Randomness in Computing

LECTURE 7

CS 537

Last time

- Randomized quicksort
- Markov's inequality
- Variance

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

Recall: variance

- The variance of a random variable X with expectation $\mathbb{E}[X] = \mu$ is

$$\mathrm{Var}[X] = \mathbb{E}[(X-\mu)^2].$$

- Equivalently, $\mathrm{Var}[X] = \mathbb{E}[X^2] - \mu^2$.

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

CS 537 Compute expectation and variance

- Fair die. Let X be the number showing on a roll of a die.

$$\mathrm{Var}[X] = E[X^2] - \mu^2$$

$$= \frac{1+4+9+16+25+36}{6} - \left(\frac{7}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

- Uniform distribution. X is uniformly distributed over $[n]$.
  - The sum of the first n squares is $\frac{n(n+1)(2n+1)}{6}$

$$\mathrm{Var}[X] = \frac{1}{n}\sum_{i\in[n]} i^2 - \left(\frac{1}{n}\sum_{i\in[n]} i\right)^2$$

$$= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}$$

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.

Solution: X_i= the indicator R.V. for person i getting their hat back.

$$X = X_1 + \cdots + X_n$$

By linearity of expectation and symmetry, $\mathbb{E}[X] = n \cdot \mathbb{E}[X_1] = n \cdot \frac{1}{n} = 1$

$$\mathbb{E}[X^2] = \mathbb{E}[(X_1 + \cdots + X_n)^2]$$
$$= n \cdot \mathbb{E}[X_1^2] + n(n-1) \cdot \mathbb{E}[X_1 \cdot X_2]$$
$$= n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n(n-1)} = 2$$

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 2 - 1 = 1$$

Random variables: covariance

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

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

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

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

- Proof: $\mathrm{Var}[X + Y]$

$$= \mathbb{E}\left[\left((X + Y) - \mathbb{E}[X + Y]\right)^2\right]$$

$$= \mathbb{E}[((X - \mu_X) + (Y - \mu_Y))^2]$$

$$= \mathbb{E}[(X - \mu_X)^2 \quad + \quad (Y - \mu_Y)^2 \quad + \quad 2(X - \mu_X)(Y - \mu_Y)]$$

$$= \mathbb{E}[(X - \mu_X)^2] + \mathbb{E}[(Y - \mu_Y)^2] + 2\mathbb{E}[(X - \mu_X)(Y - \mu_Y)]$$

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

CS 537 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,
  - $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$.
  - $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$.
  - 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 $\mathrm{Bin}(n, p)$.
- 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 $a > 0$,

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

- Proof: $\Pr[|X - \mathbb{E}[X]| \geq a] = \Pr[\underbrace{(X - \mathbb{E}[X])^2}_{Y \geq 0} \geq a^2]$

$$\leq \frac{\mathbb{E}[Y]}{a^2} \qquad \text{(by Markov)}$$

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

$$= \frac{\mathrm{Var}[X]}{a^2}$$

Chebyshev's Inequality

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

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

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

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

- 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.