EE126: Probability and Random Processes

Lecture 13: Covariance, Correlation and Iterated Expectations

Abhay Parekh

UC Berkeley

March 1, 2011

Logistics

Midterm

- Half a 8.5x11 sheet handwritten both sides
- Calculator
- Tests understanding of the material at the level of the lectures, book and HW.
- Covers the book through 3.3. Also, 4.1 (but only until page 206).
- Review session today at 7pm in 299 Cory

Continuous Bayes Rule

Y is a continuous rv:

1. We want to infer a continuous rv X:

$$f_{X|Y}(x|y) = \frac{f_X(x) f_{Y|X}(y|x)}{\int_t f_X(t) f_{Y|X}(y|t) dt}$$

2. We want to infer a discrete rv N:

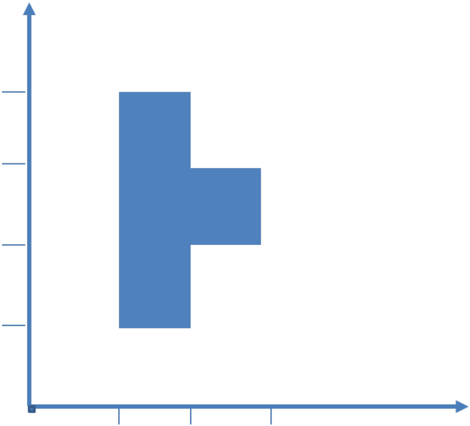
$$P(N = n | Y = y) = \frac{p_N(n) f_{Y|N}(y|n)}{\sum_i p_N(i) f_{Y|N}(y|i)}$$

Conditioning on a Random Variable

Given continuous X, Y, suppose we want $P(X \leq x|Y = y)$. Now clearly, $P(Y = y) = 0$, so what does this mean?
Example: Radar detector error $X \sim N(0, y/10)$ where Y is the speed of the car. Y is also a random variable.
Need a definition for $f_{X|Y}$:
Define

$$\boxed{f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}}$$

Example



Stick Breaking -II

Prob three pieces form triangle: First break uniform on (0,1).
Second break at Y by choosing a point uniformly over $(0, X)$.
The pieces are $y, x-y, 1-x$.
Conditions: $x > 0.5$, $x - y < 0.5$, $y < 0.5$.
Thus $0.5 < x < 1$,
$x - 0.5 < y < \min\{0.5, x\} = 0.5$
$f_X(x) = 1$, $f_{Y|X}(y|x) = \frac{1}{x} \Rightarrow f_{X,Y}(x,y) = \frac{1}{x}$,
$0 < x < 1$.

$$\text{Prob(triangle)} = \int_{.5}^{1} \int_{x-.5}^{0.5} \frac{1}{x} dx$$

$$\text{Prob(triangle)} = \int_{.5}^{1} \frac{1-x}{x} dx = \ln 2 - .5 \approx 0.193$$

What is Prob(triangle) if y is chosen uniformly over the bigger of the two pieces of length x and $1-x$?

Convolution

Let $Z = X + Y$ and assume that X, Y continuous and independent.

$$f_Z(z) = \int_x f_{X,Z}(x,z) = \int_x f(x)f(z|x)dx$$

Now

$$F_{Z|X}(z|x) = P(X+Y \leq z|X = x) = P(x+Y \leq z|X = x) = P(Y \leq z-x$$

Differentiating:

$$f_{Z|X}(x|z) = f_Y(z-x)$$

Thus

$$f_Z(z) = \int_x f_X(x)f_Y(z-x) = \int_y f_Y(y)f_X(z-y)dx$$

This is the convolution operation.

Sum of two independent Gaussians

If $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim (\mu_y, \sigma_y^2)$ then
$Z = X + Y \sim N(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$.
Why? Convolution integral is messy...but let's look at two independent standard normals.

$$
\begin{aligned}
f_Z(z) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\frac{-(z-x)^2}{2} + \frac{x^2}{2}} dx \\
&= \frac{1}{2\pi} e^{\frac{-z^2}{4}} \int_{-\infty}^{\infty} e^{-(\frac{z}{2}-x)^2} dx \\
&= \frac{1}{2\pi} e^{\frac{-z^2}{4}} \sqrt{\pi} \{ \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-(x-\frac{z}{2})^2} dx \} \\
&= \frac{1}{\sqrt{4\pi}} e^{\frac{-z^2}{4}}
\end{aligned}
$$

Thus, $Z \sim N(0, 2)$.

