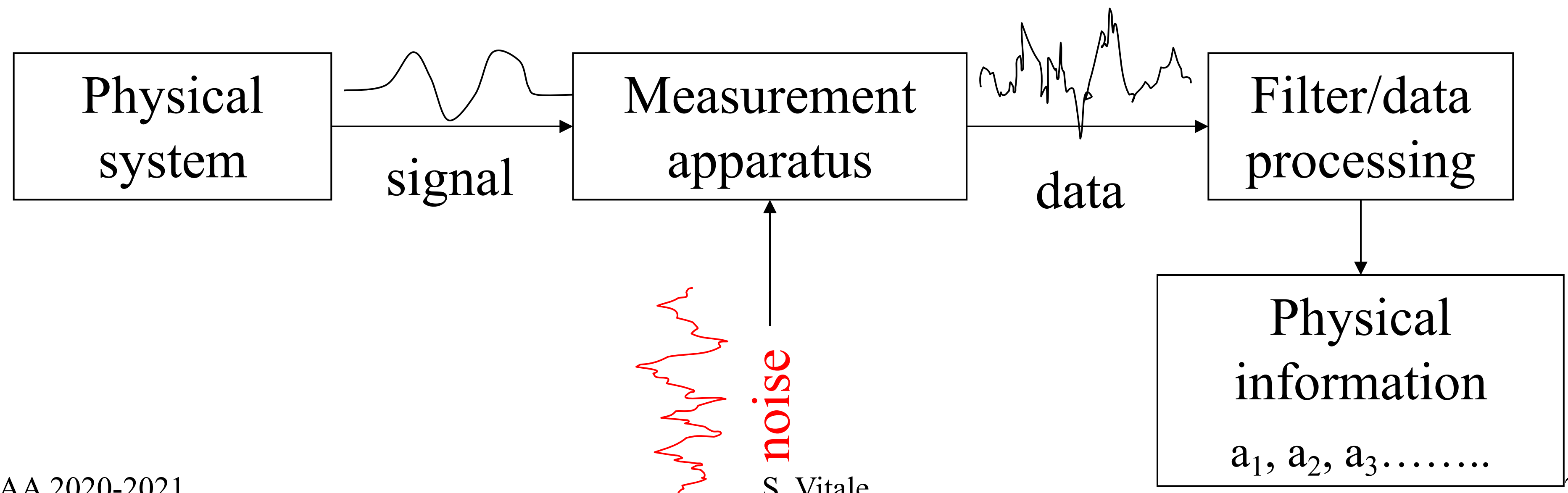
Experimental Methods
Lecture 13

October 19th, 2020

Noise

Noise in physical experiments is described as a random signal x(t):

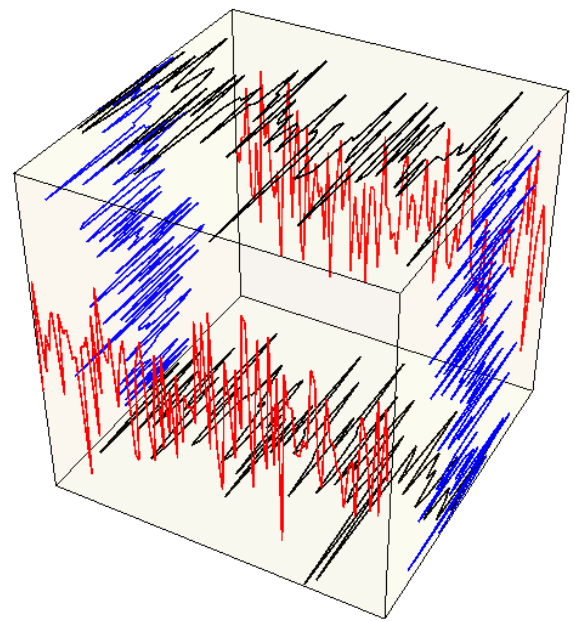
Independent repetitions of the same experiment produce different functions of time x(t)



A stochastic process

- At each outcome of an experiment I get a function of time x(t)

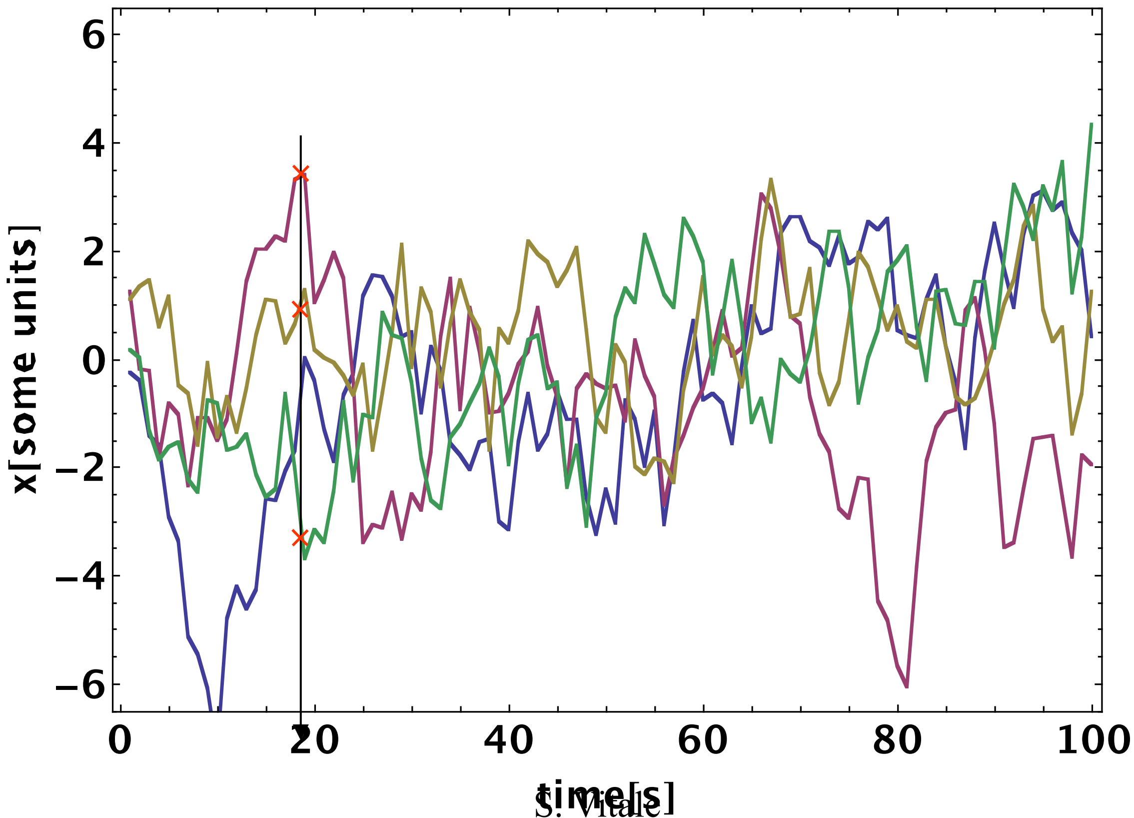
dice



Stochastic dice

A stochastic process

- For each outcome a function of time x(t)
- For each time t: a random variable



A stochastic process

- As $x(t_1)$, $x(t_2)$, $x(t_3)$….. are all different rv, a stochastic process defines infinite many rv.
- Each rv will have its own density function

$$f_{x(t)}(\chi)$$

- With mean value

$$\eta(t) \equiv \langle x(t) \rangle = \int_{-\infty}^{\infty} \chi f_{x(t)}(\chi) d\chi$$

