CS70
Concentration Inequalities, WLLN

Kelvin Lee

UC Berkeley

April 13, 2021

Overview

Covariance

Covariance

The **covariance** of X and Y is

Covariance

The **covariance** of X and Y is

$$\text{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])],$$

Covariance

The **covariance** of X and Y is

$$\text{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])],$$

which can be simplified to

Covariance

The **covariance** of X and Y is

$$\text{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])],$$

which can be simplified to

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

Covariance

The **covariance** of X and Y is

$$\mathrm{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])],$$

which can be simplified to

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

If X and Y are **independent**, then

Covariance

The **covariance** of X and Y is

$$\text{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])],$$

which can be simplified to

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

If X and Y are **independent**, then

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] \implies \text{Cov}(X, Y) = 0.$$

Correlation

Correlation

Suppose X, Y are random variables with $\sigma_X, \sigma_Y > 0$. Then the **correlation** ρ of X and Y is

Correlation

Suppose X, Y are random variables with $\sigma_X, \sigma_Y > 0$. Then the **correlation** ρ of X and Y is

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\,\text{Var}(Y)}} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Correlation

Suppose X, Y are random variables with $\sigma_X, \sigma_Y > 0$. Then the **correlation** ρ of X and Y is

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\,\text{Var}(Y)}} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

and $-1 \leq \rho(X, Y) \leq 1$.

Properties of Covariance

Properties of Covariance

Here are some important facts about covariance:

Properties of Covariance

Here are some important facts about covariance:

- $\text{Var}(X) = \text{Cov}(X, X)$.

Properties of Covariance

Here are some important facts about covariance:

- $\text{Var}(X) = \text{Cov}(X, X)$.
- X, Y **independent** $\implies \text{Cov}(X, Y) = 0$.

Properties of Covariance

Here are some important facts about covariance:

- $\text{Var}(X) = \text{Cov}(X, X)$.
- X, Y **independent** $\implies \text{Cov}(X, Y) = 0$.
- $\text{Cov}(a + X, b + Y) = \text{Cov}(X, Y)$.

Properties of Covariance

Here are some important facts about covariance:

- $\text{Var}(X) = \text{Cov}(X, X)$.
- X, Y **independent** $\implies \text{Cov}(X, Y) = 0$.
- $\text{Cov}(a + X, b + Y) = \text{Cov}(X, Y)$.
- **Bilinearity**:

Properties of Covariance

Here are some important facts about covariance:

- $\text{Var}(X) = \text{Cov}(X, X)$.
- X, Y **independent** $\implies \text{Cov}(X, Y) = 0$.
- $\text{Cov}(a + X, b + Y) = \text{Cov}(X, Y)$.
- **Bilinearity**:

$$\text{Cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \,\text{Cov}(X_i, Y_j).$$

Markov's Inequality

Markov's Inequality

Theorem (Markov's Inequality)

Markov's Inequality

Theorem (Markov's Inequality)

*For a **non-negative** random variable X with finite mean and any positive constant c,*

Markov's Inequality

Theorem (Markov's Inequality)

For a ***non-negative*** *random variable X with finite mean and any positive constant c,*

$$\mathbb{P}(X \geq c) \leq \frac{\mathbb{E}[X]}{c}.$$

Markov's Inequality

Theorem (Markov's Inequality)

*For a **non-negative** random variable X with finite mean and any positive constant c,*

$$\mathbb{P}(X \geq c) \leq \frac{\mathbb{E}[X]}{c}.$$

Proof:

Markov's Inequality

Theorem (Markov's Inequality)

For a ***non-negative*** *random variable X with finite mean and any positive constant c,*

$$\mathbb{P}(X \geq c) \leq \frac{\mathbb{E}[X]}{c}.$$

Proof:

$$\mathbb{E}[X] = \sum_{x \in \mathcal{X}} x \cdot \mathbb{P}(X = x)$$

Markov's Inequality

Theorem (Markov's Inequality)

*For a **non-negative** random variable X with finite mean and any positive constant c,*

$$\mathbb{P}(X \geq c) \leq \frac{\mathbb{E}[X]}{c}.$$

Proof:

$$\begin{aligned} \mathbb{E}[X] &= \sum_{x \in \mathcal{X}} x \cdot \mathbb{P}(X = x) \\ &\geq \sum_{x \geq c} x \cdot \mathbb{P}(X = x) \end{aligned}$$

