Statistical Tools Ft. Noise Analysis

Jiachen Jiang

University of British Columbia, Vancouver, BC

December 30, 2014

Introduction

Fourier Transforms

- Fourier series
- Fourier integral

Orthodox statistics

- Probability distributions
- Moments, moment-generating function and central limit theorem
- χ^2 distribution and student's t distribution
- Robust estimation
- Propagation of errors

Stochastic process and noise

- Stochastic process and stationary process
- Spectral density of Poisson random process and Gaussian random process
- Filtering
- Photon and thermal noise in black body radiation

Orthogonal basis functions

$$\int_{-1/2}^{1/2} cos2\pi nxsin2\pi nxdx = 0 \tag{1}$$

$$\int_{-1/2}^{1/2} cos2\pi mxcos2\pi nxdx = \frac{1}{2}\delta_{mn} \tag{2}$$

$$\int_{-1/2}^{1/2} sin2\pi mxsin2\pi nxdx = \frac{1}{2}\delta_{mn} \tag{3}$$

Series expansion f(x) in the range [-0.5:0.5]

$$f(x) = \frac{a_0}{2} + \sum_{1}^{\infty}(a_n cos2\pi nx + b_n sin2\pi nx) \tag{4}$$

Fourier series

$$f(x) = \frac{a_0}{2} + \sum_{1}^{\infty}(a_n cos2\pi nx + b_n sin2\pi nx) \tag{5}$$

$$a_n = 2\int_{-1/2}^{1/2} f(x)cos2\pi nxdx \tag{6}$$

$$a_n = 2\int_{-1/2}^{1/2} f(x)sin2\pi nxdx \tag{7}$$

Extended fourier series

$$f(x) = \frac{a_0}{2} + \sum_{1}^{\infty}(a_n cos 2\pi nx/L + b_n sin 2\pi nx/L) \tag{8}$$

$$a_n = 2\int_{-L/2}^{L/2} f(x) cos 2\pi nx/L dx \tag{9}$$

$$a_n = 2\int_{-L/2}^{L/2} f(x) sin 2\pi nx/L dx \tag{10}$$

In complex form

$$f(x) = \sum_{-\infty}^{\infty} \tilde{a_n} e^{i2\pi nx/L} \tag{11}$$

$$\tilde{a_n} = \frac{1}{2}\int_{-\infty}^{\infty} f(x) e^{-2\pi nx/L} dx \tag{12}$$

Show me an example

$$f(x) = \frac{|x|}{x} \tag{13}$$

where $0 < |x| < \pi$

$$f(x) = \frac{4}{\pi}(sinx + \frac{sin3x}{3} + \frac{sin5x}{5} + ...) \tag{14}$$

Hey,

Show me the picture on the board!

Fourier integrals

$L \to \infty$

$$f(x) = \int_{-\infty}^{\infty} F(x) e^{i2\pi sx} ds \tag{15}$$

$$F(s) = \int_{-\infty}^{\infty} f(x) e^{-i2\pi sx} dx \tag{16}$$

Points off

Coefficients are different in different fields.
Jiachen: define $-i$ transform as the forward transform and $+i$ as the reverse transform

Properties of fourier integrals

1. Symmetry properties

f(x)	F(s)
Real and even	Real and even
Real and odd	Imaginary and odd
Imaginary and even	Imaginary and even
Complex and even	Complex and even
Complex and odd	Complex and odd
Real even plus imaginary odd	Real
Real odd plus imaginary even	Imaginary

2. Scalar multiplication

$$\overrightarrow{af(x)} = a\overrightarrow{f(x)} \tag{17}$$

$$\overrightarrow{af(x) + bg(x)} = a\overrightarrow{f(x)} + b\overrightarrow{g(x)} \tag{18}$$

Dirac delta function

$$f(x) = \int_{-\infty}^{\infty} dse^{i2\pi sx} \int_{-\infty}^{\infty} f(x')e^{-i2\pi sx'}dx' \tag{19}$$

$$= \int_{-\infty}^{\infty} dx' f(x')[\int_{-\infty}^{\infty} e^{i2\pi s(x-x')}ds] \tag{20}$$

$\int_{-\infty}^{\infty} e^{i2\pi s(x-x')}ds = 0$ when $x \neq x'$

A conventional definition of the Dirac delta function is $\delta(x) = 0$ for $x \neq 0$ and yet the delta function integrates to unity

$$\int_{-\infty}^{\infty} \delta(x)dx = 1 \tag{21}$$

where

$$\delta(x) = \int_{-\infty}^{\infty} e^{i2\pi sx}ds \tag{22}$$

Parseval's theorem and a power spectrum

Parseval's theorem

$$\int_{-\infty}^{\infty} |f(t)|^2 = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} F(\nu)e^{i2\pi\nu t} d\nu \int_{-\infty}^{\infty} F^*(\nu')e^{-2i\pi\nu' t} d\nu' \tag{23}$$

$$= \int_{-\infty}^{\infty} d\nu F(\nu) F*(\nu) = \int_{-\infty}^{\infty} |F(\nu)|^2 d\nu \tag{24}$$

$\int_{-\infty}^{\infty} |f(t)|^2$ in the time domain equals $\int_{-\infty}^{\infty} |F(\nu)|^2 d\nu$ in the frequency domain. And the $|F(\nu)|^2$ is the power spectrum in the frequency domain.

