Agenda: Information, Probability, Statistical Entropy

- Information and probability
  - simple combinatorics,.
  - Probability distributions, joint probabilities,
  - Stochastic variables, correlations
  - Statistical entropy
  - Monte Carlo simulations
- Partition of probability
- Phase space evolution (Eta Theorem)
- Partition functions for different degrees of freedom
- Gibbs stability criteria, equilibrium

1

Reading Assignments
Weeks 4 & 5
LN III.1- III.6:

Kondepudi Ch. 20
Additional Material

McQuarrie & Simon
Ch. 3 & 4

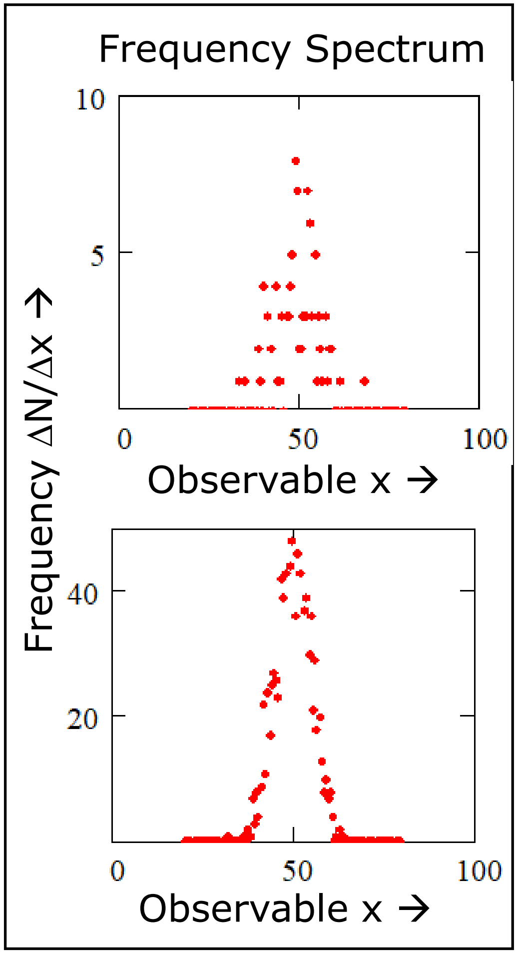
Math Chapters
MC B, E

Stochastic Variables (Observables)

Sources of stochastic observables *x* in physical sciences:

1) Quantal phenomena governed by quantal wave functions and inherent statistics.

2) Detection of processes with imperfect coverage (efficiency $\varepsilon < 1$) and finite resolution distributing sharp observable x_0 over a range in x.



Stochastic observables *x* have a range of values with frequencies determined by (normalized) probability distribution P(x). Characterize *P* by set of **moments** of *P*

$$\langle x^n \rangle = \int x^n \cdot P(x)\,dx; \quad n = 0, 1, 2, \ldots$$

with the **normalization** $\langle x^0 \rangle = 1$. First moment of P:

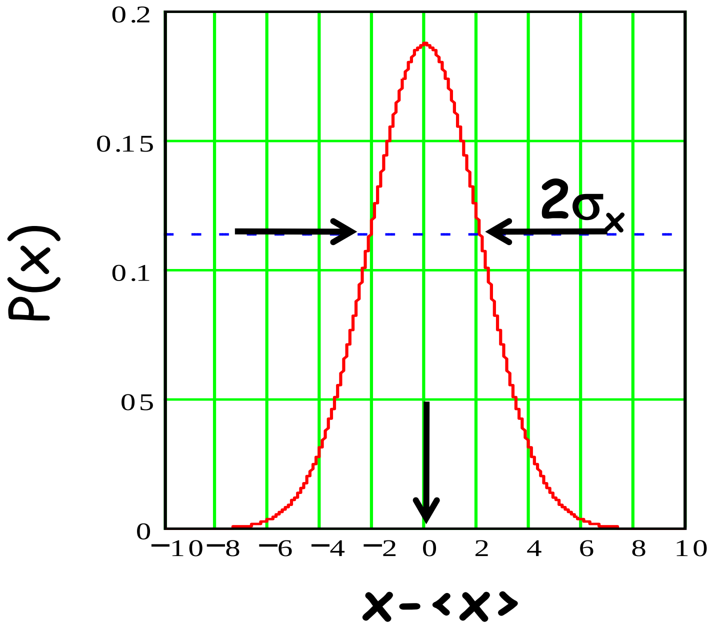
$$E(x) = \langle x \rangle = \int x \cdot P(x)\,dx$$