Markov's Inequality

Theorem (Markov's Inequality)

*For a **non-negative** random variable X with finite mean and any positive constant c,*

$$\mathbb{P}(X \geq c) \leq \frac{\mathbb{E}[X]}{c}.$$

Proof:

$$\begin{aligned} \mathbb{E}[X] &= \sum_{x \in \mathcal{X}} x \cdot \mathbb{P}(X = x) \\ &\geq \sum_{x \geq c} x \cdot \mathbb{P}(X = x) \\ &\geq \sum_{x \geq c} c \cdot \mathbb{P}(X = x) \end{aligned}$$

Markov's Inequality

Theorem (Markov's Inequality)

*For a **non-negative** random variable X with finite mean and any positive constant c,*

$$\mathbb{P}(X \geq c) \leq \frac{\mathbb{E}[X]}{c}.$$

Proof:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{x \in \mathcal{X}} x \cdot \mathbb{P}(X = x) \\
&\geq \sum_{x \geq c} x \cdot \mathbb{P}(X = x) \\
&\geq \sum_{x \geq c} c \cdot \mathbb{P}(X = x) \\
&= c \sum_{x \geq c} \mathbb{P}(X = x)
\end{aligned}$$

Markov's Inequality

Theorem (Markov's Inequality)

*For a **non-negative** random variable X with finite mean and any positive constant c,*

$$\mathbb{P}(X \geq c) \leq \frac{\mathbb{E}[X]}{c}.$$

Proof:

$$\begin{aligned}\mathbb{E}[X] &= \sum_{x \in \mathcal{X}} x \cdot \mathbb{P}(X = x) \\ &\geq \sum_{x \geq c} x \cdot \mathbb{P}(X = x) \\ &\geq \sum_{x \geq c} c \cdot \mathbb{P}(X = x) \\ &= c \sum_{x \geq c} \mathbb{P}(X = x) \\ &= c\, \mathbb{P}[X \geq c].\end{aligned}$$

Markov's Inequality

Markov's Inequality

Here's a smarter way to prove this inequality.

Markov's Inequality

Here's a smarter way to prove this inequality.

Proof:

Markov's Inequality

Here's a smarter way to prove this inequality.

Proof:

Since X is a non-negative and $c > 0$, then for all $\omega \in \Omega$

Markov's Inequality

Here's a smarter way to prove this inequality.

Proof:

Since X is a non-negative and $c > 0$, then for all $\omega \in \Omega$

$$X(\omega) \geq \mathbb{I}\{X(\omega) \geq c\}.$$

Markov's Inequality

Here's a smarter way to prove this inequality.

Proof:

Since X is a non-negative and $c > 0$, then for all $\omega \in \Omega$

$$X(\omega) \geq \mathbb{I}\{X(\omega) \geq c\}.$$

The RHS is 0 if $X(\omega) < c$ and is c if $X(\omega) \geq c$ implied by the indicator function. Taking expectations of both sides gives

Markov's Inequality

Here's a smarter way to prove this inequality.

Proof:

Since X is a non-negative and $c > 0$, then for all $\omega \in \Omega$

$$X(\omega) \geq \mathbb{I}\{X(\omega) \geq c\}.$$

The RHS is 0 if $X(\omega) < c$ and is c if $X(\omega) \geq c$ implied by the indicator function. Taking expectations of both sides gives

$$\mathbb{E}[X] \geq c\,\mathbb{E}[\mathbb{I}\{X \geq c\}] = c\,\mathbb{P}(X \geq c).$$

□

Generalized Markov's Inequality

Generalized Markov's Inequality

What if X is negative?

Generalized Markov's Inequality

What if X is negative?

Theorem (Generalized Markov's Inequality)

Generalized Markov's Inequality

What if X is negative?

Theorem (Generalized Markov's Inequality)

Let X be an arbitrary random variable with finite mean. Then, for any positive constants c and r,

Generalized Markov's Inequality

What if X is negative?

Theorem (Generalized Markov's Inequality)

Let X be an arbitrary random variable with finite mean. Then, for any positive constants c and r,

$$\mathbb{P}(|X| \geq c) \leq \frac{\mathbb{E}\left(|X|^r\right)}{c^r}.$$

Generalized Markov's Inequality

What if X is negative?

