CS 70 Discrete Mathematics for CS
Spring 2006 Vazirani

# HW 12

## Due Friday May 5

1. **Variance**

   (a) Let $X$ and $Y$ be independent random variables. Express $Var(X-Y)$ in terms of $Var(X)$ and $Var(Y)$.

   (b) Given a random variable $X$ with $E(X)=\mu$ and $Var(X)=\sigma^2$, let the random variable $Y=XZ$ where $Z$ is defined as follows: Flip a fair coin. If it comes up $H$, then $Z=1$, otherwise $Z=-1$. Find $Var(Y)$.

2. **Chebyshev Inequality**
   A friend tells you about a course called "Laziness in Modern Society" that requires almost no work. You hope to take this course so that you can devote all of your time to CS70. At the first lecture, the professor announces that grades will depend only a midterm and a final. The midterm will consist of three questions, each worth 10 points, and the final will consist of four questions, also each worth 10 points. He will give an A to any student who gets at least 60 of the possible 70 points.

   However, speaking with the professor in office hours you hear some very disturbing news. He tells you that to save time he will be grading as follows. For each student's midterm, he'll choose a real number randomly from a distribution with mean $\mu=5$ and variance $\sigma^2=1$. He'll mark each of the three questions with that score. To grade the final, he'll again choose a random number from the same distribution, independent of the first number, and will mark all four questions with that score.

   If you take the class, what will the mean and variance of your total class score be? Can you conclude that you have less than a 5% chance of getting an A? Why?

3. **Poisson Distribution** A textbook has on average one misprint per page. What is the chance that you see exactly 4 misprints on page 1? What is the chance that you see exactly 4 misprints on some page in the textbook *of 250 pages*?

4. **Significance Levels** A small college soccer team won its conference championship 9 times in the first 20 years of existence. Then for the next twenty years it won only 3 times. Is this significant (i.e. 95% significance level - recall that this means deviation greater than $1.96\sigma$)? Is it very significant (i.e. 99% significance level - recall that this means deviation greater than $2.58\sigma$)?

5. **A paradox in conditional probability?**
   Here is some on-time arrival data for two airlines, A and B, into the airports of Los Angeles and Chicago. (Predictably, both airlines perform better in LA, which is subject to less flight congestion and less bad weather.)

| | Airline A | | Airline B | |
|---|---|---|---|---|
| | # flights | # on time | #flights | # on time |
| Los Angeles | 600 | 534 | 200 | 188 |
| Chicago | 250 | 176 | 900 | 685 |

**(a)** Which of the two airlines has a better chance of arriving on time into Los Angeles? What about Chicago?

**(b)** Which of the two airlines has a better chance of arriving on time overall?

**(c)** Do the results of parts (a) and (b) surprise you? Explain the apparent paradox, and interpret it in terms of conditional probabilities.