COVARIANCE, LLSE, CONDITIONAL EXPECTATION, MARKOV CHAINS 10

COMPUTER SCIENCE MENTORS 70

November 14 to November 18, 2016

1 Covariance

1.1 Introduction

The **covariance** of two random variables X and Y is defined as:

$$\text{Cov}(X,Y) := \text{E}((X-\text{E}(X))\cdot(Y-\text{E}(Y)))$$

1.2 Warm Up

1. Prove that $\text{Cov}(X,X) = \text{Var}(X)$:

Solution:

$$\text{Cov}(X,X) = \text{E}(X\cdot X) - \text{E}(X)\cdot\text{E}(X) = \text{E}(X^2) - \text{E}(X)^2$$

2. Prove that if X and Y are independent, then $\text{Cov}(X,Y) = 0$:

Solution:

$$\text{Cov}(X,Y) = \text{E}(X\cdot Y) - \text{E}(X)\cdot\text{E}(Y)$$

Remember that a property of expectation is that if X and Y are independent, then $\mathrm{E}(XY) = \mathrm{E}(X) \cdot \mathrm{E}(Y)$, so we get 0 when we subtract

3. Prove that $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$:

Solution:

$$\begin{aligned}
\mathrm{Cov}(X + Y, Z) &= \mathrm{E}((X + Y) \cdot Z) - \mathrm{E}(X + Y) \cdot \mathrm{E}(Z) \\
&= \mathrm{E}(X \cdot Z) + \mathrm{E}(Y \cdot Z) - (\mathrm{E}(X) \cdot \mathrm{E}(Z) + \mathrm{E}(Y) \cdot \mathrm{E}(Z)) \\
&= \mathrm{E}(X \cdot Z) - \mathrm{E}(X) \cdot \mathrm{E}(Z) + \mathrm{E}(Y \cdot Z) - \mathrm{E}(Y) \cdot \mathrm{E}(Z) \\
&= \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)
\end{aligned}$$

1.3 Questions

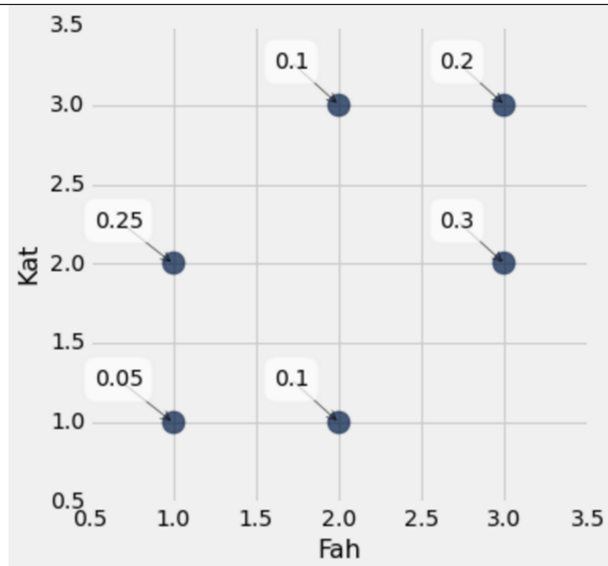
1. Roll 2 dice. Let A be the number of 6's you get, and B be the number of 5's, find $\mathrm{Cov}(A, B)$

Solution: $\mathrm{E}(A) = \frac{1}{6}$ for one die, by linearity of expectation, two dice make $\frac{1}{3}$, same for $\mathrm{E}(B)$ $\mathrm{E}(A) = \frac{1}{3}, \mathrm{E}(B) = \frac{1}{3}$
AB can be either 0 (if no 5's or 6's show up) or 1 (get a 5 and a 6).

$$\begin{aligned}
\mathrm{E}(AB) &= 1 \cdot \mathrm{P}[\text{get a 5 and a 6}] \\
&= \mathrm{P}[\text{first die = 5 and second die = 6}] + \mathrm{P}[\text{first die = 6 and second die = 5}] \\
&= \frac{1}{36} + \frac{1}{36}
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(AB) &= \mathrm{E}(AB) - \mathrm{E}(A) \cdot \mathrm{E}(B) \\
&= \frac{1}{18} - \frac{1}{9} \\
&= -\frac{1}{18}
\end{aligned}$$

2. Consider the following distribution with random variables Fah and Kat:



Find the covariance of Fah and Kat.

