Randomness and Computation
or, “Randomized Algorithms”

Mary Cryan

School of Informatics
University of Edinburgh
Coupon Collector Problem

“Coupon collecting" is the activity of buying cereal-packets, each of which will have a coupon inside. There are be $n$ different types of “coupon" (eg cards with a photo of a footballer) and the goal is to collect one copy of each … then stop buying.

On Tuesday we showed that the expected number of purchases needed $E[X]$ to collect all cards is $\sim n \ln(n)$.

Today we examine how likely a example “run" of the purchasing process is to come close to that expectation.

Results like Markov’s Inequality, Chebyshev’s Inequality and (Friday) Chernoff/Hoeffding Bounds help us show concentration about the mean.
Markov’s Inequality

The simplest one.

Theorem (3.1, Markov’s Inequality)

Let $X$ be any random variable that takes only non-negative values. Then for any $a > 0$,

$$\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.$$ 

Proof.

Define the indicator function $I = I(X)$ by

$$I(x) = \begin{cases} 
0 & x < a \\
1 & x \geq a 
\end{cases}$$

Then $X \geq a \cdot I(X)$, and hence $I(X) \leq \frac{X}{a}$.

Taking expectation of both sides, and using $\mathbb{E}[I] = \Pr[X \geq a]$, we have

$$\Pr[X \geq a] = \mathbb{E}[I] \leq \frac{1}{a} \mathbb{E}[X].$$
Markov’s Inequality
The simplest one.

**Theorem (3.1, Markov’s Inequality)**

*Let X be any random variable that takes only non-negative values. Then for any a > 0,*

\[
\Pr[X \geq a] \leq \frac{E[X]}{a}.
\]

**Proof.**

Define the indicator function \( I = I(X) \) by

\[
I(x) = \begin{cases} 
0 & x < a \\
1 & x \geq a 
\end{cases}
\]

Then \( X \geq a \cdot I(X) \), and hence \( I(X) \leq \frac{X}{a} \).

Taking expectation of both sides, and using \( E[I] = \Pr[X \geq a] \), we have

\[
\Pr[X \geq a] = E[I] \leq \frac{1}{a}E[X].
\]
Markov’s Inequality

The simplest one.

**Theorem (3.1, Markov’s Inequality)**

Let $X$ be any random variable that takes only non-negative values. Then for any $a > 0$,

$$\Pr[X \geq a] \leq \frac{E[X]}{a}.$$ 

**Proof.**

Define the indicator function $I = I(X)$ by

$$I(x) = \begin{cases} 
0 & x < a \\
1 & x \geq a 
\end{cases}$$

Then $X \geq a \cdot I(X)$, and hence $I(X) \leq \frac{X}{a}$.

Taking expectation of both sides, and using $E[I] = \Pr[X \geq a]$, we have

$$\Pr[X \geq a] = E[I] \leq \frac{1}{a}E[X].$$
Markov’s Inequality

The simplest one.

Theorem (3.1, Markov’s Inequality)

Let $X$ be any random variable that takes only non-negative values. Then for any $a > 0$,

$$
\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.
$$

Proof.

Define the indicator function $I = I(X)$ by

$$
I(x) = \begin{cases} 
0 & x < a \\
1 & x \geq a
\end{cases}
$$

Then $X \geq a \cdot I(X)$, and hence $I(X) \leq \frac{X}{a}$.

Taking expectation of both sides, and using $\mathbb{E}[I] = \Pr[X \geq a]$, we have

$$
\Pr[X \geq a] = \mathbb{E}[I] \leq \frac{1}{a} \mathbb{E}[X].
$$
Variance, Moments of a Random Variable

Definition (3.1)
The \textit{kth moment} of a random variable \(X\) is defined to be \(E[X^k]\).

Definition (3.2)
The \textit{variance} of a random variable is defined to be

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

The \textit{standard deviation} of a random variable \(X\) is defined as

\[
\sigma[X] = \sqrt{\text{Var}[X]}.
\]

(we saw why \(E[(X - E[X])^2]\) and \(E[X^2] - E[X]^2\) were equal on Tuesday)
Variance, Moments of a Random Variable

Definition (3.1)
The $k$th moment of a random variable $X$ is defined to be $E[X^k]$.

Definition (3.2)
The variance of a random variable is defined to be

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

The standard deviation of a random variable $X$ is defined as

$$\sigma[X] = \sqrt{\text{Var}[X]}.$$

(we saw why $E[(X - E[X])^2]$ and $E[X^2] - E[X]^2$ were equal on Tuesday)
Covariance of two Random Variables

Definition (3.3)
The covariance of two random variables $X$ and $Y$ is defined as

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

Theorem (3.2)
For any two random variables $X$, $Y$, we have

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y].$$

Proof.
The definition of $\text{Var}$ gives $\text{Var}[X + Y] = \mathbb{E}[(X + Y)^2] - \mathbb{E}[X + Y]^2$.
By squaring, and linearity of $\mathbb{E}$, this is

$$\mathbb{E}[X^2] + \mathbb{E}[Y^2] + \mathbb{E}[2XY] - (\mathbb{E}[X]^2 + \mathbb{E}[Y]^2 + 2\mathbb{E}[X]\mathbb{E}[Y]).$$
This is $\text{Var}[X] + \text{Var}[Y] + 2\mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y]$.
Expanding $\text{Cov}[X, Y]$, linearity of $\mathbb{E}$, gives $2\mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y]$, so
$\text{Var}[X + Y]$ equals $\text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y]$. 

