A short course on statistical inference

Satyaki Bhattacharya
Saha Institute of Nuclear Physics

HEPAP-DAS School 5-9 December 2023

Recap

- Kolmogorov's definition of probability
- Bayes theorem
- Uniform random numbers
- Distribution of random numbers
- Histogram
- Probability mass functions and density functions (mostly from Glen Cowan's book and notes)

Events, Sample space

The set of all possible outcomes of a particular experiment is called the sample space S for that experiment

Coin toss S = {H,T}
Tossing two coins S = { HH, HT, TH, TT }

Event : Any subset of S

Random variables: variables characterising events

Momentum of an electron produced in the reaction $pp \rightarrow Z \rightarrow e^+e^-$

A definition of probability

Consider a set S with subsets A, B, ...

For all $A \subset S, P(A) \geq 0$

$$P(S) = 1$$

If $A \cap B = \emptyset, P(A \cup B) = P(A) + P(B)$

Kolmogorov axioms (1933)

From these axioms we can derive further properties, e.g.

$$P(\overline{A}) = 1 - P(A)$$

$$P(A \cup \overline{A}) = 1$$

$$P(\emptyset) = 0$$

if $A \subset B$, then $P(A) \leq P(B)$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Probability mass, density functions

$$P(x \text{ found in } [x, x+dx]) = f(x)\,dx$$

→ *f(x)* = probability density function (pdf)

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$ *x* must be somewhere

Or for discrete outcome x_i with e.g. $i = 1, 2, \ldots$ we have

$$P(x_i) = p_i$$ probability mass function

$$\sum_i P(x_i) = 1$$ *x* must take on one of its possible values

Conditional probability, independence

Also define conditional probability of A given B (with $P(B) \neq 0$):

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

E.g. rolling dice: $P(n < 3 \,|\, n \text{ even}) = \frac{P((n<3) \cap n \text{ even})}{P(\text{even})} = \frac{1/6}{3/6} = \frac{1}{3}$

Subsets A, B independent if: $P(A \cap B) = P(A)P(B)$

If A, B independent, $P(A|B) = \frac{P(A)P(B)}{P(B)} = P(A)$

N.B. do not confuse with disjoint subsets, i.e., $A \cap B = \emptyset$

Bayes' theorem

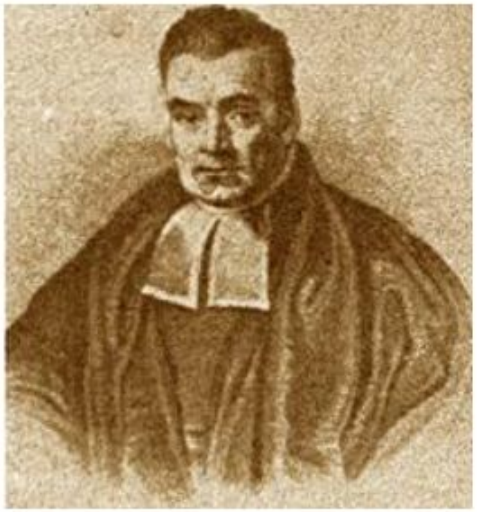
From the definition of conditional probability we have,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{and} \quad P(B|A) = \frac{P(B \cap A)}{P(A)}$$

but $P(A \cap B) = P(B \cap A)$, so

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayes' theorem



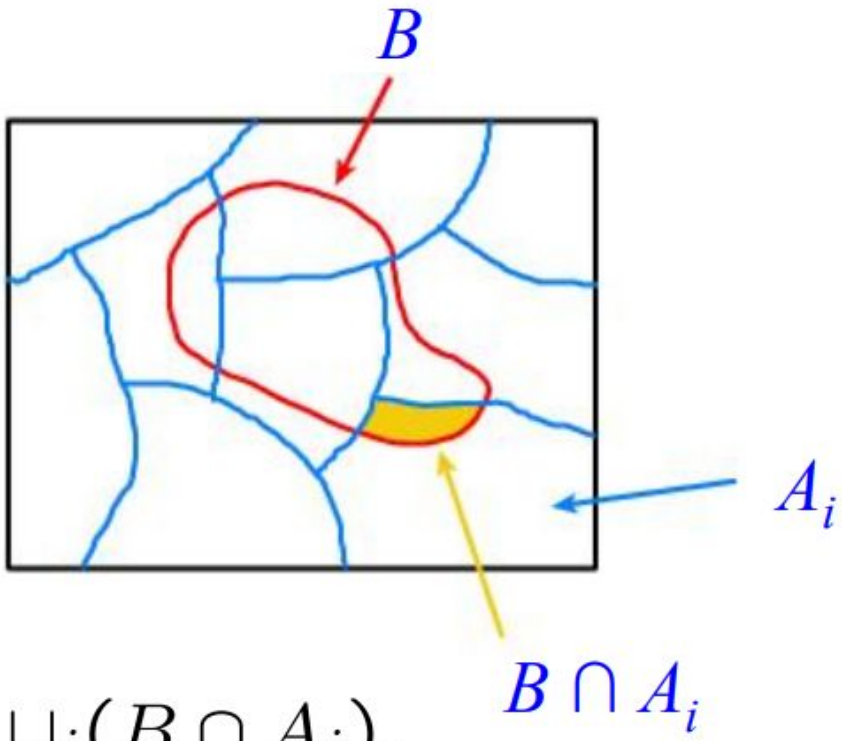
First published (posthumously) by the Reverend Thomas Bayes (1702−1761)

An essay towards solving a problem in the doctrine of chances, Philos. Trans. R. Soc. **53** (1763) 370; reprinted in Biometrika, **45** (1958) 293.

The law of total probability

Consider a subset B of the sample space S,

divided into disjoint subsets A_i such that $\cup_i A_i = S$,



$$\rightarrow \quad B = B \cap S = B \cap (\cup_i A_i) = \cup_i (B \cap A_i),$$

$$\rightarrow \quad P(B) = P(\cup_i (B \cap A_i)) = \sum_i P(B \cap A_i)$$

$$\rightarrow \quad P(B) = \sum_i P(B|A_i) P(A_i)$$ law of total probability

Bayes' theorem becomes

$$P(A|B) = \frac{P(B|A)P(A)}{\sum_i P(B|A_i)P(A_i)}$$

An example using Bayes' theorem

Suppose the probability (for anyone) to have AIDS is:

$$P(\text{AIDS}) = 0.001$$
$$P(\text{no AIDS}) = 0.999$$

← prior probabilities, i.e., before any test carried out

Consider an AIDS test: result is + or −

$$P(+|\text{AIDS}) = 0.98$$
$$P(-|\text{AIDS}) = 0.02$$

← probabilities to (in)correctly identify an infected person

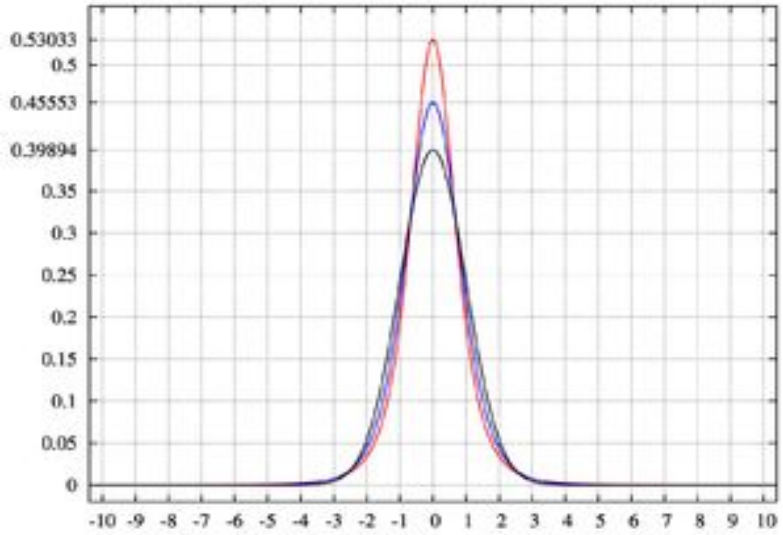
$$P(+|\text{no AIDS}) = 0.03$$
$$P(-|\text{no AIDS}) = 0.97$$

← probabilities to (in)correctly identify an uninfected person

Suppose your result is +. How worried should you be?

