Assigned: Wednesday, September 26, 2001
Due: Wednesday, October 3, 2001
Reading: Ross, Chapter 3

Reminder: Hour Exam I on Monday October 1 in class, 11:00 a.m — 11:50 a.m.

Noncredit Exercises: Ross Chapter 3, Problems: 53, 58, 59, 62, 63, 70-74, 78, 81

Problems:

1. Let **X** denote the number of Heads observed in 10 tosses of a fair coin.

(a) What kind of random variable is **X**, and what is its average value?

(b) Find $P\{X \le 5 \mid X \ge 4\}$.

(c) Given that **X** = 4, what is the (conditional) probability that the 4th toss was a Head?

(d) You have a strong suspicion that the coin that was tossed is not a fair coin. Nonetheless, your friendly neighborhood bookmaker offers 2-to-1 odds if you want to bet that the 4th toss was a Head knowing only that the event {**X** = 4} has occurred. That is, if you bet $1 that the 4th toss was a Head, then your wealth will increase by $2 if the 4th toss did in fact show a Head, and your wealth will decrease by $1 if the 4th toss resulted in a Tail. The bookie knows the outcome of the 4th toss (but cannot change it after you have placed your bet!) as well as the value of P(Head). Should you take the bet? Why or why not?

2. (Remember: $99\frac{44}{100}$% of all statistics are made up by the writer) The experiment consists of picking a flight at random from all the United Airlines and America West flights landing at Chicago, Los Angeles, Phoenix, San Diego, or San Francisco. Let U and W respectively denote the event that the chosen flight is an United Airlines or an America West flight, let C, L, X, D, and F respectively denote the event that the chosen flight is landing at Chicago, Los Angeles, Phoenix, San Diego, or San Francisco, and let T denote the event that the chosen flight is on time. The conditional probabilities of on-time arrival are as follows:

$P(T|UC) = 0.85$, $P(T|UL) = 0.92$, $P(T|UX) = 0.95$, $P(T|UD) = 0.91$, $P(T|UF) = 0.83$,
$P(T|WC) = 0.78$, $P(T|WL) = 0.88$, $P(T|WX) = 0.92$, $P(T|WD) = 0.85$ $P(T|WF) = 0.73$.

(a) Based on this data, which airline would you say has better on-time performance? Does the answer depend on which airport you are talking about?

(b) Use the fact that {C, L, X, D, F} form a partition of the sample space to show that the average on-time arrival probability $P(T|U)$ for United flights is given by

$$P(T|U) = P(T|UC)P(C|U)+P(T|UL)P(L|U)+P(T|UX)P(X|U)+P(T|UD)\,P(D|U)+P(T|UF)P(F|U)$$

where $P(C|U)$ is the conditional probability that the flight is landing at Chicago given that it is a United flight etc. State a similar expression for $P(T|W)$. (cf. Ross pp. 98-99)

(c) 60% of United Airlines flights land at its hub (snowy Chicago), 15% at each of LA and San Francisco, and 5% at each of Phoenix and San Diego. 75% of America West flights land at **its** hub (sunny Phoenix), 10% at LA, and 5% at each of the other three airports. Use these numbers in the formula of part (b) and show that $P(T|U) < P(T|W)$, i.e., United has a worse average on-time performance even though it beats America West at all the five airports! Write a short explanation of the discrepancy between the per-airport on-time performance and the overall on-time performance.

3. Ross, #12, p. 105 (5th ed.) or p. 104 (6th ed.)

4. Two of the letters from a road sign reading CHICAGO fell down, and were put back up by the friendly neighborhood drunk who is as likely to have put the letters back in their proper places as to have interchanged them.

(a) What is the probability that the sign still reads CHICAGO?

(b) Given that the sign still reads CHICAGO, *which pair* of letters is most likely to have been the ones that fell down?

(c) Given that the sign still reads CHICAGO, what is the probability that C was one of the letters that fell down? What is the probability that H was one of the letters that fell down?

