Kris Harper
STAT 24400
October 14, 2010

Homework 2

Problem 1. *This problem introduces a simple meteorological model, more complicated versions of which have been proposed in the meteorological literature. Consider a sequence of days and let R_i denote the event that it rains on day i. Suppose that $P(R_i \mid R_{i-1}) = \alpha$ and $P(R_i^c \mid R_{i-1}^c) = \beta$. Suppose further that only today's weather is relevant to predicting tomorrow's; that is $P(R_i \mid R_{i-1} \cap R_{i-2} \cap \cdots \cap R_0) = P(R_i \mid R_{i-1})$.*
(a) If the probability of rain today is p, what is the probability of rain tomorrow?
(b) What is the probability of rain the day after tomorrow?
(c) What is the probability of rain n days from now? What happens as n approaches infinity?

(a) Note that R_{i-1} and R_{i-1}^c union to the entire probability space. Therefore we can use the law of total probability to write

$$P(R_i) = P(R_i \mid R_{i-1})P(R_{i-1}) + P(R_i \mid R_{i-1}^c)P(R_{i-1}^c) = \alpha p + (1-\beta)(1-p) = p(\alpha+\beta-1) + (1-\beta).$$

(b) To find $P(R_{i+1})$ we use the same method as in part (a), but with p replaced with $P(R_i)$. Thus

$$P(R_{i+1}) = P(R_i)(\alpha+\beta-1) + (1-\beta) = (\alpha+\beta-1)^2 p + (1-\beta)(\alpha+\beta).$$

(c) Let $x = (\alpha+\beta-1)$ and $y = (1-\beta)$. We are searching for a_n such that $a_{n+1} = xa_n + y$ and $a_0 = p$. Let

$$a_n = x^n p + y\frac{x^n - 1}{x-1} = (\alpha+\beta-1)^n p + (1-\beta)\frac{(\alpha+\beta-1)^n - 1}{\alpha+\beta-2}.$$

It's clear that $a_0 = p$. On the other hand,

$$\begin{aligned}
a_{n+1} &= x^{n+1}p + y\frac{x^{n+1}-1}{x-1} \\
&= x^{n+1}p + y\frac{(x^{n+1}-x)+(x-1)}{x-1} \\
&= x^{n+1}p + y\left(\frac{x^{n+1}-x}{x-1} + 1\right) \\
&= x\left(x^n p + y\frac{x^n-1}{x-1}\right) + y \\
&= xa_n + y
\end{aligned}$$

so this fits our required relation. Note that $|x| = |\alpha+\beta-1| < 1$ so $\lim_{n\to\infty} x^n = 0$ and

$$\lim_{n\to\infty} a_n = \frac{-y}{x-1} = \frac{\beta-1}{\alpha+\beta-2}.$$

Problem 2. *A factory runs three shifts. In a given day, 1% of the items produced by the first shift are defective, 2% of the second shift's items are defective, and 5% of the third shift's items are defective. If the shifts all have the same productivity, what percentage of the items produced in a day are defective? If an item is defective, what is the probability that is was produced by the third shift?*

Suppose that each shift produces x items. Then $.01x + .02x + .05x/3x = .08/3 = 2\frac{2}{3}\%$ are defective. Let F be the event of a defective item, and let A_i be the event that it was produced on the i^{th} shift. Then by Bayes' rule we have

$$P(A_3 \mid F) = \frac{P(F \mid A_3)P(A_3)}{P(F \mid A_1)P(A_1) + P(F \mid A_2)P(A_2) + P(F \mid A_3)P(A_3)} = \frac{.05\frac{1}{3}}{.01\frac{1}{3} + .02\frac{1}{3} + .05\frac{1}{3}} = 62.5\%.$$

Problem 3. *What is the probability that the following system works if each unit fails independently with probability p (see Figure 1.5)?*

If each unit fails independently with probability p, then each unit works independently with probability $1-p$. Suppose from left to right, top to bottom A, B, C, D and E are the events that a unit works. Then, since the units are independent, the probability that the whole system works is

$$\begin{aligned}
P((A\cap B)\cup C\cup (D\cap E)) &= P(A\cap B)+P(C)+P(D\cap E)-P((A\cap B)\cap C)-P(C\cap (D\cap E))\\
&\quad -P((A\cap B)\cap (D\cap E))+P((A\cap B)\cap C\cap (D\cap E))\\
&=(1-p)^2+(1-p)+(1-p)^2-(1-p)^3-(1-p)^3-(1-p)^4+(1-p)^5\\
&=p^5-4p^4-4p^3+1.
\end{aligned}$$