RC (2017/18) – Lecture 6 – slide 5
Covariance of two Random Variables

Definition (3.3)
The covariance of two random variables $X$ and $Y$ is defined as

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

Theorem (3.2)
For any two random variables $X, Y$, we have

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y].$$

Proof.
The definition of $\text{Var}$ gives


By squaring, and linearity of exp., this is


This is

$$\text{Var}[X] + \text{Var}[Y] + 2\text{E}[XY] - 2\text{E}[X]\text{E}[Y].$$

Expanding $\text{Cov}[X, Y]$, linearity of Exp., gives $2\text{E}[XY] - 2\text{E}[X]\text{E}[Y]$, so

$$\text{Var}[X + Y]$$

equals

$$\text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y].$$

RC (2017/18) – Lecture 6 – slide 5
Covariance of two Random Variables

Definition (3.3)
The covariance of two random variables $X$ and $Y$ is defined as

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

Theorem (3.2)
For any two random variables $X$, $Y$, we have

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y].$$

Proof.
The definition of $\text{Var}$ gives \( \text{Var}[X + Y] = E[(X + Y)^2] - E[X + Y]^2 \).
By squaring, and linearity of exp., this is
\[
\]
This is $\text{Var}[X] + \text{Var}[Y] + 2E[XY] - 2E[X]E[Y].$
Expanding $\text{Cov}[X, Y]$, linearity of Exp., gives $2E[XY] - 2E[X]E[Y],$ so
$\text{Var}[X + Y]$ equals $\text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y].$

RC (2017/18) – Lecture 6 – slide 5
Covariance of two Random Variables

Definition (3.3)
The covariance of two random variables $X$ and $Y$ is defined as

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

Theorem (3.2)

For any two random variables $X$, $Y$, we have

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y].$$

Proof.
The definition of $\text{Var}$ gives $\text{Var}[X + Y] = E[(X + Y)^2] - E[X + Y]^2$. By squaring, and linearity of exp., this is


This is $\text{Var}[X] + \text{Var}[Y] + 2E[XY] - 2E[X]E[Y]$.


$\text{Var}[X + Y]$ equals $\text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y]$. 

RC (2017/18) – Lecture 6 – slide 5
Covariance of two Random Variables

**Definition (3.3)**
The covariance of two random variables $X$ and $Y$ is defined as

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

**Theorem (3.2)**
For any two random variables $X, Y$, we have

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y].$$

**Proof.**
The definition of $\text{Var}$ gives $\text{Var}[X + Y] = E[(X + Y)^2] - E[X + Y]^2$.
By squaring, and linearity of exp., this is

This is $\text{Var}[X] + \text{Var}[Y] + 2E[XY] - 2E[X]E[Y]$.

$$\text{Var}[X + Y]$$
equals $\text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y].$
Covariance of two Random Variables

Definition (3.3)
The covariance of two random variables $X$ and $Y$ is defined as
\[
\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].
\]

Theorem (3.2)
For any two random variables $X$, $Y$, we have
\[
\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y].
\]

Proof.
The definition of Var gives $\text{Var}[X + Y] = \mathbb{E}[(X + Y)^2] - \mathbb{E}[X + Y]^2$.
By squaring, and linearity of exp., this is
\[
\mathbb{E}[X^2] + \mathbb{E}[Y^2] + \mathbb{E}[2XY] - (\mathbb{E}[X]^2 + \mathbb{E}[Y]^2 + 2\mathbb{E}[X]\mathbb{E}[Y]).
\]
This is $\text{Var}[X] + \text{Var}[Y] + 2\mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y]$.
Expanding $\text{Cov}[X, Y]$, linearity of Exp., gives $2\mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y]$,
so $\text{Var}[X + Y]$ equals $\text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y]$.
Theorem (3.3)

If \( X, Y \) are a pair of independent random variables, then

\[
E[XY] = E[X] \cdot E[Y].
\]

Proof is in the book, reasoning wrt Definition 2.2 (may do on visualiser).

Corollary (3.4)

If \( X, Y \) are a pair of independent random variables, then

\[
\text{Cov}[X, Y] = 0
\]

and

\[
\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y].
\]

Proof is straightforward application of Thm 3.3.
(pairwise) Independent Random Variables

Theorem (3.3)
If $X, Y$ are a pair of independent random variables, then

$$E[XY] = E[X] \cdot E[Y].$$

Proof is in the book, reasoning wrt Definition 2.2 (may do on visualiser).

Corollary (3.4)
If $X, Y$ are a pair of independent random variables, then

$$\text{Cov}[X, Y] = 0$$

and

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y].$$

Proof is straightforward application of Thm 3.3.
Chebyshev’s Inequality

Theorem (3.2, Chebyshev’s Inequality)

For every $a > 0$,

$$\Pr[|X - \mathbb{E}[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}.$$ 

Proof.
Because the probability is of the absolute value of $X - \mathbb{E}[X]$, we know that for any $b > 0$, $|X - \mathbb{E}[X]| = b$ happens $\Leftrightarrow (X - \mathbb{E}[X])^2 = b^2$ happens.

So $\Pr[|X - \mathbb{E}[X]| \geq a] = \Pr[(X - \mathbb{E}[X])^2 \geq a^2]$.

