Week 3

Statistics 251

Zhongjian Wang

Department of Statistics
The University of Chicago

Where are we?

Random variables

Probability mass function and distribution function

Definition of expectation

Expectation of functions of a random variable

Definition of variance

Properties of variance

Decomposition into indicator variables

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
- $\mathbb{P}(X = k) = \frac{\#\{\text{sequences with } k \text{ heads}\}}{|S|} = \frac{\binom{n}{k}}{2^n}$

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

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

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

Where are we?

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 all possible values of X, we have

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

Cumulative distribution function

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

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

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 $F_X(k)$?

Where are we?

Random variables

Probability mass function and distribution function

Definition of expectation

Expectation of functions of a random variable

Definition of variance

Properties of variance

Decomposition into indicator variables

Expectation of a discrete random variable

Recall: A random variable X is a function $X : S \to \mathbb{R}$. If it is discrete, it has probability mass function $p_X(a) = \mathbb{P}(X = a)$.

The **expectation** of X is

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

It is the probability-weighted average of values of X.

Examples

Suppose $\mathbb{P}(X = 1) = \frac{1}{2}$, $\mathbb{P}(X = 2) = \frac{1}{4}$, and $\mathbb{P}(X = 3) = \frac{1}{4}$. What is $\mathbb{E}[X]$?

Suppose $\mathbb{P}(X = 1) = p$ and $\mathbb{P}(X = 0) = 1 - p$. What is $\mathbb{E}[X]$?

Examples

What is the expected value of a roll of a 6-sided die?

What is the expected number of heads in 2 coin flips?

Where are we?

Random variables

Probability mass function and distribution function

Definition of expectation

Expectation of functions of a random variable

Definition of variance

Properties of variance

Decomposition into indicator variables

Functions of a random variable

If X is a random variable and $f(x)$ is any function, we can create a new random variable $f(X)$, which maps $s \in S$ to $f(X(s)) \in \mathbb{R}$.

For $Y = f(X)$, we have

$$p_Y(y) = \mathbb{P}(f(X) = y) = \sum_{x:f(x)=y} \mathbb{P}(X = x) = \sum_{x:f(x)=y} p_X(x).$$

Suppose $\mathbb{P}(X = 0) = \frac{1}{2}$ and $\mathbb{P}(X = 1) = \frac{1}{2}$.

- We have $X = X^2$.
- For $Y = (X + 1)^2$, we have $\mathbb{P}(Y = 1) = \frac{1}{2}$ and $\mathbb{P}(Y = 4) = \frac{1}{2}$.

Expectations of functions

The expectation of $Y = f(X)$ is

$$\begin{aligned}\mathbb{E}[f(X)] &= \sum_{y:p_Y(y)>0} p_Y(y) \cdot y \\ &= \sum_{y:p_Y(y)>0} \sum_{x:f(x)=y} p_X(x) \cdot f(x) \\ &= \sum_{x:p_X(x)>0} p_X(x) \cdot f(x).\end{aligned}$$

It is the probability-weighted average value of $f(X)$.

Properties of expectation

For constants a, b, $\mathbb{E}[aX + b]$ is

$$\begin{aligned}\mathbb{E}[aX + b] &= \sum_{x:p_X(x)>0} p_X(x) \cdot (ax + b) \\ &= a \sum_{x:p_X(x)>0} p_X(x) \cdot x + b \sum_{x:p_X(x)>0} p_X(x) \\ &= a\mathbb{E}[X] + b.\end{aligned}$$

Examples

Let X be the roll of a 6-sided die. What is $\mathbb{E}[X^2]$?

Linearity of expectation

If X, Y are random variables, get new random variable $X + Y$ with

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y].$$

Also works with n variables $X_1, \ldots, X_n$:

$$\mathbb{E}[a_1 X_1 + \cdots + a_n X_n] = a_1 \mathbb{E}[X_1] + \cdots + a_n \mathbb{E}[X_n].$$

This means that $\mathbb{E}[-]$ is a **linear** function on the space of random variables.

Examples

The linearity of expectation has an important application in the method of 'decomposing into sum of indicator functions'
What is the expected number of heads in n coin flips?

Suppose we shuffle a deck of n cards. What is the expected number of cards which end up in the same position?

Suppose you put 6 pairs of socks into the laundry, but 3 socks are lost, leaving 9 total socks. What is the expected number of complete pairs? (Exercise)

Why care about expectations?

Suppose $X_1, \ldots, X_n$ are chosen independently with the same distribution as X. Law of large numbers:

$$\frac{X_1 + \cdots + X_n}{n} \overset{n \to \infty}{\to} \mathbb{E}[X].$$

- From 10^6 fair coin flips, probably $\approx$half of the flips will be heads.
- If this is not true, I should be suspicious the coin is not fair.

Later in the course, we will quantify this.

Decision making under uncertainty

Rational choice theory in economics says:

- Agents have a utility function $u(x)$ depending on the state of the world x.
- Faced with possible random states of the world $X_\cdot$, agents act (denoted as k) to maximize their **expected utility**:

$$\max_k \mathbb{E}[u(X_k)].$$

Examples:

- Netflix decides whether to invest in a new show.
- Trader decides which stock to buy.

Utility may not depend purely on monetary value.

Decision making under uncertainty

You are offered the chance to play the following game. We flip a fair coin. If it is heads, you pay me \$10. If it is tails, I pay you \$11. Do you want to play?