second **central moment** = "variance" of P(x)

$$\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2$$

Normal Distribution of a Stochastic (Random) Variable

Expected for many repeat measurements: Distribution over bins x=const.



$$P(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \cdot \exp\left\{-\frac{\left(x-\langle x\rangle\right)^2}{2\sigma_X^2}\right\}$$

$$\Gamma_{FWHM} = 2\sigma_x \cdot \sqrt{2\ln 2} = 2.35 \cdot \sigma_x$$

σ_x *is NOT = uncertainty of* $\langle x\rangle$!

Normalized (cumulative) probability

$$P(x < x_1) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \cdot \int_{-\infty}^{x_1} dx \exp\left\{-\frac{\left(x-\langle x\rangle\right)^2}{2\sigma_X^2}\right\}$$

W. Udo Schröder 2023

Fundamental: Binomial *A Priori* Distribution

Integer random variable m = number of specific binary (Yes/No) events, out of N total. Example: decay of m (from a sample of N) radioactive atoms, or m *"head" up* (out of *N flips of a coin*).

p = probability for a (one) success (coin head is up, decay of 1 exc. atom)

Choose an arbitrary sample of m trials out of N total trials (possibilities)

p^m = **probability for at least** m successes (no conditions)

$(1-p)^{N-m}$ = **probability for *N-m* failures** (coin tail, survivals)

Probability for exactly m successes out of a total of N trials

$$P(m) \propto p^m \cdot (1-p)^{N-m}$$

ways m events be 'chosen' out of N identical ? → Binomial coefficient

$$\binom{N}{m} = \frac{N!}{m!(N-m)!} = \frac{(N-m+1)\cdots N}{1\cdots m}$$

Total probability (expected success rate) for any sample of m identical events:

$$P_{binomial}(m) = \binom{N}{m} \cdot p^m \cdot (1-p)^{N-m}$$

Binomial Distribution

Probability for *m* out of N identical events

$$P_{binomial}(m) = \binom{N}{m} \cdot p^m \cdot (1-p)^{N-m}$$

Binomial coefficient

$$\binom{N}{m} = \frac{N!}{m!(N-m)!} = \frac{(N-m+1)\cdots N}{1\cdots m}$$

Total probability for any *m out* of *N* identical events:

$$\Rightarrow \textit{Proper normalization}: \sum_{m=1}^{N} P(m) = \sum_{m=1}^{N} \binom{N}{m} \cdot p^m \cdot (1-p)^{N-m} = 1$$

$$\textit{Mean value } \mu = \langle m \rangle = N \cdot p; \qquad \textit{Variance } \sigma_m^2 = N \cdot p \cdot (1-p)$$

$$\langle m^\nu \rangle = \sum_{m=0}^{N} m^\nu \cdot P(m) = \sum_{m=0}^{N} m^\nu \cdot \binom{N}{m} p^m (1-p)^{N-m}$$