- This is called the mean value of the process and is in general a function of time.
- Same for the variance of the process

$$\sigma^2(t) \equiv \langle x^2(t) \rangle - \eta(t)^2 = \int_{-\infty}^{\infty} \chi^2 f_{x(t)}(\chi) d\chi - \eta(t)^2$$

Multiple random variables

- The statistical properties of multiple random variables like the $x(t_1)$, $x(t_2)$, $x(t_3)$…… are described by joint probabilities. Take two rv x and y. Their joint probability density $f_{x,y}$ is defined by:

$$P\left\{x_o \leq x \leq x_1 \text{ and } y_o \leq y \leq y_1\right\} = \int_{x_o}^{x_1}\int_{y_o}^{y_1} f_{x,y}\left(\chi,\psi\right)d\chi\, d\psi$$

- Also this density needs to be normalized

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{x,y}\left(\chi,\psi\right)d\chi\, d\psi = 1$$

- The relation between the joint density $f_{x,y}$ and the marginal densities f_x and f_y is

$$\int_{x_o}^{x_1} f_x\left(\chi\right)d\chi = P\left\{x_o \leq x \leq x_1\right\} =$$

$$= P\left\{x_o \leq x \leq x_1 \text{ and} -\infty \leq y \leq \infty\right\} = \int_{x_o}^{x_1}\int_{-\infty}^{\infty} f_{x,y}\left(\chi,\psi\right)d\chi\, d\psi$$

- Then: $f_x\left(\chi\right) = \int_{-\infty}^{\infty} f_{x,y}\left(\chi,\psi\right)d\psi \qquad f_y\left(\psi\right) = \int_{-\infty}^{\infty} f_{x,y}\left(\chi,\psi\right)d\chi$

Multiple random variables

- Moments $\mu'_{n,m} = \langle x^n y^m \rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi^n \psi^m f_{x,y}(\chi,\psi)\, d\chi\, d\psi$
- It follows $\mu'_{n,0} = \langle x^n \rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi^n f_{x,y}(\chi,\psi)\, d\chi\, d\psi = \int_{-\infty}^{\infty} \chi^n f_x(\chi)\, d\chi = \mu'_n$
- Central moments $\mu_{m,n} = \left\langle \left(x - \langle x \rangle\right)^n \left(y - \langle y \rangle\right)^m \right\rangle$
- Specially important: the covariance $\sigma_{x,y} = \mu_{1,1} = \left\langle \left(x - \langle x \rangle\right)\left(y - \langle y \rangle\right) \right\rangle$
- Independent variables. If for any choice of x_o, x_1, y_o, and y_1

$$P\left\{x_o \le x \le x_1 \text{ and } y_o \le y \le y_1\right\} = P\left\{x_o \le x \le x_1\right\} \times P\left\{y_o \le y \le y_1\right\}$$

- i.e. if $x \in [x_1, x_2]$ and $y \in [y_1, y_2]$ are independent events x and y are independent rv. It follows

$$\int_{x_o}^{x_1}\int_{y_o}^{y_1} f_{x,y}(\chi,\psi)\, d\chi\, d\psi = \int_{x_o}^{x_1} f_x(\chi)\, d\chi \int_{y_o}^{y_1} f_y(\psi)\, d\psi \rightarrow f_{x,y}(\chi,\psi) = f_x(\chi) f_y(\psi)$$

Multiple random variables

- For independent random variables $f_{x,y}(\chi,\psi)=f_x(\chi)f_y(\psi)$

- Joint moments

$$\mu'_{n,m}=\langle x^n y^m\rangle=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\chi^n\psi^m f_{x,y}(\chi,\psi)\,d\chi\,d\psi$$
$$=\int_{-\infty}^{\infty}\chi^n f_x(\chi)\,d\chi\int_{-\infty}^{\infty}\psi^m f_y(\psi)\,d\psi=\langle x^n\rangle\langle y^m\rangle$$

- In particular you can calculate that the covariance

$$\sigma_{x,y}=\mu_{1,1}=\langle(x-\langle x\rangle)(y-\langle y\rangle)\rangle=\langle x-\langle x\rangle\rangle\langle y-\langle y\rangle\rangle=0$$

- Independent variables are "uncorrelated"
- The converse is not necessarily true.

