CS5314
Randomized Algorithms

Lecture 7: Moments and Deviations
(Markov Inequality, Variance)

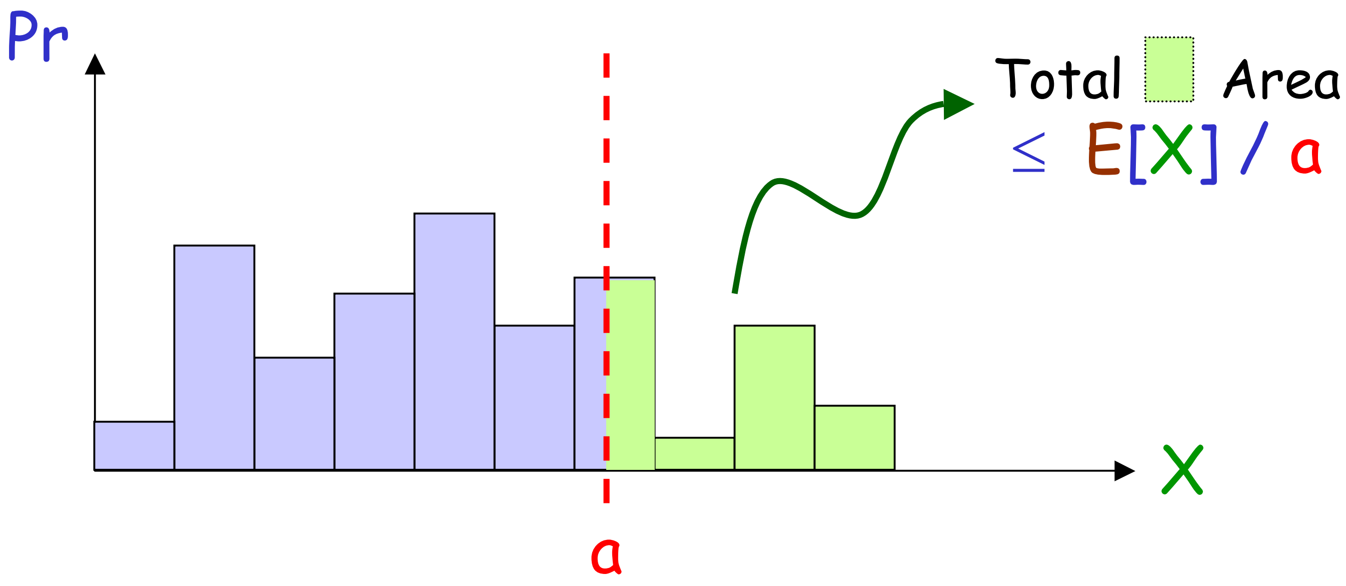
Objectives

- Introduce Markov Inequality
- Define Variance and Moments of an RV

Markov Inequality

Theorem: Let X be a random variable that takes on non-negative values only. Then, for any positive a

$$\Pr(X \geq a) \leq E[X] / a$$



First Proof (from def of E[X])

$E[X] = \Sigma_j\, j\, Pr(X=j)$

$= \Sigma_{0 \le j < a}\, j\, Pr(X=j) + \Sigma_{j \ge a}\, j\, Pr(X=j)$ [why?]

$\ge \Sigma_{0 \le j < a}\, 0\, Pr(X=j) + \Sigma_{j \ge a}\, a\, Pr(X=j)$

$= 0 + a\, \Sigma_{j \ge a}\, Pr(X=j)$

$= a\, Pr(X \ge a)$

Thus, $Pr(X \ge a) \le E[X] / a$

Second Proof (from $E[X \mid X \geq a]$)

$$E[X] = E[X \mid X < a] \Pr(X < a) + E[X \mid X \geq a] \Pr(X \geq a)$$

$$\geq 0 \Pr(X < a) + a \Pr(X \geq a)$$

$$= a \Pr(X \geq a)$$

Thus, $\Pr(X \geq a) \leq E[X] / a$

Third Proof (using indicator)

Let I be an indicator random variable with:

$I = 1$ if $X \geq a$

$I = 0$ otherwise

Recall: $E[I] = \Pr(I = 1) = \Pr(X \geq a)$

Our target is to bound $E[I]$ (w.r.t. $E[X]$).

So how is $E[I]$ related to $E[X]$??

In particular, how is I related to X ??