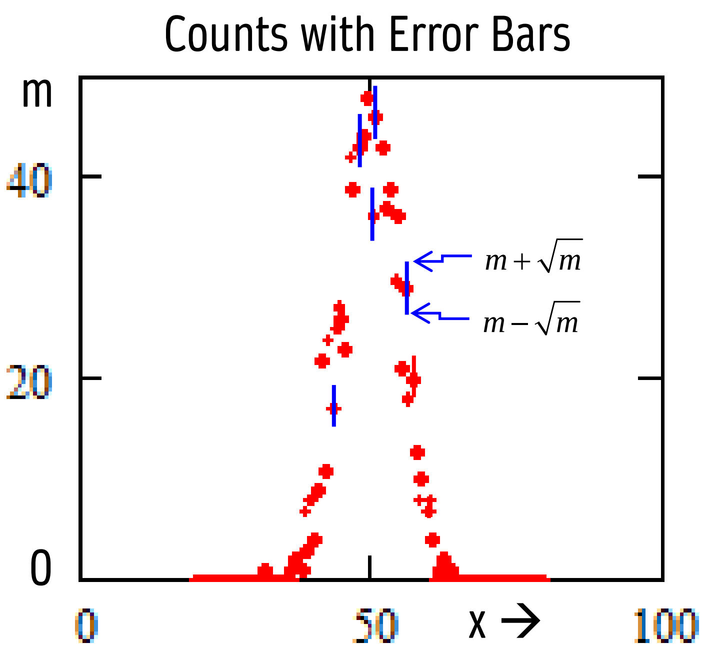
Poisson Probability Distribution

Limit of binomial distribution $Lim_{p\to 0, N\to\infty} P_{binomial}(N,m) = P_{Poisson}(\mu,m)$

Probability for observing m events when average is $\langle m\rangle = \mu$

$$P_{Poisson}(\mu,m) = \frac{\mu^m \cdot e^{-\mu}}{m!}$$

$\mu = \langle m\rangle = N\cdot p$ for $N \gg 1$
and $\sigma^2 = \mu$



For radioactive decays $[\Delta t^{-1}] \to p = \frac{Activity}{N(\#atoms)}$

$p \ll 1 \to \sigma_m^2 \approx \langle m\rangle \; \#counts$

Observe *m* counts (events) →
→ statistical uncertainty is $\pm\sigma_m = \pm\sqrt{m}$

Radioactive Decay as Poisson Process

Useful when only a mean rate is known for sample survival or decay.

^{137}Cs = unstable isotope, decays with
$t_{1/2}$ = 27 years → ***p*** = ln2/27 = 0.026/a = $\mathbf{8.2 \cdot 10^{-10} s^{-1}}$ → small

Sample of 1 μg: N = 10^{15} particles (= # trials for decay)

How many will decay?

$$\mu = N \cdot p = \mathbf{8.2 \cdot 10^{+5}\, s^{-1}}$$

Count rate estimate dN/dt = ($\mathbf{8.2 \cdot 10^{+5} \pm 905}$) $\mathbf{s^{-1}}$

fluctuation

Probability for m decays *P* (μ,m) =

$$P_{Poisson}(\mu, m) = \frac{\mu^m \cdot e^{-\mu}}{m!} = \frac{(8.52 \cdot 10^5)^m \cdot e^{-8.52 \cdot 10^5}}{m!}$$

