Discrete Mathematics and Its Applications 2 (CS147)

Lecture 12: Conditional expectation, coupon collector's problem

Fanghui Liu

Department of Computer Science, University of Warwick, UK

Definition of conditional expectation

- similar notation: from conditional probability to conditional expectation

Definition of conditional expectation

- similar notation: from conditional probability to conditional expectation

The simplest version of conditional expectation is conditioned on a single event A.

Definition of conditional expectation

- similar notation: from conditional probability to conditional expectation

The simplest version of conditional expectation is conditioned on a single event A.

Definition (Conditional expectation over an event)

For a discrete random variable X and an event A, the conditional expectation is

$$\mathbb{E}(X|A) = \sum_i x_i \Pr(\{\omega : X(\omega) = x_i\}|A)\,.$$

Definition of conditional expectation

- similar notation: from conditional probability to conditional expectation

The simplest version of conditional expectation is conditioned on a single event A.

Definition (Conditional expectation over an event)

For a discrete random variable X and an event A, the conditional expectation is

$$\mathbb{E}(X|A) = \sum_i x_i \Pr(\{\omega : X(\omega) = x_i\}|A) .$$

Information!

Example: conditional expectation over an event

Example

A dice is repeatedly thrown until it lands on a 6. Let T be the number of rolls it takes for a dice to roll a 6, and let A be the event that all dice rolls in a sequence are even. What is $\mathbb{E}(T|A)$?

Example: conditional expectation over an event

Example

A dice is repeatedly thrown until it lands on a 6. Let T be the number of rolls it takes for a dice to roll a 6, and let A be the event that all dice rolls in a sequence are even. What is $\mathbb{E}(T|A)$?

What is $\mathbb{E}(T)$?

$\Rightarrow$ Geometric distribution: $\Pr(T = k) = (1-p)^{k-1}p, \forall k \geq 1$ with $p = 1/6$.

Example: conditional expectation over an event

Example

A dice is repeatedly thrown until it lands on a 6. Let T be the number of rolls it takes for a dice to roll a 6, and let A be the event that all dice rolls in a sequence are even. What is $\mathbb{E}(T|A)$?

What is $\mathbb{E}(T)$?

$\Rightarrow$ Geometric distribution: $\Pr(T = k) = (1-p)^{k-1}p, \forall k \geq 1$ with $p = 1/6$.

Solution

We know the information the sequence are even, i.e., $2, 4, 6$.
$\mathbb{E}(T|A) \Leftrightarrow$ *find the expected number of throws until the result is 6.*
$\Rightarrow p = 1/3$
$\Rightarrow \mathbb{E}(T|A) = 1/p = 3.$

Conditioning on a random variable

◦ $\mathbb{E}(X|A)$ is a value (A is an event)
◦ $\mathbb{E}(X|Y)$: the expected value of X conditioned on Y is itself a random variable
- when $Y = y$, it takes $\mathbb{E}(X|Y = y)$

Conditioning on a random variable

○ $\mathbb{E}(X|A)$ is a value (A is an event)
○ $\mathbb{E}(X|Y)$: the expected value of X conditioned on Y is itself a random variable
- when $Y = y$, it takes $\mathbb{E}(X|Y = y)$

Informal understanding

- $E(X)$: average - the best estimate of a X given no information about it
- $\mathbb{E}(X|Y)$: we have already known the information from Y, how to give a good estimation for X? a function of Y that best approximates X

Conditioning on a random variable

○ $\mathbb{E}(X|A)$ is a value (A is an event)
○ $\mathbb{E}(X|Y)$: the expected value of X conditioned on Y is itself a random variable
- when $Y = y$, it takes $\mathbb{E}(X|Y = y)$

Informal understanding

- $E(X)$: average - the best estimate of a X given no information about it
- $\mathbb{E}(X|Y)$: we have already known the information from Y, how to give a good estimation for X? a function of Y that best approximates X

Property

- $\mathbb{E}(X|X) = X$
- $\mathbb{E}(X|Y) = \mathbb{E}(X)$ *if* X, Y *are independent.*
- $\mathbb{E}(aX + bY|Z) = a\mathbb{E}(X|Z) + b\mathbb{E}(Y|Z)$ *for two constants* a, b.

Law of total expectation

Theorem

$\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$

Law of total expectation

Theorem

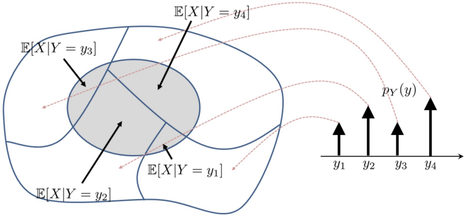
$\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$

Proof.

$$
\begin{aligned}
\mathbb{E}[\mathbb{E}[X|Y]] &= \sum_y \mathbb{E}[X|Y=y]\Pr(Y=y) \\
&= \sum_y \left(\sum_x x\Pr(X=x|Y=y)\right)\Pr(Y=y) \\
&= \sum_y \sum_x x\Pr(X=x, Y=y) \\
&= \sum_x x \sum_y \Pr(X=x, Y=y) = \sum_x x\Pr(X=x).
\end{aligned}
$$

□

total expectation theorem



Examples (I)

Example

Let X and Y be the values of independent six-sides dies. What is $\mathbb{E}(X|X+Y)$?

