SIO223A, Lecture 5, 01/21/2020

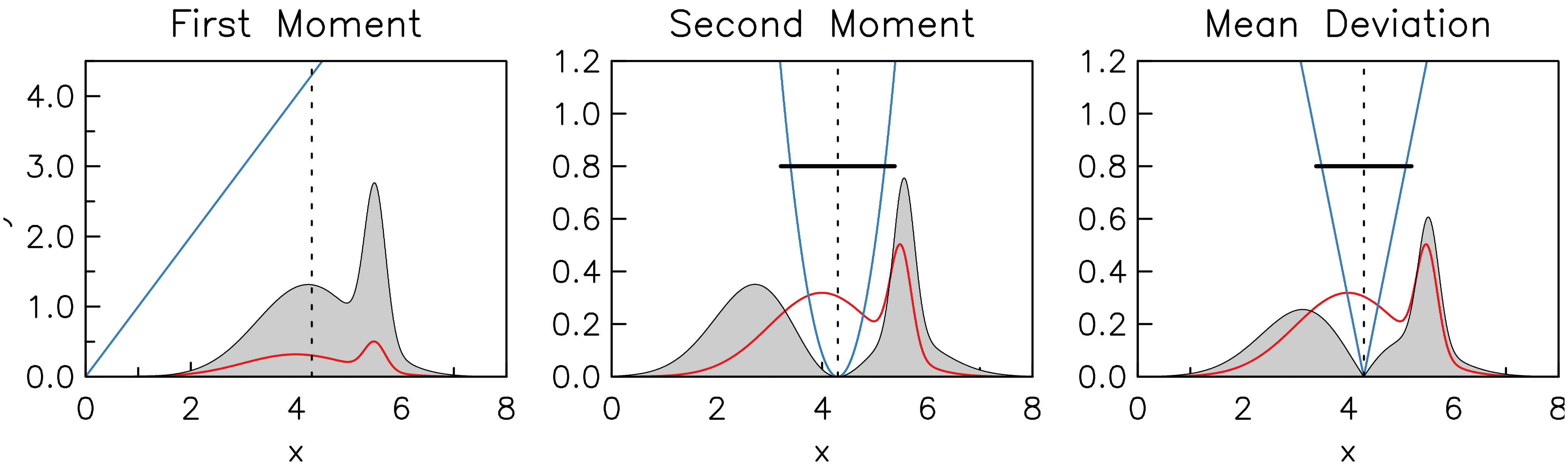
Transformations and Functions of Random Variables

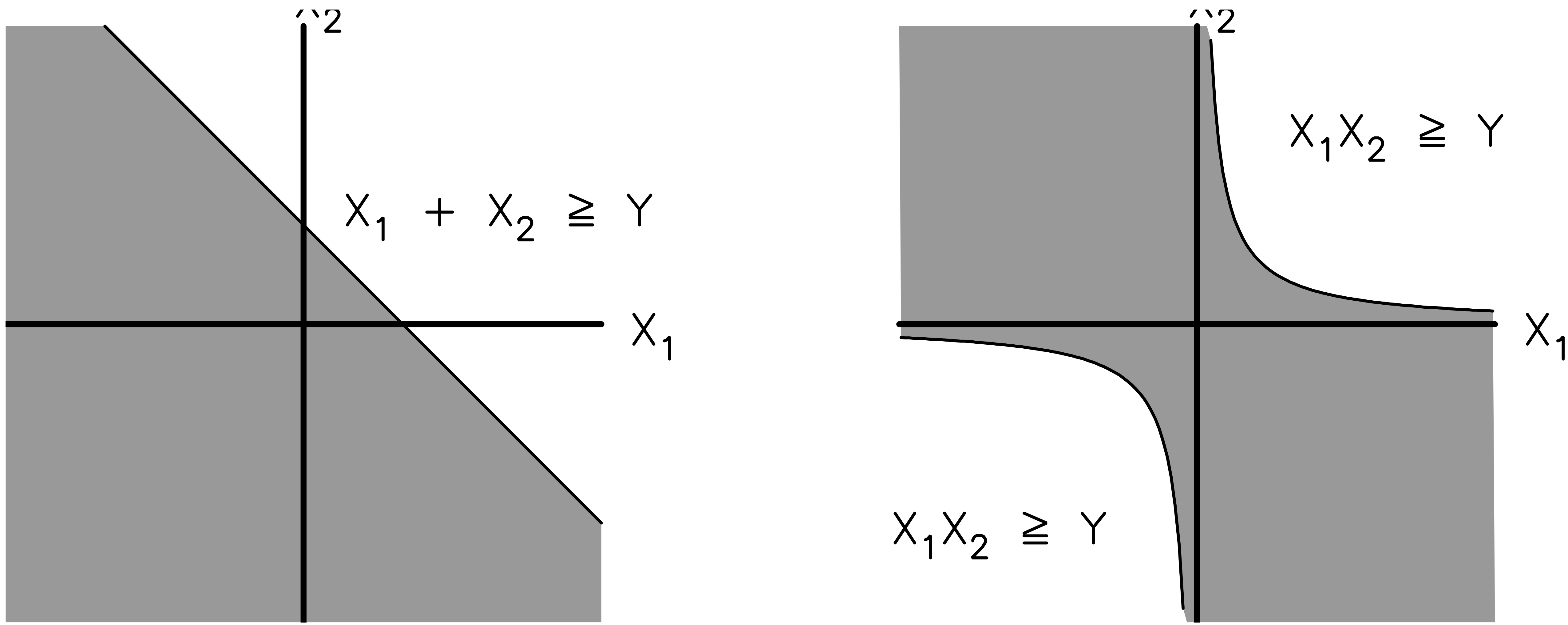
- Combinations of a random variable with conventional variables
- Joint probability distributions
- Sums and products of RVs
- The Central Limit Theorem
- Characteristic function - FT of pdf