Power specturm

Exponentially decaying oscillator

$$f(t) = e^{-\Gamma t/2} sin 2\pi\nu_0 t \quad (25)$$

$$F(\nu) = \int_{-\infty}^{\infty} f(t) e^{-2i\pi\nu t} = \int_{-\infty}^{\infty} e^{-\Gamma t/2} e^{-2i\pi\nu t} sin 2\pi\nu_0 t \quad (26)$$

$$= \frac{1}{2}\left(\frac{1}{2\pi\nu + 2\pi\nu_0 - i\Gamma/2} - \frac{1}{2\pi\nu - 2\pi\nu_0 - i\Gamma/2}\right) \quad (27)$$

Only keep the positive frequencies near ν_0 and with small damping ($\Gamma << 2\pi\nu_0$)

$$|F(\nu)|^2 \approx \frac{1}{4} \frac{1}{(2\pi\nu - 2\pi\nu_0)^2 + (\Gamma/2)^2} \quad (28)$$

Hey,

Show me the picture on the board!

Power spectrum

The power is proportional to $|f(t)|^2$ and according to Parseval's theorem to $|F(\omega)|^2$

Properties of Fourier transform

1. A narrow function in the time domain coreesponds to a broad function in the spectral domain, and vice versa.
2. And additional scaling in amplitude is required with similarity properties.
3. Translation in the time domain corresponds to a phase winding, and vice versa.

Properties of Fourier transform

Time shift

$$\mathcal{F}[x(t \pm t_0)] = X(j\omega)e^{\pm j\omega t_0}$$

Proof: Let $t' = t \pm t_0$, i.e., $t = t' \mp t_0$, we have

$$\mathcal{F}[x(t \pm t_0)] \quad : \quad \int_{-\infty}^{\infty} x(t \pm t_0) e^{-j\omega t} dt = \int_{-\infty}^{\infty} x(t') e^{-j\omega(t' \mp t_0)} dt'$$

$$: \quad e^{\pm j\omega t_0} \int_{-\infty}^{\infty} x(t') e^{-j\omega t'} dt' = X(j\omega) e^{\pm j\omega t_0}$$

Frequency shift

$$\mathcal{F}^{-1}[X(j\omega \pm \omega_0)] = x(t)e^{\mp j\omega_0 t}$$

Proof: Let $\omega' = \omega \pm \omega_0$, i.e., $\omega = \omega' \mp \omega_0$, we have

$$\mathcal{F}^{-1}[X(j(\omega \pm \omega_0))] \quad : \quad \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j(\omega \pm \omega_0)) e^{j\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\omega') e^{j\omega(\omega' \mp \omega_0)} d\omega'$$

$$: \quad e^{\mp j\omega_0 t} \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\omega') e^{j\omega'} d\omega' = x(t) e^{\mp j\omega_0 t}$$

Time and frequency scaling

$$\mathcal{F}[x(at)] = \frac{1}{a} X(\frac{\omega}{a}) \quad \text{or} \quad \mathcal{F}[ax(at)] = X(\frac{\omega}{a})$$

Proof: Let $u = at$, i.e., $t = u/a$, where $a > 0$ is a scaling factor, we have

$$\mathcal{F}[x(at)] = \int_{-\infty}^{\infty} x(at) e^{-j\omega t} dt = \int_{-\infty}^{\infty} x(u) e^{-j\omega u/a} d(u/a) = \frac{1}{a} X(\frac{\omega}{a})$$

Convolution

The convolution of two functions are defined as

$$g(x) = \int_{-\infty}^{\infty} f_1(u) f_2(x-u) du \tag{29}$$

$$g = f_1 * f_2 \tag{30}$$

Properties

$$f * g = g * f \tag{31}$$

$$f * (g + h) = f * g + f * h \tag{32}$$

$$f_1(x) * f_2(x) \leftrightarrow F_1(s) F_2(s) \tag{33}$$

$$f_1(x) f_2(x) \leftrightarrow F_1(s) * F_2(s) \tag{34}$$

Autocorrelation and Wiener-Khinchin theorem

Cross correlation of two functions are defined as

$$g(x) = \int_{-\infty}^{\infty} f_1(u) f_2(x+u) du \tag{35}$$

$$g(x) = f_1(x) \star f_2(x) \tag{36}$$

Hey,

Show me an difference from the convolution on the board!