The expected value of playing is

$$\frac{1}{2}(-\$10) + \frac{1}{2} \cdot \$11 = \$0.50.$$

What if you pay me \$100 or I pay you \$110? The expected value of playing is

$$\frac{1}{2}(-\$100) + \frac{1}{2} \cdot \$110 = \$5.$$

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

Gambler's ruin

Suppose gamblers Alice and Bob start with m 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 $\frac{n}{m+n}$. Now

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

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

Where are we?

Definition of variance

Suppose that X has expectation $\mathbb{E}[X]$. Its **variance** is

$$\mathrm{Var}(X) := \mathbb{E}[(X - \mathbb{E}[X])^2].$$

More explicitly, this is given by

$$\mathrm{Var}(X) = \sum_{x : p_X(x) > 0} p_X(x) \cdot (x - \mathbb{E}[X])^2.$$

It is the probability-weighted average deviation of X from $\mathbb{E}[X]$.

Definition of variance

Suppose that X has expectation $\mathbb{E}[X]$. Its **variance** is

$$\mathrm{Var}(X) := \mathbb{E}[(X - \mathbb{E}[X])^2].$$

Why $(X - \mathbb{E}[X])^2$ instead of $|X - \mathbb{E}[X]|$ or $(X - \mathbb{E}[X])^4$?

Alternate formula

Expanding the formula for variance using linearity of expectation:

$$\begin{aligned}\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] &= \mathbb{E}[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2] \\ &= \mathbb{E}[X^2] - 2\mathbb{E}[X] \cdot \mathbb{E}[X] + \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2.\end{aligned}$$

Variance measures the expected difference between X^2 and $\mathbb{E}[X]^2$.

Examples

If X is the outcome of a die roll, what is $\text{Var}(X)$?

If X is the number of heads in 2 coin flip, what is $\text{Var}(X)$?

Examples

Let X be the amount won by a lottery ticket with probability $\frac{1}{1000}$ of winning 1000. What are $\mathbb{E}[X]$ and $\text{Var}(X)$?

Let X be the amount won by a lottery ticket with probability $\frac{1}{1,000,000}$ of winning $1,000,000$. What are $\mathbb{E}[X]$ and $\text{Var}(X)$?

Which ticket you would like to buy?

Where are we?

Degenerate case

$\text{Var}(X) = 0$ if and only $\mathbb{P}(X = \mathbb{E}[X]) = 1$.

Affine transformations

If X is a random variable, what is $\mathrm{Var}(aX + b)$?

Recall $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$. We compute

$$\begin{aligned}
\mathrm{Var}(aX + b) &= \mathbb{E}[(aX + b)^2] - \mathbb{E}[aX + b]^2 \\
&= \mathbb{E}[a^2X^2 + 2abX + b^2] - (a\mathbb{E}[X] + b)^2 \\
&= a^2\mathbb{E}[X^2] + 2ab\mathbb{E}[X] + b^2 - a^2\mathbb{E}[X]^2 - 2ab\mathbb{E}[X] - b^2 \\
&= a^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2) \\
&= a^2\,\mathrm{Var}(X).
\end{aligned}$$

Notice that $\mathrm{Var}(X)$ scales as the square of X.

Standard deviation

The **standard deviation** of X is

$$\mathrm{SD}[X] := \sqrt{\mathrm{Var}(X)}.$$

This means that

$$\mathrm{SD}[aX + b] = \sqrt{a^2 \,\mathrm{Var}(X)} = a\mathrm{SD}[X],$$

which means that $\mathrm{SD}[X]$ is at the same scale as X.

Where are we?

Card example

Choose 5 cards from a standard deck of cards. Let X be the number of aces. What is $\mathbb{E}[X]$ and $\mathrm{Var}(X)$?

Notice that

$$\mathbb{P}(X = k) = \frac{\#\{\text{hands with } k \text{ aces}\}}{\#\{\text{hands}\}} = \frac{\binom{4}{k}\binom{48}{5-k}}{\binom{52}{5}}.$$

So in theory we can compute

$$\mathbb{E}[X] = \sum_{k=0}^{4} \mathbb{P}(X = k) \cdot k$$

$$\mathrm{Var}(X) = \sum_{k=0}^{4} \mathbb{P}(X = k) \cdot k^2 - \mathbb{E}[X]^2.$$

Card example redux

Choose 5 cards from a standard deck of cards. Let X be the number of aces. What is $\mathbb{E}[X]$ and $\mathrm{Var}(X)$?

Define the indicator variables

$$X_i = \mathbf{1}_{\text{card } i \text{ is an ace}}$$

so that

$$X = X_1 + X_2 + X_3 + X_4 + X_5.$$

Deck shuffles

Let X be the number of cards which stay in the same position when a deck of n cards is shuffled. What is $\text{Var}(X)$?

Deck shuffles

Let X be the number of cards which stay in the same position when a deck of n cards is shuffled. What is $\mathrm{Var}(X)$?

For $i = 1, \ldots, n$, let $X_i = \mathbf{1}_{\text{card } i \text{ stays in the same position}}$ so that

$$X = \sum_{i=1}^{n} X_i.$$

Notice that $\mathbb{E}[X_i^2] = \mathbb{E}[X_i] = \frac{1}{n}$ and for $i \neq j$ that $\mathbb{E}[X_i X_j] = \frac{1}{n} \cdot \frac{1}{n-1}$. This means

$$\mathbb{E}[X^2] = \sum_{i=1}^{n} \mathbb{E}[X_i^2] + 2 \sum_{i<j} \mathbb{E}[X_i X_j] = n \cdot \frac{1}{n} + 2 \cdot \frac{n(n-1)}{2} \cdot \frac{1}{n(n-1)} = 2$$

so $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 2 - 1 = 1.$