d(\chi^2) = 1$$

$$\left\langle \chi^2 \right\rangle = \nu, \quad \sigma_{\chi^2}^2 = 2\nu$$

The Gamma Function

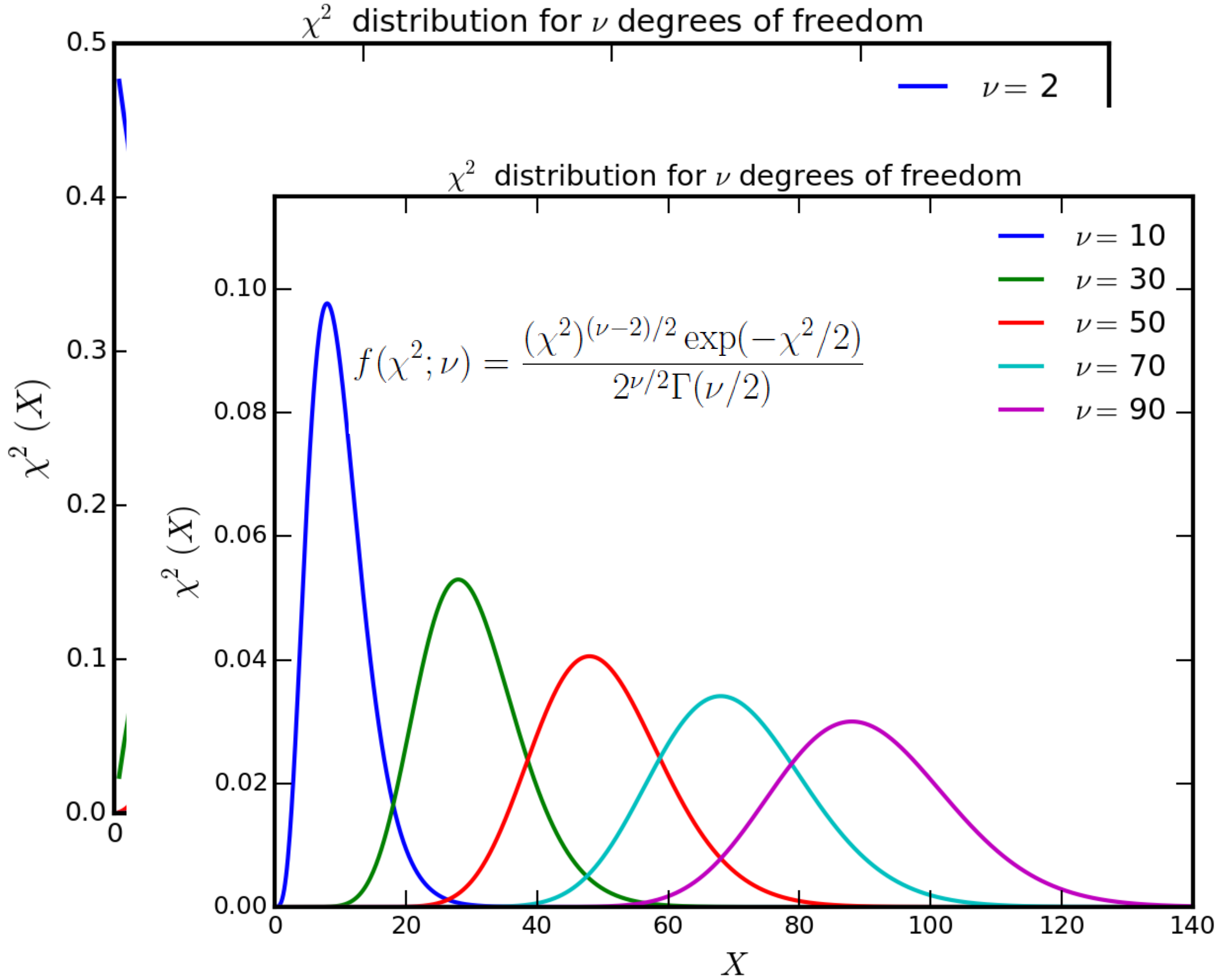
$$\Gamma(n+1) = n\Gamma(n), \quad \Gamma(1) = 1, \quad \Gamma(1/2) = \sqrt{\pi}$$

$$\Gamma(n+1) = n!, \quad n = 0, 1, 2, ...$$

$$\Gamma(n+1) = n(n-1)(n-2)...\frac{3}{2}.\frac{1}{2}\sqrt{\pi}, \quad n = 1/2, 3/2...$$

Reduced χ2

$$\chi^2_\nu = \frac{\chi^2}{\nu}, \quad \langle \chi^2_\nu \rangle = 1, \quad \sigma^2_{\chi^2_\nu} = \frac{2}{\nu}$$



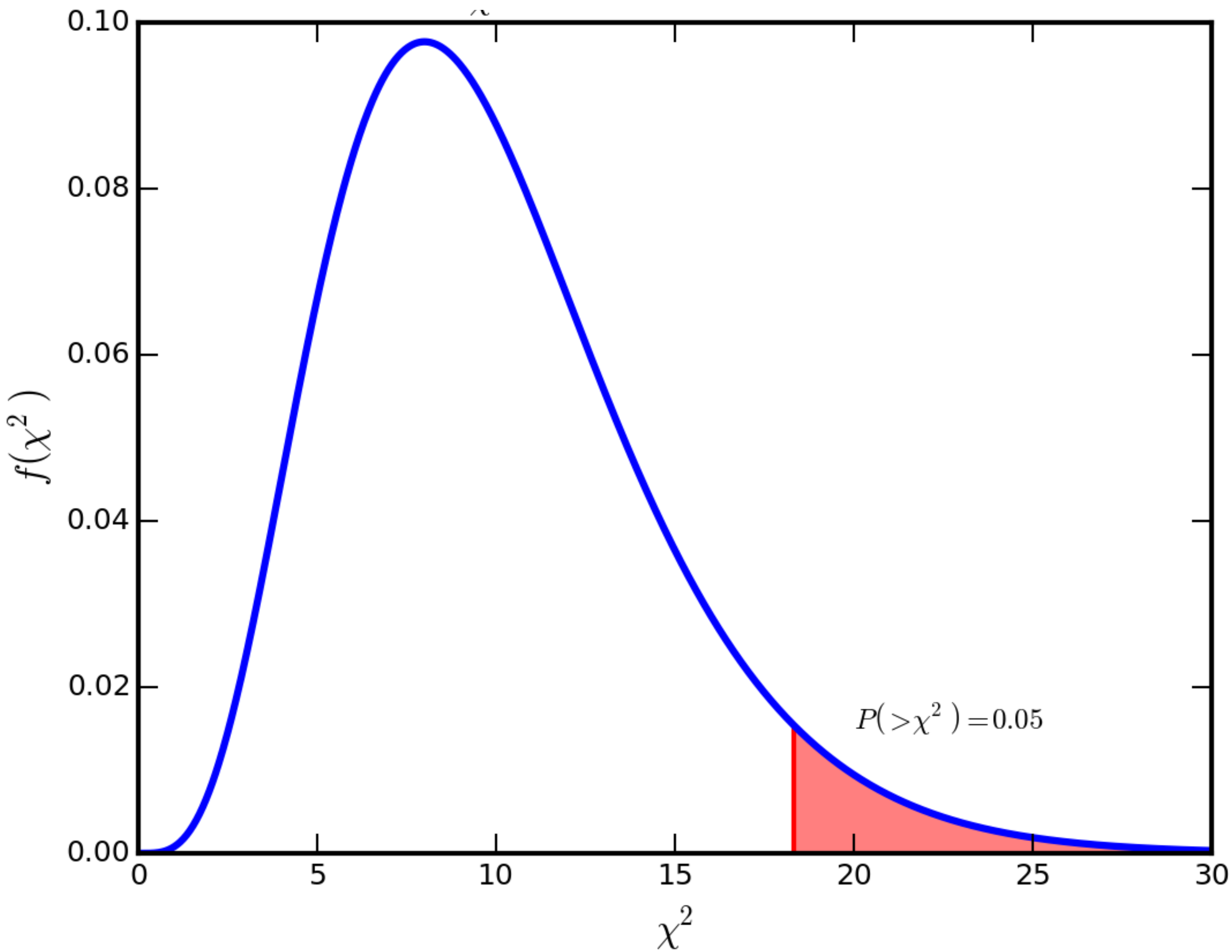
χ^2 distribution for ν degrees of freedom
$\nu = 2$
0.5
0.4
0.3
0.2
0.1
0.0
0
$\chi^2\,(X)$
χ^2 distribution for ν degrees of freedom
$\nu = 10$
$\nu = 30$
$\nu = 50$
$\nu = 70$
$\nu = 90$
$f(\chi^2;\nu) = \dfrac{(\chi^2)^{(\nu-2)/2}\exp(-\chi^2/2)}{2^{\nu/2}\Gamma(\nu/2)}$
0.10
0.08
0.06
0.04
0.02
0.00
0
20
40
60
80
100
120
140
$\chi^2\,(X)$
X

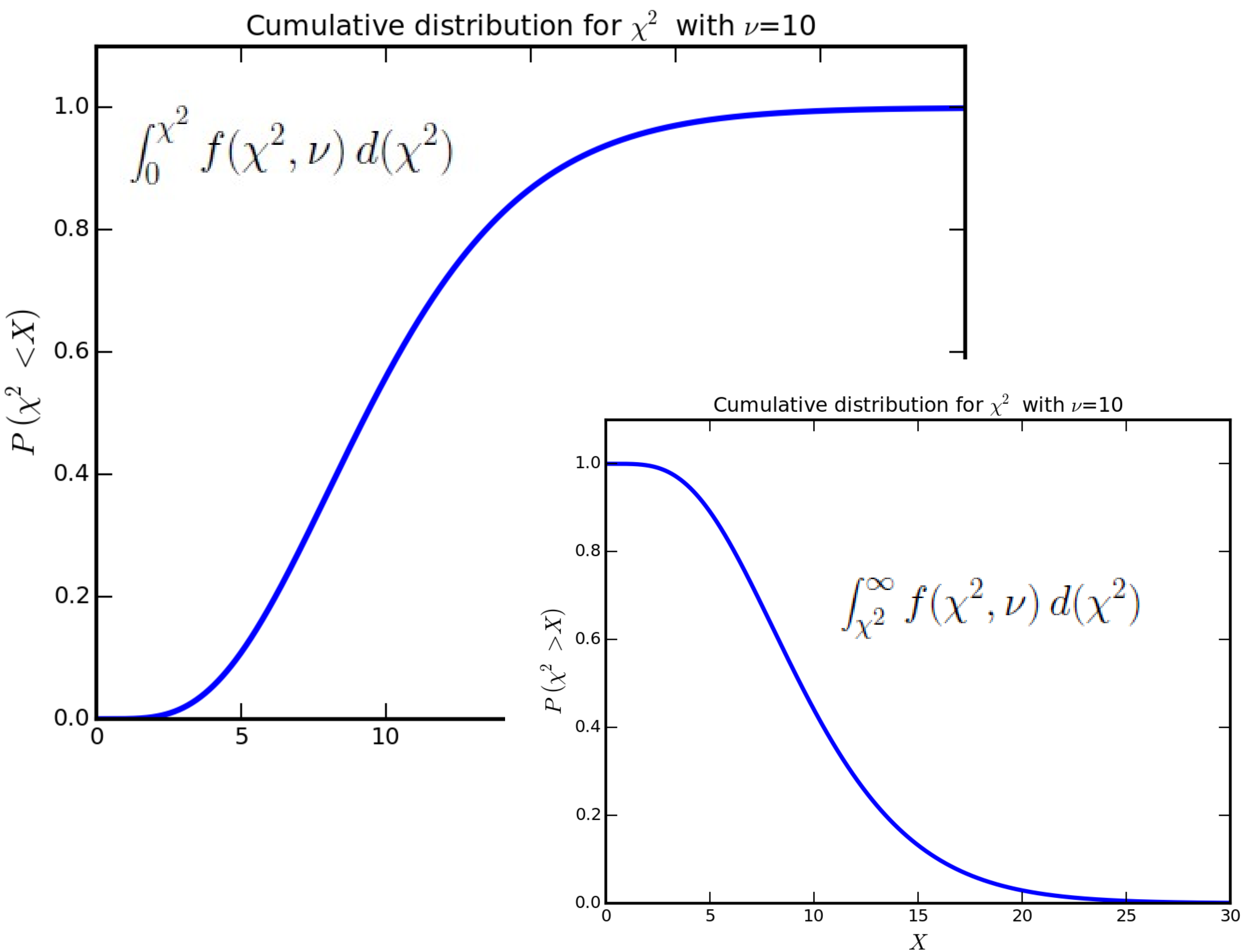
Magnitude of χ^2

- For a good $$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - y(x_i))^2}{\sigma_i^2}$$ $y_i \sim \sigma_i$, we get $\chi^2 \sim N$.
- Incorrect models can lead to large values of χ^2.
- When values $\chi^2_\nu << 1$ are obtained, it is likely that the errors are overestimated, rather than the fit being extremely good.
- Large values of χ^2 could result from underestimated values of the errors, even

Integral Probability Distribution

$$P(>\chi^2,\nu) = \int_{\chi^2}^{\infty} f(\chi^2,\nu)\, d(\chi^2)$$




Cumulative distribution for χ^2 with ν=10
$P(\chi^2 < X)$
$\int_0^{\chi^2} f(\chi^2,\nu)\,d(\chi^2)$
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
Cumulative distribution for χ^2 with ν=10
$P(\chi^2 > X)$
$\int_{\chi^2}^{\infty} f(\chi^2,\nu)\,d(\chi^2)$
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
X

Straight Line Fit

y = a + bx

Minimization of the function

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - a - bx_i)^2}{\sigma_i^2}$$

leads to the best fit values

y = a + bx

$$b = \frac{1}{\Delta}\left(\sum_i \frac{1}{\sigma_i^2}\sum_i \frac{x_i y_i}{\sigma_i^2} - \sum_i \frac{x_i}{\sigma_i^2}\sum_i \frac{y_i}{\sigma_i^2}\right)$$

$$a = \frac{1}{\Delta}\left(\sum_i \frac{x_i^2}{\sigma_i^2}\sum_i \frac{y_i}{\sigma_i^2} - \sum_i \frac{x_i}{\sigma_i^2}\sum_i \frac{x_i y_i}{\sigma_i^2}\right)$$

$$\Delta = \sum_i \frac{1}{\sigma_i^2}\sum_i \frac{x_i}{\sigma_i^2} - \left(\sum_i \frac{x_i}{\sigma_i^2}\right)^2$$

y = a + bx

Straight Line Fit

Variance and Covariance in parameters

$$\sigma_a^2 = \sum_i \left(\frac{\partial a}{\partial y_i}\right)^2 \sigma_i^2 = \frac{1}{\Delta}\sum_i \frac{x_i^2}{\sigma_i^2}$$

$$\sigma_b^2 = \sum_i \left(\frac{\partial b}{\partial y_i}\right)^2 \sigma_i^2 = \frac{1}{\Delta}\sum_i \frac{1}{\sigma_i^2}$$

$$\sigma_{ab}^2 = \sum_i \left(\frac{\partial a}{\partial y_i}\right)\left(\frac{\partial b}{\partial y_i}\right) \sigma_i^2 = -\frac{1}{\Delta}\sum_i \frac{x_i}{\sigma_i^2}$$

