MAT 395/495

Scientific Data Analysis and Computing

Summer 2020

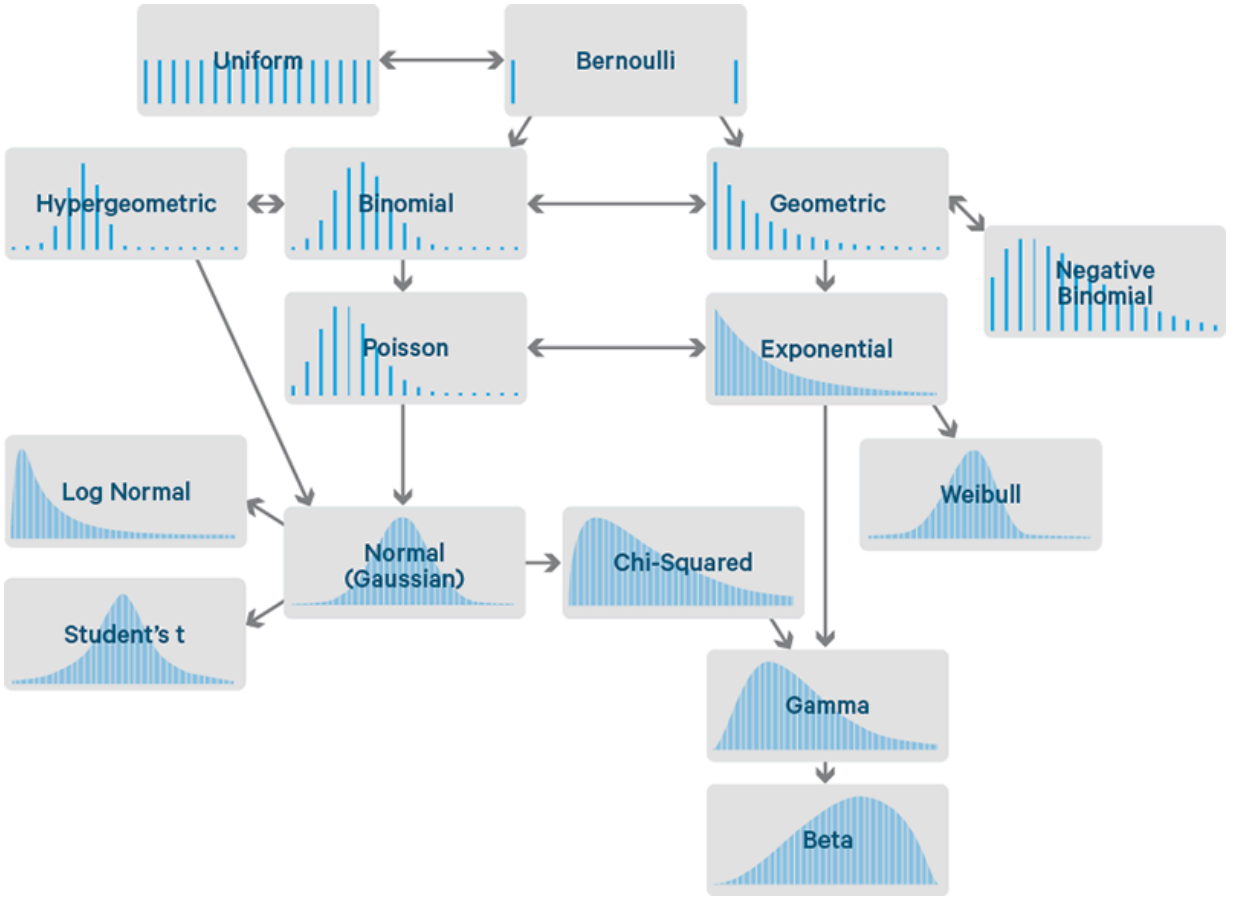
Topic 8
Understanding Distributions

SIDDHA PIMPUTKAR

ASSISTANT PROFESSOR

MATERIALS SCIENCE &
ENGINEERING DEPARTMENT

LEHIGH UNIVERSITY



Distributions

Reading

Data Reduction and Error Analysis for the Physical Sciences
by P.R. Bevington, D.K. Robinson, 3rd Ed.

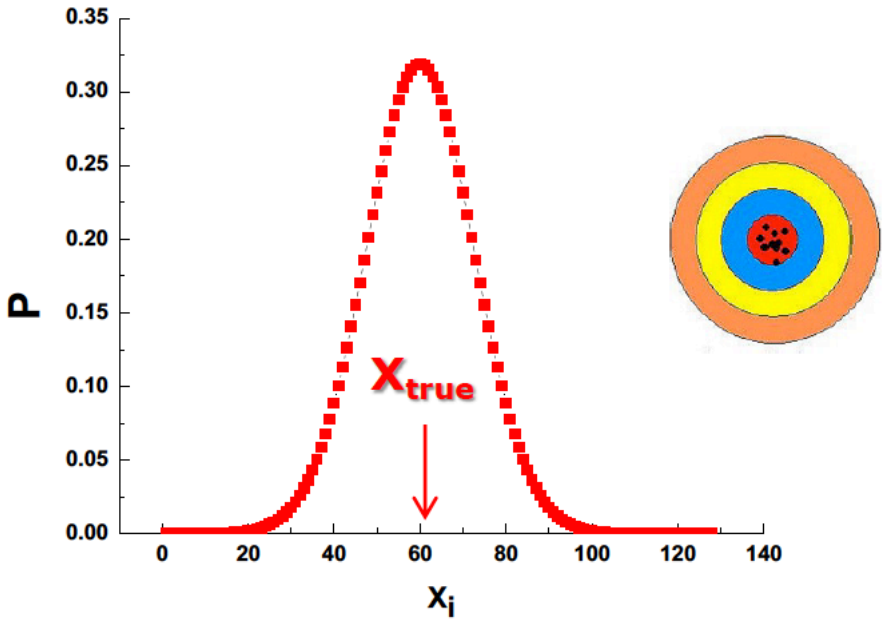
Ch. 2 Probability Distributions: pg. 17—33

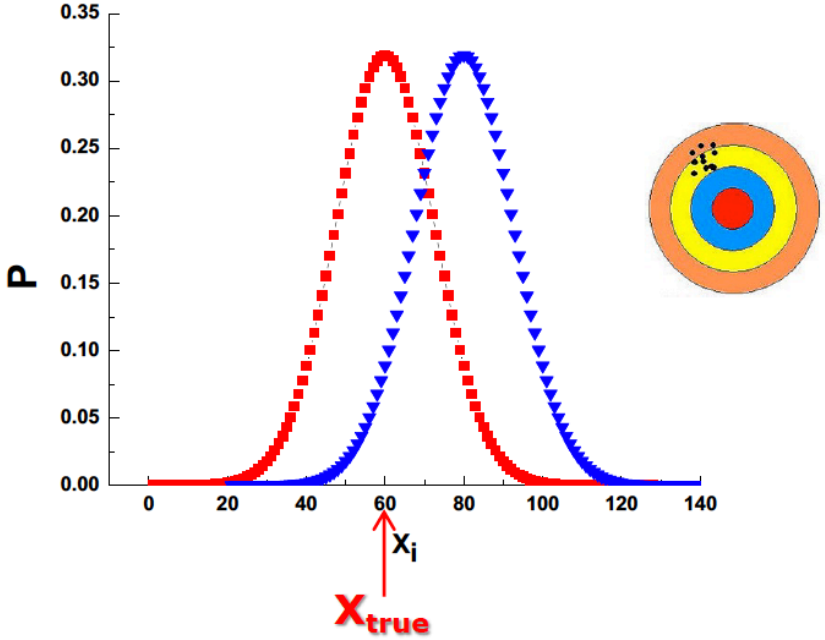
An Introduction to Error Analysis
by J.R. Taylor, 2nd Ed.

