1.(a) $\mathbf{X}$ is a binomial random variable with parameters (10, 0.5) and mean 10•0.5 = 5.

(b) $P\{\mathbf{X} \geq 4\} = 1 - P\{\mathbf{X} \leq 3\} = 1 - 2^{-10}\left[1 + \binom{10}{1} + \binom{10}{2} + \binom{10}{3}\right] = 1 - \frac{176}{1024} = \frac{848}{1024} = \frac{53}{64}$.

$P\{\mathbf{X} \leq 5 \mid \mathbf{X} \geq 4\} = \frac{P\{4 \leq \mathbf{X} \leq 5\}}{P\{\mathbf{X} \geq 4\}} = 2^{-10}\left[\binom{10}{4} + \binom{10}{5}\right]/P\{\mathbf{X} \geq 4\} = \frac{462}{1024} \times \frac{1024}{848} = \frac{231}{424}$.

(c) $P\{\text{4th toss = Head} \mid \mathbf{X} = 4\}$

$= \frac{P\{\text{4th toss = Head } \textbf{and } \mathbf{X} = 4\}}{P\{\mathbf{X} = 4\}} = \frac{P\{\text{4th toss = Head } \textbf{and} \text{ 3 Heads in other 9 tosses}\}}{P\{\mathbf{X} = 4\}}$

$= \frac{P\{\text{4th toss = Head}\}P\{\text{3 Heads in other 9 tosses}\}}{P\{\mathbf{X} = 4\}}$ (by independence of tosses) $= \frac{(1/2)\binom{9}{3}2^{-9}}{\binom{10}{4}2^{-10}} = \frac{4}{10}$.

(d) For any arbitrary value of P{Heads} = p > 0, we have that $P\{\text{4th toss = Head} \mid \mathbf{X} = 4\}$

$= \frac{P\{\text{4th toss = Head}\}P\{\text{3 Heads in other 9 tosses}\}}{P\{\mathbf{X} = 4\}} = \frac{p\binom{9}{3}p^3(1-p)^6}{\binom{10}{4}p^4(1-p)^6} = \frac{4}{10}$. Thus, not knowing p does not disadvantage me; the probability is 4/10 regardless of the value of p. Now, a fair bet should be offering 3-to-2 odds, i.e. with a bet of \$1, you win \$1.50 roughly 40% of the time and lose \$1 roughly 60% of the time. A bookie who offered 2-to-1 odds would be losing \$0.20 per dollar bet and would soon be out of business, and perhaps wearing concrete overshoes as well! The fact that he knows the outcome of the 4th toss and yet is offering such great odds to induce you to bet that a Head occurred, leads to the suspicion that the only reason the bookie can afford to offer these odds is that the 4th toss resulted in a Tail, and thus is sure that he is not going to lose. I would **not** bet on a Head.

2.(a) Obviously United Airlines has better on-time performance at all five airports.

(b) $P(T|UC)P(C|U) = \frac{P(TUC)}{P(UC)} \times \frac{P(UC)}{P(U)} = \frac{P(TUC)}{P(U)}$ and similarly for the other terms. Hence, the right side of the given expression is $\frac{P(TUC) + P(TUL) + P(TUX) + P(TUD) + P(TUF)}{P(U)}$.

But, (TUC) (TUL) (TUX) (TUD) (TUF) is a partition of the set TU, and thus the numerator above is just P(TU), and the ratio is, by definition, P(T|U). Also,

P(T|W) = P(T|WC)P(C|W)+P(T|WL)P(L|W)+P(T|WX)P(X|W)+P(T|WD) P(D|W)+P(T|WF)P(F|W)

(c) Plugging and chugging, P(T|U) = 0.8655 < 0.896 = P(T|W). The reason for the discrepancy is that America West has most of its flights into sunny Phoenix where flights have a good chance of being on time. In contrast, United has only a few flights to Phoenix and has lots of flights into snowy Chicago (and foggy San Francisco.) Note that 0.69 of the 0.896 of P(T|W) comes from Phoenix, whereas United gets only 0.0475 of the 0.8655 of P(T|U) from Phoenix (where it performs the best).

The moral for the marketing departments is that America West should advertise itself as most-often-on-time, whereas United should advertise itself as most-often-on-time-where-you-wanna-go.

Comment: This problem is based on a real-life case. The University of California at Berkeley (of all places!) was accused by the EEOC of discriminating against women applying to graduate school. Each Department actually admitted a larger fraction of its women applicants than of its men applicants, but the overall statistics showed a smaller fraction of women applicants being admitted as compared to men applicants.

3. Let A denote the event that a baby survives delivery and B the event that it is delivered by C section. Then, we are given that P(A) = 0.98, P(A|B) = 0.96, and P(B) = 0.15. The theorem of total probability tells us that $P(A) = 0.98 = P(A|B)P(B) + P(A|B^c)P(B^c) = 0.96 \times 0.15 + P(A|B^c) \times 0.85$ $P(A|B^c) \approx 0.9835$.

Quick sanity check: $P(A|B) < P(A) < P(A|B^c)$, so we haven't made an obvious calculation error.

