Lecture 10. Conditional probability and expectation

Alex Avdiushenko

Neapolis University Paphos

Bonus-2023

Conditional Probability Distribution

Conditional Probability Distribution

Definition

Let ξ and η be two discrete random variables. Then

$$P(\xi = a \mid \eta = b) = \begin{cases} \frac{P(\xi=a,\ \eta=b)}{P(\eta=b)} \\ 0, \text{ if } P(\eta = b) = 0 \end{cases}$$

Conditional Probability Distribution

Definition

Let ξ and η be two discrete random variables. Then

$$P(\xi = a \mid \eta = b) = \begin{cases} \frac{P(\xi = a,\ \eta = b)}{P(\eta = b)} \\ 0, \text{ if } P(\eta = b) = 0 \end{cases}$$

Conditional Expectation is

$$\mathbb{E}(\xi \mid \eta = b) = \sum_i a_i P(\xi = a_i \mid \eta = b)$$

Definition

Let ξ and η be two discrete random variables. Then

$$P(\xi = a \mid \eta = b) = \begin{cases} \frac{P(\xi=a,\ \eta=b)}{P(\eta=b)} \\ 0, \text{ if } P(\eta = b) = 0 \end{cases}$$

Conditional Expectation is

$$\mathbb{E}(\xi \mid \eta = b) = \sum_i a_i P(\xi = a_i \mid \eta = b)$$

Moreover

$$\sum_b \mathbb{E}(\xi \mid \eta = b) P(\eta = b) = \sum_{a,b} a P(\xi = a, \eta = b) = \sum_a a P(\xi = a) = \mathbb{E}\xi$$

Continuous Distribution

Continuous Distribution

Definition

Let $(\xi, \eta) \sim p_{\xi,\eta}(x, y)$. Then conditional density is

$$p_{\xi|\eta=y}(x) = \frac{p_{\xi,\eta}(x, y)}{p_\eta(y)}$$

Continuous Distribution

Definition

Let $(\xi, \eta) \sim p_{\xi,\eta}(x, y)$. Then conditional density is

$$p_{\xi|\eta=y}(x) = \frac{p_{\xi,\eta}(x, y)}{p_\eta(y)}$$

Conditional Expectation is

$$\mathbb{E}(\xi \mid \eta = y) = \int\limits_{\mathbb{R}} x p_{\xi|\eta=y}(x) dx$$

Continuous Distribution

Definition

Let $(\xi, \eta) \sim p_{\xi,\eta}(x, y)$. Then conditional density is

$$p_{\xi|\eta=y}(x) = \frac{p_{\xi,\eta}(x, y)}{p_\eta(y)}$$

Conditional Expectation is

$$\mathbb{E}(\xi \mid \eta = y) = \int_{\mathbb{R}} x p_{\xi|\eta=y}(x) dx$$

Moreover

$$\mathbb{E}\xi = \int_{\mathbb{R}} x p_{\xi,\eta}(x, y) dx dy = \int_{\mathbb{R}} dy \left(\int_{\mathbb{R}} x p_{\xi,\eta}(x, y) dx \right) = \int_{\mathbb{R}} \mathbb{E}(\xi \mid \eta = y) p_\eta(y) dy =$$

$$= \mathbb{E}_\eta \left(\mathbb{E}(\xi \mid \eta = y) \right)$$

Example

Let (Ω, P) is probability space and $\xi : \Omega \to \mathbb{R}$, $\chi_A : \Omega \to \mathbb{R}$ — indicator of $A \subseteq \Omega$. Find $\mathbb{E}(\xi \mid \chi_A = y)$.

Example

Let (Ω, P) is probability space and $\xi : \Omega \to \mathbb{R}$, $\chi_A : \Omega \to \mathbb{R}$ — indicator of $A \subseteq \Omega$. Find $\mathbb{E}(\xi \mid \chi_A = y)$.