Problem 4. *This problem introduces some aspects of a simple genetic model. Assume that genes in an organism occur in pairs and that each member of the pair can be either of the types a or A. The possible genotypes of an organism are then AA, Aa and aa (Aa and aA are equivalent). When two organisms mate, each independently contributes one of its two genes; either one of the pair is transmitted with probability .5.*
(a) Suppose that the genotypes of the parents are AA and Aa. Find the possible genotypes of the their offspring and the corresponding probabilities.
(b) Suppose that the probabilities of the genotypes AA, Aa and aa are p, $2q$ and r respectively, in the first generation. Find the probabilities in the second and third generations, and show that theses are the same. This result is called the Hardy-Weinberg Law.
(c) Compute the probabilities for the second and third generations as in part (b) but under the additional assumption that the probabilities that an individual of type AA, Aa or aa survives to mate are u, v and w, respectively.

Proof. (a) The possible genotypes are AA and Aa since one parent has genotype AA, so the probability that the offspring receives A is 1. There is a .5 chance the other parent contributes a, and these events are independent so there is a .5 chance of AA and a .5 chance of Aa.

(b) For each of the possible genotypes in the first generation, there is a .5 chance of a parent contributing one of the two types. Thus, the probability that an offspring has a type A is $p+.5\cdot 2q+0r=p+q$. Likewise the probability that it receives a is $q+r$. Thus the probability of genotype AA is $p_1=(p+q)^2$, Aa is $q_1=(p+q)(q+r)$ and aa is $r_1=(q+r)^2$.

For the third generation we do the exact same calculation but with p_1, q_1 and r_1 in place of p, q and r respectively. Thus

$$\begin{aligned}
p_2 &= (p_1+q_1)^2\\
&= p_1^2+2p_1q_1+q_2^2\\
&=(p+q)^4+2(p+q)^2(p+q)(q+r)+((p+q)(q+r))^2\\
&=(p+q)^2((p+q)^2+2(p+q)(q+r)+(q+r)^2)\\
&=(p+q)^2\\
&=p_1.
\end{aligned}$$

Here we've used the fact that $p+2q+r=1$ so

$$(p+2q+r)^2=((p+q)+(q+r))^2=(p+q)^2+2(p+q)(q+r)+(q+r)^2=1$$

as well. Similarly,

$$\begin{aligned}
q_2 &= (p_1+q_1)(q_1+r_1) \\
&= p_1q_1 + p_1r_1 + q_1r_1 + q_1^2 \\
&= (p+q)^2(p+q)(q+r) + (p+q)^2(q+r)^2 + (q+r)^2(p+q)(q+r) + ((p+q)(q+r))^2 \\
&= (p+q)(q+r)((p+q)^2 + 2(p+q)(q+r) + (q+r)^2) \\
&= (p+q)(q+r) \\
&= q_1
\end{aligned}$$

and

$$\begin{aligned}
r_2 &= (q_1+r_1)^2 \\
&= q_1^2 + 2q_1r_1 + r_1^2 \\
&= ((p+q)(q+r))^2 + 2(p+q)(q+r)(q+r)^2 + (q+r)^4 \\
&= (q+r)^2((p+q)^2 + 2(p+q)(q+r) + (q+r)^2) \\
&= (q+r)^2 \\
&= r_1.
\end{aligned}$$

(c) Now the probability that an offspring has an A gene is $pu + .5 \cdot 2qv + 0rw = pu + qv$. Likewise, the probability that an offspring has an a gene is $qv + rw$. Thus $p_1 = (pu+qv)^2$, $q_1 = (pu+qv)(qv+rw)$ and $r_1 = (qv+rw)^2$. To find the third generation probabilities we do the same calculation but with p_1, q_1 and r_1 in place of p, q and r respectively. Thus

$$\begin{aligned}
p_2 &= (p_1u + q_1v)^2 \\
&= (p_1u)^2 + 2p_1q_1uv + (q_1v)^2 \\
&= (pu+qv)^4u^2 + 2(pu+qv)^2(pu+qv)(qv+rw)uv + ((pu+qv)(qv+rw))^2v^2 \\
&= (pu+qv)^2((pu^2+quv)^2 + 2(pu+qv)(qv+rw)uv + (qv^2+rvw)^2),
\end{aligned}$$

$$\begin{aligned}
q_2 &= (p_1u + q_1v)(q_1v + r_1w) \\
&= p_1q_1uv + p_1r_1uw + q_1r_1vw + q_1^2v^2 \\
&= (pu+qv)^2(pu+qv)(qv+rw)uv + (pu+qv)^2(qv+rw)^2uw \\
&\qquad + (pu+qv)(qv+rw)(qv+rw)^2vw + ((pu+qv)(qv+rw))^2v^2 \\
&= (pu+qv)(qv+rw)((pu+qv)^2uv + (pu+qv)(qv+rw)uw + (pu+qv)(qv+rw)v^2 + (qv+rw)^2vw)
\end{aligned}$$