Coma Cluster Galaxies Size-Velocity Dispersion

$y=ax+b$

$a=0.23$,

$\sigma^2_a=0.04$

$b=2.17$,

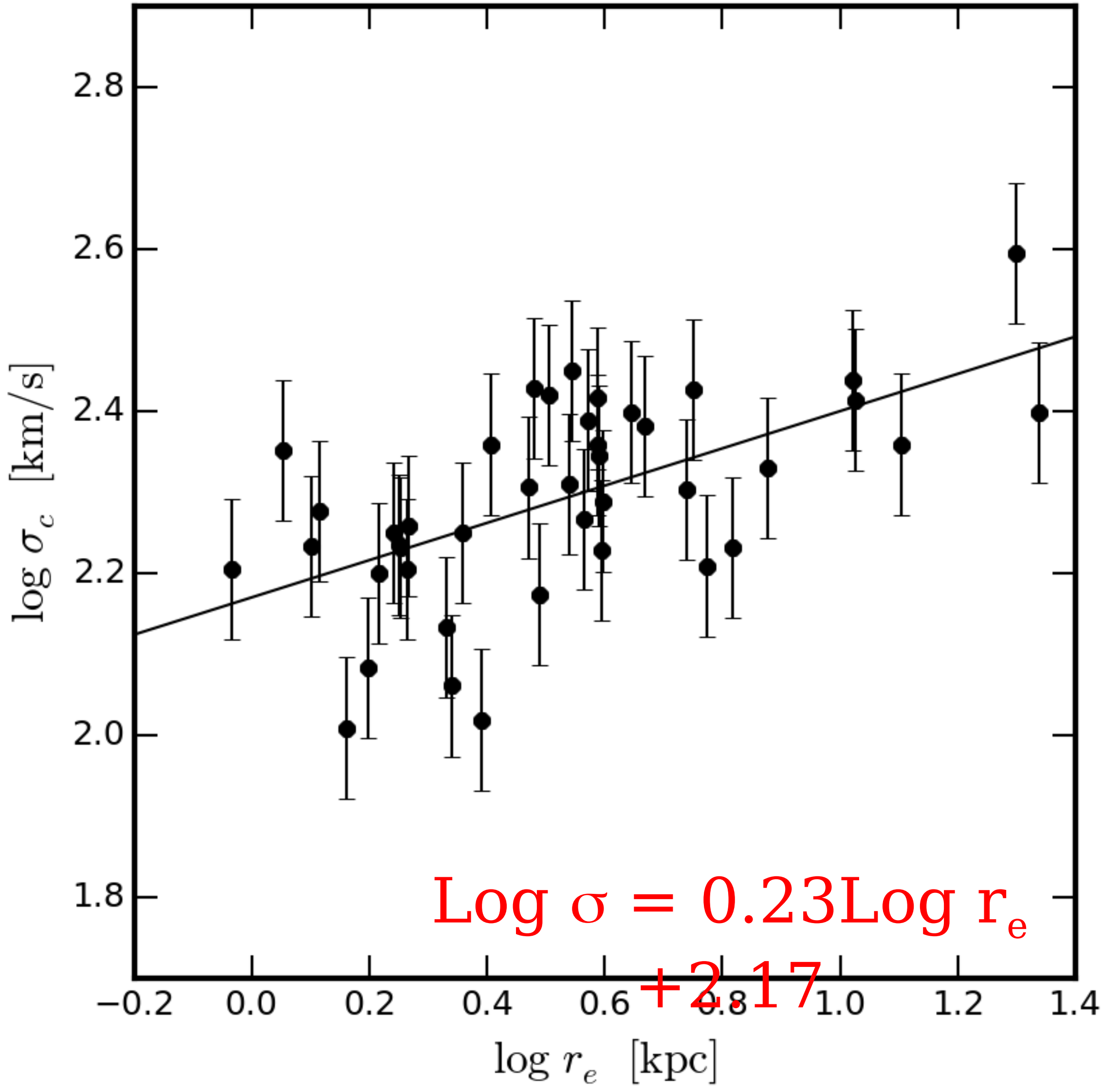
$\sigma^2_b=0.04$

$\chi^2=53.54$,

$\nu=40$

$\chi^2_\nu=1.34$

$P=0.05$



$\text{Log } \sigma = 0.23 \text{Log } r_e + 2.17$

General Linear Models

The model function, linear in the model parameters, but not necessarily in x, is

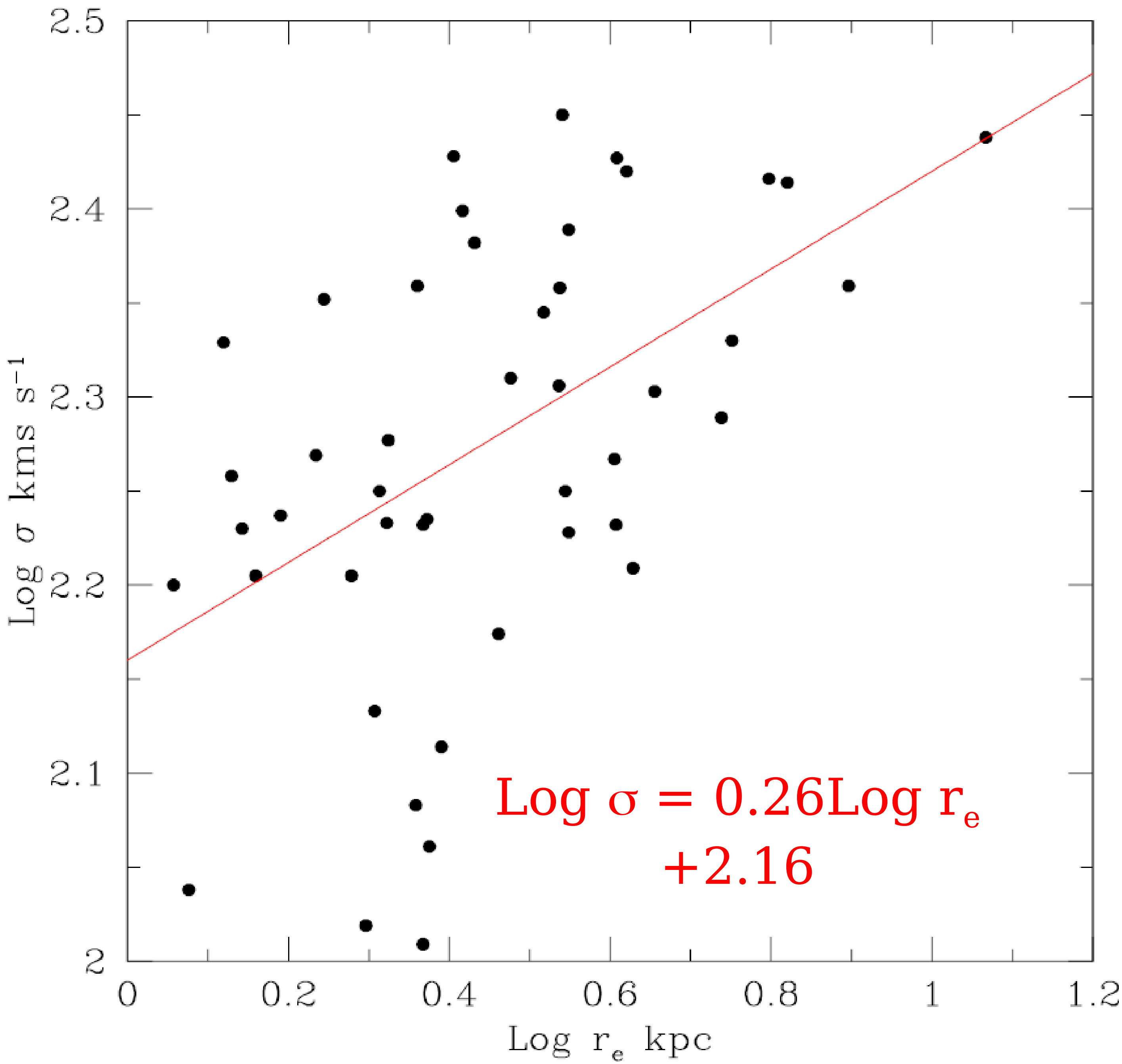
$$y(x) = \sum_{k=1}^{m} a_k f_k(x)$$

The χ^2 function is

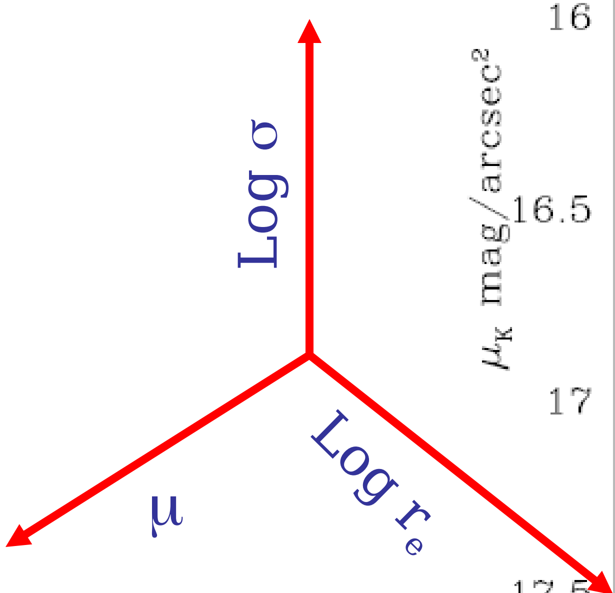
$$\chi^2 = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left[y_i - \sum_{k=1}^{m} a_k f_k(x_i) \right]^2$$

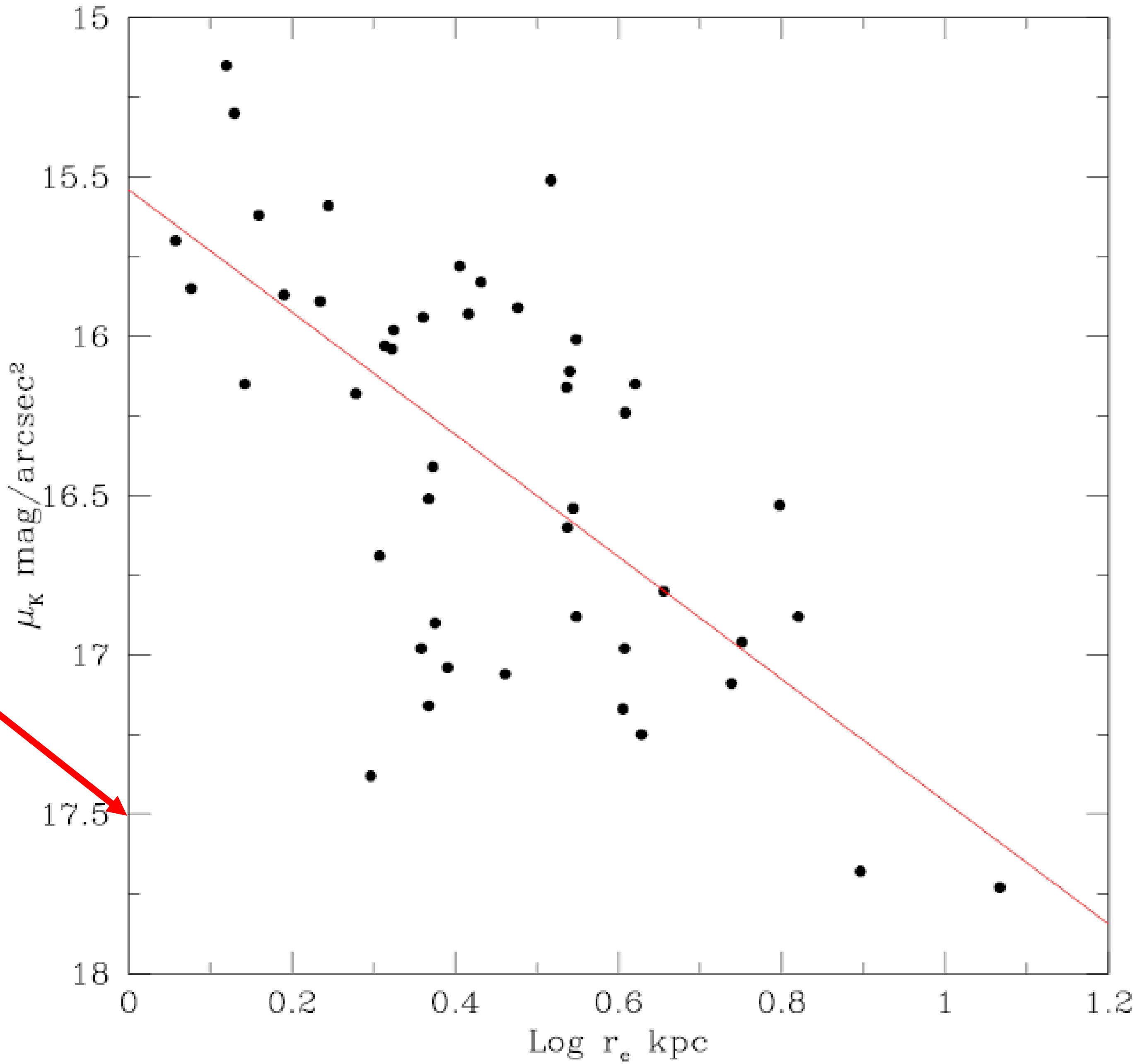
Multivariate Correlations

Coma Cluster Galaxies Size-Velocity Dispersion

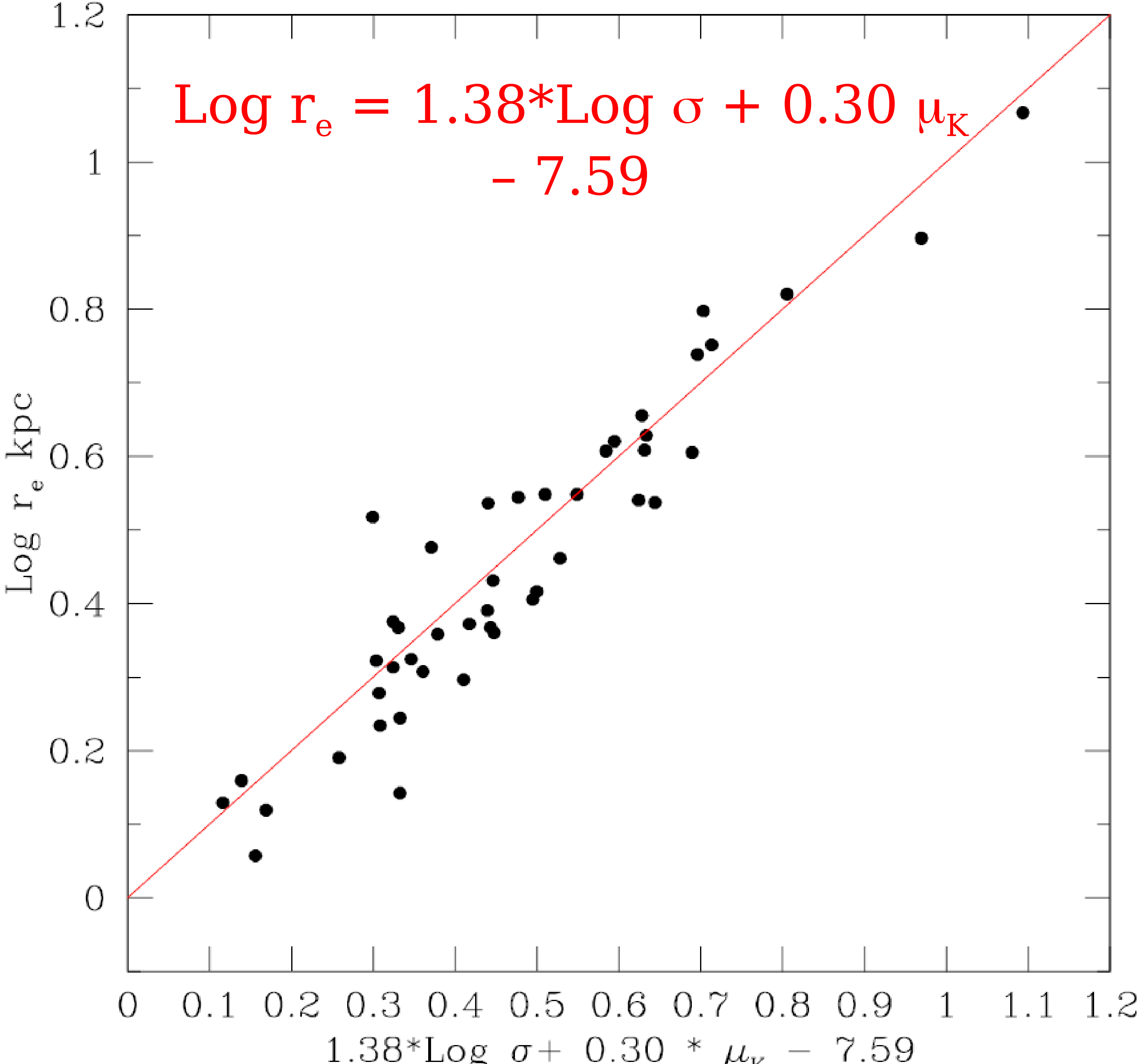


Coma Cluster Galaxies Kormendy Diagram





Coma Cluster Galaxies Fundamental Plane



Rank Correlations

Spearman Rank Correlation

This is a non-parametric test, used to see whether there is a correlation between the data points.

No model, like a straight line, is assumed, so there no parameters.

It is applied to N pairs of measurements (x_i, y_i), $i=1,N$.

Obtain the ranks $R(x_i)$ amongst the x_i, and $R(y_i)$ amongst the y_i.

Spearman Rank Correlation

Obtain the ranks $R(x_i)$ amongst the x_i, and $R(y_i)$ amongst the y_i. The Spearman rank corre

$$r_s = \frac{\operatorname{cov}\left(R(x), R(y)\right)}{\sigma_{R(x)}\sigma_{R(y)}}$$

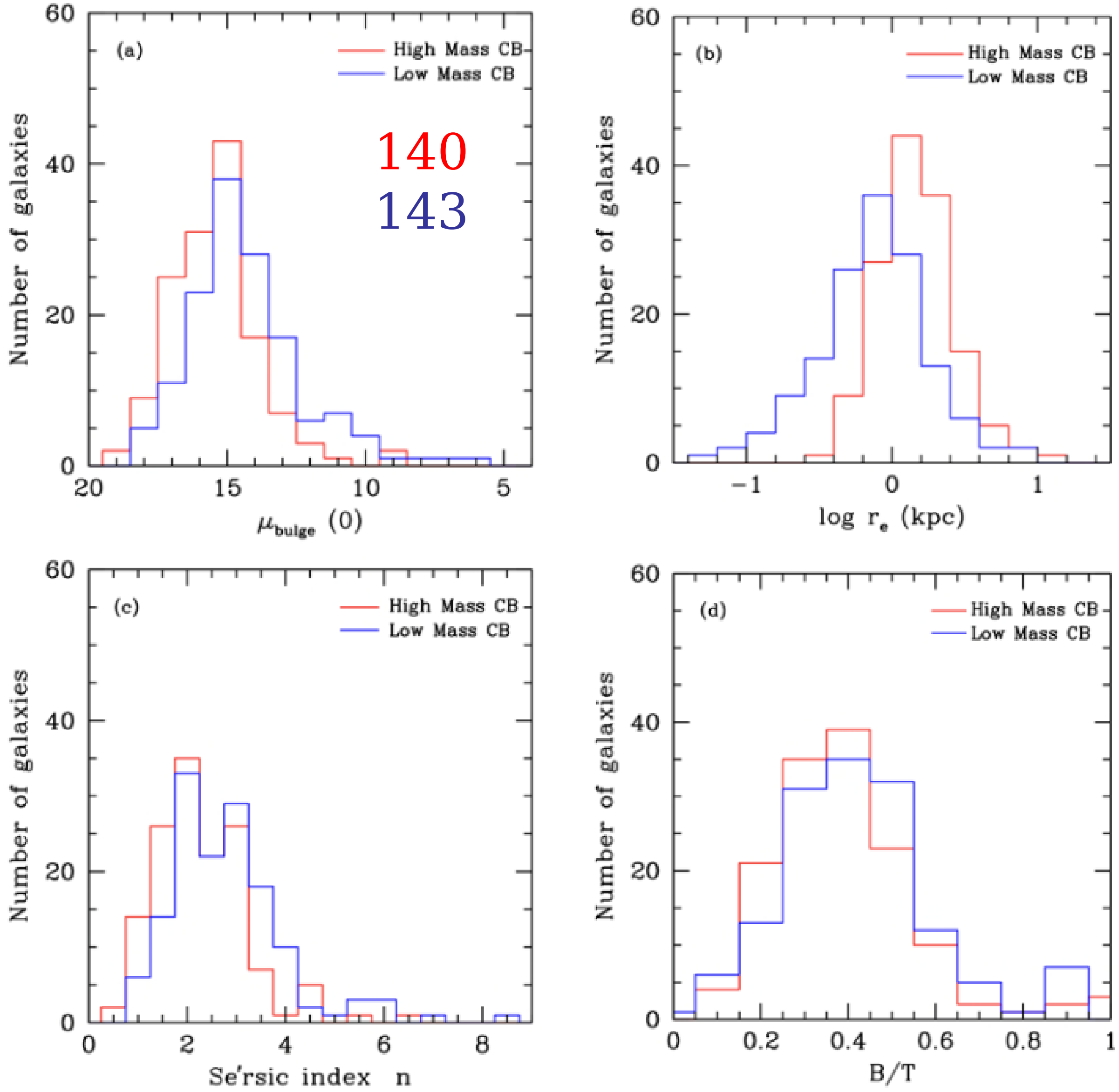
$$r_s = 1 - \frac{6\sum_i \left(R(x_i) - R(y_i)\right)^2}{N^3 - N}$$

No Ties

$$t = r_s\sqrt{\frac{N-2}{1-r_s^2}}$$

Student’s t

Parameter Distributions



Kolmogorov-Smirnov Test

$$N(< x) = \sum_{x_i < x} N(x_i)$$

Cumulative Distribution

$$p(< x) = \int_{-\infty}^{x} p(x)\, dx$$

Cumulative Probability

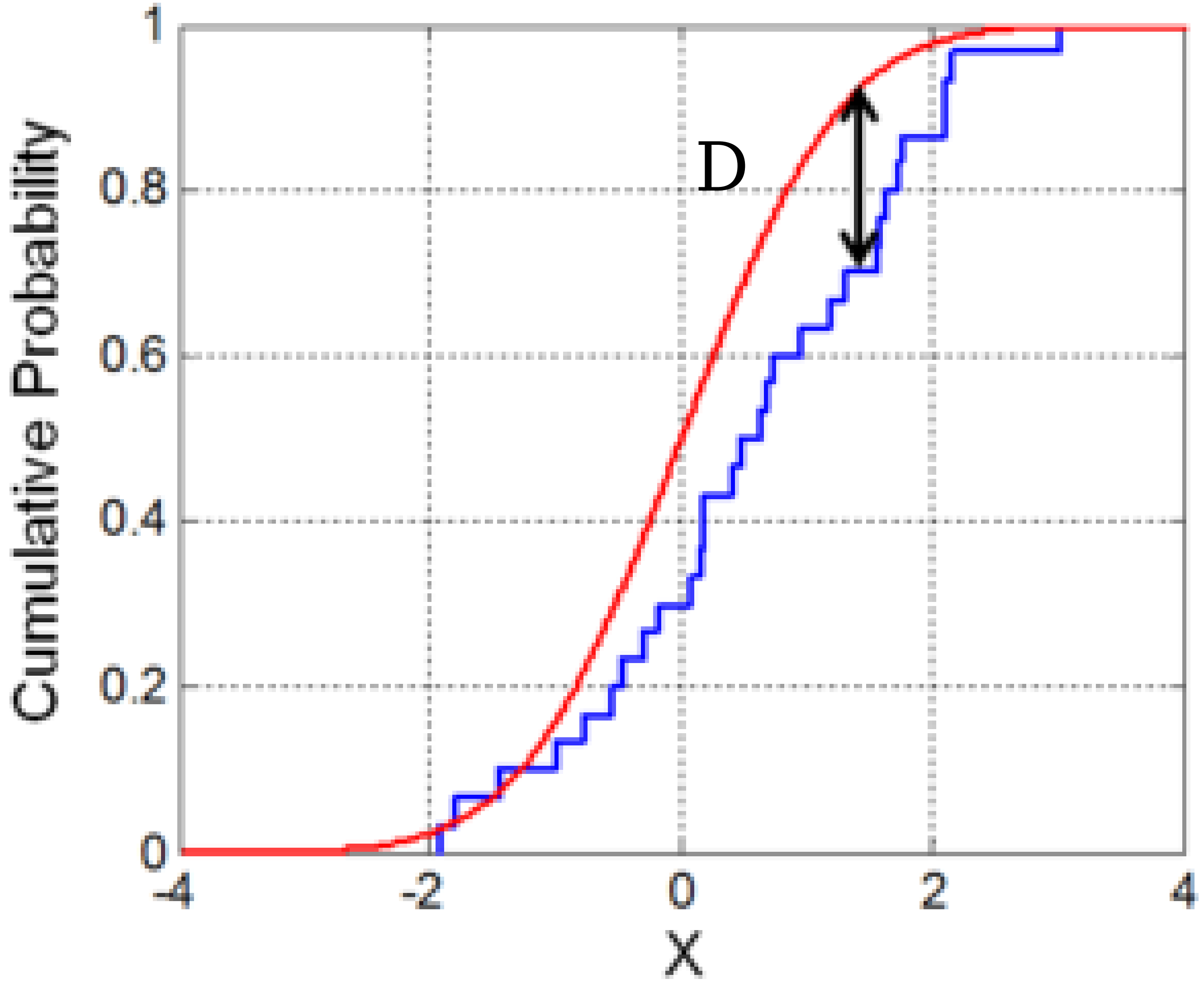
$$D = \max \left| \mathrm{N}(< \mathrm{x}) - \mathrm{p}(< \mathrm{x}) \right|$$