Autocorrelation and Wiener-Khinchin theorem

Cross correlations are not commutative
However,

$$f_1(x) \star f_2(x) \leftrightarrow F_1^*(s)F_2(s) \tag{37}$$

Particular, the autocorrelation is

$$f(x) \star f(x) \leftrightarrow |F(s)|^2 \tag{38}$$

which peaks at $s = 0$.

Common functions and Fourier pairs

- Boxcar Π $(|x| < 1/2) \leftrightarrow \delta(s)$
- Triangle Λ $(|x| < 1) \leftrightarrow sinc^2(s)$
- Step function H $(x > 0) \leftrightarrow 1/2\delta(s) - i/(2\pi s)$
- Even impulse pair II $(\frac{\delta(x+1/2)+\delta(x-1/2)}{2})$ $cos\pi x$
- Odd impulse pair II $(\frac{\delta(x+1/2)-\delta(x-1/2)}{2})$ $sin\pi x$
- Shah $\sum\delta(x-n)$ $Shah(s)$

Aliasind and Shannon's sampling theorem

Most common method to measure a continuous function is at regular intervals Δt

$$G(\nu) = F(\nu) * shah(\nu \Delta t) \tag{39}$$

If $(\Delta t)^{-1} < 2\nu_{max}$, replicas overlap and information is lost, which is known as the aliasing.
If $(\Delta t)^{-1} > 2\nu_{max}$, there is no loss.

2.1 Probability distributions

Get me prepared,
Both discrete and continuous random variables may be correlated with probability distributions.

$$\sum_i P_i(x_i) = 1 \tag{40}$$

2.1 Probability distributions

If x_i are real valued, a *cumulative* or *integral* probability $P(x)$ can be defined as

$$P(x) = \sum_{i}^{x_i < x} p_i \tag{41}$$

and the derivative of the function is the probability density

$$p(x) = \frac{dP(x)}{dx} \tag{42}$$

- For a discrete random variable, p(x) consists of a set of delta functions
$$p(x) = \sum_j p_i \delta(x - x_j) \tag{43}$$
- The quantity $p(x)dx$ ($p(x)$) corresponds to of having an event within the integral dx at x.

2.1 Probability distributions

Binomial distribution
The number of combinations of n things taken k at a time

$$C_n^k = \frac{n!}{k!(n-k)!} \tag{44}$$

Let's flip a coin, which has 50/50 chance of coming up heads, there is a total of 2^n possible outcomes. The probability of obtaining k heads is

$$p_k = C_n^k \frac{1}{2^n} (= C_n^k f^k (1-f)^k) \tag{45}$$

2.1 Probability distributions

Poisson distribution
Extend the concept of probability to MULTIPLE and DISCRETE outcomes by taking a limiting case. $n \to \infty$ while $f \to 0$ in a wa that $nf = a$. Then,

$$\lim_{n\to+\infty} C_n^k = \lim_{n\to+\infty} \frac{n!}{k!(n-k)!} = \frac{n^k}{k!} \tag{46}$$

$$\lim_{n\to+\infty}(1-f)^{n-k} = \lim_{n\to+\infty}(1-f)^n = \lim_{n\to+\infty}(1-f)^{a/f} = e^{-a} \tag{47}$$

In terms of probability density,

$$p(k) = e^{-a}\sum_{k=0}^{\infty}\frac{a^k}{k!}\delta(x-k) \tag{48}$$

The cumulative probability distribution function for the Poisson case is related to the incomplete gamma function.

2.1 Probability distribution

radioactive decay at the average rate r

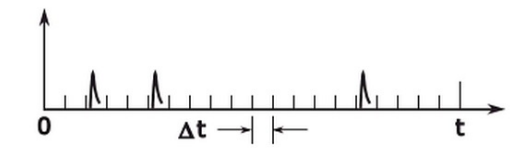


Figure: Subdivide the time from 0 to t into narrow bins with width of Δt in order that no two pulses are in the same bin.One expects to see $a = rt$ events. (Credit: E. C. Sutton)

The probability of k events is

$$p(k) = e^{-a} \frac{a^k}{k!} \tag{49}$$

Poisson noise in image

Figure: Poisson noise with different mean a= 1, 5, 99

2.1 Probability distribution

Gaussian (normal) distribution
$n \to \infty$ but with finite f so that $nf \to \infty$. The probability will peak near $k = nf \to \infty$. Stirling's formula,

$$n! \approx \sqrt{2\pi n}(\frac{n}{e})^n(1 + \frac{1}{12n} + ...) \tag{50}$$

$$p_k = C_n^k f^k (1-f)^{n-k} \approx \frac{1}{\sqrt{2\pi n}}(\frac{k}{n})^{-k-\frac{1}{2}}(\frac{n-k}{n})^{-n+k+\frac{1}{2}} f^k (1-f)^{n-k} \tag{51}$$