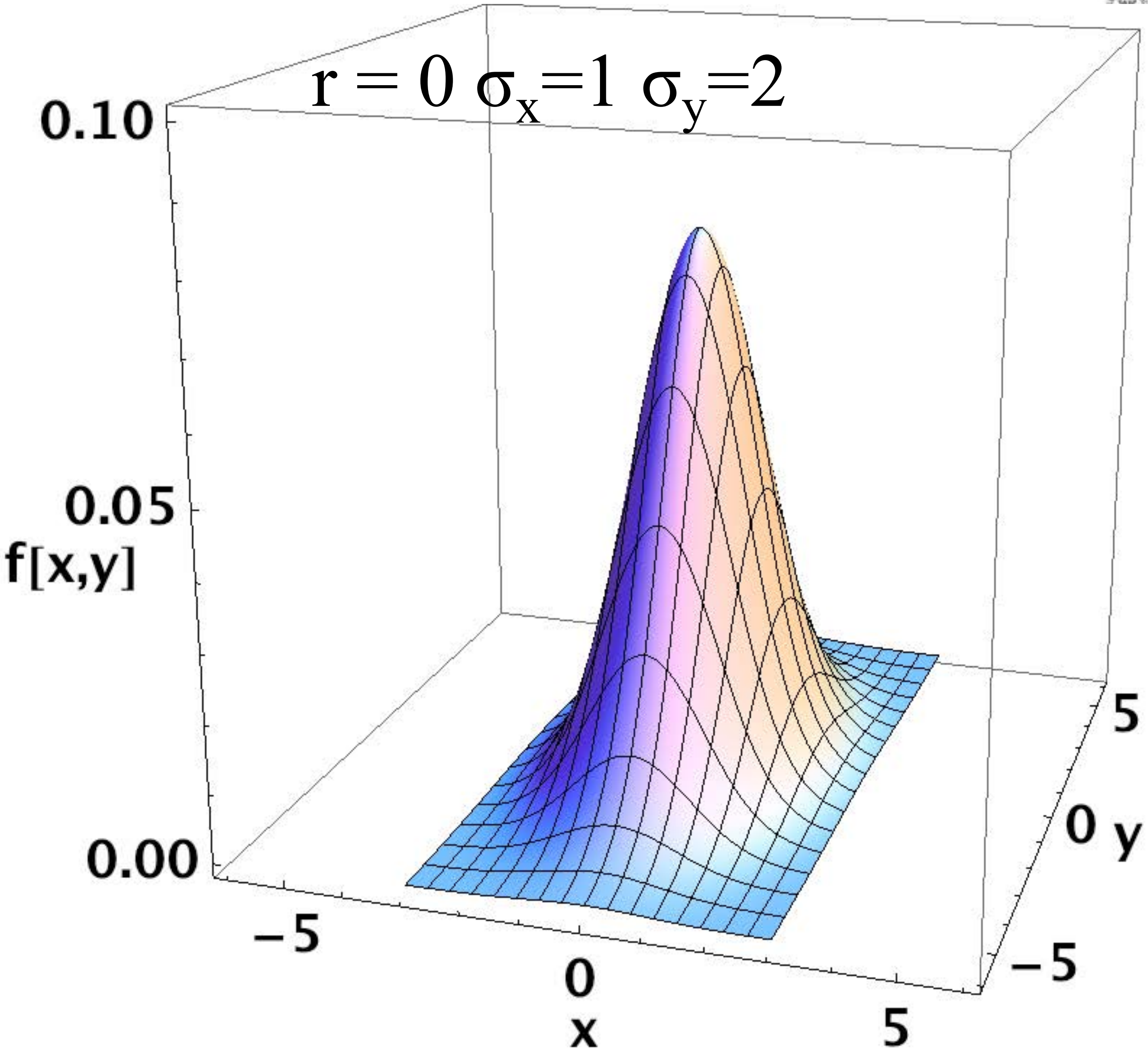
One example: a bidimensional Gaussian

- The Joint Probability Density Function

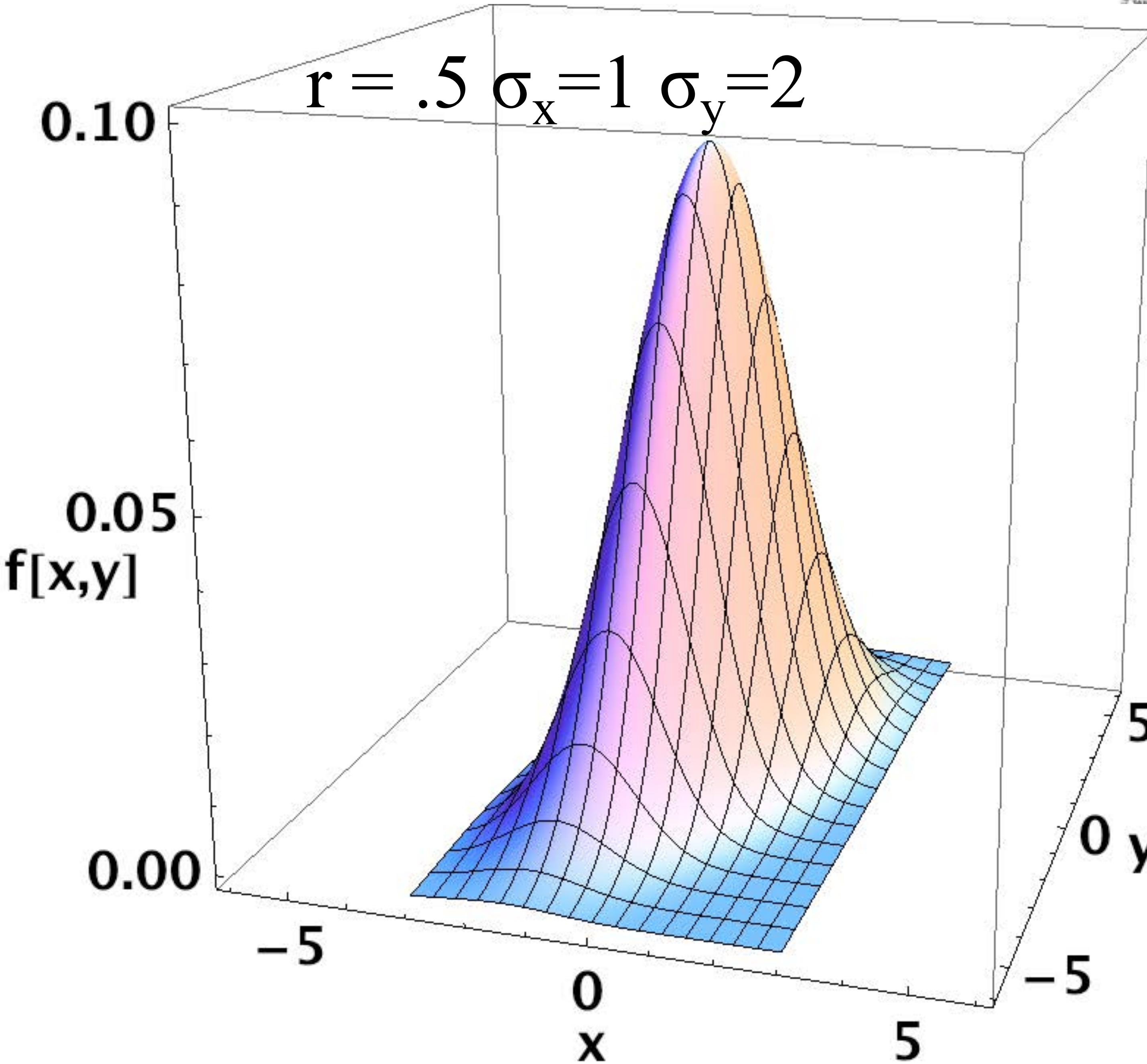
$$f[x,y]=\frac{e^{-\frac{1}{2\left(1-r^{2}\right)}\left(\frac{\left(y-y_{o}\right)^{2}}{\sigma_{y}^{2}}-\frac{2\,r}{\sigma_{x}\,\sigma_{y}}\left(x-x_{o}\right)\left(y-y_{o}\right)+\frac{\left(x-x_{o}\right)^{2}}{\sigma_{x}^{2}}\right)}}{2\,\pi\sqrt{1-r^{2}}\;\sigma_{x}\,\sigma_{y}}$$

