Lecture 8-a: Deviation from the Expectation

Instructor: Eli Upfal
Brown University Computer Science

Figure credits: Bertsekas & Tsitsiklis, *Introduction to Probability*, 2008
Pitman, *Probability*, 1999
Expectation and Variance
Markov’s Inequality
Chebyshev's Inequality
Expectation and Variance

- The **expectation or expected value** of a random variable
  
  \[ E[X] = \sum_{x \in X} x p_X(x) \]

- The **variance** is the **expected squared deviation** of a random variable from its mean (the following definitions are equivalent):
  
  \[ \text{Var}[X] = E[(X - E[X])^2] = \sum_{x \in X} (x - E[X])^2 p_X(x) \]

  \[ \text{Var}[X] = E[X^2] - E[X]^2 = \left( \sum_{x \in X} x^2 p_X(x) \right) - \left( \sum_{x \in X} x p_X(x) \right)^2 \]

- The **standard deviation** is the square root of the variance:
  
  \[ \sigma_X = \text{Std}[X] = \sqrt{\text{Var}[X]} \]
Sums of Independent Variables

- If $Z = X + Y$ and random variables $X$ and $Y$ are independent, we have:
  \[
  E[Z] = E[X] + E[Y] \quad \text{Var}[Z] = \text{Var}[X] + \text{Var}[Y]
  \]
  For any variables $X$, $Y$. Only for independent $X$, $Y$.

- Interpretation: Adding independent variables increases variance:
  \[
  \text{Var}[Z] \geq \text{Var}[X] \quad \text{and} \quad \text{Var}[Z] \geq \text{Var}[Y]
  \]

- The **standard deviation** of a sum of independent variables is then:
  \[
  \sigma_Z = \sqrt{\sigma_X^2 + \sigma_Y^2} \quad \sigma_X = \sqrt{\text{Var}[X]}, \sigma_Y = \sqrt{\text{Var}[Y]}, \sigma_Z = \sqrt{\text{Var}[Z]}
  \]

- Identity used in proof: If $X$ and $Y$ are **independent** random variables,
  \[
  E[XY] = E[X]E[Y] \quad \text{if} \quad p_{XY}(x, y) = p_X(x)p_Y(y)
  \]
  This equality does not hold for general, dependent random variables.
Reminder: Bernoulli Distribution

A Bernoulli or indicator random variable $X$ has one parameter $p$:

$$p_X(1) = p, \quad p_X(0) = 1 - p, \quad X = \{0, 1\}$$

For an indicator variable, expected values are probabilities:

$$E[X] = p$$

Variance of Bernoulli distribution:

$$\text{Var}[X] = E \left[ (X - p)^2 \right] = p(1 - p)$$

$$E[X^2] = p$$

- Fair coin ($p=0.5$) has largest variance
- Coins that always come up heads ($p=1.0$), or always come up tails ($p=0.0$), have variance 0
Suppose you flip $n$ coins with bias $p$, count number of heads

A **binomial** random variable $X$ has parameters $n, p$: $p_X(k) = \binom{n}{k} p^k (1 - p)^{n-k}$

If $X_i$ is a Bernoulli variable indicating whether toss $i$ comes up heads, then $X = \sum_{i=1}^{n} X_i$

Then because tosses are **independent**:

$E[X] = np$

$\text{Var}[X] = np(1 - p)$
CS145: Lecture 8-a Outline

- Expectation and Variance
- Markov’s Inequality
- Chebyshev's Inequality
Markov’s Inequality

**Theorem**

[Markov Inequality] For any non-negative random variable, and for all \( a > 0 \),

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

Fix some constant \( a > 0 \), and define

\[
Y_a = \begin{cases} 
0, & \text{if } X < a, \\
a, & \text{if } X \geq a.
\end{cases}
\]

\[
aP(X \geq a) = E[Y_a] \leq E[X]
\]
Markov’s Inequality

**Theorem**

[Markov Inequality] *For any non-negative random variable, and for all* \( a > 0 \),

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

Fix some constant \( a > 0 \), and define

\[
Y_a = \begin{cases} 
0, & \text{if } X < a, \\
aw, & \text{if } X \geq a.
\end{cases}
\]

- No such inequality would hold if \( X \) could take negative values. Why?
- If \( a < E[X] \), Markov’s inequality is vacuous, but no better bound is possible. Why?
CS145: Lecture 8-a Outline

- Expectation and Variance
- Markov’s Inequality
- Chebyshev's Inequality
Chebyshev’s Inequality

Theorem

For any random variable $X$, and any $a > 0$,

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

Proof.

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

By Markov inequality

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

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

Theorem

For any random variable \( X \), and any \( a > 0 \),

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

Another way of parameterizing Chebyshev’s inequality:

\[
\mu = E[X], \quad \sigma = \sqrt{Var[X]}
\]

\[
P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}
\]

Chebyshev bound is vacuous (above one) for events less than one standard deviation from the mean. But this could be likely!
Chebyshev’s Inequality

- Another way of parameterizing Chebyshev’s inequality:

\[
\mu = E[X], \quad \sigma = \sqrt{\text{Var}[X]} \\
\Pr(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}
\]

- Chebyshev bound is vacuous (above one) for events less than one standard deviation from the mean. But this could be likely!
We flip a fair coin $n$ times.
What is the probability of getting more than $3n/4$ heads?


Markov’s Inequality:

$$Pr\{X \geq \frac{3n}{4}\} \leq \frac{E[X]}{3n/4} \leq \frac{n/2}{3n/4} = \frac{2}{3}$$

Chebyshev’s Inequality:

$$Pr\{X \geq \frac{3n}{4}\} \leq Pr\{|X - \frac{n}{2}| \geq \frac{n}{4}\} \leq \frac{Var[X]}{(n/4)^2} = \frac{n/4}{n^2/16} = \frac{4}{n}$$