Statistics

2023 Lectures
Part 5 - Conditional Distributions

Institute of Economic Studies
Faculty of Social Sciences
Charles University

Conditional distributions

- so far, no additional information was available relating to the the occurrence of an event for a subset of the random variables $(X_1, \ldots, X_n)$

Example 56: Expected number of housing starts (i.e. the number of new residential construction projects that have begun during any particular month), calculated for planning purposes by building supply manufacturer, could depend on a given level of mortgage interest assumed.

- conditional distributions and conditional expectation are at the heart of regression analysis
- relates to an experiment in which n different characteristics were recorded for each outcome and we are specifically interested in analyzing a subset of these characteristics given that a fixed set of possibilities will occur with certainty for the remaining characteristics.

Discrete conditional distributions

- In the bivariate case, the question is to determine $P(X \in Q_1 | Y = y)$ or $P(Y \in Q_2 | X = x)$

a) discrete case

- if $P(Y = y_j) = 0$ then we leave the conditional probability undefined
- if $P(Y = y_j) > 0$ then

$$P(X \in Q_1 | Y = y_j) = \frac{P(X \in Q_1, Y = y_j)}{P(Y = y_j)} = \frac{\sum_{i:x_i \in Q_1} p_{ij}}{\sum_i p_{ij}} = \frac{\sum_{i:x_i \in Q_1} p_{ij}}{p_{\cdot j}}$$

- analogously,

$$P(Y \in Q_2 | X = x_i) = \frac{\sum_{j:y_j \in Q_2} p_{ij}}{\sum_j p_{ij}} = \frac{\sum_{j:y_j \in Q_2} p_{ij}}{p_{i\cdot}}$$

Example 57: Let X_1 and X_2 denote the result of the first and second toss of a die and put $Z = |X_1 - X_2|$. Determine the distribution of $Z|X_1 = 5$.

Continuous conditional distributions

Definition 25: (conditional density)
The conditional densities g_{12} and g_{21} are defined by

$$g_{12}(x|y) = \frac{f(x,y)}{f_Y(y)}, \text{ provided } f_Y(y) > 0,$$
$$g_{21}(y|x) = \frac{f(x,y)}{f_X(x)}, \text{ provided } f_X(x) > 0,$$

where f is the joint density of (X,Y) and f_X, f_Y are marginal densities.

Note that conditional densities are indeed densities:

- nonnegative in the first variable
- $$\int_{-\infty}^{\infty} g_{12}(x|y)\mathrm{d}x = \frac{1}{f_Y(y)} \int_{-\infty}^{\infty} f(x,y)\mathrm{d}x = \frac{f_Y(y)}{f_Y(y)} = 1.$$

Conditional distributions and thinning

Example 58: Let the joint density of (X, Y) be given by

$$f(x,y) = \begin{cases} cxy^2, & 0 \le x \le 1, 0 \le y \le 1, x+y \ge 1; \\ 0, & \text{otherwise.} \end{cases}$$

Determine the conditional distribution of X given $Y = 0.5$.

Example 59: An animal lays a certain number Y of eggs, where Y is a random variable with Poisson distribution with parameter $\lambda > 0$. Each egg hatches with probability $0 < p < 1$, independent of the hatching of other eggs. Determine the distribution of number of eggs X that hatch.

- process of thinning causes some of the Y objects to disappear (not be counted) due to a certain random process of selection. In this case the process is called binomial thinning.

Mixture of distributions

Example 60: Assume Y takes $n \geq 2$ values with $P(Y = y_i) = p_i$, with $\sum_i p_i = 1$. Next, assume that for a given $Y = y_i$, the random variable X has a continuous distribution with density φ_i. Thus,

$$P(a \leq X \leq b | Y = y_i) = \int_a^b \varphi_i(x) \mathrm{d}x.$$

Apply Total probability formula to find the marginal (unconditional) distribution of X:

$$P(a \leq X \leq b) = \sum_{i=1}^{n} P(a \leq X \leq b | Y = y_i) P(Y = y_i) = \int_a^b \sum_{i=1}^{n} p_i \varphi_i(x) \mathrm{d}x$$

where

$$f_X(x) = \sum_{i=1}^{n} p_i \varphi_i(x)$$

is called the mixture of densities φ_i with mixing coefficients p_i.

Conditional cdf and independence

- one can define analogously to the unconditional case the conditional cumulative distribution function. E.g. in the bivariate case:

$$F_X(t|Y \in Q_2) = P(X \leq t|Y \in Q_2).$$

- naturally, in the case of X and Y independent
  - discrete case:

$$P(X = x_i|Y = y_j) = P(X = x_i)$$

  - continuous case

$$g_{12}(x|y) = f_X(x)$$
$$g_{21}(y|x) = f_Y(y)$$

Conditional expectation

Definition 26: (conditional expectation)
For random variables X and Y, the conditional expectation $E(X|Y)$ is defined as the random variable with values $E(X|Y=y)$ when $Y=y$.

More generally, for a real-valued function g,

- in the discrete case, the values

$$E(g(X)|Y=y_j) = \sum_i g(x_i)P(X=x_i|Y=y_j)$$

  occurs with probability $P(Y=y_j)$.
- in the continuous case, the values

$$E(g(X)|Y=y) = \int_{-\infty}^{\infty} g(x)g_{12}(x|y)\mathrm{d}x, y \in \mathbb{R},$$

  are distributed with respect to $f_Y(y), y \in \mathbb{R}$.

Iterated expectation

Example 59 (cont.): Recall, $Y \sim POI(\lambda)$ and $X|Y = n \sim BIN(n,p)$. This means that the expected number of hatched eggs, provided $Y = n$, is $E(X|Y = n) = np$. So we have $E(X|Y) = Yp$.

Theorem 30: (iterated expectation theorem)
For any random variables X and Y and real-valued function g, we have

$$E(E(g(X)|Y)) = E(g(X)),$$

provided $E(g(X))$ exists.

Example 59 (cont.): Since Y has Poisson distribution with mean λ,

$$E[E(X|Y)] = E(Yp) = E(Y)p = \lambda p.$$

This also follows from the fact that $X \sim POI(\lambda p)$.

Iterated variance

Theorem 31: (without proof) For any random variables X and Y, if $E(X^2)$ exists, then

$$VarX = E[Var(X|Y)] + Var[E(X|Y)].$$

Example 59 (cont.): As $X \sim BIN(Y,p)$ and $E(X|Y) = Yp, Var(X|Y) = Yp(1-p)$; and $Y \sim POI(\lambda)$ and $E(Y) = VarY = \lambda$. Thus

$$VarX = E[Yp(1-p)] + Var(Yp) = \lambda p(1-p) + \lambda p^2 = \lambda p.$$

This is in accordance with the fact that $X \sim POI(\lambda p)$.

Substitution theorem

In cases where one is conditioning on an elementary event $Y = y$, there are useful generalizations of the conditional formulas.

Theorem 32: (without proof)(substitution theorem)
For any random variables X and Y and a real-valued mapping g, we have

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

Example 59 (cont.): Consider $Z = Y - X$ which denotes the number of unhatched eggs.

$$E(Z|Y=n) = E(Y-X|Y=n) = E(n-X|Y=n) = \\ = n - E(X|Y=n) = n(1-p).$$

In fact $Z|Y=n \sim BIN(n, 1-p)$ and $Z \sim POI(\lambda(1-p))$.