Mean, Variance, Skewness, kurtosis

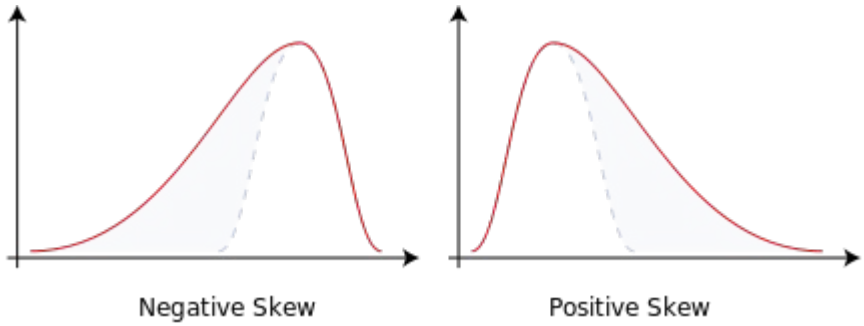
- Variance is the second standardised moment
- 3rd and 4th standardized moments (Pearson moments)



$$\gamma_1 = \mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3}{\sigma^3} = \frac{\mathrm{E}\left[(X-\mu)^3\right]}{(\mathrm{E}\left[(X-\mu)^2\right])^{3/2}} = \frac{\kappa_3}{\kappa_2^{3/2}},$$

(- 3)

$$\beta_2 = \frac{\mathrm{E}[(X-\mu)^4]}{(\mathrm{E}[(X-\mu)^2])^2} = \frac{\mu_4}{\sigma^4}$$



Covariance, correlation

$$V_{xy} = E[(x-\mu_x)(y-\mu_y)] = E[xy] - \mu_x\mu_y$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\,y\,f(x,y)\,dx\,dy - \mu_x\mu_y,$$

denoted $\operatorname{cov}[x,y]$

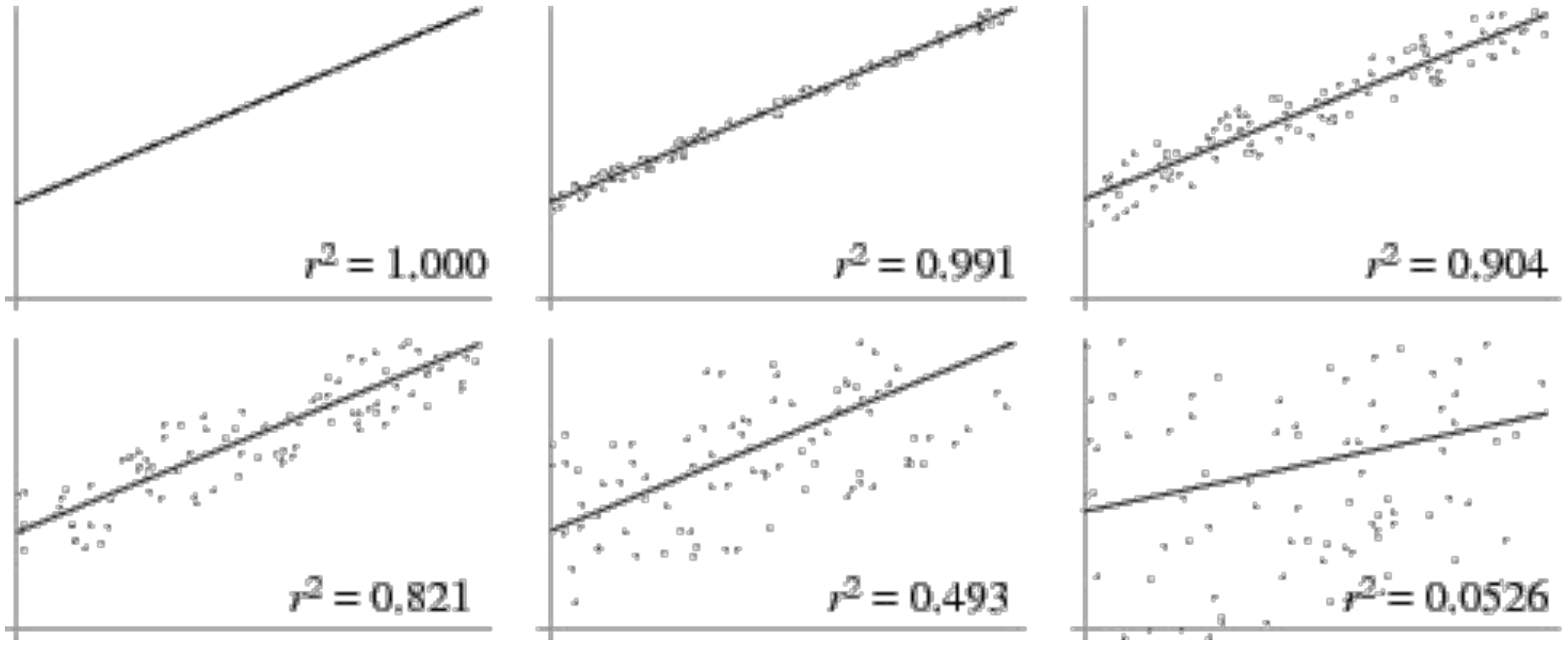
Covariance, correlation 2

$$\begin{aligned}
\mathrm{cov}[a,b] &= E[(a-\mu_a)(b-\mu_b)] \\
&= E[ab] - \mu_a\mu_b \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a\,b\,g(a,b)\,da\,db - \mu_a\mu_b \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} a(\mathbf{x})\,b(\mathbf{x})\,f(\mathbf{x})dx_1\ldots dx_n - \mu_a\mu_b
\end{aligned}$$

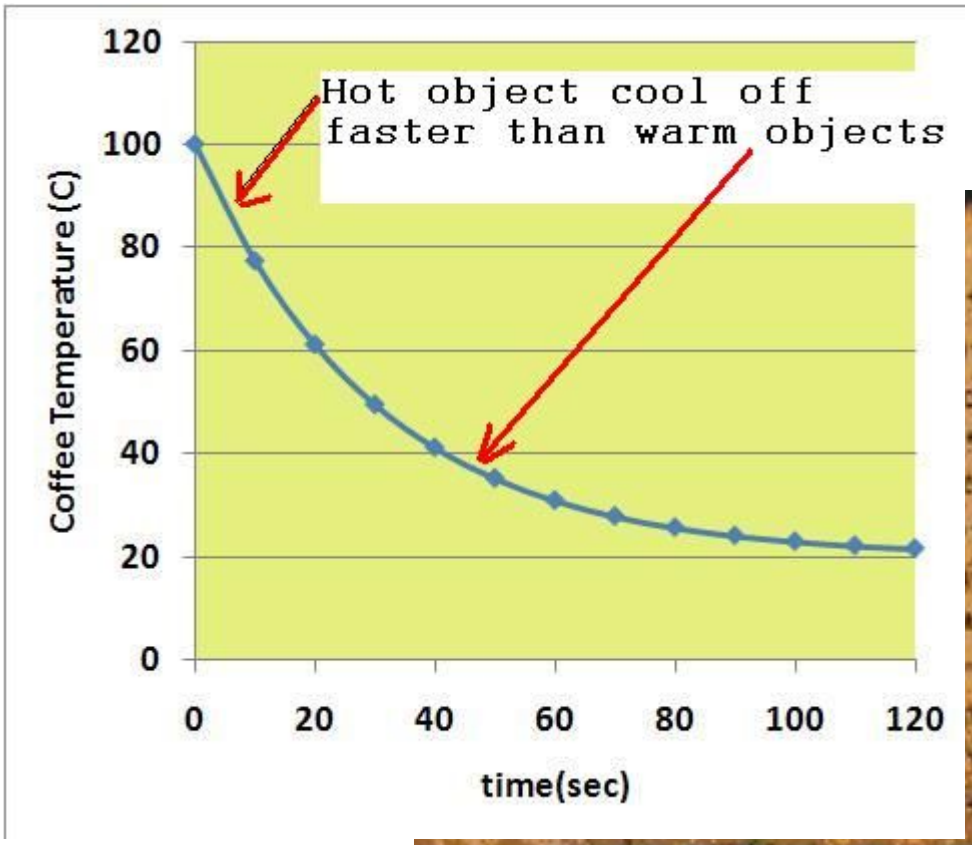
$$\rho_{xy} = \frac{V_{xy}}{\sigma_x\sigma_y}$$

