Moments and Deviations

K. Subramani

1 Lane Department of Computer Science and Electrical Engineering
West Virginia University

9 February, 2012
Outline

1. Recap
Outline

1 Recap

2 Tail bounds
Outline

1 Recap

2 Tail bounds

3 Markov’s inequality
Outline

1 Recap
2 Tail bounds
3 Markov’s inequality
4 Chebyshev’s Inequality
Main points

Probability spaces, Random Variable,
Main points

Probability spaces, Random Variable, Distribution of a random variable (pmf),
Recap

Main points
Probability spaces, Random Variable, Distribution of a random variable (pmf), Expected Value,
Main points

Probability spaces, Random Variable, Distribution of a random variable (pmf), Expected Value, Variance,
Recap

Main points

Probability spaces, Random Variable, Distribution of a random variable (pmf), Expected Value, Variance, Samples of randomized algorithms.
Tail bounds
Note

The tail bounds of a random variable $X$ are concerned with the probability that it deviates significantly from its expected value $E[X]$ on a run of the experiment.
Note

The tail bounds of a random variable $X$ are concerned with the probability that it deviates significantly from its expected value $E[X]$ on a run of the experiment.
The tail bounds of a random variable $X$ are concerned with the probability that it deviates significantly from its expected value $E[X]$ on a run of the experiment.

Example
Consider the experiment of tossing a fair coin $n$ times.
Tail bounds

**Note**

*The tail bounds of a random variable $X$ are concerned with the probability that it deviates significantly from its expected value $E[X]$ on a run of the experiment.*

**Example**

Consider the experiment of tossing a fair coin $n$ times. What is the probability that the number of heads exceeds $\frac{3}{4} \cdot n$?
Markov’s inequality
Theorem

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,
Markov’s inequality

Theorem

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

$$P(X \geq c) \leq \frac{E[X]}{c}.$$
Markov’s inequality

Theorem

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

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

Proof.

$$E[X] = \sum_{x} x \cdot P(X = x)$$
Markov’s inequality

**Theorem**

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then, $P(X \geq c) \leq \frac{E[X]}{c}$.

**Proof.**

\[
E[X] = \sum_{x} x \cdot P(X = x)
\]

\[
= \sum_{0 \leq x < c} x \cdot P(X = x) + \sum_{x \geq c} x \cdot P(X = x)
\]
Theorem

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,
$$P(X \geq c) \leq \frac{E[X]}{c}.$$ 

Proof.

$$E[X] = \sum_{x} x \cdot P(X = x)$$
$$= \sum_{0 \leq x < c} x \cdot P(X = x) + \sum_{x \geq c} x \cdot P(X = x)$$
$$\geq \sum_{x \geq c} x \cdot P(X = x)$$
Markov’s inequality

**Theorem**

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

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

**Proof.**

$$E[X] = \sum_x x \cdot P(X = x)$$

$$= \sum_{0 \leq x < c} x \cdot P(X = x) + \sum_{x \geq c} x \cdot P(X = x)$$

$$\geq \sum_{x \geq c} x \cdot P(X = x)$$

$$\geq \sum_{x \geq c} c \cdot P(X = x)$$
Theorem

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

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

Proof.

$$E[X] = \sum_x x \cdot P(X = x)$$

$$= \sum_{0 \leq x < c} x \cdot P(X = x) + \sum_{x \geq c} x \cdot P(X = x)$$

$$\geq \sum_{x \geq c} x \cdot P(X = x)$$

$$\geq \sum_{x \geq c} c \cdot P(X = x)$$

$$= c \cdot P(X \geq c)$$
Markov’s inequality

**Theorem**

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

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

**Proof.**

\[
E[X] = \sum_x x \cdot P(X = x) \\
= \sum_{0 \leq x < c} x \cdot P(X = x) + \sum_{x \geq c} x \cdot P(X = x) \\
\geq \sum_{x \geq c} x \cdot P(X = x) \\
\geq \sum_{x \geq c} c \cdot P(X = x) \\
= c \cdot P(X \geq c) \\
\Rightarrow P(X \geq c) \leq \frac{E[X]}{c}
\]
Markov's Inequality (contd.)
Markov’s Inequality (contd.)

