Correlations, Independence & Gaussians

Jonathan Pillow

Mathematical Tools for Neuroscience (NEU 314)
Fall, 2021

lecture 16

Quiz

Use the joint distribution P(x,y) shown on the left to answer the following:

P(x,y)

y \ x	0	1
1	0.1	0.3
0	0.2	0.4

1. Compute the marginal P(x)
2. Compute the marginal P(y)
3. Compute the conditional P(y | x = 0)
4. Compute the conditional P(x | y = 1)

Use the distribution P(x) shown on the left to answer:

P(x)

x	0	1
P(x)	0.2	0.8

5. What is $\mathbb{E}[x]$, expected value of x?

practice problems: Bayes rule

Consider the following model describing how a single neuron responds to houses and faces, which is given by a pair of conditional distributions:

# spikes	0	1	2	3	4
P(# spikes l "house")	0.1	0.1	0.3	0.4	0.1
P(# spikes l "face")	0.1	0.4	0.3	0.2	0

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}$$

or practically:

$$P(x|y) \propto P(y|x)P(x)$$

Furthermore, suppose P(house) = P(face) = 0.5

3) What is the most probable stimulus if you observe 3 spikes?
   • Compute P(face l 3 spikes) and P(house l 3 spikes) using Bayes' rule?
4) Is there any response for which you can be certain of what the stimulus was?
5) Re-answer #3 under the prior that P(house) = 0.2 , P(face) = 0.8

Recap:

- expectation (averages)

$$\mathbb{E}[f(x)] = \int f(x)P(x)dx$$

$$\mathbb{E}[f(x)] = \sum_i f(x_i)P(x_i)$$
$$= \vec{f}_x \cdot \vec{p}_x$$

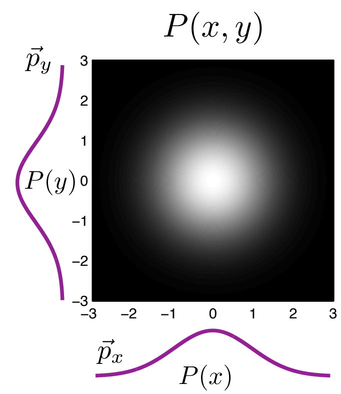
- mean and variance (moments)

$$\bar{x} = \mathbb{E}[x]$$

$$\text{var}(x) = \mathbb{E}[(x - \bar{x})^2]$$

- independence

$$P(x, y) = P(x)P(y)$$
$$= \vec{p}_y \, \vec{p}_x^{\,T}$$

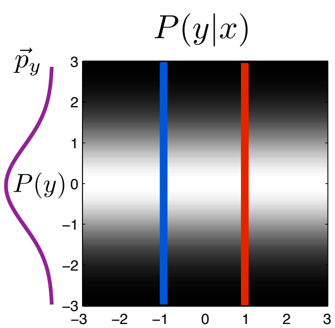


OR

$$P(y|x) = P(y)$$ for all x

or

$$P(x|y) = P(x)$$ for all y



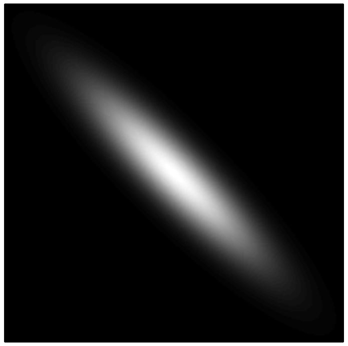
Correlation vs. Dependence

1. Correlation coefficient

$$\mathrm{corr}(x, y) = \frac{\mathbb{E}[(x - \bar{x})(y - \bar{y})]}{\sqrt{\mathrm{var}(x)\mathrm{var}(y)}}$$

positive correlation

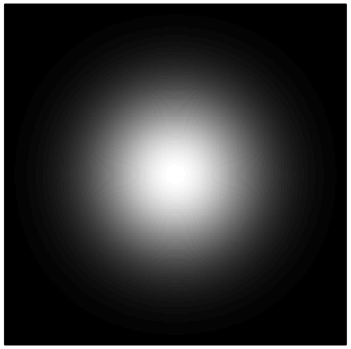
negative correlation



Linear relationship
between x and y

- value between -1 ("perfect anti-correlation") and +1 ("perfect correlation")