http://igppweb.ucsd.edu/~cathy/Classes/SIO223A/

Terminology- Lecture 5

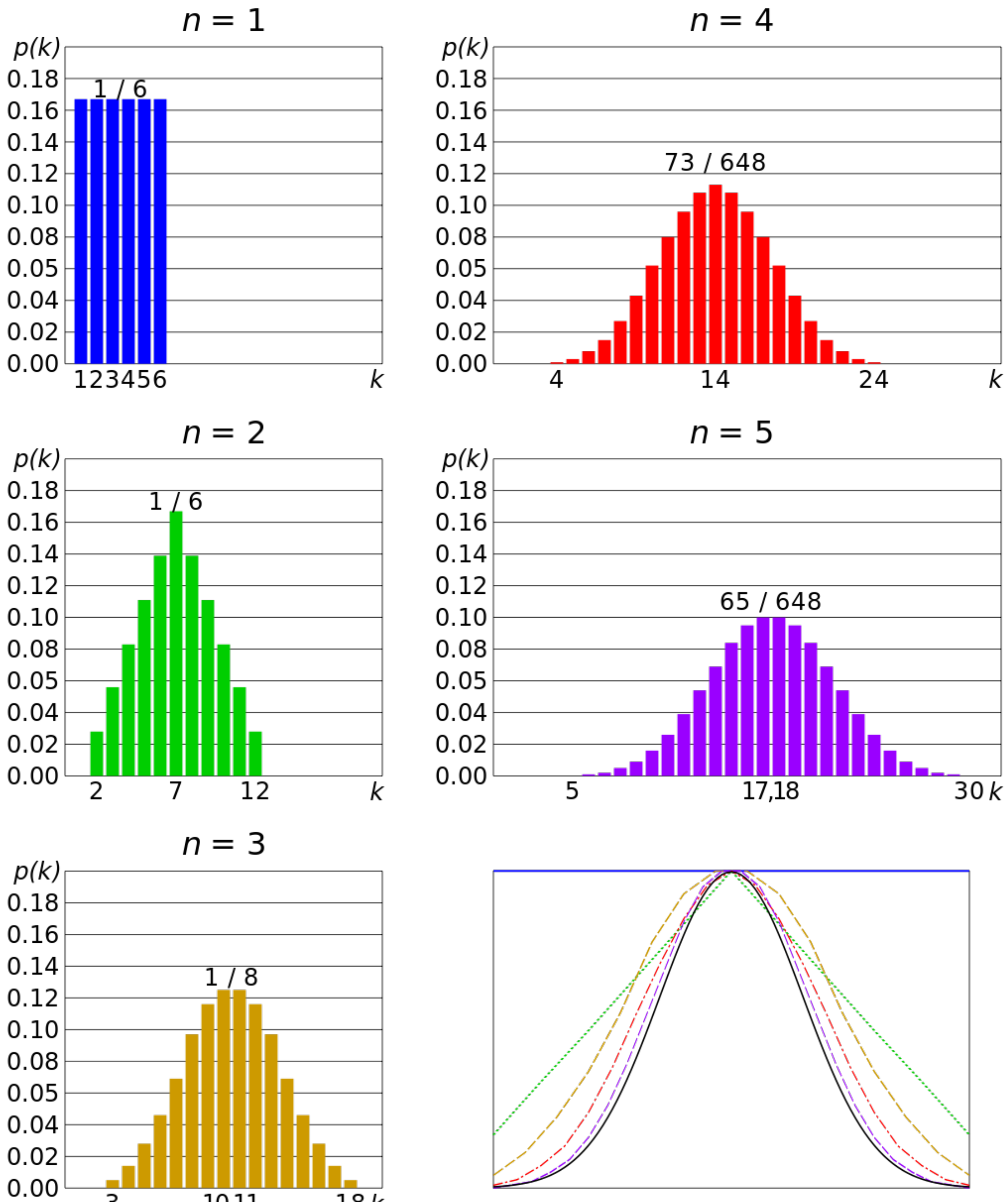
- moments
- skewness
- kurtosis
- expectation
- joint distribution
- Gaussian/Normal Distribution
- Fourier transform
- characteristic function
- Central Limit Theorem
- independent and identically distributed (iid)


First Moment
Second Moment
Mean Deviation
0.0
1.0
2.0
3.0
4.0
0.2
0.4
0.6
0.8
1.2
0
2
4
6
8
x



X1 + X2 ≧ Y
X1
X1X2 ≧ Y
X1
X1X2 ≧ Y

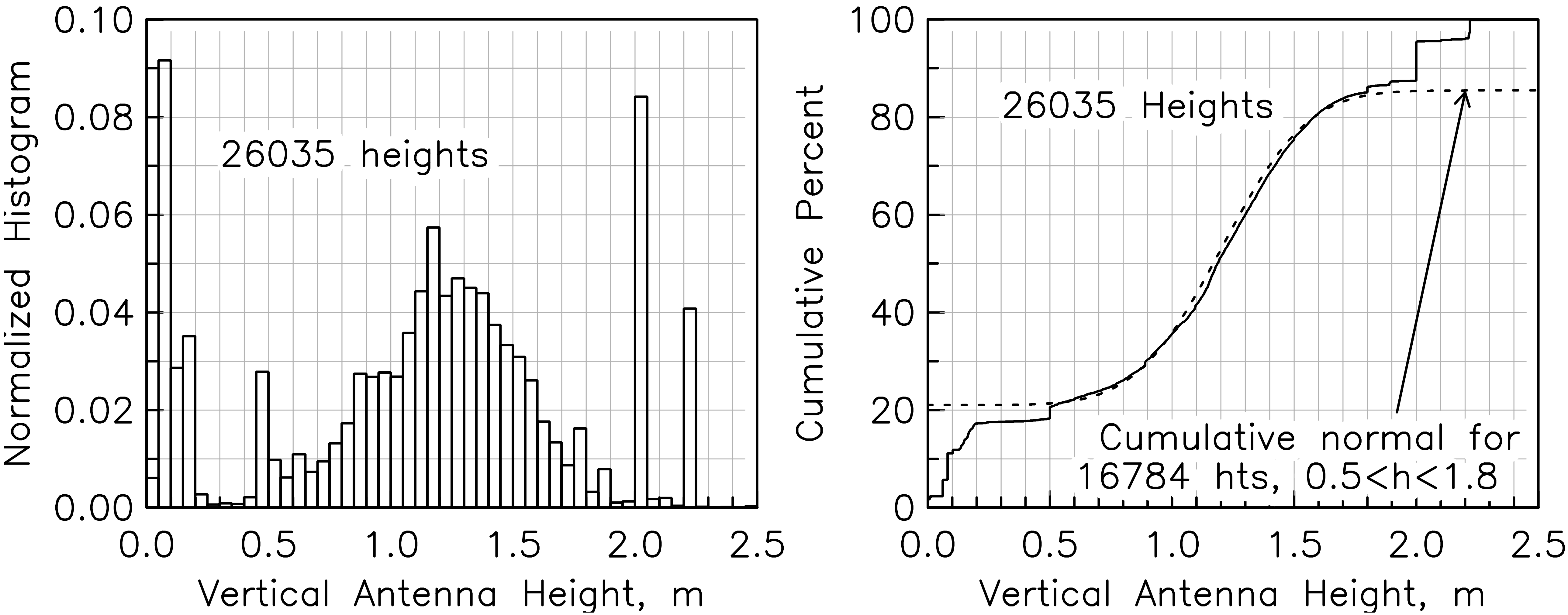
Central Limit Theorem - synthetic example



Comparison of probability density functions, p(k) for the sum of n fair 6-sided dice to show their convergence to a normal distribution with increasing n, in accordance to the central limit theorem. In the individual probability distribution functions, the minima, maxima and mods are labelled. In the bottom-right graph, smoothed profiles of the previous graphs are rescaled, superimposed and compared with a normal distribution, shown in black.

en.wikipedia.org/wiki/Central_limit_theorem#/media/File:Dice_sum_central_limit_theorem.svg

Setup Heights in SCEC GPS Data



SIO223A, Lecture 6, 01/23/2020

Univariate Distributions

- Uniform
- Normal - and generating normal deviates
- Point and renewal processes - Poisson, exponential, gamma and Weibull distributions
- Relatives of the Normal distribution - Cauchy, Chi-squared, Student's t, F, Rayleigh, Log-Normal
- von Mises, Fisher, Pareto
- General methods for deviates

http://igppweb.ucsd.edu/~cathy/Classes/SIO223A/

Terminology- Chapter 3

- location parameter
- scale parameter
- uniform/rectangular distribution
- pseudorandom numbers
- quasi-random sequence
- normal/Gaussian distribution
- deviates
- Box-Mueller transform
- Poisson process
- renewal process
- exponential distribution
- Poisson distribution
- shape parameter
- gamma distribution
- Weibull distribution
- Cauchy distribution
- Chi-squared, t, and F
- Rayleigh distribution
- Log normal, von Mises, Fisher, Distribution
- Pareto distribution

Standardizing pdf's

$$\frac{1}{cA_r(s)}\phi_s\left(\frac{x-l}{c}\right) \qquad \text{for} \quad L_b \le x \le L_c \tag{3.1}$$

In this expression the L's give the range of the variable: often from $-\infty$ to ∞, or from 0 to ∞, but sometimes over a finite range. The function ϕ_s gives the actual shape of the pdf; the constant $A_r(s)$ is the area under the function, included to normalize the integral of ϕ to unity. We call s the **shape parameter**; not all pdf's have one. But almost all pdf's do have two others:

1. A **location parameter** l, which sets the location of the pdf on the x-axis. This parameter appears mostly, though not always, for pdf's on $(-\infty, \infty)$.

2. A **scale parameter** c, which expands or contracts the scale of the x-axis. For the pdf to remain properly normalized, c also has to scale the size of the pdf, and so multiplies $A_r(s)$.

