18.600: Lecture 9

Expectations of discrete random variables

Scott Sheffield

MIT

Outline

Defining expectation

Functions of random variables

Motivation

Outline

Defining expectation

Functions of random variables

Motivation

Expectation of a discrete random variable

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

Expectation of a discrete random variable

- Recall: 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.

Expectation of a discrete random variable

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

Expectation of a discrete random variable

- Recall: 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.
- 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**.

Expectation of a discrete random variable

- Recall: 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.
- 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**.
- The **expectation** of X, written $E[X]$, is defined by

$$E[X] = \sum_{x:p(x)>0} xp(x).$$

Expectation of a discrete random variable

- Recall: 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.
- 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**.
- The **expectation** of X, written $E[X]$, is defined by

$$E[X] = \sum_{x:p(x)>0} xp(x).$$

- Represents weighted average of possible values X can take, each value being weighted by its probability.

Simple examples

- Suppose that a random variable X satisfies $P\{X = 1\} = .5$, $P\{X = 2\} = .25$ and $P\{X = 3\} = .25$.

Simple examples

- Suppose that a random variable X satisfies $P\{X = 1\} = .5$, $P\{X = 2\} = .25$ and $P\{X = 3\} = .25$.
- What is $E[X]$?

Simple examples

- Suppose that a random variable X satisfies $P\{X = 1\} = .5$, $P\{X = 2\} = .25$ and $P\{X = 3\} = .25$.
- What is $E[X]$?
- Answer: $.5 \times 1 + .25 \times 2 + .25 \times 3 = 1.75$.

Simple examples

- Suppose that a random variable X satisfies $P\{X = 1\} = .5$, $P\{X = 2\} = .25$ and $P\{X = 3\} = .25$.
- What is $E[X]$?
- Answer: $.5 \times 1 + .25 \times 2 + .25 \times 3 = 1.75$.
- Suppose $P\{X = 1\} = p$ and $P\{X = 0\} = 1 - p$. Then what is $E[X]$?

Simple examples

- Suppose that a random variable X satisfies $P\{X = 1\} = .5$, $P\{X = 2\} = .25$ and $P\{X = 3\} = .25$.
- What is $E[X]$?
- Answer: $.5 \times 1 + .25 \times 2 + .25 \times 3 = 1.75$.
- Suppose $P\{X = 1\} = p$ and $P\{X = 0\} = 1 - p$. Then what is $E[X]$?
- Answer: p.

Simple examples

- Suppose that a random variable X satisfies $P\{X = 1\} = .5$, $P\{X = 2\} = .25$ and $P\{X = 3\} = .25$.
- What is $E[X]$?
- Answer: $.5 \times 1 + .25 \times 2 + .25 \times 3 = 1.75$.
- Suppose $P\{X = 1\} = p$ and $P\{X = 0\} = 1 - p$. Then what is $E[X]$?
- Answer: p.
- Roll a standard six-sided die. What is the expectation of number that comes up?

Simple examples

- Suppose that a random variable X satisfies $P\{X = 1\} = .5$, $P\{X = 2\} = .25$ and $P\{X = 3\} = .25$.
- What is $E[X]$?
- Answer: $.5 \times 1 + .25 \times 2 + .25 \times 3 = 1.75$.
- Suppose $P\{X = 1\} = p$ and $P\{X = 0\} = 1 - p$. Then what is $E[X]$?
- Answer: p.
- Roll a standard six-sided die. What is the expectation of number that comes up?
- Answer: $\frac{1}{6}1 + \frac{1}{6}2 + \frac{1}{6}3 + \frac{1}{6}4 + \frac{1}{6}5 + \frac{1}{6}6 = \frac{21}{6} = 3.5$.

Expectation when state space is countable

- If the state space S is countable, we can give **SUM OVER STATE SPACE** definition of expectation:

$$E[X] = \sum_{s \in S} P\{s\}X(s).$$

Expectation when state space is countable

- If the state space S is countable, we can give **SUM OVER STATE SPACE** definition of expectation:

$$E[X] = \sum_{s \in S} P\{s\} X(s).$$

- Compare this to the **SUM OVER POSSIBLE** X **VALUES** definition we gave earlier:

$$E[X] = \sum_{x : p(x) > 0} x p(x).$$

Expectation when state space is countable

- If the state space S is countable, we can give **SUM OVER STATE SPACE** definition of expectation:

$$E[X] = \sum_{s \in S} P\{s\} X(s).$$

- Compare this to the **SUM OVER POSSIBLE X VALUES** definition we gave earlier:

$$E[X] = \sum_{x : p(x) > 0} x p(x).$$

- Example: toss two coins. If X is the number of heads, what is $E[X]$?

