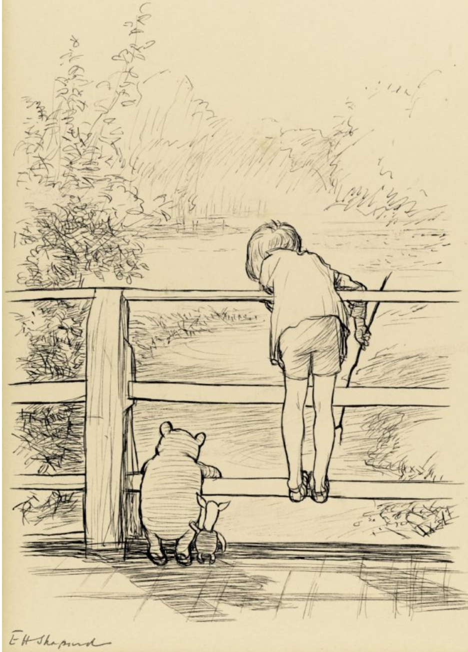
Topics in TCS

Probability overview

Raphaël Clifford


E H Shepard

Probability Overview

Prerequisites: the "Probability recap" lecture from Advanced Algorithms which is linked on the unit page. In particular you should already be comfortable with Sample Space, Events, Random Variables, Expected Value, Linearity of Expectation, Indicator Random Variables, Markov's Inequality. Also k-wise independent hash functions (Advanced Algorithms).

Probability Overview

Prerequisites: the "Probability recap" lecture from Advanced Algorithms which is linked on the unit page. In particular you should already be comfortable with Sample Space, Events, Random Variables, Expected Value, Linearity of Expectation, Indicator Random Variables, Markov's Inequality. Also k-wise independent hash functions (Advanced Algorithms).

In this probability overview we will cover:

- Markov's inequality (recap).

Probability Overview

Prerequisites: the “Probability recap” lecture from Advanced Algorithms which is linked on the unit page. In particular you should already be comfortable with Sample Space, Events, Random Variables, Expected Value, Linearity of Expectation, Indicator Random Variables, Markov's Inequality. Also k-wise independent hash functions (Advanced Algorithms).

In this probability overview we will cover:

- Markov's inequality (recap).
- Variance of a random variable.

Probability Overview

Prerequisites: the "Probability recap" lecture from Advanced Algorithms which is linked on the unit page. In particular you should already be comfortable with Sample Space, Events, Random Variables, Expected Value, Linearity of Expectation, Indicator Random Variables, Markov's Inequality. Also k-wise independent hash functions (Advanced Algorithms).

In this probability overview we will cover:

- Markov's inequality (recap).
- Variance of a random variable.
- Chebyshev's inequality.

Probability Overview

Prerequisites: the “Probability recap” lecture from Advanced Algorithms which is linked on the unit page. In particular you should already be comfortable with Sample Space, Events, Random Variables, Expected Value, Linearity of Expectation, Indicator Random Variables, Markov’s Inequality. Also k-wise independent hash functions (Advanced Algorithms).

In this probability overview we will cover:

- Markov’s inequality (recap).
- Variance of a random variable.
- Chebyshev’s inequality.
- Conditional probability.

Probability Overview

Prerequisites: the "Probability recap" lecture from Advanced Algorithms which is linked on the unit page. In particular you should already be comfortable with Sample Space, Events, Random Variables, Expected Value, Linearity of Expectation, Indicator Random Variables, Markov's Inequality. Also k-wise independent hash functions (Advanced Algorithms).

In this probability overview we will cover:

- Markov's inequality (recap).
- Variance of a random variable.
- Chebyshev's inequality.
- Conditional probability.
- Pairwise and full independence (recap).

Probability Overview

Prerequisites: the "Probability recap" lecture from Advanced Algorithms which is linked on the unit page. In particular you should already be comfortable with Sample Space, Events, Random Variables, Expected Value, Linearity of Expectation, Indicator Random Variables, Markov's Inequality. Also k-wise independent hash functions (Advanced Algorithms).

