NTNU
Norwegian University of
Science and Technology

Week 34: Lecture 2

Conditional probability and conditional expectation

Geir-Arne Fuglstad

August 19, 2020

Section 2.1: Conditional distributions

Definition

Let A and B be events. The conditional probability of A given B is defined by

$$\Pr\{A|B\} = \begin{cases} \frac{\Pr\{A\cap B\}}{\Pr\{B\}}, & \Pr\{B\} > 0, \\ \text{Not defined}, & \Pr\{B\} = 0. \end{cases}$$

Example 1

Throw one die, and let X denote the number of eyes. Find $\Pr\{X \geq 5 | X \geq 3\}$

Conditional PMFs

Definition (Conditional probability mass function (PMF))

Assume X and Y are jointly distributed random variables. The **conditional PMF** $p_{X|Y}(x|y)$ of X given Y is given by

$$p_{X|Y}(x|y) = \frac{\Pr\{X = x, Y = y\}}{\Pr\{Y = y\}} = \frac{p_{X,Y}(x,y)}{p_Y(y)}, \quad \text{if } p_Y(y) > 0.$$

Note: $\{X = x, Y = y\}$ is short-hand for $\{(X = x) \cap (Y = y)\}$.

Example 2

Throw one die and let

$$X = \text{“Number of eyes”},$$

$$Y = \begin{cases} 0, & \text{if } X \leq 2, \\ 1, & \text{if } X \geq 3. \end{cases}$$

Find the conditional PMF $p_{X|Y}$.

Joint distribution

The conditional PMF is essential to us because we can simplify the joint PMF as

$$p_{X,Y}(x,y) = \Pr\{X = x, Y = y\}$$

.

Joint distribution

The conditional PMF is essential to us because we can simplify the joint PMF as

$$\begin{aligned} p_{X,Y}(x,y) &= \Pr\{X = x, Y = y\} \\ &= \Pr\{Y = y\}\Pr\{X = x | Y = y\} \end{aligned}$$

.

Joint distribution

The conditional PMF is essential to us because we can simplify the joint PMF as

$$\begin{aligned} p_{X,Y}(x,y) &= \Pr\{X = x, Y = y\} \\ &= \Pr\{Y = y\}\Pr\{X = x|Y = y\} \\ &= p_Y(y)p_{X|Y}(x|y). \end{aligned}$$

Simplified notation

Unless it will cause confusion, we typically write

- $p(x)$ instead of $p_X(x)$
- $p(y)$ instead of $p_Y(y)$
- $p(x, y)$ instead of $p_{X,Y}(x, y)$
- $p(x|Y = y)$ instead of $p_{X|Y}(x|y)$

Marginalization

The law of total probability gives

$$\begin{aligned}\Pr\{X = x\} &= \sum_y \Pr\{X = x, Y = y\} \\ &= \sum_y \Pr\{Y = y\}\Pr\{X = x | Y = y\},\end{aligned}$$

Marginalization

The law of total probability gives

$$\Pr\{X = x\} = \sum_{y} \Pr\{X = x, Y = y\}$$
$$= \sum_{y} \Pr\{Y = y\}\Pr\{X = x | Y = y\},$$

Example 3 (Page 48 in book):
A hunter encounters N birds. For each bird, he gets one shot and either hits or misses. Assume the probability of hitting is p for each bird and that the shots are independent.

Marginalization

The law of total probability gives

$$\Pr\{X = x\} = \sum_{y} \Pr\{X = x, Y = y\}$$
$$= \sum_{y} \Pr\{Y = y\}\Pr\{X = x | Y = y\},$$

Example 3 (Page 48 in book):
A hunter encounters N birds. For each bird, he gets one shot and either hits or misses. Assume the probability of hitting is p for each bird and that the shots are independent. Additionally, assume that the number of birds encountered is Poisson distributed with mean λ, i.e., $N \sim \text{Poisson}(\lambda)$. Find the PMF of the number of birds hit.

Conditional expectation

Definition (Conditional exected value)

Let X and Y be random variables, and g a real function. The **Conditional expected value** of $g(X)$ given $Y = y$ is

$$\mathrm{E}[g(X)|Y = y] = \sum_x g(x)\Pr\{X = x|Y = y\}, \quad \text{if } \Pr\{Y = y\} > 0.$$

IMPORTANT!

Theorem (Law of iterated expectations)

Let X and Y be random variables such that $\mathrm{E}[|g(X)|] < \infty$, and let g be a real function. Then

$$\mathrm{E}[g(X)] = \mathrm{E}[\mathrm{E}[g(X)|Y]]$$

IMPORTANT!

Theorem (Law of total variance)

Let X and Y be random variables such that $\mathrm{E}[X^2] < \infty$, then

$$\mathrm{Var}[X] = \mathrm{E}[\mathrm{Var}[X|Y]] + \mathrm{Var}[\mathrm{E}[X|Y]].$$

Example 3 – Revisited

A hunter encounters N birds. For each bird, he gets one shot and either hits or misses. Assume the probability of hitting is p for each bird and that the shots are independent. Additionally, assume that the number of birds encountered is Poisson distributed with mean λ, i.e., $N \sim \text{Poisson}(\lambda)$. Find the expected value and the variance of the number of birds hit.

Formulas:

$$\mathrm{E}[X] = \mathrm{E}[\mathrm{E}[X|Y]]$$
$$\mathrm{Var}[X] = \mathrm{E}[\mathrm{Var}[X|Y]] + \mathrm{Var}[\mathrm{E}[X|Y]]$$