- 0 = no correlation

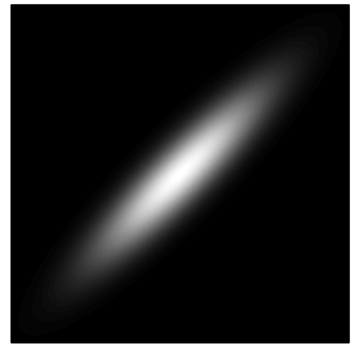


Correlation vs. Dependence

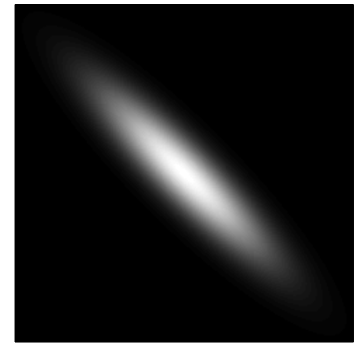
1. Correlation coefficient

$$\mathrm{corr}(x, y) = \frac{\mathbb{E}[(x - \bar{x})(y - \bar{y})]}{\sqrt{\mathrm{var}(x)\mathrm{var}(y)}}$$

positive correlation



negative correlation



Linear relationship between x and y

2. Dependence

- arises whenever $P(x, y) \neq P(x)P(y)$
- can be quantified by anything that measures mismatch e.g., mutual information: $MI(x, y) = D_{KL}(P(x, y), P(x)P(y))$

  KL divergence

- MI=0 ⇒ independence

Correlation vs. Dependence

Q: Can you draw a distribution that is *uncorrelated* but *dependent?*

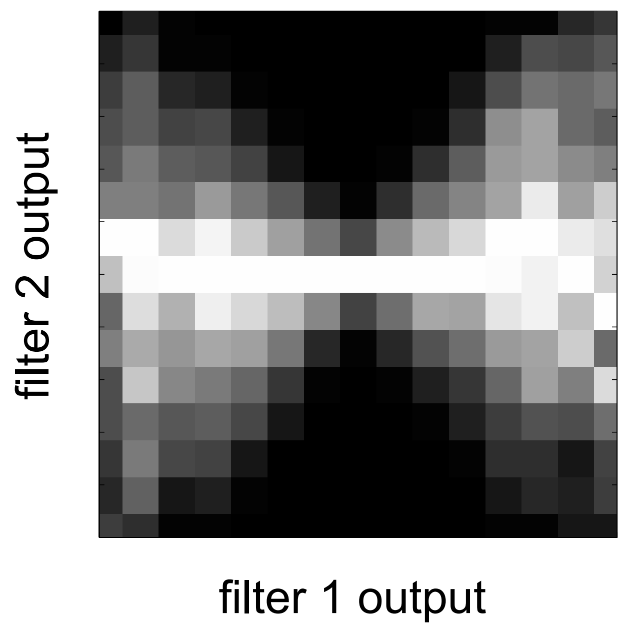
Correlation vs. Dependence

Q: Can you draw a distribution that is *uncorrelated* but *dependent?*

"Bowtie" dependencies in natural scenes:

(uncorrelated but dependent)

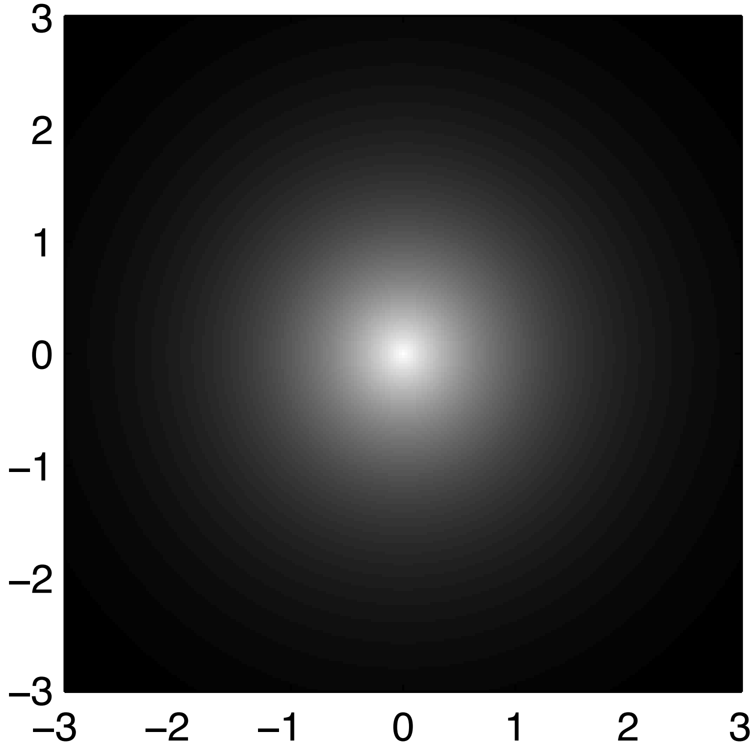
P(filter 2 output **|** filter 1 output)



[Schwartz & Simoncelli 2001]

Is this distribution independent?