If $X \subseteq \overline{A}$ then

$$P(X \mid \chi_A = 0) = \frac{P(X)}{P(\overline{A})}$$

Example

Let (Ω, P) is probability space and $\xi : \Omega \to \mathbb{R}$, $\chi_A : \Omega \to \mathbb{R}$ — indicator of $A \subseteq \Omega$. Find $\mathbb{E}(\xi \mid \chi_A = y)$.

If $X \subseteq \overline{A}$ then

$$P(X \mid \chi_A = 0) = \frac{P(X)}{P(\overline{A})}$$

So

$$\mathbb{E}(\xi \mid \chi_A = y) = \frac{\mathbb{E}(\xi, \chi_A = y)}{P(\chi_A = y)}$$

Continuous Version of Bayes Formula

Definition

Let $(\xi, \eta) \sim p_{\xi,\eta}(x, y)$. Then conditional density

$$p_{\xi \mid \eta = y}(x) = p(x \mid y) = \frac{p(y \mid x) p(x)}{p(y)}$$

$$P(A \cap B \cap C) = P(A \mid B \cap C) P(B \mid C) P(C)$$

For $\xi_1, \ldots, \xi_n$ we have density $p(x_1, \ldots, x_n)$ and

$$p(x_1 \; x_2 \; \mid \; x_3 \; x_4) = \frac{p(x_1 \; x_2 \; x_3 \; x_4)}{p(x_3 \; x_4)}$$

and hence

$$p(x_1 \; x_2 \; x_3) = p(x_1 \mid x_2 \; x_3) p(x_2 \mid x_3) p(x_3)$$

Task 1

Suppose we have a dice with n faces, results are uniformly distribute. We roll the dice until the sum of all the rolled results is greater than n. So

$$\Omega = \left\{ (a_1, \ldots, a_k) \mid \sum_{i=1}^{k} a_i \geq n, \sum_{i=1}^{k-1} a_i < n \right\}$$

$$P(a_1, \ldots, a_k) = \frac{1}{n^k}$$

$\xi(a_1, \ldots, a_k) = k$ — number of dice rolls, $\mathbb{E}\xi$–?

Task 1

Suppose we have a dice with n faces, results are uniformly distribute. We roll the dice until the sum of all the rolled results is greater than n. So

$$\Omega = \left\{ (a_1, \ldots, a_k) \mid \sum_{i=1}^{k} a_i \geq n, \sum_{i=1}^{k-1} a_i < n \right\}$$

$$P(a_1, \ldots, a_k) = \frac{1}{n^k}$$

$$\xi(a_1, \ldots, a_k) = k \text{ — number of dice rolls, } \mathbb{E}\xi\text{–?}$$

Consider new random variable $\eta_1 =$ result of first dice roll

$$\mathbb{E}\xi = \mathbb{E}_{\eta_1}\left(\mathbb{E}(\xi \mid \eta_1 = a)\right) = \sum_{k=1}^{n} \mathbb{E}(\xi \mid \eta_1 = k) P(\eta_1 = k)$$

Task 1

Suppose we have a dice with n faces, results are uniformly distribute. We roll the dice until the sum of all the rolled results is greater than n. So

$$\Omega = \left\{ (a_1, \ldots, a_k) \mid \sum_{i=1}^{k} a_i \geq n, \sum_{i=1}^{k-1} a_i < n \right\}$$

$$P(a_1, \ldots, a_k) = \frac{1}{n^k}$$

$$\xi(a_1, \ldots, a_k) = k \text{ — number of dice rolls, } \mathbb{E}\xi \text{–?}$$

Consider new random variable $\eta_1 =$ result of first dice roll

$$\mathbb{E}\xi = \mathbb{E}_{\eta_1}\left(\mathbb{E}(\xi \mid \eta_1 = a)\right) = \sum_{k=1}^{n} \mathbb{E}(\xi \mid \eta_1 = k) P(\eta_1 = k)$$