Solution: E(Fah) $= 1 \cdot .3 + 2 \cdot .2 + 3 \cdot .5 =$ **2.2**
E(Kat) = $1 \cdot .15 + 2 \cdot .55 + 3 \cdot .3 =$ **2.15**
E(KatFah) = $1 \cdot 1 \cdot .05 + 1 \cdot 2 \cdot .25 + 2 \cdot 1 \cdot .1 + 2 \cdot 3 \cdot .1 + 3 \cdot 2 \cdot .3 + 3 \cdot 3 \cdot .2 =$ **4.95**
cov(Kat, Fah) = $4.95 - 2.2 \cdot 2.15 =$ **0.22**

2 LLSE

2.1 Introduction

Theorem: Consider two random variables, X, Y with a given distribution $\mathrm{P}[X = x, Y = y]$. Then

$$\mathrm{L}[Y|X] = \mathrm{E}(Y) + \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}(X - \mathrm{E}(X))$$

2.2 Questions

1. Assume that

$$Y = \alpha X + Z$$

where X and Z are independent and $\mathrm{E}(X) = \mathrm{E}(Z) = 0$. Find $\mathrm{L}[X|Y]$.

Solution:

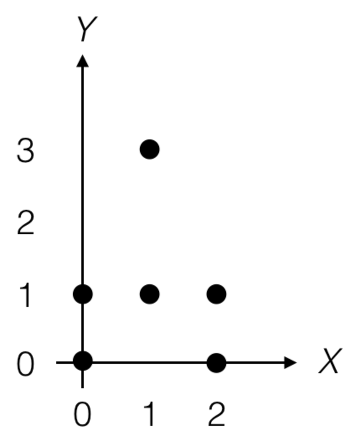
$$\begin{aligned}\text{Cov}(X,Y) &= \text{E}(XY) - \text{E}(X)\text{E}(Y) \\ &= \text{E}(X \cdot (\alpha X + Z)) = \alpha \text{E}(X^2)\end{aligned}$$

$$\begin{aligned}\text{Var}(Y) &= \alpha^2 \text{Var}(X) + \text{Var}(Z) \\ &= \alpha^2 \text{E}(X^2) + \text{E}(Z^2)\end{aligned}$$

Therefore,

$$\text{L}[X|Y] = \frac{\alpha \text{E}(X^2)}{\alpha^2 \text{E}(X^2) + \text{E}(Z^2)} \cdot Y$$

2. The figure below shows the six equally likely values of the random pair (X,Y). Specify the functions of:
   - $L[Y \mid X]$
   - $E(X \mid Y)$
   - $L[X \mid Y]$
   - $E(Y \mid X)$



Solution: Let's calculate some useful properties of the distribution first and then see how we can use them to calculate the estimates.

$$\begin{aligned}
|\Omega| = 6 &\implies P[\text{one point}] = \frac{1}{6} \\
\mathrm{E}(X) &= 0\left(\frac{2}{6}\right) + 1\left(\frac{2}{6}\right) + 2\left(\frac{2}{6}\right) \\
&= 1 \\
\mathrm{E}(Y) &= 0\left(\frac{2}{6}\right) + 1\left(\frac{3}{6}\right) + 30\left(\frac{1}{6}\right) \\
&= 1 \\
\mathrm{E}(XY) &= 0\left(\frac{3}{6}\right) + 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) \\
&= 1 \\
\mathrm{Cov}(X,Y) = 0 &\implies \mathrm{L}[Y|X] = E(Y)
\end{aligned}$$