Theorem (Generalized Markov's Inequality)

Let X be an arbitrary random variable with finite mean. Then, for any positive constants c and r,

$$\mathbb{P}(|X| \geq c) \leq \frac{\mathbb{E}(|X|^r)}{c^r}.$$

Proof:

Generalized Markov's Inequality

What if X is negative?

Theorem (Generalized Markov's Inequality)

Let X be an arbitrary random variable with finite mean. Then, for any positive constants c and r,

$$\mathbb{P}(|X| \geq c) \leq \frac{\mathbb{E}\left(|X|^r\right)}{c^r}.$$

Proof:

For $c > 0$ and $r > 0$, we have

Generalized Markov's Inequality

What if X is negative?

Theorem (Generalized Markov's Inequality)

Let X be an arbitrary random variable with finite mean. Then, for any positive constants c and r,

$$\mathbb{P}(|X| \geq c) \leq \frac{\mathbb{E}\left(|X|^r\right)}{c^r}.$$

Proof:

For $c > 0$ and $r > 0$, we have

$$|X|^r \geq |X|^r \cdot \mathbb{I}\{|X| \geq c\} \geq c^r \mathbb{I}\{|X| \geq c\}.$$

Generalized Markov's Inequality

What if X is negative?

Theorem (Generalized Markov's Inequality)

Let X be an arbitrary random variable with finite mean. Then, for any positive constants c and r,

$$\mathbb{P}(|X| \geq c) \leq \frac{\mathbb{E}\left(|X|^r\right)}{c^r}.$$

Proof:

For $c > 0$ and $r > 0$, we have

$$|X|^r \geq |X|^r \cdot \mathbb{I}\{|X| \geq c\} \geq c^r \mathbb{I}\{|X| \geq c\}.$$

(Note that the last inequality would not hold if r were negative.) Taking expectations of both sides gives

Generalized Markov's Inequality

What if X is negative?

Theorem (Generalized Markov's Inequality)

Let X be an arbitrary random variable with finite mean. Then, for any positive constants c and r,

$$\mathbb{P}(|X| \geq c) \leq \frac{\mathbb{E}\left(|X|^r\right)}{c^r}.$$

Proof:

For $c > 0$ and $r > 0$, we have

$$|X|^r \geq |X|^r \cdot \mathbb{I}\{|X| \geq c\} \geq c^r \mathbb{I}\{|X| \geq c\}.$$

(Note that the last inequality would not hold if r were negative.) Taking expectations of both sides gives

$$\mathbb{E}\left[|X|^r\right] \geq c^r \cdot \mathbb{E}[\mathbb{I}\{|X| \geq c\}] = c^r \mathbb{P}(|X| \geq c).$$

Chebyshev's Inequality

Chebyshev's Inequality

Theorem (Chebyshev's Inequality)

Chebyshev's Inequality

Theorem (Chebyshev's Inequality)

For a random variable X with finite expectation $\mathbb{E}[X] = \mu$ and any positive constant c,

Chebyshev's Inequality

Theorem (Chebyshev's Inequality)

For a random variable X with finite expectation $\mathbb{E}[X] = \mu$ and any positive constant c,

$$\mathbb{P}(|X - \mu| \geq c) \leq \frac{\mathrm{Var}(X)}{c^2}.$$

Chebyshev's Inequality

Theorem (Chebyshev's Inequality)

For a random variable X with finite expectation $\mathbb{E}[X] = \mu$ and any positive constant c,

$$\mathbb{P}(|X - \mu| \geq c) \leq \frac{\text{Var}(X)}{c^2}.$$

Proof:

Chebyshev's Inequality

Theorem (Chebyshev's Inequality)

For a random variable X with finite expectation $\mathbb{E}[X] = \mu$ and any positive constant c,

$$\mathbb{P}(|X - \mu| \geq c) \leq \frac{\text{Var}(X)}{c^2}.$$

Proof:

Define $Y = (X - \mu)^2$ and so $\mathbb{E}[Y] = \mathbb{E}\left[(X - \mu)^2\right] = \text{Var}(X)$.

Chebyshev's Inequality

Theorem (Chebyshev's Inequality)

For a random variable X with finite expectation $\mathbb{E}[X] = \mu$ and any positive constant c,

$$\mathbb{P}(|X - \mu| \geq c) \leq \frac{\text{Var}(X)}{c^2}.$$

Proof:

Define $Y = (X - \mu)^2$ and so $\mathbb{E}[Y] = \mathbb{E}\left[(X - \mu)^2\right] = \text{Var}(X)$.We are interested in $|X - \mu| \geq c$, which is equivalent to $Y = (X - \mu)^2 \geq c^2$.