In this probability overview we will cover:

- Markov's inequality (recap).
- Variance of a random variable.
- Chebyshev's inequality.
- Conditional probability.
- Pairwise and full independence (recap).
- The Chernoff bound.

Probabilistic bounds

Recall that we can give a probabilistic bound for any non-negative random variable whose mean we can compute.

Theorem (Markov's inequality)

If X is a non-negative r.v., then for all $a > 0$,

$$\Pr(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

Probabilistic bounds

Recall that we can give a probabilistic bound for any non-negative random variable whose mean we can compute.

Theorem (Markov's inequality)

If X is a non-negative r.v., then for all $a > 0$,

$$\Pr(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

Example of Markov's inequality I

Suppose the average mark on a CS exam is 60%. Give an upper bound on the proportion of students who get at least 90%.

Probabilistic bounds

Recall that we can give a probabilistic bound for any non-negative random variable whose mean we can compute.

Theorem (Markov's inequality)

If X is a non-negative r.v., then for all $a > 0$,

$$\Pr(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

Example of Markov's inequality I

Suppose the average mark on a CS exam is 60%. Give an upper bound on the proportion of students who get at least 90%.

$$\Pr(X \geq 90) \leq \frac{\mathbb{E}(X)}{90} = \frac{60}{90} = \frac{2}{3}$$

How? $2/3$ could get 90 and then $1/3$ would have to get 0!

Probabilistic bounds

Recall that we can give a probabilistic bound for any non-negative random variable whose mean we can compute.

Theorem (Markov's inequality)

If X is a non-negative r.v., then for all $a > 0$,

$$\Pr(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

Example of Markov's inequality II

A coin has probability of landing on heads of 20%. If the coin is tossed 20 times, find a bound for the probability of getting at least 16 heads.

Probabilistic bounds

Recall that we can give a probabilistic bound for any non-negative random variable whose mean we can compute.

Theorem (Markov's inequality)

If X is a non-negative r.v., then for all $a > 0$,

$$\Pr(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

Example of Markov's inequality II

A coin has probability of landing on heads of 20%. If the coin is tossed 20 times, find a bound for the probability of getting at least 16 heads.

If the r.v. X is the number of heads then $\mathbb{E}(X) = 20/5 = 4$.

$$\Pr(X \geq 16) \leq \frac{\mathbb{E}(X)}{16} = \frac{4}{16} = \frac{1}{4}$$

Variance

The **mean** or **expectation** of random variable is defined to be:

$$\mathbb{E}(Y) = \sum_{x \in S} Y(x) \cdot \Pr(x)$$

where S is the sample space.

Variance

The **mean** or **expectation** of random variable is defined to be:

$$\mathbb{E}(Y) = \sum_{x \in S} Y(x) \cdot \Pr(x)$$

where S is the sample space.
The **variance** of a random variable is defined to be:

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

where $\mu_X = \mathbb{E}(X)$.

Variance

The **mean** or **expectation** of random variable is defined to be:

$$\mathbb{E}(Y) = \sum_{x \in S} Y(x) \cdot \Pr(x)$$

where S is the sample space.

The **variance** of a random variable is defined to be:

$$\mathrm{var}(X) = \mathbb{E}[(X - \mu_X)^2]$$

where $\mu_X = \mathbb{E}(X)$.

Example

Consider a r.v. X with $\Pr(X = 1) = p$ and $\Pr(X = 0) = 1 - p = q$.

$$\mu_X = 1 \cdot p + 0 \cdot q = p$$

$$\mathrm{var}(X) = (1 - p)^2 . p + (0 - p)^2 . q = pq$$

Variance

Definition of the variance

The **variance** of a random variable is defined to be:

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

or equivalently

$$\mathrm{var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

Variance

Definition of the variance

The **variance** of a random variable is defined to be:

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

or equivalently

$$\mathrm{var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

Example

Consider a r.v. X with $\Pr(X = -1) = P(X = 1) = \frac{1}{50}$, $P(X = 0) = \frac{24}{25}$.