5. ["Take me out to the ball game"] A baseball pitcher's repertoire is limited to fastballs (event F), curveballs (event C), or sliders (event S). It is known that P(C) = 2P(F). The event H that the batter hits the ball has probabilities $P(H|F) = 2/5$, $P(H|C) = 1/4$, and $P(H|S) = 1/6$.

(a) If P(H) = 1/4, what is P(C)?

(b) A fan sitting in the bleachers sees the batter getting a hit, i.e. the event H. He knows the values of $P(H|F)$, $P(H|C)$, and $P(H|S)$, but is sitting too far away to tell whether the pitch was a fastball, a curveball or a slider. What is his maximum-likelihood decision as to what kind of pitch it was?

(c) After cheering the hit, the fan finds that P(F), P(C), and P(S) are listed in the program guide. What is his maximum *a posteriori* probability decision as to the kind of pitch it was?

6. ["Give me an F!" shouted the cheerleader…] H_0, H_1, and H_2 respectively denote the hypotheses that a student is excellent, good, or average (there are no poor students). The number of grade points earned by the student in a course is a random variable **X** that takes on values 3, 6, 9, and 12 only. The professor knows that the pmf of **X** when H_0 is true is

$p_0(12) = 0.75$, $p_0(9) = 0.15$, $p_0(6) = 0.08$, $p_0(3) = 0.02$,

that is, an excellent student has 75% chance of doing well enough on the exam to get an A, 15% chance of a B, etc. Similarly, when H_1 is the true hypothesis, the pmf of **X** is

$p_1(12) = 0.15$, $p_1(9) = 0.6$, $p_1(6) = 0.15$, $p_1(3) = 0.1$

and $p_2(12) = 0.05$, $p_2(9) = 0.1$, $p_2(6) = 0.65$, $p_2(3) = 0.2$.

The professor observes **X** and must decide which of the hypotheses H_0, H_1 and H_2 is true.

(a) What is the professor's maximum-likelihood decision rule?

(b) What is the probability that an excellent student is mistakenly labeled as good? What is the probability that an excellent student is mistakenly labeled as average? What is the probability that an average student is classified either as good or as excellent?

(c) If $P(H_0) = 0.2$, $P(H_1) = 0.55$, and $P(H_2) = 0.25$, what is the probability that the maximum-likelihood decision rule mis-classifies students?

(d) What is the Bayes' decision rule corresponding to these probabilities and what is the probability that the Bayes' decision rule mis-classifies students?

(e) At the Lake Wobegon campus of the University, 95% of students are excellent and 5% are good (and thus they are all above average!) What is Bayes' decision rule in this case? That is, what does the Bayesian professor decide about a student based on the four possible results of the student's exam?

7. ["Have I got a deal for you!"] You are a CEO considering replacing an IC production line, 10% of whose output is defective. A salesman from the Flyby Knight Co. offers a brand new machine which he claims produces only 5% defective items. You are not sure if this claim is true, and suspect that the new machine may not be any better than your current one

(a) Consider the hypotheses: "The new machine has a 5% defect rate" and "The new machine has a 10% defect rate." Which of these should be chosen as the null hypothesis and which as the alternative hypothesis? Explain your choice.

(b) A test run of 100 chips on the new machine produces 7 defectives. What is the value of the likelihood ratio for this observation, and what is the maximum-likelihood decision as to which hypothesis should be believed?

(c) State the maximum-likelihood decision rule in terms of a test on the number of observed defectives in the form: "If the number of defectives in a run of 100 exceeds N, do not buy the new machine." Here you are being asked to compute the value of N.

(d) For a test run of 100 chips, what is the false alarm probability P_{FA}? What is the probability of missed detection P_{MD}?

(e) Repeat parts (c) and (d) for a test run of 1000 chips, and compare your answers to those in part (d). What happens to P_{FA} and P_{MD} as the test run size increases?