Applying Markov’s Ineq. to the random variable $(X - \mathbb{E}[X])^2$, we know

$$\Pr[(X - \mathbb{E}[X])^2 \geq a^2] \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{a^2}.$$ 

and by definition of $\text{Var}(\cdot)$, this gives

$$\Pr[|X - \mathbb{E}[X]| \geq a] = \Pr[(X - \mathbb{E}[X])^2 \geq a^2] \leq \frac{\text{Var}[X]}{a^2}.$$ 

RC (2017/18) – Lecture 6 – slide 7
Chebyshev’s Inequality

**Theorem (3.2, Chebyshev’s Inequality)**

For every \( a > 0 \),

\[
\Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}.
\]

**Proof.**

Because the probability is of the *absolute value* of \( X - E[X] \), we know that for any \( b > 0 \), \(|X - E[X]| = b\) happens if and only if \((X - E[X])^2 = b^2\) happens.

So \( \Pr[|X - E[X]| \geq a] = \Pr[(X - E[X])^2 \geq a^2] \).

Applying Markov’s Ineq. to the random variable \((X - E[X])^2\), we know

\[
\Pr[(X - E[X])^2 \geq a^2] \leq \frac{\text{E}[(X - E[X])^2]}{a^2}.
\]

and by definition of \( \text{Var}(\cdot) \), this gives

\[
\Pr[|X - E[X]| \geq a] = \Pr[(X - E[X])^2 \geq a^2] \leq \frac{\text{Var}[X]}{a^2}.
\]
Chebyshev’s Inequality

Theorem (3.2, Chebyshev’s Inequality)

For every \( a > 0 \),

\[
\Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}.
\]

Proof.
Because the probability is of the \textit{absolute value} of \( X - E[X] \), we know that for any \( b > 0 \), \( |X - E[X]| = b \) happens \( \Leftrightarrow (X - E[X])^2 = b^2 \) happens.
So \( \Pr[|X - E[X]| \geq a] = \Pr[(X - E[X])^2 \geq a^2] \).

Applying Markov’s Ineq. to the random variable \((X - E[X])^2\), we know

\[
\Pr[(X - E[X])^2 \geq a^2] \leq \frac{E[(X - E[X])^2]}{a^2}.
\]

and by definition of \text{Var}(\cdot), this gives

\[
\Pr[|X - E[X]| \geq a] = \Pr[(X - E[X])^2 \geq a^2] \leq \frac{\text{Var}[X]}{a^2}.
\]

RC (2017/18) – Lecture 6 – slide 7
Chebyshev’s Inequality

Theorem (3.2, Chebyshev’s Inequality)

For every $a > 0$,

$$\Pr[|X - \mathbb{E}[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}.$$ 

Proof.

Because the probability is of the absolute value of $X - \mathbb{E}[X]$, we know that for any $b > 0$, $|X - \mathbb{E}[X]| = b$ happens $\iff (X - \mathbb{E}[X])^2 = b^2$ happens.

So $\Pr[|X - \mathbb{E}[X]| \geq a] = \Pr[(X - \mathbb{E}[X])^2 \geq a^2]$.

Applying Markov’s Ineq. to the random variable $(X - \mathbb{E}[X])^2$, we know

$$\Pr[(X - \mathbb{E}[X])^2 \geq a^2] \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{a^2}.$$ 

and by definition of $\text{Var}(\cdot)$, this gives

$$\Pr[|X - \mathbb{E}[X]| \geq a] = \Pr[(X - \mathbb{E}[X])^2 \geq a^2] \leq \frac{\text{Var}[X]}{a^2}.$$
Chebyshev’s Inequality

Theorem (3.2, Chebyshev’s Inequality)

For every $a > 0$,

$$\Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}.$$ 

Proof.

Because the probability is of the absolute value of $X - E[X]$, we know that for any $b > 0$, $|X - E[X]| = b$ happens $\iff (X - E[X])^2 = b^2$ happens.

So $\Pr[|X - E[X]| \geq a] = \Pr[(X - E[X])^2 \geq a^2]$.

Applying Markov’s Ineq. to the random variable $(X - E[X])^2$, we know

$$\Pr[(X - E[X])^2 \geq a^2] \leq \frac{E[(X - E[X])^2]}{a^2}.$$ 

and by definition of Var($\cdot$), this gives

$$\Pr[|X - E[X]| \geq a] = \Pr[(X - E[X])^2 \geq a^2] \leq \frac{\text{Var}[X]}{a^2}.$$
Chebyshev’s Inequality

Theorem (3.2, Chebyshev’s Inequality)

For every $a > 0$,

$$\Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}.$$ 

Proof.
Because the probability is of the absolute value of $X - E[X]$, we know that for any $b > 0$, $|X - E[X]| = b$ happens $\Leftrightarrow (X - E[X])^2 = b^2$ happens.
So $\Pr[|X - E[X]| \geq a] = \Pr[(X - E[X])^2 \geq a^2]$.

Applying Markov’s Ineq. to the random variable $(X - E[X])^2$, we know

$$\Pr[(X - E[X])^2 \geq a^2] \leq \frac{\text{E}[(X - E[X])^2]}{a^2}.$$ 

and by definition of $\text{Var}(\cdot)$, this gives

$$\Pr[|X - E[X]| \geq a] = \Pr[(X - E[X])^2 \geq a^2] \leq \frac{\text{Var}[X]}{a^2}.$$ 

RC (2017/18) – Lecture 6 – slide 7
Bounding Coupon Collector purchases - Markov

Remember $X$ are the number of packets bought until we have all $n$ different cards, $E[X] = n \ln(n) + \Theta(n)$ is the expected number.