$$\mu_X = -1 \cdot \frac{1}{50} + 0 \cdot \frac{24}{25} + 1 \cdot \frac{1}{50} = 0$$

$$\mathrm{var}(X) = (-1)^2 \frac{1}{50} + 0^2 \cdot \frac{24}{25} + 1^2 \cdot \frac{1}{50} = \frac{1}{25}$$

Variance

Definition of the variance

The **variance** of a random variable is defined to be:

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

or equivalently

$$\mathrm{var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

Example

For independent r.v. X, Y

$$\begin{aligned}
\mathbb{E}(XY) &= \mathbb{E}(X)\mathbb{E}(Y) \\
\mathrm{var}(X+Y) &= \mathbb{E}((X+Y)^2) - (\mathbb{E}(X+Y))^2 \\
&= \mathrm{var}(X) + \mathrm{var}(Y)
\end{aligned}$$

Chebyshev's inequality

With the variance we can often improve on Markov's inequality.

Theorem (Chebyshev's inequality)

If X is a real valued r.v., then for all k,

$$\Pr(|X - \mathbb{E}(X)| \geq k) \leq \frac{\mathrm{var}(X)}{k^2}$$

Chebyshev's inequality

With the variance we can often improve on Markov's inequality.

Theorem (Chebyshev's inequality)

If X is a real valued r.v., then for all k,

$$\Pr(|X - \mathbb{E}(X)| \geq k) \leq \frac{\mathrm{var}(X)}{k^2}$$

Example

Toss 100 fair coins, let X be the number of heads. Markov's inequality gives $\Pr(X \geq 75) \leq 2/3$. Using Chebyshev:

$$\begin{aligned} \mathrm{var}(X) = 100pq = 100 \cdot \frac{1}{2} \cdot \frac{1}{2} &= 25 \\ \Pr(|X - 50| \geq 25) \leq \frac{25}{25^2} &= \frac{1}{25} \end{aligned}$$

Proof of Chebyshev's inequality

Theorem (Chebyshev's inequality)

If X is a real valued r.v., then for all k,

$$\Pr(|X - \mathbb{E}(X)| \geq k) \leq \frac{\mathrm{var}(X)}{k^2}$$

Proof of Chebyshev's inequality

THEOREM (Chebyshev's inequality)

If X is a real valued r.v., then for all k,

$$\Pr(|X - \mathbb{E}(X)| \geq k) \leq \frac{\mathrm{var}(X)}{k^2}$$

Proof:
Let $Y = (X - \mathbb{E}(X))^2$. Y is non-negative and $\mathbb{E}(Y) = \mathrm{var}(X)$. By Markov's inequality,

$$\Pr(Y \geq k^2) \leq \frac{\mathbb{E}(Y)}{k^2} = \frac{\mathrm{var}(X)}{k^2}$$

Proof of Chebyshev's inequality

Theorem (Chebyshev's inequality)

If X is a real valued r.v., then for all k,

$$\Pr(|X - \mathbb{E}(X)| \geq k) \leq \frac{\operatorname{var}(X)}{k^2}$$

Proof:
Let $Y = (X - \mathbb{E}(X))^2$. Y is non-negative and $\mathbb{E}(Y) = \operatorname{var}(X)$. By Markov's inequality,

$$\Pr(Y \geq k^2) \leq \frac{\mathbb{E}(Y)}{k^2} = \frac{\operatorname{var}(X)}{k^2}$$

Notice that the event $Y \geq k^2$ is the same as $|X - \mathbb{E}(X)| \geq k$, so

$$\Pr(|X - \mathbb{E}(X)| \geq k) \leq \frac{\operatorname{var}(X)}{k^2}.$$

$\square$

Conditional probability and independence

Let B be an event such that $\Pr(B) \neq 0$. The *conditional probability* of an event A given B is defined as

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A, B)}{\Pr(B)}$$

It follows that $\Pr(A, B) = \Pr(A)\Pr(B \mid A) = \Pr(B)\Pr(A \mid B)$

Conditional probability and independence

Let B be an event such that $\Pr(B) \neq 0$. The *conditional probability* of an event A given B is defined as