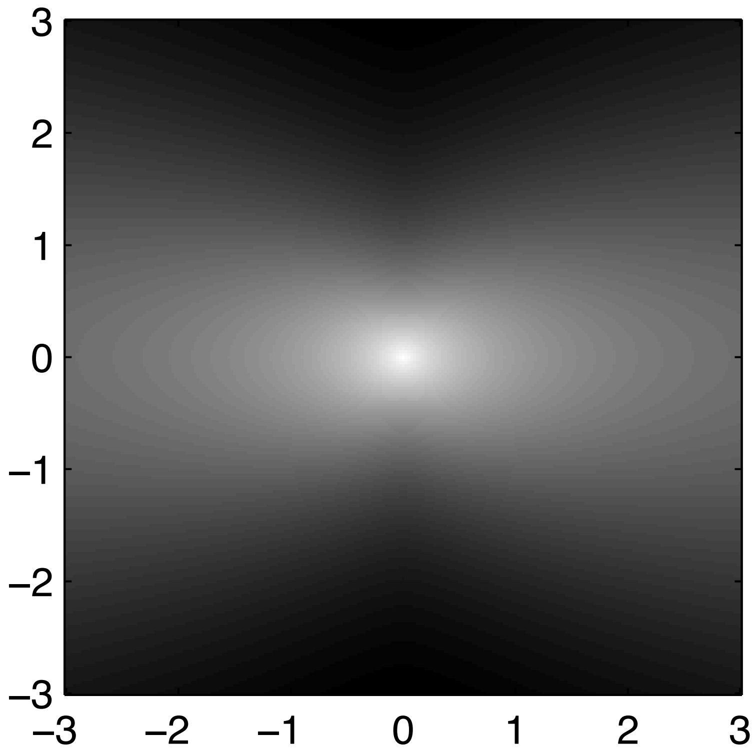
$P(x, y)$



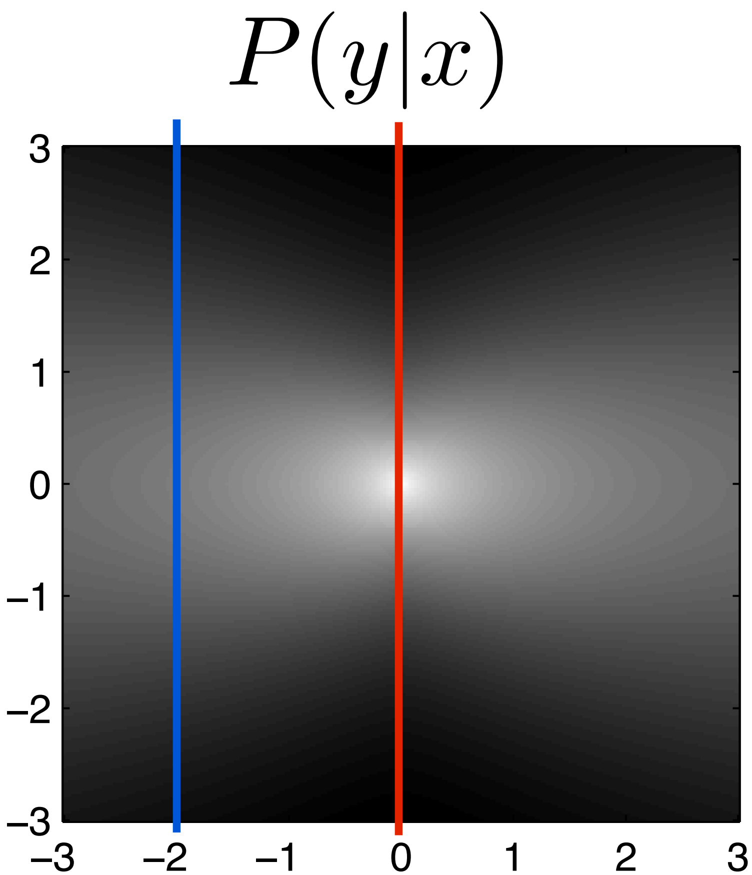
Is this distribution independent?