Chebyshev's Inequality

Theorem (Chebyshev's Inequality)

For a random variable X with finite expectation $\mathbb{E}[X] = \mu$ and any positive constant c,

$$\mathbb{P}(|X - \mu| \geq c) \leq \frac{\text{Var}(X)}{c^2}.$$

Proof:

Define $Y = (X - \mu)^2$ and so $\mathbb{E}[Y] = \mathbb{E}\left[(X - \mu)^2\right] = \text{Var}(X)$.We are interested in $|X - \mu| \geq c$, which is equivalent to $Y = (X - \mu)^2 \geq c^2$. Therefore, $\mathbb{P}(|X - \mu| \geq c) = \mathbb{P}\left(Y \geq c^2\right)$.

Chebyshev's Inequality

Theorem (Chebyshev's Inequality)

For a random variable X with finite expectation $\mathbb{E}[X] = \mu$ and any positive constant c,

$$\mathbb{P}(|X - \mu| \geq c) \leq \frac{\text{Var}(X)}{c^2}.$$

Proof:

Define $Y = (X - \mu)^2$ and so $\mathbb{E}[Y] = \mathbb{E}\left[(X - \mu)^2\right] = \text{Var}(X)$.We are interested in $|X - \mu| \geq c$, which is equivalent to $Y = (X - \mu)^2 \geq c^2$. Therefore, $\mathbb{P}(|X - \mu| \geq c) = \mathbb{P}\left(Y \geq c^2\right)$. Moreover, Y is obviously non-negative, so we can apply Markov's inequality. □

Weak Law of Large Numbers

Weak Law of Large Numbers

Theorem (Weak Law of Large Numbers)

Weak Law of Large Numbers

Theorem (Weak Law of Large Numbers)

Let $X_1, X_2, \ldots, X_n$ be i.i.d random variables with the same distribution and mean μ. Then, for all $\varepsilon > 0$,

Weak Law of Large Numbers

Theorem (Weak Law of Large Numbers)

Let $X_1, X_2, \ldots, X_n$ be i.i.d random variables with the same distribution and mean μ. Then, for all $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \to 0, \text{ as } n \to \infty$$

Weak Law of Large Numbers

Theorem (Weak Law of Large Numbers)

Let $X_1, X_2, \ldots, X_n$ be i.i.d random variables with the same distribution and mean μ. Then, for all $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \to 0, \text{ as } n \to \infty$$

Proof:

Weak Law of Large Numbers

Theorem (Weak Law of Large Numbers)

Let $X_1, X_2, \ldots, X_n$ be i.i.d random variables with the same distribution and mean μ. Then, for all $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \to 0, \text{ as } n \to \infty$$

Proof:

Let $Y_n = \frac{X_1 + \cdots + X_n}{n}$. Then

Weak Law of Large Numbers

Theorem (Weak Law of Large Numbers)

Let $X_1, X_2, \ldots, X_n$ be i.i.d random variables with the same distribution and mean μ. Then, for all $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \to 0, \text{ as } n \to \infty$$

Proof:

Let $Y_n = \frac{X_1+\cdots+X_n}{n}$. Then

$$\mathbb{P}(|Y_n - \mu| \geq \varepsilon) \leq \frac{\mathrm{Var}(Y_n)}{\varepsilon^2} = \frac{\mathrm{Var}(X_1 + \cdots + X_n)}{n^2\varepsilon^2}$$

Weak Law of Large Numbers

Theorem (Weak Law of Large Numbers)

Let $X_1, X_2, \ldots, X_n$ be i.i.d random variables with the same distribution and mean μ. Then, for all $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \to 0, \text{ as } n \to \infty$$

Proof:

Let $Y_n = \frac{X_1 + \cdots + X_n}{n}$. Then

$$\mathbb{P}\left(|Y_n - \mu| \geq \varepsilon\right) \leq \frac{\operatorname{Var}(Y_n)}{\varepsilon^2} = \frac{\operatorname{Var}(X_1 + \cdots + X_n)}{n^2\varepsilon^2}$$

$$= \frac{n\operatorname{Var}(X_1)}{n^2\varepsilon^2} = \frac{\operatorname{Var}(X_1)}{n\varepsilon^2} \to 0, \text{ as } n \to \infty$$

Problem Time!