1-SampleTtest

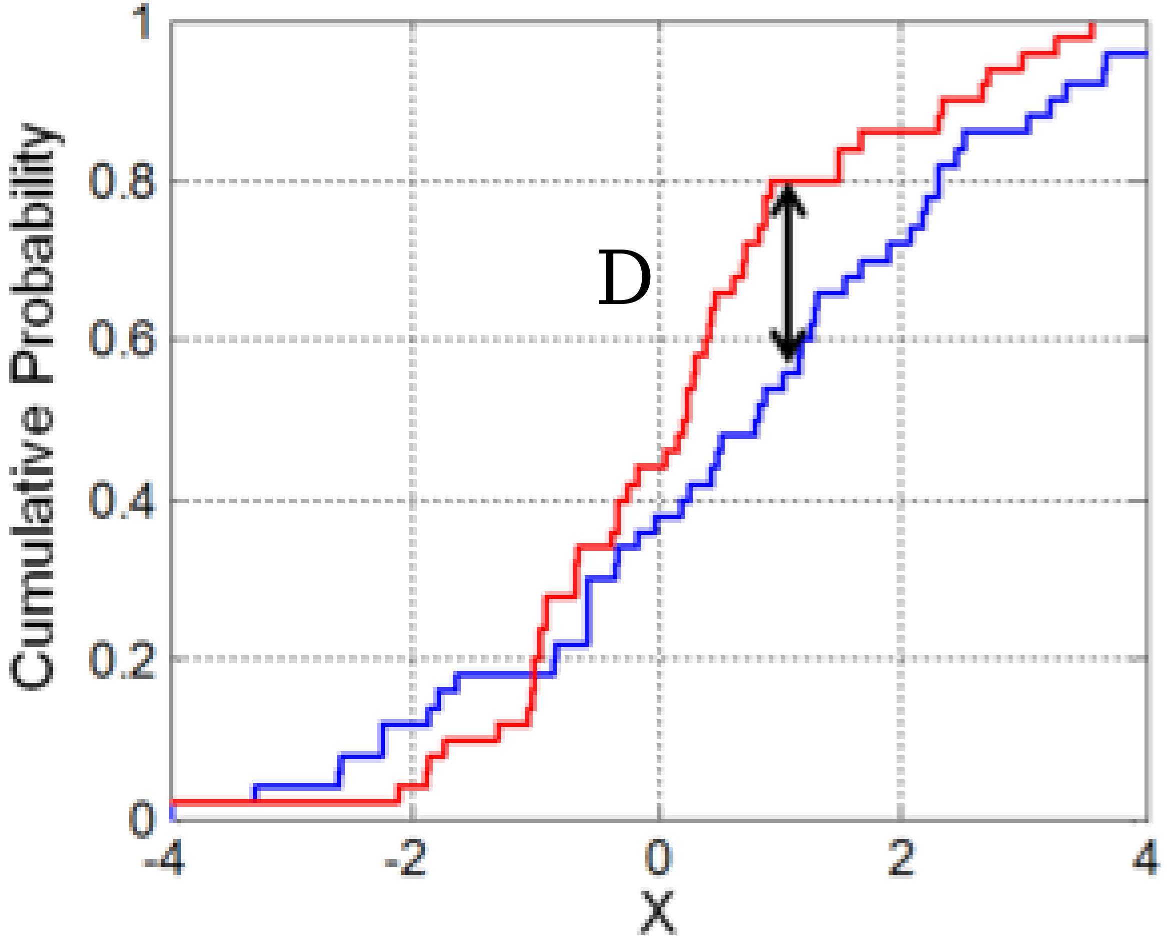
$$D = \max \left| \mathrm{N}_1(< \mathrm{x}) - \mathrm{N}_2(< \mathrm{x}) \right|$$

2-SampleTtest

Kolmogorov-Smirnov 1-Sample Test



Kolmogorov-Smirnov 2-Samples Test



Kolmogorov-Smirnov 2-Sample Test

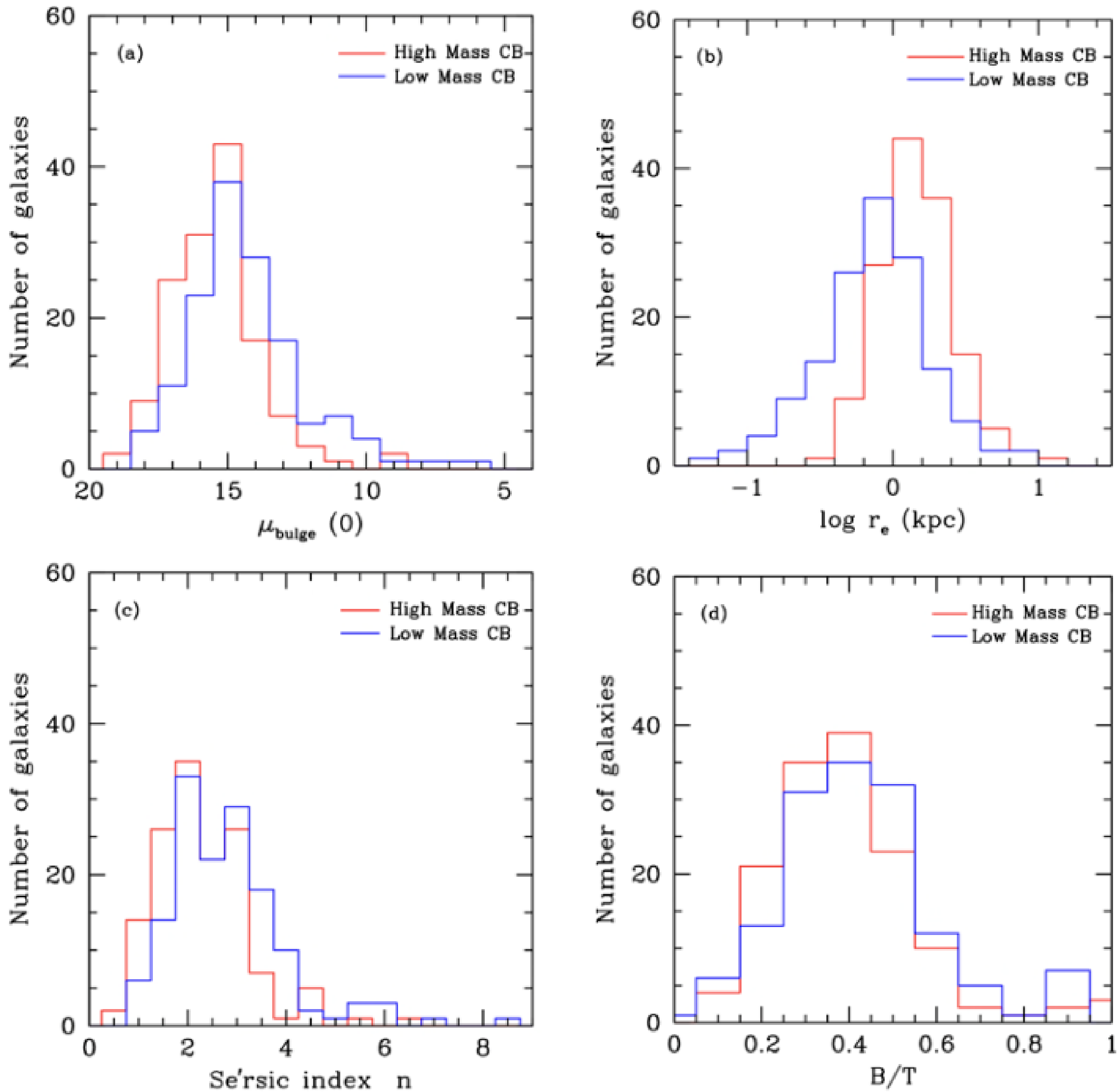
Galaxies: Sample 1, 140; Sample 2, 143

Parameters	D	P (> D)
$\log r_e$	0.414	2.701×10^{-11}
$\log n$	0.207	0.004
$< \mu_b (< r_e) >$	0.189	0.010
B/T	0.131	0.162

Parameter Distributions With KS Test

0.19
0.01

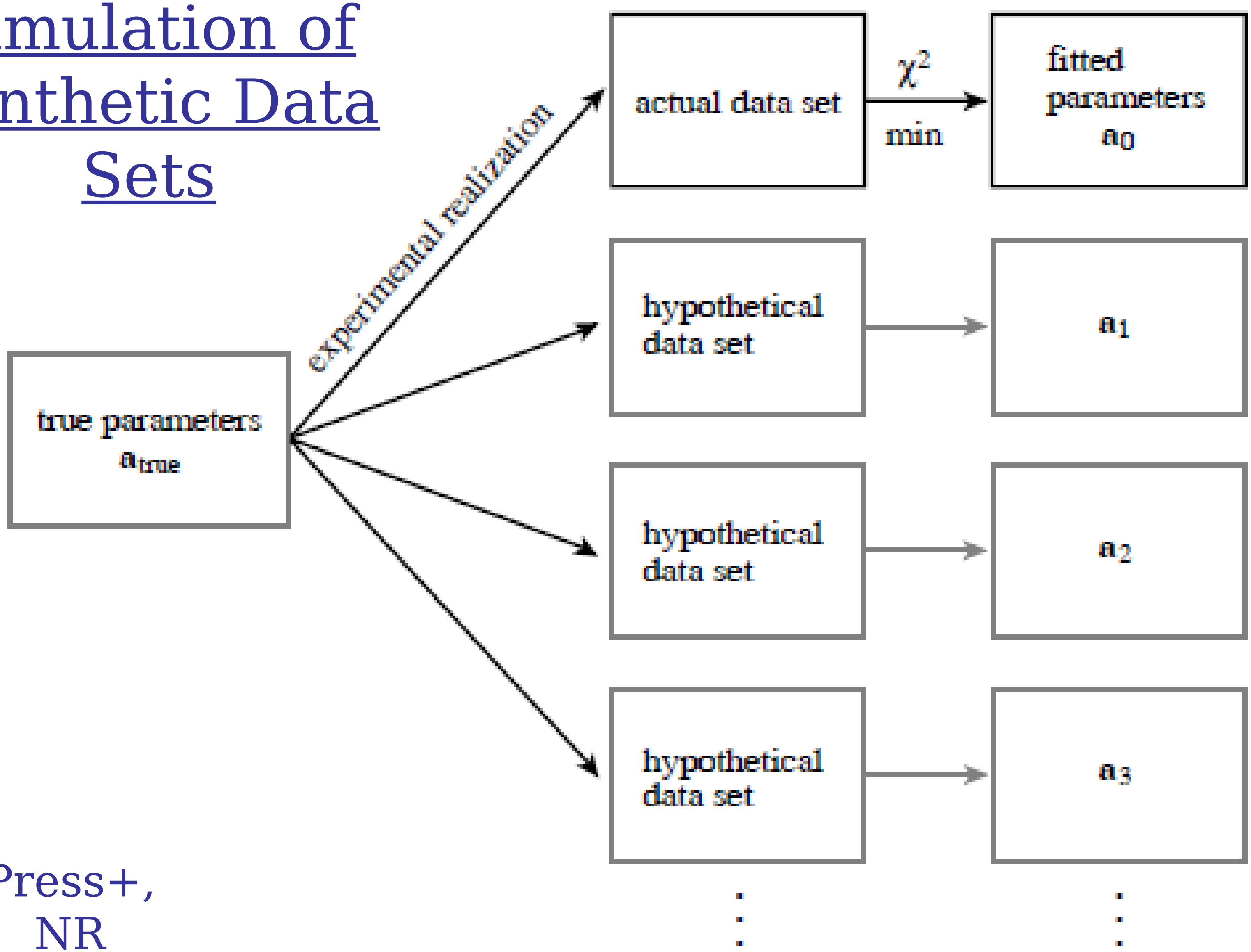
D=0.41
P(>D)=2.7x10^{-11}



0.20,
0.004

0.13
,
0.16

Simulation of Synthetic Data Sets



Press+, NR

Figure 15.6.1. A statistical universe of data sets from an underlying model. True parameters a_{true} are realized in a data set, from which fitted (observed) parameters a_0 are obtained. If the experiment were repeated many times, new data sets and new values of the fitted parameters would be obtained.

Simulation of Synthetic Data Sets

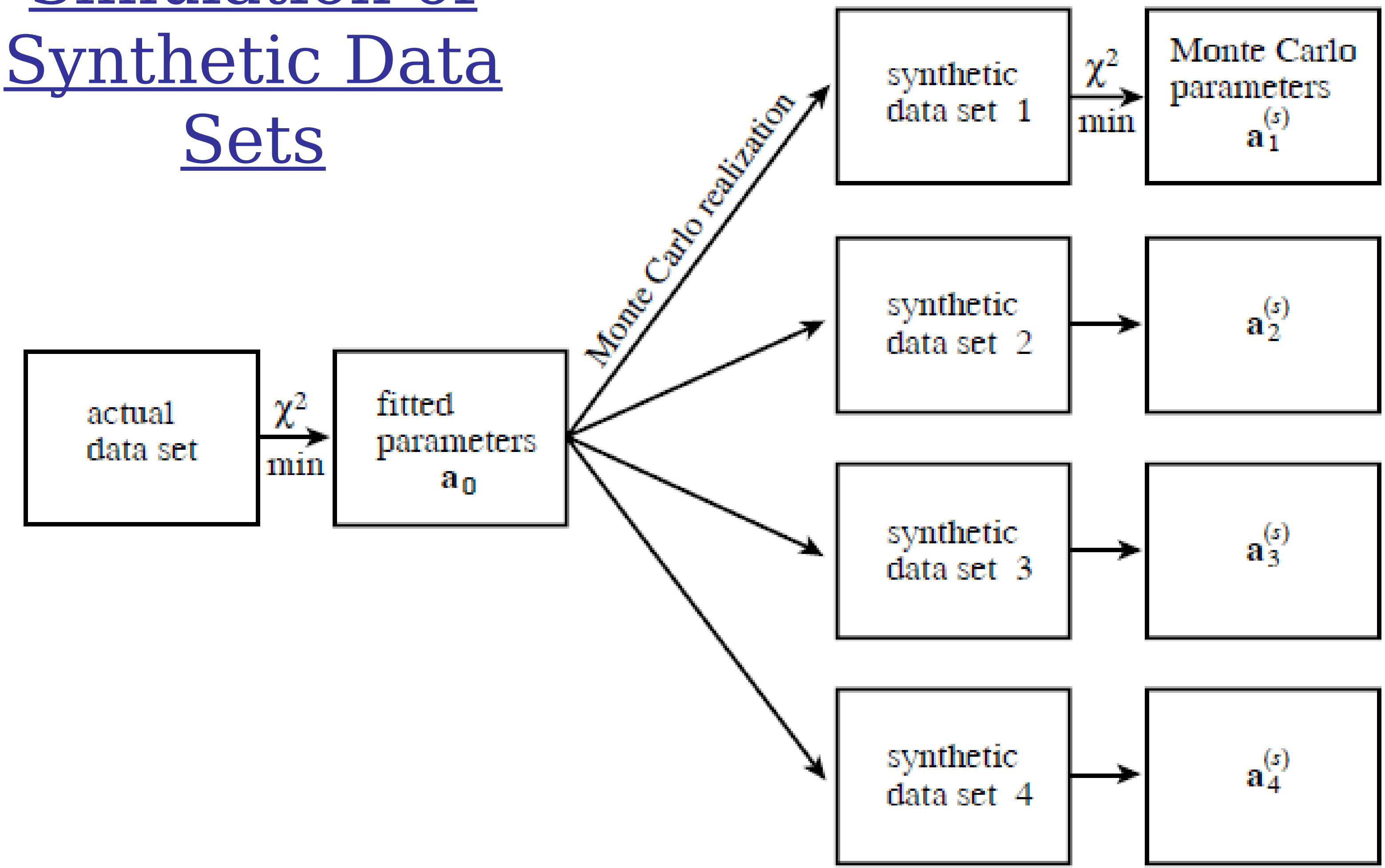


Figure 15.6.2. Monte Carlo simulation of an experiment. The fitted parameters from an actual experiment are used as surrogates for the true parameters. Computer-generated random numbers are used to simulate many synthetic data sets. Each of these is analyzed to obtain its fitted parameters. The distribution of these fitted parameters around the (known) surrogate true parameters is thus studied.