Alternative Form
Alternative Form

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,
Alternative Form

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

$$P(X \geq c \cdot E[X]) \leq \frac{1}{c}.$$
Markov’s Inequality (contd.)

Alternative Form

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then, $P(X \geq c \cdot E[X]) \leq \frac{1}{c}$.

Example (Application to coin tossing problem)
Markov’s Inequality (contd.)

Alternative Form

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

$$ P(X \geq c \cdot E[X]) \leq \frac{1}{c}. $$

Example (Application to coin tossing problem)

$$ P(X \geq \frac{3n}{4}) = P(X \geq \frac{3}{2} \cdot \frac{n}{2}) $$
Markov’s Inequality (contd.)

**Alternative Form**

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then, $P(X \geq c \cdot E[X]) \leq \frac{1}{c}$.

**Example (Application to coin tossing problem)**

\[
P(X \geq \frac{3n}{4}) = P(X \geq \frac{3}{2} \cdot \frac{n}{2}) \leq \frac{1}{\frac{3}{2}}
\]
Markov’s Inequality (contd.)

Alternative Form

Let $X$ be a non-negative random variable and let $c > 0$ be a positive constant. Then,

$$P(X \geq c \cdot E[X]) \leq \frac{1}{c}.$$ 

Example (Application to coin tossing problem)

$$P(X \geq \frac{3n}{4}) = P(X \geq \frac{3}{2} \cdot \frac{n}{2})$$

$$\leq \frac{1}{\frac{3}{2}}$$

$$= \frac{2}{3}$$
Chebyshev’s Inequality
Chebyshev’s Inequality

**Theorem**

*Let X be a random variable (not necessarily positive).*
Theorem

Let $X$ be a random variable (not necessarily positive). Then, $P(|X - E[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2}$. 
**Chebyshev’s Inequality**

**Theorem**

Let $X$ be a random variable (not necessarily positive). Then, $P(|X - E[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2}$.

**Proof.**

$$P(|X - E[X]| \geq a) =$$
Chebyshev’s Inequality

**Theorem**

Let $X$ be a random variable (not necessarily positive). Then, $P(|X - E[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2}$.

**Proof.**

$$P(|X - E[X]| \geq a) = P(|X - E[X]|^2 \geq a^2)$$
Chebyshev’s Inequality

Theorem

Let $X$ be a random variable (not necessarily positive). Then, $P(|X - E[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2}$.

Proof.

\[
P(|X - E[X]| \geq a) = P(|X - E[X]|^2 \geq a^2) \leq \frac{E[(X - E[X])^2]}{a^2}, \text{ Markov’s inequality}
\]
Chebyshev’s Inequality

Theorem

Let $X$ be a random variable (not necessarily positive). Then, $P(|X - E[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2}$.

Proof.

$$P(|X - E[X]| \geq a) = P(|X - E[X]|^2 \geq a^2) \leq \frac{E[(X - E[X])^2]}{a^2}, \text{ Markov's inequality}$$

$$= \frac{\text{Var}[X]}{a^2}$$
**Chebyshev’s Inequality**

**Theorem**

Let $X$ be a random variable (not necessarily positive). Then, $P(|X - E[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2}$.

**Proof.**

\[
P(|X - E[X]| \geq a) = P(|X - E[X]|^2 \geq a^2) \leq \frac{E[(X - E[X])^2]}{a^2}, \text{ Markov’s inequality} \]

\[
= \frac{\text{Var}[X]}{a^2}
\]

**Note**

Chebyshev’s theorem is alternatively stated as:
Chebyshev’s Inequality

**Theorem**

Let $X$ be a random variable (not necessarily positive). Then, $P(|X - E[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2}$.