Generating Random Samples

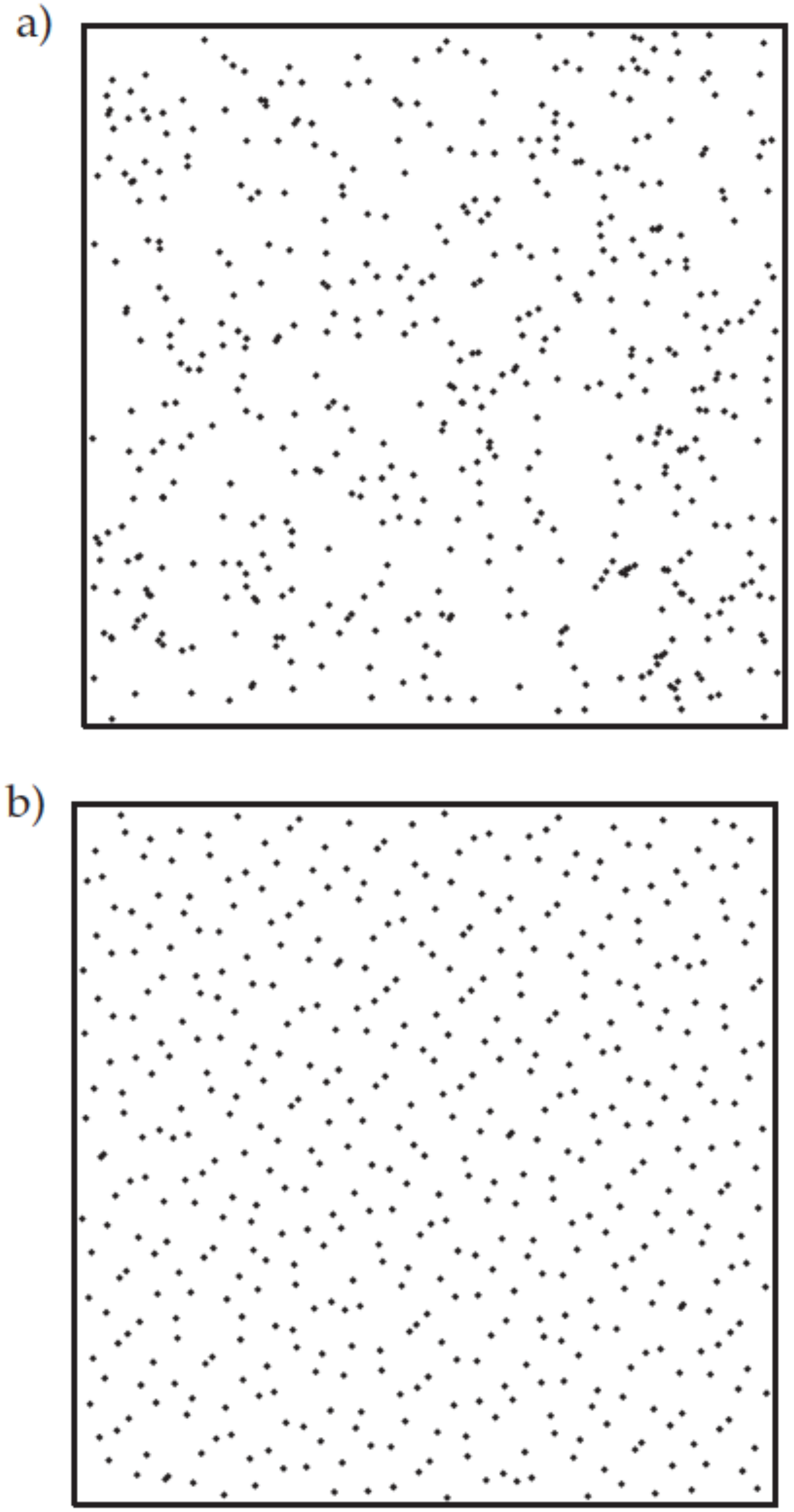


Figure 5. (a) The 1OOO uniform points generated with a pseudorandom number generator, showing typical clustering and uneveness. (b) The 1OOO uniform points generated with two sequences of quasi-random numbers. Note the considerably more uniform appearance and even density compared to Figure 5a.

a) Pseudo random number generator

John von Neumann cautioned about the misinterpretation of a PRNG as a truly random generator, and joked that "Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin." https://en.wikipedia.org/wiki/Pseudorandom_number_generator

b) Quasi random sequences

Sambridge & Mosegaard, 2002

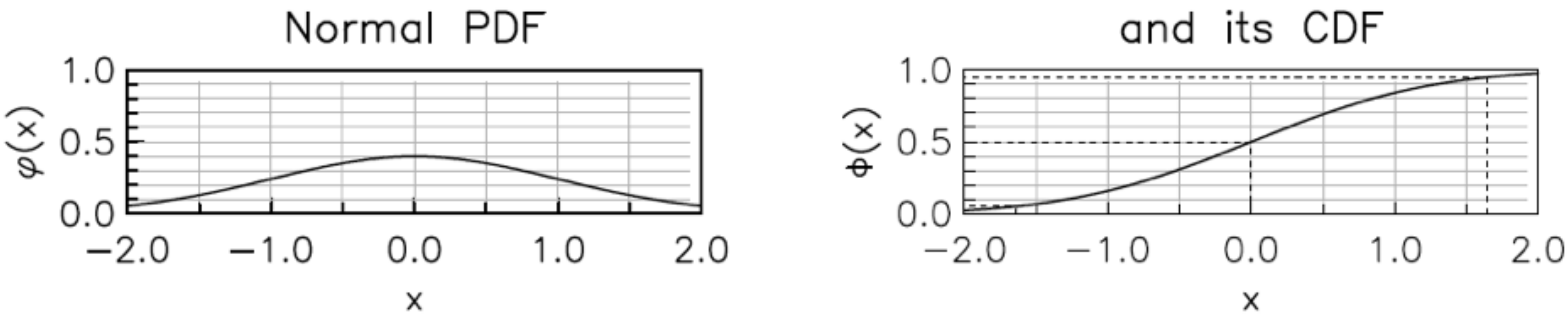


Figure 3.1: A Normal, or Gaussian, pdf and its cdf, plotted for zero mean and unit variance. In this figure, unlike most of the others, the y axis is not exaggerated relative to the x axis.

x	± 1.00	± 1.65	± 1.96	± 2.58	± 3.29	± 3.90
Mass fraction	0.68	0.90	0.95	0.99	0.999	0.9999

3.3 The Normal (Gaussian) Distribution

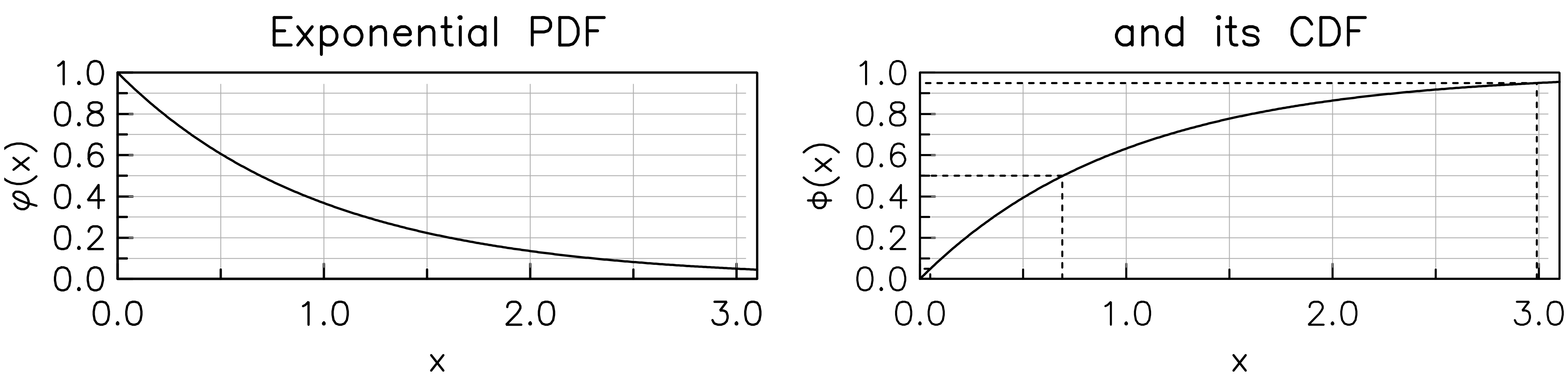
We have already met this pdf,[1] but present it again to illustrate our different ways of writing a pdf:

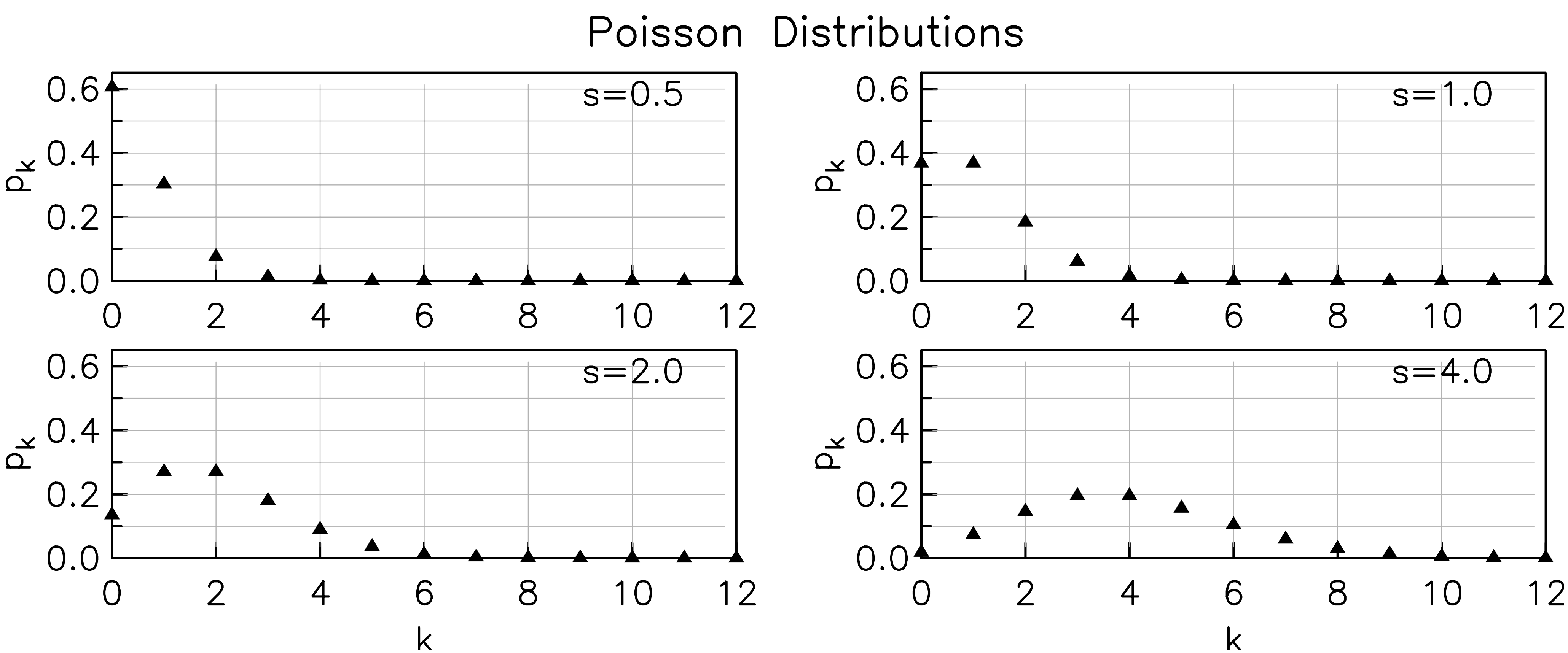
$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad = \frac{1}{c\sqrt{2\pi}} e^{-(x-l)^2/2c^2} \qquad = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \qquad (3.4)$$

Generating Normal Deviates

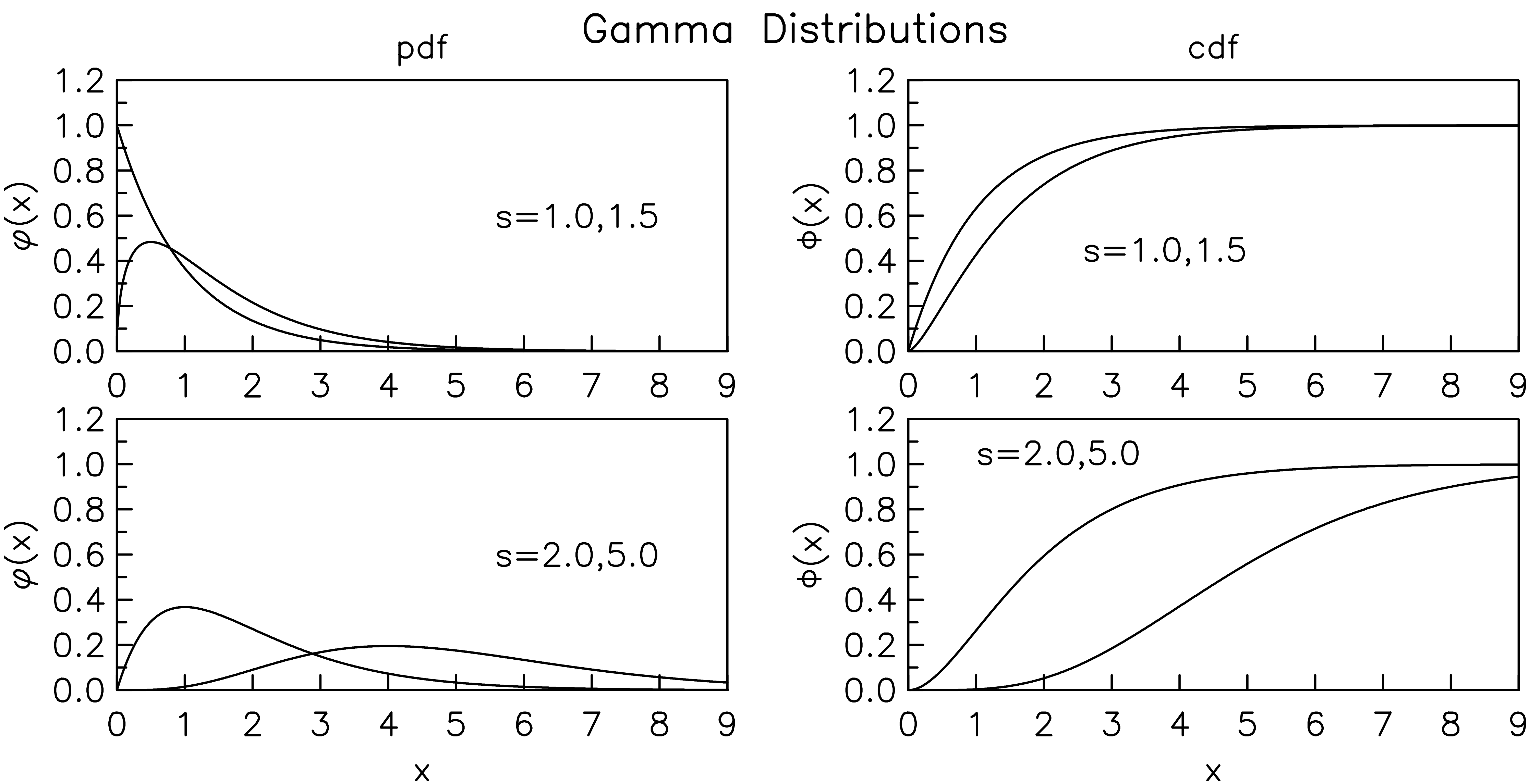
- Box Mueller transformation
- Rejection from square to circle

