Week 3 Math Homework Solution:
Allman Ch 4.3 Ex. 1,3,4,6,7,9 and Allman Ch 6.2 Ex. 1,6,15,16

4.3. Conditional Probabilities

4.3.1. a. $\{(F, F), (F, M), (M, F), (M, M)\}$, all with probability .25
b. 3/4
c. 1/2
d. 2/3
e. 1
f. No. Knowledge that one child is female effects the likelihood that the youngest child is female, since $1/2 \neq 2/3$. Alternately, $3/4 \neq 1$ shows that knowledge that the youngest child is female affects the likelihood that one child is female.

4.3.2. a. $\mathcal{P}(E_{odd} \cap E_{\leq 2}) = \mathcal{P}(E_1) = \frac{1}{6} = \mathcal{P}(E_{odd})\mathcal{P}(E_{\leq 2}) = \frac{1}{2} \cdot \frac{1}{3}$
b. $\mathcal{P}(E_{odd} \cap E_{\leq 3}) = \mathcal{P}(E_1 \cup E_3) = \frac{1}{3} \neq \mathcal{P}(E_{odd})\mathcal{P}(E_{\leq 3}) = \frac{1}{2} \cdot \frac{1}{2}$
c. $E_{\leq 2} = \{1, 2\}$ and $E'_{\leq 2} = \{3, 4, 5, 6\}$. Both $E_{\leq 2}$ and $E'_{\leq 2}$ contain equal numbers of evens and odds. However, $E_{\leq 3}$ contains two odds and one even, while $E'_{\leq 3}$ contains one odd and two evens. Knowledge of whether the roll is less than or equal to 3 effects the probability that the roll is even or odd.

4.3.3. a. Sensitivity is $\mathcal{P}(+ \text{ result} \mid \text{disease})$; Specificity is $\mathcal{P}(- \text{ result} \mid \text{no disease})$.
b. False positive: $\mathcal{P}(+ \text{ result} \mid \text{no disease})$; False negative: $\mathcal{P}(- \text{ result} \mid \text{disease})$.
c. Sensitivity $= 22/30 = .7333$; Specificity $= 1739/1790 = .9715$.

4.3.4. a.

	Healthy Persons	Diseased Persons
Negative Result	98901	1
Positive Result	999	99

b. $\mathcal{P}(\text{Diseased} \mid +) = 99/(999 + 99) = .0902$. Thus only about 9 of every 100 individuals testing positive actually have the disease, despite the high specificity of the test.

4.3.6. a. The diagonal entries correspond to no mutation occurring. These are likely to be the largest, since point mutations are rare.
b. Transitions: entries $(1, 2)$, $(2, 1)$, $(3, 4)$, $(4, 3)$; Transversions: entries $(1, 3)$, $(1, 4)$, $(2, 3)$, $(2, 4)$, $(3, 1)$, $(3, 2)$, $(4, 1)$, $(4, 2)$. This table does not support the hypothesis that transitions are more common than transversions.

4.3.7. a. The distribution of bases in S_0 is estimated by $p_A = .225$, $p_G = .275$, $p_C = .275$, $p_T = .225$.
b. The distribution of bases in S_1 is estimated by $p_A = .225$, $p_G = .3$, $p_C = .275$, $p_T = .2$.

4.3.9. a. Since there is no relationship between the two sequences, knowing information about one should convey nothing about the other.
b. All the columns would be the same.

6.2. Probability Distributions in Genetics

6.2.1. $HHHTT$, $HHTHT$, $HTHHT$, $THHHT$, $HHTTH$, $HTHTH$, $THHTH$, $HTTHH$, $THTHH$, $TTHHH$; $\binom{5}{3} = \frac{5\cdot4\cdot3\cdot2\cdot1}{(3\cdot2\cdot1)(2\cdot1)} = \frac{5\cdot4}{2} = 10$.

6.2.6. a. $\binom{6}{4}(\frac{1}{2})^4(\frac{1}{2})^2 = 15\frac{1}{64} = \frac{15}{64} \approx .2344$

b. $\mathcal{P}$(exactly i boys in 6 children) $= \binom{6}{i}(\frac{1}{2})^6$, so for $i = 0, 1, 2, \ldots 6$, the values are: .0156, .0938, .2344, .3125, .2344, .0938, .0156.

$\mathcal{P}$(exactly i girls in 6 children) has exactly the same values.

c. The expected number of boys is $\sum_{i=0}^{6} i\mathcal{P}$(exactly i boys in 6 children) $= 0(.0156) + 1(.0938) + 2(.2344) + 3(.3125) + 4(.2344) + 5(.0938) + 6(.0156) = 3$. Alternately, for a binomial distribution, the expected value is $n \cdot p = 6 \cdot \frac{1}{2} = 3$.

d. $\mathcal{P}$(4 or more girls of 6 children) $= \mathcal{P}$(4 girls) $+ \mathcal{P}$(5 girls) $+ \mathcal{P}$(6 girls) $= .2344 + .0938 + .0156 = .3438$.

6.2.15. a. Recall from problem 6.1.10, that the probability an offspring is yellow is $2/3$. Then the probability 5 of 12 have normal coloring is $\binom{12}{5}(1/3)^5(2/3)^7 \approx .1908$.

b. $\sum_{i=10}^{12}\binom{12}{i}(2/3)^i(1/3)^{12-i} \approx .1811$

c. $\sum_{i=0}^{3}\binom{12}{i}(2/3)^i(1/3)^{12-i} \approx .0039$

6.2.16. a. Since the probability that any given child in the family will develop Huntington disease is $1/2$, the probability that none of 4 do is $\binom{4}{0}(1/2)^0(1/2)^4 = 1/16$.

b. The probability that at least one of the 4 develops the disease is $1 - \mathcal{P}(\text{none of } 4) = 1 - 1/16 = 15/16$.

c. The probability that 3 or more develop the disease is $\binom{4}{3}(1/2)^3(1/2)^1 + \binom{4}{4}(1/2)^4(1/2)^0 = 5/16$.