Consider at small deviations ϵ around nf, let $k = nf + \epsilon$ where $\epsilon << nf$,

$$p_k \approx \frac{1}{\sqrt{2\pi n}} \frac{1}{\sqrt{f(1-f)}} exp[-\frac{\epsilon^2}{2nf(1-f)}] \tag{52}$$

2.1 Probability distribution

Conventional form $\sigma^2 = nf(1-f)$ and a mean value μ and pass from the discrete to the continuous,

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \tag{53}$$

Cumulative probability

$$P(x) = \int_{-\infty}^{x} p(k)dk = \frac{1}{2}[1 + erf(\frac{x-\mu}{\sqrt{2}\sigma})] \tag{54}$$

where the error function is defined as

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{55}$$

2.1 Probability distributions

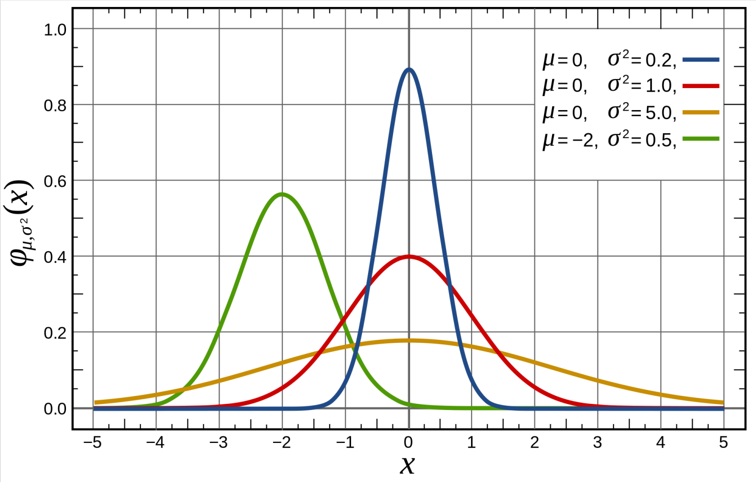


Figure: Probability density of Gaussian distributions for different μ and σ.

Gaussian noise in image

Figure: Gaussian noise with zero mean and different deviation $\sigma = 0.5, 1, 5$.

2.2.1 Moments of a distribution

Known distribution p(x)

0th order

The law of probability distributions

$$1 = \int_{-\infty}^{+\infty} p(x)dx \tag{56}$$

1st order

The mean of the distribution is

$$\mu =< x >= \int_{-\infty}^{+\infty} xp(x)dx \tag{57}$$

2nd order

The variance of the distribution is

$$\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 p(x) dx \quad (58)$$

- The normal distribution is a two independent parameter distribution with a mean μ and a standard deviation σ.
- The Poisson distribution is a one parameter distribution with a mean $\mu = a$ and a standard deviation $\sigma = \sqrt{a}$.

CAUSE SOME PROBLEM $\rightarrow$ Student's t distribution

higher order

Higher orders are no longer robust indicators. (Don't hurry! We will talk about it in minutes!)

2.2.2 Moment-generating function

Let's do Fourier transform!

$$\phi(k) = \int_{-\infty}^{+\infty} p(x) e^{ikx} dx \tag{59}$$

$$= \int_{-\infty}^{+\infty} p(x)[1 + ikx - \frac{1}{2}k^2x^2 - \frac{i}{3!}k^3x^3 + ...]dx \tag{60}$$

$$= i + ik < x > + \frac{(ik)^2}{2!} < x^2 > + \frac{(ik)^3}{3!} < x^3 > + ... \tag{61}$$

$\phi(k)$ generates the moments of the distribution, known as the moment-generating function.

2.2.2 Moment-generating function

Normal distribution

$$\phi(k) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} e^{ikx} dx \quad (62)$$

Gaussian's Fourier transform is still a Gaussian multiplied by a complex factor.

$f(x+y) = x + y$

$$\phi_z(k) = \int\int e^{ik(x+y)} p(x) dx p(y) dy = \phi_x(k)\phi_y(k) \quad (63)$$

Moment-generating function of the sum of two independent random variables is the product of their individual functions.

2.2.3 Central limit theorem

Consider a random variable x with a probability density $p(x)$, mean μ, variance σ_x^2 and unspecified higher moments.

$$\phi_{x-\mu_x}(k) = \int e^{ik(x-\mu_x)} p(x) dx \tag{64}$$

Zero mean?!