Consider how likely we are to need twice the expected number of purchases ($2E[X]$). By Markov’s Ineq.,

$$\Pr[X \geq 2E[X]] \leq \frac{E[X]}{2E[X]} = \frac{1}{2}.$$

Or, if we are willing to spend $10E[X]$ (ie, $10n(\ln(n) + 1)$), there is at most $1/10$ probability we fail to get all cards.

Very boring! (Markov’s Ineq)

We can do much better with Chebyshev’s Ineq. . . .

RC (2017/18) – Lecture 6 – slide 8
Bounding Coupon Collector purchases - Markov

Remember $X$ are the number of packets bought until we have all $n$ different cards, $\mathbb{E}[X] = n \ln(n) + \Theta(n)$ is the expected number.

Consider how likely we are to need twice the expected number of purchases ($2\mathbb{E}[X]$). By Markov’s Ineq.,

$$\Pr[X \geq 2\mathbb{E}[X]] \leq \frac{\mathbb{E}[X]}{2\mathbb{E}[X]} = \frac{1}{2}.$$  

Or, if we are willing to spend $10\mathbb{E}[X]$ (ie, $10n(\ln(n) + 1)$), there is at most $1/10$ probability we fail to get all cards.

Very boring! (Markov’s Ineq)
We can do much better with Chebyshev’s Ineq. …
Bounding Coupon Collector purchases - Markov

Remember $X$ are the number of packets bought until we have all $n$ different cards, $E[X] = n \ln(n) + \Theta(n)$ is the expected number.

Consider how likely we are to need twice the expected number of purchases ($2E[X]$). By Markov's Ineq.,

$$\Pr[X \geq 2E[X]] \leq \frac{E[X]}{2E[X]} = \frac{1}{2}.$$

Or, if we are willing to spend $10E[X]$ (ie, $10n(\ln(n) + 1)$), there is at most $1/10$ probability we fail to get all cards.

Very boring! (Markov’s Ineq)

We can do much better with Chebyshev’s Ineq. ...
Bounding Coupon Collector purchases - Markov

Remember $X$ are the number of packets bought until we have all $n$ different cards, $E[X] = n \ln(n) + \Theta(n)$ is the expected number.

Consider how likely we are to need twice the expected number of purchases ($2E[X]$). By Markov’s Ineq.,

$$\Pr[X \geq 2E[X]] \leq \frac{E[X]}{2E[X]} = \frac{1}{2}.$$  

Or, if we are willing to spend $10E[X]$ (ie, $10n(\ln(n) + 1)$), there is at most $1/10$ probability we fail to get all cards.

Very boring! (Markov’s Ineq)
We can do much better with Chebyshev’s Ineq. . .
Bounding Coupon Collector purchases - Markov

Remember $X$ are the number of packets bought until we have all $n$ different cards, $E[X] = n \ln(n) + \Theta(n)$ is the expected number.

Consider how likely we are to need *twice* the expected number of purchases ($2E[X]$). By Markov’s Ineq.,

$$\Pr[X \geq 2E[X]] \leq \frac{E[X]}{2E[X]} = \frac{1}{2}.$$

Or, if we are willing to spend $10E[X]$ (ie, $10n(\ln(n) + 1)$), there is at most $1/10$ probability we fail to get all cards.

Very boring! (Markov’s Ineq)

We can do much better with Chebyshev’s Ineq. ...
Bounding Coupon Collector purchases - Markov

Remember $X$ are the number of packets bought until we have all $n$ different cards, $E[X] = n \ln(n) + \Theta(n)$ is the expected number.

Consider how likely we are to need twice the expected number of purchases ($2E[X]$). By Markov’s Ineq.,

$$
\Pr[X \geq 2E[X]] \leq \frac{E[X]}{2E[X]} = \frac{1}{2}.
$$

Or, if we are willing to spend $10E[X]$ (ie, $10n(\ln(n) + 1)$), there is at most $1/10$ probability we fail to get all cards.

Very boring! (Markov’s Ineq)

We can do much better with Chebyshev’s Ineq. . . .
Bounding Coupon Collector purchases - Chebyshev

\[ \Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}. \]

- Need to evaluate \( \text{Var}[X] \), which is \( \text{Var}[X_1 + \ldots X_n] \).
- Looking back at Corollary 3.4, see that for independent \( Y, Z \),
  \( \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \).
- Recall that \( X_i \), the number of packets bought to get the \( i \)-th new card, is independent of the value of \( X_{i-1} \) or any of the earlier \( X_n \) values. \( X_i \) only depends on the values \( n \) and \( i \).
- Hence the random variables \( X_1, \ldots, X_n \) are all mutually independent.
- So
  \[ \text{Var}[X] = \text{Var}[X_1] + \text{Var}[X_2] + \ldots + \text{Var}[X_n]. \]
Bounding Coupon Collector purchases - Chebyshev

\[ \Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}. \]

▶ Need to evaluate \( \text{Var}[X] \), which is \( \text{Var}[X_1 + \ldots + X_n] \).

▶ Looking back at Corollary 3.4, see that for independent \( Y, Z \),
\[ \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]. \]

▶ Recall that \( X_i \), the number of packets bought to get the \( i \)-th new card, is independent of the value of \( X_{i-1} \) or any of the earlier \( X_h \) values. \( X_i \) only depends on the values \( n \) and \( i \).