Third Proof (using indicator)

Note: $I = 1$ if $X \geq a$

$I = 0$ otherwise

Also $X \geq 0$, so that we always have

$$I \leq X/a$$

Thus, $E[I] \leq E[X/a] = E[X]/a$ [why?]

Combining, we have,

$$\Pr(X \geq a) \leq E[X] / a$$

Example

Let us flip a fair coin n times

Question: Can we bound the probability of getting more than $3n/4$ heads ?

Let X = number of heads

So, $E[X] = n/2$

By Markov Inequality,

$$\Pr(X \geq 3n/4) \leq E[X] / (3n/4)$$

$$= (n/2) / (3n/4) = 2/3$$

Variance and Moments

Definition: The k^{th} **moment** of a random variable X is defined as $E[X^k]$

Definition: The **variance** of a random variable X is defined as

$$Var[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

Definition: The **standard deviation** of a random variable X is defined as

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

Linearity of Variance?

Recall: For any random variables X and Y,

$$E[X + Y] = E[X] + E[Y]$$

Is it still true for variance? That is,

$$Var[X + Y] = Var[X] + Var[Y] ?$$

Covariance

Answer: No! In fact, an extra term, called covariance, will be involved…

Definition: The covariance of two random variables X and Y is defined as

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$

Covariance (2)

Theorem: For any random variables X and Y,

$$Var[X + Y] = Var[X] + Var[Y] + 2Cov(X,Y)$$

Proof: Let $A = X - E[X]$ and $B = Y - E[Y]$.

$Var[X+Y] = E[(X+Y - E[X+Y])^2]$

$= E[(X+Y - E[X] - E[Y])^2]$

$= E[(A+B)^2] = E[A^2 + B^2 + 2AB]$

$= E[A^2] + E[B^2] + 2E[AB]$

$= Var[X] + Var[Y] + 2Cov(X,Y)$

Covariance (3)

Theorem (generalized version): For any finite number of random variables X_1, X_2, … , X_k,

$$Var[\sum_j X_j] = (\sum_j Var[X_j]) + 2\sum_{i<j} Cov(X_i,X_j)$$

Proof: By induction (try this at home!)

More on Covariance

Recall: If random variables X and Y are independent, then for every values a and b,

$$\Pr((X=a) \cap (Y=b)) = \Pr(X=a)\Pr(Y=b)$$

Theorem: If X and Y are independent random variables, then

$$E[XY] = E[X]\,E[Y]$$

Proof

$E[XY]$

$= \sum_a \sum_b ab \Pr((X=a) \cap (Y=b))$

$= \sum_a \sum_b ab \Pr(X=a) \Pr(Y=b)$

$= \sum_a a \Pr(X=a) \sum_b b \Pr(Y=b)$

$= E[X] E[Y]$

More on Covariance (2)

Lemma: If X and Y are independent random variables, then

$$Cov(X,Y) = 0$$

Proof: $Cov(X,Y)$

$= E[(X - E[X]) (Y - E[Y])]$

$= E[X - E[X]] \; E[Y - E[Y]]$ ...... [why?]

$= (E[X] - E[X]) \; E[Y - E[Y]]$

$= 0$

More on Covariance (3)

Corollary: If X and Y are independent random variables, then

$$Var[X + Y] = Var[X] + Var[Y]$$

Corollary: If $X_1, X_2, \dots, X_k$, are pairwise independent random variables, then

$$Var[\sum_j X_j] = \sum_j Var[X_j]$$

Pairwise? Mutually?

Recall: For random variables $X_1, X_2, \ldots, X_k$, they are mutually independent if for any subset $I \subseteq [1,k]$, and any values x_i

$$\Pr\left(\bigcap_{i\in I} X_i=x_i\right) = \prod_{i\in I} \Pr(X_i=x_i)$$

But for random variables $X_1, X_2, \ldots, X_k$ to be pairwise independent, we only need each pair of X_i, X_j to be independent

Thus, mutually independent implies pairwise independent, but not the other way round

Any Example?

Suppose we roll two fair dice. Let X, Y, Z be indicator random variables, such that

-- $X = 0$ if first die is even,
   $X = 1$ otherwise;
-- $Y = 0$ if second die is even,
   $Y = 1$ otherwise;
-- $Z = 0$ if total sum is even,
   $Z = 1$ otherwise