Press+, NR

Straight Line Model Simulations

1. Choose some values of 'x'. Since we have assumed it to be error free these will be same for all the data sets.

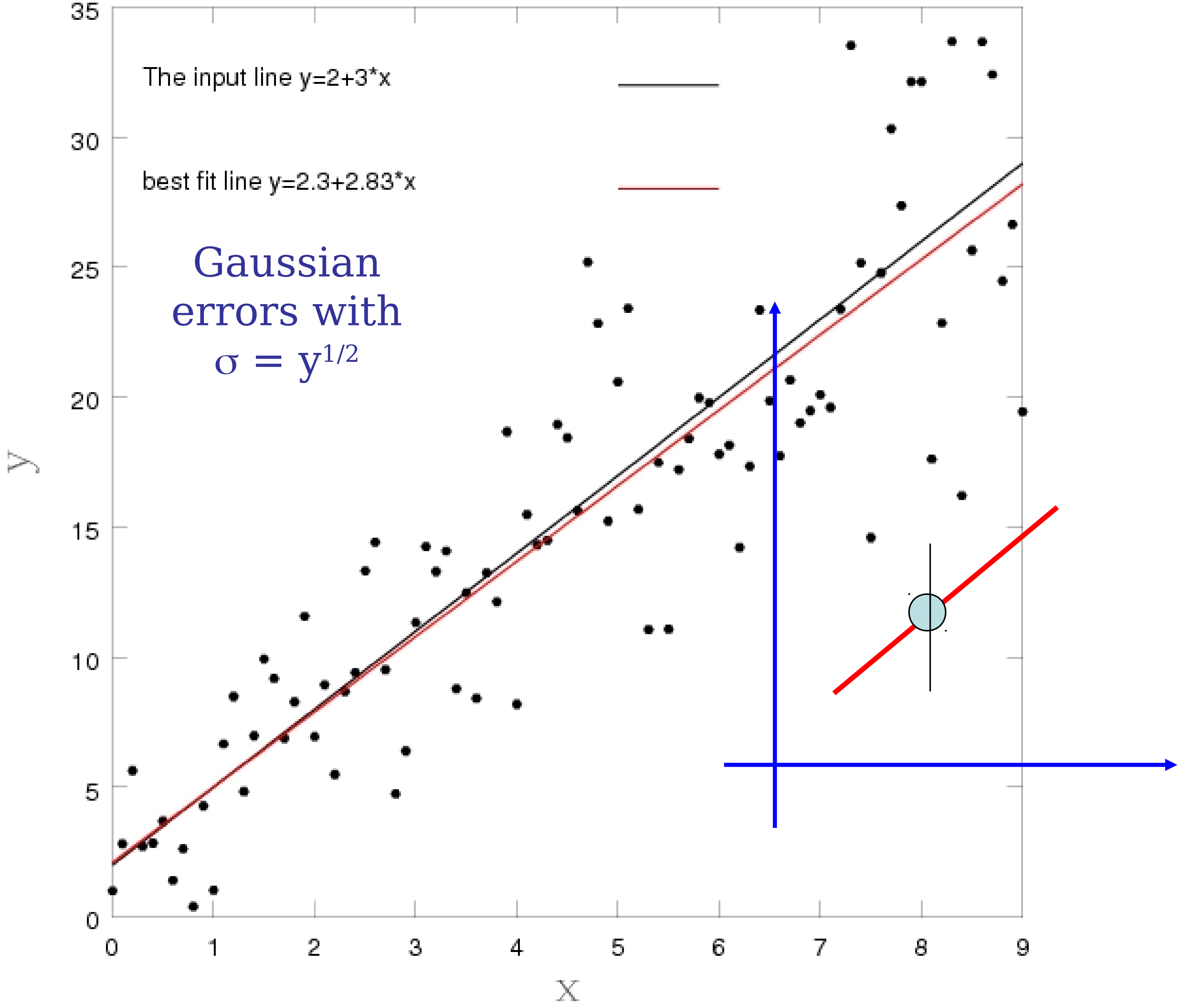
2. Choose some true values for 'a' and 'b'. (2 and 3, in our case.)

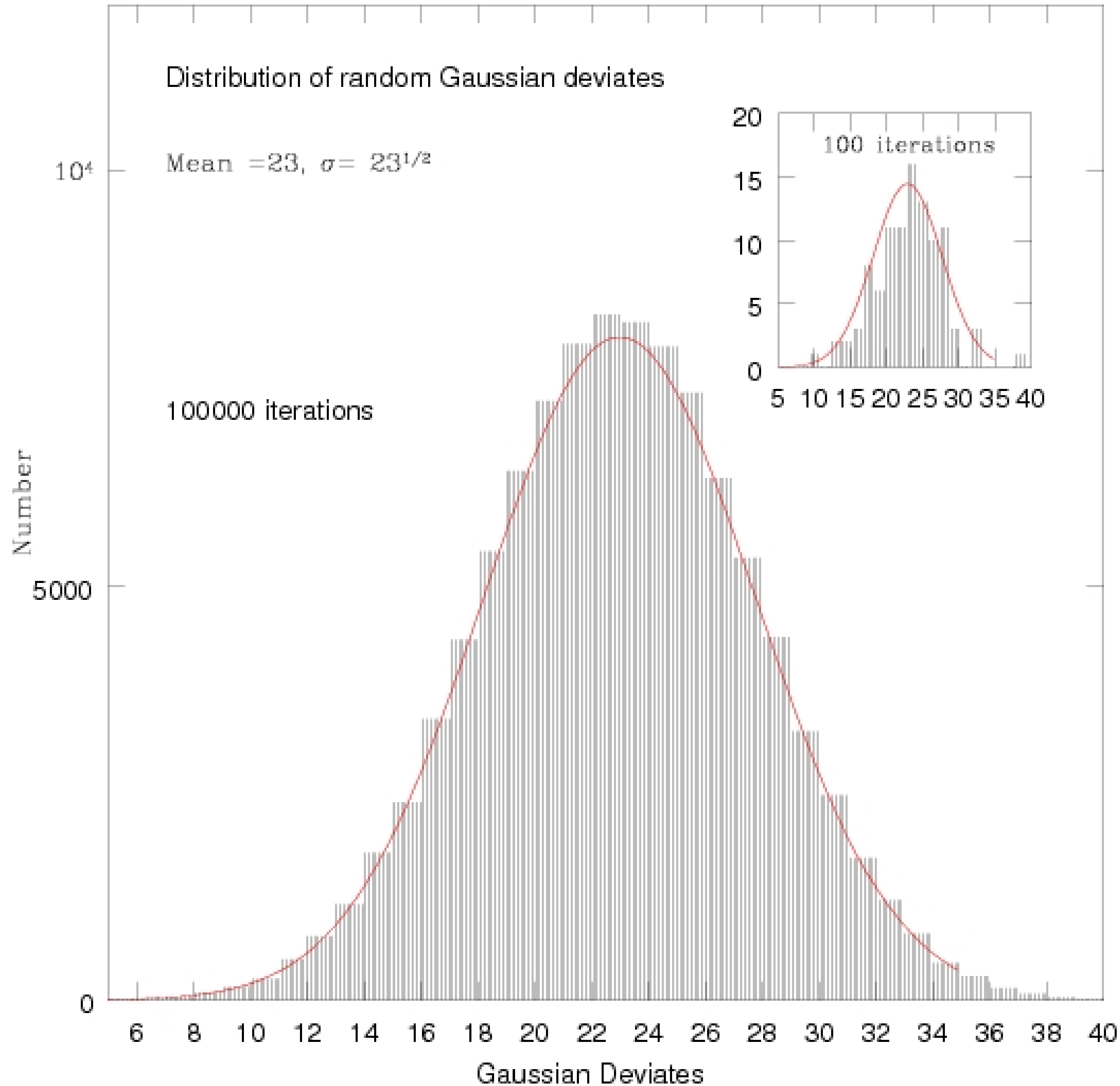
3. For each value of 'x', compute the true value of 'y'.

4. From a Gaussian centered on $y_{\mathbf{true}}$, with a σ of $\sqrt{y}$, we draw a random variable. This becomes the measured 'y'. We choose $\sqrt{y}$ to approximate a Poissonian distribution but this is not necessary. We might have as well have chosen a constant σ.

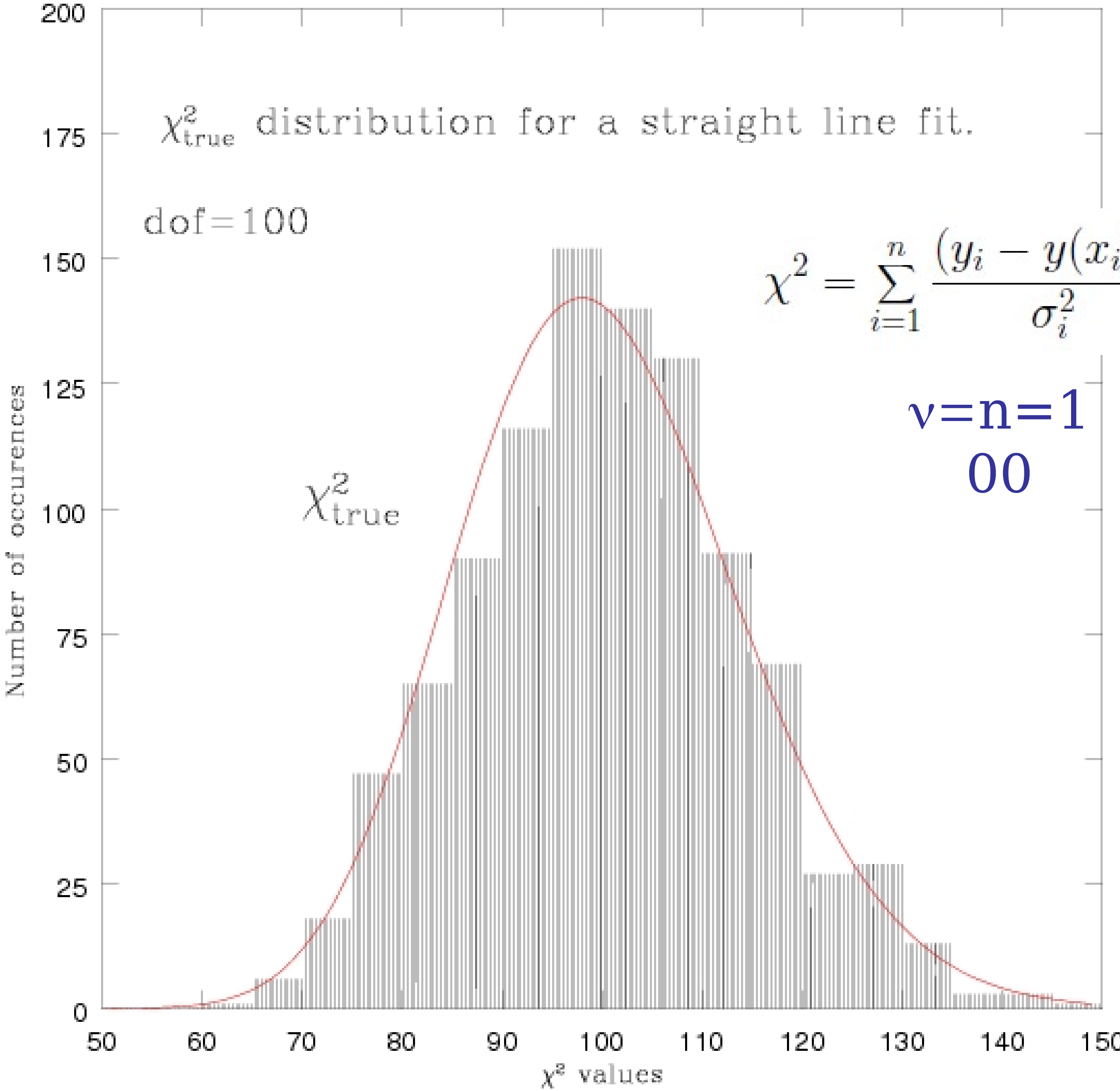
5. This way we have a synthetic data set. For this set, we can fit a straight line and determine 'a' and 'b' and their standard deviations. We can compute χ^2 and χ^2_ν and their respective standard deviations as well. The goal will be to study these distributions and verify the theory.

6. Steps 3 - 5 can be repeated as many times as the number of synthetic data sets we need to construct.


The input line y=2+3*x
best fit line y=2.3+2.83*x
Gaussian errors with σ = y^1/2
y
x
0
5
10
15
20
25
30
35
0
1
2
3
4
5
6
7
8
9

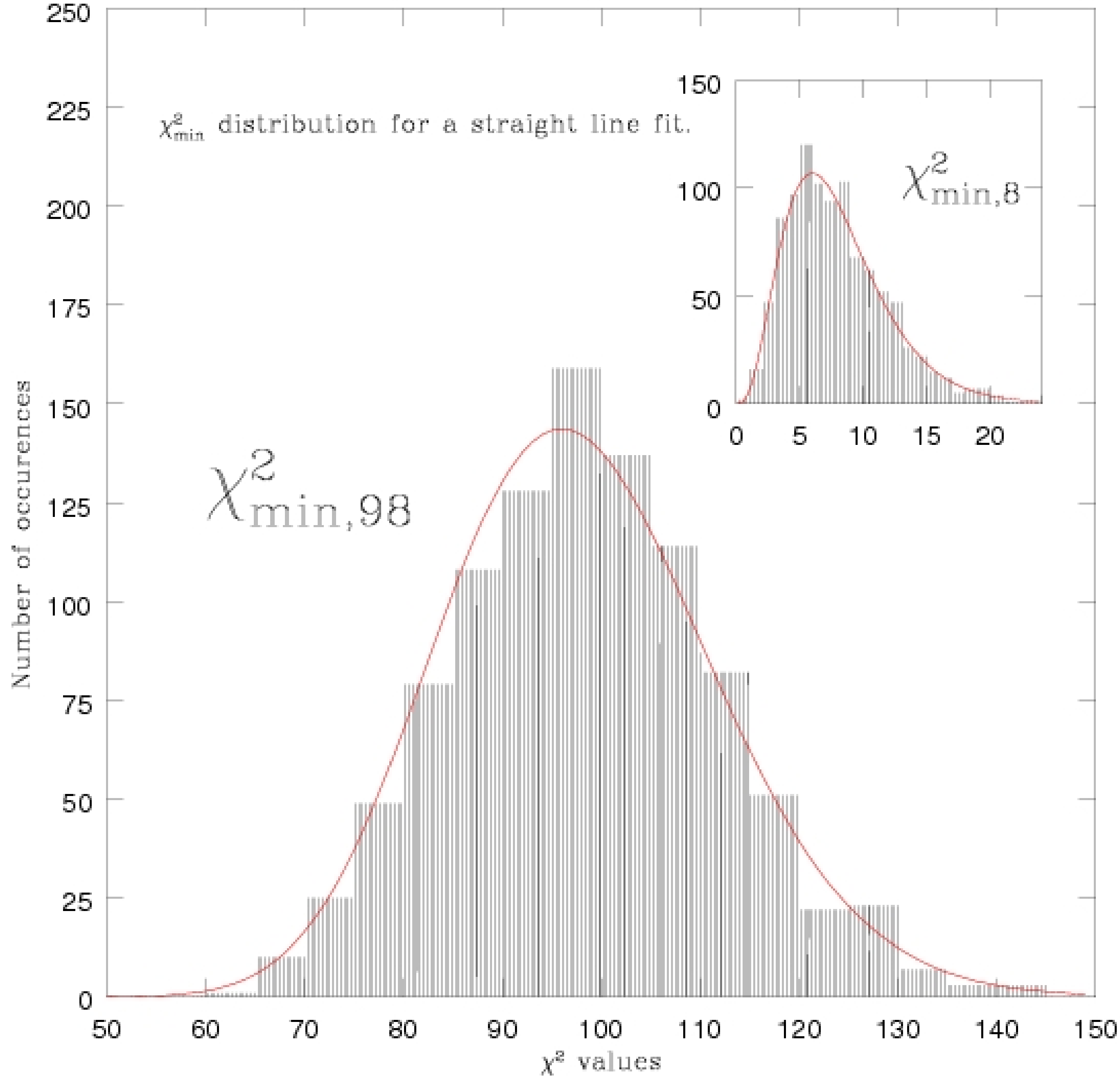


Distribution of random Gaussian deviates
Mean =23, σ= 23^1/2
100000 iterations
100 iterations
Number
Gaussian Deviates

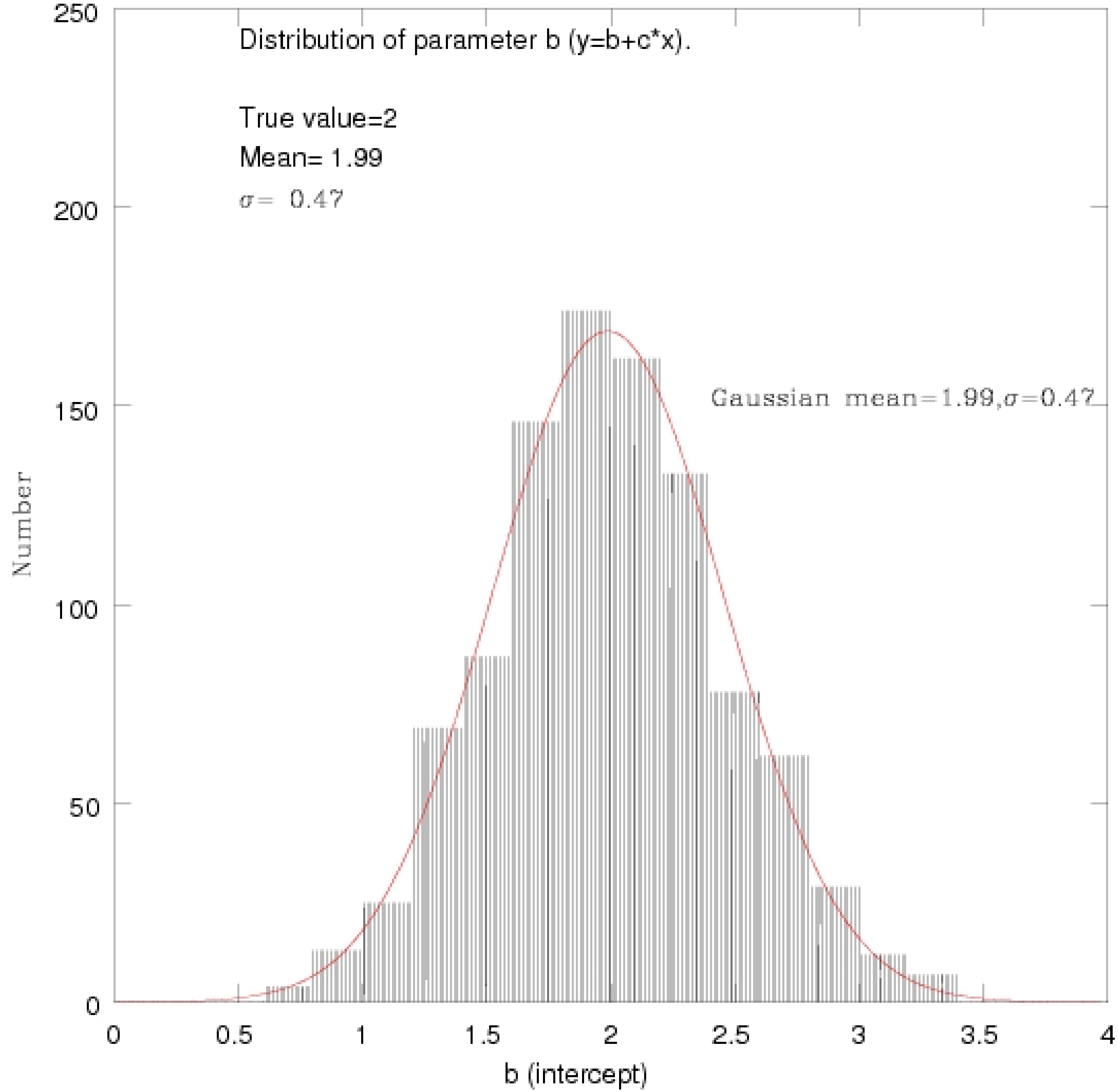


$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - y(x_i))^2}{\sigma_i^2}$$

ν=n=1
00



χ^2_{min} distribution for a straight line fit.
$\chi^2_{min,98}$
$\chi^2_{min,8}$
Number of occurences
χ^2 values
0
25
50
75
100
125
150
175
200
225
250
50
60
70
80
90
100
110
120
130
140
150
0
50
100
150
0
5
10
15
20


Distribution of parameter b (y=b+c*x).
True value=2
Mean= 1.99
σ= 0.47
Gaussian mean=1.99,σ=0.47
Number
b (intercept)
0
50
100
150
200
250
0
0.5
1
1.5
2
2.5
3
3.5
4

