Moments and Deviations

K. Subramani[1]

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

9 February, 2012

Outline

Outline

Outline

Outline

Recap

Main points

Probability spaces, Random Variable,

Recap

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,

Recap

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

Tail bounds

Note

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

Tail bounds

Note

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

Tail bounds

Note

The tail bounds of a random variable X are concerned with the probability that it deviates signficantly 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 signficantly 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

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.

$$\begin{aligned} E[X] &= \sum_x x \cdot P(X = x) \\ &= \sum_{0 \leq x < c} x \cdot P(X = x) + \sum_{x \geq c} x \cdot P(X = x) \end{aligned}$$

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.

$$\begin{aligned} E[X] &= \sum_{x} x \cdot P(X = x) \\ &= \sum_{0 \leq x < c} x \cdot P(X = x) + \sum_{x \geq c} x \cdot P(X = x) \\ &\geq \sum_{x \geq c} x \cdot P(X = x) \end{aligned}$$

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.

$$\begin{aligned} E[X] &= \sum_{x} x \cdot P(X = x) \\ &= \sum_{0 \leq x < c} x \cdot P(X = x) + \sum_{x \geq c} x \cdot P(X = x) \\ &\geq \sum_{x \geq c} x \cdot P(X = x) \\ &\geq \sum_{x \geq c} c \cdot P(X = x) \end{aligned}$$

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.

$$
\begin{aligned}
E[X] &= \sum_{x} x \cdot P(X = x) \\
&= \sum_{0 \leq x < c} x \cdot P(X = x) + \sum_{x \geq c} x \cdot P(X = x) \\
&\geq \sum_{x \geq c} x \cdot P(X = x) \\
&\geq \sum_{x \geq c} c \cdot P(X = x) \\
&= c \cdot P(X \geq c)
\end{aligned}
$$

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.

$$\begin{aligned} E[X] &= \sum_x x \cdot P(X = x) \\ &= \sum_{0 \leq x < c} x \cdot P(X = x) + \sum_{x \geq c} x \cdot P(X = x) \\ &\geq \sum_{x \geq c} x \cdot P(X = x) \\ &\geq \sum_{x \geq c} c \cdot P(X = x) \\ &= c \cdot P(X \geq c) \\ \Rightarrow P(X \geq c) &\leq \frac{E[X]}{c} \end{aligned}$$

□

Markov's Inequality (contd.)

Markov's Inequality (contd.)

Alternative Form

Markov's Inequality (contd.)

Alternative Form

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

Markov's Inequality (contd.)

Alternative Form

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

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)

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

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)

$$\begin{aligned} P(X \geq \frac{3n}{4}) &= P(X \geq \frac{3}{2} \cdot \frac{n}{2}) \\ &\leq \frac{1}{\frac{3}{2}} \end{aligned}$$

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)

$$\begin{aligned} P(X \geq \frac{3n}{4}) &= P(X \geq \frac{3}{2} \cdot \frac{n}{2}) \\ &\leq \frac{1}{\frac{3}{2}} \\ &= \frac{2}{3} \end{aligned}$$

Chebyshev's Inequality

Chebyshev's Inequality

Theorem

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

Chebyshev's Inequality

Theorem

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

Chebyshev's Inequality

Theorem

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

Proof.

$$P(|X - E[X]| \geq a) \quad =$$

Chebyshev's Inequality

Theorem

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

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, $P(|X - E[X]| \geq a) \leq \frac{\mathbf{Var}[X]}{a^2}$.

Proof.

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

Chebyshev's Inequality

Theorem

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

Proof.

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

□

Chebyshev's Inequality

Theorem

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

Proof.

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

□

Note

Chebyshev's theorem is alternatively stated as:

Chebyshev's Inequality

Theorem

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

Proof.

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

□

Note

Chebyshev's theorem is alternatively stated as:
$P(|X - E[X]| \geq a \cdot E[X]) \leq \frac{\textbf{Var}[X]}{(a \cdot E[X])^2}$.

Chebyshev's inequality (contd.)

Example (Application to coin tossing problem)

Chebyshev's inequality (contd.)

Example (Application to coin tossing problem)

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

Chebyshev's inequality (contd.)

Example (Application to coin tossing problem)

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

Chebyshev's inequality (contd.)