Now consider random variables
$\xi_m =$ number of dice rolls to get m in sum, so $\xi = \xi_n$

$$\mathbb{E}(\xi_n \mid \eta_1 = k) = \mathbb{E}(1 + \xi_{n-k} \mid \eta_1 = k) = \{\text{as } \xi_{n-k} \perp\!\!\!\perp \eta_1\} = \mathbb{E}(1 + \xi_{n-k})$$

$$\Rightarrow \mathbb{E}\xi_m = \frac{\mathbb{E}(1 + \xi_{m-1}) + \mathbb{E}(1 + \xi_{m-2}) + \dots + \mathbb{E}(1 + \xi_{m-n})}{n} =$$

$$= \frac{n + \mathbb{E}\xi_{m-1} \dots + \mathbb{E}\xi_1}{n}$$

$$\mathbb{E}(\xi_n \mid \eta_1 = k) = \mathbb{E}(1 + \xi_{n-k} \mid \eta_1 = k) = \{\text{as } \xi_{n-k} \perp\!\!\!\perp \eta_1\} = \mathbb{E}(1 + \xi_{n-k})$$

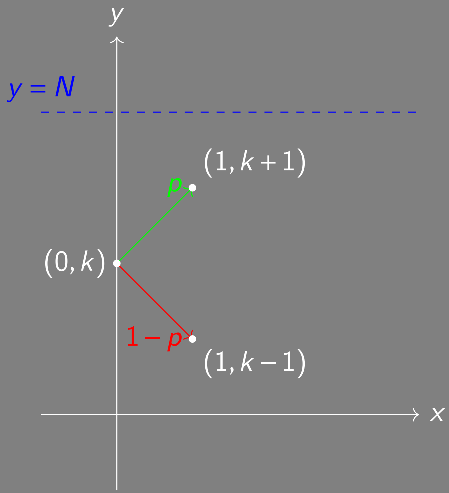
$$\Rightarrow \mathbb{E}\xi_m = \frac{\mathbb{E}(1 + \xi_{m-1}) + \mathbb{E}(1 + \xi_{m-2}) + \cdots + \mathbb{E}(1 + \xi_{m-n})}{n} =$$

$$= \frac{n + \mathbb{E}\xi_{m-1} \cdots + \mathbb{E}\xi_1}{n}$$

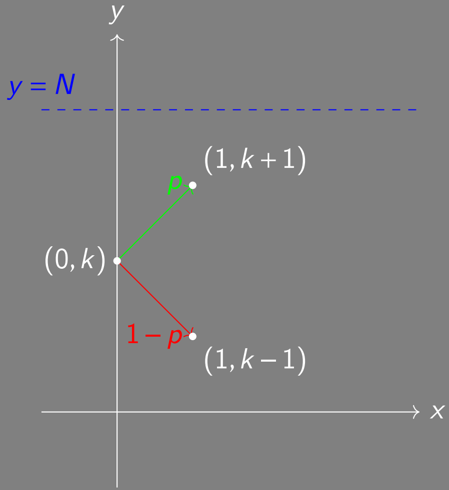
If $a_k = \mathbb{E}\xi_k, \; a_1 = \mathbb{E}\xi_1 = 1$

$$a_k = 1 + \frac{a_1 + \cdots + a_{k-1}}{n} \Rightarrow a_k - a_{k-1} = \frac{a_{k-1}}{n} \Rightarrow a_k = \left(1 + \frac{1}{n}\right) a_{k-1} = \left(1 + \frac{1}{n}\right)^{k-1}$$

Task 2: casino



If you are lucky to score N, then you win and leave. If you lose all the money to zero, then the game also ends.
What is Ω here? All the paths!



If you are lucky to score N, then you win and leave. If you lose all the money to zero, then the game also ends.
What is Ω here? All the paths!

Question

Find the average length of game $\mathbb{E}\xi_k$.