**Proof.**

\[
P(|X - E[X]| \geq a) = P(|X - E[X]|^2 \geq a^2) \\
\leq \frac{E[(X - E[X])^2]}{a^2}, \text{ Markov’s inequality} \\
= \frac{\text{Var}[X]}{a^2}
\]

**Note**

Chebyshev’s theorem is alternatively stated as:

\[
P(|X - E[X]| \geq a \cdot E[X]) \leq \frac{\text{Var}[X]}{(a \cdot E[X])^2}.
\]
Chebyshev’s inequality (contd.)

Example (Application to coin tossing problem)
Example (Application to coin tossing problem)

\[ P(X \geq \frac{3n}{4}) = \]
Chebyshev’s inequality (contd.)

Example (Application to coin tossing problem)

\[ P(X \geq \frac{3n}{4}) = P(X - \frac{n}{2} \geq \frac{n}{4}) \]
Example (Application to coin tossing problem)

\[ P(X \geq \frac{3n}{4}) = P(X - \frac{n}{2} \geq \frac{n}{4}) \leq P(|X - \frac{n}{2}| \geq \frac{n}{4}) \]
Example (Application to coin tossing problem)

\[ P(X \geq \frac{3n}{4}) = P(X - \frac{n}{2} \geq \frac{n}{4}) \leq P(|X - \frac{n}{2}| \geq \frac{n}{4}) = P(|X - E[X]| \geq \frac{1}{2}E[X]) \]
Example (Application to coin tossing problem)

\[ P(X \geq \frac{3n}{4}) = P(X - \frac{n}{2} \geq \frac{n}{4}) \]
\[ \leq P(|X - \frac{n}{2}| \geq \frac{n}{4}) \]
\[ = P(|X - E[X]| \geq \frac{1}{2}E[X]) \]
\[ \leq \frac{n}{4} \cdot \left( \frac{1}{2} \right)^2 \cdot \left( \frac{n}{2} \right)^2 \]
Chebyshev’s inequality (contd.)

Example (Application to coin tossing problem)

\[
P(X \geq \frac{3n}{4}) = P(X - \frac{n}{2} \geq \frac{n}{4})
\]
\[
\leq P(|X - \frac{n}{2}| \geq \frac{n}{4})
\]
\[
= P(|X - E[X]| \geq \frac{1}{2}E[X])
\]
\[
\leq \frac{n}{4} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{n}{2}\right)^2
\]
\[
= \frac{4}{n}
\]
The coupon collecting problem
The coupon collecting problem

Restatement
The coupon collecting problem

Restatement

You are required to collect coupons in a series of iterations.
The coupon collecting problem

Restatement
You are required to collect coupons in a series of iterations. Assume that each coupon belongs to one of \( n \) types, where \( n \) is a fixed number.
The coupon collecting problem

Restatement

You are required to collect coupons in a series of iterations. Assume that each coupon belongs to one of $n$ types, where $n$ is a fixed number. The coupons are drawn uniformly and at random from the $n$ coupon types.
The coupon collecting problem

Restatement

You are required to collect coupons in a series of iterations. Assume that each coupon belongs to one of $n$ types, where $n$ is a fixed number. The coupons are drawn uniformly and at random from the $n$ coupon types. What is the expected number of coupons to be collected, to ensure that each coupon type has been collected?
The coupon collecting problem

Restatement

You are required to collect coupons in a series of iterations. Assume that each coupon belongs to one of $n$ types, where $n$ is a fixed number. The coupons are drawn uniformly and at random from the $n$ coupon types. What is the expected number of coupons to be collected, to ensure that each coupon type has been collected?

Remark

Let $X$ denote the number of coupons to be collected in order to ensure that we have one coupon of each type.
The coupon collecting problem

**Restatement**

You are required to collect coupons in a series of iterations. Assume that each coupon belongs to one of $n$ types, where $n$ is a fixed number. The coupons are drawn uniformly and at random from the $n$ coupon types. What is the expected number of coupons to be collected, to ensure that each coupon type has been collected?

