18.440: Lecture 8

Discrete random variables

Scott Sheffield

MIT

Outline

Outline

Defining random variables

Probability mass function and distribution function

Recursions

Random variables

- A random variable X is a function from the state space to the real numbers.

Random variables

- A random variable X is a function from the state space to the real numbers.
- Can interpret X as a quantity whose value depends on the outcome of an experiment.

Random variables

- A random variable X is a function from the state space to the real numbers.
- Can interpret X as a quantity whose value depends on the outcome of an experiment.
- Example: toss n coins (so state space consists of the set of all 2^n possible coin sequences) and let X be number of heads.

Random variables

- A random variable X is a function from the state space to the real numbers.
- Can interpret X as a quantity whose value depends on the outcome of an experiment.
- Example: toss n coins (so state space consists of the set of all 2^n possible coin sequences) and let X be number of heads.
- Question: What is $P\{X = k\}$ in this case?

Random variables

- A random variable X is a function from the state space to the real numbers.
- Can interpret X as a quantity whose value depends on the outcome of an experiment.
- Example: toss n coins (so state space consists of the set of all 2^n possible coin sequences) and let X be number of heads.
- Question: What is $P\{X = k\}$ in this case?
- Answer: $\binom{n}{k}/2^n$, if $k \in \{0, 1, 2, \dots, n\}$.

Independence of multiple events

- In n coin toss example, knowing the values of some coin tosses tells us nothing about the others.

Independence of multiple events

- In n coin toss example, knowing the values of some coin tosses tells us nothing about the others.
- Say $E_1 \ldots E_n$ are independent if for each $\{i_1, i_2, \ldots, i_k\} \subset \{1, 2, \ldots n\}$ we have $P(E_{i_1} E_{i_2} \ldots E_{i_k}) = P(E_{i_1}) P(E_{i_2}) \ldots P(E_{i_k})$.

Independence of multiple events

- In n coin toss example, knowing the values of some coin tosses tells us nothing about the others.
- Say $E_1 \ldots E_n$ are independent if for each $\{i_1, i_2, \ldots, i_k\} \subset \{1, 2, \ldots n\}$ we have $P(E_{i_1} E_{i_2} \ldots E_{i_k}) = P(E_{i_1}) P(E_{i_2}) \ldots P(E_{i_k})$.
- In other words, the product rule works.

Independence of multiple events

- In n coin toss example, knowing the values of some coin tosses tells us nothing about the others.
- Say $E_1 \dots E_n$ are independent if for each $\{i_1, i_2, \dots, i_k\} \subset \{1, 2, \dots n\}$ we have $P(E_{i_1} E_{i_2} \dots E_{i_k}) = P(E_{i_1}) P(E_{i_2}) \dots P(E_{i_k})$.
- In other words, the product rule works.
- Independence implies $P(E_1 E_2 E_3 | E_4 E_5 E_6) = \frac{P(E_1)P(E_2)P(E_3)P(E_4)P(E_5)P(E_6)}{P(E_4)P(E_5)P(E_6)} = P(E_1 E_2 E_3)$, and other similar statements.

Independence of multiple events

- In n coin toss example, knowing the values of some coin tosses tells us nothing about the others.
- Say $E_1 \dots E_n$ are independent if for each $\{i_1, i_2, \dots, i_k\} \subset \{1, 2, \dots n\}$ we have $P(E_{i_1} E_{i_2} \dots E_{i_k}) = P(E_{i_1}) P(E_{i_2}) \dots P(E_{i_k})$.
- In other words, the product rule works.
- Independence implies $P(E_1 E_2 E_3 | E_4 E_5 E_6) = \frac{P(E_1)P(E_2)P(E_3)P(E_4)P(E_5)P(E_6)}{P(E_4)P(E_5)P(E_6)} = P(E_1 E_2 E_3)$, and other similar statements.
- Does pairwise independence imply independence?