$\mathbb{E}\xi_0 = \mathbb{E}\xi_N = 0$

$$\begin{aligned}
\mathbb{E}\xi_k &= \mathbb{E}(\xi_k \mid \text{win})P(\text{win}) + \mathbb{E}(\xi_k \mid \text{lose})P(\text{lose}) = \\
&= \mathbb{E}(1 + \xi_{k+1} \mid \text{win})p + \mathbb{E}(1 + \xi_{k-1} \mid \text{lose})q = \\
&= 1 + p\mathbb{E}(\xi_{k+1}) + q\mathbb{E}(\xi_{k-1})
\end{aligned}$$

$$\mathbb{E}\xi_0 = \mathbb{E}\xi_N = 0$$

$$\begin{aligned}\mathbb{E}\xi_k &= \mathbb{E}(\xi_k \mid \text{win})P(\text{win}) + \mathbb{E}(\xi_k \mid \text{lose})P(\text{lose}) = \\ &= \mathbb{E}(1 + \xi_{k+1} \mid \text{win})p + \mathbb{E}(1 + \xi_{k-1} \mid \text{lose})q = \\ &= 1 + p\mathbb{E}(\xi_{k+1}) + q\mathbb{E}(\xi_{k-1})\end{aligned}$$

Let column $x = (x_0, \dots, x_N) = (\mathbb{E}\xi_0, \dots, \mathbb{E}\xi_N)$

$$\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ \vdots \\ x_{N-1} \\ x_N \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & \dots & 0 & 0 \\ q & 0 & p & 0 & \dots & 0 \\ 0 & q & 0 & p & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & q & 0 & p \\ 0 & 0 & \dots & \dots & 0 & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ \vdots \\ x_{N-1} \\ x_N \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ \vdots \\ \vdots \\ 1 \\ 0 \end{bmatrix}$$

$\mathbb{E}\xi_0 = \mathbb{E}\xi_N = 0$

$$\begin{aligned}\mathbb{E}\xi_k &= \mathbb{E}(\xi_k \mid \text{win})P(\text{win}) + \mathbb{E}(\xi_k \mid \text{lose})P(\text{lose}) = \\ &= \mathbb{E}(1+\xi_{k+1} \mid \text{win})p + \mathbb{E}(1+\xi_{k-1} \mid \text{lose})q = \\ &= 1 + p\mathbb{E}(\xi_{k+1}) + q\mathbb{E}(\xi_{k-1})\end{aligned}$$

Let column $x = (x_0, \ldots, x_N) = (\mathbb{E}\xi_0, \ldots, \mathbb{E}\xi_N)$

$$\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ \vdots \\ x_{N-1} \\ x_N \end{bmatrix} = \begin{bmatrix} 0 & 0 & \ldots & \ldots & 0 & 0 \\ q & 0 & p & 0 & \ldots & 0 \\ 0 & q & 0 & p & \ldots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & q & 0 & p \\ 0 & 0 & \ldots & \ldots & 0 & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ \vdots \\ x_{N-1} \\ x_N \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ \vdots \\ \vdots \\ 1 \\ 0 \end{bmatrix}$$

$x = Ax + b \Rightarrow \boxed{x = (I - A)^{-1} b}$

Distribution of sum

We can find it only for independent random variables $\xi \perp\!\!\!\perp \eta$, or we must know joint probability distribution.

Distribution of sum

We can find it only for independent random variables $\xi \perp\!\!\!\perp \eta$, or we must know joint probability distribution.
In discrete case

$$P(\xi + \eta = a) = \sum_{u+v=a} P(\xi = u) P(\eta = v)$$

And for integer values we have convolution

$$P(\xi + \eta = k) = \sum_{i \in \mathbb{Z}} p_i q_{k-i}$$

For continuous case we have $\xi \sim p(x), \ \eta \sim q(y)$

$$F_{\xi+\eta}(t) = P(\xi + \eta \leq t) = \int\limits_{x+y\leq t} p(x)q(y)dxdy =$$

$$= \{x + y = u, \ y = v\} = \int\limits_{u\leq t} p(u - v)q(v)\det\begin{bmatrix}1 & -1\\ 0 & 1\end{bmatrix} dudv =$$

$$= \int\limits_{-\infty}^{t} du \int\limits_{\mathbb{R}} p(u - v)q(v)dv$$

$$F'_{\xi+\eta}(t) = p_{\xi+\eta}(t) = \int\limits_{\mathbb{R}} p(t - v)q(v)dv = \int\limits_{\mathbb{R}} p(u)q(t - u)du$$