**Remark**

Let $X$ denote the number of coupons to be collected in order to ensure that we have one coupon of each type. We have shown that $E[X] = n \cdot H_n$, where $H_n$ is the $n^{th}$ harmonic number.
Tail bounds for coupon collecting
## Tail bounds for coupon collecting

<table>
<thead>
<tr>
<th>Markov</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X \geq 2 \cdot n \cdot H_n)$</td>
</tr>
</tbody>
</table>

- **Recap**
- **Tail bounds**
- **Markov's inequality**
- **Chebyshev's inequality**
Markov

\[ P(X \geq 2 \cdot n \cdot H_n) \leq \frac{1}{2} \]
Tail bounds for coupon collecting

Markov

\[ P(X \geq 2 \cdot n \cdot H_n) \leq \frac{1}{2} \]

Chebyshev

What do we need?
### Markov

\[ P(X \geq 2 \cdot n \cdot H_n) \leq \frac{1}{2} \]

### Chebyshev

What do we need? \( \text{Var}[X] \).
### Markov

$$P(X \geq 2 \cdot n \cdot H_n) \leq \frac{1}{2}$$

### Chebyshev

What do we need? $\text{Var}[X]$. Observe that $\text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i]$, where $X_i$ is the random variable which counts the number of coupons to be drawn assuming that $(i-1)$ distinct types have already been drawn, in order to draw a coupon of a new type.
### Tail bounds for coupon collecting

#### Markov

\[ P(X \geq 2 \cdot n \cdot H_n) \leq \frac{1}{2} \]

#### Chebyshev

What do we need? \( \text{Var}[X] \). Observe that \( \text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i] \), where \( X_i \) is the random variable which counts the number of coupons to be drawn assuming that \((i - 1)\) distinct types have already been drawn, in order to draw a coupon of a new type.

For a geometric variable \( X_i \) with parameter \( p \), we know that \( \text{Var}[X_i] = \frac{1-p_i}{p_i^2} \).
Tail bounds for coupon collecting

Markov

\[ P(X \geq 2 \cdot n \cdot H_n) \leq \frac{1}{2} \]

Chebyshev

What do we need? \( \text{Var}[X] \). Observe that \( \text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i] \), where \( X_i \) is the random variable which counts the number of coupons to be drawn assuming that \((i-1)\) distinct types have already been drawn, in order to draw a coupon of a new type.

For a geometric variable \( X_i \) with parameter \( p \), we know that \( \text{Var}[X_i] = \frac{1-p_i}{p_i^2} \leq \frac{1}{p_i^2} \).
Tail bounds for coupon collecting

Markov

\[ P(X \geq 2 \cdot n \cdot H_n) \leq \frac{1}{2} \]

Chebyshev

What do we need? \( \text{Var}[X] \). Observe that \( \text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i] \), where \( X_i \) is the random variable which counts the number of coupons to be drawn assuming that \( (i-1) \) distinct types have already been drawn, in order to draw a coupon of a new type.

For a geometric variable \( X_i \) with parameter \( p \), we know that \( \text{Var}[X_i] = \frac{1-p_i}{p_i^2} \leq \frac{1}{p_i^2} \).

But recall that \( p_i = \frac{n-i+1}{n} \).
Tail bounds for coupon collecting

**Markov**

\[ P(X \geq 2 \cdot n \cdot H_n) \leq \frac{1}{2} \]

**Chebyshev**

What do we need? \( \text{Var}[X] \). Observe that \( \text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i] \), where \( X_i \) is the random variable which counts the number of coupons to be drawn assuming that \( (i-1) \) distinct types have already been drawn, in order to draw a coupon of a new type.