Independence of multiple events

- In n coin toss example, knowing the values of some coin tosses tells us nothing about the others.
- Say $E_1 \dots E_n$ are independent if for each $\{i_1, i_2, \dots, i_k\} \subset \{1, 2, \dots n\}$ we have $P(E_{i_1} E_{i_2} \dots E_{i_k}) = P(E_{i_1}) P(E_{i_2}) \dots P(E_{i_k})$.
- In other words, the product rule works.
- Independence implies $P(E_1 E_2 E_3 | E_4 E_5 E_6) = \frac{P(E_1)P(E_2)P(E_3)P(E_4)P(E_5)P(E_6)}{P(E_4)P(E_5)P(E_6)} = P(E_1 E_2 E_3)$, and other similar statements.
- Does pairwise independence imply independence?
- No. Consider these three events: first coin heads, second coin heads, odd number heads. Pairwise independent, not independent.

Examples

- Shuffle n cards, and let X be the position of the jth card. State space consists of all $n!$ possible orderings. X takes values in $\{1, 2, \ldots, n\}$ depending on the ordering.

Examples

- Shuffle n cards, and let X be the position of the jth card. State space consists of all $n!$ possible orderings. X takes values in $\{1, 2, \ldots, n\}$ depending on the ordering.
- Question: What is $P\{X = k\}$ in this case?

Examples

- Shuffle n cards, and let X be the position of the jth card. State space consists of all $n!$ possible orderings. X takes values in $\{1, 2, \ldots, n\}$ depending on the ordering.
- Question: What is $P\{X = k\}$ in this case?
- Answer: $1/n$, if $k \in \{1, 2, \ldots, n\}$.

Examples

- Shuffle n cards, and let X be the position of the jth card. State space consists of all $n!$ possible orderings. X takes values in $\{1, 2, \ldots, n\}$ depending on the ordering.
- Question: What is $P\{X = k\}$ in this case?
- Answer: $1/n$, if $k \in \{1, 2, \ldots, n\}$.
- Now say we roll three dice and let Y be sum of the values on the dice. What is $P\{Y = 5\}$?

Indicators

- Given any event E, can define an **indicator** random variable, i.e., let X be random variable equal to 1 on the event E and 0 otherwise. Write this as $X = 1_E$.

Indicators

- Given any event E, can define an **indicator** random variable, i.e., let X be random variable equal to 1 on the event E and 0 otherwise. Write this as $X = 1_E$.
- The value of 1_E (either 1 or 0) *indicates* whether the event has occurred.

Indicators

- Given any event E, can define an **indicator** random variable, i.e., let X be random variable equal to 1 on the event E and 0 otherwise. Write this as $X = 1_E$.
- The value of 1_E (either 1 or 0) *indicates* whether the event has occurred.
- If $E_1, E_2, \dots E_k$ are events then $X = \sum_{i=1}^k 1_{E_i}$ is the number of these events that occur.

Indicators

- Given any event E, can define an **indicator** random variable, i.e., let X be random variable equal to 1 on the event E and 0 otherwise. Write this as $X = 1_E$.
- The value of 1_E (either 1 or 0) *indicates* whether the event has occurred.
- If $E_1, E_2, \ldots E_k$ are events then $X = \sum_{i=1}^k 1_{E_i}$ is the number of these events that occur.
- Example: in n-hat shuffle problem, let E_i be the event ith person gets own hat.

Indicators

- Given any event E, can define an **indicator** random variable, i.e., let X be random variable equal to 1 on the event E and 0 otherwise. Write this as $X = 1_E$.
- The value of 1_E (either 1 or 0) *indicates* whether the event has occurred.
- If $E_1, E_2, \ldots E_k$ are events then $X = \sum_{i=1}^{k} 1_{E_i}$ is the number of these events that occur.
- Example: in n-hat shuffle problem, let E_i be the event ith person gets own hat.
- Then $\sum_{i=1}^{n} 1_{E_i}$ is total number of people who get own hats.

