Math 3215: Lecture 10

Will Perkins

February 9, 2012

1 Variance

The variance of a random variable is defined as:

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

this is the same as:

$$var(X) = \mathbb{E}(X^2) - (\mathbb{E}X)^2$$

prove that.
Calculate some variances:

- Let $Y = 7$ with probability 1. $var(Y) = ?$
- Let X_i be the indicator r.v. of a getting heads with a fair coin. $var(X_i) = ?$
- X_i is the indicator of a p-biased coin. $var(X_i) = ?$
- If X and Y are independent, prove that $\mathbb{E}[XY] = \mathbb{E}X\mathbb{E}Y$. (*note*: this does not go the other way)
- what is $var(aX)$?
- what is $var(X + b)$?

Let Y be the number of heads in 10 flips of a fair coin.

- What is $\mathbb{E}Y$?
- What is $var(Y)$?

2 What does variance mean?

We can describe random variables by their expectation and variance.

The $\mathbb{E}X$, also called the mean of X, is a measurement of how large X is on average. The variance of X is a measurement of how much X varies from its mean.

3 How to calculate variance?

We have the formula $var(X) = \mathbb{E}(X^2) - (\mathbb{E}X)^2$. You know how to calculate the mean, so how do we calculate $\mathbb{E}(X^2)$?

1. Sum over all outcomes: $\mathbb{E}(X^2) = \sum_{x \in S} p(x)X(x)^2$
2. Sum over all possible values of X: $\mathbb{E}(X^2) = \sum_{t: P_X(t) \neq 0} P_X(t)t^2$
3. break up X into a sum (of indicator random variables if X is a count): $X = \sum_i X_i$, then$\mathbb{E}(X^2) = \sum_{i,j} \mathbb{E}(X_i X_j)$

4. in particular if X is the sum of indicator rv's $X_i = I_{A_i}$, then $\mathbb{E}(X_i^2) = \mathbb{E}X_i = \Pr(A_i$ and $\mathbb{E}(X_i X_j) = \Pr[A_i \cap A_j]$.

Examples: calculate the following variances by breaking up into a sum:

1. $var(X)$ where X is the number of 6's in 10 rolls of a fair die
2. $var(Y)$ where Y is the number of red cards in a hand of 5 cards
3. $var(Z)$ where Z is the number of Kings in a hand of 5 cards

4 Chebyshev's Inequality

If a store sold an average of $10 worth of umbrellas a month in 2010, what is the most they could have sold in a single month?

Using the same logic, prove the following:

- If X is always non-negative, then $\mathbb{E}X \geq \Pr[X \geq 1]$
- If X is always non-negative, then $\Pr[X \geq t] \leq \frac{\mathbb{E}X}{t}$
- $var(X) \geq t^2 \Pr[|X - \mathbb{E}X| \geq t]$

You've proved Chebyshev's Inequality:

$$\Pr[|X - \mathbb{E}X| \geq t] \leq \frac{var(X)}{t^2}$$

5 Law of Large Numbers

Say 60% of people prefer Clinton to Dole. What's an upper bound on the probability that we sample n people, and less that 55% of them prefer Clinton?

A sequence of i.i.d. random variables $X_1, X_2, \ldots X_n$ is a collection of random variables that are *independent* and *identically distributed*. We usually let $S_n = \sum_{i=1}^n X_i$ be the sum of the sequence. We are concerned with the emperical average $\frac{S_n}{n}$ in cases like surveys, medical experiments, data collection.

Law of Large Numbers

Let $X_1, \ldots X_n$ be an i.i.d. sequence of random variables with $\mathbb{E}X_i^2 < \infty$. Let $S_n = \sum_{i=1}^n X_i$. Then

$$\Pr\left[\left|\frac{S_n}{n} - \mathbb{E}X_i\right| \geq t\right] \leq \frac{var(X_i)}{t^2 n}$$

In particular, for t as small as we like, say .01, we can pick n large enough so that this probability is as small as we like, say .001.