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A, B)}{\Pr(B)}$$

It follows that $\Pr(A, B) = \Pr(A)\Pr(B \mid A) = \Pr(B)\Pr(A \mid B)$

Conditional probability example

Suppose you pick a uniformly random integer from $\{1, \ldots, 100\}$. If A is the event that the last digit is a 3 then $\Pr(A) = 1/10$. If B is the event that the number is prime then $\Pr(B) = 1/4$.

$$\Pr(A \mid B) = \frac{\Pr(A, B)}{\Pr(B)} = \frac{7/100}{1/4} = \frac{28}{100}$$

Conditional probability and independence

Let B be an event such that $\Pr(B) \neq 0$. The *conditional probability* of an event A given B is defined as

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A, B)}{\Pr(B)}$$

It follows that $\Pr(A, B) = \Pr(A) \Pr(B \mid A) = \Pr(B) \Pr(A \mid B)$

2,3,5,7,11,13,17,19,23,29,31,37,41,43,47,53,59,61,67,71,73,79,83,89,97

Conditional probability example

Suppose you pick a uniformly random integer from $\{1, \ldots, 100\}$. If A is the event that the last digit is a 3 then $\Pr(A) = 1/10$. If B is the event that the number is prime then $\Pr(\text{B}) = 1/4$.

$$\Pr(A \mid B) = \frac{\Pr(A, B)}{\Pr(B)} = \frac{7/100}{1/4} = \frac{28}{100}$$

Conditional probability and independence

Let B be an event such that $\Pr(B) \neq 0$. The *conditional probability* of an event A given B is defined as

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A, B)}{\Pr(B)}$$

It follows that $\Pr(A, B) = \Pr(A)\Pr(B \mid A) = \Pr(B)\Pr(A \mid B)$

Conditional probability example

Suppose you pick a uniformly random integer from $\{1, \dots, 100\}$. If A is the event that the last digit is a 3 then $\Pr(A) = 1/10$. If B is the event that the number is prime then $\Pr(B) = 1/4$.

2,3,5,7,11,13,17,19,23,29,31,37,41,43,47,53,59,61,67,71,73,79,83,89,97

$$\Pr(A \mid B) = \frac{\Pr(A, B)}{\Pr(B)} = \frac{7/100}{1/4} = \frac{28}{100}$$

3,13,23,43,53,73,83

Conditional probability and independence

Let B be an event such that $\Pr(B) \neq 0$. The *conditional probability* of an event A given B is defined as

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A, B)}{\Pr(B)}$$

It follows that $\Pr(A, B) = \Pr(A) \Pr(B \mid A) = \Pr(B) \Pr(A \mid B)$

Two r.v.s A and B are said to be *independent* if

$$\Pr(A, B) = \Pr(A) \Pr(B)$$

If $\Pr(B) \neq 0$, this is equivalent to

$$\Pr(A \mid B) = \Pr(A)$$

Pairwise and full independence

Consider the sample space

$$S = \{abc, acb, cab, cba, bca, bac, aaa, bbb, ccc\}$$

Suppose that each of the nine elementary events in S occurs with equal probability $\frac{1}{9}$. Let A_k be the event that the kth letter is a.

Pairwise and full independence

Consider the sample space

$$S = \{abc, acb, cab, cba, bca, bac, aaa, bbb, ccc\}$$

Suppose that each of the nine elementary events in S occurs with equal probability $\frac{1}{9}$. Let A_k be the event that the kth letter is a.

$$\Pr(A_1|A_2) = \frac{1}{3} = \Pr(A_1) = \frac{3}{9} \qquad (A_1 \text{ and } A_2 \text{ are independent})$$

$$\Pr(A_2|A_3) = \frac{1}{3} = \Pr(A_2) = \frac{3}{9} \qquad (A_2 \text{ and } A_3 \text{ are independent})$$

$$\Pr(A_1|A_3) = \frac{1}{3} = \Pr(A_1) = \frac{3}{9} \qquad (A_1 \text{ and } A_3 \text{ are independent})$$