Indicators

- Given any event E, can define an **indicator** random variable, i.e., let X be random variable equal to 1 on the event E and 0 otherwise. Write this as $X = 1_E$.
- The value of 1_E (either 1 or 0) *indicates* whether the event has occurred.
- If $E_1, E_2, \ldots E_k$ are events then $X = \sum_{i=1}^{k} 1_{E_i}$ is the number of these events that occur.
- Example: in n-hat shuffle problem, let E_i be the event ith person gets own hat.
- Then $\sum_{i=1}^{n} 1_{E_i}$ is total number of people who get own hats.
- Writing random variable as sum of indicators: frequently useful, sometimes confusing.

Outline

Outline

Defining random variables

Probability mass function and distribution function

Recursions

Probability mass function

- Say X is a **discrete** random variable if (with probability one) it takes one of a countable set of values.

Probability mass function

- Say X is a **discrete** random variable if (with probability one) it takes one of a countable set of values.
- For each a in this countable set, write $p(a) := P\{X = a\}$. Call p the **probability mass function**.

Cumulative distribution function

- Write $F(a) = P\{X \leq a\} = \sum_{x \leq a} p(x)$.

Cumulative distribution function

- Write $F(a) = P\{X \leq a\} = \sum_{x \leq a} p(x)$.
- Example: Let $T_1, T_2, T_3, \dots$ be sequence of independent fair coin tosses (each taking values in $\{H, T\}$) and let X be the smallest j for which $T_j = H$.

Cumulative distribution function

- Write $F(a) = P\{X \leq a\} = \sum_{x \leq a} p(x)$.
- Example: Let $T_1, T_2, T_3, \ldots$ be sequence of independent fair coin tosses (each taking values in $\{H, T\}$) and let X be the smallest j for which $T_j = H$.
- What is $p(k) = P\{X = k\}$ (for $k \in \mathbb{Z}$) in this case?

Cumulative distribution function

- Write $F(a) = P\{X \leq a\} = \sum_{x \leq a} p(x)$.
- Example: Let $T_1, T_2, T_3, \ldots$ be sequence of independent fair coin tosses (each taking values in $\{H, T\}$) and let X be the smallest j for which $T_j = H$.
- What is $p(k) = P\{X = k\}$ (for $k \in \mathbb{Z}$) in this case?
- What is F?

Another example

- Another example: let X be non-negative integer such that $p(k) = P\{X = k\} = e^{-\lambda}\lambda^k/k!$.

Another example

- Another example: let X be non-negative integer such that $p(k) = P\{X = k\} = e^{-\lambda}\lambda^k/k!$.
- Recall Taylor expansion $\sum_{k=0}^{\infty} \lambda^k/k! = e^{\lambda}$.

Another example

- Another example: let X be non-negative integer such that $p(k) = P\{X = k\} = e^{-\lambda}\lambda^k/k!$.
- Recall Taylor expansion $\sum_{k=0}^{\infty} \lambda^k/k! = e^{\lambda}$.
- In this example, X is called a **Poisson** random variable with intensity λ.

Another example

- Another example: let X be non-negative integer such that $p(k) = P\{X = k\} = e^{-\lambda}\lambda^k/k!$.
- Recall Taylor expansion $\sum_{k=0}^{\infty} \lambda^k/k! = e^{\lambda}$.
- In this example, X is called a **Poisson** random variable with intensity λ.
- Question: what is the state space in this example?

Another example

- Another example: let X be non-negative integer such that $p(k) = P\{X = k\} = e^{-\lambda}\lambda^k/k!$.
- Recall Taylor expansion $\sum_{k=0}^{\infty} \lambda^k/k! = e^{\lambda}$.
- In this example, X is called a **Poisson** random variable with intensity λ.
- Question: what is the state space in this example?
- Answer: Didn't specify. One possibility would be to define state space as $S = \{0, 1, 2, \dots\}$ and define X (as a function on S) by $X(j) = j$. The probability function would be determined by $P(S) = \sum_{k \in S} e^{-\lambda}\lambda^k/k!$.