Point processes- Exponential and Poisson Distributions

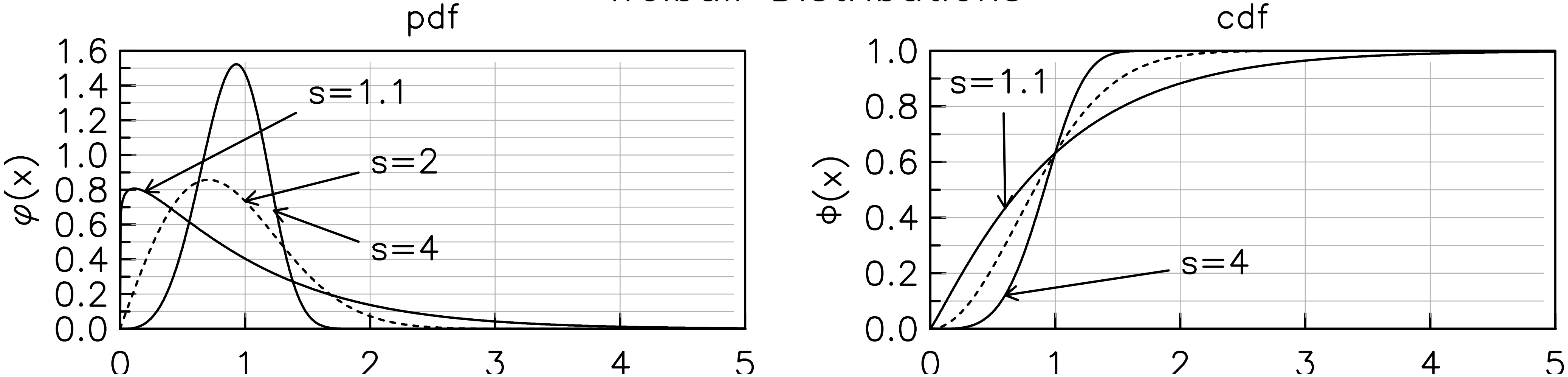




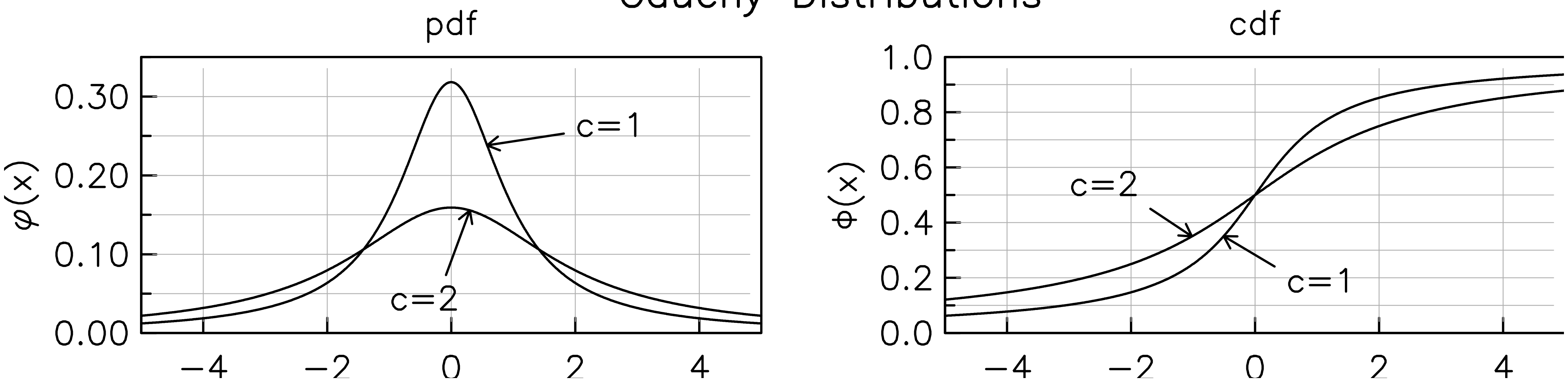
Gamma Distribution



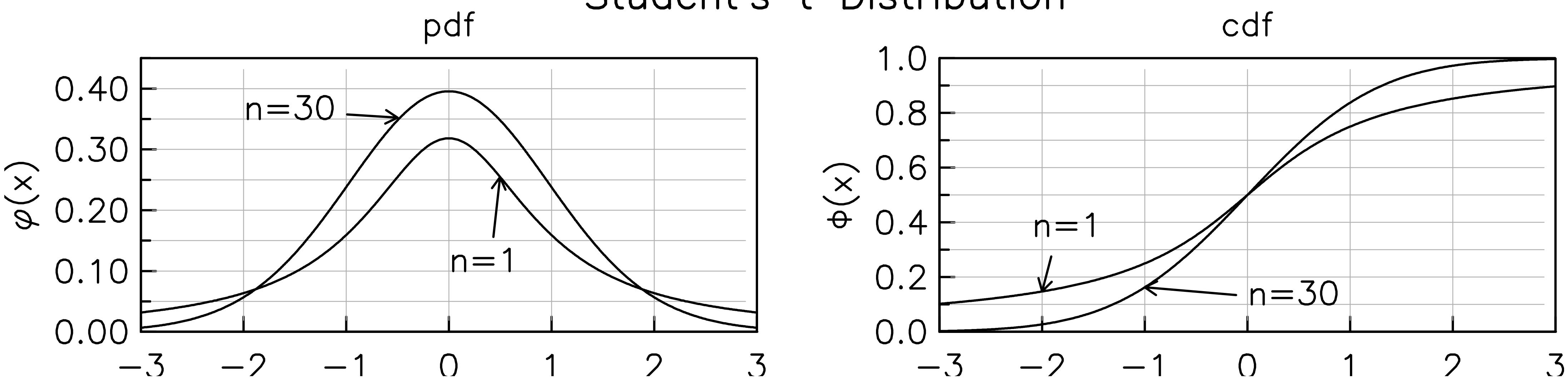
Weibull Distributions



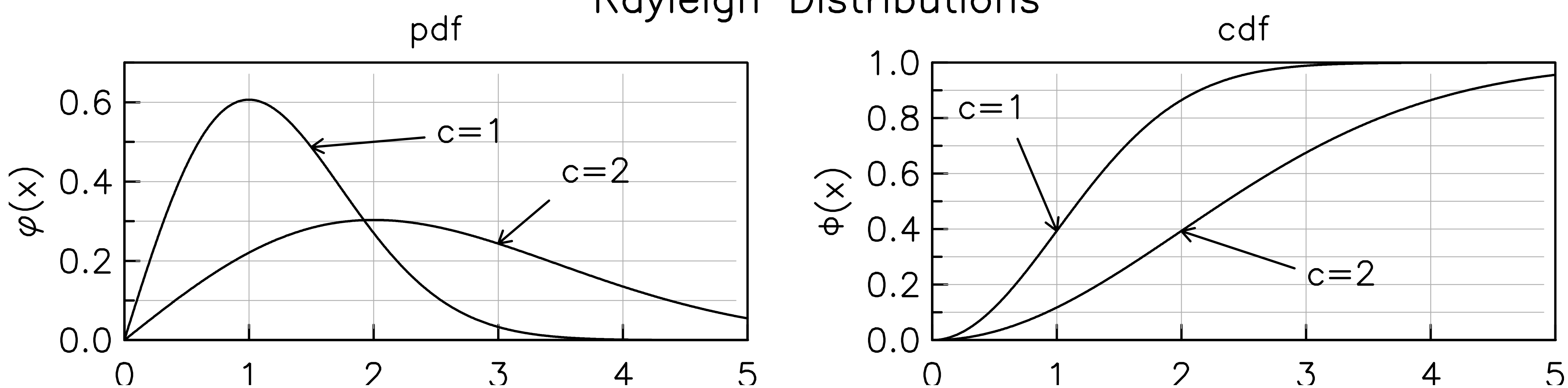
Cauchy Distributions



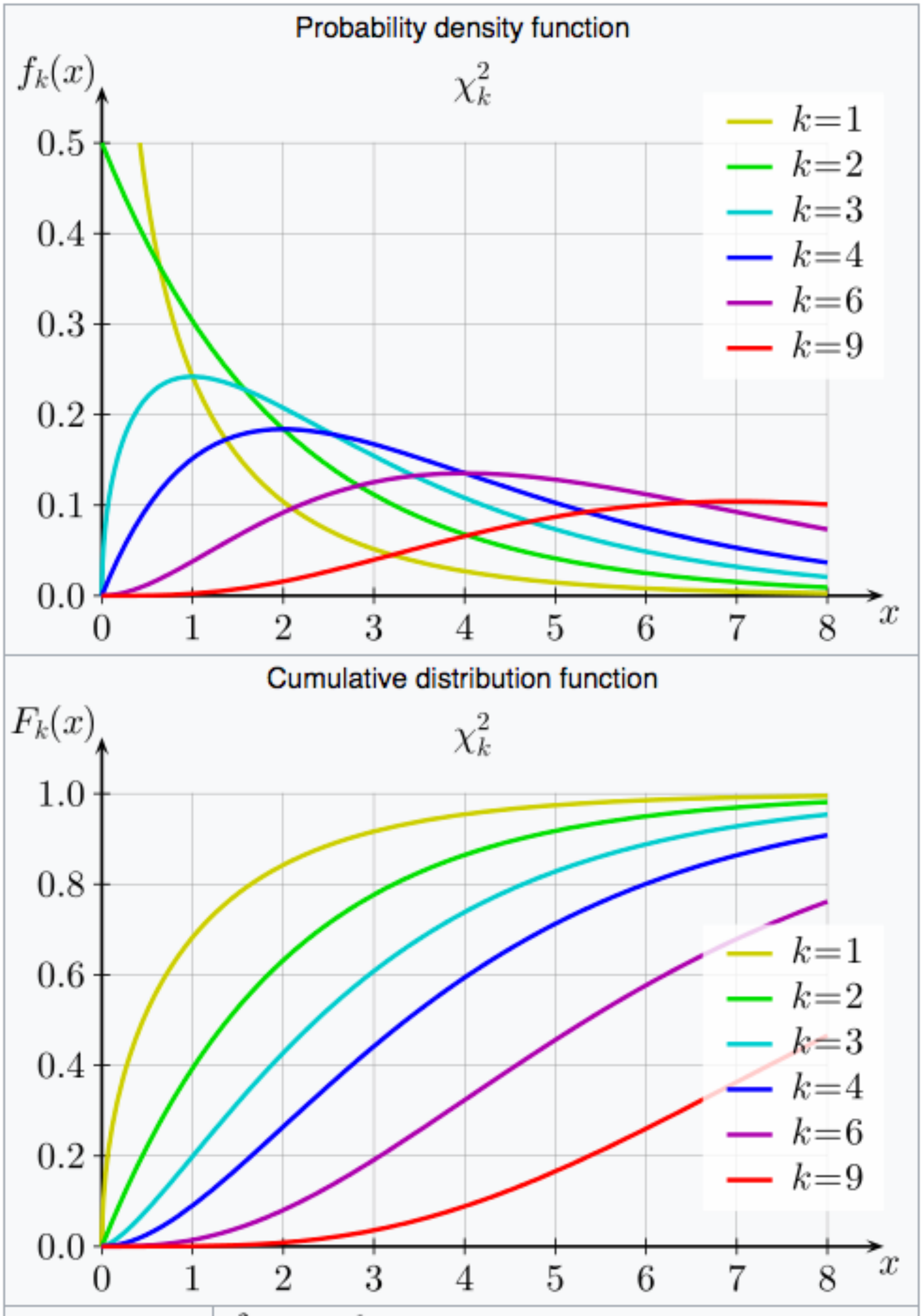
Student's t Distribution



Rayleigh Distributions



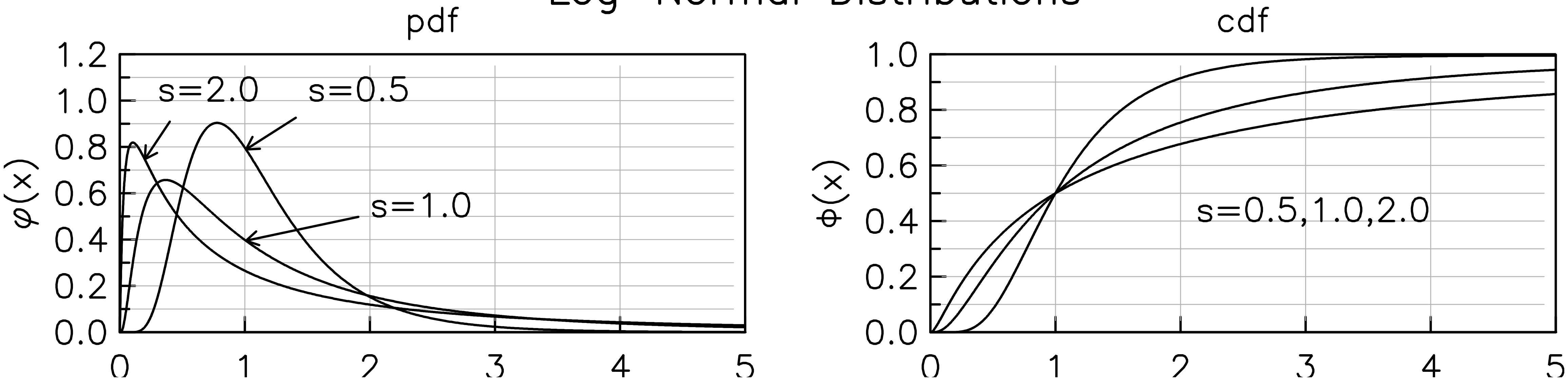
chi-squared



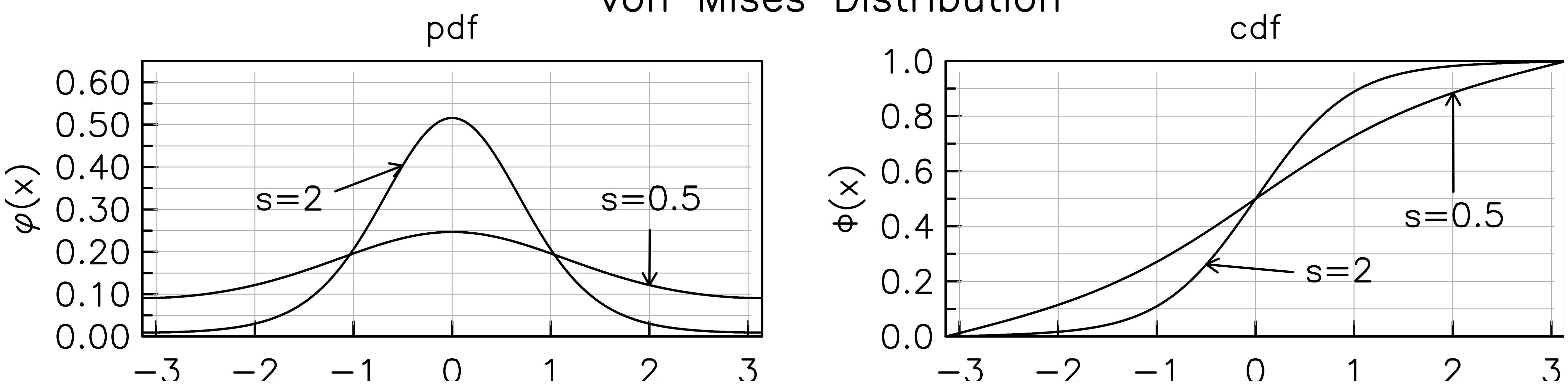
Properties

Notation	$\chi^2(k)$ or χ^2_k
Parameters	$k \in \mathbb{N}_{>0}$ (known as "degrees of freedom")
Support	$x \in [0, +\infty)$
PDF	$\frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}\, x^{\frac{k}{2}-1} e^{-\frac{x}{2}}$
CDF	$\frac{1}{\Gamma\left(\frac{k}{2}\right)}\,\gamma\left(\frac{k}{2}, \frac{x}{2}\right)$
Mean	k
Median	$\approx k\left(1 - \frac{2}{9k}\right)^3$
Mode	$\max\{k-2, 0\}$
Variance	$2k$
Skewness	$\sqrt{8/k}$
Ex. kurtosis	$12/k$
Entropy	$\frac{k}{2} + \ln(2\Gamma(\frac{k}{2})) + (1-\frac{k}{2})\psi(\frac{k}{2})$ (nats)
MGF	$(1-2t)^{-k/2}$ for $t < ½$
CF	$(1-2it)^{-k/2}$ [1]