2.2.3 Central limit theorem

n measurements of x and form the sum and average, from the transform of $f(x+y) = x + y$

$$a = \frac{s}{n} = \frac{x_1 + x_2 + \ldots}{n} \tag{65}$$

$$\phi_{a-\mu_x}(k) = [\phi_{x-\mu_x}(\frac{k}{n})]^n = [1 - \frac{1}{2}\frac{k^2\sigma_x^2}{n^2} + O(\frac{k^3}{n^3})]^n \tag{66}$$

When $n \to \infty$

$$\phi_{a-\mu_x}(k) = e^{-k^2\sigma_x^2/2n} \tag{67}$$

$$p(a) = \frac{\sqrt{n}}{\sqrt{2\pi}\sigma_x} e^{-n(a-\mu_x)^2/2\sigma_x^2} \tag{68}$$

This is normal distribution with mean $\mu_a = \mu_x$ and a standard deviation $\sigma_a = \sigma_x/\sqrt{n}$, no matter the initial distribution $p(x)$.

2.2.3 Central limit theorem

Rainfall

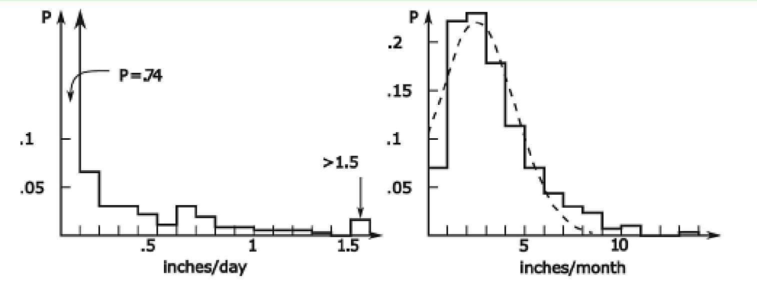


Figure: Left panel: 0.74 days in 1995 Urbana, Illinois had less than 0.1 inches of rain. Six days had more than 1.5 inches. Right panel: Monthly rainfall statistics for 25 years (1984-2008) is fit using Gaussian shown in a dashed line. (Credit: E. C. Sutton)

2.2.3 Central limit theorem

- Well-defined mean and standard deviation
- Whenever you add or average large numbers of data, the sum or average of measurement approaches a normal distribution. The width gets narrower with variance reduced by n and the standard deviation reduced by $\sqrt{n}$.
- It only works for random errors but not systematic errors, generated by some specified physics process. And other mathematical operators such as power, exponentiation or even multitude doesn't work.

2.3.1 χ^2 distribution

Consider a variable x is described by a probability distribution $p(x)$ with known μ and σ. We make n measurements of x. Define a quantity known as

$$\chi^2 = \sum_{i=1}^{n} (\frac{x_i - \mu}{\sigma})^2 \tag{69}$$

to how well this set of data is described by a normal distribution. Also defined reduced χ^2 as

$$\chi_v^2 = \frac{\chi^2}{v} \tag{70}$$

where $v = n - m$ and m is the number of degree.

2.3.1 χ^2 distribution

If we assume $p(x)$ is a Gaussian, we can calculate the probability density of χ^2

$$p_n\chi^2 = \frac{1}{2^{n/2}\Gamma(n/2)}(\chi^2)^{n/2-1}e^{-\chi^2/2} \tag{71}$$

The expectation value of χ^2 is the mean of the distribution

$$<\chi^2>= n \tag{72}$$

m=2 if the mean and the standard deviation are derived from the data instead od being assumed a priori. Then

