

COMP2610 / COMP6261 - Information Theory

# Assignment Project Exam Help

Lecture 9: Probabilistic Inequalities

<https://eduassistpro.github.io>



Australian  
National  
University

Add WeChat edu\_assist\_pro

20 August, 2018

Last time

# Assignment Project Exam Help

Mutual information chain rule

Jensen'

<https://eduassistpro.github.io>

"Informa

Data processing inequality

Add WeChat edu\_assist\_pro

## Review: Data-Processing Inequality

Theorem

$i: X \rightarrow Y \rightarrow Z$  then  $I(X; Y) \geq I(X; Z)$

- $X$  is t

s the

pro

<https://eduassistpro.github.io>

- No “clever” manipulation of the data can impro  
can be made from the data

Add WeChat edu\_assist\_pro

- No processing of  $Y$ , deterministic or random, can increase the information that  $Y$  contains about  $X$

This time

# Assignment Project Exam Help

- Markov's inequality,

- Ch <https://eduassistpro.github.io>

- Add WeChat edu\_assist\_pro

- Law of large numbers

# Assignment Project Exam Help

② Mark

③ Cheb <https://eduassistpro.github.io>

④ Law of large numbers

## Add WeChat edu\_assist\_pro

⑤ Wrapping Up

1 Properties of expectation and variance

# Assignment Project Exam Help

2 Markov's inequality

3 Cheb

<https://eduassistpro.github.io>

4 Law of large numbers

Add WeChat edu\_assist\_pro

5 Wrapping Up

## Expectation and Variance

Let  $X$  be a random variable over  $\mathcal{X}$ , with probability distribution  $p$

# Assignment Project Exam Help

Expected value:

Variance

<https://eduassistpro.github.io>

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

Add WeChat [edu\\_assist\\_pro](https://edu_assist_pro)

Standard deviation is  $\sqrt{\mathbb{V}[X]}$

## Properties of expectation

A key property of expectations is **linearity**:

Assignment Project Exam Help

$$\mathbb{E} \quad X_i = \mathbb{E}[X_i]$$

<https://eduassistpro.github.io>

$$x_1 \in \mathcal{X}_1$$

$$x_n \in \mathcal{X}_n$$

$$i=1$$

This holds even if the variables are dependent!

Add WeChat edu\_assist\_pro

We have for any  $a \in \mathbb{R}$ ,

$$\mathbb{E}[aX] = a \cdot \mathbb{E}[X].$$

## Properties of variance

We have linearity of variance for independent random variables:

Assignment Project Exam Help

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

Does not hold for dependent variables

<https://eduassistpro.github.io>

(prove this: expand the definition of variance and rely upon the fact that  $\text{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]$ )

Add WeChat edu\_assist\_pro

We have for any  $a \in \mathbb{R}$ ,

$$\text{Var}(aX) = a^2 \cdot \text{Var}(X).$$

1 Properties of expectation and variance

2 Markov's inequality

3 Cheb

<https://eduassistpro.github.io>

4 Law of large numbers

Add WeChat edu\_assist\_pro

5 Wrapping Up

## Markov's Inequality

### Motivation

# Assignment Project Exam Help

1000 school students sit an examination

The busy p

<https://eduassistpro.github.io>

The principal wants to estimate the maximum poss

who scored more than 80

- A question about the minimum number of s

Add WeChat [edu\\_assist\\_pro](https://edu_assist_pro)

# Markov's Inequality

## Motivation

Call  $x$  the number of students who score  $> 80$

Assignment Project Exam Help

We know:

$40 \cdot 1$  <https://eduassistpro.github.io/markov-inequality.html>

Exam scores are nonnegative, so certainly

Add WeChat [edu\\_assist\\_pro](https://edu_assist_pro)

Thus,  $80x < 40 \cdot 1000$ , or

$$x < 500.$$

Can we formalise this more generally?

