Probability Calculus

Anna Janicka

lecture X, 16.12.2021

CONDITIONAL EXPECTATION

LINEAR REGRESSION

Plan for Today

1. Conditional expectation
2. Linear regression

Conditional Expectations

1. Intuition
2. Definition in the discrete case:

Let (X, Y) *be a discrete random vector such that* $\mathbb{E}Y$ *exists.*
For any $x \in \mathbb{R}$ *such that* $\mathbb{P}(X = x) > 0$, *we define the* **conditional expected value of variable** Y **given** $X = x$ *as the expected value of a random variable with distribution* $\mu(A) = \mathbb{P}(Y \in A | X = x)$.
That is, if $S_x = \{y \in \mathbb{R} : \mathbb{P}(X = x, Y = y) > 0\}$, *we have* $\mathbb{E}(Y|X = x) = \sum_{y \in S_x} y\mathbb{P}(Y = y|X = x)$.

Conditional Expected Value of discrete RV

3. Example:
   - double 0-1
   - function of X
4. Transformations

Let (X, Y) *be a discrete random vector, and* $\varphi : \mathbb{R} \to \mathbb{R}$ *a Borel function such that* $\mathbb{E}|\varphi(Y)| < \infty$. *We then have that for any* x *such that* $\mathbb{P}(X = x) > 0$*:*
$\mathbb{E}(\varphi(Y)|X = x) = \sum_{y \in S_x} \varphi(y)\mathbb{P}(Y = y|X = x)$,
where $S_x = \{y \in \mathbb{R} : \mathbb{P}(X = x, Y = y) > 0\}$.

Conditional density

5. Definition

Let (X, Y) *be a continuous random vector with density* $g : \mathbb{R}^2 \to [0, \infty)$. *Let* $g_X(x) = \int_{-\infty}^{\infty} g(x, y)dy > 0$ *be the marginal density of* X. *For all* $x \in \mathbb{R}$, *we define* **conditional density** *of variable* Y *given* $X = x$ *as the function*

$$g_{Y|X}(y|x) = \begin{cases} \frac{g(x,y)}{g_X(x)} & \text{if } g_X(x) > 0 \\ f(y) & \text{otherwise,} \end{cases}$$

where $f : \mathbb{R} \to [0, \infty)$ *is any density function of our choice.*

Conditional density – cont.

6. Properties:
   - density
   - corresponds to conditional probability
   - different functions possible
   - OK for independent variables
7. Examples:
   - uniform distribution over square
   - “chain rule”

Conditional Expected value of continuous RV

8. Definition

Let (X,Y) *be a continuous random vector with density* $g:\mathbb{R}^2 \to [0,\infty)$, *such that* $\mathbb{E}|Y| < \infty$. *For all* $x \in \mathbb{R}$ *we define the* **conditional expected value of variable** Y **given** $X = x$ *as the expected value of a random variable with density* $f_x(y) = g_{Y|X}(y|x)$, *i.e.* $\mathbb{E}(Y|X = x) = \int_{-\infty}^{\infty} y g_{Y|X}(y|x)dy$.

9. Example

Conditional Expected value of continuous RV – cont.

10. Transformations:

Let (X, Y) *be a continuous random vector with density* $g : \mathbb{R}^2 \to [0, \infty)$, *and* $\varphi : \mathbb{R} \to \mathbb{R}$ *be a Borel function such that* $\mathbb{E}|\varphi(Y)| < \infty$. *Then, we have that for any* $x \in \mathbb{R}$, $\mathbb{E}(\varphi(Y)|X = x) = \int_{-\infty}^{\infty} \varphi(y) g_{Y|X}(y|x) dy$.

Conditional Expectation

11. General definition of conditional expectation

Let (X,Y) *be a random vector, such that* $\mathbb{E}|Y| < \infty$. *The* **conditional expected value of** Y **given** X, *demoted as* $\mathbb{E}(Y|X)$, *is a random variable such that* $\mathbb{E}(Y|X) = m(X)$, *where* $m(x) = \mathbb{E}(Y|X = x)$.

12. Examples

Properties of Conditional Expectations

13. Properties of expected values

Let $X, Y, Z : \Omega \to \mathbb{R}$ *be random variables such that* $\mathbb{E}|X|, \mathbb{E}|Y| < \infty$. *We have:*

(i) *If* $X \geqslant 0$, *then* $\mathbb{E}(X|Z) \geqslant 0$.

(ii) $|\mathbb{E}(X|Z)| \leqslant \mathbb{E}(|X||Z)$.

(iii) *For any* $a, b \in \mathbb{R}$ *we have*

$$\mathbb{E}(aX + bY|Z) = a\mathbb{E}(X|Z) + b\mathbb{E}(Y|Z).$$

Properties of Conditional Expectations – cont.

13. Specific properties

Let $X, Y : \Omega \to \mathbb{R}$ *be random variables such that* $\mathbb{E}|Y| < \infty$. *We have that*

(i) $\mathbb{E}|\mathbb{E}(Y|X)| < \infty$ *and* $\mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}Y$.

(ii) *If* X *and* Y *are independent, then* $\mathbb{E}(Y|X) = \mathbb{E}Y$.

(iii) *If* $h(X)$ *is a limited random variable, then* $\mathbb{E}(h(X) \cdot Y|X) = h(X)\mathbb{E}(Y|X)$.

Conditional Probability

14. Definition

Let X *be a random variable. For any event* $A \in \mathcal{F}$, *we define*

$$\mathbb{P}(A|X) = \mathbb{E}(1_A|X)$$

Linear regression

1. Best (in terms of average square deviation) **linear** approximation of variable Y with variable X, i.e. aX+b:

minimizes

$$f(a,b) = \mathbb{E}(Y - aX - b)^2$$

solution:

$$a = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X} \qquad b = \mathbb{E}Y - \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}X}\mathbb{E}X$$

$$a = \rho_{X,Y} \cdot \frac{\sigma_Y}{\sigma_X}$$

Conditional Expectation as an approximation

1. Theorem:

Let $X, Y : \Omega \to \mathbb{R}$ *be random variables such that* $\mathbb{E}Y^2 < \infty$. *Then, the function* $\varphi^* : \mathbb{R} \to \mathbb{R}$, *such that* $\varphi^*(x) = \mathbb{E}(Y|X = x)$, *satisfies:*

$$\mathbb{E}(Y - \varphi^*(X))^2$$

$$= \min\{\mathbb{E}(Y - \varphi(X))^2 : \varphi \text{ is a Borel function} : \mathbb{R} \to \mathbb{R}\}.$$