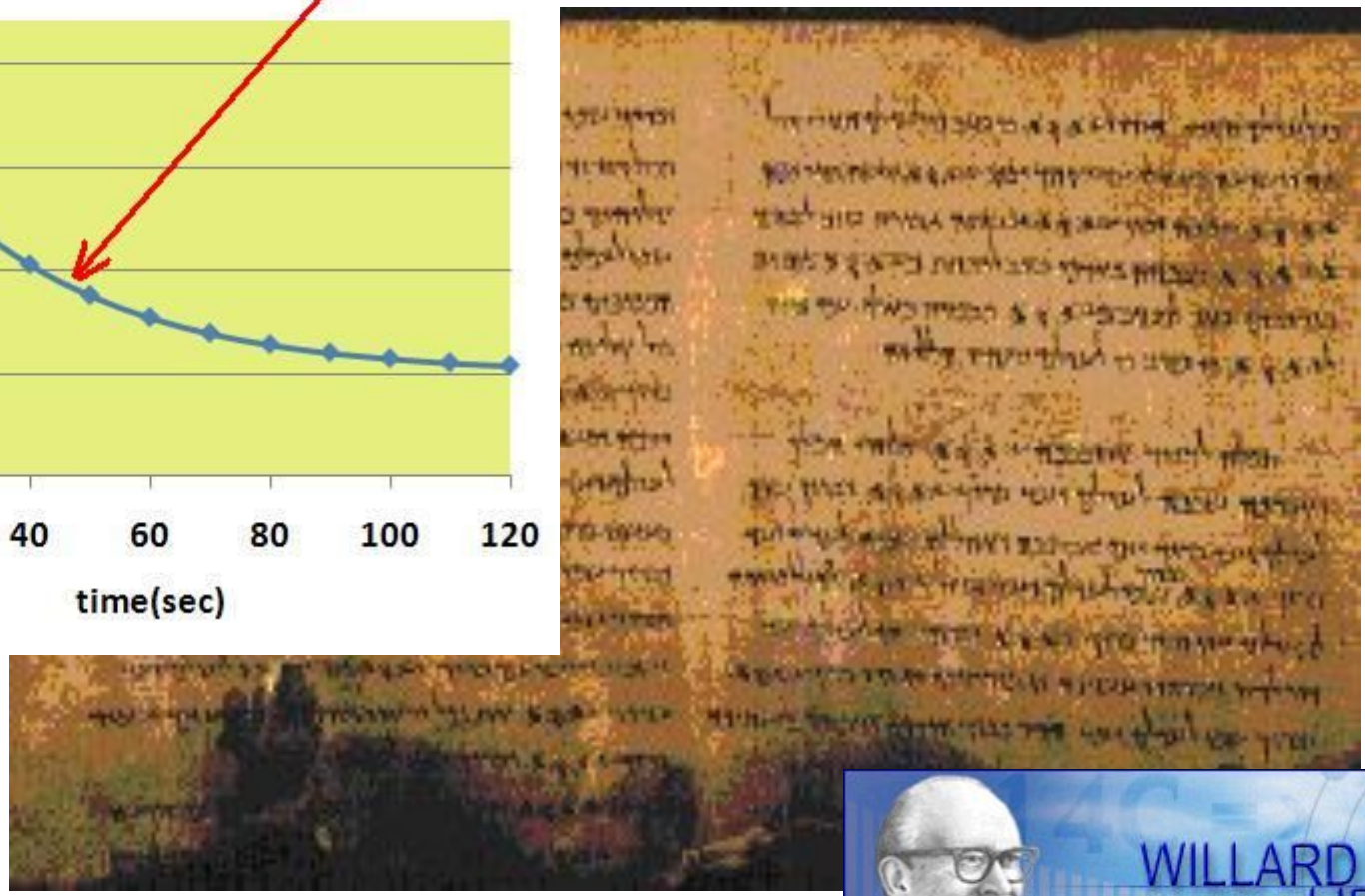
Prove $-1 \le \rho_{xy} \le 1$

Scatter plot and regression



exponential





Exponential 2

$$f(t;\tau) = \frac{1}{\tau} e^{-t/\tau} \qquad (\tau = \text{mean lifetime})$$

- E[t] = τ
- V[t] = τ^2
- Memoryless: $f(t - t_0 | t \geq t_0) = f(t)$
- Cosmic muon lifetime experiment
- Homework: Check the statements

binomial

- Out of ten generated uniform random numbers, in {0,1} how many will fall in the interval {0,0.1}?
  - Random, with average 1
- What is the probability that exactly two uniform random numbers out of ten generated urn's in {0,1}will fall between {0,0.1}
- $({}^{N}C_{n})p^{n}(1\text{-}p)^{(N-n)}$,p =0.1, N =10, n = 2

$$E[n] = \sum_{n=0}^{\infty} n \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n} = Np$$

$$\begin{aligned} V[n] &= E[n^2] - (E[n])^2 \\ &= Np(1-p). \end{aligned}$$

Multinomial distribution

Like binomial but now m outcomes instead of two, probabilities are

$$\vec{p} = (p_1, \ldots, p_m) , \quad \text{with} \quad \sum_{i=1}^{m} p_i = 1 .$$

For N trials we want the probability to obtain:

n_1 of outcome 1,
n_2 of outcome 2,
…
n_m of outcome m.

This is the multinomial distribution for $\vec{n} = (n_1, \ldots, n_m)$

$$f(\vec{n}; N, \vec{p}) = \frac{N!}{n_1! n_2! \cdots n_m!} p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m}$$

poisson

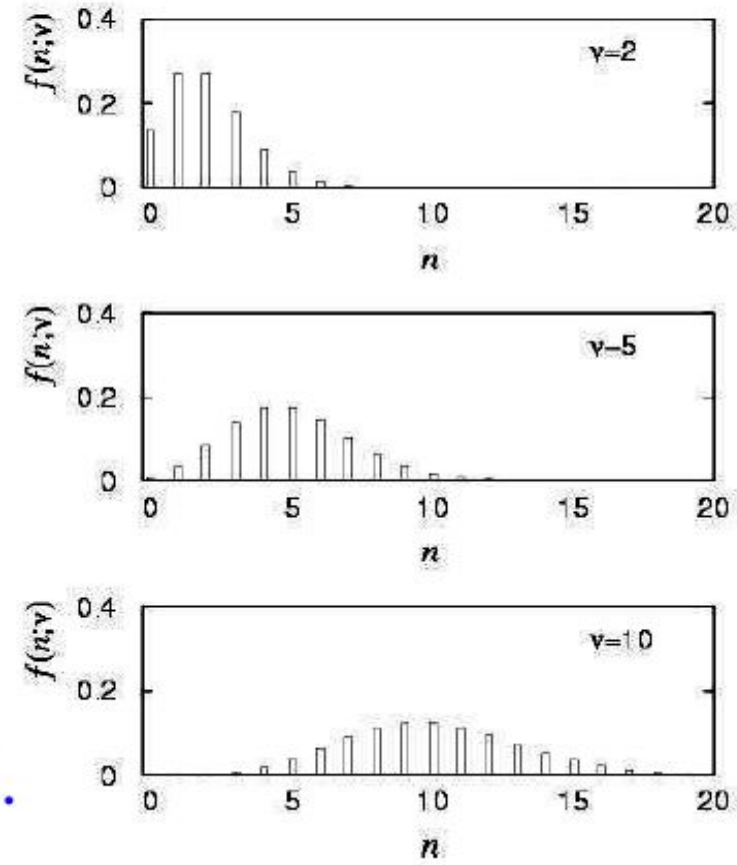
Consider binomial n in the limit

$$N \to \infty, \qquad p \to 0, \qquad E[n] = Np \to \nu \; .$$

$\to$ n follows the Poisson distribution:

$$f(n;\nu) = \frac{\nu^n}{n!} e^{-\nu} \quad (n \geq 0)$$

$$E[n] = \nu \, , \quad V[n] = \nu \; .$$



Example: number of scattering events n with cross section σ found for a fixed integrated luminosity, with $\nu = \sigma \int L \, dt$.

Gaussian (normal)

Multivariate Gaussian pdf for the vector $\vec{x} = (x_1, \ldots, x_n)$:

$$f(\vec{x}; \vec{\mu}, V) = \frac{1}{(2\pi)^{n/2}|V|^{1/2}} \exp\left[-\frac{1}{2}(\vec{x} - \vec{\mu})^T V^{-1}(\vec{x} - \vec{\mu})\right]$$

$\vec{x}$, $\vec{\mu}$ are column vectors, $\vec{x}^T$, $\vec{\mu}^T$ are transpose (row) vectors,

$$E[x_i] = \mu_i, \ , \quad \text{cov}[x_i, x_j] = V_{ij} \ .$$

For $n = 2$ this is

$$f(x_1, x_2, ; \mu_1, \mu_2, \sigma_1, \sigma_2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

$$\times \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]\right\}$$

where $\rho = \text{cov}[x_1, x_2]/(\sigma_1\sigma_2)$ is the correlation coefficient.