For a geometric variable \( X_i \) with parameter \( p \), we know that \( \text{Var}[X_i] = \frac{1-p_i}{p_i^2} \leq \frac{1}{p_i^2} \).

But recall that \( p_i = \frac{n-i+1}{n} \). Therefore, \( \frac{1}{p_i} = \frac{n}{n-i+1} \).
Tail bounds for coupon collecting

**Markov**

\[ P(X \geq 2 \cdot n \cdot H_n) \leq \frac{1}{2} \]

**Chebyshev**

What do we need? \( \text{Var}[X] \). Observe that \( \text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i] \), where \( X_i \) is the random variable which counts the number of coupons to be drawn assuming that \((i - 1)\) distinct types have already been drawn, in order to draw a coupon of a new type.

For a geometric variable \( X_i \) with parameter \( p \), we know that \( \text{Var}[X_i] = \frac{1-p_i}{p_i^2} \leq \frac{1}{p_i^2} \).

But recall that \( p_i = \frac{n-i+1}{n} \). Therefore, \( \frac{1}{p_i} = \frac{n}{n-i+1} \). Hence,

\[ \text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i] \]
### Markov

\[ P(X \geq 2 \cdot n \cdot H_n) \leq \frac{1}{2} \]

### Chebyshev

What do we need? \( \text{Var}[X] \). Observe that \( \text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i] \), where \( X_i \) is the random variable which counts the number of coupons to be drawn assuming that \( (i - 1) \) distinct types have already been drawn, in order to draw a coupon of a new type.

For a geometric variable \( X_i \) with parameter \( p \), we know that \( \text{Var}[X_i] = \frac{1-p_i}{p_i^2} \leq \frac{1}{p_i^2} \).

But recall that \( p_i = \frac{n-i+1}{n} \). Therefore, \( \frac{1}{p_i} = \frac{n}{n-i+1} \). Hence,

\[
\text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i] \\
\leq \sum_{i=1}^{n} \frac{1}{p_i^2}
\]
Tail bounds for coupon collecting (contd.)

Chebyshev (contd.)
Tail bounds for coupon collecting (contd.)

Chebyshev (contd.)

\[ \text{Var}[X] \leq \sum_{i=1}^{n} \left( \frac{n}{n - i + 1} \right)^2 \]
Tail bounds for coupon collecting (contd.)

Chebyshev (contd.)

\[ \text{Var}[X] \leq \sum_{i=1}^{n} \left( \frac{n}{n-i+1} \right)^2 \]

\[ = n^2 \cdot \sum_{i=1}^{n} \left( \frac{1}{n-i+1} \right)^2 \]
Tail bounds for coupon collecting (contd.)

Chebyshev (contd.)

\[
\Var[X] \leq \sum_{i=1}^{n} \left( \frac{n}{n-i+1} \right)^2 \\
= n^2 \cdot \sum_{i=1}^{n} \left( \frac{1}{n-i+1} \right)^2 \\
= n^2 \cdot \sum_{i=1}^{n} \frac{1}{i^2}
\]
Tail bounds for coupon collecting (contd.)

Chebyshev (contd.)

$$\text{Var}[X] \leq \sum_{i=1}^{n} \left( \frac{n}{n-i+1} \right)^2$$

$$= n^2 \cdot \sum_{i=1}^{n} \left( \frac{1}{n-i+1} \right)^2$$

$$= n^2 \cdot \sum_{i=1}^{n} \frac{1}{i^2}$$

$$\leq n^2 \cdot \frac{\pi^2}{6}$$
Tail bounds for coupon collecting (contd.)

Analysis (contd.)
It follows that

\[ P(X \geq 2 \cdot n \cdot H_n) = \]
Tail bounds for coupon collecting (contd.)

Analysis (contd.)