▶ Hence the random variables \( X_1, \ldots, X_n \) are all mutually independent.

▶ So
\[ \text{Var}[X] = \text{Var}[X_1] + \text{Var}[X_2] + \ldots + \text{Var}[X_n]. \]
Bounding Coupon Collector purchases - Chebyshev

\[ \Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}. \]

- Need to evaluate \( \text{Var}[X] \), which is \( \text{Var}[X_1 + \ldots X_n] \).
- Looking back at Corollary 3.4, see that for independent \( Y, Z \),
  \( \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \).
- Recall that \( X_i \), the number of packets bought to get the \( i \)-th new card, is independent of the value of \( X_{i-1} \) or any of the earlier \( X_n \) values. \( X_i \) only depends on the values \( n \) and \( i \).
- Hence the random variables \( X_1, \ldots, X_n \) are all mutually independent.
- So
  \[ \text{Var}[X] = \text{Var}[X_1] + \text{Var}[X_2] + \ldots + \text{Var}[X_n]. \]
Bounding Coupon Collector purchases - Chebyshev

\[ \Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}. \]

▶ Need to evaluate \( \text{Var}[X] \), which is \( \text{Var}[X_1 + \ldots + X_n] \).

▶ Looking back at Corollary 3.4, see that for independent \( Y, Z \),
  \( \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \).

▶ Recall that \( X_i \), the number of packets bought to get the \( i \)-th new card, is independent of the value of \( X_{i-1} \) or any of the earlier \( X_h \) values. \( X_i \) only depends on the values \( n \) and \( i \).

▶ Hence the random variables \( X_1, \ldots, X_n \) are all mutually independent.

▶ So

\[ \text{Var}[X] = \text{Var}[X_1] + \text{Var}[X_2] + \ldots + \text{Var}[X_n]. \]
Bounding Coupon Collector purchases - Chebyshev

\[ \Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}. \]

- Need to evaluate \( \text{Var}[X] \), which is \( \text{Var}[X_1 + \ldots + X_n] \).
- Looking back at Corollary 3.4, see that for independent \( Y, Z \), \( \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \).
- Recall that \( X_i \), the *number of packets* bought to get the \( i \)-th new card, is independent of the value of \( X_{i-1} \) or any of the earlier \( X_h \) values. \( X_i \) only depends on the values \( n \) and \( i \).
- Hence the random variables \( X_1, \ldots, X_n \) are all mutually independent.
- So

\[ \text{Var}[X] = \text{Var}[X_1] + \text{Var}[X_2] + \ldots + \text{Var}[X_n]. \]
Bounding Coupon Collector purchases - Chebyshev

\[
\Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}.
\]

▶ Need to evaluate \(\text{Var}[X]\), which is \(\text{Var}[X_1 + \ldots + X_n]\).

▶ Looking back at Corollary 3.4, see that for independent \(Y, Z\),
  \(\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]\).

▶ Recall that \(X_i\), the \textit{number of packets} bought to get the \(i\)-th new card, is independent of the value of \(X_{i-1}\) or any of the earlier \(X_h\) values. \(X_i\) only depends on the values \(n\) and \(i\).

▶ Hence the random variables \(X_1, \ldots, X_n\) are all mutually independent.

▶ So

\[
\text{Var}[X] = \text{Var}[X_1] + \text{Var}[X_2] + \ldots + \text{Var}[X_n].
\]

\(RC (2017/18) – Lecture 6 – slide 9\)
Bounding Coupon Collector purchases - Chebyshev

Each $X_i$ is a geometric random variable with parameter $\frac{n-(i-1)}{n}$.

Lemma (3.8)
For any geometric random variable $X$ with parameter $p$, $E[X] = p^{-1}$ and $\text{Var}[X] = \frac{1-p}{p^2}$.

Proof.
We have $\text{Var}[X] = E[X^2] - E[X]^2$. For geometric variable, $E[X]^2 = p^{-2}$. Working with $E[X^2] = \sum_{j=1}^{\infty} j^2 \cdot \Pr[X = j]$, we have

$$
E[X^2] = \sum_{j=1}^{\infty} j^2 \cdot (1-p)^{j-1}p \quad \text{a geom. variable}
$$

$$
= \frac{p}{1-p} \sum_{j=1}^{\infty} j^2 \cdot (1-p)^j
$$

$$
= \frac{p}{1-p} \frac{(1-p)^2 + (1-p)}{p^3} \quad \text{formula for } \sum_{j=1}^{\infty} i^2 x^i.
$$
Bounding Coupon Collector purchases - Chebyshev

Each $X_i$ is a geometric random variable with parameter $\frac{n-(i-1)}{n}$.

**Lemma (3.8)**

*For any geometric random variable $X$ with parameter $p$, $\mathbb{E}[X] = p^{-1}$ and $\text{Var}[X] = \frac{1-p}{p^2}$.***

**Proof.**

We have $\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. For geometric variable, $\mathbb{E}[X]^2 = p^{-2}$.

Working with $\mathbb{E}[X^2] = \sum_{j=1}^{\infty} j^2 \cdot \Pr[X = j]$, we have

$$
\mathbb{E}[X^2] = \sum_{j=1}^{\infty} j^2 \cdot (1 - p)^{j-1} p \quad \text{a geom. variable}
$$

$$
= \frac{p}{1 - p} \sum_{j=1}^{\infty} j^2 \cdot (1 - p)^j
$$

$$
= \frac{p}{1 - p} \left( \frac{(1-p)^2 + (1-p)}{p^3} \right) \quad \text{formula for } \sum_{j=1}^{\infty} i^2 x^i.
$$
Bounding Coupon Collector purchases - Chebyshev

Each $X_i$ is a geometric random variable with parameter $\frac{n-(i-1)}{n}$.