Example (Application to coin tossing problem)

$$\begin{aligned} P(X \geq \frac{3n}{4}) &= P(X - \frac{n}{2} \geq \frac{n}{4}) \\ &\leq P(|X - \frac{n}{2}| \geq \frac{n}{4}) \end{aligned}$$

Chebyshev's inequality (contd.)

Example (Application to coin tossing problem)

$$
\begin{aligned}
P(X \geq \frac{3n}{4}) &= P(X - \frac{n}{2} \geq \frac{n}{4}) \\
&\leq P(|X - \frac{n}{2}| \geq \frac{n}{4}) \\
&= P(|X - E[X]| \geq \frac{1}{2}E[X])
\end{aligned}
$$

Chebyshev's inequality (contd.)

Example (Application to coin tossing problem)

$$\begin{aligned} P(X \geq \frac{3n}{4}) &= P(X - \frac{n}{2} \geq \frac{n}{4}) \\ &\leq P(|X - \frac{n}{2}| \geq \frac{n}{4}) \\ &= P(|X - E[X]| \geq \frac{1}{2}E[X]) \\ &\leq \frac{\frac{n}{4}}{(\frac{1}{2})^2 \cdot (\frac{n}{2})^2} \end{aligned}$$

Chebyshev's inequality (contd.)

Example (Application to coin tossing problem)

$$\begin{aligned}
P(X \geq \frac{3n}{4}) &= P(X - \frac{n}{2} \geq \frac{n}{4}) \\
&\leq P(|X - \frac{n}{2}| \geq \frac{n}{4}) \\
&= P(|X - E[X]| \geq \frac{1}{2}E[X]) \\
&\leq \frac{\frac{n}{4}}{(\frac{1}{2})^2 \cdot (\frac{n}{2})^2} \\
&= \frac{4}{n}
\end{aligned}$$

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

Markov

$$P(X \geq 2 \cdot n \cdot H_n)$$

Tail bounds for coupon collecting

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?

Tail bounds for coupon collecting

Markov

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

Chebyshev

What do we need? $\mathbf{Var}[X]$.

Tail bounds for coupon collecting

Markov

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

Chebyshev

What do we need? $\mathbf{Var}[X]$. Observe that $\mathbf{Var}[X] = \sum_{i=1}^{n} \mathbf{Var}[X_i]$, 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? $\mathbf{Var}[X]$. Observe that $\mathbf{Var}[X] = \sum_{i=1}^{n} \mathbf{Var}[X_i]$, 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 $\mathbf{Var}[X_i] = \frac{1-p_i}{p_i^2}$

Tail bounds for coupon collecting

Markov

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

Chebyshev

What do we need? $\mathbf{Var}[X]$. Observe that $\mathbf{Var}[X] = \sum_{i=1}^{n} \mathbf{Var}[X_i]$, 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 $\mathbf{Var}[X_i] = \frac{1-p_i}{p_i^2} \leq \frac{1}{p_i^2}$.

Tail bounds for coupon collecting

Markov

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

Chebyshev

What do we need? $\mathbf{Var}[X]$. Observe that $\mathbf{Var}[X] = \sum_{i=1}^{n} \mathbf{Var}[X_i]$, 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 $\mathbf{Var}[X_i] = \frac{1-p_i}{p_i^2} \leq \frac{1}{p_i^2}$.
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? $\textbf{Var}[X]$. Observe that $\textbf{Var}[X] = \sum_{i=1}^{n} \textbf{Var}[X_i]$, 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 $\textbf{Var}[X_i] = \frac{1-p_i}{p_i^2} \leq \frac{1}{p_i^2}$.

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? $\mathbf{Var}[X]$. Observe that $\mathbf{Var}[X] = \sum_{i=1}^{n} \mathbf{Var}[X_i]$, 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 $\mathbf{Var}[X_i] = \frac{1-p_i}{p_i^2} \leq \frac{1}{p_i^2}$.
But recall that $p_i = \frac{n-i+1}{n}$. Therefore, $\frac{1}{p_i} = \frac{n}{n-i+1}$. Hence,

$$\mathbf{Var}[X] \quad = \quad \sum_{i=1}^{n} \mathbf{Var}[X_i]$$

Tail bounds for coupon collecting

Markov

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

Chebyshev

What do we need? $\mathbf{Var}[X]$. Observe that $\mathbf{Var}[X] = \sum_{i=1}^{n} \mathbf{Var}[X_i]$, 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 $\mathbf{Var}[X_i] = \frac{1-p_i}{p_i^2} \leq \frac{1}{p_i^2}$.
But recall that $p_i = \frac{n-i+1}{n}$. Therefore, $\frac{1}{p_i} = \frac{n}{n-i+1}$. Hence,

$$\begin{aligned}\mathbf{Var}[X] &= \sum_{i=1}^{n} \mathbf{Var}[X_i] \\ &\leq \sum_{i=1}^{n} \frac{1}{{p_i}^2}\end{aligned}$$