Examples (I)

Example

Let X and Y be the values of independent six-sides dies. What is $\mathbb{E}(X|X+Y)$?

Intuition: we know the information of $X+Y$ and want to estimate X.

- $\mathbb{E}(X+Y|X+Y) = X+Y$
- $\mathbb{E}(X+Y|X+Y) = \mathbb{E}(X|X+Y) + \mathbb{E}(Y|X+Y) = 2\mathbb{E}(X|X+Y)$ by symmetry

Then we have $\mathbb{E}(X|X+Y) = (X+Y)/2$. $\square$

Examples (II)

Example

We roll two standard 6-sided dice, let X_1 and X_2 be the numbers we obtain and $X = X_1 + X_2$. Compute $\mathbb{E}[X_1|X = 8]$.

Solution

$$\mathbb{E}[X_1|X=8] = \sum_{i=1}^{6} i\Pr(X_1 = i|X=8) = \sum_{i=2}^{6} i\Pr(X_1 = i|X=8)\,,$$

where $X_1 \neq 1$ for the condition $X = X_1 + X_2 = 8$. Otherwise $X_2 = 7$, which is unrealistic.

Examples (II)

Example

We roll two standard 6-sided dice, let X_1 and X_2 be the numbers we obtain and $X = X_1 + X_2$. Compute $\mathbb{E}[X_1|X = 8]$.

Solution

$$\mathbb{E}[X_1|X = 8] = \sum_{i=1}^{6} i\Pr(X_1 = i|X = 8) = \sum_{i=2}^{6} i\Pr(X_1 = i|X = 8)\,,$$

where $X_1 \neq 1$ for the condition $X = X_1 + X_2 = 8$. Otherwise $X_2 = 7$, which is unrealistic. Then, the event $\{X_1 = i|X = 8\}$ for any $i \in \{2, 3, 4, 5, 6\}$ is an equal-probability event, so

$$\mathbb{E}[X_1|X = 8] = \sum_{i=1}^{6} i\Pr(X_1 = i|X = 8) = \sum_{i=2}^{6} i\Pr(X_1 = i|X = 8) = \frac{1}{5}(2 + 3 + 4 + 5 + 6) = 4\,.$$

Coupon collector's problem

Problem

We repeatedly sample from a set of N distinct objects until at least one copy of each distinct object is obtained. Denote T as the number of draws until the every $\{1, 2, \cdots, N\}$ is seen, what is $\mathbb{E}(T)$?

Coupon collector's problem

Problem

We repeatedly sample from a set of N distinct objects until at least one copy of each distinct object is obtained. Denote T as the number of draws until the every $\{1, 2, \cdots, N\}$ is seen, what is $\mathbb{E}(T)$?

Recall Geometric distribution:

- Fails in the first $n-1$ times

$$\Pr(X = n) = (1-p)^{n-1} p$$

- Success at the n-th time

Illustration

- Sample a new object: ✓
- Sample a repeated object: ✗

$\Rightarrow$ success to sample a new object before previous (repeated) objects
$\Rightarrow$ the first occurrence of success for a new object

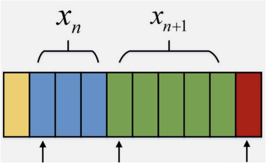
example with 5 coupons: [3] [1] [3] [5] [5] [3] [1] [1] [3] [2] [3] [2] [2] [4]

Illustration

- Sample a new object: ✓
- Sample a repeated object: ✗

$\Rightarrow$ success to sample a new object before previous (repeated) objects
$\Rightarrow$ the first occurrence of success for a new object

example with 5 coupons: [3] [1] [3] [5] [5] [3] [1] [1] [3] [2] [3] [2] [2] [4]



Solution

- $\Pr(\text{find the first unique coupon}) = \frac{N}{N} = 1$
- $\Pr(\text{find the second unique coupon}) = \frac{N-1}{N}$
- $\Pr(\text{find the } n\text{-th unique coupon}) = \frac{N-(n-1)}{N}$

Solution

- $\Pr(\text{find the first unique coupon}) = \frac{N}{N} = 1$
- $\Pr(\text{find the second unique coupon}) = \frac{N-1}{N}$
- $\Pr(\text{find the } n\text{-th unique coupon}) = \frac{N-(n-1)}{N}$

Solution

Let x_n as how many times we need to collect the n-th unique coupon after collecting $(n-1)$-th unique coupons.

$\Rightarrow x_n \sim \mathit{Geo}(p_n)$.

$\Rightarrow p_n = \frac{N-n+1}{N}$.

$$\mathbb{E}(T) = \mathbb{E}(\sum_{n=1}^{N} x_n) = \sum_{n=1}^{N} \mathbb{E}(x_n) = \frac{N}{N} + \frac{N}{N-1} + \ldots + \frac{N}{1}$$
$$= N(1 + \frac{1}{2} + \ldots + \frac{1}{N}) = NH_N \approx N(\log N + \gamma)\,.$$

Next lecture...

Tail probability: We wish to create an upper bound R such that T exceed R with a low probability

$$\Pr(T \geq R) \leq \textcolor{red}{small}\,.$$