- For r=0 x and y are independent

Wolfram Mathematica 7


r = 0 σx=1 σy=2
0.10
0.05
0.00
f[x,y]
−5
0
5
x
5
0
−5
y

$r = .5\ \sigma_x=1\ \sigma_y=2$



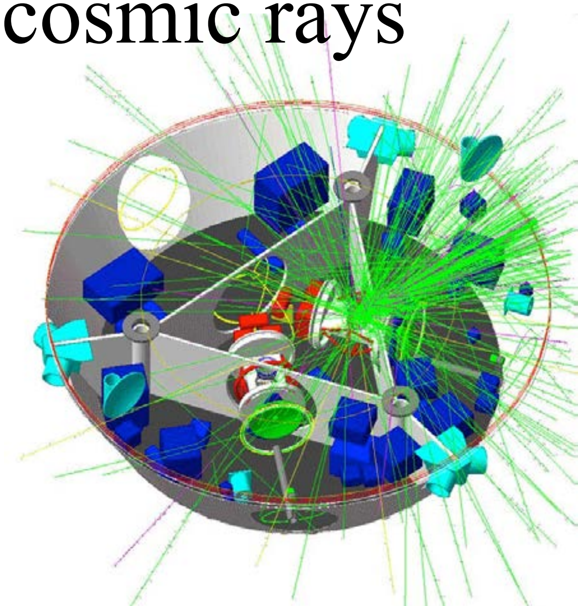
Auto-correlation of a stochastic process

- Key high order moments of a stochastic processes are
  - The auto-correlation
    $$R(t,t')=\langle x(t)x(t')\rangle \quad =\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\chi\psi f_{x(t),x(t')}(\chi,\psi)d\chi\, d\psi$$
  - And the auto-covariance
    $$C(t,t')=\langle [x(t)-\eta(t)][x(t')-\eta(t')]\rangle = R(t,t')-\eta(t)\eta(t')$$
  - Notice that a stochastic process can always be written as its mean value, a deterministic function of time, plus the zero-mean process
    $$\tilde{x}(t)=x(t)-\eta(t)\rightarrow x(t)=\eta(t)+\tilde{x}(t)$$
  - The auto-covariance is the autocorrelation of the "noisy" part of the process
  - Auto-covariance expresses the statistical memory of the process. To be further discussed

An important example: Poisson process and shot noise

- A stochastic process based model, for all phenomena where carriers arrive at random at a collector/detector:
  - Photons in a laser beam;
  - Charge carriers across a junction;
  - Particles from a radioactive source;

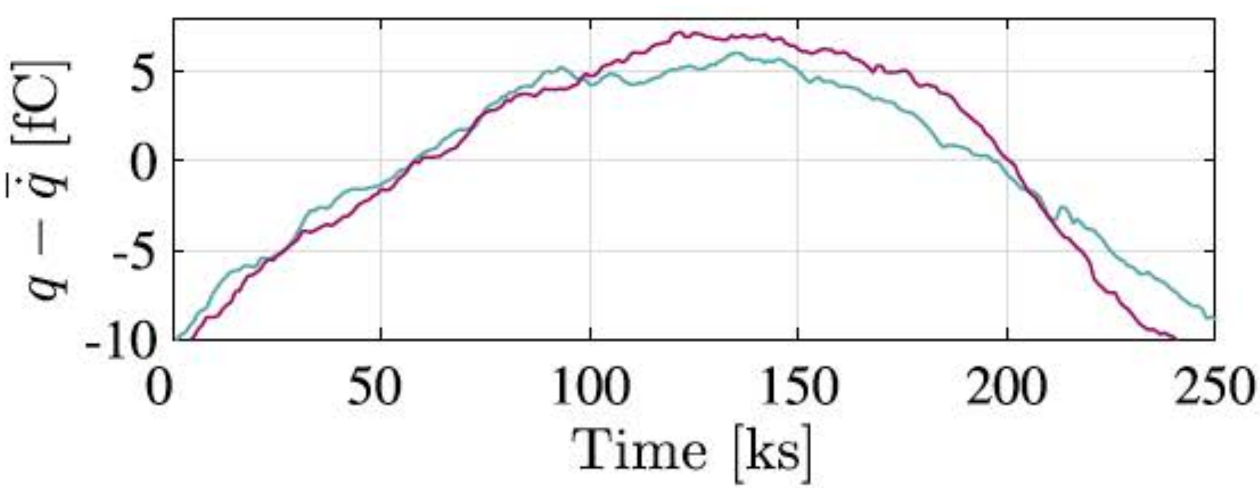
One example: random arrival of charge on LISA Pathfinder test-masses because of cosmic rays

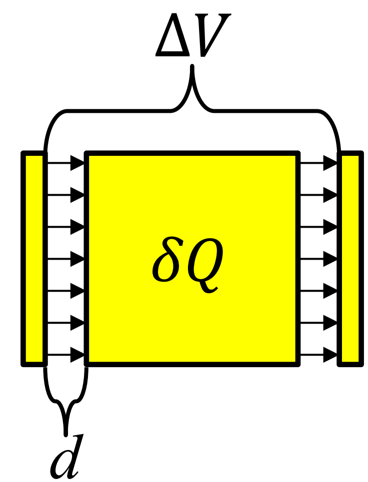


- Cosmic rays keep charging up the test-mass
- Random charge δQ produces force noise