Another example

- Another example: let X be non-negative integer such that $p(k) = P\{X = k\} = e^{-\lambda}\lambda^k/k!$.
- Recall Taylor expansion $\sum_{k=0}^{\infty} \lambda^k/k! = e^{\lambda}$.
- In this example, X is called a **Poisson** random variable with intensity λ.
- Question: what is the state space in this example?
- Answer: Didn't specify. One possibility would be to define state space as $S = \{0, 1, 2, \ldots\}$ and define X (as a function on S) by $X(j) = j$. The probability function would be determined by $P(S) = \sum_{k \in S} e^{-\lambda}\lambda^k/k!$.
- Are there other choices of S and P — and other functions X from S to P — for which the values of $P\{X = k\}$ are the same?

Another example

- Another example: let X be non-negative integer such that $p(k) = P\{X = k\} = e^{-\lambda}\lambda^k/k!$.
- Recall Taylor expansion $\sum_{k=0}^{\infty} \lambda^k/k! = e^{\lambda}$.
- In this example, X is called a **Poisson** random variable with intensity λ.
- Question: what is the state space in this example?
- Answer: Didn't specify. One possibility would be to define state space as $S = \{0, 1, 2, \dots\}$ and define X (as a function on S) by $X(j) = j$. The probability function would be determined by $P(S) = \sum_{k \in S} e^{-\lambda}\lambda^k/k!$.
- Are there other choices of S and P — and other functions X from S to P — for which the values of $P\{X = k\}$ are the same?
- Yes. "X is a Poisson random variable with intensity λ" is statement only about the *probability mass function* of X.

Outline

Defining random variables

Probability mass function and distribution function

Recursions

Outline

Defining random variables

Probability mass function and distribution function

Recursions

Using Bayes' rule to set up recursions

- Gambler one has integer m dollars, gambler two has integer n dollars. Take turns making one dollar bets until one runs out of money. What is probability first gambler runs out of money first?

Using Bayes' rule to set up recursions

- Gambler one has integer m dollars, gambler two has integer n dollars. Take turns making one dollar bets until one runs out of money. What is probability first gambler runs out of money first?
- **Gambler's ruin:** what if gambler one has an unlimited amount of money?

Using Bayes' rule to set up recursions

- Gambler one has integer m dollars, gambler two has integer n dollars. Take turns making one dollar bets until one runs out of money. What is probability first gambler runs out of money first?
- **Gambler's ruin:** what if gambler one has an unlimited amount of money?
- **Problem of points:** in sequence of independent fair coin tosses, what is probability $P_{n,m}$ to see n heads before seeing m tails?

Using Bayes' rule to set up recursions

- Gambler one has integer m dollars, gambler two has integer n dollars. Take turns making one dollar bets until one runs out of money. What is probability first gambler runs out of money first?
- **Gambler's ruin:** what if gambler one has an unlimited amount of money?
- **Problem of points:** in sequence of independent fair coin tosses, what is probability $P_{n,m}$ to see n heads before seeing m tails?
- Observe: $P_{n,m}$ is equivalent to the probability of having n or more heads in first $m+n-1$ trials.

Using Bayes' rule to set up recursions

- Gambler one has integer m dollars, gambler two has integer n dollars. Take turns making one dollar bets until one runs out of money. What is probability first gambler runs out of money first?
- **Gambler's ruin:** what if gambler one has an unlimited amount of money?
- **Problem of points:** in sequence of independent fair coin tosses, what is probability $P_{n,m}$ to see n heads before seeing m tails?
- Observe: $P_{n,m}$ is equivalent to the probability of having n or more heads in first $m+n-1$ trials.
- Probability of exactly n heads in $m+n-1$ trials is $\binom{m+n-1}{n}$.

Using Bayes' rule to set up recursions

- Gambler one has integer m dollars, gambler two has integer n dollars. Take turns making one dollar bets until one runs out of money. What is probability first gambler runs out of money first?
- **Gambler's ruin:** what if gambler one has an unlimited amount of money?
- **Problem of points:** in sequence of independent fair coin tosses, what is probability $P_{n,m}$ to see n heads before seeing m tails?
- Observe: $P_{n,m}$ is equivalent to the probability of having n or more heads in first $m+n-1$ trials.
- Probability of exactly n heads in $m+n-1$ trials is $\binom{m+n-1}{n}$.
- Famous correspondence by Fermat and Pascal. Led Pascal to write *Le Triangle Arithmétique*.