Characteristic functions

$$\phi_x(k) = E[e^{ikx}] = \int_{-\infty}^{\infty} e^{ikx} f(x) dx.$$

- Characteristic functions are fourier transforms of pdf, in one to one correspondence with the pdf's
- For sum of random variables $z = \sum_i x_i$

$$\begin{aligned}
\phi_z(k) &= \int \dots \int \exp\left(ik \sum_{i=1}^{n} x_i\right) f_1(x_1) \dots f_n(x_n) dx_1 \dots dx_n \\
&= \int e^{ikx_1} f_1(x_1) dx_1 \dots \int e^{ikx_n} f_n(x_n) dx_n \\
&= \phi_1(k) \dots \phi_n(k). \longrightarrow \boxed{f(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_z(k)\, e^{-ikz}\, dk.}
\end{aligned}$$

Characteristic functions 2

Distribution	p.d.f.	$\phi(k)$
Binomial	$f(n;N,p) = \frac{N!}{n!(N-n)!}\, p^n\,(1-p)^{N-n}$	$[p(e^{ik}-1)+1]^N$
Poisson	$f(n;\nu) = \frac{\nu^n}{n!}e^{-\nu}$	$\exp[\nu(e^{ik}-1)]$
Uniform	$f(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta-\alpha} & \alpha \le x \le \beta \\ 0 & \text{otherwise} \end{cases}$	$\frac{e^{i\beta k}-e^{i\alpha k}}{(\beta-\alpha)ik}$
Exponential	$f(x;\xi) = \frac{1}{\xi}e^{-x/\xi}$	$\frac{1}{1-ik\xi}$
Gaussian	$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$	$\exp(i\mu k - \frac{1}{2}\sigma^2 k^2)$
Chi-square	$f(z;n) = \frac{1}{2^{n/2}\Gamma(n/2)} z^{n/2-1}e^{-z/2}$	$(1-2ik)^{-n/2}$
Cauchy	$f(x) = \frac{1}{\pi}\,\frac{1}{1+x^2}$	$e^{-\lvert k\rvert}$

Characteristic functions 3

$$\left.\frac{d^m}{dk^m}\phi_z(k)\right|_{k=0} = \left.\frac{d^m}{dk^m}\int e^{ikz} f(z)dz\right|_{k=0}$$

$$= i^m \int z^m f(z)dz$$

$$= i^m \mu'_m$$

$$E[x] = \left.-i\frac{d}{dk}[\exp(i\mu k - \tfrac{1}{2}\sigma^2 k^2)]\right|_{k=0} = \mu,$$

$$V[x] = E[x^2] - (E[x])^2$$

$$= \left.-\frac{d^2}{dk^2}[\exp(i\mu k - \tfrac{1}{2}\sigma^2 k^2)]\right|_{k=0} - \mu^2 = \sigma^2.$$

Characteristic function 4

$$\phi(k) = [p(e^{ik} - 1) + 1]^N.$$

Taking the limit $p \to 0$, $N \to \infty$ with $\nu = pN$ constant gives

$$\phi(k) = \left(\frac{\nu}{N}(e^{ik} - 1) + 1\right)^N \to \exp[\nu(e^{ik} - 1)],$$

Cauchy

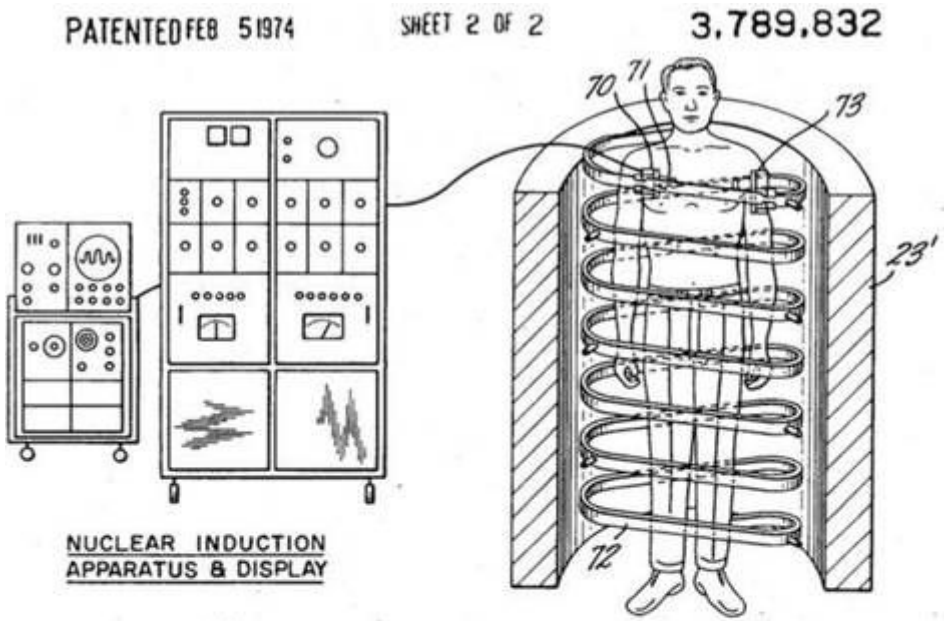


FIG. 2

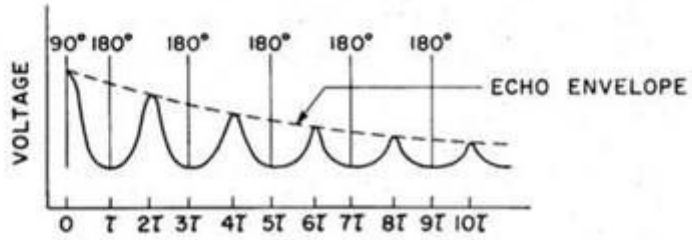
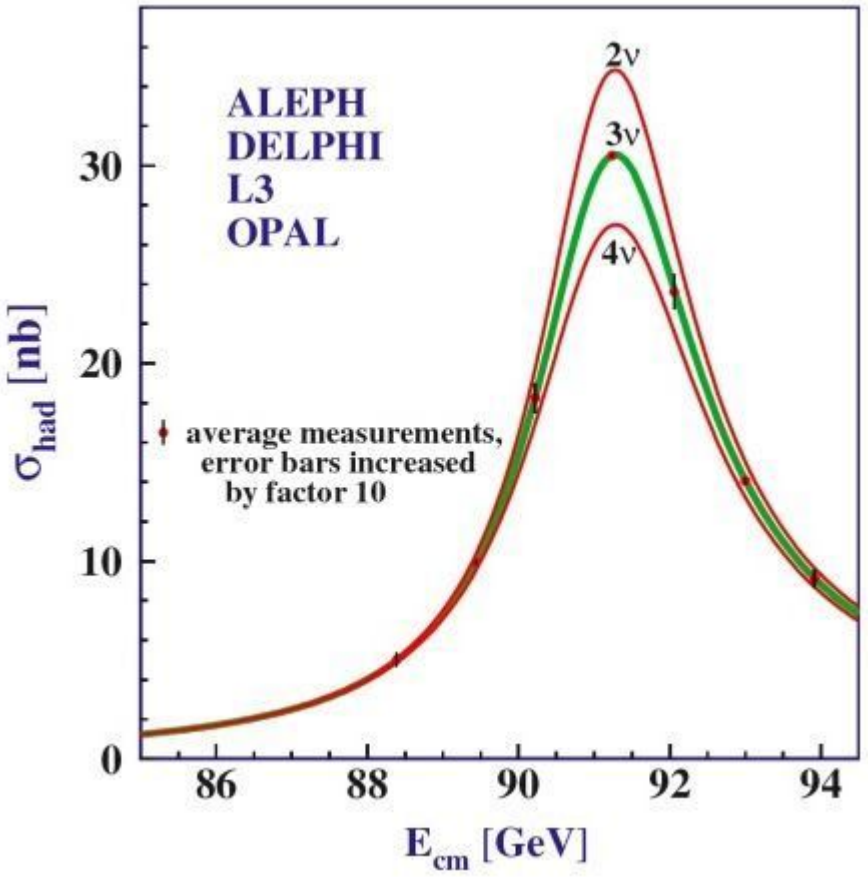
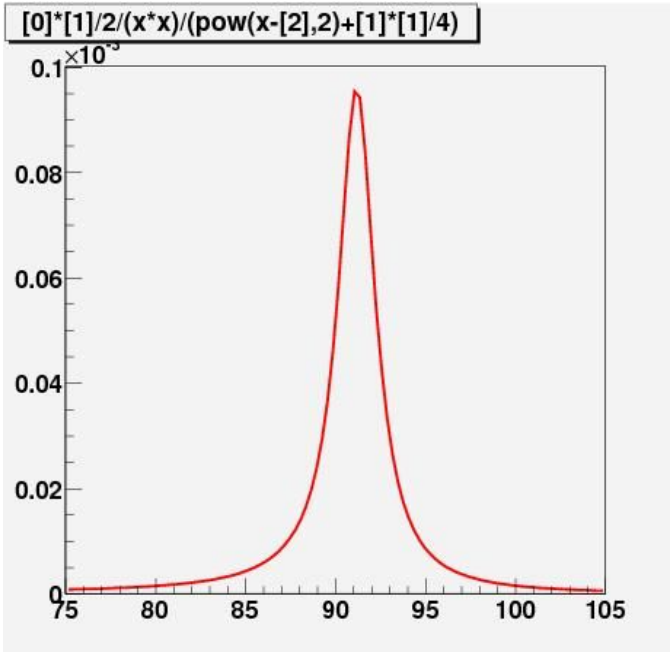


FIG. 3



Cauchy



$$f(x;\Gamma,x_0) = \frac{1}{\pi}\frac{\Gamma/2}{\Gamma^2/4 + (x-x_0)^2}$$

$(\Gamma = 2, x_0 = 0$ is the Cauchy pdf.)