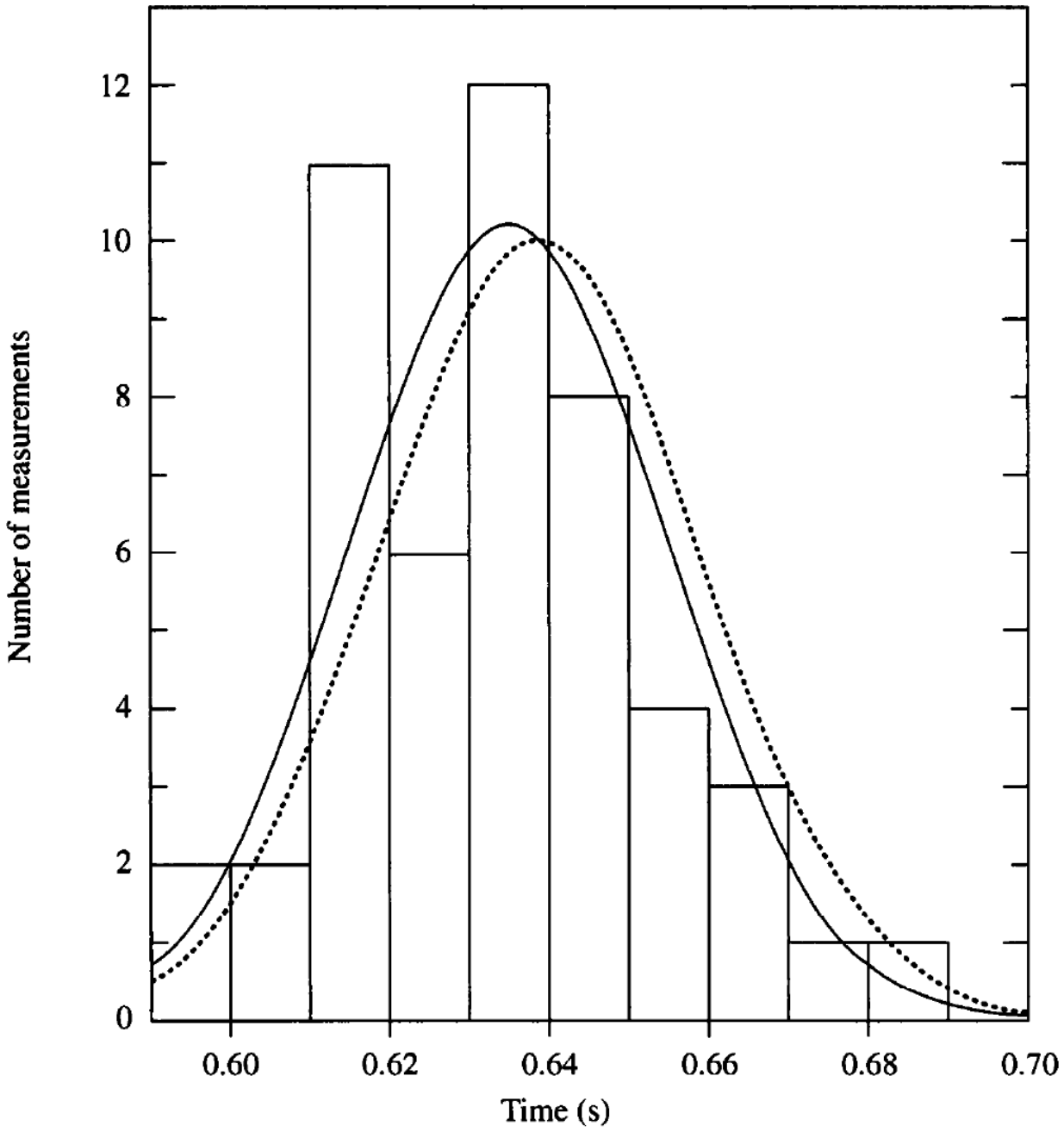
Ch. 5 Normal Distribution

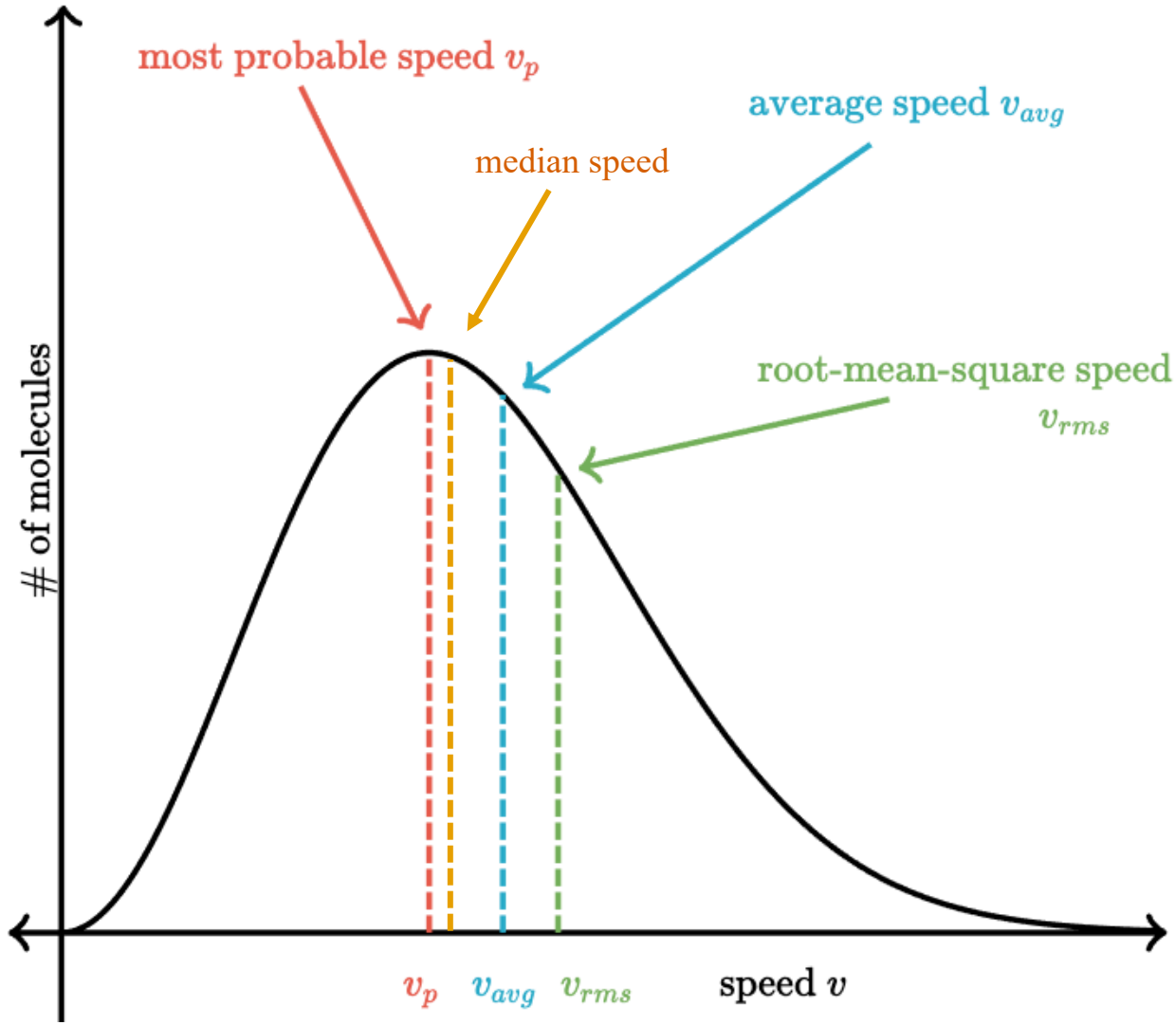
Ch. 10 Binomial Distribution

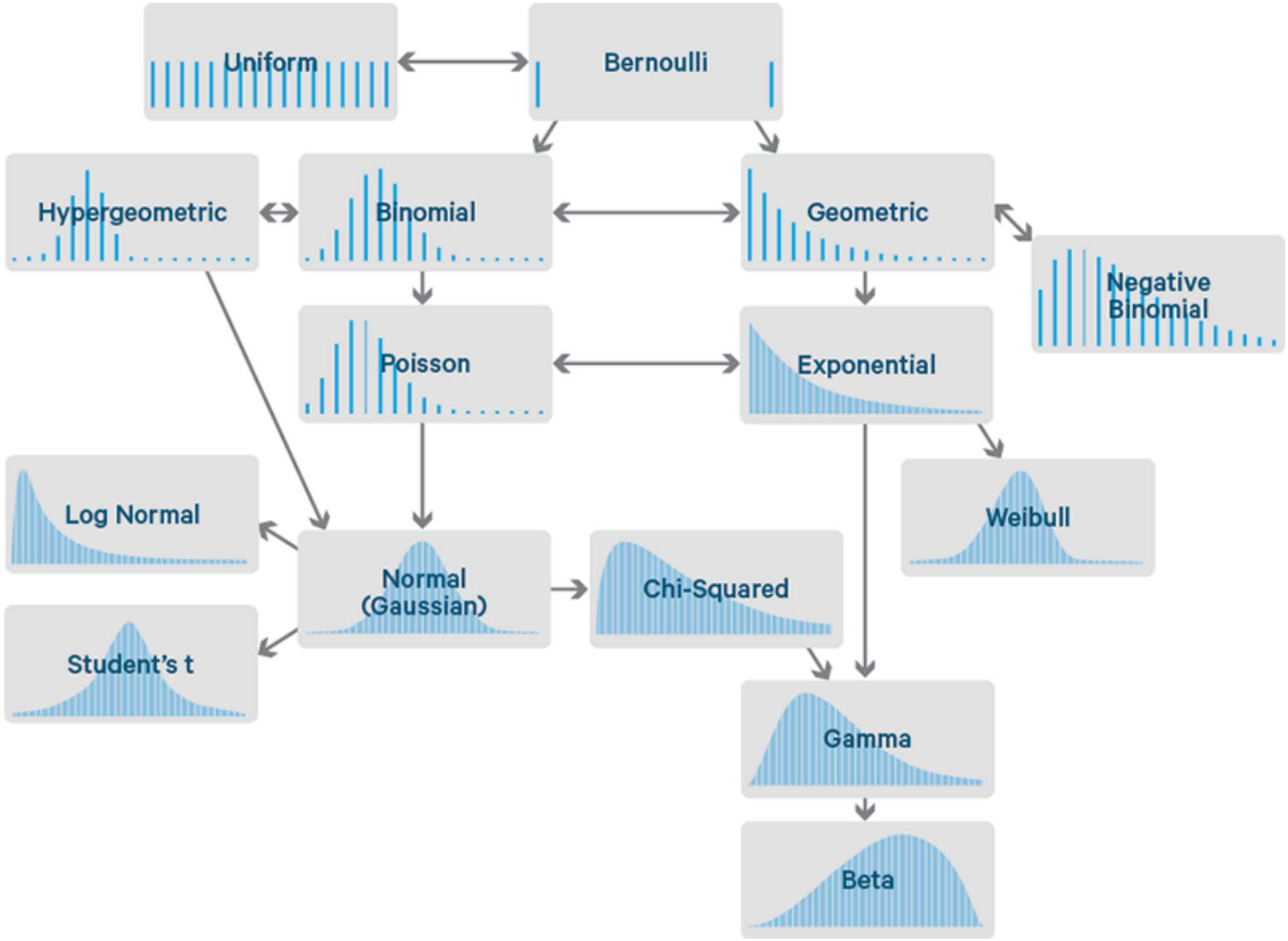
Ch. 11 Poisson Distribution

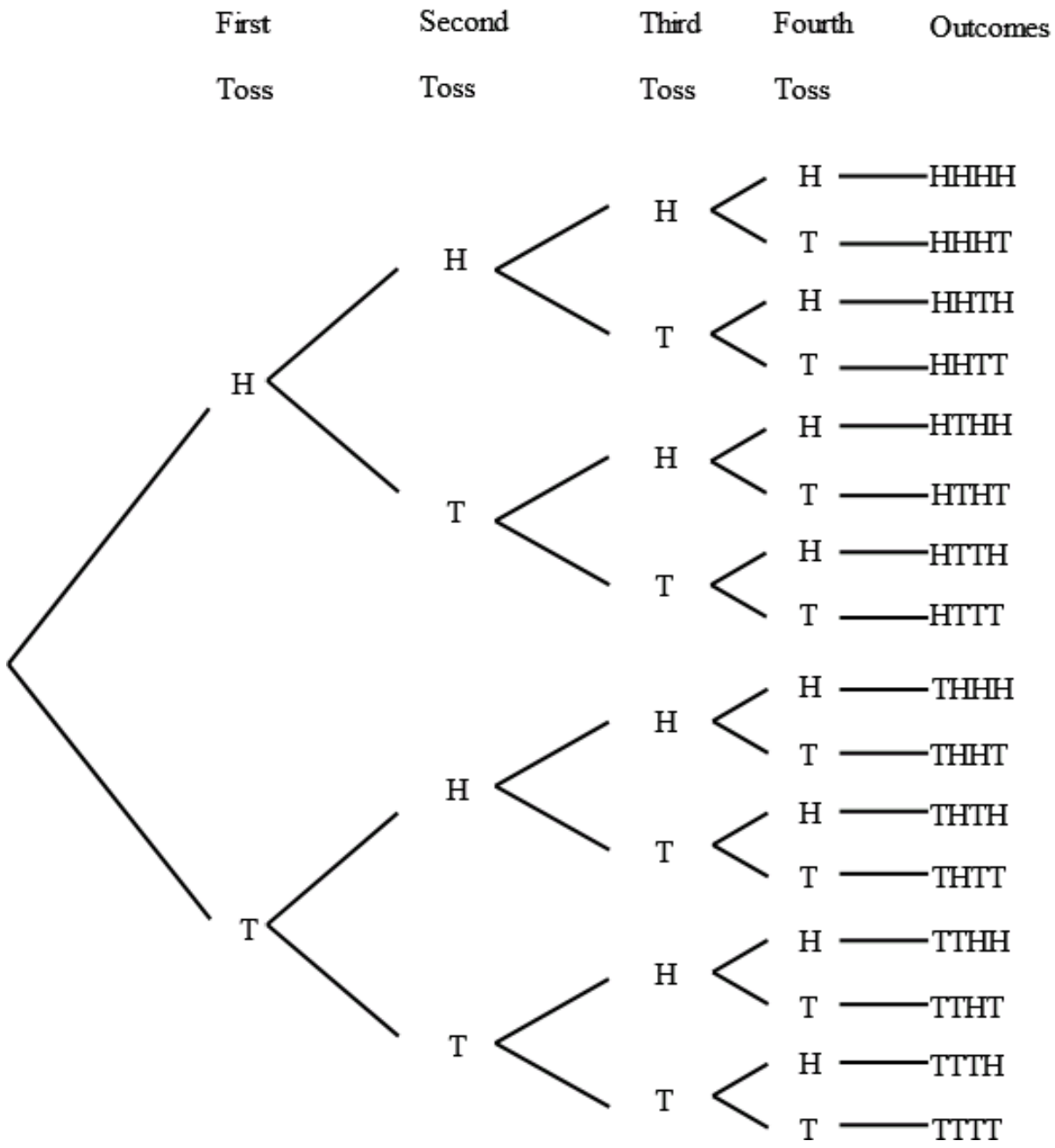

P
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
20
40
60
80
100
120
140
X_i
X_{true}