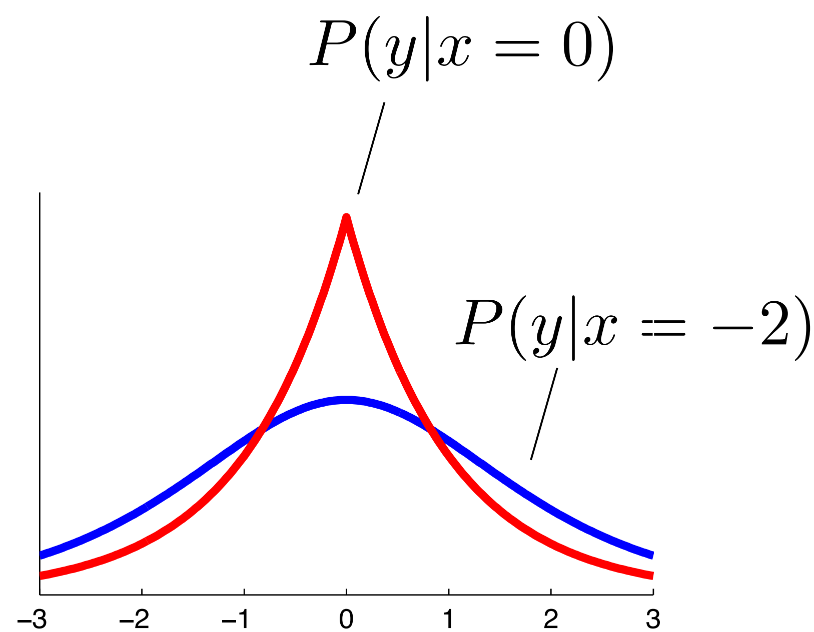
$$P(y|x)$$

(each column is a conditional distribution)



Is this distribution independent?





No! Conditionals over y are different for different x!

FUN FACT:

Gaussian is the only distribution that can be both:

- independent (equal to the product of its marginals)
- spherically symmetric: $P(\vec{x}) = P(U\vec{x})$

orthogonal matrix

Corollary: circular scatter / contour plot not sufficient to show independence!

the amazing Gaussian

What else about Gaussians is awesome?

Gaussian family closed under many operations:

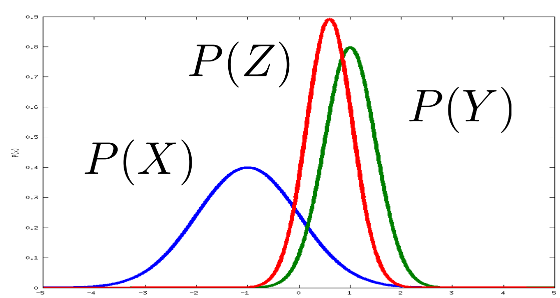
1. scaling: $X \sim Gaussian \implies aX$ is Gaussian
2. sums: $X, Y \sim Gaussian \implies X + Y$ is Gaussian

(thus, any linear function Gaussian RVs is Gaussian)