Tail bounds for coupon collecting (contd.)

Chebyshev (contd.)

Tail bounds for coupon collecting (contd.)

Chebyshev (contd.)

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

Tail bounds for coupon collecting (contd.)

Chebyshev (contd.)

$$\begin{aligned}\textbf{Var}[X] &\leq \sum_{i=1}^{n}\left(\frac{n}{n-i+1}\right)^2 \\ &= n^2 \cdot \sum_{i=1}^{n}\left(\frac{1}{n-i+1}\right)^2\end{aligned}$$

Tail bounds for coupon collecting (contd.)

Chebyshev (contd.)

$$\begin{aligned}\textbf{Var}[X] &\leq \sum_{i=1}^{n}\left(\frac{n}{n-i+1}\right)^2 \\ &= n^2 \cdot \sum_{i=1}^{n}\left(\frac{1}{n-i+1}\right)^2 \\ &= n^2 \cdot \sum_{i=1}^{n}\frac{1}{i^2}\end{aligned}$$

Tail bounds for coupon collecting (contd.)

Chebyshev (contd.)

$$
\begin{aligned}
\mathbf{Var}[X] &\leq \sum_{i=1}^{n} \left(\frac{n}{n-i+1}\right)^2 \\
&= n^2 \cdot \sum_{i=1}^{n} \left(\frac{1}{n-i+1}\right)^2 \\
&= n^2 \cdot \sum_{i=1}^{n} \frac{1}{i^2} \\
&\leq n^2 \cdot \frac{\pi^2}{6}
\end{aligned}
$$

Tail bounds for coupon collecting (contd.)

Analysis (contd.)

Tail bounds for coupon collecting (contd.)

Analysis (contd.)

It follows that

$$P(X \geq 2 \cdot n \cdot H_n) \quad =$$

Tail bounds for coupon collecting (contd.)

Analysis (contd.)

It follows that

$$P(X \geq 2 \cdot n \cdot H_n) = P((X - n \cdot H_n) \geq n \cdot H_n)$$

Tail bounds for coupon collecting (contd.)

Analysis (contd.)

It follows that

$$\begin{aligned} P(X \geq 2 \cdot n \cdot H_n) &= P((X - n \cdot H_n) \geq n \cdot H_n) \\ &\leq P(|X - n \cdot H_n| \geq n \cdot H_n) \end{aligned}$$

Tail bounds for coupon collecting (contd.)

Analysis (contd.)

It follows that

$$\begin{aligned} P(X \geq 2 \cdot n \cdot H_n) &= P((X - n \cdot H_n) \geq n \cdot H_n) \\ &\leq P(|X - n \cdot H_n| \geq n \cdot H_n) \\ &\leq \frac{n^2 \cdot \frac{\pi^2}{6}}{(n \cdot H_n)^2} \end{aligned}$$

Tail bounds for coupon collecting (contd.)

Analysis (contd.)

It follows that

$$\begin{aligned} P(X \geq 2 \cdot n \cdot H_n) &= P((X - n \cdot H_n) \geq n \cdot H_n) \\ &\leq P(|X - n \cdot H_n| \geq n \cdot H_n) \\ &\leq \frac{n^2 \cdot \frac{\pi^2}{6}}{(n \cdot H_n)^2} \\ &\in O(\frac{1}{\ln^2 n}) \end{aligned}$$

Tail bounds (first principles)

Tail bounds (first principles)

First principles

Focus on a coupon of type i.

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?

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?

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

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?

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

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?

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

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?

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

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?

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

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?

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

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)

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?

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

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

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?

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

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 $e^{-\ln n}$

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?

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

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 $e^{-\ln n} = \frac{1}{n}$.

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?

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

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:

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?

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

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.