Lemma (3.8)

For any geometric random variable $X$ with parameter $p$, $E[X] = p^{-1}$ and $\text{Var}[X] = \frac{1-p}{p^2}$.

Proof.

We have $\text{Var}[X] = E[X^2] - E[X]^2$. For geometric variable, $E[X]^2 = p^{-2}$.

Working with $E[X^2] = \sum_{j=1}^{\infty} j^2 \cdot \Pr[X = j]$, we have

$$E[X^2] = \sum_{j=1}^{\infty} j^2 \cdot (1 - p)^{j-1} p \quad \text{a geom. variable}$$

$$= \frac{p}{1-p} \sum_{j=1}^{\infty} j^2 \cdot (1 - p)^{j}$$

$$= \frac{p}{1-p} \frac{(1-p)^2 + (1-p)}{p^3} \quad \text{formula for } \sum_{j=1}^{\infty} i^2 x^i.$$
Bounding Coupon Collector purchases - Chebyshev

Each $X_i$ is a geometric random variable with parameter $\frac{n-(i-1)}{n}$.

Lemma (3.8)

For any geometric random variable $X$ with parameter $p$, $E[X] = p^{-1}$ and $\text{Var}[X] = \frac{1-p}{p^2}$.

Proof.

We have $\text{Var}[X] = E[X^2] - E[X]^2$. For geometric variable, $E[X]^2 = p^{-2}$. Working with $E[X^2] = \sum_{j=1}^{\infty} j^2 \cdot \Pr[X = j]$, we have

$$E[X^2] = \sum_{j=1}^{\infty} j^2 \cdot (1-p)^{j-1} p$$

a geom. variable

$$= \frac{p}{1-p} \sum_{j=1}^{\infty} j^2 \cdot (1-p)^j$$

$$= \frac{p}{1-p} \left( \frac{(1-p)^2 + (1-p)}{p^3} \right)$$

formula for $\sum_{j=1}^{\infty} j^2 x^i$. 

RC (2017/18) – Lecture 6 – slide 10
Bounding Coupon Collector purchases - Chebyshev

Each $X_i$ is a geometric random variable with parameter $\frac{n-(i-1)}{n}$.

Lemma (3.8)

For any geometric random variable $X$ with parameter $p$, $E[X] = p^{-1}$ and $\text{Var}[X] = \frac{1-p}{p^2}$.

Proof.

We have $\text{Var}[X] = E[X^2] - E[X]^2$. For geometric variable, $E[X]^2 = p^{-2}$. Working with $E[X^2] = \sum_{j=1}^{\infty} j^2 \cdot \Pr[X = j]$, we have

$$
E[X^2] = \sum_{j=1}^{\infty} j^2 \cdot (1 - p)^{j-1} p \quad \text{a geom. variable}
$$

$$
= \frac{p}{1-p} \sum_{j=1}^{\infty} j^2 \cdot (1 - p)^j
$$

$$
= \frac{p}{1-p} \frac{(1 - p)^2 + (1 - p)}{p^3} \quad \text{formula for } \sum_{j=1}^{\infty} i^2 x^i.
$$
Bounding Coupon Collector purchases - Chebyshev

Each $X_i$ is a geometric random variable with parameter $\frac{n-(i-1)}{n}$.

**Lemma (3.8)**

For any geometric random variable $X$ with parameter $p$, $E[X] = p^{-1}$ and $\text{Var}[X] = \frac{1-p}{p^2}$.

**Proof.**

We have $\text{Var}[X] = E[X^2] - E[X]^2$. For geometric variable, $E[X]^2 = p^{-2}$.

Working with $E[X^2] = \sum_{j=1}^{\infty} j^2 \cdot \text{Pr}[X = j]$, we have

\[
E[X^2] = \sum_{j=1}^{\infty} j^2 \cdot (1 - p)^{j-1} p \quad \text{a geom. variable}
\]

\[
= \frac{p}{1-p} \sum_{j=1}^{\infty} j^2 \cdot (1 - p)^j
\]

\[
= \frac{p}{1-p} \frac{(1-p)^2 + (1-p)}{p^3} \quad \text{formula for } \sum_{j=1}^{\infty} i^2 x^i.
\]
Lemma (3.8)

For any geometric random variable $X$ with parameter $p$, $E[X] = p^{-1}$ and $\text{Var}[X] = \frac{1-p}{p^2}$.

Proof (cont’d).

So $E[X^2] = \frac{p}{1-p} \frac{(1-p)(2-p)}{p^3} = \frac{2-p}{p^2}$, hence

$$E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2}$$

$$= \frac{(2-p) - 1}{p^2}$$

$$= \frac{1-p}{p^2},$$

as claimed.
Lemma (3.8)

For any geometric random variable $X$ with parameter $p$, $E[X] = p^{-1}$ and $\text{Var}[X] = \frac{1-p}{p^2}$.

Proof (cont’d).

So $E[X^2] = \frac{p}{1-p} \frac{(1-p)(2-p)}{p^3} = \frac{2-p}{p^2}$, hence

$$E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{(2-p) - 1}{p^2} = \frac{1-p}{p^2},$$

as claimed.
Lemma (3.8)

For any geometric random variable $X$ with parameter $p$, $E[X] = p^{-1}$ and $\text{Var}[X] = \frac{1-p}{p^2}$.