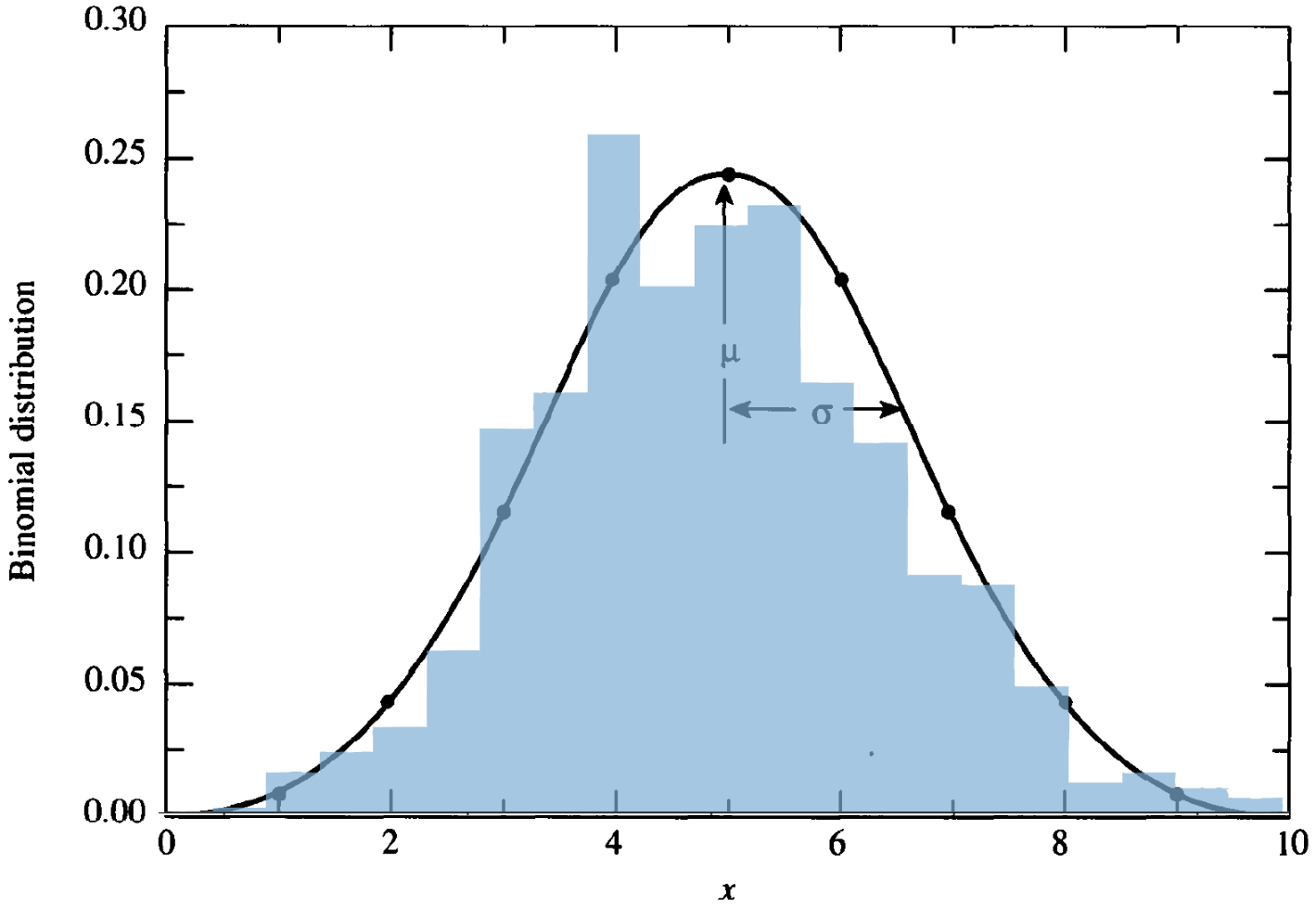

P
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
20
40
60
80
100
120
140
X_i
X_{true}


Number of measurements
0
2
4
6
8
10
12
0.60
0.62
0.64
0.66
0.68
0.70
Time (s)

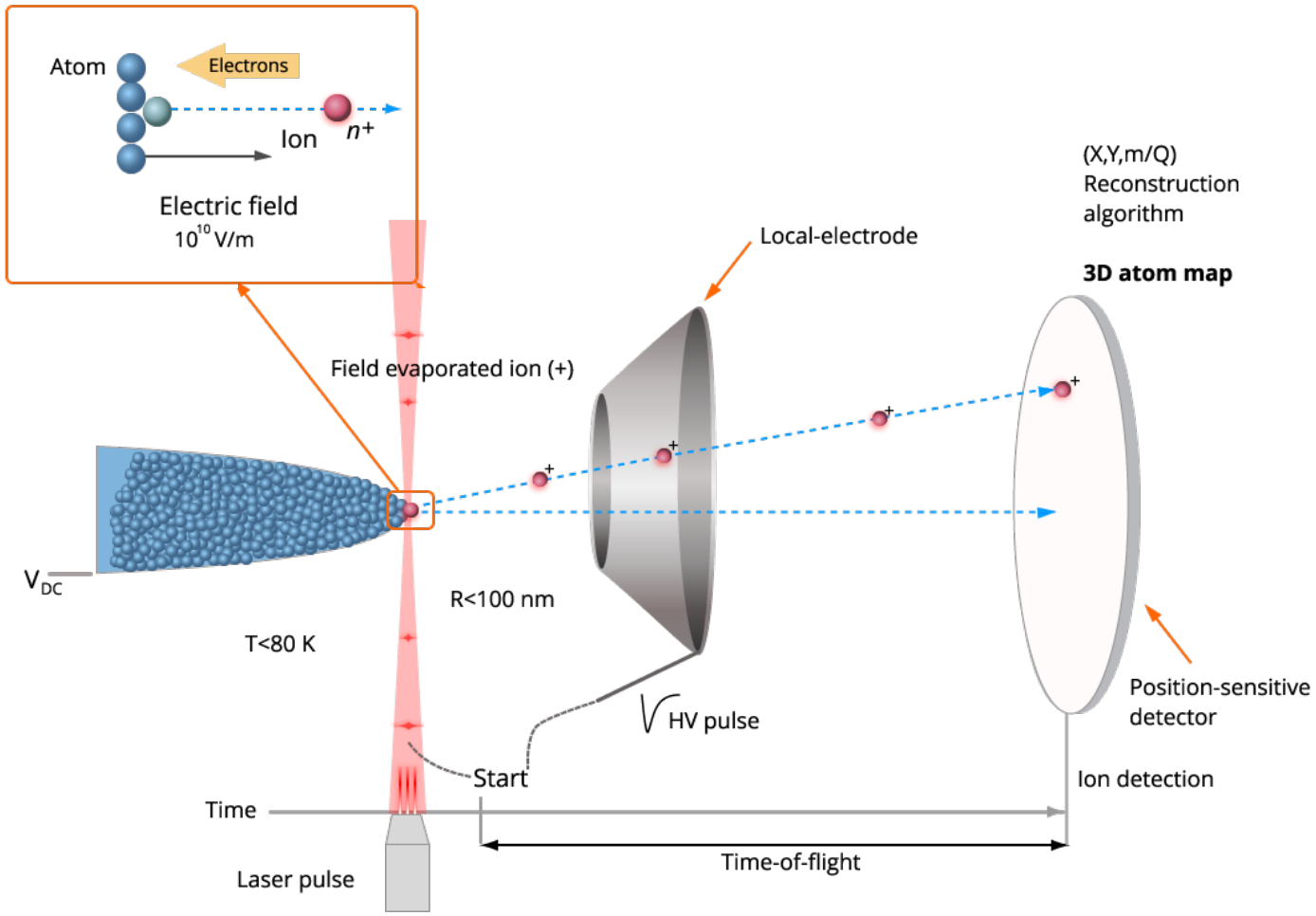

most probable speed v_p
median speed
average speed v_{avg}
root-mean-square speed
v_{rms}
# of molecules
v_p
v_{avg}
v_{rms}
speed v

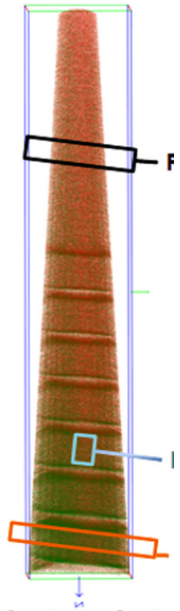

Uniform
Bernoulli
Hypergeometric
Binomial
Geometric
Negative Binomial
Poisson
Exponential
Log Normal
Weibull
Normal (Gaussian)
Chi-Squared
Student's t
Gamma
Beta