Poisson Probability Distribution

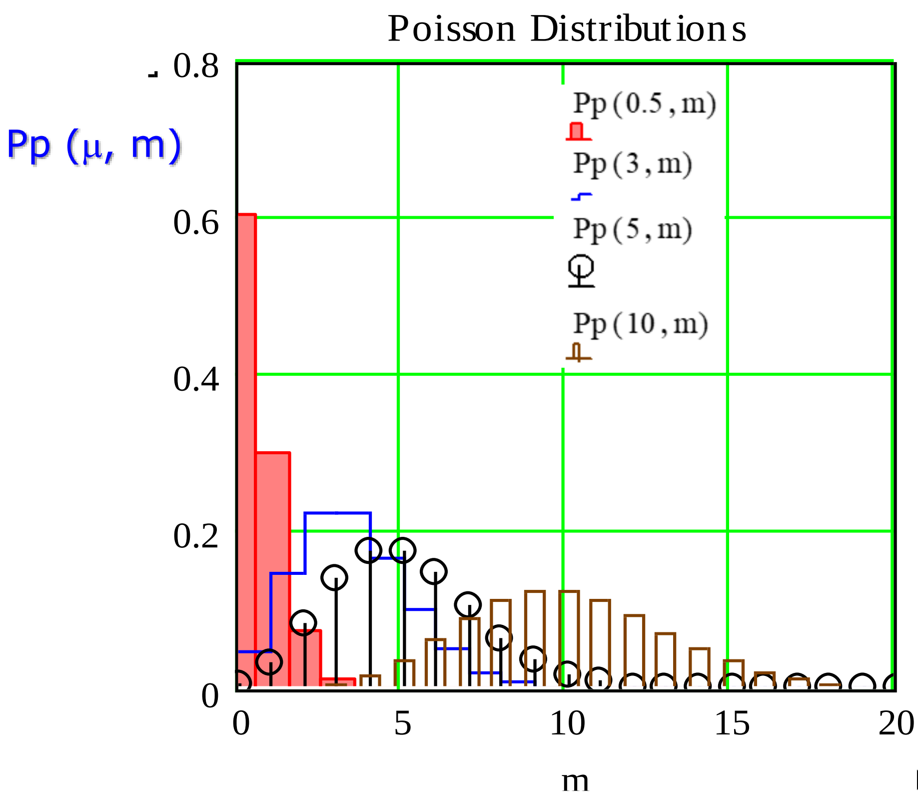
Binomial distribution in the limit of small p and large N (N·p > 0)

$$\lim_{\substack{p\to 0\\ and\ N\to\infty}} P_{binomial}(N,m) = P_{Poisson}(\mu,m)$$

Probability for observing m events when average is $<m> = \mu$

$$P_{Poisson}(\mu,m) = \frac{\mu^m \cdot e^{-\mu}}{m!}$$

m= 0,1,2, …



$\mu = <m> = N\cdot p$ and $\sigma^2 = \mu$

is the mean, the average number of successes in N trials.

Observe *N* counts (events) → → uncertainty is $\sigma = \sqrt{\mu}$

Unlike the binomial distribution, the Poisson distribution does not depend explicitly on p or N !

→ With increasing p (<1.0):

Poisson → Gaussian (Normal Distribution)

Distribution Moments and Limits

$$P_{binomial}(N,m,p) = \binom{N}{m} p^m (1-p)^{N-m}$$

Probability for m "successes" out of N trials, individual probability p

Measurement mean and variance ('uncertainty')

$$\bar{m} = N \cdot p \approx N_{obs} \quad and \quad \sigma_m^2 = N \cdot p \cdot (1-p) \approx N_{obs}$$

$$N_{obs} = \#\ of\ \text{"counts" observed}, \quad p \ll 1.0$$

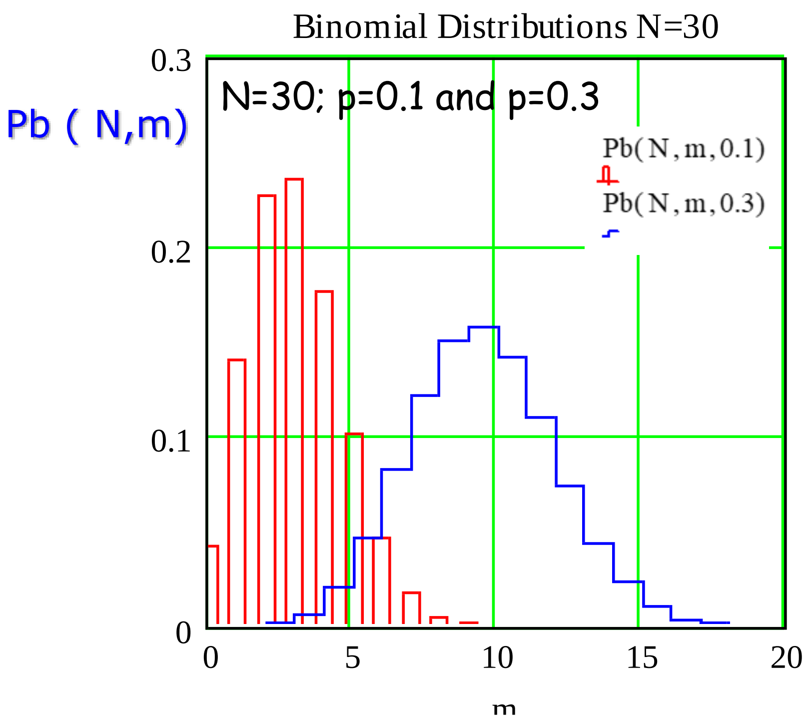
$$Statistical\ \text{"error"}\ of\ N_{obs} : \sigma_m \approx \sqrt{N_{obs}}$$