Graphical Convolution

X, Y uniform $[0, 1]$ $Z = X + Y$, $W = X - Y$.

Covariance

We showed earlier for any two random variables X and Y:

$$var(X + Y) = var(X) + var(Y) - 2\underbrace{E[XY] - E[X]E[Y]}_{\text{Covariance}}$$

$$E[X]E[Y] = E[XE[Y]] = E[YE[X]]$$

$$cov(X, Y) = E\big[XY - YE[X] - XE[Y] + E[X]E[Y]\big] = E[(X - E[X])(Y - E[$$

Given X, Y:

$$cov(X, Y) = E[XY] - E[X]E[Y] = E[(X - E[X])(Y - E[Y])]$$

Properties

Given X, Y:

$$cov(X, Y) = E[XY] - E[X]E[Y] = E[(X - E[X])(Y - E[Y])]$$

- $cov(X, X) = var(X)$
- $cov(aX + b, Y) =$
- $cov(a, Y) = 0$
- $cov(X, Y + Z) = cov(X, Y) + cov(X, Z)$

Interpretation of $cov(X, Y)$

Given X, Y:

$$cov(X, Y) = E[XY] - E[X]E[Y] = E[(X - E[X])(Y - E[Y])]$$

Roughly speaking:

- $cov(X, Y) > 0$, Y tends to increase with X
- $cov(X, Y) < 0, Y$ tends to decrease with increasing X
- $cov(X, Y) = 0$ Y tends to move unpredictably with X

Covariance and Independence

If X and Y are independent then, $E[XY] = E[X]E[Y]$ so $cov(X, Y) = 0$.
But, what if $cov(X, Y) = 0$ then X and Y **may not be independent**.
Example: $p_{X,Y}(x, y) = \frac{1}{4}$ if $(x, y) \in \{(1, 0), (-1, 0), (0, 1), (0, -1)\}$.
Reason: The condition for zero correlation is

$$\boxed{E[X|Y] = E[X]}$$

as opposed to $p_{X,Y}(x, y) = p_X(x)p_Y(y)$:

Correlation and Independence

The condition for zero correlation is

$$\boxed{E[X|Y] = E[X]}$$

as opposed to $p_{X,Y}(x,y) = p_X(x)p_Y(y)$:
For discrete rvs:

$$\begin{aligned}
E[XY] &= \sum_{xy} xyp_{X,Y}(x,y) = \sum_{x,y} xyp_{X|Y}(y)p_Y(y) \\
&= \sum_{y} yp_Y(y) \sum_{x} xp_{X|Y}(y) \\
&= \sum_{y} yp_Y(y)E[X|Y=y] \\
&= \sum_{y} yp_Y(y)E[X] \\
&= E[X]E[Y]
\end{aligned}$$

Correlation Coefficient

We want to normalize $cov(X, Y)$ by $f(x, y) > 0$ so that

$$-1 \leq \rho(X, Y) = \frac{cov(X, Y)}{f(X, Y)} \leq +1$$

Define $\hat{X} = X - E[X]$ and $\hat{Y} = Y - E[Y]$.

$$-1 \leq \rho(X, Y) = \frac{E[\hat{X}\hat{Y})}{f(X, Y)} \leq +1$$

$$\rho(X, Y)^2 = \frac{E[\hat{X}\hat{Y}]^2}{f(X, Y)^2} \leq 1$$

Cauchy-Schwartz Inequality says that for any two rvs X, Y:

$$E[XY]^2 \leq E[X^2]E[Y^2]$$

Thus we can pick:

$$f(X, Y)^2 = E[\hat{X}^2]E[\hat{Y}^2] = E[(X - E[X])^2]E[(Y - E[Y])^2] = var(X)var(Y)$$

and $f(X, Y) = \sqrt{var(X)var(Y)}$ (assuming $\mathrm{var}(X)\mathrm{var}(Y) \neq 0$.)