- E(x) is not well defined
- V(x) tends to infinity
- Mode x0
- Fwhm: Γ

Student's t

$$f(x;\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}\left(1+\frac{x^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$$

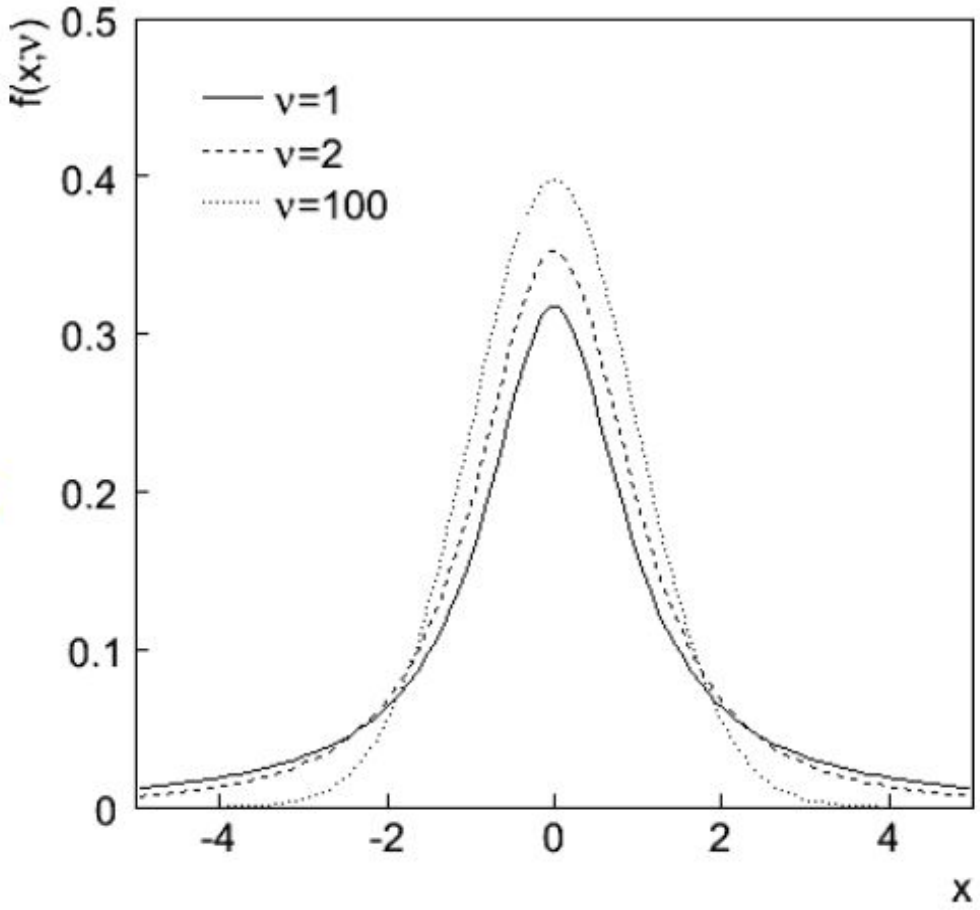
$$E[x] = 0 \quad (\nu > 1)$$

$$V[x] = \frac{\nu}{\nu-2} \quad (\nu > 2)$$

ν = number of degrees of freedom
(not necessarily integer)

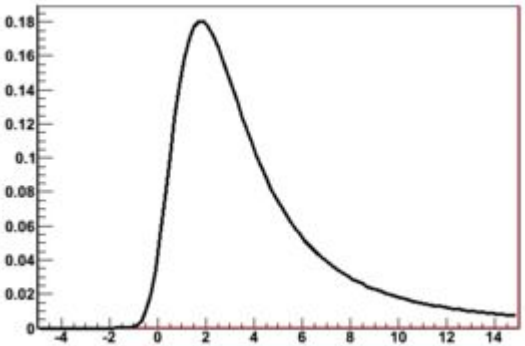
$\nu = 1$ gives Cauchy,

$\nu \to \infty$ gives Gaussian.



Ionization loss

- Charged particle ionization is used to detect radiation
  - Cosmic radiation was detected with Wulf electrometer
  - Mars rover RAD detects radiation using ionization
- Delta rays give a long tail in energy loss due to ionization - Landau

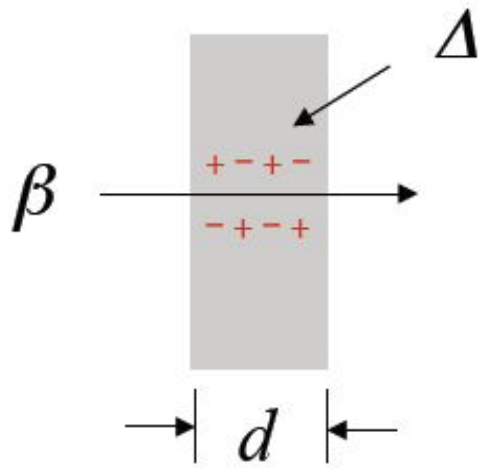


Landau distribution

$$f(\Delta;\beta) = \frac{1}{\xi}\phi(\lambda)\ ,$$

$$\phi(\lambda) = \frac{1}{\pi}\int_0^\infty \exp(-u\ln u - \lambda u)\sin\pi u\, du\ ,$$

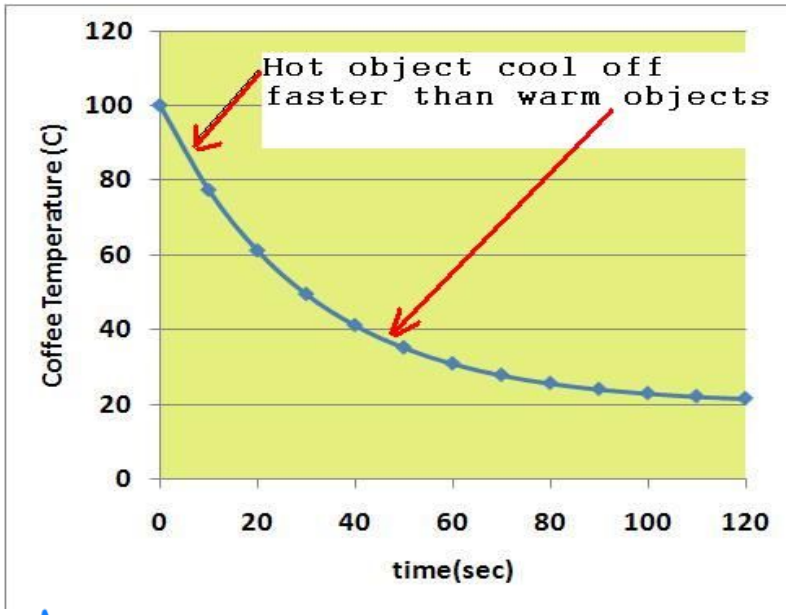
$$\lambda = \frac{1}{\xi}\left[\Delta - \xi\left(\ln\frac{\xi}{\epsilon'} + 1 - \gamma_{\mathrm{E}}\right)\right]\ ,$$

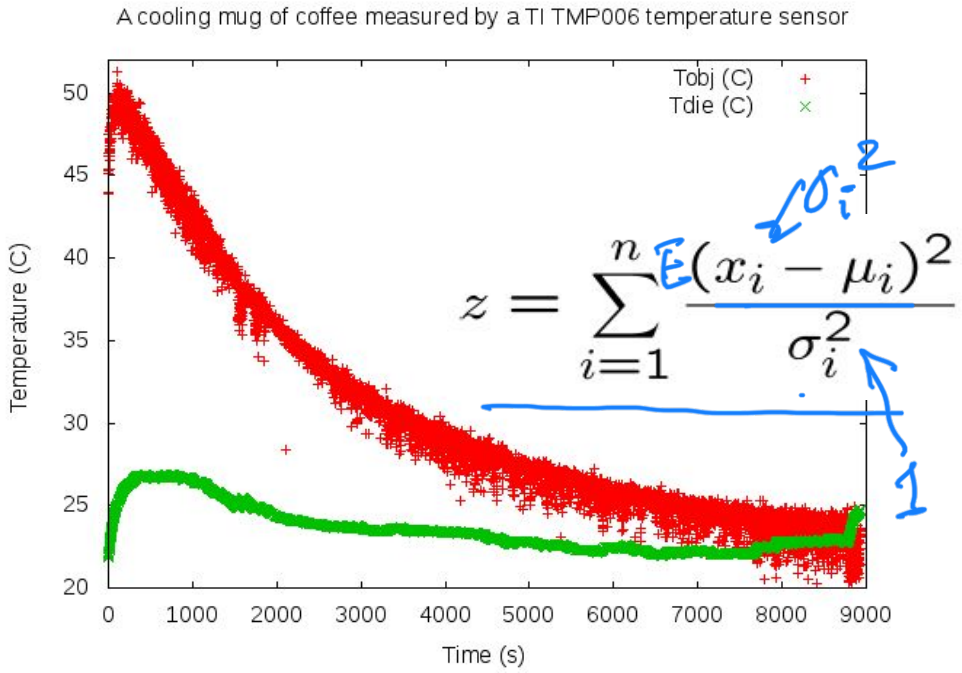


$$\xi = \frac{2\pi N_{\mathrm{A}} e^4 z^2 \rho \sum Z}{m_{\mathrm{e}} c^2 \sum A}\frac{d}{\beta^2}\ , \qquad \epsilon' = \frac{I^2 \exp\beta^2}{2m_{\mathrm{e}}c^2\beta^2\gamma^2}\ .$$