## Markov's Inequality

# Assignment Project Exam Help

Theorem

Let  $X$  be a nonnegative random variable. Then, for any  $\lambda > 0$ ,

<https://eduassistpro.github.io>

Bounds probability of observing a large outcome

Add WeChat edu\_assist\_pro

Vacuous if  $\lambda < \mathbb{E}[X]$

## Markov's Inequality

### Alternate Statement

# Assignment Project Exam Help

#### Corollary

Let  $X$  be

<https://eduassistpro.github.io>

Observations of nonnegative random variable  $u$

than expected value

Add WeChat edu\_assist\_pro

Vacuous if  $\lambda < 1$

## Markov's Inequality

### Proof

Assignment Project Exam Help

$$\mathbb{E}[X] = \sum x \cdot p(x)$$

<https://eduassistpro.github.io>

$$\geq \lambda \cdot p(X \geq \lambda)$$

Add WeChat edu\_assist\_pro

$$\geq \sum_{x \geq \lambda} x \cdot p(x)$$

$$= \lambda \cdot p(X \geq \lambda).$$

## Markov's Inequality

Illustration from

<https://justindomke.wordpress.com/2008/06/19/markovs-inequality/>

# Assignment Project Exam Help

<https://eduassistpro.github.io>

Add WeChat edu\_assist\_pro

## Markov's Inequality

Illustration from

<https://justindomke.wordpress.com/2008/06/19/markovs-inequality/>

# Assignment Project Exam Help

<https://eduassistpro.github.io>

Add WeChat edu\_assist\_pro

## Markov's Inequality

Illustration from

<https://justindomke.wordpress.com/2008/06/19/markovs-inequality/>

# Assignment Project Exam Help

<https://eduassistpro.github.io>

Add WeChat edu\_assist\_pro

## Markov's Inequality

Illustration from

<https://justindomke.wordpress.com/2008/06/19/markovs-inequality/>

# Assignment Project Exam Help

<https://eduassistpro.github.io>

Add WeChat edu\_assist\_pro

1 Properties of expectation and variance

2 Markov's inequality

# Assignment Project Exam Help

3 Cheb

<https://eduassistpro.github.io>

4 Law of large numbers

Add WeChat edu\_assist\_pro

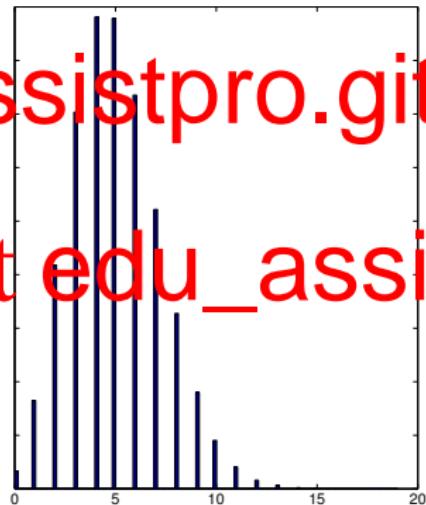
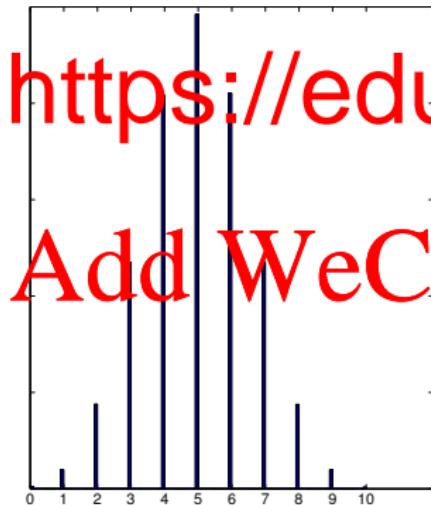
5 Wrapping Up

# Chebyshev's Inequality

## Motivation

Markov's inequality only uses the **mean** of the distribution

Assignment Project Exam Help  
What about the spread of the distribution (**Variance**)?



<https://eduassistpro.github.io>  
Add WeChat `edu_assist_pro`

## Chebyshev's Inequality

# Assignment Project Exam Help

Let  $X$  be a random variable with  $\mathbb{E}[X] < \infty$ . Then, for any  $\lambda > 0$ ,

<https://eduassistpro.github.io/>

Bounds the probability of observing an “unexpect

Does not require non-negativity

Add WeChat edu\_assist\_pro

Two-sided bound

## Chebyshev's Inequality

Alternate Statement

# Assignment Project Exam Help

Corollary

Let  $X$  be

<https://eduassistpro.github.io>

$$(|x - \bar{x}|) \geq \lambda \cdot \sqrt{V(x)}.$$

Observations are unlikely to occur several standard deviations from the mean

Add WeChat `edu_assist_pro`

## Chebyshev's Inequality

Proof

Define

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

Then, by Markov's inequality, for any  $y > 0$ ,

<https://eduassistpro.github.io>

But,

$$\mathbb{E}[Y] = \mathbb{V}[X]$$

Also, [Add WeChat edu\\_assist\\_pro](#)

Thus, setting  $\lambda = \sqrt{\nu}$ ,

$$p(|X - \mathbb{E}[X]| \geq \lambda) \leq \frac{\mathbb{V}[X]}{\lambda^2}.$$

## Chebyshev's Inequality

Illustration

For a binomial  $X$  with  $N$  trials and success probability  $\theta$ , we have e.g.