$$\delta F_Q = \frac{\delta Q \Delta V}{d}$$

PRL **118**, 171101 (2017) PHYSICAL REVIEW LETTERS week ending 28 APRIL 2017





Basics of the Poisson process

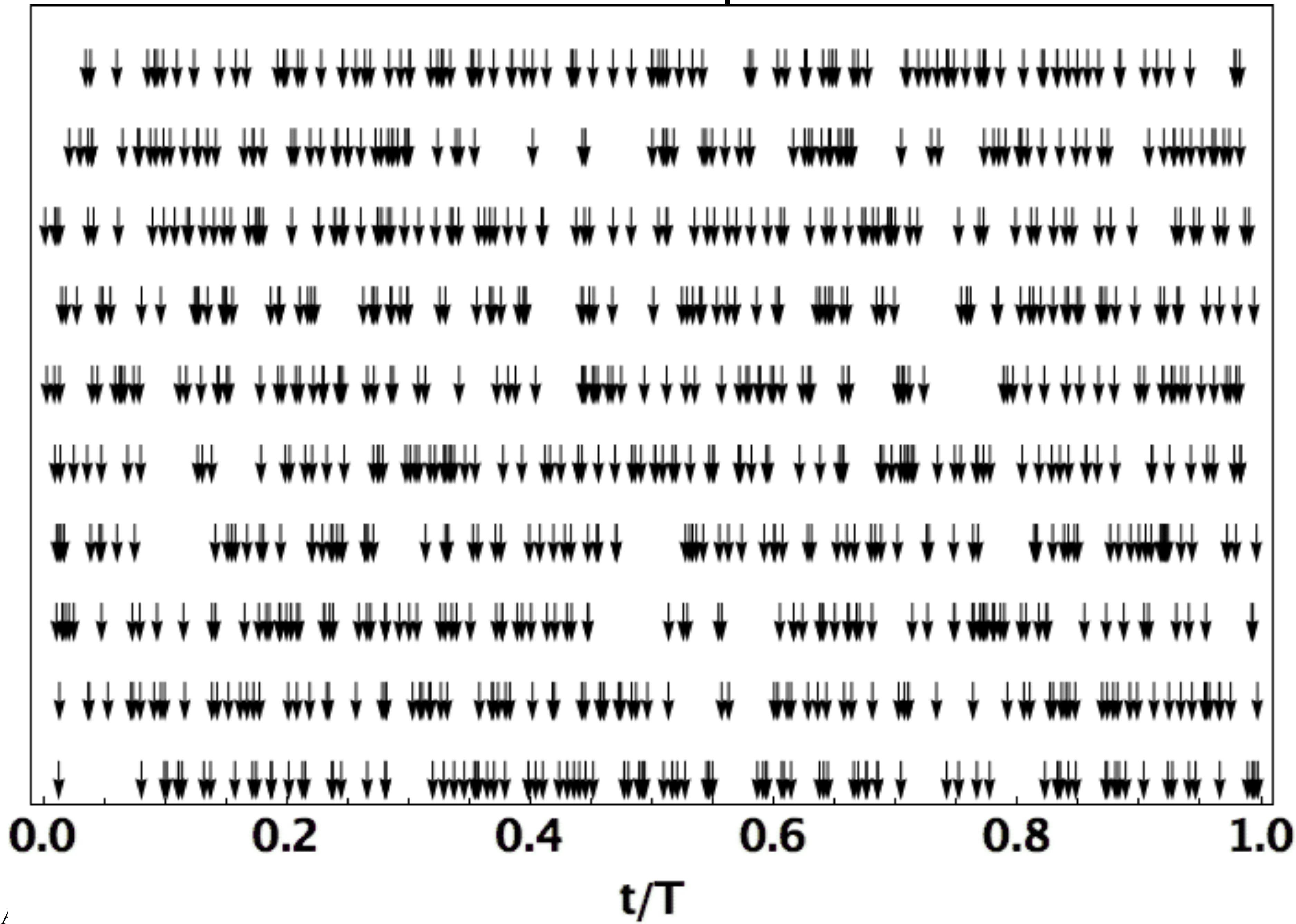
- Take a time interval with $0 \le t \le T$
- Pick N time instants at random, that is:
  - Instant t_i and instant t_k, for $i \neq k$, and $1 \le i$ and $k \le N$, are independently picked.
  - The probability density function for the random variable t_i is given by

$$f_{t_i}(x) = \frac{1}{T}\Pi\left(\frac{t}{T} - \frac{1}{2}\right)$$

- Conditions above mean that for $i \neq k$, $1 \le i$ and $k \le N$, $t_1 \le t_2$, and $t_3 \le t_4$:

$$P\left\{t_1 \le t_i \le t_2\right\} = \frac{t_2 - t_1}{T} \qquad P\left\{t_1 \le t_i \le t_2 \;\&\; t_3 \le t_j \le t_4\right\} = \frac{t_2 - t_1}{T}\frac{t_4 - t_3}{T}$$

Numerical example N=100



Probability of finding k points in a given interval

- Probability for one point of falling between t_1 and t_2 ($t_1 < t_2$)

$$P\{t_1 \le t \le t_2\} = \frac{t_2 - t_1}{T} \equiv p$$

- The events $t \in [t_1,t_2]$ and $t \notin [t_1,t_2]$ are mutually exclusive with probability p and 1-p respectively. Thus the probability of k points ($k \in$ Integers, $0 \le k \le N$) falling between t_1 and t_2 ($t_1<t_2$) (Binomial distribution) is

$$P\left(k, t \in \left[t_1, t_2\right]\right) = \frac{N!}{k!\left(N-k\right)!} p^k \left(1-p\right)^{N-k}$$

- Note: k is a random variable. Its probability density is

$$f_k\left(x\right) = \sum_{k=0}^{N} \frac{N!}{k!\left(N-k\right)!} p^k \left(1-p\right)^{N-k} \delta\left(x-k\right)$$

The Poisson limit

- Binomial distribution

$$P\left(k, t \in \left[t_1, t_2\right]\right) = \frac{N!}{k!\left(N-k\right)!} p^k \left(1-p\right)^{N-k}$$

- Now try the following limit: $N \rightarrow \infty$, $T \rightarrow \infty$, $N/T \rightarrow \lambda$. Then

$$p = \left(t_2 - t_1\right)/T \rightarrow 0 \qquad Np = N\left(t_2 - t_1\right)/T \rightarrow \lambda\left(t_2 - t_1\right)$$

- In the next slide it is shown that in this limit (Poisson)

$$P\left(k, t \in \left[t_1, t_2\right]\right) = \frac{\left[\lambda\left(t_2 - t_1\right)\right]^k}{k!} e^{-\lambda\left(t_2 - t_1\right)}$$

- You can check that

$$P\left\{0 \leq k \leq \infty\right\} = e^{-\lambda\left(t_2 - t_1\right)} \sum_{k=0}^{\infty} \frac{\left[\lambda\left(t_2 - t_1\right)\right]^k}{k!} = e^{-\lambda\left(t_2 - t_1\right)} e^{+\lambda\left(t_2 - t_1\right)} = 1$$

The Poisson Limit

Let's now calculate the limit for $N \to \infty$, $T \to \infty$, $N/T \to \lambda$ of the binomial formula

$$P_k = \frac{N!}{k!\,(N-k)!}\, p^k (1-p)^{N-k}$$

First rewrite probability as $p = (N\ (t_2 - t_1)/T)/N = \lambda(t_2 - t_1)/N$ then:

$$P_k = \frac{N!}{k!\,(N-k)!}\left[\frac{\lambda(t_2-t_1)}{N}\right]^k \left[1-\frac{\lambda(t_2-t_1)}{N}\right]^{N-k} =$$

$$= \frac{N!}{N^k (N-k)!}\,\frac{[\lambda(t_2-t_1)]^k}{k!}\left[1-\frac{\lambda(t_2-t_1)}{N}\right]^{N-k}$$

Now consider the limits of the various parts

$$\lim_{N\to\infty}\left[1-\frac{\lambda(t_2-t_1)}{N}\right]^N = e^{-\lambda(t_2-t_1)} \qquad \lim_{N\to\infty}\left[1-\frac{\lambda(t_2-t_1)}{N}\right]^{-k} = 1$$

$$\lim_{n\to\infty}\frac{N!}{N^k (N-k)!} = \lim_{N\to\infty}\frac{N^k + O\left[N^{k-1}\right]}{N^k} = 1$$