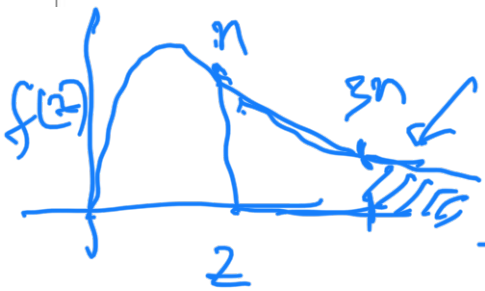
- All moments infinite!

Chi-square (X^2)





$$z = \sum_{i=1}^{n} \frac{(x_i - \mu_i)^2}{\sigma_i^2}$$



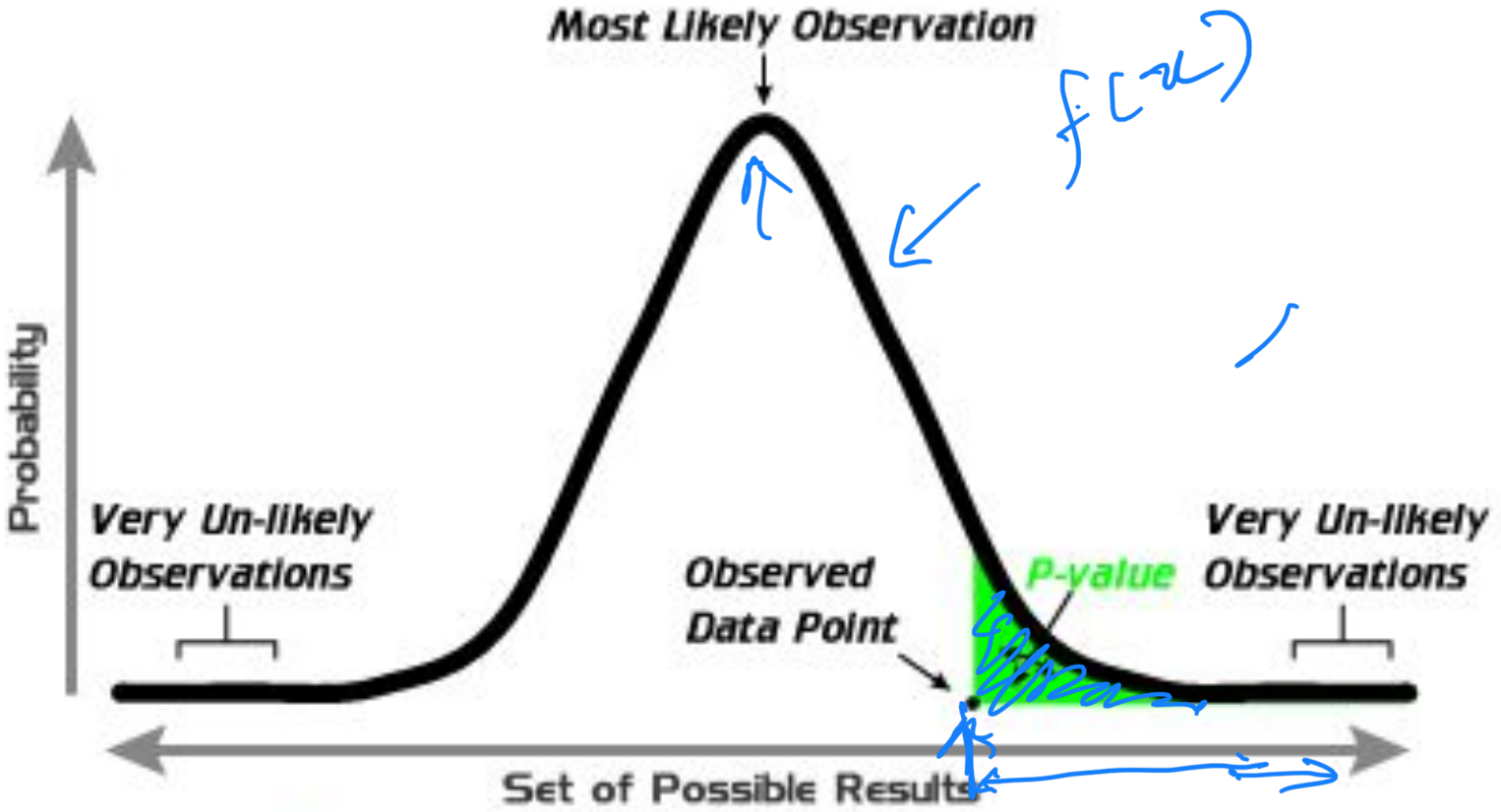
$$f(z;n) = \frac{1}{2^{n/2}\Gamma(n/2)} z^{n/2-1} e^{-z/2}$$

$n = 1, 2, \ldots =$ number of 'degrees of freedom' (dof)

$\chi^2/\text{ndof} \approx 1$

$$E[z] = n, \quad V[z] = 2n\,.$$

p-value



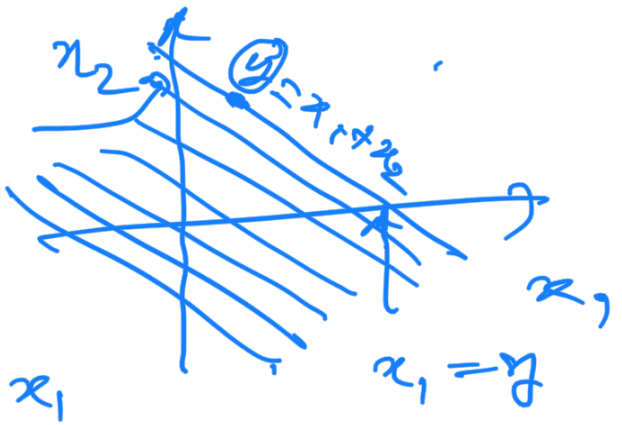
A p-value (shaded green area) is the probability of an observed (or more extreme) result arising by chance

Source: wikipedia

Sum of random numbers?

- Question: How is the sum of two uniform random numbers distributed?
- How is the sum of many uniform random numbers distributed?

$$F(y) = \int_{-\infty}^{\infty} f_1(x_1)\, F(y - x_1)\, dx_1$$



$$\int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{y-x_1} f_1(x_1) \cdot f_2(x_2)\, dx_1 dx_2$$

Central limit theorem

- Consider sum (xj) with mean mu_j and standard deviation sigma_j

$$y_j = \frac{x_j - \mu_j}{\sqrt{n}}$$

$$\phi_j(k) = \sum_{m=0}^{\infty} \left. \frac{d^m \phi}{dk^m} \right|_{k=0} \frac{k^m}{m!}$$

$$= \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} E[y^m]$$

$$= 1 - \frac{k^2}{2n}\sigma^2 - \frac{ik^3}{3!} \frac{E[(x_j - \mu_j)^3]}{n^{3/2}} + \ldots,$$

Central limit theorem 2

$$z = \sum_j y_j$$

$$\phi_z(k) = \prod_{j=1}^{n} \phi_j(k) = \prod_{j=1}^{n} \left(1 - \frac{k^2}{2n}\sigma^2 - \frac{ik^3}{3!}\frac{E[(x_j - \mu_j)^3]}{n^{3/2}} + \ldots\right)$$

- Neglecting higher orders

$$\phi_z(k) \approx \left(1 - \frac{k^2}{2n}\sigma^2\right)^n \rightarrow \exp\left(-\tfrac{1}{2}\sigma^2 k^2\right)$$

- Gaussian with mean 0 and variance sigma^2
- Transforming back, one gets Gaussian with sum (mu_j) and variance n*sigma^2
- Important: sample mean is Gaussian distributed
- Fails for Landau, Cauchy