$$\frac{\sigma_m}{\bar{m}} = \frac{\sqrt{N \cdot p \cdot (1-p)}}{N \cdot p} \approx \frac{1}{\sqrt{N_{obs}}}$$

$\rightarrow$ *more counts = smaller error*

Observe Poisson → Gaussian/Normal

$$\lim_{\substack{p \to 1 \\ N \to \infty}} P_{bin}(N,m,p) = \frac{1}{\sqrt{2\pi\sigma_m^2}} \cdot \exp\left\{ -\frac{(x - \langle m \rangle)^2}{2\sigma_m^2} \right\}$$



Distributions P(m) approximates Gaussian very fast, already good for p=0.2-0.3

Central-Limit Theorem

Means (averages) of different samples of a large population data set cluster together closely. → general property of samples of stochastic variables

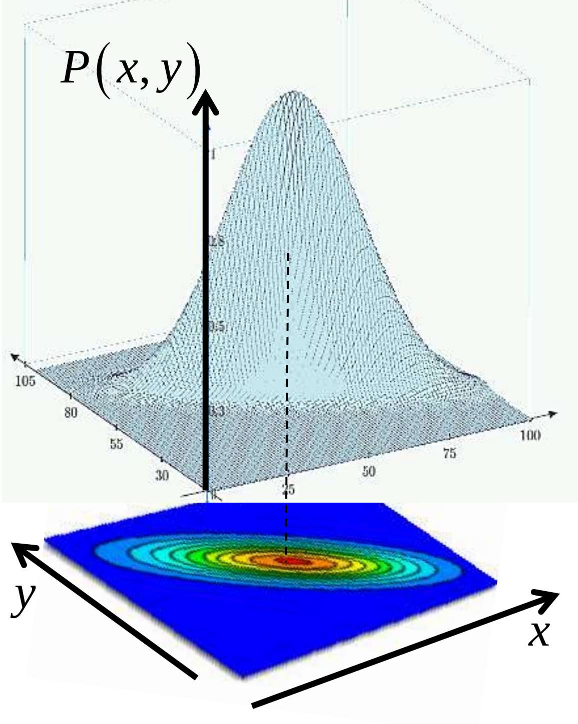
The distribution of the **sample means** approaches a Gaussian normal distribution with increasing size of the sample, *regardless of the form of the original (population) distribution*.

The mean (average) of a distribution of stochastic data does not contain information on the actual shape of the distribution.

The average of any truly random sample of a population is already close to the true population average. Considering many samples, or large samples, narrows the choices. The Gaussian width becomes narrower for larger samples. → The standard error of the mean decreases as the sample size increases.

Continuous Probability Distributions

Degrees of freedom $\{x, y, ...\} \rightarrow$ *Joint Probability* $P(x, y, ...)$



Normalized Probability

$$\iiint dx\, dy \cdots\cdots P(x, y, ...) \equiv 1$$

Partial or conditional probability

$$P(x) = \iiint \cancel{dx}\, dy \cdots P(x, y, ...) < 1$$

$$P(x = a) = \iiint dx\, dy \cdots P(x, y, ...)\, \delta(x - a) < 1$$

$$\text{Average } \langle x \rangle_P = \int \int_0^\infty x \cdot P(x, y, ...) dx \cdot dy ...$$