It follows that

\[ P(X \geq 2 \cdot n \cdot H_n) = P((X - n \cdot H_n) \geq n \cdot H_n) \]
It follows that

\[ P(X \geq 2 \cdot n \cdot H_n) = P((X - n \cdot H_n) \geq n \cdot H_n) \leq P(|X - n \cdot H_n| \geq n \cdot H_n) \]
Recap
Tail bounds
Markov’s inequality
Chebyshev’s Inequality

Tail bounds for coupon collecting (contd.)

Analysis (contd.)

It follows that

\[ P(X \geq 2 \cdot n \cdot H_n) = P((X - n \cdot H_n) \geq n \cdot H_n) \leq P(|X - n \cdot H_n| \geq n \cdot H_n) \leq \frac{n^2 \cdot \pi^2}{6(n \cdot H_n)^2} \]
It follows that

\[ P(X \geq 2 \cdot n \cdot H_n) = P((X - n \cdot H_n) \geq n \cdot H_n) \leq P(|X - n \cdot H_n| \geq n \cdot H_n) \leq \frac{n^2 \cdot \pi^2}{6 (n \cdot H_n)^2} \leq O\left(\frac{1}{\ln^2 n}\right) \]
Tail bounds (first principles)
Tail bounds (first principles)

First principles

Focus on a coupon of type $i$. 

Focus on a coupon of type $i$. What is the probability that a coupon of type $i$ has not been drawn after $n \cdot \ln n + c \cdot n$ trials?
Focus on a coupon of type $i$. What is the probability that a coupon of type $i$ has not been drawn after $n \cdot \ln n + c \cdot n$ trials?

$$(1 - \frac{1}{n})^{n \cdot \ln n + c \cdot n} =$$
Tail bounds (first principles)

First principles

Focus on a coupon of type $i$. What is the probability that a coupon of type $i$ has not been drawn after $n \cdot \ln n + c \cdot n$ trials?

$$(1 - \frac{1}{n})^{n \cdot \ln n + c \cdot n} = (1 - \frac{1}{n})^{n \cdot (\ln n + c)}$$
Tail bounds (first principles)

Focus on a coupon of type $i$. What is the probability that a coupon of type $i$ has not been drawn after $n \cdot \ln n + c \cdot n$ trials?

$$(1 - \frac{1}{n})^{n \cdot \ln n + c \cdot n} = (1 - \frac{1}{n})^{n \cdot (\ln n + c)} \leq e^{-1 \cdot (\ln n + c)}$$
Tail bounds (first principles)

Focus on a coupon of type $i$. What is the probability that a coupon of type $i$ has not been drawn after $n \cdot \ln n + c \cdot n$ trials?

\[
(1 - \frac{1}{n})^{n \cdot \ln n + c \cdot n} = (1 - \frac{1}{n})^{n \cdot (\ln n + c)}
\]
\[
\leq e^{-1 \cdot (\ln n + c)}
\]
\[
= \frac{1}{e^c \cdot n}
\]
Tail bounds (first principles)

First principles

Focus on a coupon of type $i$. What is the probability that a coupon of type $i$ has not been drawn after $n \cdot \ln n + c \cdot n$ trials?

$$
(1 - \frac{1}{n})^{n\cdot\ln n + c \cdot n} = (1 - \frac{1}{n})^{n \cdot (\ln n + c)} \\
\leq e^{-1 \cdot (\ln n + c)} \\
= \frac{1}{e^c \cdot n}
$$

What is the probability that a coupon of any type has not been drawn after $n \cdot \ln n + c \cdot n$ trials?
Tail bounds (first principles)

First principles

Focus on a coupon of type \( i \). What is the probability that a coupon of type \( i \) has not been drawn after \( n \cdot \ln n + c \cdot n \) trials?

\[
\left( 1 - \frac{1}{n} \right)^{n \cdot \ln n + c \cdot n} = \left( 1 - \frac{1}{n} \right)^{n \cdot (\ln n + c)} \\
\leq e^{-1 \cdot (\ln n + c)} = \frac{1}{e^{c \cdot n}}
\]

