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:

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

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

5. What is $\mathbb{E}[x]$, expected value of $x$?
Consider the following model describing how a single neuron responds to houses and faces, which is given by a pair of conditional distributions:

<table>
<thead>
<tr>
<th># spikes</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>P( # spikes I “house”)</td>
<td>0.1</td>
<td>0.1</td>
<td>0.3</td>
<td>0.4</td>
<td>0.1</td>
</tr>
<tr>
<td>P( # spikes I “face”)</td>
<td>0.1</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0</td>
</tr>
</tbody>
</table>

Furthermore, suppose $P(\text{house}) = P(\text{face}) = 0.5$

3) What is the most probable stimulus if you observe 3 spikes?
   - Compute $P(\text{face I 3 spikes})$ and $P(\text{house I 3 spikes})$ 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 $P(\text{house}) = 0.2$, $P(\text{face}) = 0.8$
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**
  \[ P(x, y) = P(x) P(y) \]
  \[ = \vec{p}_y \cdot \vec{p}_x^T \]
  OR
  \[ P(y|x) = P(y) \quad \text{for all } x \]
  \[ P(x|y) = P(x) \quad \text{for all } y \]
Correlation vs. Dependence

1. Correlation coefficient

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

- value between -1 (“perfect anti-correlation”) and +1 (“perfect correlation”)
- 0 = no correlation

positive correlation

negative correlation

Linear relationship between x and y
Correlation vs. Dependence

1. Correlation coefficient

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

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: \( MI(x, y) = D_{KL}(P(x, y), P(x)P(y)) \)

• \( MI=0 \Rightarrow \text{independence} \)

KL divergence
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: \( P(\vec{x}) = P(U\vec{x}) \)

**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: \( X \sim \text{Gaussian} \implies aX \) is Gaussian

2. sums: \( X, Y \sim \text{Gaussian} \implies X + Y \) is Gaussian
   
   (thus, any linear function Gaussian RVs is Gaussian)

3. products of Gaussian distributions

\[
X, Y \sim \text{Gaussian} \implies P(X)P(Y) \propto P(Z)
\]
the amazing Gaussian

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

Central Limit Theorem: \[ S_n = \frac{X_1 + \cdots + X_n}{n} \] is Gaussian

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

coin flipping:

http://statwiki.ucdavis.edu/Textbook_Maps/General_Statistics/Shafie_and_Zhang's_Introductory_Statistics/06%3A_Sampling_Distributions/6.2_The_Sampling_Distribution_of_the_Sample_Mean
the amazing Gaussian

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

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

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

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

\[ \tilde{X} \sim \mathcal{N}(\mu, C) \]

(“\(X\) is distributed according to a Gaussian distribution”)

\[
P(\tilde{X} = \tilde{x}) = \frac{1}{\sqrt{|2\pi C|}} e^{-\frac{1}{2} (\tilde{x} - \mu)^\top C^{-1} (\tilde{x} - \mu)}
\]
what is covariance?

\[ \mathbb{E}[\bar{x}] \quad \text{mean} \]

\[ \text{cov}(\bar{x}) = \mathbb{E}[(\bar{x} - \bar{x})(\bar{x} - \bar{x})^T] \quad \text{• n x n matrix} \]

• the i,j’th element of the matrix:

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

• 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:

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

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]