Correlation Coefficient

For any two random variables, X and Y with non zero variance, the correlation coefficient $\rho(X, Y)$ is

$$\rho(X, Y) = \frac{cov(X, Y)}{\sqrt{var(X)var(Y)}}.$$

ρ measures linearity not dependence

Suppose $Y = aX + b$. Then,

$$\begin{aligned} cov(X,Y) &= E[XY] - E[X]E[Y] \\ &= aE[X^2] + bE[X] - aE[X]^2 - bE[X] \\ &= a\ \mathrm{var}(X) \end{aligned}$$

$$\boxed{\rho(X,Y) = \frac{a\mathrm{var}(X)}{\sqrt{\mathrm{var}(X)a^2\mathrm{var}(X)}} = \frac{a}{|a|}}$$

Now let $Z = X^2$. Z is entirely determined by X but $cov(X,Z) \neq 1$.
Example: X: $p_X(1) = .2, p_X(2) = .8$
$E[X] = 1.8, E[X^2] = 3.4, E[X^3] = 6.6, E[X^4] = 13.$
$var(X) = 3.4 - 1.8^2 = 0.16,\ var(X^2) = 13 - 11.56 = 1.44$

$$cov(X,Z) = E[X^3] - E[X]E[X^2] = 6.6 - (1.8)(3.4) = 0.28$$

$$\sqrt{\mathrm{var}(X)\mathrm{var}(Z)} = \sqrt{\mathrm{var}(X)\mathrm{var}(X^2)} = \sqrt{(0.16)(1.44)} = 0.4 * 1.2 = .48$$

$$\boxed{\rho(X,Y) = 0.583}$$

Summary

Special cases for the correlation coefficient, $\rho(X, Y)$:

$$\rho(X, Y) = \begin{cases} 1, & Y = aX + b \; a > 0; \\ -1, & Y = aX + b, \; a < 0; \\ 0, & E[X|Y] = E[X]. \end{cases}$$

Sum of Variances

Given $X_1, ..., X_n$:

$$var(\sum_i X_n) = \sum_i var(X_i) + \sum_i \sum_{j \neq i} cov(X_i, X_j)$$

Example: Seats on a plane
Expected number of passengers in their original seat is 1. What about the variance?
$X_i = 1$ if passenger i is in his seat and 0 otherwise.
$E[X_i] = \frac{1}{N}$ and $var(X_i) = \frac{1}{N}\frac{N-1}{N}$. Now for the covariances:

$$cov(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j].$$

$$E[X_i X_j] = P(X_i = 1, X_j = 1) = \frac{1}{N}\frac{1}{N-1}.$$

$$cov(X_i, X_j) = \frac{1}{N(N-1)} - \frac{1}{N^2} = \frac{1}{N^2(N-1)}$$

Putting it all together...

$$var(\sum_i X_i) = \frac{N-1}{N} + \frac{1}{N} = 1$$

Iterated Expectations

Given $E[X|Y]$ what is $E[E[X|Y]]$?
$E[X|Y]$ gives a real numbered value for any value of $Y = y$. Thus it is a function of y.
$E[E[X|Y]]$ is the expectation of $E[X|Y]$ with respect to Y:
For continuous rvs:

$$
\begin{aligned}
E[E[X|Y]] &= \int_y E[X|Y = y] f_Y(y) dy \\
&= \int_y \int_x x f_{X|Y}(x|y) f_Y(y) dx\ dy \\
&= \int_y \int_x x f_{X,Y}(x, y) dx dy \\
&= \int_x x f_X(x) dx = E[X]
\end{aligned}
$$

Example: Breaking a Stick

A stick of length L is broken uniformly at a point x. The left end of the stick is broken again at y. What is the expected length of piece we are left with?

Y is uniform on $[0, X]$. We need $E[Y]$. Clearly, $E[X] = \frac{L}{2}$ and $E[Y|X] = \frac{X}{2}$. Thus

$$E[E[Y|X]] = E[\frac{X}{2}] = \frac{L}{4}$$

But since $E[Z] = E[E[Y|X]]$, $E[Z] = \frac{L}{4}$.