Proof (cont’d).

So $E[X^2] = \frac{p(1-p)(2-p)}{1-p} = \frac{2-p}{p^2}$, hence

$$E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{(2-p) - 1}{p^2} = \frac{1-p}{p^2},$$

as claimed.
Bounding Coupon Collector purchases - Chebyshev

Lemma (3.8)

For any geometric random variable $X$ with parameter $p$, $\mathbb{E}[X] = p^{-1}$ and $\text{Var}[X] = \frac{1-p}{p^2}$.

Proof (cont’d).

So $\mathbb{E}[X^2] = \frac{p}{1-p} \frac{(1-p)(2-p)}{p^3} = \frac{2-p}{p^2}$, hence

$$\mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{(2-p) - 1}{p^2} = \frac{1-p}{p^2},$$

as claimed. \[
\]

RC (2017/18) – Lecture 6 – slide 11
Lemma (3.8)

*For any geometric random variable $X$ with parameter $p$, $E[X] = p^{-1}$ and $\text{Var}[X] = \frac{1-p}{p^2}$.\*

**Proof (cont’d).**

So $E[X^2] = \frac{p}{1-p} \frac{(1-p)(2-p)}{p^3} = \frac{2-p}{p^2}$, hence

$$
E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{(2-p) - 1}{p^2} = \frac{1-p}{p^2},
$$

as claimed.
Bounding Coupon Collector purchases - Chebyshev

\[
\Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2} = \frac{\sum_{j=1}^{n} \text{Var}[X_j]}{a^2}. 
\]

Each individual \(X_j\) is geometric with parameter \(\frac{n-(j-1)}{n}\). So each \(X_j\) has

\[
\text{Var}[X_j] = \frac{j-1}{n} \left( \frac{n}{n+1-j} \right)^2 \leq \left( \frac{n}{n+1-j} \right)^2.
\]

Hence

\[
\text{Var}[X] \leq n^2 \sum_{j=1}^{n} \left( \frac{1}{n+1-j} \right)^2 = n^2 \sum_{j=n}^{1} \left( \frac{1}{j} \right)^2 = \frac{\pi^2 n^2}{6}.
\]

(using Euler’s series for the \(\frac{\pi}{6}\), see page 5 of “TCS cheat sheet”).
Bounding Coupon Collector purchases - Chebyshev

$$\Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2} = \frac{\sum_{j=1}^{n} \text{Var}[X_j]}{a^2}.$$ 

Each individual $X_j$ is geometric with parameter $\frac{n-(j-1)}{n}$. So each $X_j$ has 

$$\text{Var}[X_j] = \frac{j-1}{n} \left( \frac{n}{n+1-j} \right)^2 \leq \left( \frac{n}{n+1-j} \right)^2.$$ 

Hence 

$$\text{Var}[X] \leq n^2 \sum_{j=1}^{n} \left( \frac{1}{n+1-j} \right)^2 = n^2 \sum_{j=n}^{1} \left( \frac{1}{j} \right)^2 = \frac{\pi^2 n^2}{6}.$$ 

(using Euler’s series for the $\frac{\pi}{6}$, see page 5 of “TCS cheat sheet”).
Bounding Coupon Collector purchases - Chebyshev

\[ \Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2} \leq \frac{\sum_{j=1}^{n} \text{Var}[X_j]}{a^2}. \]

Each individual \( X_j \) is geometric with parameter \( \frac{n-(j-1)}{n} \). So each \( X_j \)

has

\[ \text{Var}[X_j] = \frac{j-1}{n} \left( \frac{n}{n+1-j} \right)^2 \leq \left( \frac{n}{n+1-j} \right)^2. \]

Hence

\[ \text{Var}[X] \leq n^2 \sum_{j=1}^{n} \left( \frac{1}{n+1-j} \right)^2 = n^2 \sum_{j=n}^{1} \left( \frac{1}{j} \right)^2 = \frac{\pi^2 n^2}{6}. \]

(using Euler’s series for the \( \frac{\pi}{6} \), see page 5 of “TCS cheat sheet”).
Bounding Coupon Collector purchases - Chebyshev

$$\Pr[|X - E[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2} = \frac{\sum_{j=1}^{n} \text{Var}[X_j]}{a^2}.$$  

Each individual $X_j$ is geometric with parameter $\frac{n-(j-1)}{n}$. So each $X_j$ has

$$\text{Var}[X_j] = \frac{j-1}{n} \left(\frac{n}{n+1-j}\right)^2 \leq \left(\frac{n}{n+1-j}\right)^2.$$  

Hence

$$\text{Var}[X] \leq n^2 \sum_{j=1}^{n} \left(\frac{1}{n+1-j}\right)^2 = n^2 \sum_{j=n}^{1} \left(\frac{1}{j}\right)^2 = \frac{\pi^2 n^2}{6}. $$

(using Euler’s series for the $\frac{\pi}{6}$, see page 5 of “TCS cheat sheet”).
We know \( \text{Var}[X] \leq \frac{\pi^2 n^2}{6} \) for our coupon collector process.

Suppose we are willing to make \( 2\text{E}[X] \) (about \( 2n \ln(n) \)) purchases.

