STAT 516

Conditional Distributions and Conditional Expectations

Prof. Michael Levine

April 12, 2020

Motivation

- Let (X, Y) have the joint pmf $f(x, y)$.
- The conditional distribution of X given $Y = y$ is defined to be

$$p(x|y) = P(X = x|Y = y) = \frac{p(x, y)}{p_Y(y)}$$

- The conditional expectation of X given $Y = y$ is defined ot be

$$E(X|Y = y) = \sum_x xp(x|y) = \frac{\sum_x xp(x, y)}{p_Y(y)} = \frac{\sum_x xp(x, y)}{\sum_x p(x, y)}$$

Basic properties of conditional expectation

- Often, $E(X|y)$ is used a shorthand for $E(X|Y=y)$
- Let X and Y be random variables defined on a common sample space Ω. Then,
  1. $$E(g(Y)|Y=y)=g(y)$$
     for any function g
  2. $$E(Xg(Y)|Y=y)=g(y)E(X|Y=y)$$
     for any y and any function g

Independence of two discrete random variables

- The following are several equivalent definitions that can be used
  1. $p(x|y) = p_X(x)$ for any x, y such that $p_Y(y) > 0$
  2. $p(y|x) = p_Y(y)$ for any x, y such that $p_X(x) > 0$
  3. $p(x, y) = p_X(x)p_Y(y)$ for any x, y
  4. $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$ for any x, y
- The version (3) is usually the most convenient
- Suppose X and Y are independent random variables. Then, for any appropriate function $g(X)$ and any y,

$$E[g(X)|Y = y] = E[g(X)]$$

Example I

- Consider the joint distribution of (X, Y) where X is the larger and Y is the smaller of the two dice rolls
- By definition,

$$P(Y = 1|X = 2) = \frac{P(Y = 1, X = 2)}{P(X = 2)} = \frac{1/18}{1/18 + 1/36} = \frac{2}{3}$$

- In the same way,

$$P(Y = 2|X = 2) = \frac{P(Y = 2, X = 2)}{P(X = 2)} = \frac{1/36}{1/18 + 1/36} = \frac{1}{3}$$

- In an automatic way, $P(Y = 1|X = 1) = 1$

Example II

- Suppose X and Y are the binary variables, each taking only values 0 and 1
- The joint distribution is

	Y	
X	0	1
0	s	t
1	u	v

Example II

- By definition,

$$E(X|Y=0) = \frac{0 \times p(0,0) + 1 \times p(1,0)}{p(0,0) + p(1,0)} = \frac{u}{s+u}$$

- $$E(X|Y=1) = \frac{0 \times p(0,1) + 1 \times p(1,1)}{p(0,1) + p(1,1)} = \frac{v}{t+v}$$

- Therefore,

$$E(X|Y=1) - E(X|Y=0) = \frac{v}{t+v} - \frac{u}{s+u} = \frac{vs - ut}{(t+v)(s+u)}$$

- The above implies the single formula:

$$E(X|Y=y) = \frac{u}{s+u} + \frac{vs - ut}{(t+v)(s+u)} y$$

- Whenever X and Y are both binary variables, $E(X|Y=y)$ is linear in y

Conditional variance

- Let (X, Y) have the joint pmf $p(x, y)$. Let $\mu_X(y) = E(X|Y = y)$. The conditional variance of X given $Y = y$ is

$$Var(X|Y = y) = E[(X - \mu_X)^2|Y = y] = \sum_x (x - \mu_X(y))^2 p(x|y)$$

- The usual shorthand for $Var(X|Y = y)$ is $Var(X|y)$.

Example

- Again, let X be the max and Y be the min of two rolls of a die.
- For example, if $y = 3$, then one can compute $\mu_X(y) = E(X|Y = y) = E(X|Y = 3) = 4.71$
- Therefore,

$$
\begin{aligned}
Var(X|y) &= \sum_x (x - 4.71)^2 p(x|3) \\
&= \frac{(3-4.71)^2 \times \frac{1}{36} + (4-4.71)^2 \times \frac{1}{18}}{\frac{1}{36} + \frac{1}{18} + \frac{1}{18} + \frac{1}{18}} \\
&+ \frac{(5-4.71)^2 \times \frac{1}{18} + (6-4.71)^2 \times \frac{1}{18}}{\frac{1}{36} + \frac{1}{18} + \frac{1}{18} + \frac{1}{18}} \\
&= 1.06
\end{aligned}
$$

Poisson conditional distribution

- Let X and Y be the independent Poisson random variables with means λ and μ.
- The conditional distribution of X given $X + Y = t$ is $Bin(t, p)$ where $p = \frac{\lambda}{\lambda+\mu}$
- This result is very important in many applications

Proof

- Clearly, $P(X = x|X + Y = t) = 0$ if $x > t$
- For any $x \leq t$,

$$
\begin{aligned}
P(X = x|X + Y = t) &= \frac{P(X = x, X + Y = t)}{P(X + Y = t)} \\
&= \frac{P(X = x, Y = t - x)}{P(X + Y = t)} \\
&= \frac{e^{-\lambda}\lambda^x}{x!} \frac{e^{-\mu}\mu^{t-x}}{(t - x)!} \frac{t!}{e^{-(\lambda+\mu)}(\lambda + \mu)!} \\
&= \frac{t!}{x!(t - x)!} \frac{\lambda^x \mu^{t-x}}{(\lambda + \mu)!} \\
&= \binom{t}{x} \left(\frac{\lambda}{\lambda + \mu}\right)^x \left(\frac{\mu}{\lambda + \mu}\right)^{t-x}
\end{aligned}
$$

- In the above, we used the fact that $X + Y \sim Poi(\lambda + \mu)$