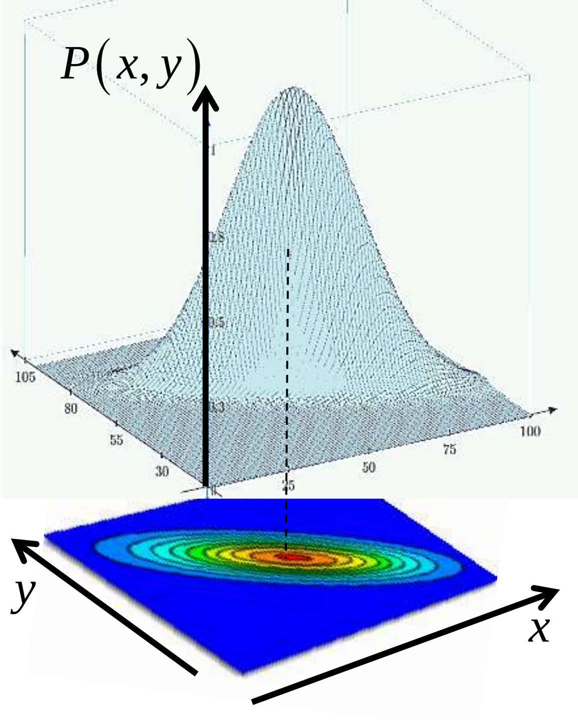
$$\text{and } \sigma_x^2 = \left\langle \left(x - \langle x \rangle \right)^2 \right\rangle_P$$

Define normalized Gaussian probability

$$P(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \cdot \exp\left\{ -\frac{\left(x - \langle x \rangle \right)^2}{2\sigma_x^2} \right\}$$

Continuous Probability Distributions

Degrees of freedom $\{x, y, \ldots\} \rightarrow$ *Joint Probability* $P(x, y, \ldots)$



Normalized Probability

$$\iiint dx\, dy \cdots\cdots P(x, y, \ldots) \equiv 1$$

Partial or conditional probability

$$P(y) = \iiint dx \cdots P(x, y, \ldots) < 1$$

$$P(y = b) = \iiint dx\, dy \cdots P(x, y, \ldots)\, \delta(y - b) < 1$$

$$\text{Average } \langle y \rangle_P = \int \int_0^\infty y \cdot P(x, y, \ldots) dx \ldots$$

$$\text{and } \sigma_y^2 = \left\langle \left(y - \langle y \rangle \right)^2 \right\rangle_P$$

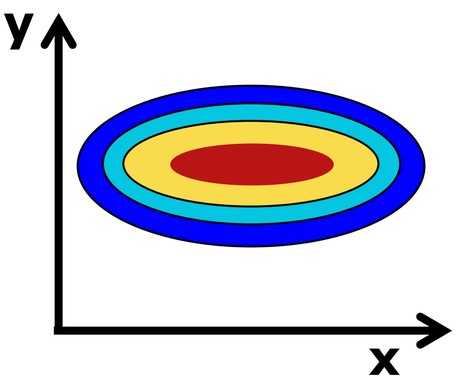
Define normalized Gaussian probability

$$P(y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \cdot \exp\left\{ -\frac{\left(y - \langle y \rangle \right)^2}{2\sigma_y^2} \right\}$$

Correlations in Joint Distributions

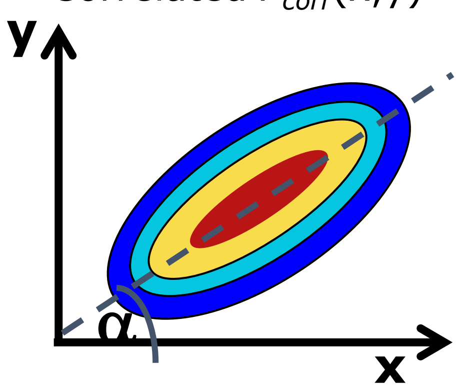
For independent events (*d.o.f*) probabilities multiply. Correlations within distributions. Example: *y* increases with increasing *x*