Pairwise and full independence

Consider the sample space

$$S = \{abc, acb, cab, cba, bca, bac, aaa, bbb, ccc\}$$

Suppose that each of the nine elementary events in S occurs with equal probability $\frac{1}{9}$. Let A_k be the event that the kth letter is a.

$$\Pr(A_1|A_2) = \frac{1}{3} = \Pr(A_1) = \frac{3}{9} \qquad (A_1 \text{ and } A_2 \text{ are independent})$$
$$\Pr(A_2|A_3) = \frac{1}{3} = \Pr(A_2) = \frac{3}{9} \qquad (A_2 \text{ and } A_3 \text{ are independent})$$
$$\Pr(A_1|A_3) = \frac{1}{3} = \Pr(A_1) = \frac{3}{9} \qquad (A_1 \text{ and } A_3 \text{ are independent})$$

BUT: $\Pr(A_1|A_2, A_3) = 1 \neq \Pr(A_1)$. (A_1, A_2, A_3) are pairwise independent but not (fully) independent.

Key facts about Expectation and Variance

1. $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$. (Linearity of expectation).

Key facts about Expectation and Variance

1. $\mathbb{E}(X+Y)=\mathbb{E}(X)+\mathbb{E}(Y)$. (Linearity of expectation).
2. $\mathbb{E}(aX)=a\mathbb{E}(X)$ for $a \in \mathbb{R}$.

Key facts about Expectation and Variance

1. $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$. (Linearity of expectation).

2. $\mathbb{E}(aX) = a\mathbb{E}(X)$ for $a \in \mathbb{R}$.

3. Let $A_1, \ldots, A_n$ be disjoint nonempty events that form a partition of Ω, then

$$\mathbb{E}(X) = \sum_{i=1}^{n} \Pr(A_i)\mathbb{E}(X \mid A_i) \qquad \text{(Law of total expectation)}$$

Key facts about Expectation and Variance

1. $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$. (Linearity of expectation).

2. $\mathbb{E}(aX) = a\mathbb{E}(X)$ for $a \in \mathbb{R}$.

3. Let $A_1, \ldots, A_n$ be disjoint nonempty events that form a partition of Ω, then

$$\mathbb{E}(X) = \sum_{i=1}^{n} \Pr(A_i)\mathbb{E}(X \mid A_i) \qquad \text{(Law of total expectation)}$$

4. If X and Y are independent then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Key facts about Expectation and Variance

1. $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$. (Linearity of expectation).
2. $\mathbb{E}(aX) = a\mathbb{E}(X)$ for $a \in \mathbb{R}$.
3. Let $A_1, \ldots, A_n$ be disjoint nonempty events that form a partition of Ω, then

$$\mathbb{E}(X) = \sum_{i=1}^{n} \Pr(A_i)\mathbb{E}(X \mid A_i) \qquad \text{(Law of total expectation)}$$

4. If X and Y are independent then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.
5. Variance **cannot** be negative.

Key facts about Expectation and Variance

1. $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$. (Linearity of expectation).

2. $\mathbb{E}(aX) = a\mathbb{E}(X)$ for $a \in \mathbb{R}$.

3. Let $A_1, \ldots, A_n$ be disjoint nonempty events that form a partition of Ω, then

$$\mathbb{E}(X) = \sum_{i=1}^{n} \Pr(A_i)\mathbb{E}(X \mid A_i) \qquad \text{(Law of total expectation)}$$

4. If X and Y are independent then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

5. Variance **cannot** be negative.

6. If X and Y are independent then $\mathrm{var}(X + Y) = \mathrm{var}(X) + \mathrm{var}(Y)$.

Key facts about Expectation and Variance

1. $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$. (Linearity of expectation).
2. $\mathbb{E}(aX) = a\mathbb{E}(X)$ for $a \in \mathbb{R}$.
3. Let $A_1, \ldots, A_n$ be disjoint nonempty events that form a partition of Ω, then

   $$\mathbb{E}(X) = \sum_{i=1}^{n} \Pr(A_i)\mathbb{E}(X \mid A_i) \qquad \text{(Law of total expectation)}$$

4. If X and Y are independent then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.
5. Variance **cannot** be negative.
6. If X and Y are independent then $\mathrm{var}(X + Y) = \mathrm{var}(X) + \mathrm{var}(Y)$.
7. $\mathrm{var}(aX) = a^2\,\mathrm{var}(X)$ for $a \in \mathbb{R}$.