Log–Normal Distributions



von Mises Distribution



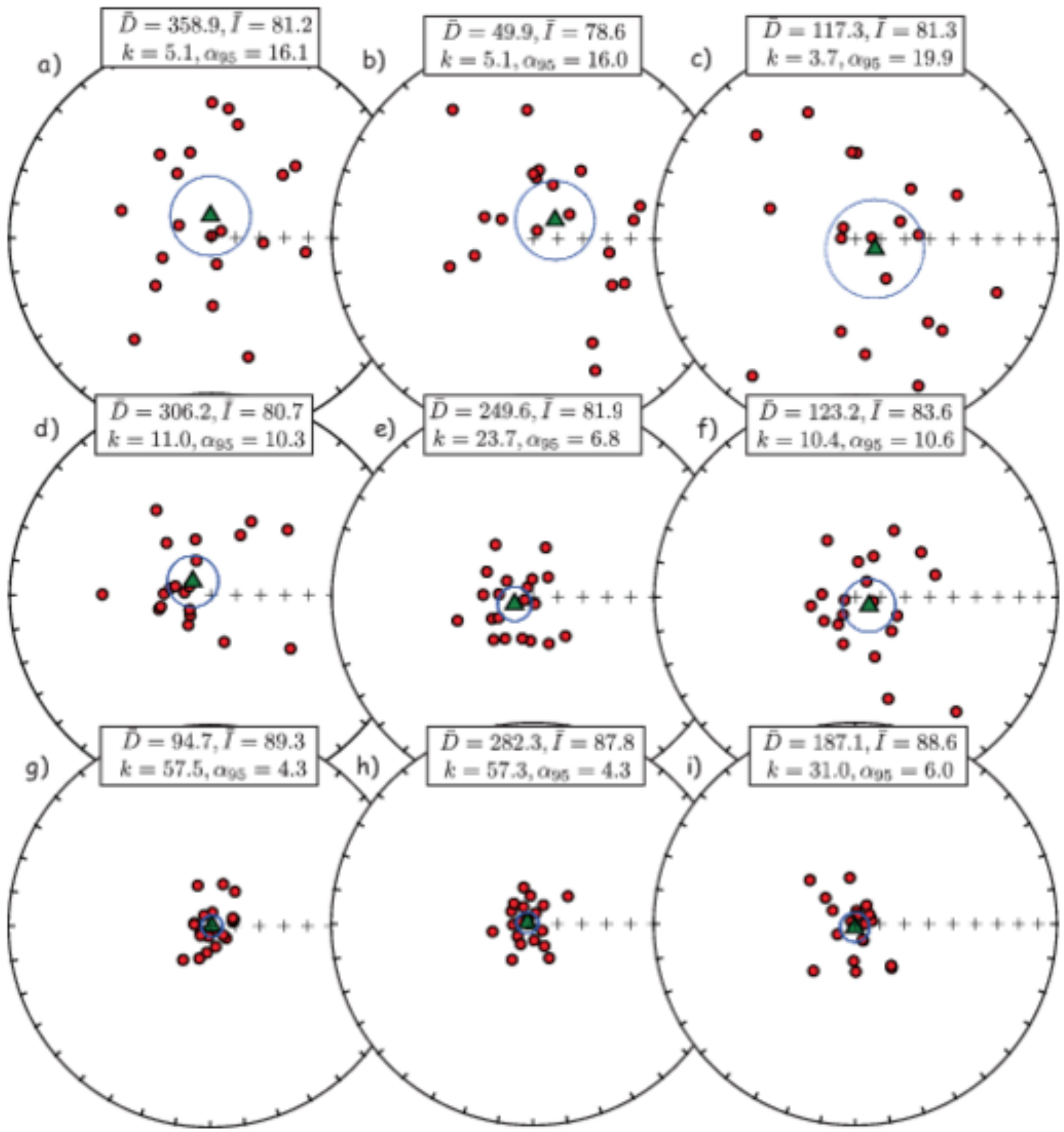


Figure 11.2: Hypothetical data sets drawn from Fisher distributions

$$\phi(\Delta,\theta) = \frac{s}{4\pi \sinh s} e^{s\cos\Delta} = \frac{\kappa}{4\pi \sinh\kappa} e^{\kappa\cos\Delta}$$