$$<\chi^2_v>=\frac{<\chi^2>}{v}=\frac{v}{v}=1 \tag{73}$$

and the variance

$$\sigma^2(\chi^2)=2v \tag{74}$$

ν	P 0.99	0.98	0.95	0.90	0.80	0.70	0.60	0.50
1	0.00016	0.00063	0.00393	0.0158	0.0642	0.148	0.275	0.455
2	0.0100	0.0202	0.0515	0.105	0.223	0.357	0.511	0.693
3	0.0383	0.0617	0.117	0.195	0.335	0.475	0.623	0.789
4	0.0742	0.107	0.178	0.266	0.412	0.549	0.688	0.839
5	0.111	0.150	0.229	0.322	0.469	0.600	0.731	0.870
6	0.145	0.189	0.273	0.367	0.512	0.638	0.762	0.891
7	0.177	0.223	0.310	0.405	0.546	0.667	0.785	0.907
8	0.206	0.254	0.342	0.436	0.574	0.691	0.803	0.918
9	0.232	0.281	0.369	0.463	0.598	0.710	0.817	0.927
10	0.256	0.306	0.394	0.487	0.618	0.727	0.830	0.934
11	0.278	0.328	0.416	0.507	0.635	0.741	0.840	0.940
12	0.298	0.348	0.436	0.525	0.651	0.753	0.848	0.945
13	0.316	0.367	0.453	0.542	0.664	0.764	0.856	0.949
14	0.333	0.383	0.469	0.556	0.676	0.773	0.863	0.953
15	0.349	0.399	0.484	0.570	0.687	0.781	0.869	0.956
16	0.363	0.413	0.498	0.582	0.697	0.789	0.874	0.959
17	0.377	0.427	0.510	0.593	0.706	0.796	0.879	0.961
18	0.390	0.439	0.522	0.604	0.714	0.802	0.883	0.963
19	0.402	0.451	0.532	0.613	0.722	0.808	0.887	0.965
20	0.413	0.462	0.543	0.622	0.729	0.813	0.890	0.967
22	0.434	0.482	0.561	0.638	0.742	0.823	0.897	0.970
24	0.452	0.500	0.577	0.652	0.753	0.831	0.902	0.972
26	0.469	0.516	0.592	0.665	0.762	0.838	0.907	0.974
28	0.484	0.530	0.605	0.676	0.771	0.845	0.911	0.976
30	0.498	0.544	0.616	0.687	0.779	0.850	0.915	0.978
32	0.511	0.556	0.627	0.696	0.786	0.855	0.918	0.979
34	0.523	0.567	0.637	0.704	0.792	0.860	0.921	0.980
36	0.534	0.577	0.646	0.712	0.798	0.864	0.924	0.982
38	0.545	0.587	0.655	0.720	0.804	0.868	0.926	0.983
40	0.554	0.596	0.663	0.726	0.809	0.872	0.928	0.983
42	0.563	0.604	0.670	0.733	0.813	0.875	0.930	0.984
44	0.572	0.612	0.677	0.738	0.818	0.878	0.932	0.985
46	0.580	0.620	0.683	0.744	0.822	0.881	0.934	0.986
48	0.587	0.627	0.690	0.749	0.825	0.884	0.936	0.986
50	0.594	0.633	0.695	0.754	0.829	0.886	0.937	0.987
60	0.625	0.662	0.720	0.774	0.844	0.897	0.944	0.989
70	0.649	0.684	0.739	0.790	0.856	0.905	0.949	0.990
80	0.669	0.703	0.755	0.803	0.865	0.911	0.952	0.992
90	0.686	0.718	0.768	0.814	0.873	0.917	0.955	0.993
100	0.701	0.731	0.779	0.824	0.879	0.921	0.958	0.993
120	0.724	0.753	0.798	0.839	0.890	0.928	0.962	0.994
140	0.743	0.770	0.812	0.850	0.898	0.934	0.965	0.995
160	0.758	0.784	0.823	0.860	0.905	0.938	0.968	0.996

Figure: Values of χ_v^2 for which the probability of χ_v^2 exceeding that probability at the top of each column.

χ^2 distribution

χ^2 is also used to fit a function to a set of data. Consider independent variable x_i to be error free and variable y_i to have uncertainties σ_i.

$$\chi^2 = \sum_{i=1}^{n} (\frac{f(x_i) - y_i}{\sigma})^2 \tag{75}$$

The functional parameters are chosen to minimize χ^2.

χ^2 distribution

Linear fitting

$$S = \sum \frac{(N_j - \alpha N_j')^2}{\sigma} \tag{76}$$

α is used to minimize L_r.

$$\alpha = \frac{\sum N_j' N_j}{\sum \frac{(N_j')^2}{\sigma}} \tag{77}$$

χ^2 distribution

able 3-1 Critical Values of the χ^2 Distribution

	P									
df	0.995	0.975	0.9	0.5	0.1	0.05	0.025	0.01	0.005	df
1	.000	.000	0.016	0.455	2.706	3.841	5.024	6.635	7.879	1
2	0.010	0.051	0.211	1.386	4.605	5.991	7.378	9.210	10.597	2
3	0.072	0.216	0.584	2.366	6.251	7.815	9.348	11.345	12.838	3
4	0.207	0.484	1.064	3.357	7.779	9.488	11.143	13.277	14.860	4
5	0.412	0.831	1.610	4.351	9.236	11.070	12.832	15.086	16.750	5
6	0.676	1.237	2.204	5.348	10.645	12.592	14.449	16.812	18.548	6
7	0.989	1.690	2.833	6.346	12.017	14.067	16.013	18.475	20.278	7
8	1.344	2.180	3.490	7.344	13.362	15.507	17.535	20.090	21.955	8
9	1.735	2.700	4.168	8.343	14.684	16.919	19.023	21.666	23.589	9
10	2.156	3.247	4.865	9.342	15.987	18.307	20.483	23.209	25.188	10
11	2.603	3.816	5.578	10.341	17.275	19.675	21.920	24.725	26.757	11
12	3.074	4.404	6.304	11.340	18.549	21.026	23.337	26.217	28.300	12
13	3.565	5.009	7.042	12.340	19.812	22.362	24.736	27.688	29.819	13
14	4.075	5.629	7.790	13.339	21.064	23.685	26.119	29.141	31.319	14
15	4.601	6.262	8.547	14.339	22.307	24.996	27.488	30.578	32.801	15

Figure: Values of χ^2 for which the probability of χ^2 exceeding that probability at the top of wach column.