Expectation when state space is countable

- If the state space S is countable, we can give **SUM OVER STATE SPACE** definition of expectation:

$$E[X] = \sum_{s \in S} P\{s\} X(s).$$

- Compare this to the **SUM OVER POSSIBLE** X **VALUES** definition we gave earlier:

$$E[X] = \sum_{x:p(x)>0} x p(x).$$

- Example: toss two coins. If X is the number of heads, what is $E[X]$?
- State space is $\{(H,H),(H,T),(T,H),(T,T)\}$ and summing over state space gives $E[X] = \frac{1}{4}2 + \frac{1}{4}1 + \frac{1}{4}1 + \frac{1}{4}0 = 1$.

A technical point

- If the state space S is countable, is it possible that the sum $E[X] = \sum_{s \in S} P(\{s\})X(s)$ somehow depends on the order in which $s \in S$ are enumerated?

A technical point

- If the state space S is countable, is it possible that the sum $E[X] = \sum_{s \in S} P(\{s\})X(s)$ somehow depends on the order in which $s \in S$ are enumerated?
- In principle, yes... We only say expectation is defined when $\sum_{s \in S} P(\{x\})|X(s)| < \infty$, in which case it turns out that the sum does not depend on the order.

Outline

Defining expectation

Functions of random variables

Motivation

Outline

Defining expectation

Functions of random variables

Motivation

Expectation of a function of a random variable

- If X is a random variable and g is a function from the real numbers to the real numbers then $g(X)$ is also a random variable.

Expectation of a function of a random variable

- If X is a random variable and g is a function from the real numbers to the real numbers then $g(X)$ is also a random variable.
- How can we compute $E[g(X)]$?

Expectation of a function of a random variable

- If X is a random variable and g is a function from the real numbers to the real numbers then $g(X)$ is also a random variable.
- How can we compute $E[g(X)]$?
- **SUM OVER STATE SPACE:**

$$E[g(X)] = \sum_{s \in S} P(\{s\})g(X(s)).$$

Expectation of a function of a random variable

- If X is a random variable and g is a function from the real numbers to the real numbers then $g(X)$ is also a random variable.
- How can we compute $E[g(X)]$?
- **SUM OVER STATE SPACE:**

$$E[g(X)] = \sum_{s \in S} P(\{s\}) g(X(s)).$$

- **SUM OVER X VALUES:**

$$E[g(X)] = \sum_{x : p(x) > 0} g(x) p(x).$$

Expectation of a function of a random variable

- If X is a random variable and g is a function from the real numbers to the real numbers then $g(X)$ is also a random variable.
- How can we compute $E[g(X)]$?
- **SUM OVER STATE SPACE:**

$$E[g(X)] = \sum_{s \in S} P(\{s\})g(X(s)).$$

- **SUM OVER X VALUES:**

$$E[g(X)] = \sum_{x: p(x) > 0} g(x)p(x).$$

- Suppose that constants a, b, μ are given and that $E[X] = \mu$.

Expectation of a function of a random variable

- If X is a random variable and g is a function from the real numbers to the real numbers then $g(X)$ is also a random variable.
- How can we compute $E[g(X)]$?
- **SUM OVER STATE SPACE:**

$$E[g(X)] = \sum_{s \in S} P(\{s\}) g(X(s)).$$

- **SUM OVER X VALUES:**

$$E[g(X)] = \sum_{x:p(x)>0} g(x)p(x).$$

- Suppose that constants a, b, μ are given and that $E[X] = \mu$.
- What is $E[X + b]$?

Expectation of a function of a random variable

- If X is a random variable and g is a function from the real numbers to the real numbers then $g(X)$ is also a random variable.
- How can we compute $E[g(X)]$?
- **SUM OVER STATE SPACE:**

$$E[g(X)] = \sum_{s \in S} P(\{s\}) g(X(s)).$$

- **SUM OVER X VALUES:**

$$E[g(X)] = \sum_{x : p(x) > 0} g(x) p(x).$$

- Suppose that constants a, b, μ are given and that $E[X] = \mu$.
- What is $E[X + b]$?
- How about $E[aX]$?

Expectation of a function of a random variable

- If X is a random variable and g is a function from the real numbers to the real numbers then $g(X)$ is also a random variable.
- How can we compute $E[g(X)]$?
- **SUM OVER STATE SPACE:**

$$E[g(X)] = \sum_{s \in S} P(\{s\})g(X(s)).$$

- **SUM OVER X VALUES:**

$$E[g(X)] = \sum_{x:p(x)>0} g(x)p(x).$$

- Suppose that constants a, b, μ are given and that $E[X] = \mu$.
- What is $E[X + b]$?
- How about $E[aX]$?
- Generally, $E[aX + b] = aE[X] + b = a\mu + b$.