As a consequence

$$\operatorname*{Lim}_{N\to\infty, T\to\infty, N/T\to\lambda} P_k = \frac{[\lambda(t_2-t_1)]^k}{k!}\, e^{-\lambda(t_2-t_1)}$$

The Poisson limit

- Thus in the limit: $N\to\infty$, $T\to\infty$, $N/T\to\lambda$, k has probability density:

$$f_k(x)=e^{-\lambda(t_2-t_1)}\sum_{m=0}^{\infty}\frac{\left[\lambda(t_2-t_1)\right]^m}{m!}\delta(x-m)$$

- Mean value of which is

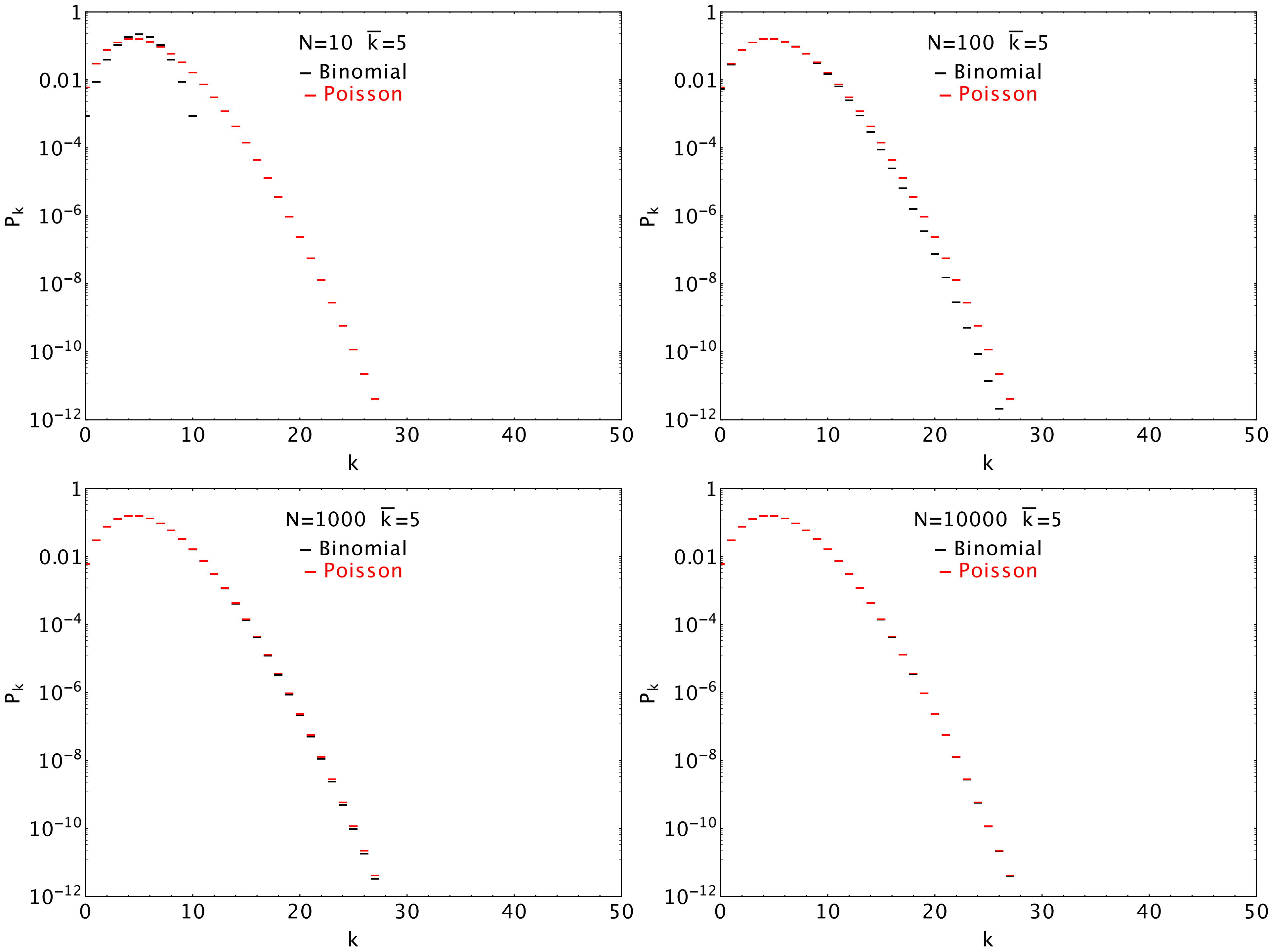
$$\bar{k}=\int_{-\infty}^{\infty}xf_k(x)dx=e^{-\lambda(t_2-t_1)}\sum_{m=0}^{\infty}m\frac{\left[\lambda(t_2-t_1)\right]^m}{m!}\ =e^{-\lambda(t_2-t_1)}\sum_{m=1}^{\infty}m\frac{\left[\lambda(t_2-t_1)\right]^m}{m!}$$

- By expanding

$$\bar{k}=e^{-\lambda(t_2-t_1)}\sum_{m=1}^{\infty}\frac{\left[\lambda(t_2-t_1)\right]^m}{(m-1)!}\ =e^{-\lambda(t_2-t_1)}\lambda(t_2-t_1)\sum_{m=1}^{\infty}\frac{\left[\lambda(t_2-t_1)\right]^{m-1}}{(m-1)!}$$

$$=e^{-\lambda(t_2-t_1)}\lambda(t_2-t_1)\sum_{n=0}^{\infty}\frac{\left[\lambda(t_2-t_1)\right]^n}{n!}=e^{-\lambda(t_2-t_1)}\lambda(t_2-t_1)e^{\lambda(t_2-t_1)}\ =\lambda(t_2-t_1)$$

- Then $\qquad P_k=e^{-\bar{k}}\dfrac{\bar{k}^m}{m!} \qquad f_k(x)=e^{-\bar{k}}\displaystyle\sum_{m=0}^{\infty}\frac{\bar{k}^m}{m!}\delta(x-m)$



N=10 $\bar{k}$=5
Binomial
Poisson
N=100 $\bar{k}$=5
Binomial
Poisson
N=1000 $\bar{k}$=5
Binomial
Poisson
N=10000 $\bar{k}$=5
Binomial
Poisson
P_k
k
1
0.01
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
0
10
20
30
40
50