Uncorrelated P_{unc}(x,y)



$$P_{unc}(x,y) = P(x)\cdot P(y) = \\ = \frac{1}{\sqrt{4\pi^2\sigma_x^2}}\cdot \exp\left\{-\left[\frac{\left(x-\langle x\rangle\right)^2}{2\sigma_x^2} + \frac{\left(y-\langle y\rangle\right)^2}{2\sigma_y^2}\right]\right\}$$

Correlated P_{corr}(x,y)



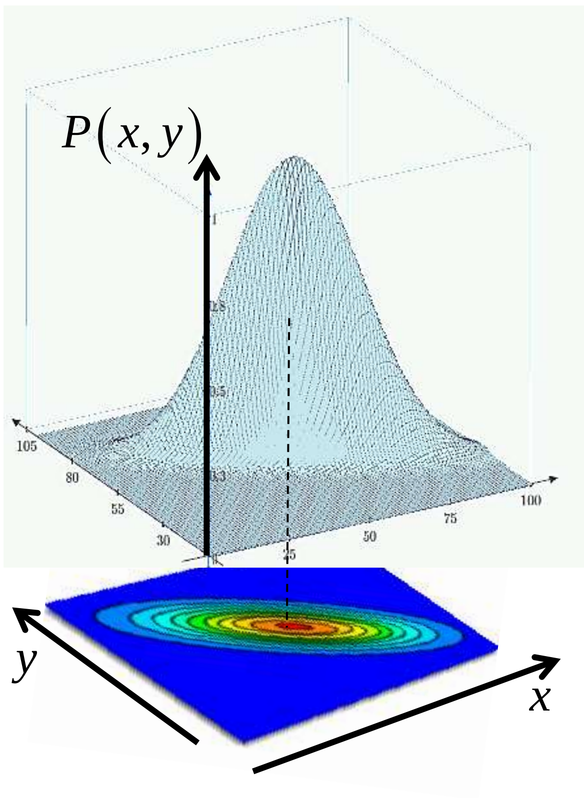
$$P_{corr}(x,y) = \frac{1}{2\pi\sqrt{\sigma_x^2\sigma_y^2 - \sigma_{xy}^2}}\cdot \\ \cdot \exp\left\{-\frac{\left(x-\langle x\rangle\right)^2\sigma_x^2 + \left(y-\langle y\rangle\right)\ \sigma_y^2 - 2\left(x-\langle x\rangle\right)\left(y-\langle y\rangle\right)\sigma_{xy}}{2\left(\sigma_x^2\sigma_y^2 - \sigma_{xy}^2\right)}\right\}$$

$$Covariance:\ \sigma_{\mathbf{xy}} = \int \left(x-\langle x\rangle\right)\cdot\left(y-\langle y\rangle\right)\cdot P(x,y)\,dx\,dy$$

$$\cot\alpha = \frac{1}{2\sigma_{xy}}\left\{\left(\sigma_x^2-\sigma_y^2\right)+\sqrt{\left(\sigma_x^2-\sigma_y^2\right)^2+4\sigma_{xy}^2}\right\}$$

correlation coefficient $r_{xy} = \sigma_{xy}/\left(\sigma_x\sigma_y\right);\ -1 \le r_{xy} \le 1$

Probability Generating Functions



Example: $1-$*dimensional system* $P(x)$

Characteristic functions of $P(x)$
(same info)
Laplace transformation of P

$$\Lambda(s) := \int dx\, e^{-s\cdot x} P(x)$$

Derivatives of $\Lambda(s)$:

$$\frac{d^n}{ds^n}\Lambda(s) := \frac{d^n}{ds^n}\int_0^\infty dx\, e^{-s\cdot x} P(x)$$

$$= (-)^n \int_0^\infty dx\, x^n e^{-s\cdot x} P(x)$$

(Set s=0)

$$-\left(\frac{d}{ds}\Lambda(s)\right)_{s=0} = \int dx\, x \cdot P(x) = \langle x \rangle$$

$\rightarrow$ *similar*:

$$\langle x^n \rangle = (-1)^n \left(\frac{d^n}{ds^n}\Lambda(s)\right)_{s=0}$$

End

Probability & Stochastic Vrbls