> Large χ^2 value probably means the incorrect choice of the function.

Wait a minute!

More generally, we have variable x_i described by probability density $p(x)$ with unknown moments μ and σ.

If we know, why we need to measure it...

Do n measurements x_i, $i = 1, 2, 3...n$. Then, Best estimate of mean μ is $\bar{x} = \frac{\sum_n x_i}{n}$, which we call sample mean. Best estimate is variance σ^2 is $s^2 = \frac{\sum_n (x_i - \bar{x})^2}{n-1}$, which we can sample variance.

A set of measurement $2, 3, 4$

Sample mean: $\bar{x} = \frac{\sum_n x_i}{n} = 3$. Sample variance: $s^2 = \frac{\sum_n (x_i - \bar{x})^2}{n-1} = 1$ What is the probability that the mean of this distribution $\mu > 4$?

$$\int_{4}^{+\infty} \frac{e^{-(x-3)^2/2s^2}}{\sqrt{2\pi}s} dx = 0.159 \tag{78}$$

$$\int_{4}^{+\infty} \frac{\sqrt{3} \times e^{-3\times(x-3)^2/2s^2}}{\sqrt{2\pi}s} dx = 0.159 \tag{79}$$

WRONG!

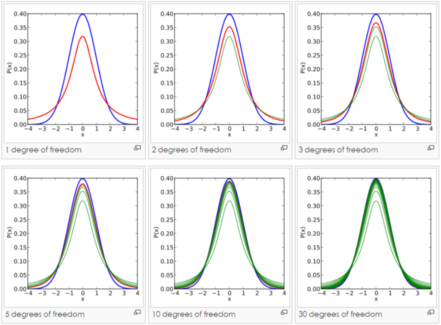
Cause s is the only estimate of σ.

Student's t distribution

$$t = \frac{\bar{x} - \mu}{x/\sqrt{n}} \tag{80}$$

$$p(t, \upsilon) = \frac{\Gamma(\frac{\upsilon+1}{2})}{\sqrt{\upsilon\pi}\Gamma\upsilon/2}(1 + \frac{t^2}{\upsilon})^{-(\upsilon+1)/2} \tag{81}$$

Density of the *t*-distribution (red) for 1, 2, 3, 5, 10, and 30 degrees of freedom compared to the standard normal distribution (blue).
Previous plots shown in green.



2.4 Robust estimation

What about the sample mean $\bar{x}$? Is it the best estimator of the central moment μ? It is the best if we assume the probability distribution is a Gaussian. (minimum χ^2). However it is NOT when there are large fluctuations. Errors which are not Gaussian distributed are called "outliers".

Figure: An outlier may greatly affect the average. Credit: E. C. Sutton

2.4 Robust estimation

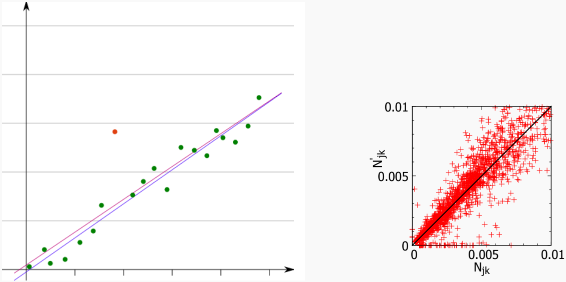


Figure: If there are large fluctuations, it will lead to improper functional parameter by minimizing χ^2. Left panel: Straight line fits with or without outlier. Right panel: Linear fitting with large group of data.

We need a penalty function, which reflects the outliers' probability.

2.4 Robust estimation

Possible choices are

$$\sum_i \left|\frac{x_i - \bar{x}}{\sigma_i}\right| \tag{82}$$

$$\sum_i log(1 + \frac{1}{2}(\frac{x_i - \bar{x}}{\sigma_i})^2) \tag{83}$$

2.5 Propagation of errors

x=f(u,v...)

$$\sigma_x^2 = \sigma_u^2(\frac{\partial x}{\partial u})^2 + \sigma_v^2(\frac{\partial x}{\partial v})^2 + 2 \times \sigma_{uv}^2(\frac{\partial x}{\partial u})(\frac{\partial x}{\partial v}) \tag{84}$$

If u and v are uncorrelated, cross term vanishes.

$$\sigma_x^2 = \sigma_u^2(\frac{\partial x}{\partial u})^2 + \sigma_v^2(\frac{\partial x}{\partial v})^2 \tag{85}$$

Show me examples

- x=au+bv

$$\sigma_x^2 = a^2\sigma_u^2 + b^2\sigma_v^2 \tag{86}$$

- x=auv

$$\sigma_x^2/x^2 = \sigma_u^2/u^2 + \sigma_v^2/v^2 \tag{87}$$

3.1 Stochastic process

Time-dependent random variable $x(\epsilon, t)$, where ϵ is the possible realization of x.