4. There are $\binom{11}{2} = \frac{11 \times 10}{1 \times 2} = 55$ pairs of letters that might fall off.

The 55 pairs can be classified into the pair OO, 5 *other* pairs from the set {H, O, O, N}, 4×7 = 28 pairs consisting of one letter from {H, O, O, N} and one from {C, A, T, T, A, G, A}, 3 pairs AA and 1 pair TT, and 17 *other* pairs from {C, A, T, T, A, G, A}.

If OO fell off, the sign will always read correctly.

If one of 5 other pairs from {H, O, O, N} fell off, the letters will always look right side up, but will be interchanged with probability 1/2.

For the 28 pairs consisting of one letter from {H, O, O, N} and one from {C, A, T, T, A, G, A}, the restored sign is correct with probability 1/4, incorrect but readable with probability 1/4, and one letter

appears inverted with probability 1/2.
For the AA and TT pairs, the sign is correct with probability 1/4, has one letter inverted with probability 1/2 and has both letters inverted with probability 1/4.
The remaining 17 pairs are put back correctly with probability 1/8, interchanged but upright with probability 1/8, have one letter upside down with probability 1/2, and have both letters upside down with probability 1/4. This gives

(a) P(sign reads CHATTANOOGA) = $\left[1 + 5\times\frac{1}{2} + 28\times\frac{1}{4} + 4\times\frac{1}{4} + 17\text{–}\frac{1}{8}\right]\times\frac{1}{55} = \frac{109}{440} \approx 25\%$. Not Bad!

(b) P(sign readable but not CHATTANOOGA) = $\left[5\times\frac{1}{2} + 28\times\frac{1}{4} + 17\times\frac{1}{8}\right]\times\frac{1}{55} = \frac{93}{440}$

(c) P(one letter seems to be upside down) = $\left[28\times\frac{1}{2} + 4\times\frac{1}{2} + 17\times\frac{1}{2}\right]\times\frac{1}{55} = \frac{196}{440} = \frac{49}{110}$

(d) P(two letters seem to be upside down) = $\left[4\times\frac{1}{4} + 17\times\frac{1}{4}\right]\times\frac{1}{55} = \frac{42}{440} = \frac{21}{220}$

Sanity check: The sum of the probabilities is (109+93+196+42)/440 = 1.

(e) All letters seem to be right side up includes the case when the sign reads CHATTANOOGA! Now,

P(at least one vowel|all seem to be right side up) = $\dfrac{\text{P(all seem to be right side up} \cap \text{ at least one vowel)}}{\text{P(all seem to be right side up)}}$

We know from (a) and (b) that the denominator is (109+93)/440 = 101/220.
Now, our list above can be further classified into
1 pair OO and 2 pairs each of the forms HO and NO (for these the letters always appear to be right side up), 14 of the form O and one from {CATTAGA} and 3 each of the form HA and NA (for these, the probability that the letters seem right side up is 1/2),
3 of the form AA, and 3 each of the form CA and GA, and 6 of the form TA (for these the probability that tbe letters seem to be right side up is 1/4). All other pairs {from the set CHTTNG} do not include a vowel.
Check: P(at least one vowel) = 1 – P(no vowel) = 1 – [(6×5)/(1×2)]/55 = 40/55
= (1 + 2 + 2 + 14 + 3 + 3 + 3 + 3 + 3 + 6)/55 so we got 'em all!
From all this, we can readily compute P(all seem to be right side up ∩ at least one vowel)
= [(1+2+2)×1 + (14 + 3 + 3)×1/2 + (3 + 3 + 3 + 6)×(1/4)]/55 = 75/220.

Hence, P(at least one vowel | all seem to be right side up) = $\frac{75}{220} / \frac{101}{220} = \frac{75}{101}$.

(f) The designated driver can correctly identify the letters that fell down if and only if the driver can see that two letters are obviously switched (possibly being turned upside down) or if two letters both appear to be upside down but in their correct places. Note that these are disjoint possibilities. For the various pairs of different types, we see that if OO fell down, the sign will still read CHATTANOOGA and the driver will not be able to identify with certainty which letters fell down.. If one of the 5 other pairs from {H, O, O, N} fell down, the driver can identify them if and only if they have been switched (probability 1/2); whether they have been flipped over doesn't matter. For the 28 pairs consisting of one letter from {H, O, O, N} and one from {C, A, T, T, A, G, A}, the two letters must be swapped in order to be identifiable (probability 1/2), while for the 3 + 1 pairs AA and TT, both letters must be upside down to be identifiable with certainty. Finally, the remaining 17 *other* pairs from {C, A, T, T, A, G, A}, either both letters must be in the wrong place (probability 1/2) or they must be in the right place and both upside down (probability 1/8).

Hence, P{driver can identify correctly} = $\left[\frac{1}{55}\times 0 + \frac{5}{55}\times\frac{1}{2} + \frac{28}{55}\times\frac{1}{2} + \frac{4}{55}\times\frac{1}{4} + \frac{17}{55}\times\frac{5}{8}\right] = \frac{225}{440} = \frac{45}{88} > 50\%$!!