The Poisson limit

- k has probability density function:

$$f_k(x) = e^{-\bar{k}} \sum_{m=0}^{\infty} \frac{\bar{k}^m}{m!} \delta(x - m)$$

- With mean value

$$\bar{k} = \lambda(t_2 - t_1)$$

- Variance. Begin by calculating

$$\langle k^2 \rangle = \int_{-\infty}^{\infty} x^2 f_k(x) dx = e^{-\bar{k}} \sum_{m=0}^{\infty} m^2 \frac{\bar{k}^m}{m!} = e^{-\bar{k}} \sum_{m=1}^{\infty} m^2 \frac{\bar{k}^m}{m!} = e^{-\bar{k}} \bar{k} \sum_{m=1}^{\infty} m \frac{\bar{k}^{m-1}}{(m-1)!}$$

$$= \bar{k}\left[e^{-\bar{k}} \sum_{n=0}^{\infty} (n+1) \frac{\bar{k}^n}{n!} \right] = \bar{k}(\bar{k}+1)$$

- Then the variance is

$$\sigma_k^2 = \langle k^2 \rangle - \langle k \rangle^2 = \bar{k}(\bar{k}+1) - \bar{k}^2 = \bar{k}$$

- And the standard deviation is $\sigma_k = \sqrt{\bar{k}}$
- Notice

$$\sigma_k / \bar{k} = 1 / \sqrt{\bar{k}}$$

Numbers of instants falling in non-overlapping time intervals are not independent rv

- Consider 2 time intervals $[t_1,t_2]$ and $[t_3,t_4]$ between 0 and T that have no common points. Take $t_2<t_3$

$$\underbrace{\overbrace{\qquad\qquad\qquad}^{\left[t_1,t_2\right]}\overbrace{\qquad\qquad\qquad}^{\left[t_3,t_4\right]}\overbrace{\qquad\qquad\qquad}^{\left[0,T\right]-\left[t_1,t_2\right]-\left[t_3,t_4\right]}}_{\left[0,T\right]}$$

- The events $t\in[t_1,t_2]$, $t\in[t_3,t_4]$, and events $t\in[0,T]-[t_1,t_2]-[t_3,t_4]$ are mutually exclusive. The probability of k_1 time-instants falling into $[t_1,t_2]$, k_2 instants falling into $[t_3,t_4]$ and $N-k_1-k_2$ falling into $[0,T]-[t_1,t_2]-[t_3,t_4]$ is given by:

$$P\left[k_1\in\left[t_1,t_2\right]\text{and}\,k_2\in\left[t_3,t_4\right]\right]=$$

$$=\frac{N!}{k_1!k_2!\left(N-k_1-k_2\right)!}\left(\frac{t_2-t_1}{T}\right)^{k_1}\left(\frac{t_4-t_3}{T}\right)^{k_2}\left(1-\frac{t_2-t_1}{T}-\frac{t_4-t_3}{T}\right)^{N-k_1-k_2}$$

Numbers of instants in non-overlapping time intervals are not independent rv

$$\underbrace{\overbrace{\qquad\qquad\qquad}^{\left[t_1,t_2\right]}\overbrace{\qquad\qquad\qquad}^{\left[t_3,t_4\right]}\overbrace{\qquad\qquad\qquad\qquad\qquad}^{\left[0,T\right]-\left[t_1,t_2\right]-\left[t_3,t_4\right]}}_{\left[0,T\right]}$$

- Notice

$$P\left[k_1 \in \left[t_1,t_2\right] \text{and}\, k_2 \in \left[t_3,t_4\right]\right]=$$

$$=\frac{N!}{k_1!k_2!\left(N-k_1-k_2\right)!}\left(\frac{t_2-t_1}{T}\right)^{k_1}\left(\frac{t_4-t_3}{T}\right)^{k_2}\left(1-\frac{t_2-t_1}{T}-\frac{t_4-t_3}{T}\right)^{N-k_1-k_2} \neq$$

$$\neq \frac{N!}{k_1!\left(N-k_1\right)!}\left(\frac{t_2-t_1}{T}\right)^{k_1}\left(1-\frac{t_2-t_1}{T}\right)^{N-k_1} \times$$

$$\times \frac{N!}{k_2!\left(N-k_2\right)!}\left(\frac{t_4-t_3}{T}\right)^{k_2}\left(1-\frac{t_4-t_3}{T}\right)^{N-k_2} = P\left[k_1 \in \left[t_1,t_2\right]\right]P\left[k_2 \in \left[t_3,t_4\right]\right]$$

Numbers of instants in non-overlapping time intervals are not independent rv

$$\underbrace{\overbrace{\qquad\qquad\qquad}^{\left[t_1,t_2\right]}\overbrace{\qquad\qquad\qquad}^{\left[t_3,t_4\right]}\overbrace{\qquad\qquad\qquad\qquad\qquad}^{\left[0,T\right]-\left[t_1,t_2\right]-\left[t_3,t_4\right]}}_{\left[0,T\right]}$$

- As

$$P\left[k_1 \in \left[t_1,t_2\right] \text{and}\, k_2 \in \left[t_3,t_4\right]\right] \neq P\left[k_1 \in \left[t_1,t_2\right]\right] P\left[k_2 \in \left[t_3,t_4\right]\right]$$

- k_1 and k_2 are not independent.
- They only *become independent* in the now famous limit $N\rightarrow\infty$, $T\rightarrow\infty$, $N/T\rightarrow\lambda$ (see next 2 pages)

The limit $N\to\infty$, $T\to\infty$, $N/T\to\lambda$

Start with the formula:

$$\frac{N!}{k_1!\,k_2!\,(N-k_1-k_2)!}\left(\frac{t_2-t_1}{T}\right)^{k_1}\left(\frac{t_4-t_3}{T}\right)^{k_2}\left(1-\frac{t_2-t_1}{T}-\frac{t_4-t_3}{T}\right)^{N-k_1-k_2}$$

Use the same trick as before $(t_2-t_1)/\mathrm{T}=[(\mathrm{N/T})(t_2-t_1)/\mathrm{N}]=\lambda(t_2-t_1)$ and same for t_3 and t_4. Then

$$=\frac{N!}{k_1!\,k_2!\,(N-k_1-k_2)!}\,(\lambda(t_2-t_1)/N)^{k_1}\,(\lambda(t_4-t_3)/N)^{k_2}\times$$
$$\times(1-\lambda(t_2-t_1)/N-\lambda(t_4-t_3)/N)^{N-k_1-k_2}$$

You can proceed as before

$$=\frac{N!}{N^{k_1}\,N^{k_2}(N-k_1-k_2)!}\left[(\lambda(t_2-t_1))^{k_1}/k_1!\right]\left[(\lambda(t_4-t_3))^{k_2}/k_2!\right]\times$$
$$\times(1-\lambda(t_2-t_1)/N-\lambda(t_4-t_3)/N)^{N}\,(1-\lambda(t_2-t_1)/N-\lambda(t_4-t_3)/N)^{-k_1-k_2}$$