> We cannot assign $p(\epsilon)$ or $p(\epsilon)d\epsilon$!

Instead, for a fixed time t, we have $p(x, t)$

3.1.1 Stochastic process

Get me prepared The ensemble mean (mean over the ensemble ϵ)

$$\eta(t) = Ex(t) = \int_{-\infty}^{+\infty} x(\epsilon, t)p(x, t)dx \tag{88}$$

The ensemble variance

$$\sigma^2(t) = E(x(t))^2 - (Ex(t))^2 = \int_{-\infty}^{+\infty} [x(\epsilon, t)]^2 p(x, t)dx - (\eta(t))^2 \tag{89}$$

The autocorrealton

$$R(t_1, t_2) = Ex(t_1), x(t_2) = \int\int_{-\infty}^{+\infty} x_1 x_2 p(x_1, x_2, t_1, t_2) dx_1 dx_2 \tag{90}$$

The autocovariance

$$C(t_1, t_2) = E[x(t_1) - \eta_1][x(t_2) - \eta_2] \tag{91}$$

$$= \int\int_{-\infty}^{+\infty} (x_1 - \eta_1)(x_2 - \eta_2) p(x_1, x_2, t_1, t_2) dx_1 dx_2 \tag{92}$$

$$= R(t_1, t_2) - \eta(t_1)\eta(t_2) \tag{93}$$

$$\sigma^2(t) = C(t, t) = R(t, t) - \eta^2(t) \tag{94}$$

3.1.2 Stationary Process p(x)

$$R(t_1, t_2) = R(\tau) \tag{95}$$

Even function!
R(0) corresponds the average power of the process;
C(t,t) corresponds the covariance.

3.2 Poisson random process

Recall the realization of Poisson process

$\tau > 0$

$$R(\tau) = c_1 \tag{96}$$

$\tau = 0$

$$R(\tau) = c_1 + c_2\delta(\tau) \tag{97}$$

3.2 Poisson random process

Autocorrelation of poisson random process to be taken in the limit $\delta t \to 0$
SHOW ME AN EXAMPLE!

$$R(t_1, t_2) = \int\int_{-\infty}^{+\infty} x_1 x_2 p(x_1, x_2, t_1, t_2) dx_1 dx_2 \tag{98}$$

$\tau > 0$

$$R = \Sigma\Sigma(\Delta t)^{-1}(\Delta t)^{-1}(r\Delta t)^2 = r^2 \tag{99}$$

$\tau = 0$

$$R = \Sigma\Sigma(\Delta t)^{-1}(\Delta t)^{-1}(r\Delta t) = r(\Delta t)^{-1} \tag{100}$$

Finally in the limit of $\Delta t \to 0$,

$$R(\tau) = r^2 + r\delta(\tau) \tag{101}$$

$$S(\nu) = r^2\delta\nu + r \tag{102}$$

White noise

$$S(\nu) = r^2\delta\nu + r \tag{103}$$

The spectrum of a Poisson random noise (short noise) is white (flat) except $\nu = 0$.
Cosmic noise...

Help focus? Yes it does to me!
Help sleep? Yes it does to me!

Filtering

Let's skip something!

$$y(f) = x(f)h(f) \tag{104}$$

For example,
low pass filter $h(f) = \Pi(\frac{f}{2f_c})$

Noise

Blackbody radiation intensity

$$I_\nu = \frac{2h\nu^3}{c^2(e^{h\nu/kT} - 1)} \tag{105}$$

$$dE = I_\nu dA dt d\Omega d\nu \tag{106}$$

Noise

Consider etendue of $A\Omega = \lambda^2$ and a single polarization equivalent to considering a single mode, the mean power is

$$< P(\nu) >= \frac{h\nu}{e^{h\nu/kT} - 1} \tag{107}$$

,energy per photon times the photon occupation number.

$$< [P(\nu)- < P(\nu) >]^2 >=< P > h\nu[1 + \frac{1}{e^{h\nu/kT} - 1}] \tag{108}$$

Noise

Photon noise $h\nu >> kT$

$$< \Delta P^2 >=< p > h\nu \tag{109}$$

fluctuation

$$< \Delta n^2 >=< n > \tag{110}$$

Thermal noise $h\nu << kT$

$$< \Delta P^2 >=< p > kT =< p >^2 \tag{111}$$

fluctuation

$$< \Delta n^2 >=< n >^2 \tag{112}$$

The blackbody power spectral (power per mode) of thermal noise is $P(\nu) \approx kT$ In real life, an RC filter

$$h(\nu) = \frac{1}{1 + (\frac{\nu}{\nu_c})^2} \tag{113}$$

$$\sigma^2 = kT\pi\nu_c \tag{114}$$

Johnson noise...

Thank you!