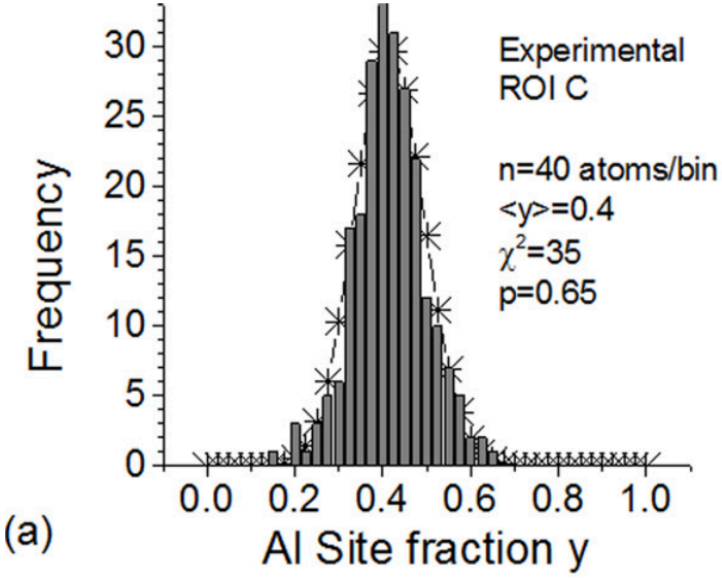

First Toss
Second Toss
Third Toss
Fourth Toss
Outcomes
H
T
H
T
H
T
H
T
H
T
H
T
H
T
H
T
H
T
H
T
H
T
H
T
H
T
HHHH
HHHT
HHTH
HHTT
HTHH
HTHT
HTTH
HTTT
THHH
THHT
THTH
THTT
TTHH
TTHT
TTTH
TTTT


Binomial distribution
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
x
μ
σ

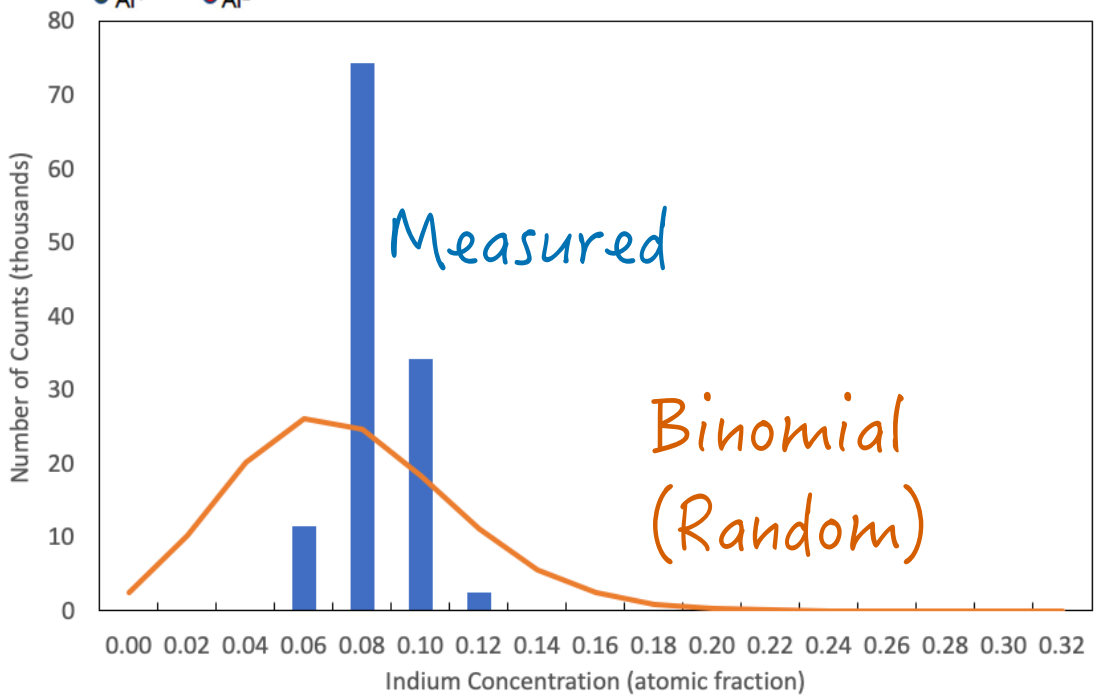
Example: Atomic distribution of In/Ga on Group III site of InGaN. Use of atom probe:

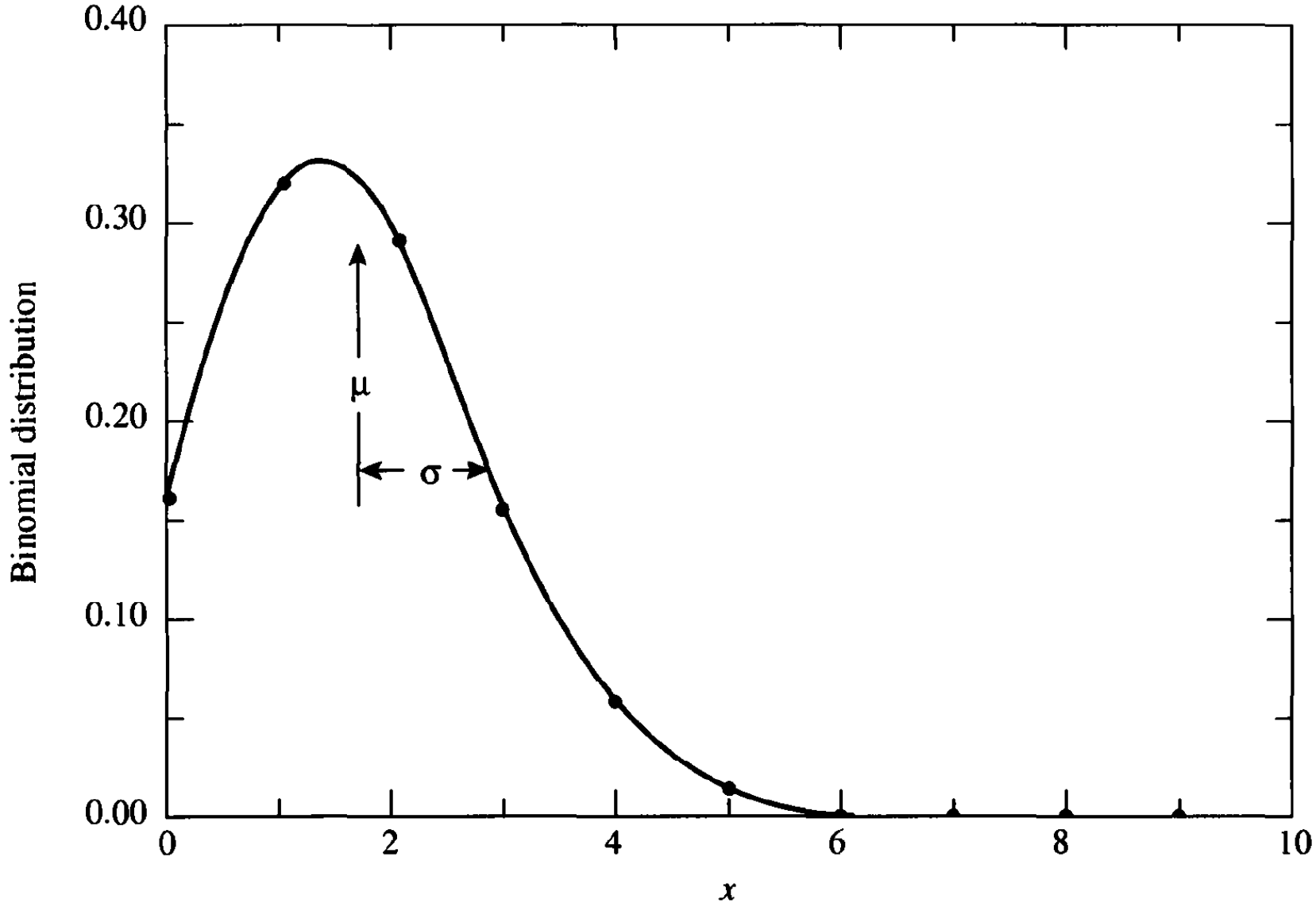


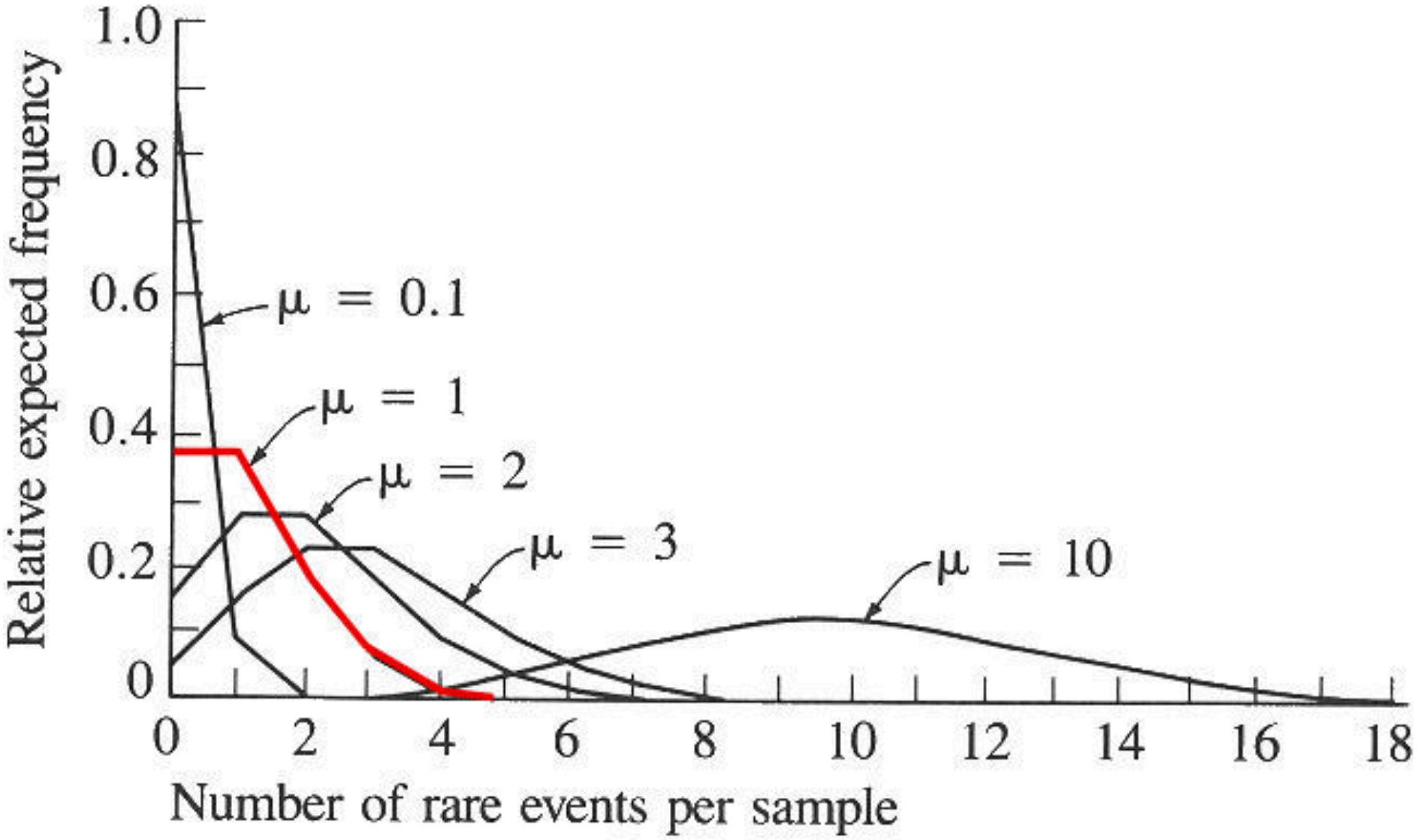


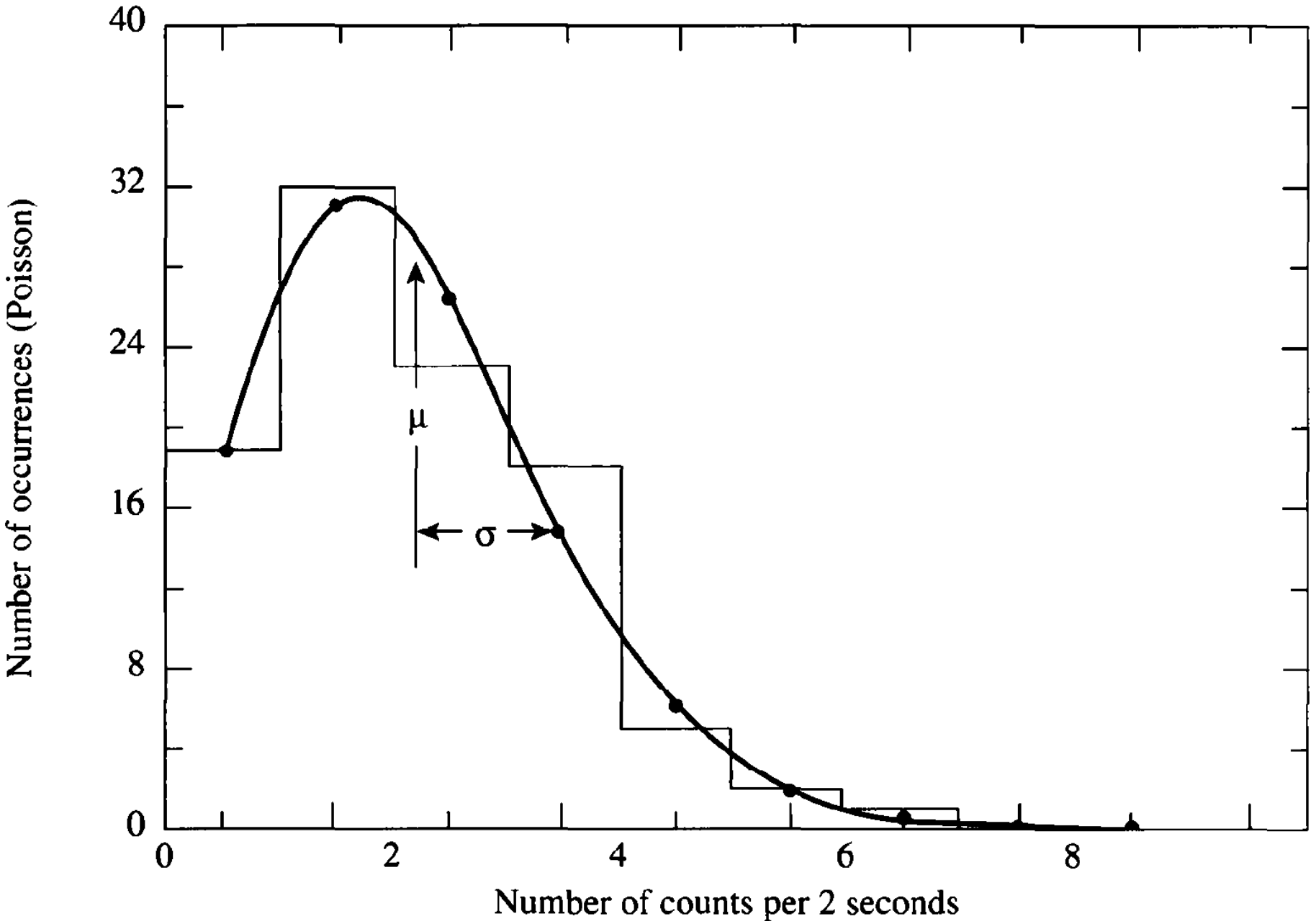


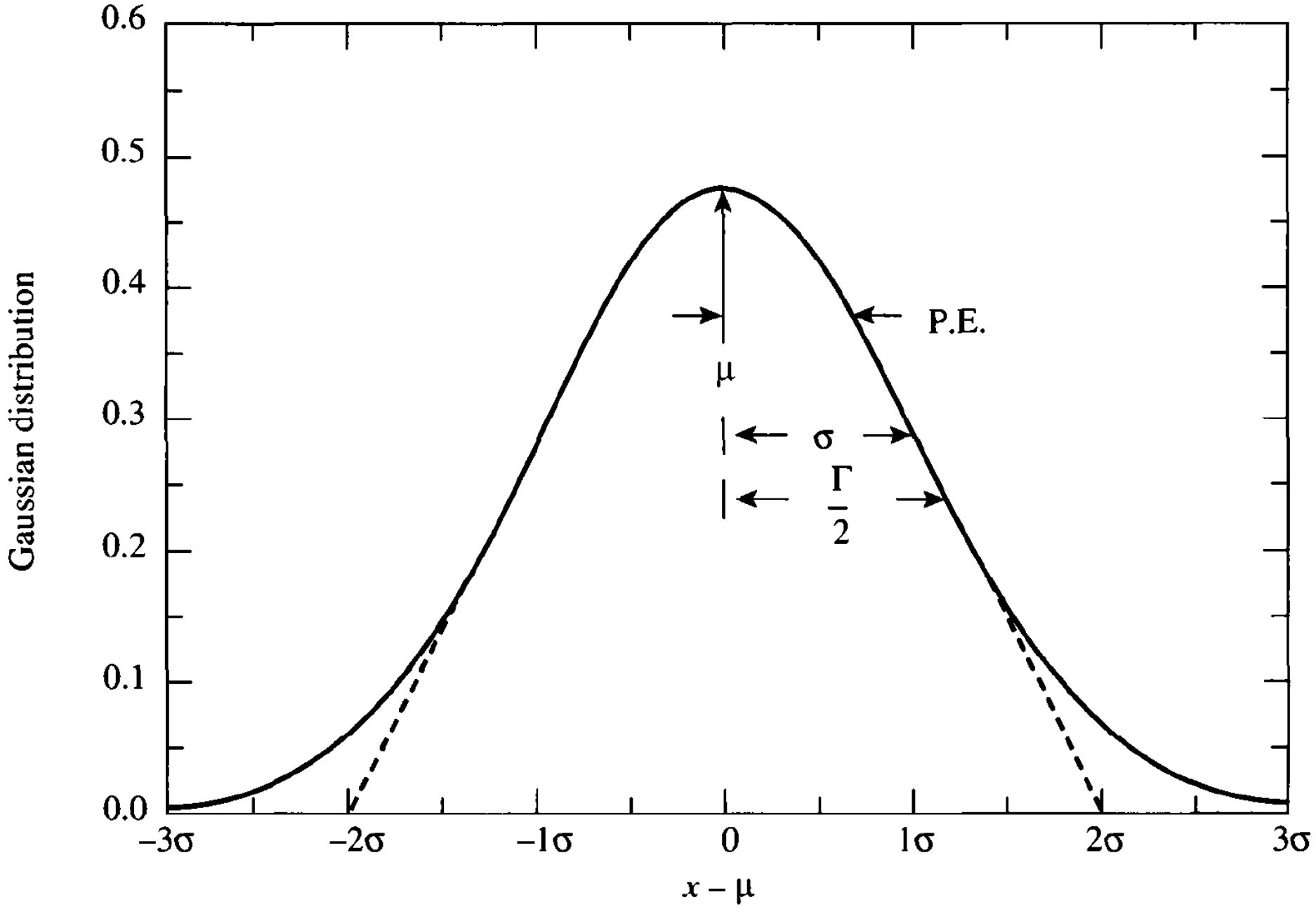
Atom Probe Data of InGaN Layer

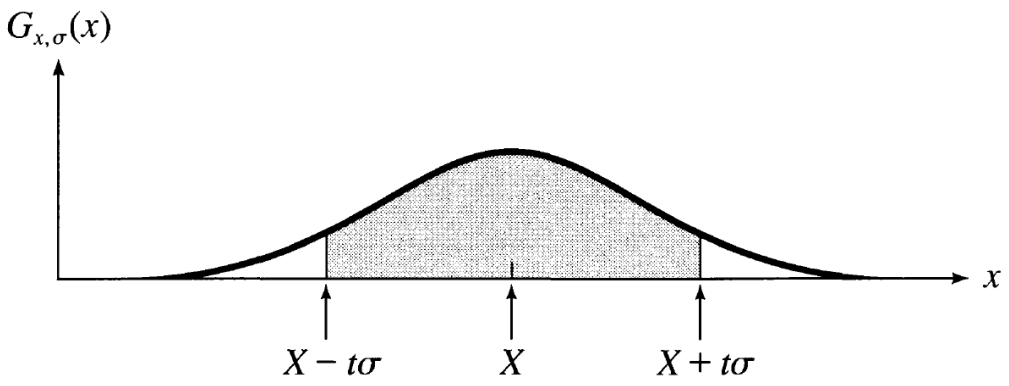



0.40
0.30
0.20
0.10
0.00
Binomial distribution
μ
σ
0
2
4
6
8
10
x

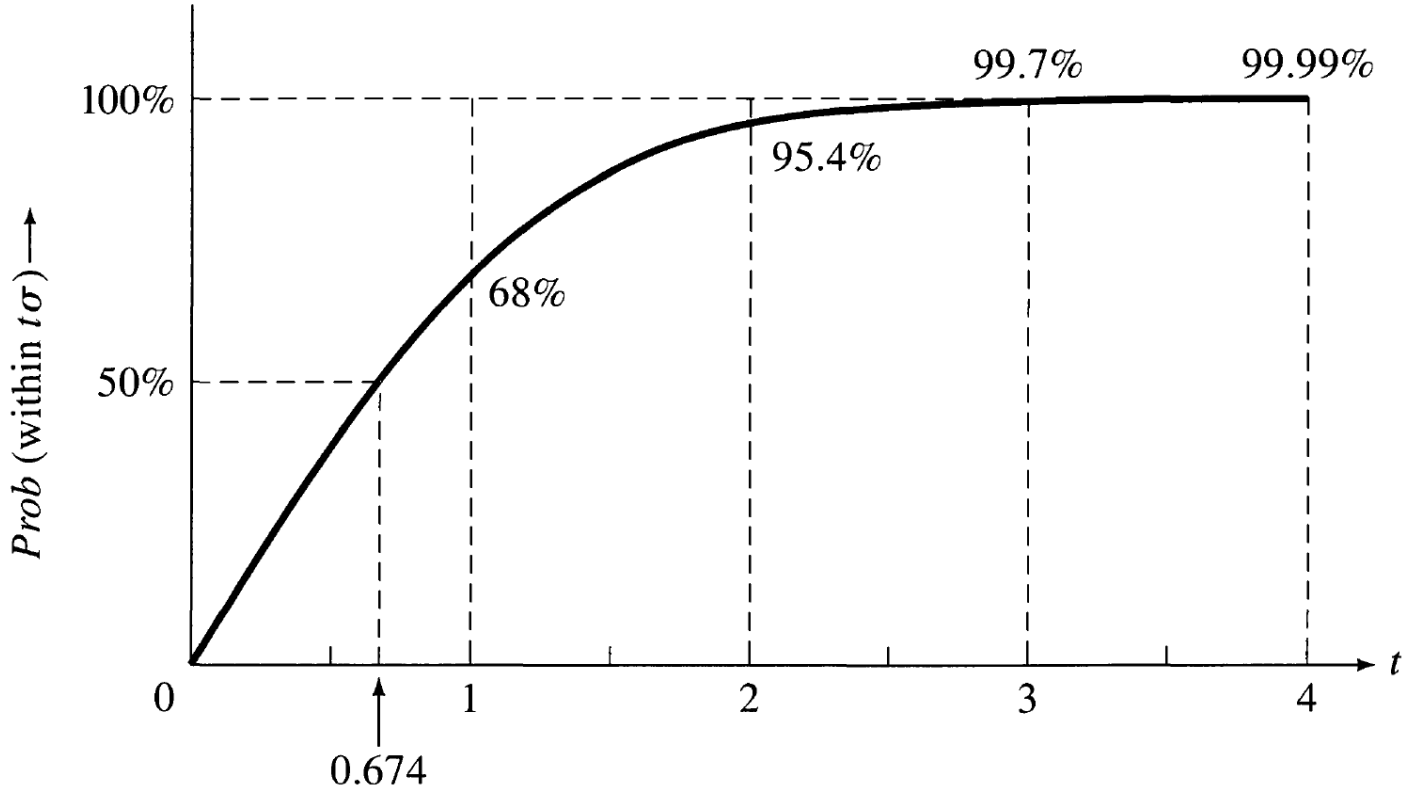

Relative expected frequency
1.0
0.8
0.6
0.4
0.2
0
μ = 0.1
μ = 1
μ = 2
μ = 3
μ = 10
0
2
4
6
8
10