3. products of Gaussian distributions

Gaussian density

$$X, Y \sim Gaussian \implies P(X)P(Y) \propto P(Z)$$



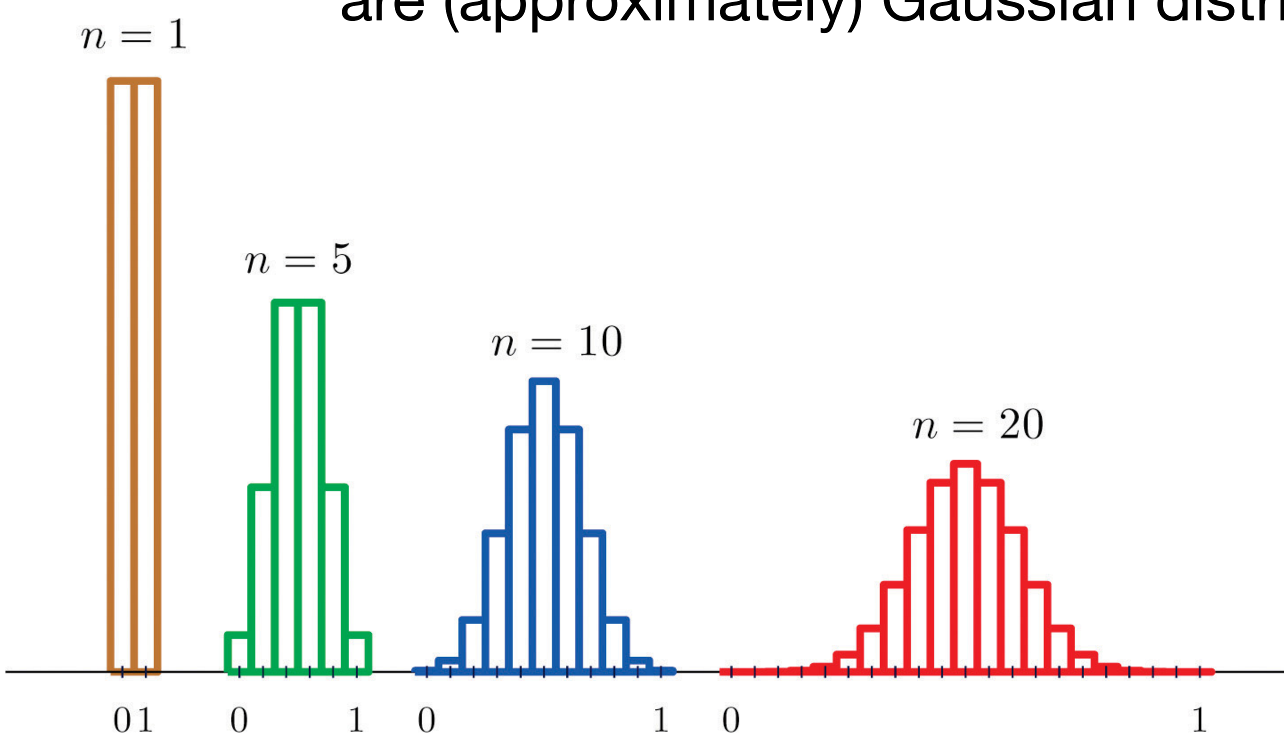
the amazing Gaussian

4. Average of many (non-Gaussian) RVs is Gaussian!

Central Limit Theorem: $S_n = \dfrac{X_1 + \cdots X_n}{n}$ is Gaussian

• explains why many things in the world are (approximately) Gaussian distributed

coin flipping:



http://statwiki.ucdavis.edu/Textbook_Maps/General_Statistics/Shafer_and_Zhang's_Introductory_Statistics/06%3A_Sampling_Distributions/6.2_The_Sampling_Distribution_of_the_Sample_Mean

the amazing Gaussian

Multivariate Gaussians: $\vec{X} \sim \mathcal{N}(\vec{\mu}, C)$

mean → $\vec{\mu}$, cov → C

(The random variable X is distributed according to a Gaussian distribution)

5. Marginals and conditionals ("slices") are Gaussian

6. Linear projections: $\vec{Y} = A\vec{X} \implies \vec{Y} \sim \mathcal{N}(A\vec{\mu}, ACA^\top)$

multivariate Gaussian

mean → $\vec{\mu}$, covariance → C

$$\vec{X} \sim \mathcal{N}(\vec{\mu}, C)$$

("X is distributed according to a Gaussian distribution")

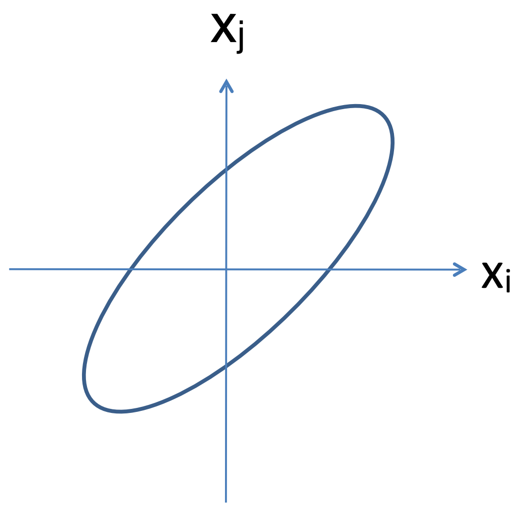
$$P(\vec{X} = \vec{x}) = \frac{1}{\sqrt{|2\pi C|}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^\top C^{-1}(\vec{x}-\vec{\mu})}$$

what is covariance?