What is the probability that a coupon of any type has not been drawn after \( n \cdot \ln n + c \cdot n \) trials? At most \( e^{-c} \).
Tail bounds (first principles)

<table>
<thead>
<tr>
<th>First principles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus on a coupon of type $i$. What is the probability that a coupon of type $i$ has not been drawn after $n \cdot \ln n + c \cdot n$ trials?</td>
</tr>
</tbody>
</table>

\[
(1 - \frac{1}{n})^{n \cdot \ln n + c \cdot n} = (1 - \frac{1}{n})^{n \cdot (\ln n + c)} \\
\leq e^{-1 \cdot (\ln n + c)} \\
= \frac{1}{e^c \cdot n}
\]

What is the probability that a coupon of any type has not been drawn after $n \cdot \ln n + c \cdot n$ trials? At most $e^{-c}$. Hence, the probability that a coupon of some type is not picked after $2 \cdot n \cdot \ln n$ trials is at most
Focus on a coupon of type $i$. What is the probability that a coupon of type $i$ has not been drawn after $n \cdot \ln n + c \cdot n$ trials?

$$
(1 - \frac{1}{n})^{n \cdot \ln n + c \cdot n} = (1 - \frac{1}{n})^{n \cdot (\ln n + c)}
\leq e^{-1 \cdot (\ln n + c)} = \frac{1}{e^c \cdot n}
$$

What is the probability that a coupon of any type has not been drawn after $n \cdot \ln n + c \cdot n$ trials? At most $e^{-c}$. Hence, the probability that a coupon of some type is not picked after $2 \cdot n \cdot \ln n$ trials is at most $e^{-\ln n}$.
First principles

Focus on a coupon of type $i$. What is the probability that a coupon of type $i$ has not been drawn after $n \cdot \ln n + c \cdot n$ trials?

\[
(1 - \frac{1}{n})^{n \cdot \ln n + c \cdot n} = (1 - \frac{1}{n})^{n \cdot (\ln n + c)} \\
\leq e^{-1 \cdot (\ln n + c)} \\
= \frac{1}{e^c \cdot n}
\]

What is the probability that a coupon of any type has not been drawn after $n \cdot \ln n + c \cdot n$ trials? At most $e^{-c}$. Hence, the probability that a coupon of some type is not picked after $2 \cdot n \cdot \ln n$ trials is at most $e^{-\ln n} = \frac{1}{n}$.
Focus on a coupon of type $i$. What is the probability that a coupon of type $i$ has not been drawn after $n \cdot \ln n + c \cdot n$ trials?

$$(1 - \frac{1}{n})^{n \cdot \ln n + c \cdot n} = (1 - \frac{1}{n})^{n \cdot (\ln n + c)} \leq e^{-1 \cdot (\ln n + c)} = \frac{1}{e^{c} \cdot n}$$

What is the probability that a coupon of any type has not been drawn after $n \cdot \ln n + c \cdot n$ trials? At most $e^{-c}$. Hence, the probability that a coupon of some type is not picked after $2 \cdot n \cdot \ln n$ trials is at most $e^{-\ln n} = \frac{1}{n}$. Moral of the story:
First principles

Focus on a coupon of type $i$. What is the probability that a coupon of type $i$ has not been drawn after $n \cdot \ln n + c \cdot n$ trials?

\[
(1 - \frac{1}{n})^{n \cdot \ln n + c \cdot n} = (1 - \frac{1}{n})^{n \cdot (\ln n + c)} \leq e^{-1 \cdot (\ln n + c)} = \frac{1}{e^c \cdot n}
\]

What is the probability that a coupon of any type has not been drawn after $n \cdot \ln n + c \cdot n$ trials? At most $e^{-c}$. Hence, the probability that a coupon of some type is not picked after $2 \cdot n \cdot \ln n$ trials is at most $e^{-\ln n} = \frac{1}{n}$. Moral of the story: First principle bounds are always better than cookie-cutter bounds.