Assignment Project Exam Help

<https://eduassistpro.github.io>

Add WeChat edu\_assist\_pro

## Chebyshev's Inequality

### Example

Suppose we have a coin with bias  $\theta$ , i.e.  $p(X = 1) = \theta$

Say we flip the coin  $n$  times, and observe  $x_1, \dots, x_n \in \{0, 1\}$

We use th

<https://eduassistpro.github.io>

Estimate how large  $n$  should be such that

Add WeChat edu\_assist\_pro  
 $p(|\hat{\theta}_n - \theta| \geq 0.05) \leq 0.01$

1% probability of a 5% error

(Aside: the need for two parameters here is generic: “Probably Approximately Correct”)

## Chebyshev's Inequality

Example

Observe that

$$\text{Assignment Project Exam Help}$$
$$\frac{\mathbb{E}[\theta_n] = \frac{\sum_{i=1}^n \mathbb{E}[x_i]}{n} - \theta}{\frac{n \mathbb{V}[x_i]}{\theta(1-\theta)}}$$

Thus, app <https://eduassistpro.github.io>

Add WeChat edu\_assist\_pr

We are guaranteed this is less than 0.01 if

$$n \geq \frac{\theta(1-\theta)}{(0.05)^2(0.01)}.$$

When  $\theta = 0.5$ ,  $n \geq 10,000$  (!)

1 Properties of expectation and variance

2 Markov's inequality

3 Chebyshev's inequality

<https://eduassistpro.github.io>

4 Law of large numbers

Add WeChat edu\_assist\_pro

5 Wrapping Up

## Independent and Identically Distributed

Let  $X_1, \dots, X_n$  be random variables such that:

- Each  $X_i$  is independent of  $X_j$

# Assignment Project Exam Help

- The <https://eduassistpro.github.io>

Then, we say that  $X_1, \dots, X_n$  are **independ**  
(or **iid**)

Add WeChat edu\_assist\_pro

Example: For  $n$  independent flips of an unbiased coin from Bernoulli( $\frac{1}{2}$ )

1            n      re iid

## Law of Large Numbers

### Theorem

Let  $X_1, \dots, X_n$  be a sequence of iid random variables, with

$$\mathbb{E}[X_i] = \mu$$

and  $\mathbb{V}[X_i]$

<https://eduassistpro.github.io>

$$n = \frac{\text{sum of } X_i}{n}.$$

Then, for any  $\epsilon > 0$ ,

Add WeChat edu\_assist\_pro

$$\lim_{n \rightarrow \infty} p(|\bar{X}_n - \mu| > \epsilon) = 0.$$

Given enough trials, the empirical “success frequency” will be close to the expected value

## Law of Large Numbers

### Proof

Since the  $X_i$ 's are identically distributed,

**Assignment Project Exam Help**

Since the

<https://eduassistpro.github.io>

$$\begin{aligned} \frac{\mathbb{V}[X_1 + \dots + X_n]}{n^2} &= \frac{n\sigma^2}{n^2} \\ &= \frac{\sigma^2}{n}. \end{aligned}$$

Add WeChat `edu_assist_pro`

## Law of Large Numbers

Proof

Applying Chebyshev's inequality to  $\bar{X}_n$

$$\underline{\mathbb{V}[\bar{X}_n]}$$

<https://eduassistpro.github.io>

For any fixed  $\epsilon > 0$ , as  $n \rightarrow \infty$ , the right-hand side

Add WeChat edu\_assist\_pro

Thus,

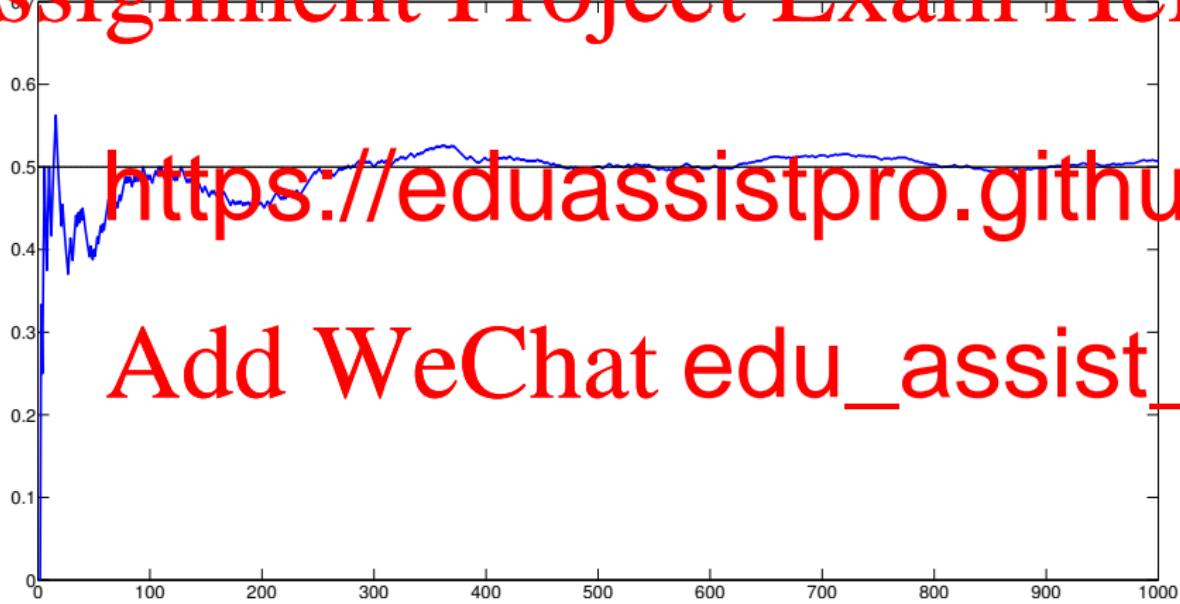
$$p(|\bar{X}_n - \mu| < \epsilon) \rightarrow 1.$$

