Chapter 7.4: Expected Value and Variance

Monday, August 3

Summary

- Indictator Variables: $I_E(s) = \begin{cases} 1 & s \in E \\ 0 & s \notin E \end{cases}$
- Variance: $Var(X) = E([X - E(X)]^2) = E(X^2) - E(X)^2$
- If X and Y are independent then $Var(X+Y) = Var(X) + Var(Y)$
- If $X_1, \ldots, X_n$ are pairwise independent then $Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i)$.
- $Var(aX+b) = a^2 \cdot Var(X)$.
- Markov's Inequality: If X is non-negative then $p(X \geq a) \leq E(X)/a$.
- Chebyshev's Inequality: $p(|X - E(X)| \geq r) \leq Var(X)/r^2$
- Covariance: $Cov(X,Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$
- $Cov(X,Y) = Cov(Y,X)$, $Cov(X, Y+Z) = Cov(X,Y) + Cov(X,Z)$
- E and F are positively correlated if $p(E \cap F) > p(E)p(F)$.

Variance

1. (★) If X is the sum of a rolled pair of dice, what is $Var(X)$?
2. (★) If a coin with a 25% chance of landing on heads and X is the number of heads that result from 50 flips of the coin, what is $Var(X)$?
3. If E is some event, what is $Var(I_E)$?
4. Prove by induction: If $X_1, \ldots, X_n$ are *mutually* independent random variables, prove by induction that $Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i)$.
5. Why does induction not work if the variables are only pairwise independent?

Chebyshev's Inequality

1. (★) If a fair coin is flipped 100 times, use Chebyshev's inequality to bound the probability of getting at least 55 or at most 45 heads.

2. Bound the probability of getting at least 65 or at most 35 heads.

3. Prove: if you flip N fair coins and X_N is the number of heads, then $\lim_{n\to\infty} p(|X - N/2| \geq \epsilon N) = 0$ for any $\epsilon > 0$.

4. Prove: If $X_1, X_2, \ldots$ are independent identically distributed (i.i.d.) random variables with finite variance and expected value μ, and $\mu_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ (the average of the first n variables), then $\lim_{n\to\infty} p(|\mu_n - \mu| > \epsilon) = 0$ for any $\epsilon > 0$.

Covariance

1. Show that if k is any constant (or more precisely, a random variable that always takes on the value k) then $Cov(k, X) = 0$.

2. (★) Show that $Cov(aX + b, Y) = a \cdot Cov(X, Y)$.

3. If $p(E) = .6$ and $p(F) = .8$, what can we say about the correlation of E and F?

4. (★) Flip 100 coins and let X be the number of times HT appears. Find $Var(X)$.

5. Flip 100 coins. Let X be the number of times HHH appears and let Y be the number of times HHT appears. Find $Cov(X, Y)$.

6. Prove that if $P(E|F) > P(E)$ then $P(E|\overline{F}) < P(E)$ (in other words, if E and F are positively correlated then E and $\overline{F}$ are negatively correlated).