Lecture 3b: Conditioning

Math 586

January 29, 2009

Two random variables: X, Y.
What is the behavior of Y given X?
Example: X= porosity, Y= permeability.

Let $f(x, y)$= joint density, $f_X(x), f_Y(y)$ - marginals. We have $f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy$ and similarly for $f_Y(y)$.

Define **conditional density** $f(y \mid x) = \dfrac{f(x,y)}{f_X(x)}$ then $f(x, y) = f(y \mid x)f(x)$

$\mathbb{E}[Y \mid X = x]$ = **conditional expected value** (can be modeled by local regression).

$$\mathbb{E}[Y \mid X = x] = \int_{-\infty}^{\infty} y\, f(y|x)\, dy = H(x) \quad \text{some function}$$

$$Var[Y \mid X = x] = \int_{-\infty}^{\infty} [y - H(x)]^2\, f(y|x)\, dy$$

Discrete: sums are used.

<u>General idea</u>: fix one variable. Examine the other. Then, "unfix" the first variable.

Example X = Exposed Surface Area
Y = Amount absorbed

Suppose that $\mathbb{E}(Y \mid X) = 0.2X + \sqrt{X}$, and X has pdf

$$f_X(x) = \begin{cases} 6x(1-x), & 0 \le x \le 1 \\ 0 & otherwise \end{cases}$$

Then

$$\mathbb{E}(Y) = \int_0^1 (0.2x + \sqrt{x}) \cdot 6x(1-x)\, dx = ...$$

$Var(Y)$ is more complex.

Prediction

For two r.v. X, Y, what is the "best" prediction of Y given X?
Depends on what one means by "best". One possibility: minimum MSE (mean square error). That is, if $\mathcal{F}(X)$ is the predictor of Y then find $\mathcal{F}$ that minimizes MSE $= \mathbb{E}[(Y - \mathcal{F}(X))^2]$.

First, consider a simpler question: for a single r.v. X, what is the best predictor that minimizes MSE, that is, find constant a such that $\mathbb{E}[(X-a)^2] \mapsto \min$.
Answer:

$$\mathbb{E}[(X-a)^2] = \mathbb{E}[(X-\mu+\mu-a)^2] = \mathbb{E}[(X-\mu)^2] + 2\mathbb{E}[(X-\mu)(\mu-a)] + (\mu-a)^2 =$$

$$= \mathbb{E}[(X-\mu)^2] + (\mu-a)^2,$$

where $\mu = \mathbb{E}[X]$. The min is reached when $a = \mu$.

A similar argument shows that the best predictor of Y given X is conditional expectation:

$$\mathcal{F}(X) = \mathbb{E}[Y \mid X]$$

Also, the same extends to vectors (when X is replaced by $\mathbf{X}$).

Other criteria are possible: e.g. MAE (mean absolute error) is more robust against outliers. In this case, *median* is a better estimator than the mean.

Example (Discrete case): Given probability table

Y	10	12	13
X = 6	.3	.1	.1
X = 8	.1	.2	.2

What is $P(Y = y \mid X)$?
What is $\mathbb{E}[Y \mid X = 6]$? $\mid X = 8$?

Note that the conditional expectation is an unbiased predictor:

$$\mathbb{E}_X[\mathbb{E}_Y[Y \mid X]] \equiv \mathbb{E}_X[\mathcal{F}(X)] = \mathbb{E}[Y]$$

(if no information is given, your best strategy is to guess $\mathbb{E}[Y]$).