Randomness in Computing

CS 537

LECTURE 8

Last time

- Randomized quicksort
- Markov's inequality
- Variance
- **Today**
- Variance, covariance
- Chebyshev's inequality
- Variance of Binomial and Geometric RVs

Random variables: variance

- The variance of a random variable X with expectation $E[X] = \mu$ is
$$\mathrm{Var}[X] = E[(X - \mu)^2].$$
- Equivalently, $\mathrm{Var}[X] = E[X^2] - \mu^2$.

- The standard deviation of X is $\sigma[X] = \sqrt{\mathrm{Var}[X]}$.

Exercise

Compute expectation and variance

- Number of fixed points of a permutation. Let X be the number of students that get their hats back when n students randomly switch hats, so that every permutation of hats is equally likely.

Random variables: covariance

- The covariance of two random variables X and Y with expectations $E[X] = \mu_X$ and $E[Y] = \mu_Y$ is

$$\mathrm{Cov}(X, Y) = E\left[(X - \mu_X)(Y - \mu_y)\right].$$

- Theorem. For any two random variables X and Y,

$$\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\,\mathrm{Cov}(X, Y).$$

Independent RVs

- Random variables X and Y on the same probability space are independent if for all values a and b,

  the events $X = a$ and $Y = b$ are independent.

  Equivalently, for all a, b,

$$\Pr[X = a \wedge Y = b] = \Pr[X = a] \cdot \Pr[Y = b].$$

- Theorem. For independent random variables X and Y,
  - $E[XY] = E[X] \cdot E[Y]$.
  - $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$.
  - Cov(X,Y)=0.

Example: n coin tosses

- Let X be the number of HEADS in n tosses of a biased coin with HEADS probability p.
- We know: X has binomial distribution $\text{Bin}(n, p)$.
- What is the variance of X?

Answer: $np(1-p)$.

Example: Geometric RV

- Let X be the # of coin tosses until the first HEADS of a biased coin with HEADS probability p.
- We know: X has geometric distribution Geom(p).
- What is the variance of X?

Answer: $\frac{1-p}{p^2}$.

Variance: additional facts

- Theorem. For $a, b \in \mathbb{R}$ and a random variable X,
$$\mathrm{Var}[aX + b] = a^2 \mathrm{Var}[X].$$
- Theorem. If $X_1, \ldots, X_n$ are pairwise independent random variables, then
$$\mathrm{Var}[X_1 + \cdots + X_n] = \mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_n].$$

Chebyshev's Inequality

- Theorem. For a random variable X and $a > 0$,

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

- Alternatively: Then, for all $t > 1$,

$$\Pr[|X - E[X]| \geq t \cdot \sigma[X]] \leq \frac{1}{t^2}.$$

- Example 1: $X \sim \mathrm{Bin}(n, 1/2)$.

  Bound $\Pr\left[X > \frac{3n}{4}\right]$ using Markov and Chebyshev.

- Example 2: Coupon Collector Problem.

  Bound $\Pr[X > 2nH_n]$ using Markov and Chebyshev.