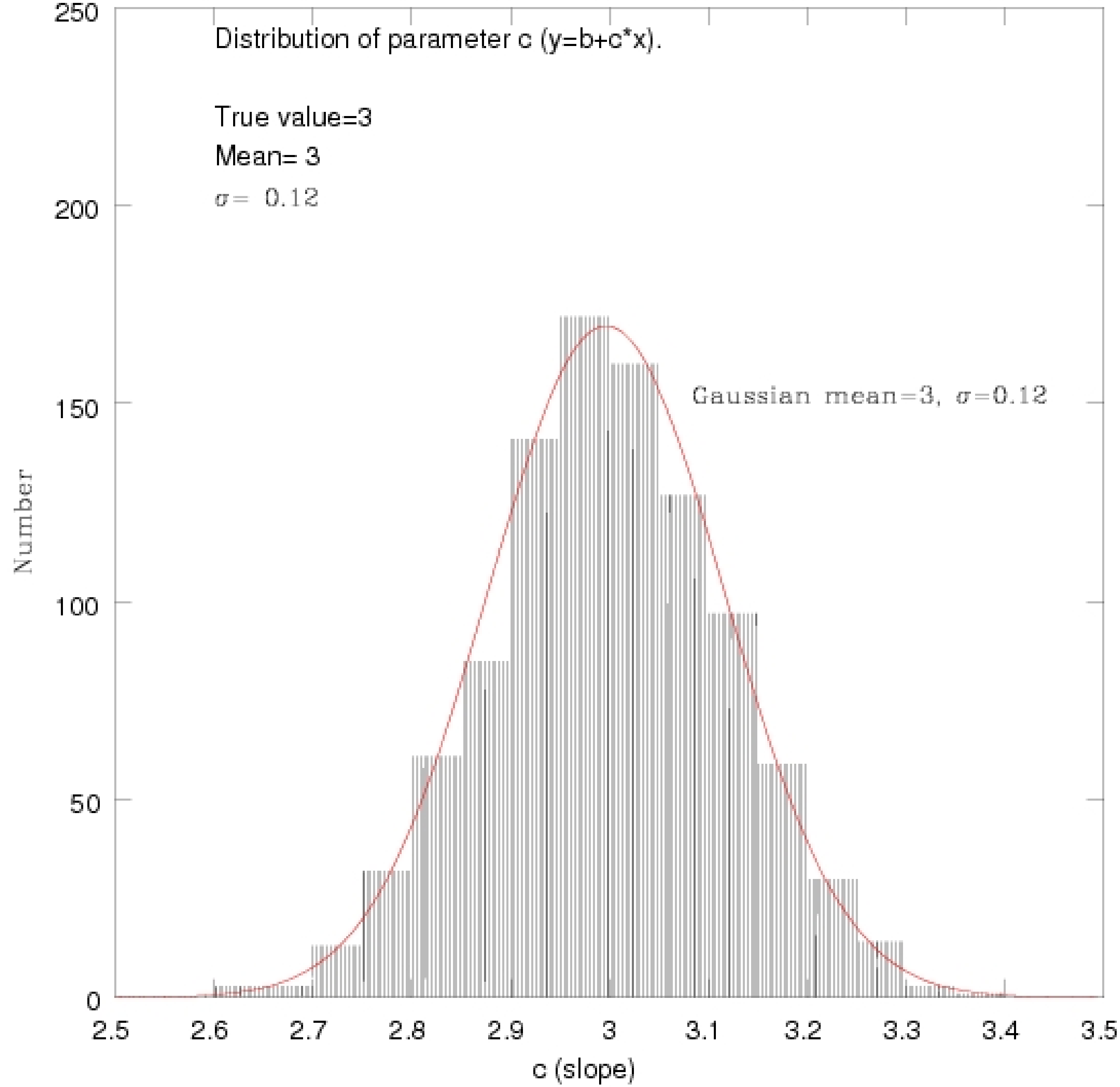

Distribution of parameter c (y=b+c*x).
True value=3
Mean= 3
σ= 0.12
Gaussian mean=3, σ=0.12
Number
c (slope)
0
50
100
150
200
250
2.5
2.6
2.7
2.8
2.9
3
3.1
3.2
3.3
3.4
3.5

Table 1: Parameters, their means and standard deviations

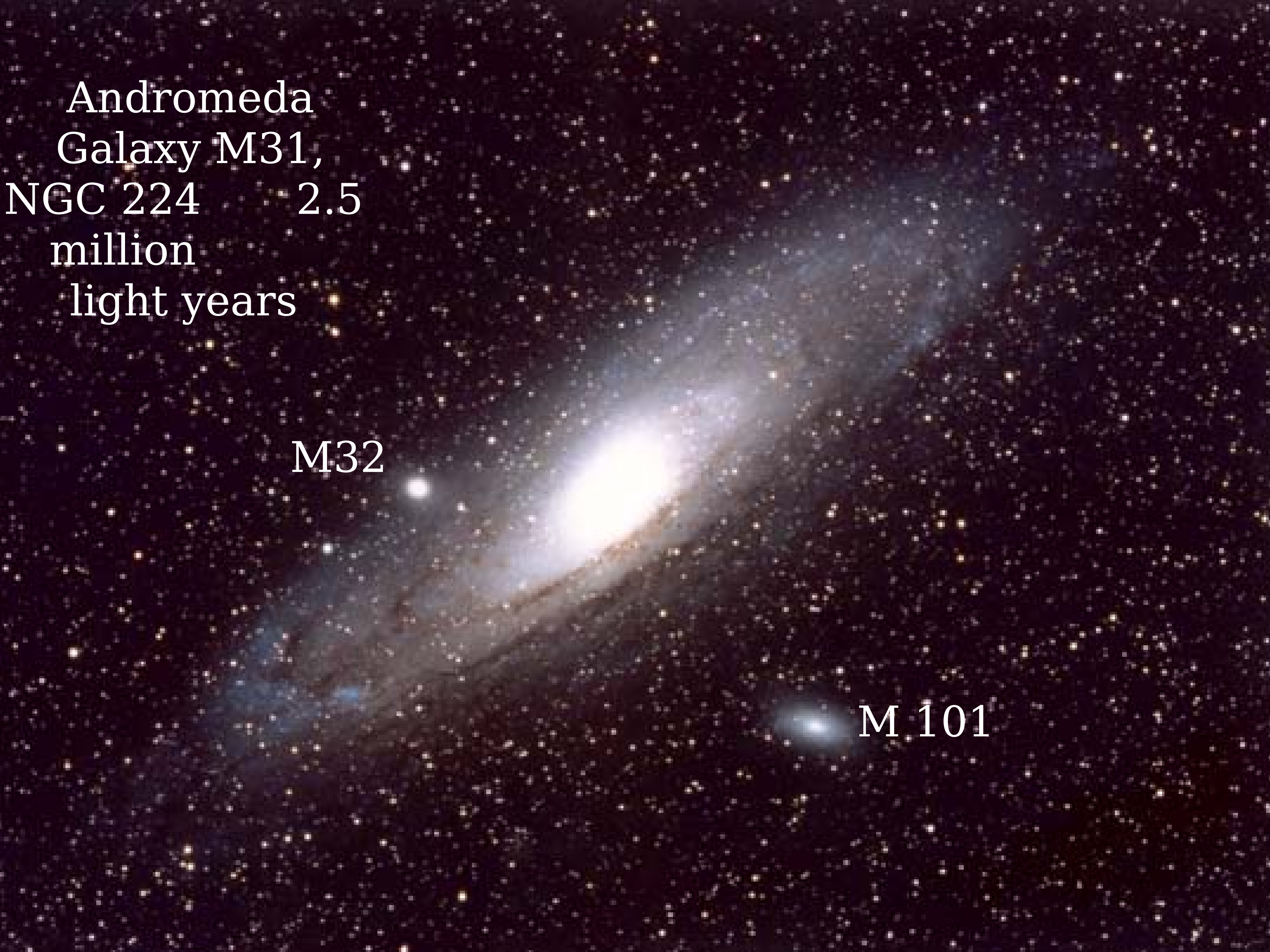
Parameter	Mean (theory)	Mean (simulations)	Std Dev (theory)	Std Dev (simulations)
a	2.0	2.0061	0.470	0.468
b	3.0	2.9986	0.119	0.119
χ^2_{min}	98	98.259	14.0	14.007
χ^2_ν	1	1.002	0.143	0.143

Table 1: The errors in parameters b and c of a straight line $y = b + c \times x$ obtained in 5 different ways. The realisations were obtained using a dispersion $\sigma = y_{true}^{1/2}$.

Errors obtained using:	σ_b (mean b=2)	σ_c (mean c= 3)	σ_{bc} (covariance)
Theoretical calculation	0.47	0.11	-0.04
Simulations	0.47	0.11	
Minimising over the other parameter(Analytic)	0.46	0.105	
Minimising over the other parameter(Fitted)	0.465	0.115	
Error ellipse projections	0.57	0.17	

Thank You!

Galaxies


Andromeda
Galaxy M31,
NGC 224 2.5
million
light years
M32
M 101

M 51

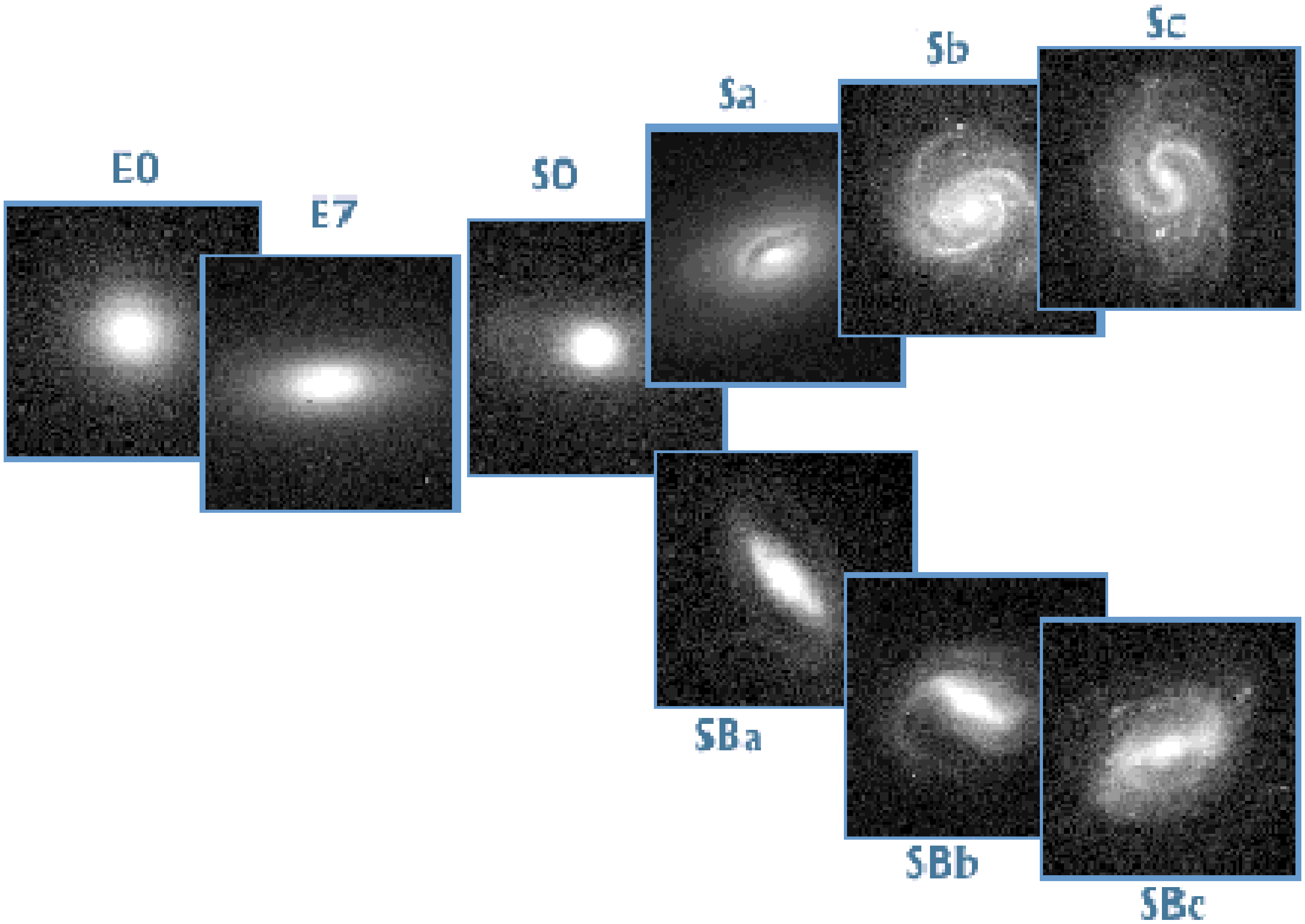
Distance ~31 Mly

First object to be recognized as a spiral, by Lord Ross 1845.

Elliptical Galaxy M 87

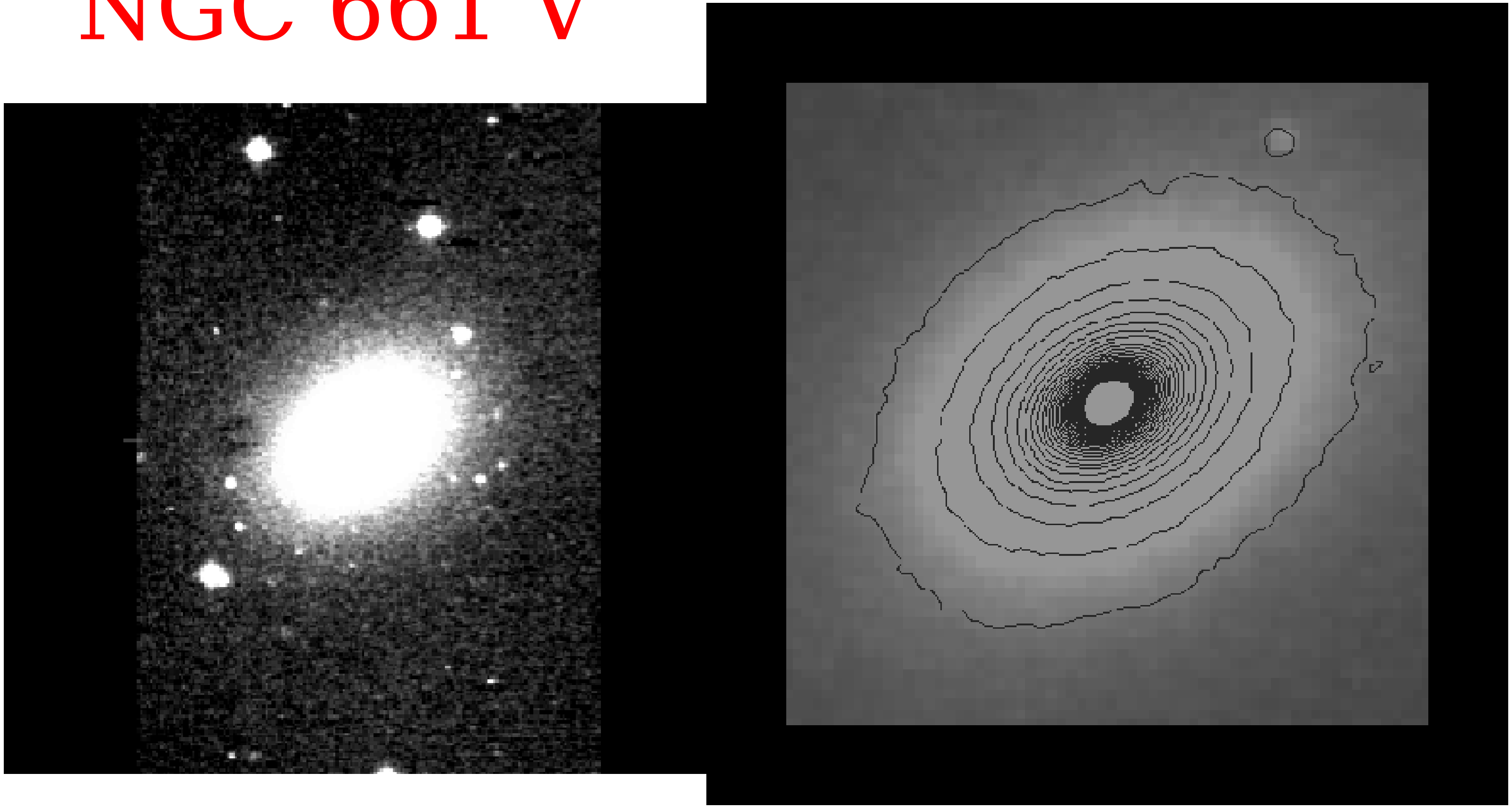
Largest and brightest galaxy in the Virgo Cluster, distance 55 Mlyr

Hubble's Tuning Fork

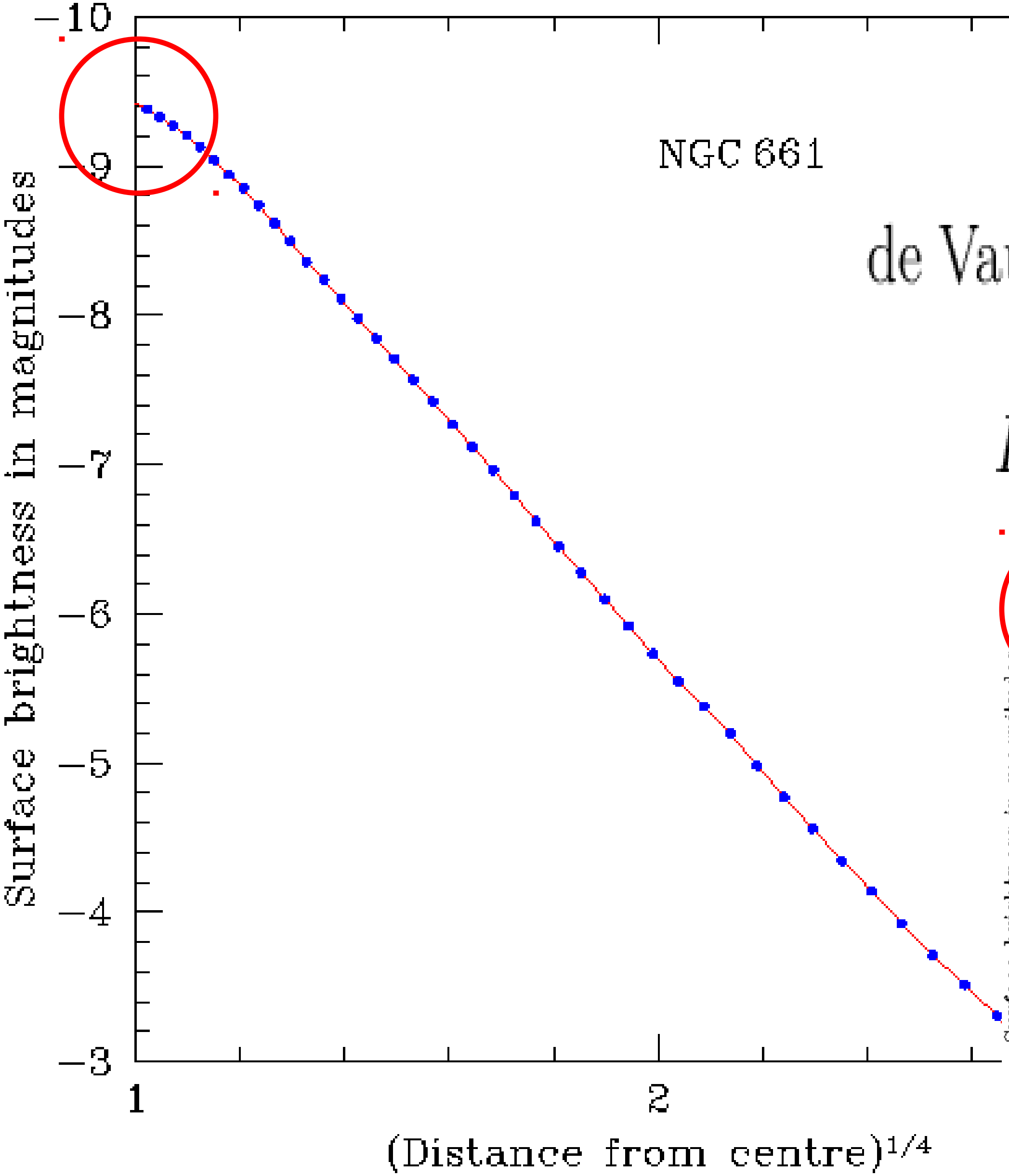


Surface Brightness Distribution

NGC 661 V

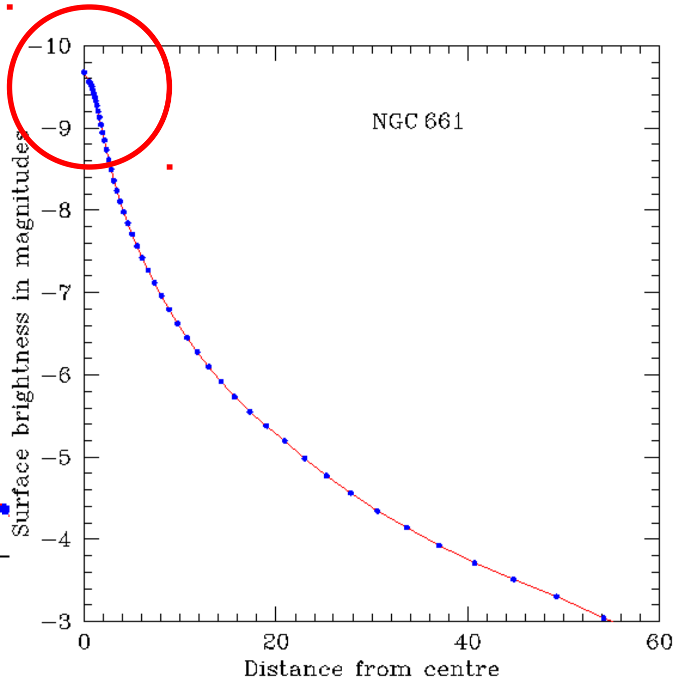


Surface Brightness Profile



de Vaucouleurs' law:

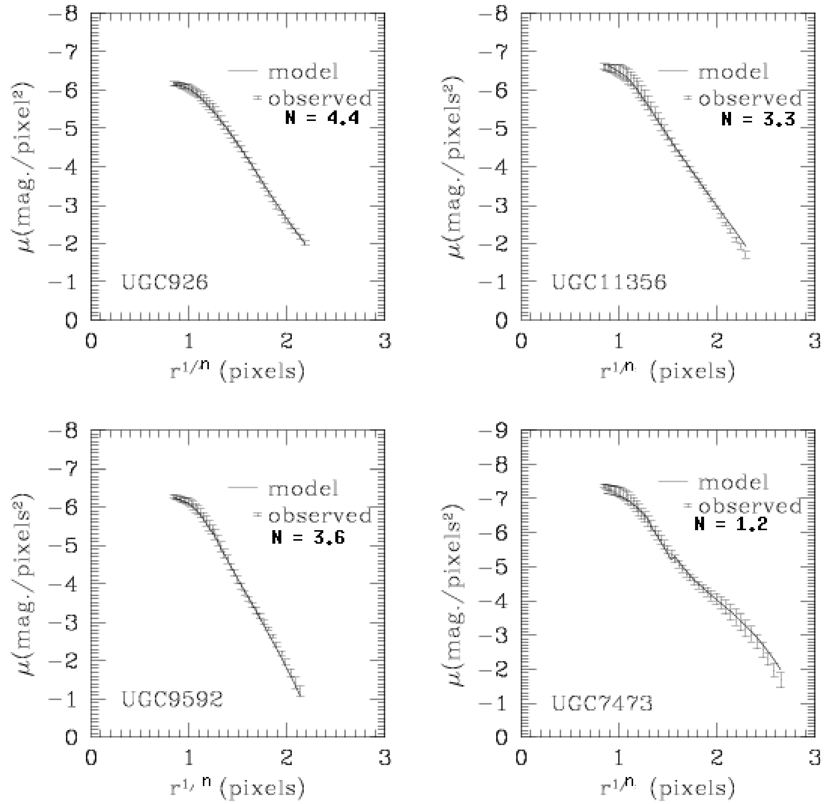
$$I_b(r) = I_b(0)10^{-3.33(r/r_e)^{1/4}}$$



Galaxy Surface Brightness

Surface Brightness

$$I(r) = \delta(r) + I_{\rm b}(r) + I_{\rm d}(r)$$



de Vaucouleurs' law:

$$I_{\rm b}(r) = I_{\rm b}(0) 10^{-3.33 (r/r_{\rm e})^{1/4}}$$

Sersic law:

$$I_{\rm b}(r) = I_{\rm b}(0) 10^{-c_n \left(\frac{r}{r_{\rm e}}\right)^{1/n}}$$

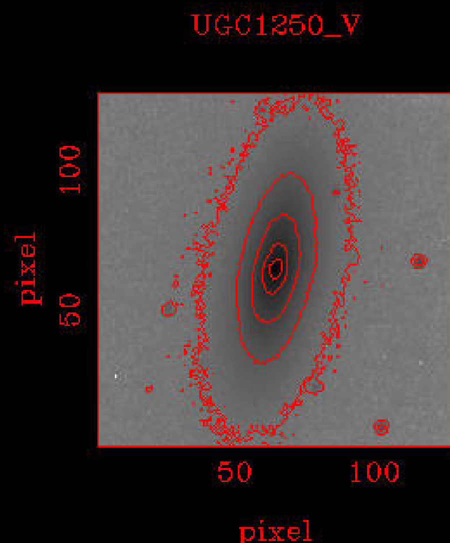
Disk surface brightness:

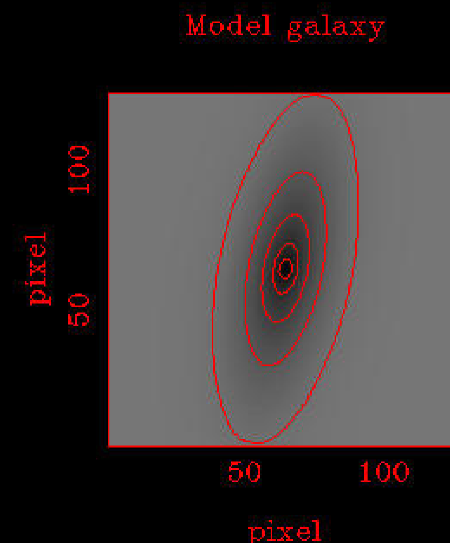
$$I_{\rm d}(r) = I_{\rm d}(0) e^{-(r/r_{\rm d})}$$

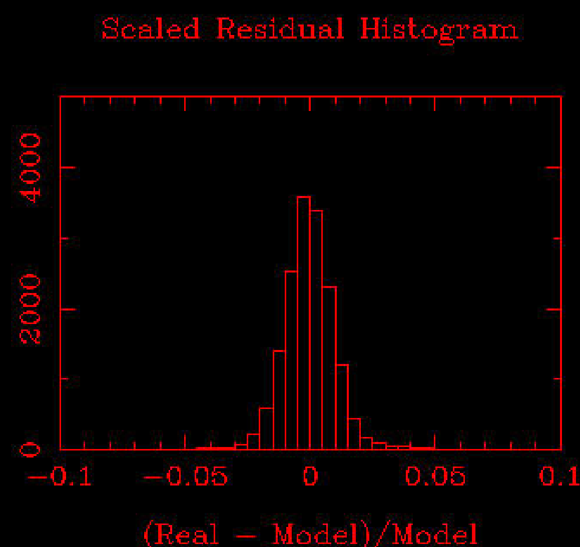
Bulge –Disk Decomposition

$$\chi^2 = \Sigma(O_i - m_i)^2/\sigma_i^2$$

UGC 1250







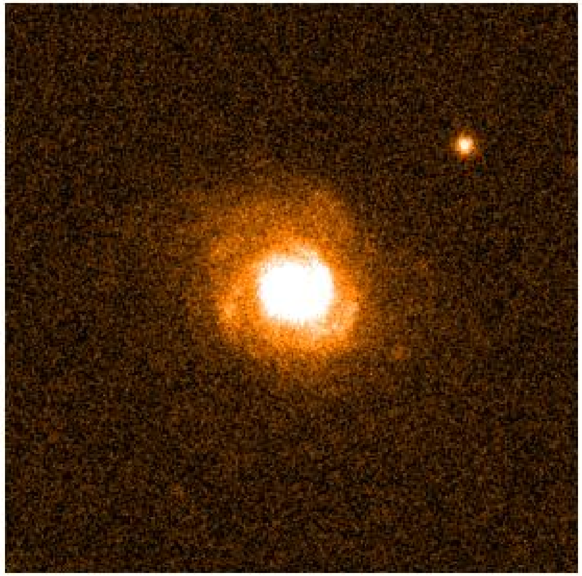
Iteration Number: 910

	BULGE	DISK
Intensity:	899380	5681.56
Scale:	12.8185	14.9379
Ellipticity:	0.52	0.67
N:	3.35107	

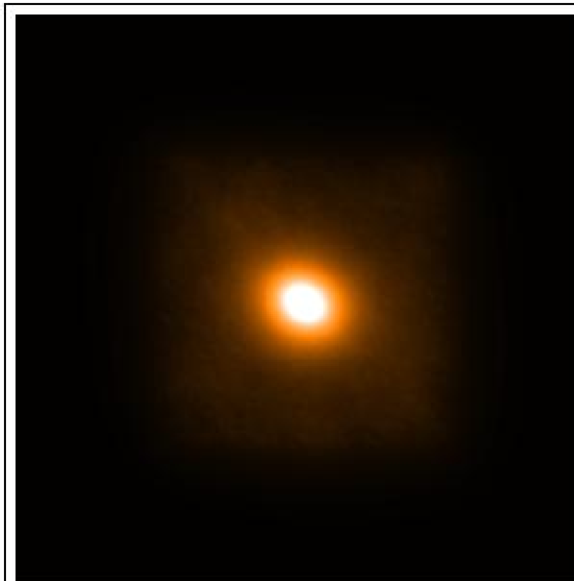
P. Intensity: 2.66454e−15
D/B: 1.51527

Reduced χ^2 0.900344

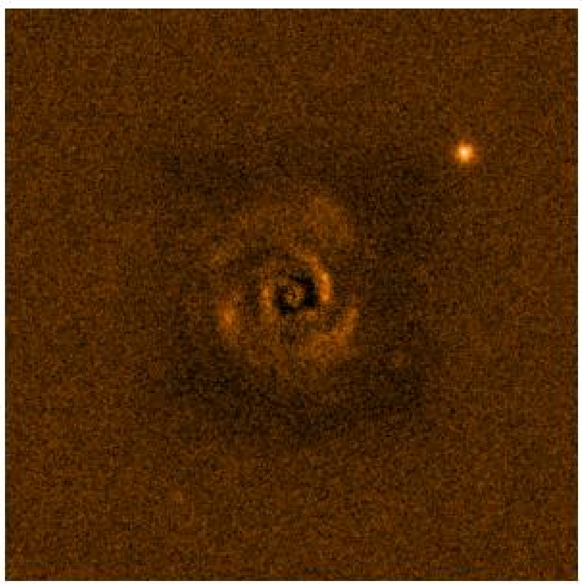
Bulge-Disk Decomposition



Galaxy 2401283 z=0.5487



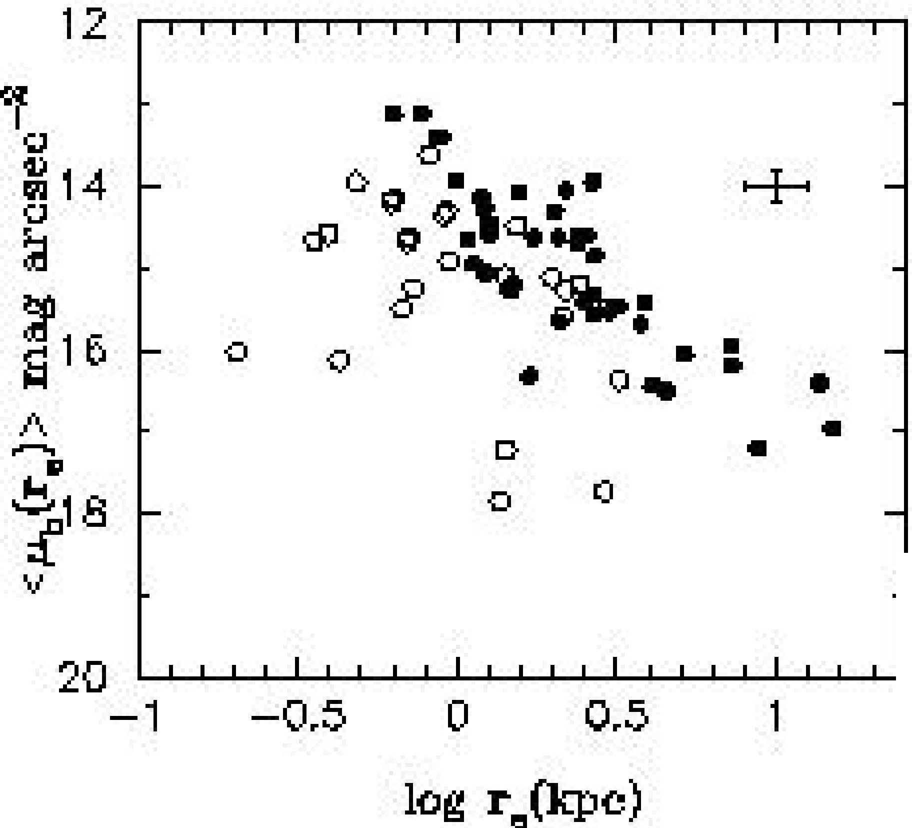
Model image

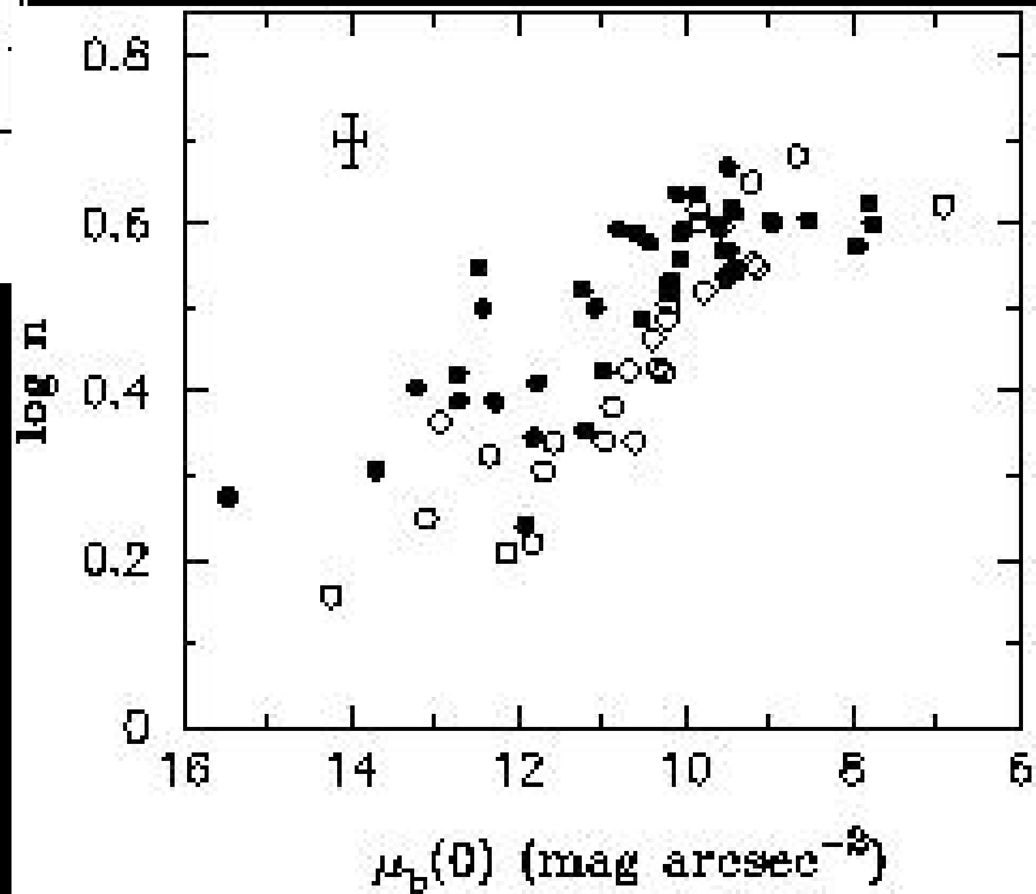


Residual image

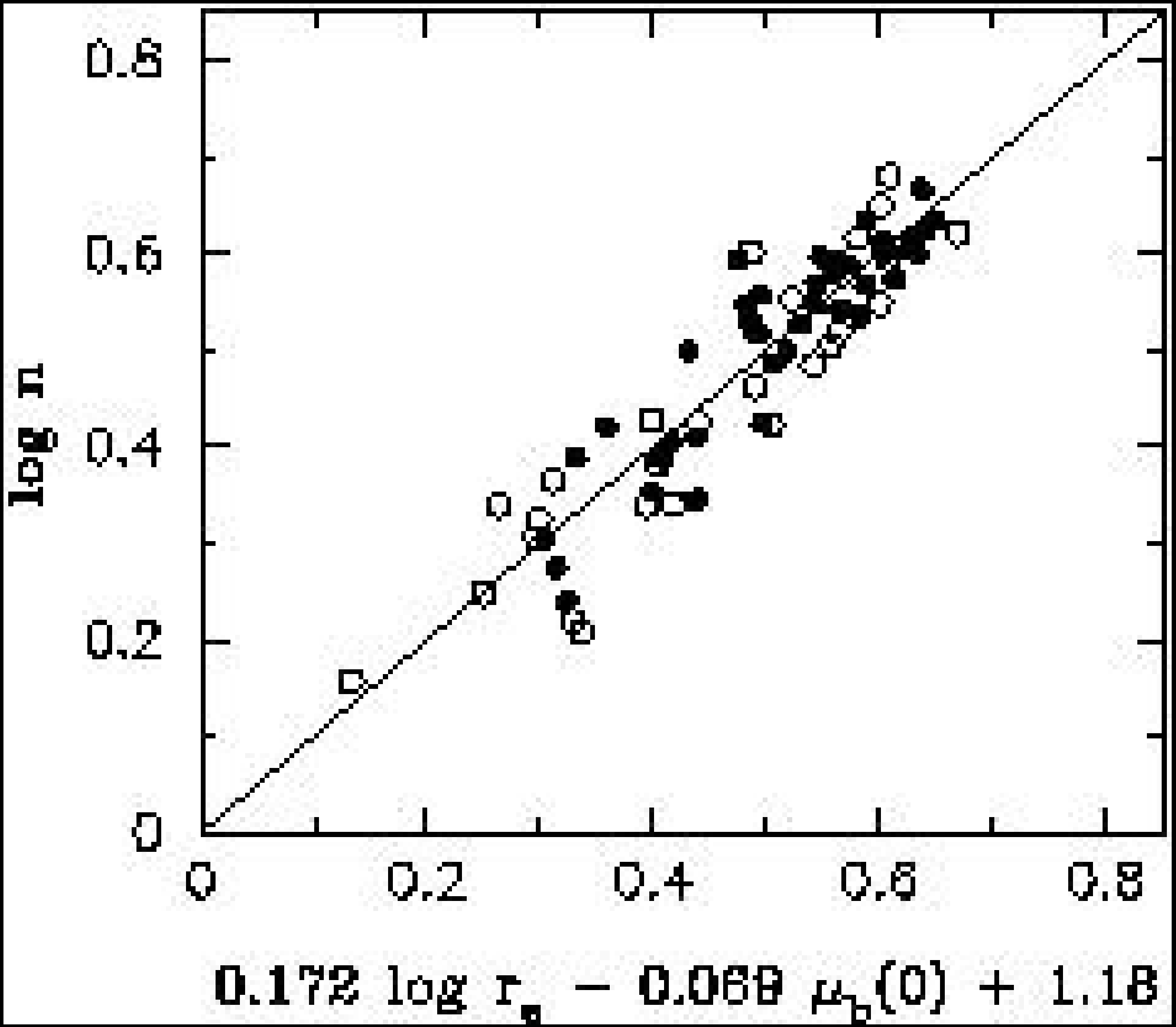
The Photometric Plane

2-D Correlations





Ellipticals and Early Type Bulges



Photometric Plane

Ellipticals and Early Type Bulges

A Mass Fundamental Plane for SuperMassive Black Holes

Table II. Complete List of SBH Mass Detection Based on Resolved Dynamical Studies