Tighter bounds - Chernoff

Theorem (Chernoff bound)

Consider fully independent indicator r.v.s $X_1, X_2, \ldots, X_n$ and $X = \sum_{i=1}^{n} X_i$. Let $\mu = \mathbb{E}(X)$. For any $\delta > 0$,

$$\Pr[X \geq (1+\delta)\mu] \leq \exp\left(-\delta^2\mu/3\right)$$

$$\Pr[X \leq (1-\delta)\mu] \leq \exp\left(-\delta^2\mu/2\right)$$

Tighter bounds - Chernoff

Theorem (Chernoff bound)

Consider fully independent indicator r.v.s $X_1, X_2, \ldots, X_n$ and $X = \sum_{i=1}^{n} X_i$. Let $\mu = \mathbb{E}(X)$. For any $\delta > 0$,

$$\Pr[X \geq (1+\delta)\mu] \leq \exp\left(-\delta^2\mu/3\right)$$
$$\Pr[X \leq (1-\delta)\mu] \leq \exp\left(-\delta^2\mu/2\right)$$

Chernoff bound example I

Toss 100 fair coins and let X be the number of heads.

1. Markov: $\Pr(X \geq 75) \leq \mathbb{E}(X)/75 = 2/3$.
2. Chebyshev: $\Pr(X \geq 75) \leq \Pr((|X - \mathbb{E}(X)| \geq 25) \leq \frac{1}{25}$.
3. Chernoff: $\Pr(X \geq (1+1/2)50) \leq e^{-50/(4\cdot 3)} \approx 1/64$
4. True answer: ≈ 0.00000028.

Tighter bounds - Chernoff

Theorem (Chernoff bound)

Consider fully independent indicator r.v.s $X_1, X_2, \ldots, X_n$ and $X = \sum_{i=1}^{n} X_i$. Let $\mu = \mathbb{E}(X)$. For any $\delta > 0$,

$$\Pr[X \geq (1+\delta)\mu] \leq \exp\left(-\delta^2\mu/3\right)$$
$$\Pr[X \leq (1-\delta)\mu] \leq \exp\left(-\delta^2\mu/2\right)$$

Chernoff bound example II

Toss 1000 coins and let X be the number of heads.

1. Markov: $\Pr(X \geq 750) \leq 2/3$.
2. Chebyshev: $\Pr(X \geq 750) \leq \frac{1}{250}$.
3. Chernoff:
   $\Pr(X \geq (1+1/2)500) \leq e^{-500/(4\cdot 3)} \approx 8.024 \cdot 10^{-19}$
4. True answer: $\approx 6.738 \cdot 10^{-59}$

Summary

- You will need the probability slides from AA and this set of slides for this unit.

Summary

- You will need the probability slides from AA and this set of slides for this unit.
- In these slides the new topics we covered are:

Summary

- You will need the probability slides from AA and this set of slides for this unit.

- In these slides the new topics we covered are:
  1. Variance of a random variable

Summary

- You will need the probability slides from AA and this set of slides for this unit.

- In these slides the new topics we covered are:
  1. Variance of a random variable

  2. Chebyshev's inequality

Summary

- You will need the probability slides from AA and this set of slides for this unit.

- In these slides the new topics we covered are:
  1. Variance of a random variable
  2. Chebyshev's inequality
  3. Conditional probability

Summary

- You will need the probability slides from AA and this set of slides for this unit.
- In these slides the new topics we covered are:
  1. Variance of a random variable
  2. Chebyshev's inequality
  3. Conditional probability
  4. The Chernoff bound.