More examples

- Let X be the number that comes up when you roll a standard six-sided die. What is $E[X^2]$?

More examples

- Let X be the number that comes up when you roll a standard six-sided die. What is $E[X^2]$?
- $\frac{1}{6}(1 + 4 + 9 + 16 + 25 + 36) = 91/6$

More examples

- Let X be the number that comes up when you roll a standard six-sided die. What is $E[X^2]$?
- $\frac{1}{6}(1 + 4 + 9 + 16 + 25 + 36) = 91/6$
- Let X_j be 1 if the jth coin toss is heads and 0 otherwise. What is the expectation of $X = \sum_{i=1}^{n} X_j$?

More examples

- Let X be the number that comes up when you roll a standard six-sided die. What is $E[X^2]$?
- $\frac{1}{6}(1 + 4 + 9 + 16 + 25 + 36) = 91/6$
- Let X_j be 1 if the jth coin toss is heads and 0 otherwise. What is the expectation of $X = \sum_{i=1}^{n} X_j$?
- Can compute this directly as $\sum_{k=0}^{n} P\{X = k\}k$.

More examples

- Let X be the number that comes up when you roll a standard six-sided die. What is $E[X^2]$?
- $\frac{1}{6}(1 + 4 + 9 + 16 + 25 + 36) = 91/6$
- Let X_j be 1 if the jth coin toss is heads and 0 otherwise. What is the expectation of $X = \sum_{i=1}^{n} X_j$?
- Can compute this directly as $\sum_{k=0}^{n} P\{X = k\}k$.
- Alternatively, use symmetry. Expected number of heads should be same as expected number of tails.

More examples

- Let X be the number that comes up when you roll a standard six-sided die. What is $E[X^2]$?
- $\frac{1}{6}(1+4+9+16+25+36) = 91/6$
- Let X_j be 1 if the jth coin toss is heads and 0 otherwise. What is the expectation of $X = \sum_{i=1}^{n} X_j$?
- Can compute this directly as $\sum_{k=0}^{n} P\{X = k\}k$.
- Alternatively, use symmetry. Expected number of heads should be same as expected number of tails.
- This implies $E[X] = E[n - X]$. Applying $E[aX + b] = aE[X] + b$ formula (with $a = -1$ and $b = n$), we obtain $E[X] = n - E[X]$ and conclude that $E[X] = n/2$.

Additivity of expectation

- If X and Y are distinct random variables, then can one say that $E[X+Y] = E[X] + E[Y]$?

Additivity of expectation

- If X and Y are distinct random variables, then can one say that $E[X+Y] = E[X] + E[Y]$?
- Yes. In fact, for real constants a and b, we have $E[aX+bY] = aE[X] + bE[Y]$.

Additivity of expectation

- If X and Y are distinct random variables, then can one say that $E[X+Y] = E[X] + E[Y]$?
- Yes. In fact, for real constants a and b, we have $E[aX + bY] = aE[X] + bE[Y]$.
- This is called the **linearity of expectation**.

Additivity of expectation

- If X and Y are distinct random variables, then can one say that $E[X+Y] = E[X] + E[Y]$?
- Yes. In fact, for real constants a and b, we have $E[aX + bY] = aE[X] + bE[Y]$.
- This is called the **linearity of expectation**.
- Another way to state this fact: given sample space S and probability measure P, the expectation $E[\cdot]$ is a **linear** real-valued function on the space of random variables.

Additivity of expectation

- If X and Y are distinct random variables, then can one say that $E[X+Y] = E[X] + E[Y]$?
- Yes. In fact, for real constants a and b, we have $E[aX + bY] = aE[X] + bE[Y]$.
- This is called the **linearity of expectation**.
- Another way to state this fact: given sample space S and probability measure P, the expectation $E[\cdot]$ is a **linear** real-valued function on the space of random variables.
- Can extend to more variables $E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n]$.

More examples

- Now can we compute expected number of people who get own hats in n hat shuffle problem?

More examples

- Now can we compute expected number of people who get own hats in n hat shuffle problem?
- Let X_i be 1 if ith person gets own hat and zero otherwise.

More examples

- Now can we compute expected number of people who get own hats in n hat shuffle problem?
- Let X_i be 1 if ith person gets own hat and zero otherwise.
- What is $E[X_i]$, for $i \in \{1, 2, \ldots, n\}$?

More examples

- Now can we compute expected number of people who get own hats in n hat shuffle problem?
- Let X_i be 1 if ith person gets own hat and zero otherwise.
- What is $E[X_i]$, for $i \in \{1, 2, \ldots, n\}$?
- Answer: $1/n$.