and

$$\begin{aligned}
r_2 &= (q_1v + r_1w)^2 \\
&= q_1^2v^2 + 2q_1r_1vw + r_1^2w^2 \\
&= ((pu+qv)(qv+rw))^2v^2 + 2(pu+qv)(qv+rw)(qv+rw)^2vw + (qv+rw)^4w^2 \\
&= (qv+rw)^2((puv+qv^2)^2 + 2(pu+qv)(qv+rw)vw + (qvw+rw^2)^2
\end{aligned}$$

$\square$

Problem 5. *If a parent has genotype Aa, he transmits either A or a to an offspring (each with a $\frac{1}{2}$ chance). The gene he transmits to one offspring is independent of the one he transmits to another. Consider a parent with three children and the following evens: $A = \{$children 1 and 2 have the same gene$\}$, $B = \{$children 1 and 3 have the same gene$\}$, $C = \{$childrend 2 and 3 have the same gene$\}$. Show that these events are pairwise independent but not mutually independent.*

Proof. First note that $P(A) = P(B) = P(C) = \frac{1}{2}$ since there's a $\frac{1}{2}$ chance of receiving either gene. Consider $P(A \cap B)$. This is the probability that children 1 and 2 have the same gene and children 1 and 3 have the same gene. This must mean all three children have the same gene, and the probability of this is $\frac{2}{8} = \frac{1}{4} = P(A)P(B)$. A similar argument holds for $P(A \cap C)$ and $P(B \cap C)$. So we see that all three events are pairwise independent.

But then note that $P(A \cap B \cap C)$ is the probability that children 1 and 2 and children 1 and 3 and children 2 and 3 all have the same gene. Thus, this is again the probability that all three children have the same gene so $P(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8} = P(A)P(B)P(C)$. So these events are not mutually independent. $\square$

Problem 6. *Which is more likely: 9 hears in 10 tosses of a pair coin or 18 heads in 20 tosses?*

The first probability is

$$\binom{10}{9}\left(\frac{1}{2}\right)^9\left(\frac{1}{2}\right) = \frac{5}{512} \approx 9.8 \times 10^{-3}$$

and the second is

$$\binom{20}{18}\left(\frac{1}{2}\right)^{18}\left(\frac{1}{2}\right)^2 = \frac{95}{524288} \approx 1.8 \times 10^{-4}.$$

Thus the first event is more likely.

Problem 7. *Two boys play basketball in the following way. They take turns shooting and stop when a basket is made. Player A goes first and has probability p_1 of making a basket on any throw. Player B, who shoots second, has probability p_2 of making a basket. The outcomes of the successive trials are assumed to be independent.*
(a) Find the frequency function for the total number of attempts.
(b) What is the probability that player A wins?

(a) Since the players alternate, the probabilities can be modeled as

$$p(k) = \begin{cases} p_1(1-p_1)^{\frac{k+1}{2}-1} & k \text{ odd} \\ p_2(1-p_2)^{\frac{k}{2}-1} & k \text{ even} \end{cases}.$$

(b) The probability that player A wins is given by taking a sum of the normal geometric distribution with p_1 for the odd entries and p_2 for the even entries. Thus

$$P(A) = \sum_{k \text{ odd}}^{\infty} p_1(1-p_1)^{k-1} + \sum_{k \text{ even}}^{\infty} (1-p_2)^k = \sum_{k=1}^{\infty} p_1(1-p_1)^{(2k-1)-1} + \sum_{k=1}^{\infty}(1-p_2)^{2k} = \frac{1}{2-p_1} - \frac{(p_2-1)^2}{(p_2-2)p_2}.$$

Problem 8. *Three identical fair coins are thrown simultaneously until all three show the same face. What is the probability that they are thrown more than three times?*

The probability that on any one toss all three show the same face is $2/8 = 1/4$. Using the geometric distribution we see that $p(X = 1) = (3/4)^0(1/4) = 1/4$, $p(X = 2) = (3/4)^1(1/4) = 3/16$ and $p(X = 3) = (3/4)^2(1/4) = 9/64$. Assuming that the probability that they eventually all show the same face is 1, the probability that it takes more than three times is

$$1 - \frac{1}{4} - \frac{3}{16} - \frac{9}{64} = \frac{27}{64}.$$