$$\phi(\Delta) = \frac{\kappa}{2\sinh\kappa} e^{\kappa\cos\Delta} \sin\Delta$$

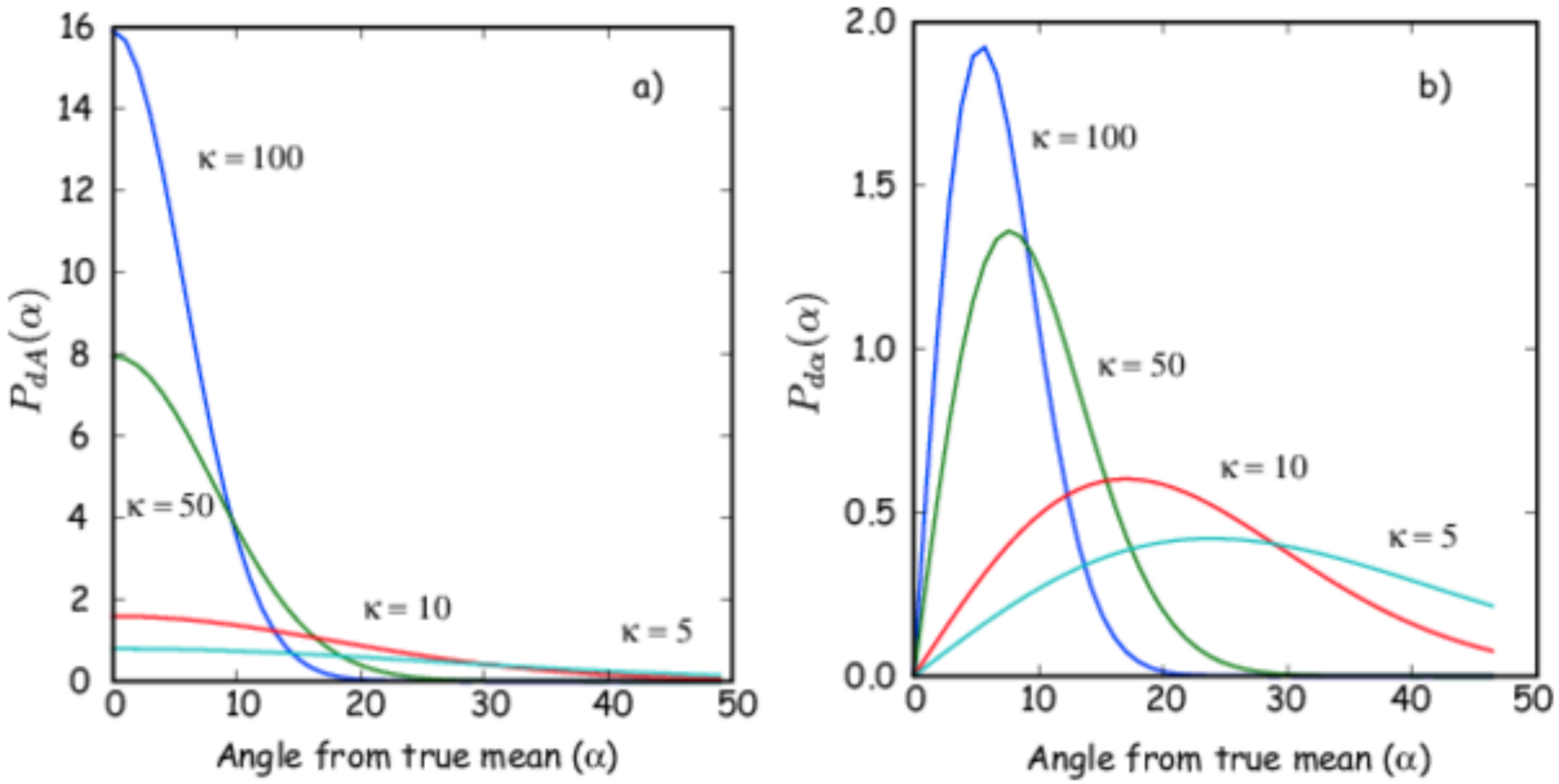


Figure 11.3: a) Probability of finding a direction within an angular area, dA centered at an angle α from the true mean. b) Probability of finding a direction at angle α away from the true mean direction.

$$\phi(x) = s l^s x^{-(s+1)}$$

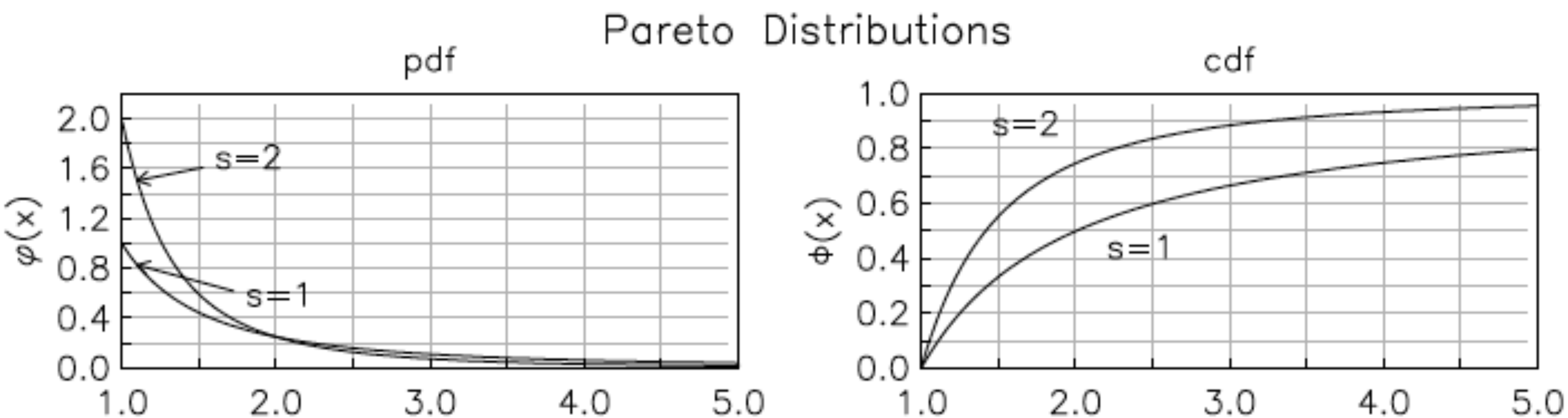


Figure 3.11: Pareto distributions for different shape factors.

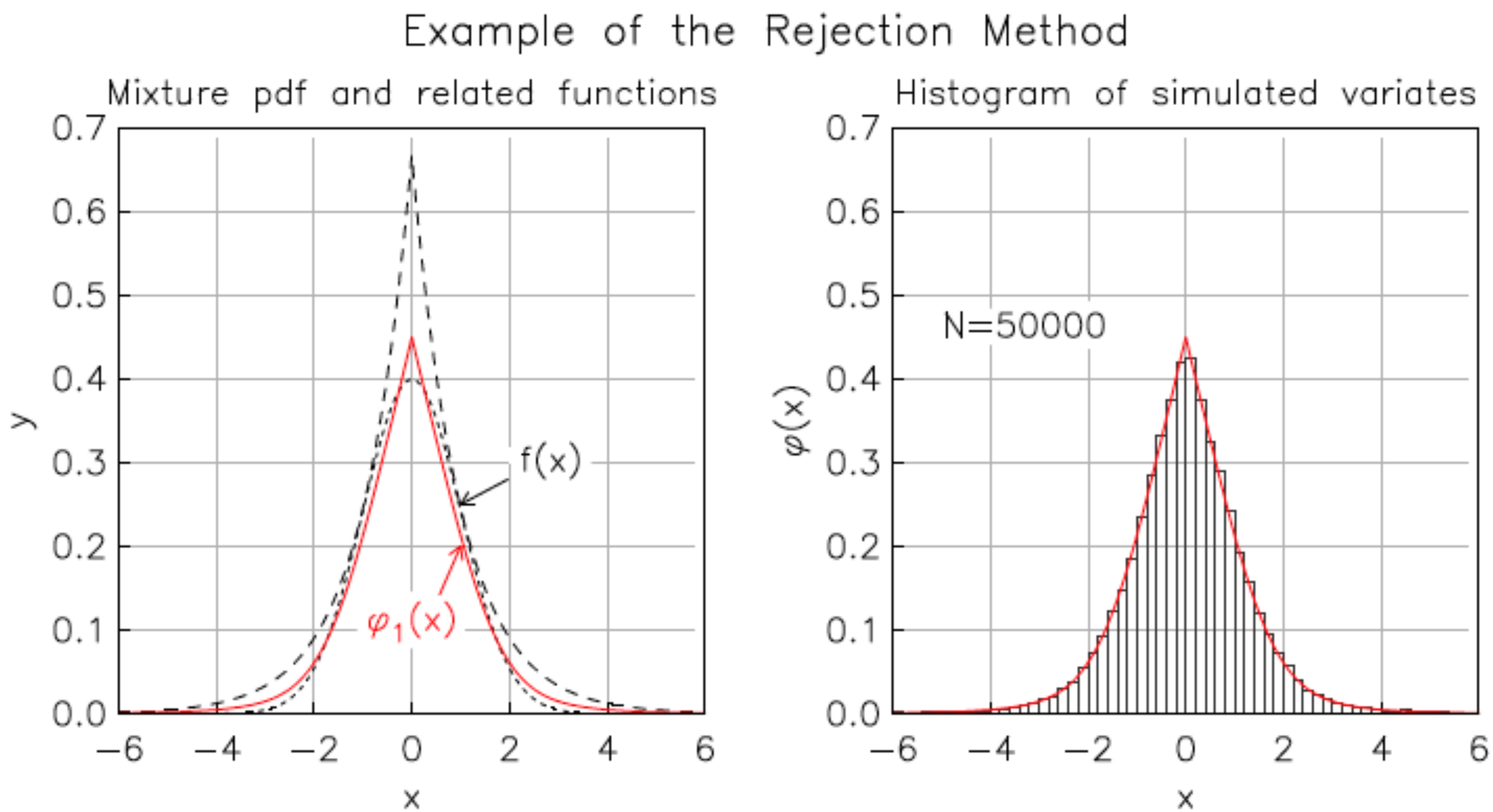


Figure 3.12: In the left panel, the dotted line shows a Gaussian pdf; the solid red line is the mixture pdf we wish to generate variates for, and the dashed line is the function used to produce variates for rejection. The right panel compares the target pdf (again in red) with a histogram of 50,000 variates produced using the rejection method.