The limit N→∞, T→∞, N/T→λ

Now take the limits of the various parts

$$\frac{N!}{N^{k_1} N^{k_2}(N - k_1 - k_2)!} \rightarrow \frac{N^{k_1+k_2} + O\left[N^{k_1+k_2-1}\right]}{N^{k_1+k_2}} \rightarrow 1$$

$$(1 - \lambda(t_2 - t_1)/N - \lambda(t_4 - t_3)/N)^{-k_1-k_2} \rightarrow 1$$

$$(1 - \lambda(t_2 - t_1)/N - \lambda(t_4 - t_3)/N)^{N} \rightarrow e^{-\lambda(t_2-t_1)-\lambda(t_4-t_3)}$$

then

$$= \frac{N!}{N^{k_1} N^{k_2}(N-k_1-k_2)!}\left[(\lambda(t_2 - t_1))^{k_1}/k_1!\right]\left[(\lambda(t_4 - t_3))^{k_2}/k_2!\right]\times$$

$$\times(1 - \lambda(t_2 - t_1)/N - \lambda(t_4 - t_3)/N)^{N}(1 - \lambda(t_2 - t_1)/N - \lambda(t_4 - t_3)/N)^{-k_1-k_2}$$

$$\rightarrow \left[(\lambda(t_2 - t_1))^{k_1}/k_1!\right]\left[(\lambda(t_4 - t_3))^{k_2}/k_2!\right] e^{-\lambda(t_2-t_1)} e^{-\lambda(t_4-t_3)}$$

That is

$$P[k_1 \in [t_1, t_2]\ and\ k_2 \in [t_3, t_4]] = P[k_1 \in [t_1, t_2]]\, P[k_2 \in [t_3, t_4]]$$

Basics of the Poisson process

- In the limit $N\rightarrow\infty$, $T\rightarrow\infty$, $N/T\rightarrow\lambda$ k_1 and k_2, numbers of points within disjoint intervals $[t_1,t_2]$ and $[t_3,t_4]$, are independent random variable

$$\underbrace{\overbrace{\qquad\qquad}^{\left[t_1,t_2\right]}\overbrace{\qquad\qquad}^{\left[t_3,t_4\right]}\overbrace{\qquad\qquad\qquad\qquad}^{\left[0,T\right]-\left[t_1,t_2\right]-\left[t_3,t_4\right]}}_{\left[0,T\right]}$$

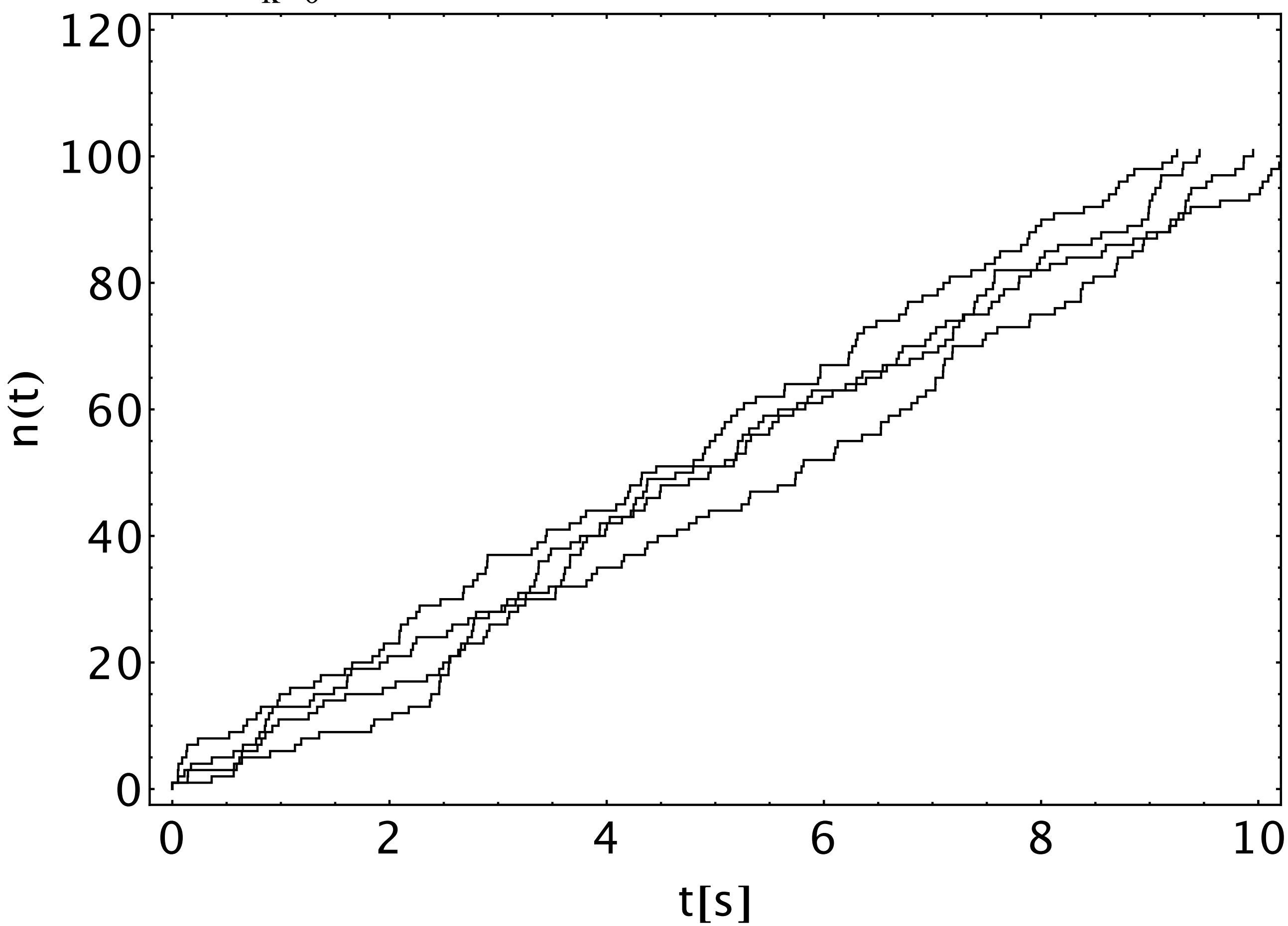
$$f_{k_1,k_2}\left(x,y\right)=\left\{e^{-\bar{k}_1}\sum_{m=0}^{\infty}\frac{\bar{k}_1^m}{m!}\delta\left(x-m\right)\right\}\left\{e^{-\bar{k}_2}\sum_{m=0}^{\infty}\frac{\bar{k}_2^m}{m!}\delta\left(x-m\right)\right\}=f_{k_1}\left(x\right)f_{k_2}\left(y\right)$$

The Poisson stochastic process

- Assume that the time instants above represent the time of arrival of some carrier at a given point. We define the process n(t) as the number of carriers that have already arrived at time t:

$$n(t) = \sum_{k=0}^{\infty} \Theta(t - t_k)$$

Example with λ=10 s^{-1}. Different lines corresponds to different outcomes of the experiment



Geiger counting