# Law of Large Numbers

## Illustration

$N = 1000$  trials with Bernoulli random variable with parameter  $\frac{1}{2}$

# Assignment Project Exam Help

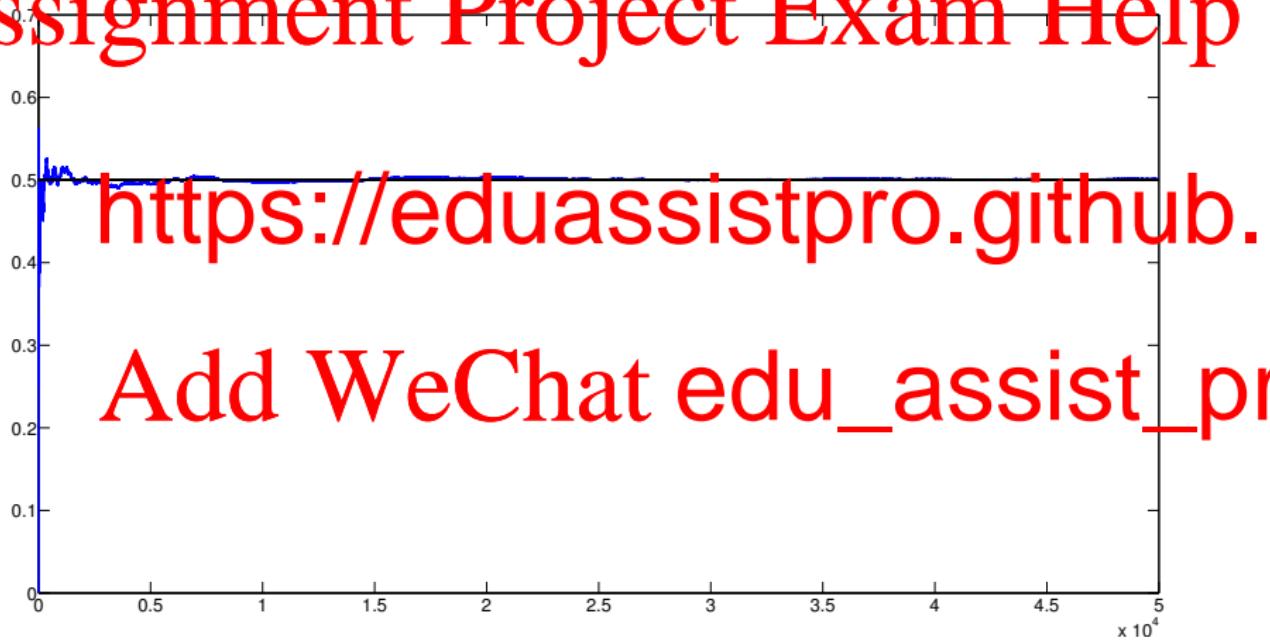


# Law of Large Numbers

## Illustration

$N = 50000$  trials with Bernoulli random variable with parameter  $\frac{1}{2}$

Assignment Project Exam Help



1 Properties of expectation and variance

2 Markov's inequality

3 Chebyshev's inequality

<https://eduassistpro.github.io>

4 Law of large numbers

Add WeChat edu\_assist\_pro

5 Wrapping Up

# Assignment Project Exam Help

- Markov's inequality,

- Ch <https://eduassistpro.github.io>

- Add WeChat edu\_assist\_pro

- Law of large numbers

Next time

# Assignment Project Exam Help

- Ensembles and sequences

- Typ <https://eduassistpro.github.io>

- Add WeChat edu\_assist\_pro

- Approximation Equipartition (AEP)