- $L[Y \mid X]$: Using the LLSE formula: $\mathrm{L}[Y \mid X] = \mathrm{E}[Y] + \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}(Y - \mathrm{E}[Y]) = \mathrm{E}[Y]$. Therefore $\boxed{L[Y \mid X] = 1}$
- $E[X \mid Y]$: Notice the symmetry across $X = 1$. For all values of y, $E[X|Y = y]$ is the same; therefore $\boxed{E[X|Y] = E[X] = 1}$.
- $L[X \mid Y]$: The MMSE estimator for X given Y is a linear function, therefore $\boxed{\mathrm{L}[X \mid Y] = \mathrm{E}[X \mid Y] = 1}$
- $E[Y \mid X]$ For this one we can't make use of symmetry or directly apply what we calculated above. We must go back to the definition of conditional expectation. We can calculate $\mathrm{E}[Y \mid X = x]$ for every point x, and that entirely defines the expression:

$$\mathrm{E}(Y \mid X = x) = \begin{cases} \frac{1}{2} & \text{if } x = 0 \\ 2 & \text{if } x = 1 \\ \frac{1}{2} & \text{if } x = 2 \end{cases}$$

The above equation is sufficient, but we can go further by realizing that these points are part of a flipped absolute value function centered around $x = 1$: $\boxed{\mathrm{E}[Y \mid X] = \frac{-3}{2}|X - 1| + 2}$. Indeed, this is not linear, which is why $\mathrm{L}[Y \mid X] \neq \mathrm{E}[Y \mid X]$.

3 Conditional Expectation

3.1 Introduction

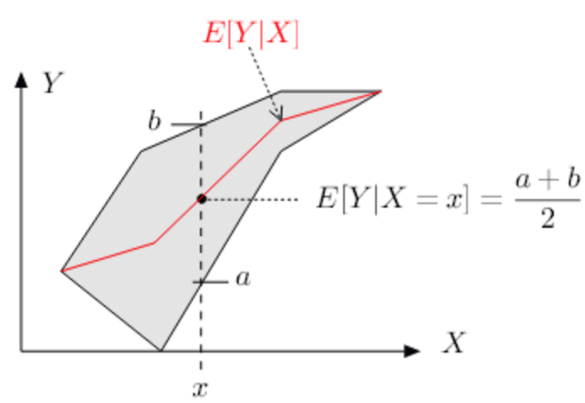
The **conditional expectation** of Y given X is defined by

$$\mathrm{E}[Y|X=x] = \sum_y y \cdot \mathrm{P}[Y=y|X=x] = \sum_y y \cdot \frac{\mathrm{P}[X=x, Y=y]}{\mathrm{P}[X=x]}$$

Properties of Conditional Expectation

$$\begin{aligned}
\mathrm{E}(a|Y)) &= a \\
\mathrm{E}(aX+bZ|Y) &= a \cdot \mathrm{E}(X|Y) + b \cdot \mathrm{E}(Z|Y) \\
\mathrm{E}(X|Y) &\geq 0 \text{ if } X \geq 0 \\
\mathrm{E}(X|Y) &= \mathrm{E}(X) \text{ if } X, Y \text{ independent} \\
\mathrm{E}(\mathrm{E}(X|Y)) &= \mathrm{E}(X)
\end{aligned}$$

Solution: Here is a picture that shows that conditioning creates a new random variable with a new distribution. Figure 9 of note 26 does so.



3.2 Questions

1. Prove $\mathrm{E}(\mathrm{E}(Y|X)) = \mathrm{E}(Y)$

Solution:

$$\mathrm{E}(\mathrm{E}(Y|X)) = \sum_x \mathrm{E}(Y|X=x) \cdot \mathrm{P}[X=x]$$

$$
\begin{aligned}
&= \sum_x (\sum_y y \cdot \mathrm{P}[Y = y | X = x]) \cdot \mathrm{P}[X = x] \\
&= \sum_y y \cdot \sum_x y \cdot \mathrm{P}[X = x | Y = y]) \cdot \mathrm{P}[Y = y] \\
&= \sum_y y \cdot \mathrm{P}[Y = y] \cdot \sum_x \mathrm{P}[X = x | Y = y]) \\
&= \sum_y y \cdot \mathrm{P}[Y = y] = E[Y]
\end{aligned}
$$

2. Prove $\mathrm{E}(h(X) \cdot Y | X) = h(X) \cdot \mathrm{E}(Y | X)$

Solution:

$$
\begin{aligned}
\mathrm{E}(h(X) \cdot Y | X) &= \sum_y h(X) \cdot y \cdot \mathrm{P}[Y = y | X] \\
&= h(X) \sum_y y \cdot \mathrm{P}[Y = y | X] \\
&= h(X) \cdot \mathrm{E}[Y | X]
\end{aligned}
$$