12
14
16
18
Number of rare events per sample


40
32
24
16
8
0
0
2
4
6
8
Number of occurrences (Poisson)
Number of counts per 2 seconds
μ
σ

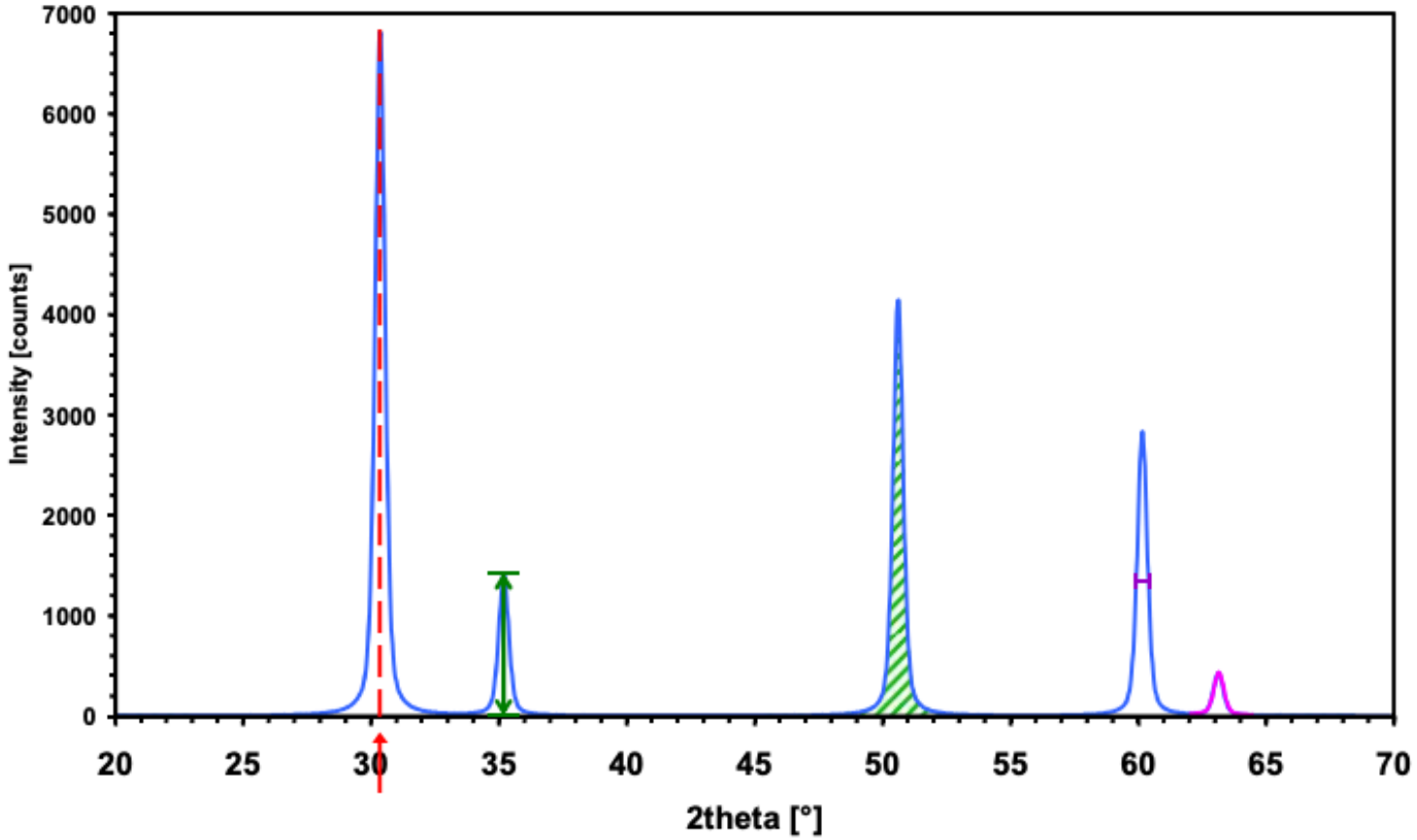

Gaussian distribution
0.6
0.5
0.4
0.3
0.2
0.1
0.0
P.E.
μ
σ
Γ/2
−3σ
−2σ
−1σ
0
1σ
2σ
3σ
x − μ

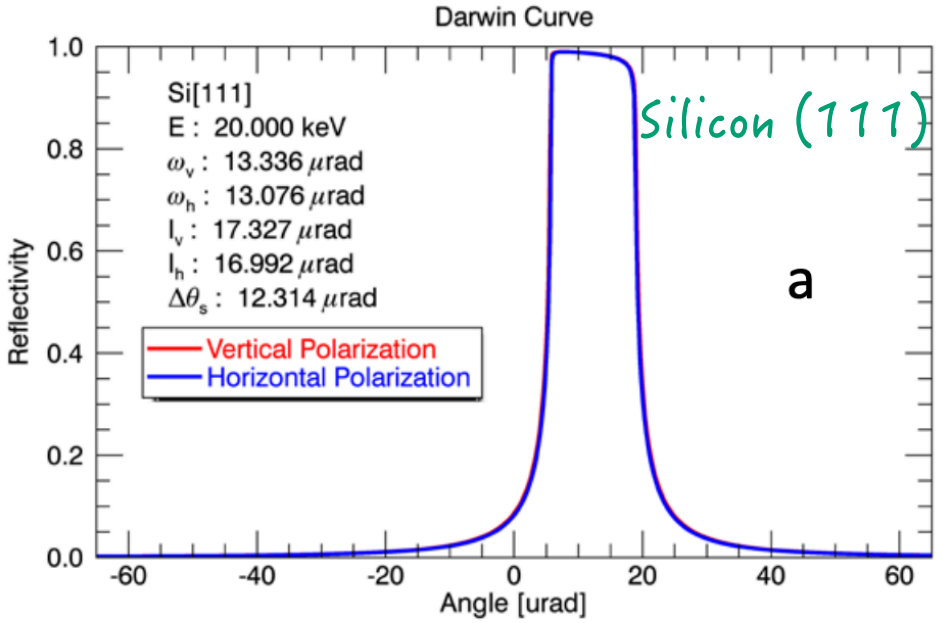


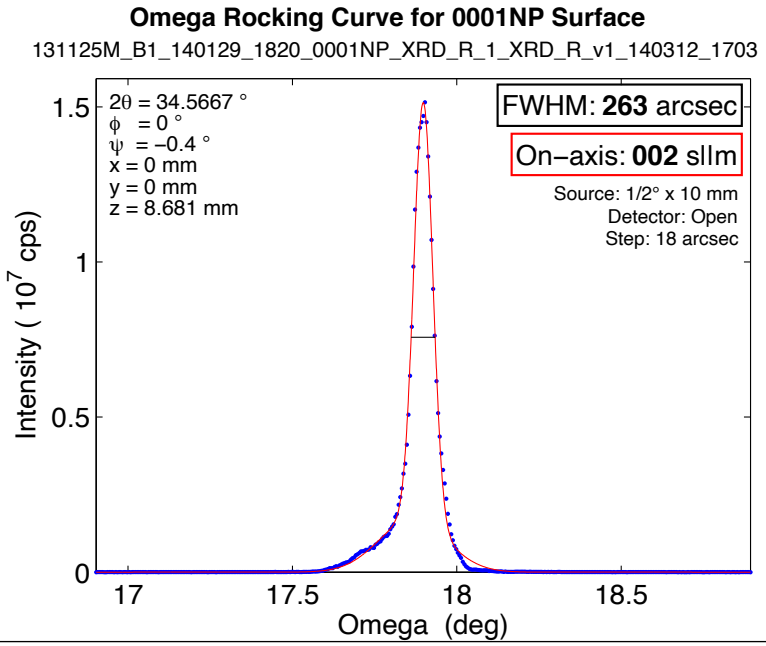
$$Prob(\text{within } t\sigma) = \frac{1}{\sqrt{2\pi}} \int_{-t}^{t} e^{-z^2/2}\, dz.$$