Object	Hubble Type	Distance (Mpc)	$M_\bullet$ ($10^8 M_\odot$)	$M_\bullet$ Ref. & Method	σ (km s^{-1})	$M^0_{B,T}$ (mag)	$L_{B,bulge}/L_{B,total}$	r_h/r_{res}
MW	SbI-II	0.008	$0.040^{+0.003}_{-0.003}$	1,PM	100±20	-20.08±0.50	0.34	1700
N4258	SAB(s)bc	7.2	$0.390^{+0.034}_{-0.034}$	2,MM	138±18	-20.76±0.15	0.16	880
N4486	E0pec	16.1	$35.7^{+10.2}_{-10.2}$	3,GD	345±45	-21.54±0.16	1.0	34.6
N3115	S0	9.7	$9.2^{+3.0}_{-3.0}$	4,SD	278±36	-20.19±0.20	0.64	22.8
I1459	E3	29.2	$26.0^{+11.0}_{-11.0}$	5,SD	312±41	-21.50±0.32	1.0	17.0
N4374	E1	18.7	$17^{+12}_{-6.7}$	6,GD	286±37	-21.31±0.13	1.0	10.3
N4697	E6	11.7	$1.7^{+0.2}_{-0.3}$	7,SD	163±21	-20.34±0.18	1.0	10.2
N4649	E2	16.8	$20.0^{+4.0}_{-6.0}$	7,SD	331±43	-21.43±0.16	1.0	10.1
N221	cE2	0.8	$0.025^{+0.005}_{-0.005}$	8,SD	76±10	-15.76±0.18	1.0	10.1
N5128	S0pec	4.2	$2.0^{+3.0}_{-1.4}$	9,GD	145±25	-20.78±0.15	0.64	8.41
M81	SA(s)ab	3.9	$0.70^{+0.2}_{-0.1}$	10,GD	174±17	-20.42±0.26	0.33	5.50
N4261	E2	31.6	$5.4^{+1.2}_{-1.2}$	11,GD	290±38	-21.14±0.20	1.0	3.77
N4564	E6	15.0	$0.56^{+0.03}_{-0.08}$	7,SD	153±20	-19.00±0.18	1.0	2.96
CygA	E	240	$25.0^{+7.0}_{-7.0}$	12,GD	270±87	-20.03±0.27	1.0	2.65
N2787	SB(r)0	7.5	$0.90^{+6.89}_{-0.69}$	13,GD	210±23	-18.12±0.39	0.64	2.53
N3379	E1	10.6	$1.35^{+0.73}_{-0.73}$	14,SD	201±26	-19.94±0.20	1.0	2.34
N5845	E*	25.9	$2.4^{+0.4}_{-1.4}$	7,SD	275±36	-18.80±0.25	1.0	2.28
N3245	SB(s)b	20.9	$2.1^{+0.5}_{-0.5}$	15,GD	211±19	-20.01±0.25	0.33	2.10
N4473	E5	15.7	$1.1^{+0.5}_{-0.8}$	7,SD	188±25	-19.94±0.14	1.0	1.84
N3608	E2	22.9	$1.9^{+1.0}_{-0.6}$	7,SD	206±27	-20.11±0.17	1.0	1.82
N4342	S0	16.7	$3.3^{+1.9}_{-1.1}$	16,GD	261±34	-17.74±0.20	0.64	1.79
N7052	E	66.1	$3.7^{+2.6}_{-1.5}$	17,GD	261±34	-21.33±0.38	1.0	1.53
N4291	E3	26.2	$3.1^{+0.8}_{-2.3}$	7,SD	269±35	-19.82±0.35	1.0	1.52
N6251	E	104	$5.9^{+2.0}_{-2.0}$	18,GD	297±39	-21.94±0.28	1.0	1.19
N3384	SB(s)0-	11.6	$0.16^{+0.01}_{-0.02}$	7,SD	151±20	-19.59±0.15	0.64	1.12
N7457	SA(rs)0-	13.2	$0.035^{+0.011}_{-0.014}$	7,SD	73±10	-18.74±0.24	0.64	0.92
N1023	S0	11.4	$0.44^{+0.06}_{-0.06}$	7,SD	201±14	-20.20±0.17	0.64	0.89
N821	E6	24.1	$0.37^{+0.24}_{-0.08}$	7,SD	196±26	-20.50±0.21	1.0	0.74
N3377	E5	11.2	$1.00^{+0.9}_{-0.1}$	7,SD	131±17	-19.16±0.13	1.0	0.74
N2778	E	22.9	$0.14^{+0.08}_{-0.09}$	7,SD	171±22	-18.54±0.33	1.0	0.39

SMBH Systematics

All nearby galaxies with a significant bulge contain super-massive black holes

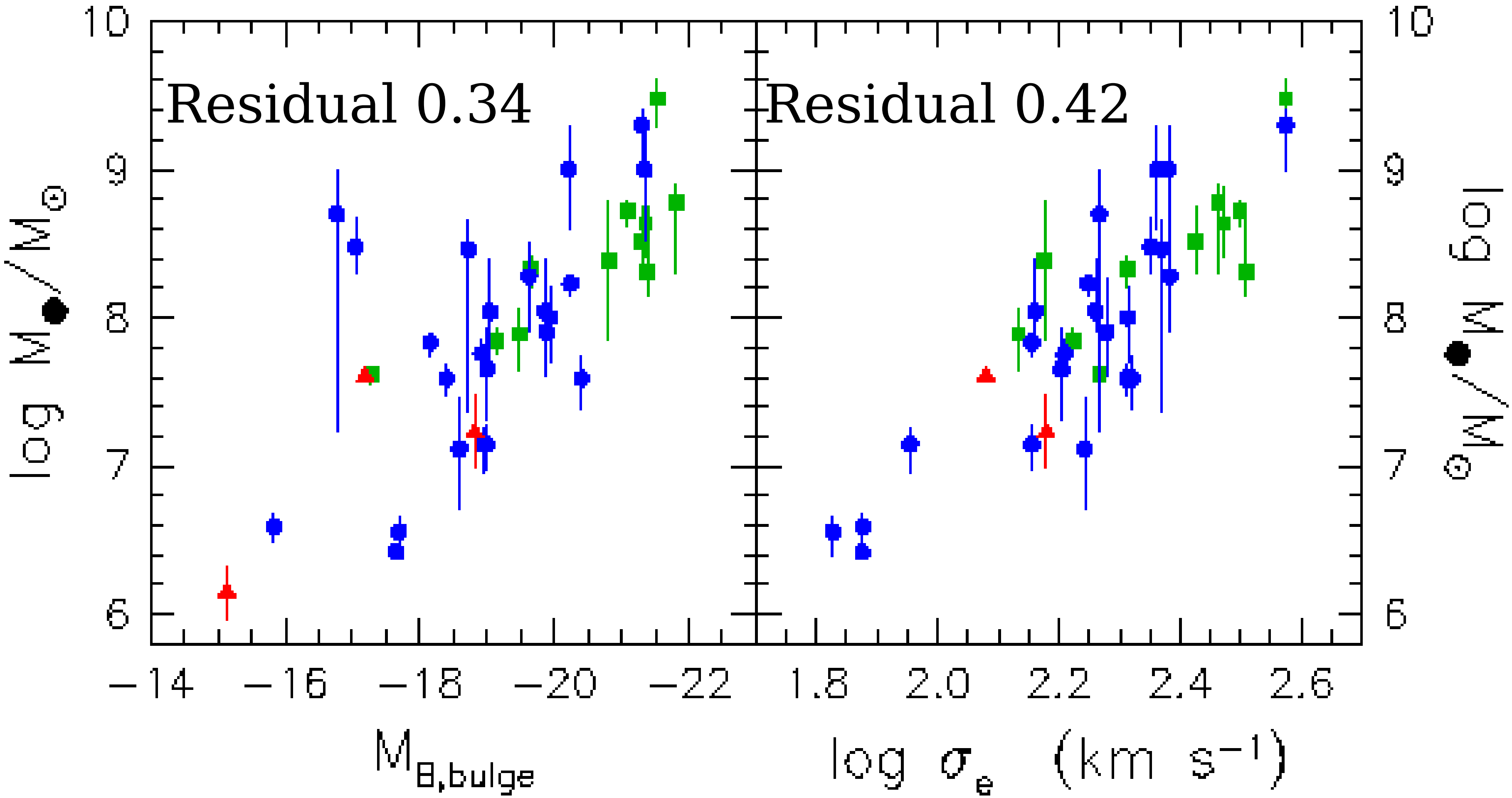
The black hole mass is proportional to the bulge mass, and to the fourth power of the central velocity dispersion

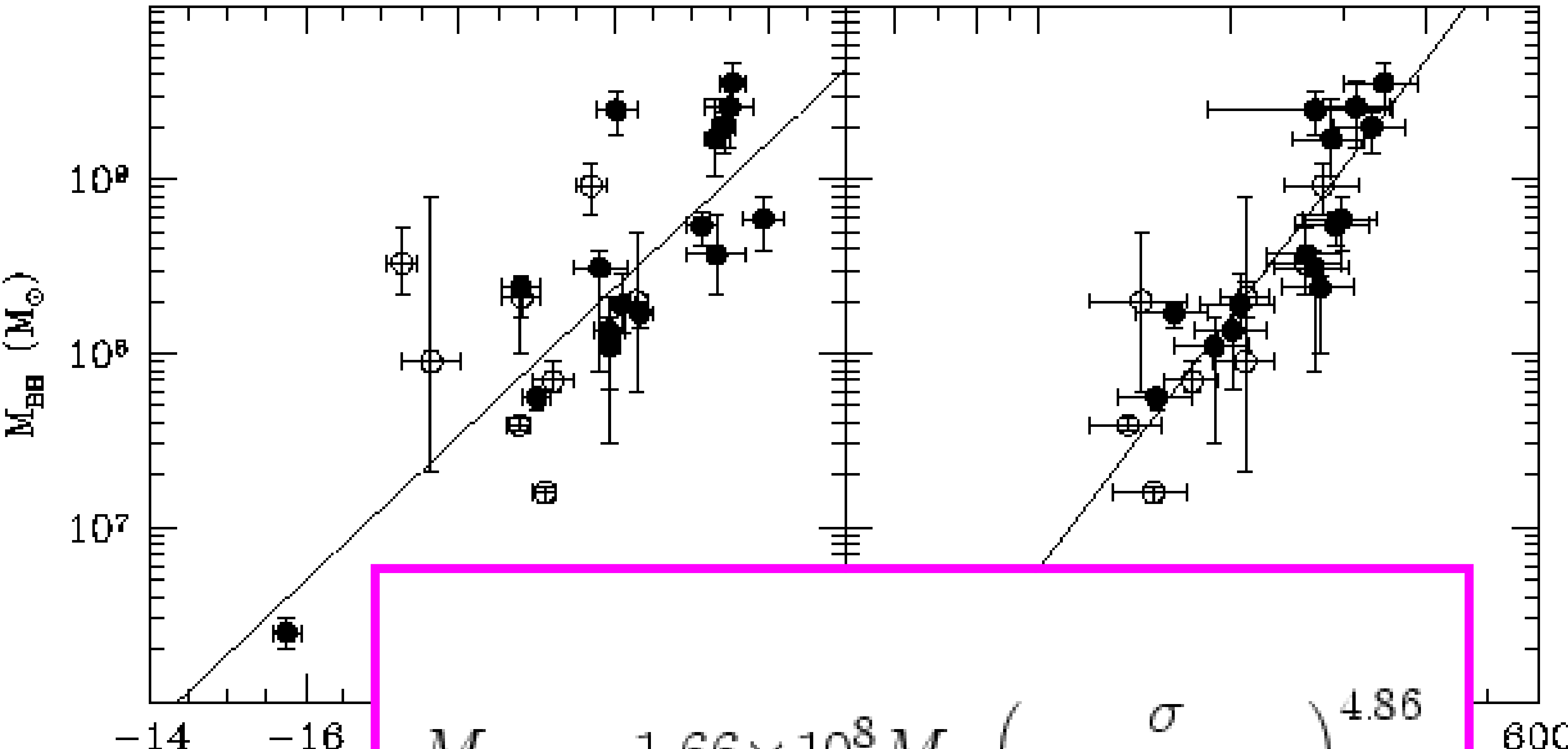
The relation has been extended to lower masses, as in globular clusters

$$\mathfrak{M} \sim \sigma^4$$

SuperMassive Black Hole Systematics

Ferrarese et al





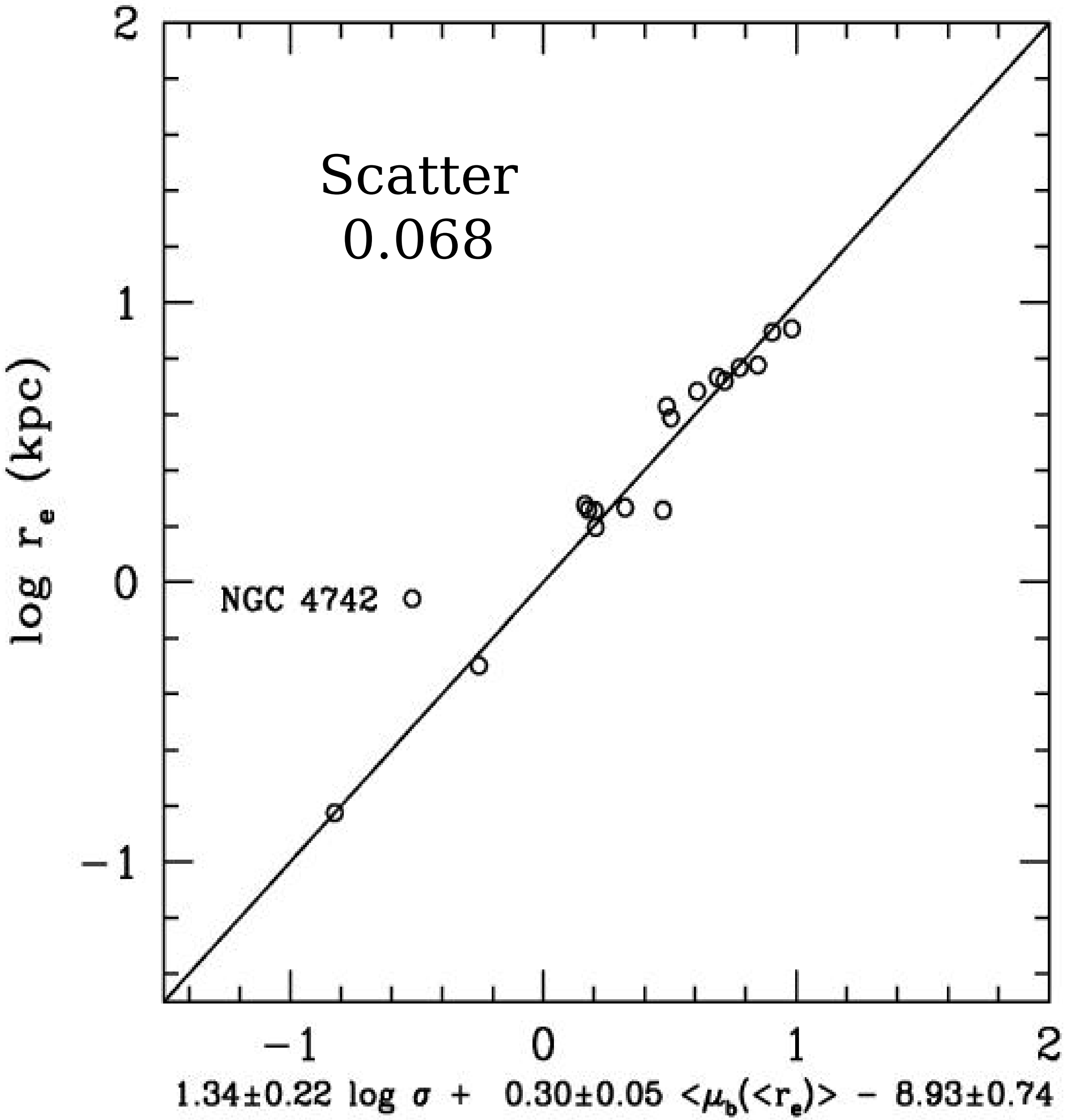
$$M_{\mathrm{BH}} = 1.66\times10^{8}M_\odot\left(\frac{\sigma}{200\,\mathrm{km\,s^{-1}}}\right)^{4.86}$$

SMBH Host Galaxies

TABLE 1. BASIC PARAMETERS FOR ELLIPTICAL GALAXIES WITH MEASURED BLACK HOLE MASS.

