Lecture 8: Variance and standard deviation

Statistics 251

Yi Sun and Zhongjian Wang

Department of Statistics
The University of Chicago

Before lectures

Somebody asks a very good question during yesterdays office hour. How to *understand* $\mathbb{P}(\cup_{i=1}^{\infty} A_i)$?

Lecture Outline

Definition of variance

Properties of variance

Decomposition into indicator variables

Where are we?

Reminder of expectation

If X is a discrete random variable, its expectation is

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

for probability mass function $p_X(x)$.

Definition of variance

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

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

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 $\mathbb{E}[X]$.

Alternate formula

Expanding the formula for variance using linearity of expectation:

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

Variance measures the 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 $\text{Var}(X)$?

Where are we?

Definition of variance

Properties of variance

Decomposition into indicator variables

Affine transformations

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

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?

Definition of variance

Properties of variance

Decomposition into indicator variables

Card example

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

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 $\mathrm{Var}(X)$?

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

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

so that

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

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

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

so that

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

This means that $\mathbb{E}[X_i] = \frac{1}{13}$ and hence

$$\mathbb{E}[X] = 5 \cdot \mathbb{E}[X_1] = \frac{5}{13}.$$

Card example redux

(Alternative approach with indicator functions)
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 $X_i = 1_{\text{card } i \text{ is an ace}}$ so that $X = X_1 + X_2 + X_3 + X_4 + X_5$.

Card example redux

(Alternative approach with indicator functions)
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 $X_i = 1_{\text{card } i \text{ is an ace}}$ so that $X = X_1 + X_2 + X_3 + X_4 + X_5$. We now find that

$$\mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2 + X_3 + X_4 + X_5)^2]$$
$$= \mathbb{E}\left[\sum_{i=1}^{5} X_i^2 + 2\sum_{i<j} X_i X_j\right].$$

We have $\mathbb{E}[X_i^2] = \frac{1}{13}$ and $\mathbb{E}[X_i X_j] = \frac{1}{13} \cdot \frac{3}{51} = \frac{1}{13 \cdot 17}$. This means $\mathbb{E}[X^2] = \frac{5}{13} + \frac{20}{13 \cdot 17} = \frac{105}{13 \cdot 17}$. This means

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{105}{13 \cdot 17} - \frac{25}{13^2}.$$

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 $\mathsf{Var}(X)$?

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 $X_i = 1_{\text{card } i \text{ stays in the same position}}$ 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 $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 2 - 1 = 1.$