Central limit theorem 2

$$z = \sum_j y_j$$

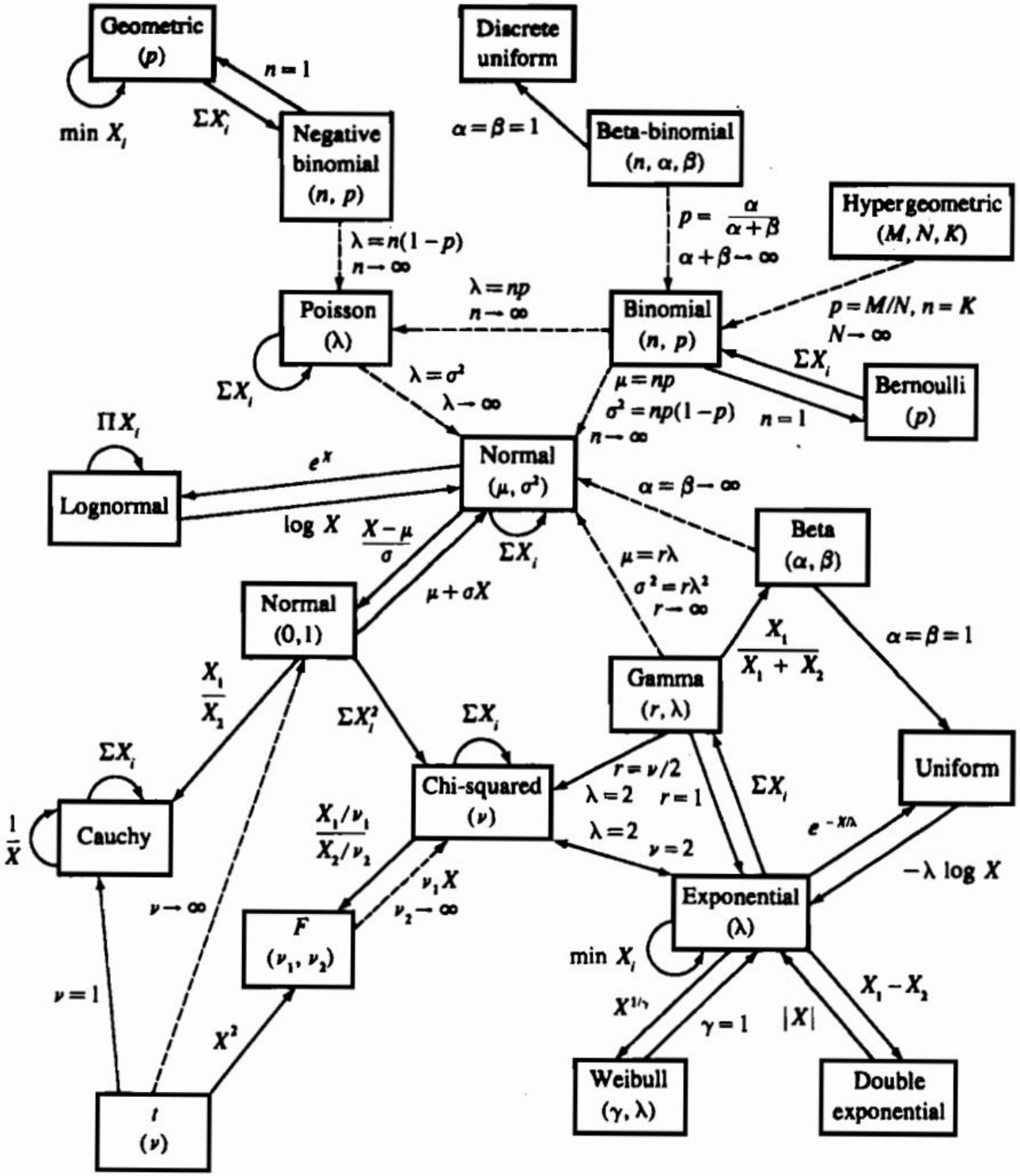
$$\phi_z(k) = \prod_{j=1}^{n} \phi_j(k) = \prod_{j=1}^{n} \left(1 - \frac{k^2}{2n}\sigma^2 - \frac{ik^3}{3!}\frac{E[(x_j - \mu_j)^3]}{n^{3/2}} + \ldots\right)$$

- Neglecting higher orders

$$\phi_z(k) \approx \left(1 - \frac{k^2}{2n}\sigma^2\right)^n \to \exp\left(-\tfrac{1}{2}\sigma^2 k^2\right)$$

$\sum_{i=1}^{10} \mu_j = n\sigma^2$

- Gaussian with mean 0 and variance sigma^2
- Transforming back, one gets Gaussian with sum (mu_j) and variance n*sigma^2
- Important: sample mean is Gaussian distributed
- Fails for Landau, Cauchy


Geometric (p)
min X_i
$n = 1$
ΣX_i
Negative binomial (n, p)
Discrete uniform
$\alpha = \beta = 1$
Beta-binomial (n, α, β)
$p = \frac{\alpha}{\alpha+\beta}$
$\alpha+\beta \to \infty$
Hypergeometric (M, N, K)
$\lambda = n(1-p)$
$n \to \infty$
Poisson (λ)
$\lambda = np$
$n \to \infty$
Binomial (n, p)
$p = M/N, n = K$
$N \to \infty$
ΣX_i
$\lambda = \sigma^2$
$\lambda \to \infty$
$\mu = np$
$\sigma^2 = np(1-p)$
$n \to \infty$
$n = 1$
Bernoulli (p)
ΠX_i
e^X
Lognormal
log X
Normal (μ, σ²)
$\frac{X-\mu}{\sigma}$
$\mu + \sigma X$
$\alpha = \beta \to \infty$
Beta (α, β)
$\mu = r\lambda$
$\sigma^2 = r\lambda^2$
$r \to \infty$
Normal (0,1)
$\frac{X_1}{X_2}$
ΣX_i^2
ΣX_i
$\frac{X_1}{X_1 + X_2}$
Gamma (r, λ)
$\alpha = \beta = 1$
Uniform
ΣX_i
Cauchy
$\frac{1}{X}$
Chi-squared (ν)
$r = \nu/2$
$\lambda = 2$
$r = 1$
$\lambda = 2$
$\nu = 2$
ΣX_i
$e^{-X/\lambda}$
$-\lambda \log X$
$\frac{X_1/\nu_1}{X_2/\nu_2}$
$\nu_1 X$
$\nu_2 \to \infty$
$\nu \to \infty$
F (ν_1, ν_2)
Exponential (λ)
min X_i
$\nu = 1$
X^2
$X^{1/\gamma}$
$\gamma = 1$
$|X|$
$X_1 - X_2$
t (ν)
Weibull (γ, λ)
Double exponential

Lecture 3

Linear Congruential Generator

- Goal: Generate Un uniform in the interval [0,1)
- Generate Xn in [0,m), Un = Xn/m
- Xn+1 = (a*Xn + c)%m – Linear congruential series
- Four constants required
- X0 (starting value/seed), a (multiplier), c (increment/bias),m (modulus)
- X0 = a = c = 7, m = 10 will give 7, 6, 9, 0, 7, 6, 9, 0, ...
  - Four magic numbers required:

Linear Congruential Generator 2

- Xn+1 =(65539*Xn)%pow(2,31)
- This is essentially RANDU, most popular generator for many years
  - Multiplicative congruential method (Lehman's original method)
  - Mixed congruential method C != 0
- For the math (number theory): http://www.math.cornell.edu/~mec/Winter2009/Luo/Linear%20Congruential%20Generator/linear%20congruential%20gen1.html