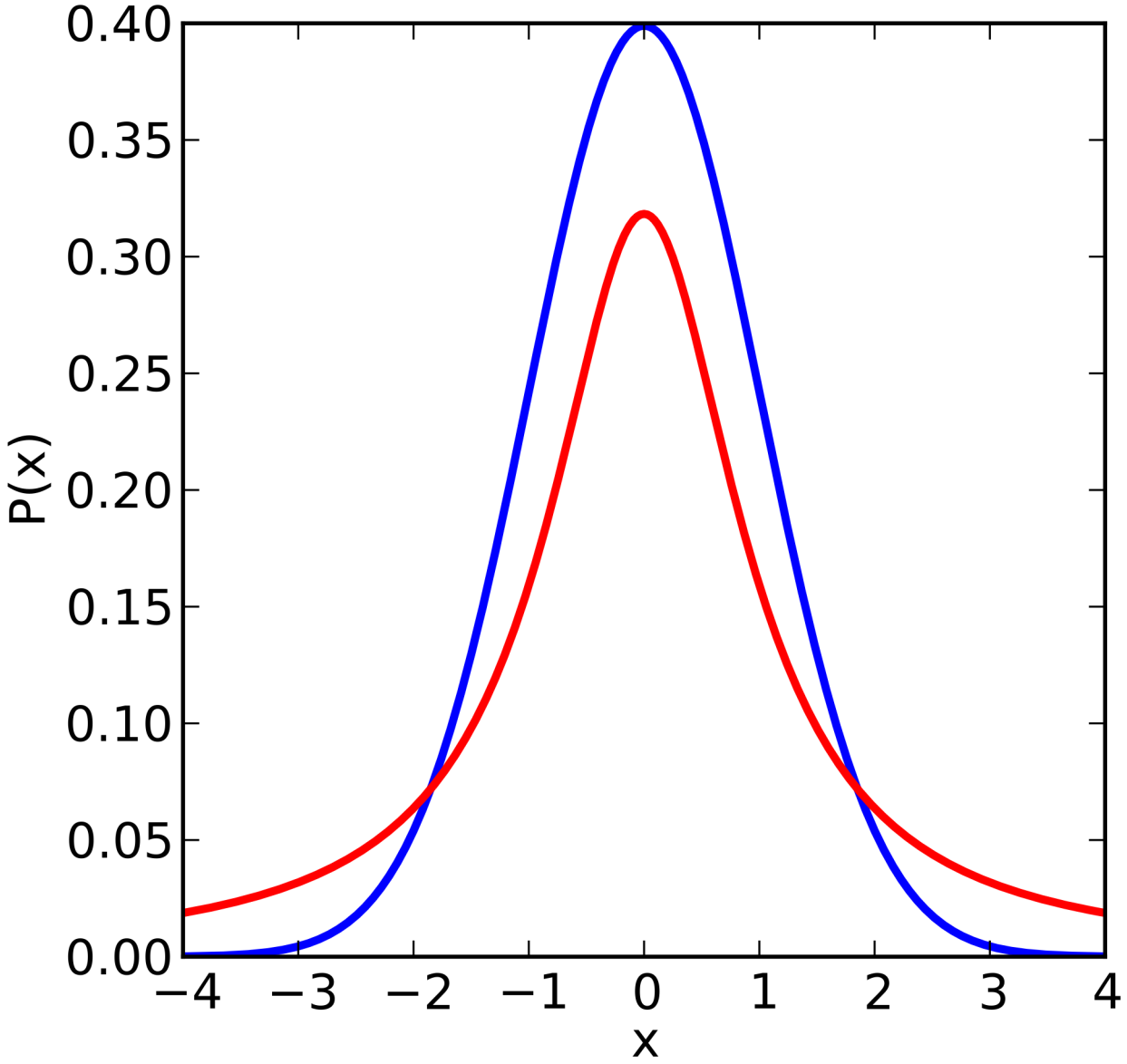


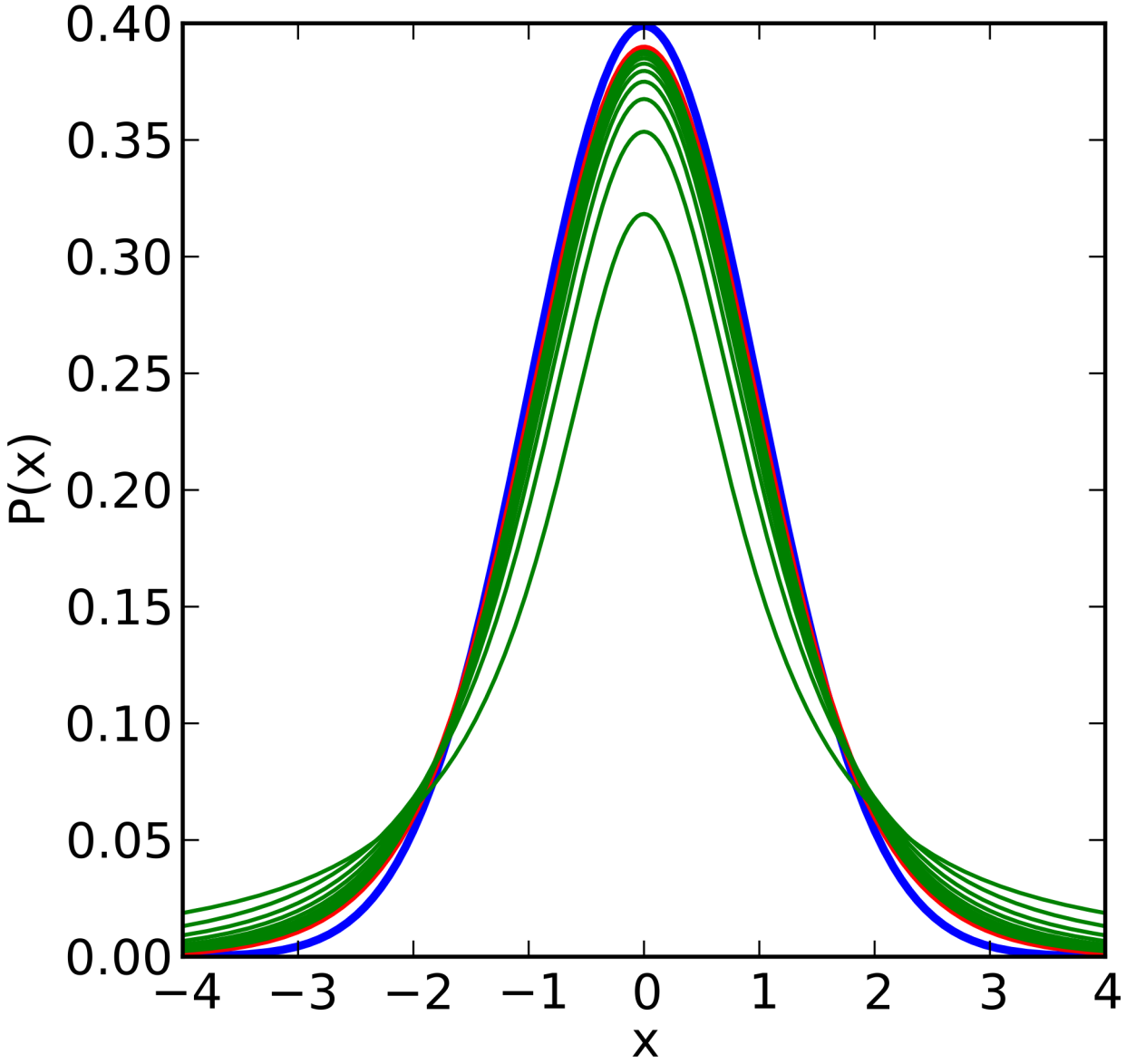
t	0	0.25	0.5	0.75	1.0	1.25	1.5	1.75	2.0	2.5	3.0	3.5	4.0
Prob (%)	0	20	38	55	68	79	87	92	95.4	98.8	99.7	99.95	99.99

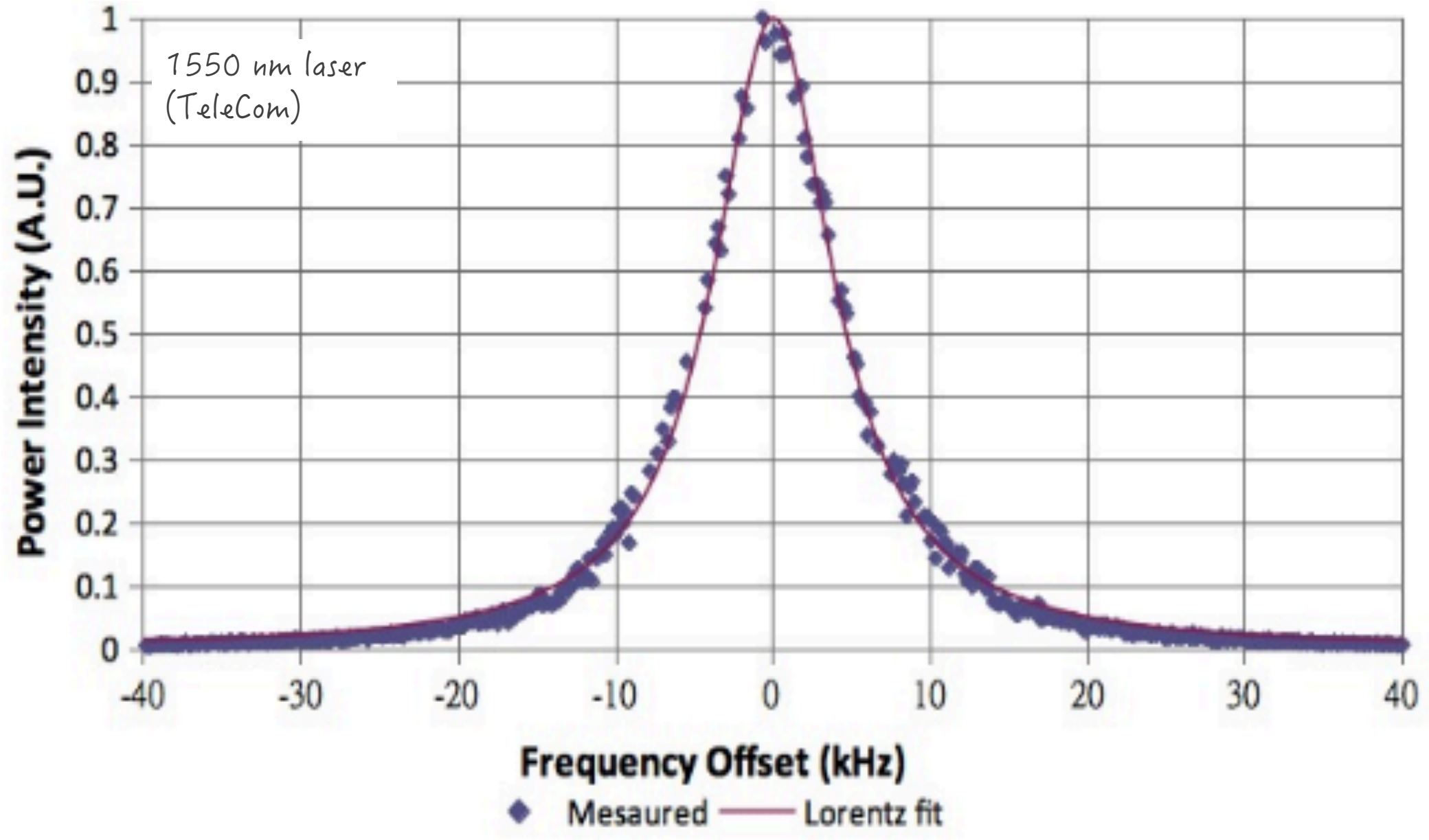

Intensity [counts]
2theta [°]

