Lecture 6: Discrete random variables

Statistics 251

Yi Sun and Zhongjian Wang

Department of Statistics
The University of Chicago

Lecture Outline

Random variables

Probability mass function and distribution function

Where are we?

Random variables

Probability mass function and distribution function

Random variables

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

- Interpretation: X is a quantity depending on the outcome in S

Random variables

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

- Interpretation: X is a quantity depending on the outcome in S

Example: Toss n coins so S is the set of 2^n coin sequences.

- $X = \{\text{number of heads}\}$ maps a sequence to the num. of heads
- What is $\mathbb{P}(X = k)$?

Random variables

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

- Interpretation: X is a quantity depending on the outcome in S

Example: Toss n coins so S is the set of 2^n coin sequences.

- $X = \{\text{number of heads}\}$ maps a sequence to the num. of heads
- What is $\mathbb{P}(X = k)$?
- Answer: $\frac{\#\{\text{sequences with } k \text{ heads}\}}{|S|} = \frac{\binom{n}{k}}{2^n}$

Examples

Shuffle n cards and let X be the position of card 1.

- X has values in $\{1, \ldots, n\}$
- What is $\mathbb{P}(X = k)$?

Roll 3 dice and let Y be the sum of the values. What is $\mathbb{P}(Y = 5)$?

Indicator random variables

The **indicator random variable** of an event E is

$$1_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}.$$

If $E_1, \ldots, E_k$ are events, then $X = \sum_{i=1}^{k} 1_{E_i}$ is the number of events which occur.

Where are we?

Random variables

Probability mass function and distribution function

Probability mass function

A random variable X is **discrete** if its values lie in a countable set.

The **probability mass function** is $p_X(a) = \mathbb{P}(X = a)$.

If A is the countable set of values of X, we have

$$\sum_{a \in A} p_X(a) = 1.$$

Cumulative distribution function

The **cumulative distribution function** (CDF) of X is

$$F_X(a) := \mathbb{P}(X \leq a) = \sum_{x \leq a} p_X(x).$$

Notice that $F_X(a)$ is non-decreasing and

$$\lim_{a \to -\infty} F_X(a) = 0$$

and

$$\lim_{a \to \infty} F_X(a) = 1.$$

Examples

Let T_1, $T_2, \ldots$ be a sequence of independent fair coin tosses in $\{T, H\}$. Let X be the smallest i for which $T_i = H$.

What is $p_X(k)$?

What is $F_X(k)$?

Examples

The probability mass function of a random variable X is given by $p(i) = c\lambda i/i!$, $i = 0, 1, 2, cdots$, where λ is some positive value. Find (a) $\mathbb{P}\{X = 0\}$ and (b) $\mathbb{P}\{X > 2\}$.

Gambler's ruin

Suppose gamblers Alice and Bob start with m and n dollars and take turns making fair \$1 bets.
What is the probability that Alice runs out of money first?

Gambler's ruin

Suppose gamblers Alice and Bob start with m and n dollars and take turns making fair \$1 bets.
What is the probability that Alice runs out of money first?

- Let X_m be the event that Alice runs out of money first starting with m dollars.
- Let F be the event that Alice wins the first flip.

Gambler's ruin

Suppose gamblers Alice and Bob start with m and n dollars and take turns making fair \$1 bets.
What is the probability that Alice runs out of money first?

- Let X_m be the event that Alice runs out of money first starting with m dollars.
- Let F be the event that Alice wins the first flip.
- By law of total probability, we have

$$\mathbb{P}(X_m) = \mathbb{P}(X_m \mid F)\mathbb{P}(F) + \mathbb{P}(X_m \mid F^c)\mathbb{P}(F^c),$$

  where $\mathbb{P}(X_m \mid F) = \mathbb{P}(X_{m+1})$ and $\mathbb{P}(X_m \mid F^c) = \mathbb{P}(X_{m-1})$.
- Setting $p_m = \mathbb{P}(X_m)$, this means that

$$p_m = \frac{1}{2}p_{m-1} + \frac{1}{2}p_{m+1}.$$

Gambler's ruin

Suppose gamblers Alice and Bob start with l and n dollars and take turns making fair \$1 bets.
What is the probability that Alice runs out of money first?

- Let X_m be the event that Alice runs out of money first starting with m dollars (then Bob start with $l + n - m$).
- For $p_m = \mathbb{P}(X_m)$, we have

$$p_m = \frac{1}{2}p_{m-1} + \frac{1}{2}p_{m+1} \iff p_m - p_{m-1} = p_{m+1} - p_m.$$

Gambler's ruin

Suppose gamblers Alice and Bob start with l and n dollars and take turns making fair \$1 bets.
What is the probability that Alice runs out of money first?

- Let X_m be the event that Alice runs out of money first starting with m dollars (then Bob start with $l+n-m$).
- For $p_m = \mathbb{P}(X_m)$, we have

$$p_m = \frac{1}{2}p_{m-1} + \frac{1}{2}p_{m+1} \iff p_m - p_{m-1} = p_{m+1} - p_m.$$

- Notice that $p_0 = 1$ and $p_{l+n} = 0$, so $p_0, p_1, \ldots, p_{l+n}$ are evenly spaced along $[0,1]$.
- This implies $p_k = \frac{l+n-k}{l+n}$, hence $p_l = \frac{n}{l+n}$.

Gambler's ruin

Suppose gamblers Alice and Bob start with l and ∞ **dollars** and take turns making fair \$1 bets.
What is the probability that Alice runs out of money first?

Gambler's ruin

Suppose gamblers Alice and Bob start with l and ∞ **dollars** and take turns making fair \$1 bets.
What is the probability that Alice runs out of money first?

- If Bob starts with n dollars, answer is $\frac{n}{l+n}$. Now

$$\lim_{n\to\infty} \frac{n}{l+n} = 1.$$

- If n is large, the house (Bob) always wins.