Code for linear congruential generator

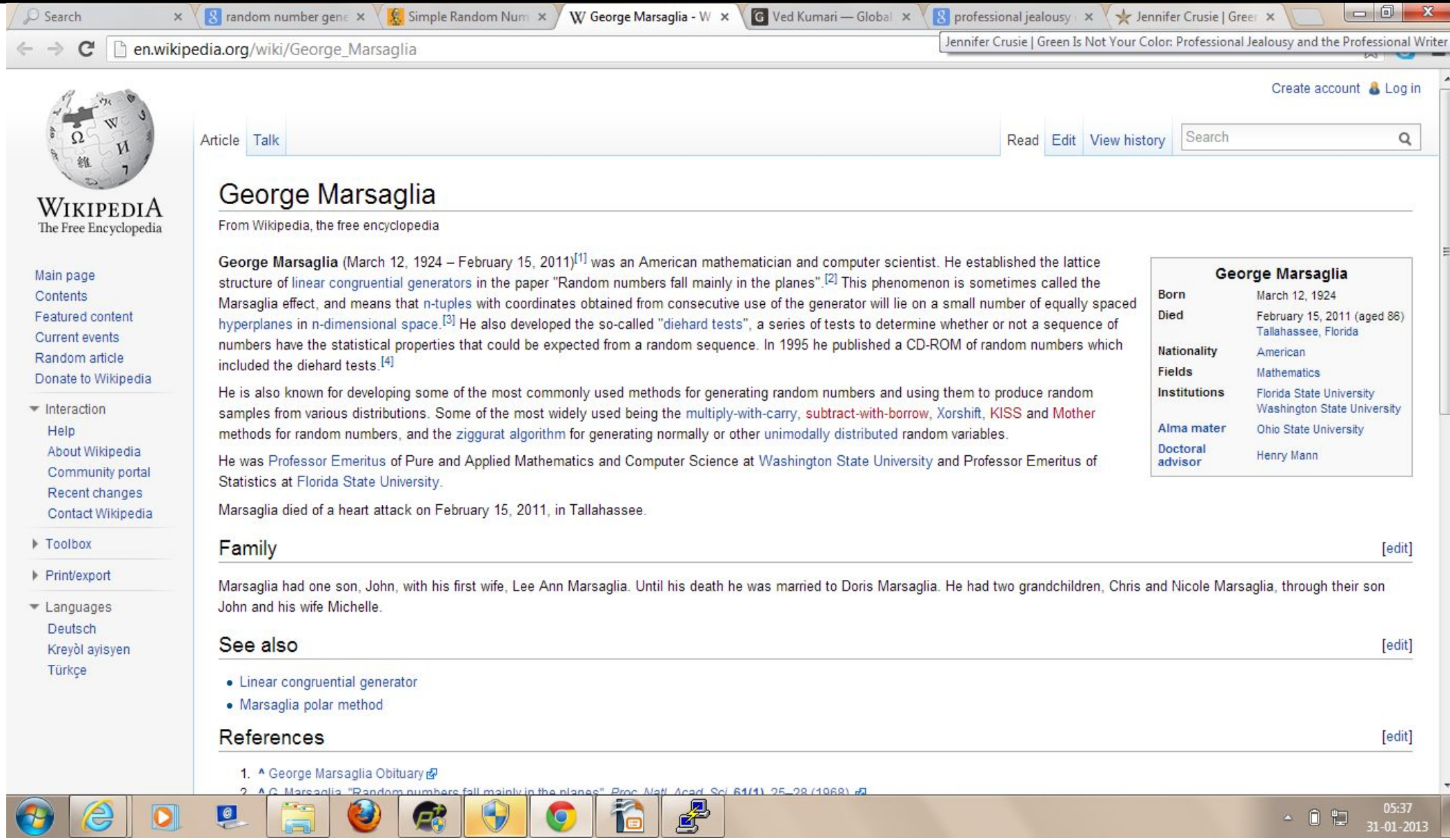
```
#include <iostream>
#include <ostream>
#include <cmath>
#include <TMath.h>
#include <TRandom2.h>
#include <TH1.h>
#include <TH1D.h>
#include <TCanvas.h>
#include <TStyle.h>


double GetUniform()
{
  Static int X0 = 12345, m = 0, Xn =
0;
  m = pow(2,31);
  Xn=X0;
  Xn = (65539*Xn)%m;
  return (double)Xn/(double)m;
}

int main(){

  std::cout<<GetUniform()<<std::endl;
```

Marsaglia


Search
random number gene
Simple Random Num
George Marsaglia - W
Ved Kumari — Global
professional jealousy
Jennifer Crusie | Gree
Jennifer Crusie | Green Is Not Your Color: Professional Jealousy and the Professional Writer
en.wikipedia.org/wiki/George_Marsaglia
Create account
Log in
Article
Talk
Read
Edit
View history
Search
WIKIPEDIA
The Free Encyclopedia
Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia
Toolbox
Print/export
Languages
Deutsch
Kreyòl ayisyen
Türkçe
George Marsaglia
From Wikipedia, the free encyclopedia
George Marsaglia (March 12, 1924 – February 15, 2011)[1] was an American mathematician and computer scientist. He established the lattice structure of linear congruential generators in the paper "Random numbers fall mainly in the planes".[2] This phenomenon is sometimes called the Marsaglia effect, and means that n-tuples with coordinates obtained from consecutive use of the generator will lie on a small number of equally spaced hyperplanes in n-dimensional space.[3] He also developed the so-called "diehard tests", a series of tests to determine whether or not a sequence of numbers have the statistical properties that could be expected from a random sequence. In 1995 he published a CD-ROM of random numbers which included the diehard tests.[4]
He is also known for developing some of the most commonly used methods for generating random numbers and using them to produce random samples from various distributions. Some of the most widely used being the multiply-with-carry, subtract-with-borrow, Xorshift, KISS and Mother methods for random numbers, and the ziggurat algorithm for generating normally or other unimodally distributed random variables.
He was Professor Emeritus of Pure and Applied Mathematics and Computer Science at Washington State University and Professor Emeritus of Statistics at Florida State University.
Marsaglia died of a heart attack on February 15, 2011, in Tallahassee.
George Marsaglia
Born March 12, 1924
Died February 15, 2011 (aged 86) Tallahassee, Florida
Nationality American
Fields Mathematics
Institutions Florida State University Washington State University
Alma mater Ohio State University
Doctoral advisor Henry Mann
Family [edit]
Marsaglia had one son, John, with his first wife, Lee Ann Marsaglia. Until his death he was married to Doris Marsaglia. He had two grandchildren, Chris and Nicole Marsaglia, through their son John and his wife Michelle.
See also [edit]
Linear congruential generator
Marsaglia polar method
References [edit]
1. ^ George Marsaglia Obituary
2. ^ G. Marsaglia, "Random numbers fall mainly in the planes", Proc. Natl. Acad. Sci. 61(1), 25–28 (1968)
05:37
31-01-2013

Random numbers stay mainly in the plane

Search × | random number gene × | Simple Random Num × | George Marsaglia - W × | RANDOM NUMBERS | × | linear congruential ge × | SectionII: Linear Cong ×
www.pnas.org/content/61/1/25.full.pdf+html

RANDOM NUMBERS FALL MAINLY IN THE PLANES

By George Marsaglia

MATHEMATICS RESEARCH LABORATORY, BOEING SCIENTIFIC RESEARCH LABORATORIES, SEATTLE, WASHINGTON

Communicated by G. S. Schairer, June 24, 1968

Virtually all the world's computer centers use an arithmetic procedure for generating random numbers. The most common of these is the multiplicative congruential generator first suggested by D. H. Lehmer. In this method, one merely multiplies the current random integer I by a constant multiplier K and keeps the remainder after overflow:

$$\text{new } I = K \times \text{old } I \quad \text{modulo } M.$$

The apparently haphazard way in which successive multiplications by a large integer K produce remainders after overflow makes the resulting numbers work surprisingly well for many Monte Carlo problems. Scores of papers have reported favorably on cycle length and statistical properties of such generators.

The purpose of this note is to point out that all multiplicative congruential

The purpose of this note is to point out that all multiplicative congruential random number generators have a defect—a defect that makes them unsuitable for many Monte Carlo problems and that cannot be removed by adjusting the starting value, multiplier, or modulus. The problem lies in the "crystalline" nature of multiplicative generators—if n-tuples $(u_1,u_2,\ldots,u_n)$, $(u_2,u_3,\ldots,u_{n+1}),\ldots$ of uniform variates produced by the generator are viewed as points in the unit cube of n dimensions, then *all* the points will be found to lie in a relatively small number of parallel hyperplanes. Furthermore, there are many systems of parallel hyperplanes which contain all of the points; the points are about as randomly spaced in the unit n-cube as the atoms in a perfect crystal at absolute zero.

One can readily think of Monte Carlo problems where such regularity in "random" points in n-space would be unsatisfactory; more disturbing is the possibility that for the past 20 years such regularity might have produced bad, but unrecognized, results in Monte Carlo studies which have used multiplicative generators.

Multiply with carry

```
uint GetUint()
{
    m_z = 36969 * (m_z & 65535) + (m_z >> 16);
    m_w = 18000 * (m_w & 65535) + (m_w >> 16);
    return (m_z << 16) + m_w;
}
```

Test of randomness

- Diehard tests (marsaglia 1995)
- Birthday spacings, parking lot test, the craps test, monkey tests (based on infinite monkey theorem), count the 1's,...
- See eg:http://en.wikipedia.org/wiki/Diehard_tests

Test with linear congruential numbers

lecture-3

Other distributions from uniform variate

- Uniform random numbers can be used to generate other distributions
- Let x be uniform in (0.,1.), we want a new random number a in (a1,a2) distributed as g(a)
- Conservation of probability:
- g(a)da = f(x)dx; f(x) = 1.
- g(a) = |dx/da|
- If g(a) is desired to be exponential then:
- (1/D)*exp(-a/D) = |dx/da| (D = const parameter)

Usfulness of randomness

- What is the probability of getting two sixes in 10 throws of a fair dice?
- Example code dicethrow
- What is the probability of two successive sixes in 10 throws of a dice?
- Modify dicethrow