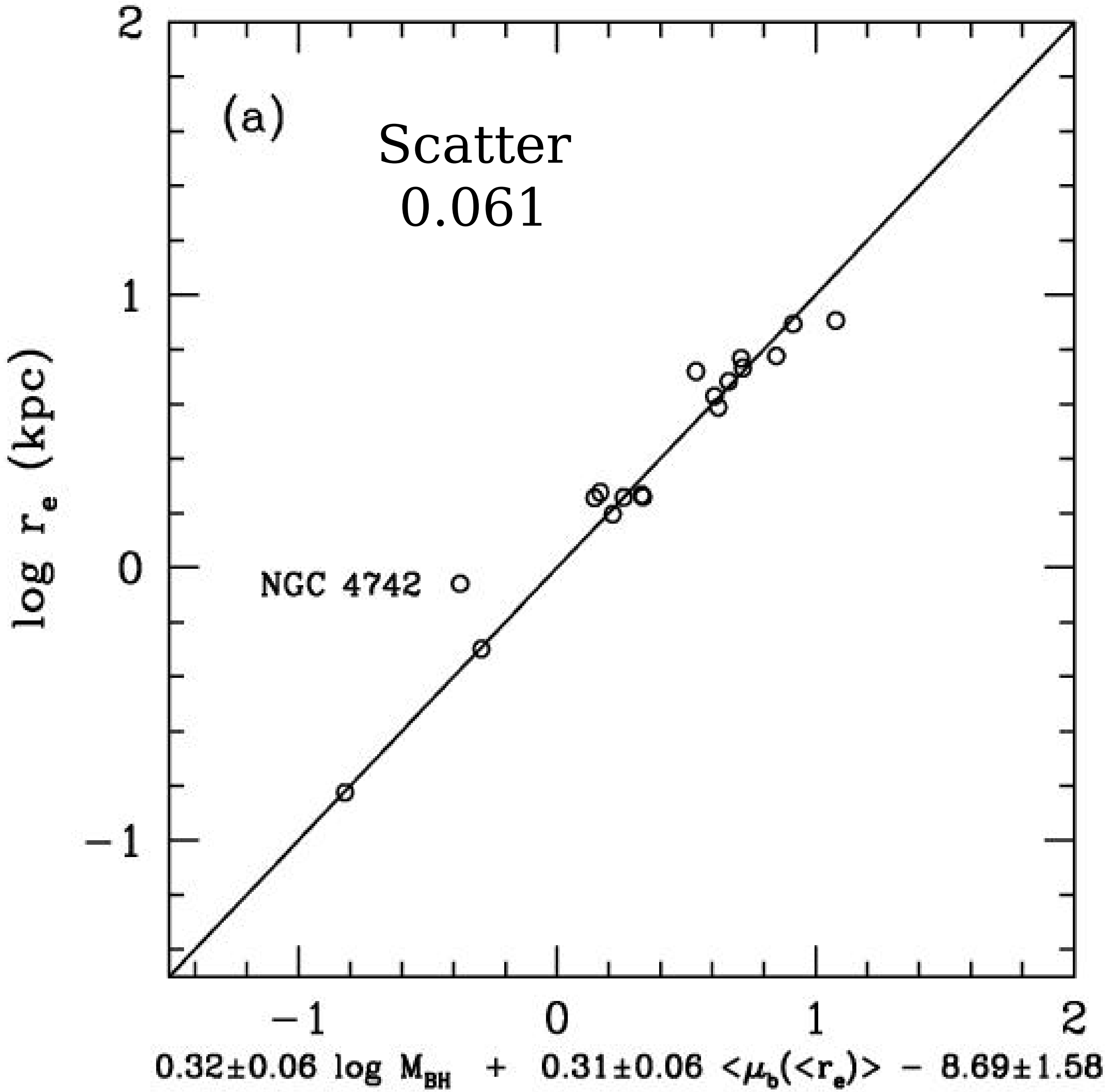
Object	Type	Distance (Mpc)	M_{BH} ($10^8\ M_\odot$)	σ (km s^{-1})	L_B (mag)	log r_e (kpc)	$\langle\mu_B(<r_e)\rangle$ (mag arcsec^{-2})
NGC 221/M32	-6.0	0.80	$2.5^{+0.5}_{-0.5} \times 10^6$	75 ± 10	-15.80 ± 0.18	-0.83	18.69
NGC 821	-5.0	24.1	$3.7^{+2.4}_{-0.8} \times 10^7$	209 ± 26	-20.42 ± 0.21	0.72	21.85
NGC 2778	-5.0	22.9	$1.4^{+0.8}_{-0.9} \times 10^7$	175 ± 22	-18.58 ± 0.33	0.26	21.38
NGC 3377	-5.0	11.2	$1.0^{+0.9}_{-0.1} \times 10^8$	145 ± 17	-19.18 ± 0.13	0.26	20.76
NGC 3379	-5.0	10.6	$1.0^{+0.6}_{-0.5} \times 10^8$	206 ± 26	-19.81 ± 0.20	0.26	20.16
NGC 3608	-5.0	22.9	$1.9^{+1.0}_{-0.6} \times 10^8$	182 ± 27	-20.07 ± 0.17	0.59	21.41
NGC 4261	-5.0	31.6	$5.2^{+1.0}_{-1.1} \times 10^8$	315 ± 38	-21.23 ± 0.20	0.77	21.25
NGC 4291	-5.0	26.2	$3.1^{+0.8}_{-2.3} \times 10^8$	242 ± 35	-19.72 ± 0.35	0.27	20.25
NGC 4374/M84	-5.0	18.4	$1.0^{+2.0}_{-0.6} \times 10^9$	296 ± 37	-21.40 ± 0.31	0.68	20.81
NGC 4473	-5.0	15.7	$1.1^{+0.4}_{-0.8} \times 10^8$	190 ± 25	-19.86 ± 0.14	0.28	20.19
NGC 4486/M87	-4.0	16.1	$3.4^{+1.0}_{-1.0} \times 10^9$	375 ± 45	-21.71 ± 0.16	0.91	21.60
NGC 4564	-5.0	15.0	$5.6^{+0.3}_{-0.8} \times 10^7$	162 ± 20	-18.94 ± 0.18	0.19	20.64
NGC 4697	-5.0	11.7	$1.7^{+0.2}_{-0.1} \times 10^8$	177 ± 10	-20.20 ± 0.18	0.63	21.41
NGC 4649/M60	-5.0	16.8	$2.0^{+0.4}_{-0.6} \times 10^9$	385 ± 43	-21.30 ± 0.16	0.78	21.10
NGC 4742	-5.0	15.5	$1.4^{+0.4}_{-0.5} \times 10^7$	90 ± 05	-19.03 ± 0.10	-0.06	19.36
NGC 5845	-5.0	25.9	$2.4^{+0.4}_{-1.4} \times 10^8$	234 ± 36	-18.92 ± 0.25	-0.30	18.38
NGC 7052	-5.0	71.4	$4.0^{+2.8}_{-1.6} \times 10^8$	266 ± 34	-21.43 ± 0.38	0.89	22.01
IC 1459	-5.0	29.2	$1.5^{+1.0}_{-1.0} \times 10^9$	340 ± 41	-21.45 ± 0.32	0.73	20.81
NGC 6251	-5.0	107.0	$6.1^{+2.0}_{-2.1} \times 10^8$	290 ± 39	-21.96 ± 0.28	1.31	—
CygA	-5.0	240.0	$2.9^{+0.7}_{-0.7} \times 10^9$	270 ± 90	-20.03 ± 0.27	—	—

The Fundamental Plane for SMBH Galaxies

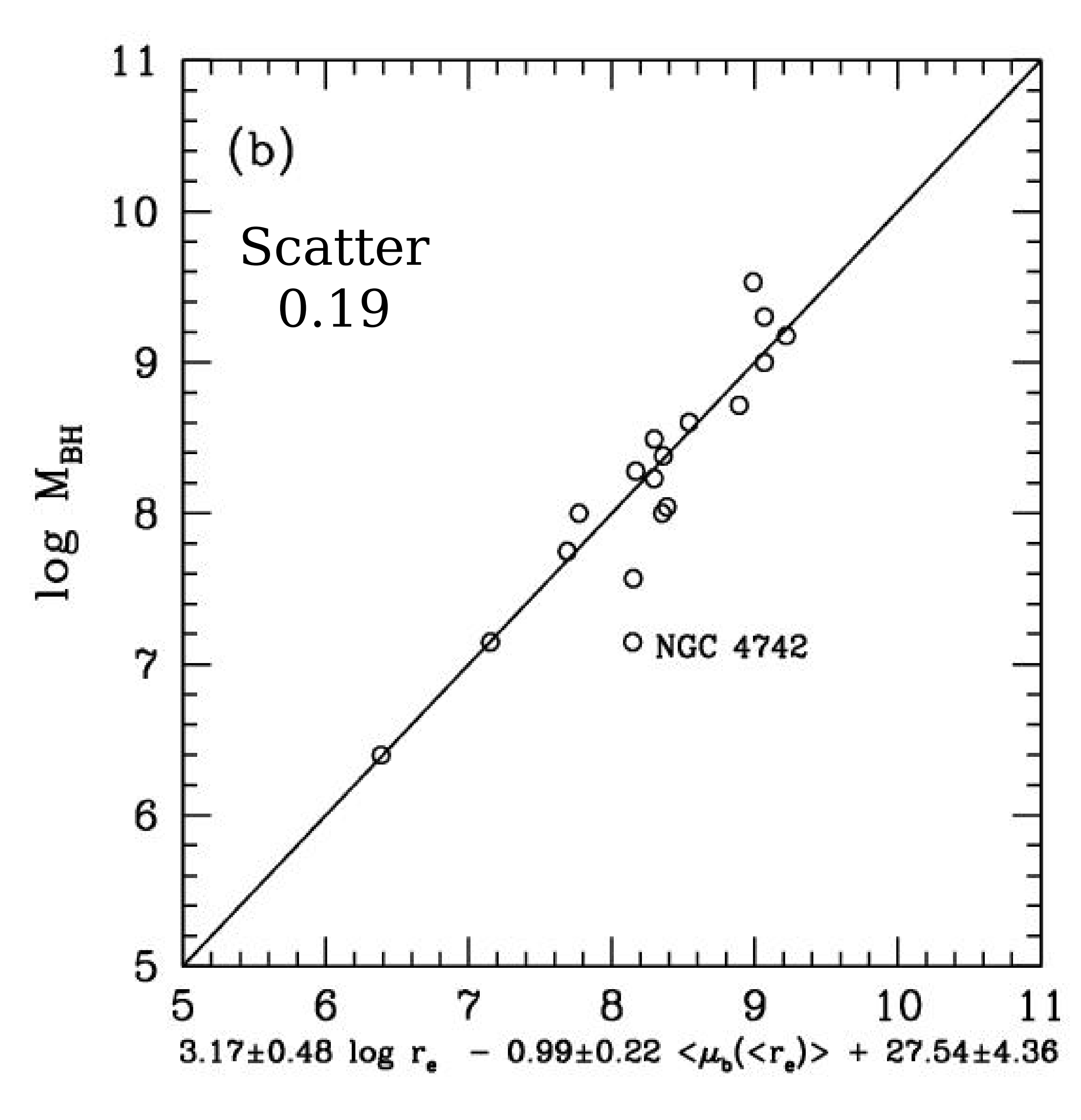


Scatter
0.068

The Mass Fundamental Plane for SMBH



Black Hole Mass from the Mass Fundamental Plane



The Fundamental Plane

The Fundamental Plane

I_e + const

Observed global properties of elliptical galaxies form a two-dimensional family.

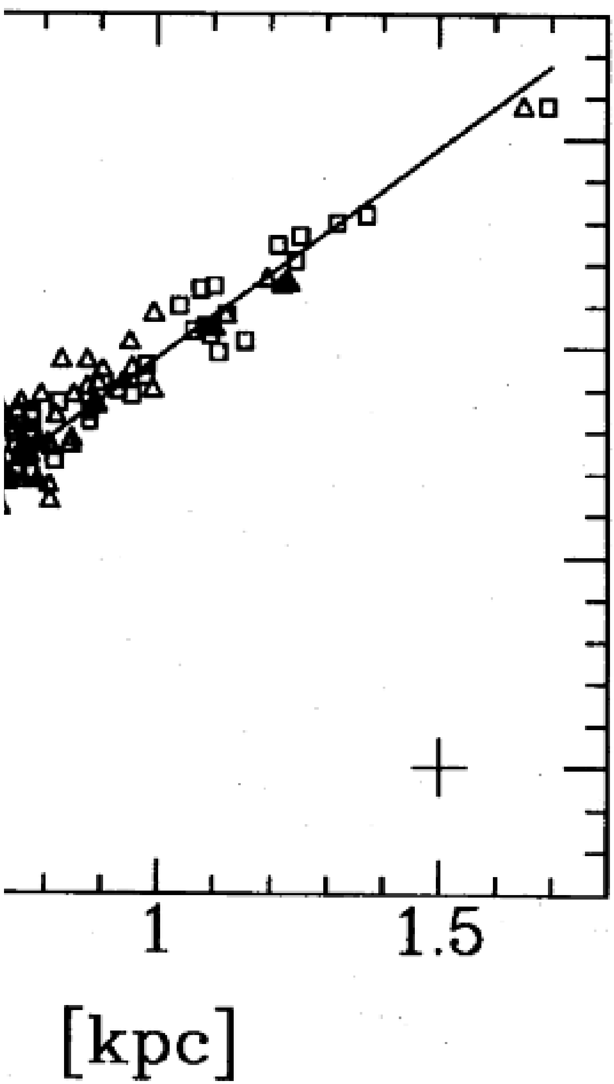
The best representation of this surface is:

$$r_e \sim \sigma^A \langle I \rangle_e^B$$

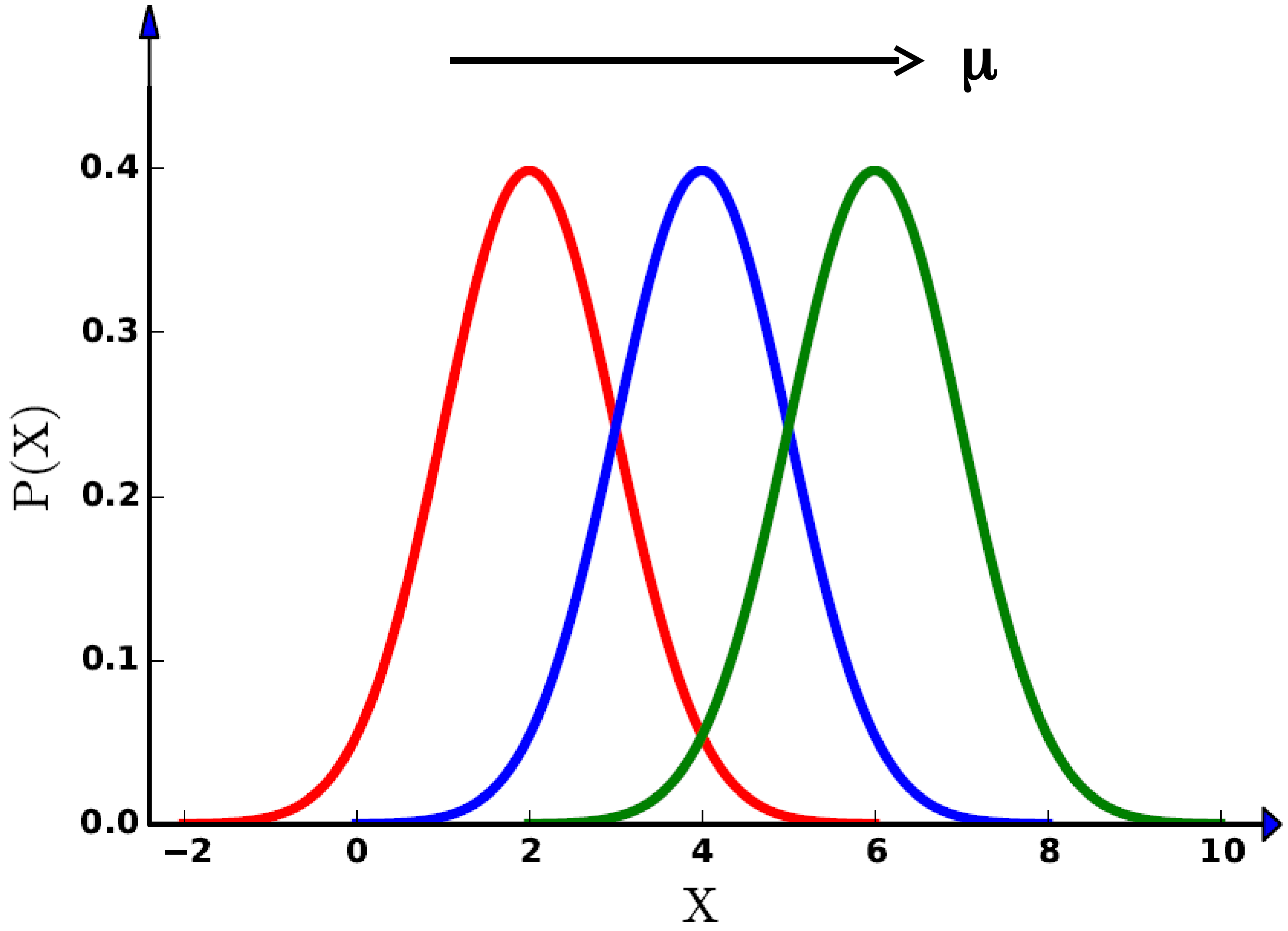
The distribution is a plane in log space.

Thickness of observed plane is due to measurement errors.

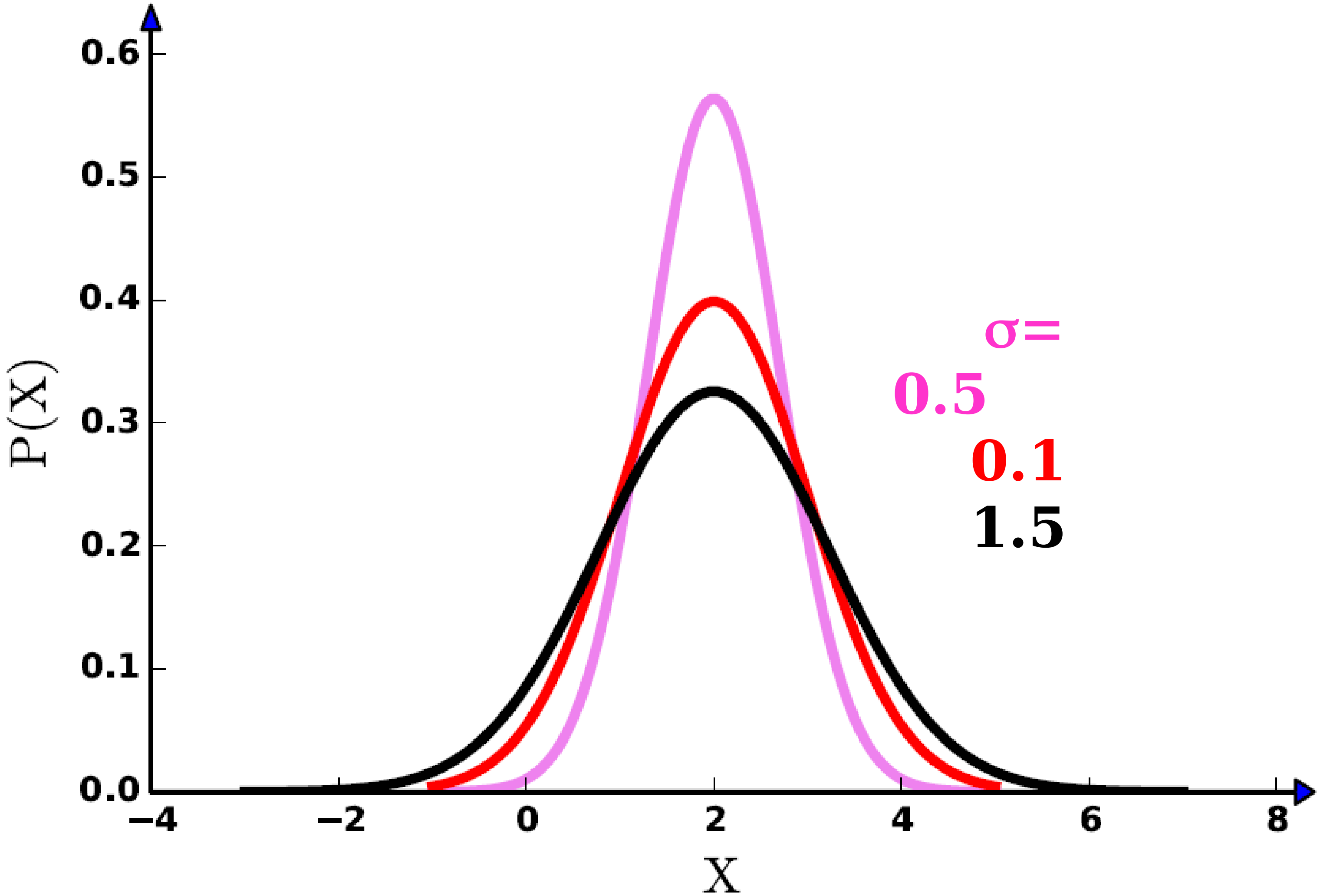
Other global galaxy parameter correlations follow from this fundamental plane.



The Gaussian Distribution



The Gaussian Distribution



Maximum Likelihood Method

For very small probabilities, the chosen parameter values are probably not right. Coversely, the given data set should not be too improbable for the correct set of parameters.

The probability of obtaining the data, given the parameters, is identified with the *likelihood* of the parameters, given the data.

The best fit parameters are accepted as those values which maximise the likelihood of the parameters, given the data. This is the Maximum Likelihood Method.