$$\mathbb{E}[\vec{x}]$$ mean

$$\text{cov}(\vec{x}) = \mathbb{E}[(\vec{x} - \bar{x})(\vec{x} - \bar{x})^T]$$

- n x n matrix



- the i,j'th element of the matrix:

$$\text{cov}(x_i, x_j) = \mathbb{E}[(x_i - \bar{x}_i)(\vec{x}_j - \bar{x}_j)^T]$$

- diagonal elements are variances

$$\text{cov}(x_i, x_i) = \text{var}(x_i) = \mathbb{E}[(x_i - \bar{x}_i)^2]$$

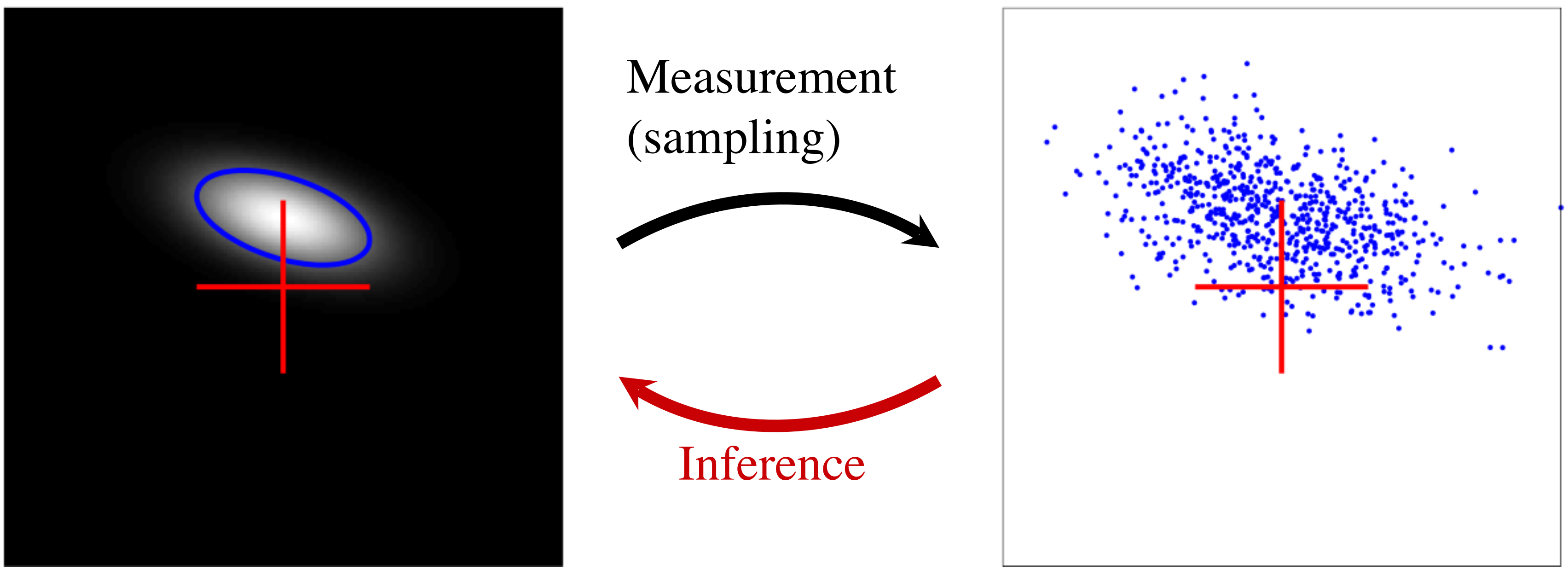
- relationship to correlation coefficient:

$$\text{corr}(x, y) = \frac{\mathbb{E}[(x - \bar{x})(y - \bar{y})]}{\sqrt{\text{var}(x)\text{var}(y)}} = \frac{\text{cov(x, y)}}{\sqrt{\text{var}(x)\text{var}(y)}}$$

sampling vs. inference

bivariate normal density

700 samples



true mean: [0 0.8]
true cov: [1.0 -0.25
-0.25 0.3]

sample mean: [-0.05 0.83]
sample cov: [0.95 -0.23
-0.23 0.29]