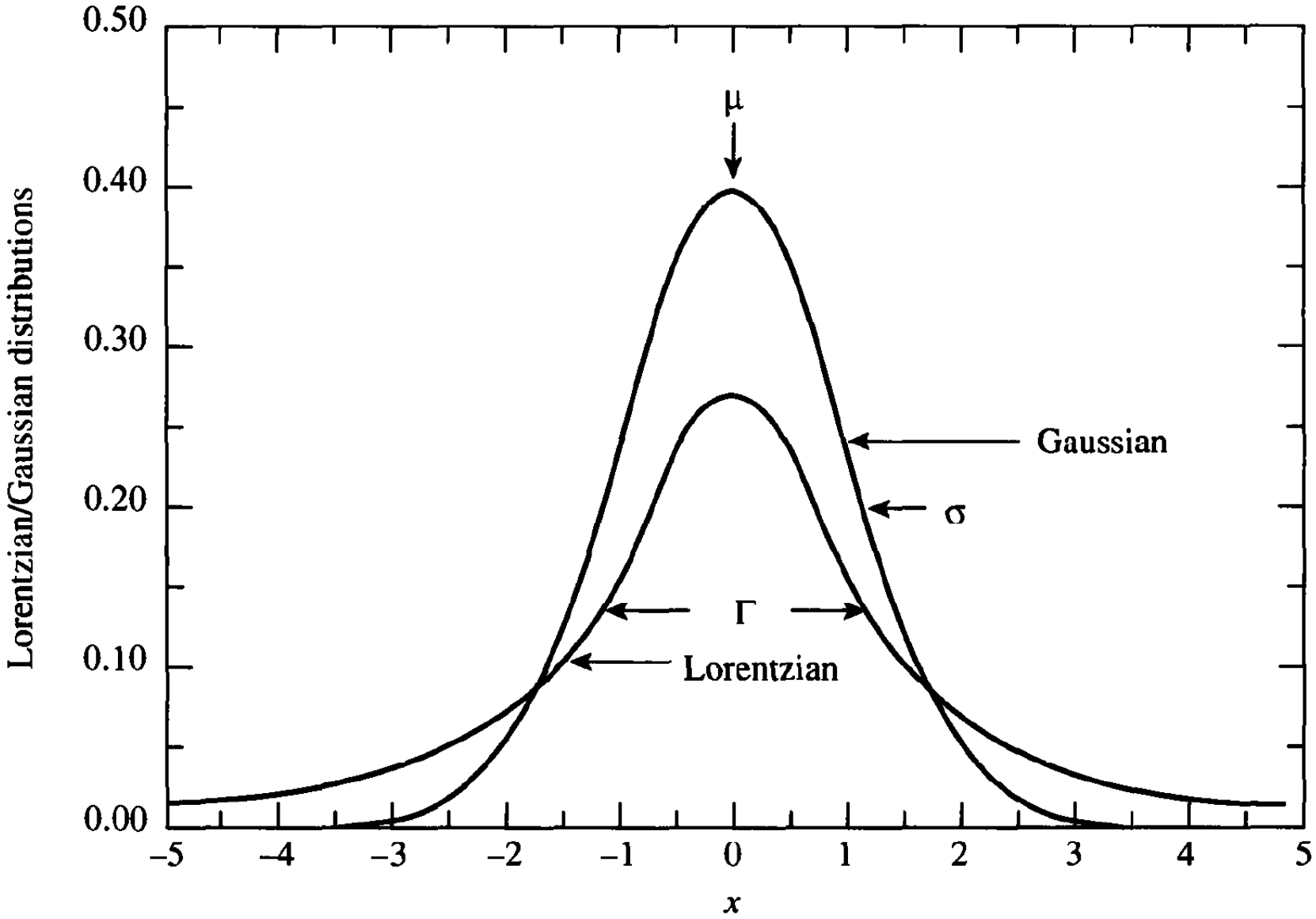

Darwin Curve
Si[111]
E : 20.000 keV
ω_v : 13.336 μrad
ω_h : 13.076 μrad
I_v : 17.327 μrad
I_h : 16.992 μrad
$\Delta\theta_s$: 12.314 μrad
Vertical Polarization
Horizontal Polarization
Reflectivity
Angle [urad]
Silicon (111)
a

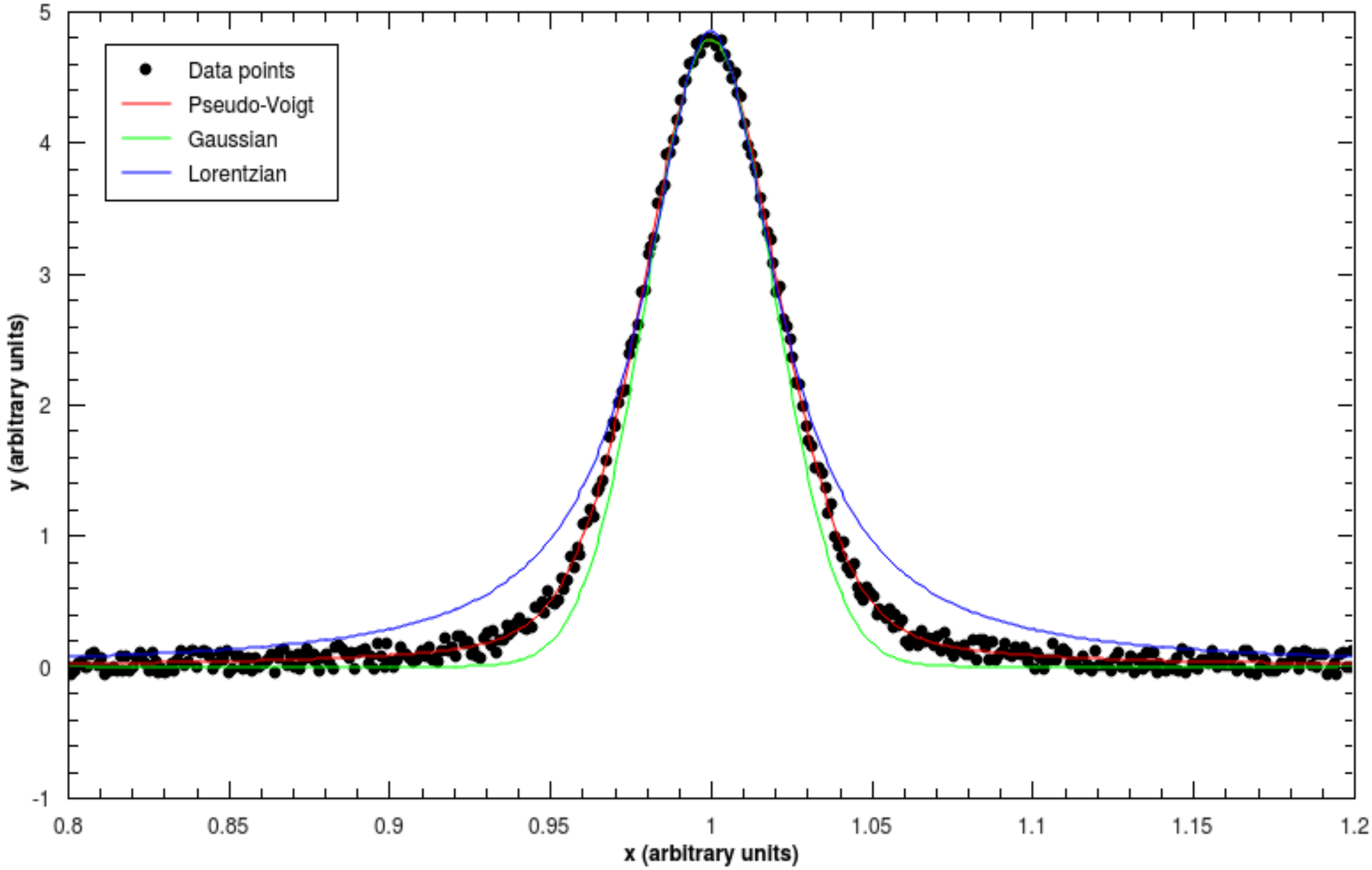

Omega Rocking Curve for 0001NP Surface
131125M_B1_140129_1820_0001NP_XRD_R_1_XRD_R_v1_140312_1703
2θ = 34.5667 °
ϕ = 0 °
ψ = –0.4 °
x = 0 mm
y = 0 mm
z = 8.681 mm
FWHM: 263 arcsec
On–axis: 002 sllm
Source: 1/2° x 10 mm
Detector: Open
Step: 18 arcsec
Intensity (10^7 cps)
Omega (deg)


0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
P(x)
−4 −3 −2 −1 0 1 2 3 4
x


0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
P(x)
−4 −3 −2 −1 0 1 2 3 4
x


1550 nm laser
(TeleCom)
Power Intensity (A.U.)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-40
-30
-20
-10
0
10
20
30
40
Frequency Offset (kHz)
Mesaured
Lorentz fit


Lorentzian/Gaussian distributions
0.50
0.40
0.30
0.20
0.10
0.00
μ
Gaussian
σ
Γ
Lorentzian
−5
−4
−3
−2
−1
0
1
2
3
4
5
x


Data points
Pseudo-Voigt
Gaussian
Lorentzian
y (arbitrary units)
x (arbitrary units)
5
4
3
2
1
0
-1
0.8
0.85
0.9
0.95
1
1.05
1.1
1.15
1.2

Example: You are measuring photon emissions from a sample at a specific wavelength. Within a 2-hour period you record 6 photons. What is the probability of measuring one additional photon in the next 15 min?

As a large team working at CERN you as the team leaders is trying to determine which three people must run the nightshift with you to monitor and revise an ongoing experiment. The names of all senior team members are thrown into a hat and the first three that are drawn will be joining you. The names are not replaced once they are drawn (i.e. one person cannot assume the responsibilities of two team members). Everyone either likes beer or wine. You want to know the probability of all three individuals being beer drinkers as you have an awesome homebrew bottle you want to share.

Does this problem follow the binomial distribution?

You are synthesizing nanoparticles. The diameter of the nanoparticles follows a normal distribution due to the synthesis processes you used. The mean diameter you measure is 5.85 nm and you observe a standard deviation of 0.24 nm.

a) Find the probability that a randomly selected particle will have a diameter of 6.0 nm. Plot the probability curve.

b) The middle 20% of nanoparticles will have diameters between what sizes?

c) Find the 90^{th} percentile for the diameters for the nanoparticles and interpret it in a sentence.