Buying this number of packets, the probability we fail to get all cards is

\[
\Pr[X > 2\text{E}[X]] = \Pr[X - \text{E}[X] > \text{E}[X]] \leq \Pr[|X - \text{E}[X]| > \text{E}[X]]
\]

We can upper bound the probability of the bad event \( |X - \text{E}[X]| > \text{E}[X] \) (aka “didn’t get all cards”) using Chebyshev’s Inequality with \( a = \text{E}[X] \):

\[
\Pr[|X - \text{E}[X]| \geq \text{E}[X]] \leq \frac{\text{Var}[X]}{\text{E}[X]^2} \leq \frac{\pi^2 n^2}{6n^2 H(n)^2}.
\]

This value simplifies to \( \frac{\pi^2}{6H(n)^2} \), which is less than \( \frac{2}{\ln(n)^2} \), much better probability than \( 1/2 \) (given by Markov for \( 2\text{E}[X] \) purchases).

*RC (2017/18) – Lecture 6 – slide 13*
We know \( \text{Var}[X] \leq \frac{\pi^2 n^2}{6} \) for our coupon collector process.

Suppose we are willing to make \( 2E[X] \) (about \( 2n \ln(n) \)) purchases.

Buying this number of packets, the probability we fail to get all cards is

\[
\begin{align*}
\Pr[X > 2E[X]] &= \Pr[X - E[X] > E[X]] \\
&\leq \Pr[|X - E[X]| > E[X]]
\end{align*}
\]

We can upper bound the probability of the bad event \( |X - E[X]| > E[X] \) (aka “didn’t get all cards”) using Chebyshev’s Inequality with \( a = E[X] \):

\[
\Pr[|X - E[X]| \geq E[X]] \leq \frac{\text{Var}[X]}{E[X]^2} \leq \frac{\pi^2 n^2}{6n^2 H(n)^2}.
\]

This value simplifies to \( \frac{\pi^2}{6H(n)^2} \), which is less than \( \frac{2}{\ln(n)^2} \), much better probability than \( 1/2 \) (given by Markov for \( 2E[X] \) purchases).
Bounding Coupon Collector purchases - Chebyshev

We know $\text{Var}[X] \leq \frac{\pi^2 n^2}{6}$ for our coupon collector process.

Suppose we are willing to make $2E[X]$ (about $2n\ln(n)$) purchases.

Buying this number of packets, the probability we fail to get all cards is

$$\Pr[X > 2E[X]] = \Pr[X - E[X] > E[X]] \leq \Pr[|X - E[X]| > E[X]]$$

We can upper bound the probability of the bad event $|X - E[X]| > E[X]$ (aka “didn’t get all cards”) using Chebyshev’s Inequality with $a = E[X]$:

$$\Pr[|X - E[X]| \geq E[X]] \leq \frac{\text{Var}[X]}{E[X]^2} \leq \frac{\pi^2 n^2}{6n^2 H(n)^2}.$$ 

This value simplifies to $\frac{\pi^2}{6H(n)^2}$, which is less than $\frac{2}{\ln(n)^2}$, much better probability than $1/2$ (given by Markov for $2E[X]$ purchases).
Bounding Coupon Collector purchases - Chebyshev

We know \( \text{Var}[X] \leq \frac{\pi^2 n^2}{6} \) for our coupon collector process.

Suppose we are willing to make \( 2E[X] \) (about \( 2n \ln(n) \)) purchases.

Buying this number of packets, the probability we fail to get all cards is

\[
\Pr[X > 2E[X]] = \Pr[X - E[X] > E[X]] \leq \Pr[|X - E[X]| > E[X]]
\]

We can upper bound the probability of the bad event \( |X - E[X]| > E[X] \) (aka “didn’t get all cards”) using Chebyshev’s Inequality with \( a = E[X] \):

\[
\Pr[|X - E[X]| \geq E[X]] \leq \frac{\text{Var}[X]}{E[X]^2} \leq \frac{\pi^2 n^2}{6n^2 H(n)^2}.
\]

This value simplifies to \( \frac{\pi^2}{6H(n)^2} \), which is less than \( \frac{2}{\ln(n)^2} \), much better probability than 1/2 (given by Markov for \( 2E[X] \) purchases).
If we are willing to make as many as $10\mathbb{E}[X]$ purchases, then we will upper-bound (probability of the "bad" scenario) $\Pr[|X - \mathbb{E}[X]| \geq 9\mathbb{E}[X]]$ by setting $a = 9\mathbb{E}[X]$ in Chebyshev’s Inequality:

$$\Pr[|X - \mathbb{E}[X]| \geq 9\mathbb{E}[X]] \leq \frac{\text{Var}[X]}{(9\mathbb{E}[X])^2} \leq \frac{\pi^2 n^2}{6 \cdot 81 \cdot n^2 H(n)^2}.$$ 

Working details with $\pi$, the right-hand side is at most $\frac{1}{49 \cdot H(n)^2}$, much much less/better than the $\frac{1}{10}$ we got with Markov’s Ineq.
Wrapping up today

On Tuesday next we will continue the theme of “bounding deviation from the mean" by introducing some stronger concentration inequalities called Chernoff/Hoeffding bounds (which hold for iterations of independent Poisson trials, and related distributions).

- I have finally distributed the spec for coursework, deadline is now 4pm, Thursday, 15th Feb, 2018, and will return feedback within 2 weeks.
- Tutorials are starting next week. I have distributed the first tutorial sheet today.
- Anyone who wants to prepare in advance for next week, take a look at the early sections of Chapter 4.