

Stat 531 / Math 547, Spring 2007, Homework Set I

Due Friday February 8, 2008

1. Durrett Exercise 1.2 of Chapter 4 (Bayes' Formula).
2. Define a probability model for the following conditional probability statement. Your model should contain a probability triple, $(\Omega, \mathcal{F}, \mathbf{P})$ a subfield $\mathcal{F} \subseteq \mathcal{F}_0$ of information and an event A . Then you should compute the conditional probability $\mathbf{P}(A | \mathcal{F})$.

You find a strange-looking coin. Having no idea what is the probability p of flipping HEADS with this coin, you guess p to be uniformly distributed on $[0, 1]$. You flip once and it comes up HEADS. What is the chance that your next flip is HEADS?

3. Fix integers $-M < k < M$. Let $\{S_n\}$ be a simple random walk whose increments $\{X_n\}$ are IID fair ± 1 coin flips, started from $S_0 = k$. Let

$$\tau := \inf\{n : S_n = \pm M\}$$

be the hitting time of $\{-M, M\}$. Use Bayes' rule to compute $\mathbf{P}(X_1 = 1 | S_\tau = M)$. What is the limit as $M \rightarrow \infty$? Now assume $k > 0$ and let

$$\rho := \inf\{n : S_n \in \{0, M\}\}$$

be the time to hit 0 or M . Compute $\mathbf{P}(X_1 = 1 | S_\rho = M)$. Again, identify the limit as $M \rightarrow \infty$, which characterizes **simple random walk conditioned to escape to $+\infty$** .

4. (a) Durrett Exercise 1.9 (decomposing the variance).
(b) A data set includes a measurement of a certain blood protein and a task performance score for each of 100 subjects. Make a probability model (a space, sigma field, measure, sub-sigma field, random variable, etc.) in which the statement "three quarters of the variation in performance may be explained by the

level of blood protein A” may be made into a mathematical statement. Do not make more assumptions than necessary about the distributions of the random variables.

- (c) Now assume that protein level and performance are jointly normal with some unknown covariance matrix. State (but you need not justify) a procedure for estimating the portion of variance of performance explained by the level of blood protein A, given only the above data (blood levels and performance scores on 100 subjects).

5. Durrett Exercise 1.10 (Variance of a random sum)

6. Suppose that a finite population of people is partitioned into households. Let α be the law of a uniformly selected person. Let β be the law of a person selected by first selecting a household uniformly at random and then selecting a person uniformly from that household. Some surveys re-weight their sample by giving each person x a weight of $d\alpha/d\beta(x)$. What is $d\alpha/d\beta$ and why do they re-weight like this?