More examples

- Now can we compute expected number of people who get own hats in n hat shuffle problem?
- Let X_i be 1 if ith person gets own hat and zero otherwise.
- What is $E[X_i]$, for $i \in \{1, 2, \ldots, n\}$?
- Answer: $1/n$.
- Can write total number with own hat as $X = X_1 + X_2 + \ldots + X_n$.

More examples

- Now can we compute expected number of people who get own hats in n hat shuffle problem?
- Let X_i be 1 if ith person gets own hat and zero otherwise.
- What is $E[X_i]$, for $i \in \{1, 2, \ldots, n\}$?
- Answer: $1/n$.
- Can write total number with own hat as $X = X_1 + X_2 + \ldots + X_n$.
- Linearity of expectation gives $E[X] = E[X_1] + E[X_2] + \ldots + E[X_n] = n \times 1/n = 1$.

Outline

Defining expectation

Functions of random variables

Motivation

Outline

Defining expectation

Functions of random variables

Motivation

Why should we care about expectation?

- **Laws of large numbers:** choose lots of independent random variables with same probability distribution as X — their average tends to be close to $E[X]$.

Why should we care about expectation?

- **Laws of large numbers:** choose lots of independent random variables with same probability distribution as X — their average tends to be close to $E[X]$.
- Example: roll $N = 10^6$ dice, let Y be the sum of the numbers that come up. Then Y/N is probably close to 3.5.

Why should we care about expectation?

- **Laws of large numbers:** choose lots of independent random variables with same probability distribution as X — their average tends to be close to $E[X]$.
- Example: roll $N = 10^6$ dice, let Y be the sum of the numbers that come up. Then Y/N is probably close to 3.5.
- **Economic theory of decision making:** Under "rationality" assumptions, each of us has utility function and tries to optimize its expectation.

Why should we care about expectation?

- **Laws of large numbers:** choose lots of independent random variables with same probability distribution as X — their average tends to be close to $E[X]$.
- Example: roll $N = 10^6$ dice, let Y be the sum of the numbers that come up. Then Y/N is probably close to 3.5.
- **Economic theory of decision making:** Under "rationality" assumptions, each of us has utility function and tries to optimize its expectation.
- **Financial contract pricing:** under "no arbitrage/interest" assumption, price of derivative equals its expected value in so-called **risk neutral probability**.

Why should we care about expectation?

- **Laws of large numbers:** choose lots of independent random variables with same probability distribution as X — their average tends to be close to $E[X]$.
- Example: roll $N = 10^6$ dice, let Y be the sum of the numbers that come up. Then Y/N is probably close to 3.5.
- **Economic theory of decision making:** Under "rationality" assumptions, each of us has utility function and tries to optimize its expectation.
- **Financial contract pricing:** under "no arbitrage/interest" assumption, price of derivative equals its expected value in so-called **risk neutral probability**.
- **Comes up everywhere** probability is applied.

Expected utility when outcome only depends on wealth

- Contract one: I'll toss 10 coins, and if they all come up heads (probability about one in a thousand), I'll give you 20 billion dollars.

Expected utility when outcome only depends on wealth

- Contract one: I'll toss 10 coins, and if they all come up heads (probability about one in a thousand), I'll give you 20 billion dollars.
- Contract two: I'll just give you ten million dollars.

Expected utility when outcome only depends on wealth

- Contract one: I'll toss 10 coins, and if they all come up heads (probability about one in a thousand), I'll give you 20 billion dollars.
- Contract two: I'll just give you ten million dollars.
- What are expectations of the two contracts? Which would you prefer?

Expected utility when outcome only depends on wealth

- Contract one: I'll toss 10 coins, and if they all come up heads (probability about one in a thousand), I'll give you 20 billion dollars.
- Contract two: I'll just give you ten million dollars.
- What are expectations of the two contracts? Which would you prefer?
- Can you find a function $u(x)$ such that given two random wealth variables W_1 and W_2, you prefer W_1 whenever $E[u(W_1)] < E[u(W_2)]$?

Expected utility when outcome only depends on wealth

- Contract one: I'll toss 10 coins, and if they all come up heads (probability about one in a thousand), I'll give you 20 billion dollars.
- Contract two: I'll just give you ten million dollars.
- What are expectations of the two contracts? Which would you prefer?
- Can you find a function $u(x)$ such that given two random wealth variables W_1 and W_2, you prefer W_1 whenever $E[u(W_1)] < E[u(W_2)]$?
- Let's assume $u(0) = 0$ and $u(1) = 1$. Then $u(x) = y$ means that you are indifferent between getting 1 dollar no matter what and getting x dollars with probability $1/y$.