(g) Since all pairs are equally likely to fall down, the Bayes' decision is the same as the maximum-likelihood decision. Since P(CHATTANOOGA | OO fell down) = 1 and P(CHATTANOOGA | XY fell down) < 1 for any other choices of X and Y, the Bayes' (and maximum likelihood) decision favors OO falling down whenever the sign reads CHATTANOOGA. From part (a), the conditional probability that the decision is correct when the sign reads CHATTANOOGA is P(OO | CHATTANOOGA)
= P(CHATTANOOGA | OO)P(OO)/P(CHATTANOOGA) = 1×(1/55)/(109/440) = 8/109.

5.(a) Since the pitcher can only throw fastballs, curveballs or sliders, P(F) + P(C) + P(S) = 1. Also, P(C) = 2P(F) so we conclude that P(S) = 1 – 3P(F). Now, 1/4 = P(H|C)P(C) + P(H|F)P(F) + P(H|S)P(S)
= P(H|C)2P(F)+ P(H|F)P(F) + P(H|S)(1 – 3P(F)) = P(F)[2/4 + 2/5 –3/6] which gives P(F) = 5/24,
P(C) = 10/24 = 5/12, and P(S) = 9/24 = 3/8.

(b) The likelihood of a hit is P(H|F) = 2/5 or P(H|C) = 1/4, or P(H|S) = 1/6 depending on which hypothesis is true. Since P(H|F) = 2/5 is the largest, the maximum-likelihood decision is that it was a fastball.

(c) Now we compare the joint probabilities P(H|F)P(F) = (2/5)(5/24) = 2/24, P(H|C)P(C) = (1/4)(10/24) = 2.5/24, and P(H|S)P(S) = (1/6)(9/24) = 1.5/24 and get the Bayesian decision that it was a curveball. Note that it is not necessary to find P(F|H), P(C|H), and P(S|H) explicitly; the joint probabilities suffice.

6.(a) The likelihood matrix is as shown.

Hypothesis	**X** = 3	**X** = 6	**X** = 9	**X** = 12
H_0: excellent student	0.02	0.08	0.15	**0.75**
H_1: good student	0.1	0.15	**0.6**	0.15
H_2: average student	**0.2**	**0.65**	0.1	0.05

Looking in each column, it is obvious that observing {**X** = 12} should lead to classification as an excellent student, observing {**X** = 9} should lead to classification as a good student, while observing {**X** = 6} or {**X** = 3} should lead to classification as an average student. The corresponding entries are in boldface.

(b) An execellent student who gets a B is labeled as good. The probability of this occurring is 0.15. Similarly, such a student has 0.08 + 0.02 = 0.1 probability of being labeled average. The probability that an average student is classified as being above average is 0.1 + 0.05 = 0.15.

(c) The joint probability matrix is found by multiplying each row of the likelihood matrix by the probabilities of the hypothesis. This gives the matrix shown below.

Hypothesis	**X** = 3	**X** = 6	**X** = 9	**X** = 12	Row sums
H_0: excellent student	0.004	0.0160	0.030	**0.15**	0.2
H_1: good student	0.055	0.0825	**0.330**	0.0825	0.55
H_2: average student	**0.050**	**0.1625**	0.025	0.0125	0.25
Column sums	0.109	0.261	0.385	0.2450	1.00

Note that the row sums are the probabilities of the hypotheses while the column sums are the probabilities of the various types of performance. The maximum-likelihood decision rule is still shown by the boldface entries. Notice that P{correct decision} for Professor Max L. Hood is just the sum of the boldface entries which is 0.15 + 0.33 + 0.1625 + 0.05 = 0.6925. Hence, P{mis-classification} = 1 – 0.6925 = 0.3075. Hey, correct classification in better than 2 out of 3 cases ain't bad!

(d) Bayes' rule is as shown below.

Hypothesis	**X** = 3	**X** = 6	**X** = 9	**X** = 12	Row sums
H_0: excellent student	0.004	0.0160	0.030	**0.15**	0.2
H_1: good student	**0.055**	0.0825	**0.330**	0.0825	0.55
H_2: average student	0.050	**0.1625**	0.025	0.0125	0.25
Column sums	0.109	0.261	0.385	0.2450	1.00

Note that students with D's are being classified as good (i.e. better than average), while students with C's are being classified as average. Holy capricious grading complaint, Batman!
Note that Prof. (Joan) Baez classifies students correctly with slightly larger probability 0.6975 (and hence slightly smaller error probability 0.3025) than does Prof. Hood, but Prof. Baez spends 16.25% of her time defending her grading practices before various committees!

(e) Now the joint probability matrix looks like this and all students are classified as excellent regardless of their exam performance. As a result, Prof. Garrison Keillor no longer gives exams — they are a waste of everybody's time! All 5% of the good students are mis-classified as excellent, and the error probability is thus 0.05. Note that the maximum-likelihood decision rule has error probability 1–0.7125 = 0.2875.

Hypothesis	**X** = 3	**X** = 6	**X** = 9	**X** = 12	Row sums
H_0: excellent student	**0.019**	**0.076**	**0.1425**	**0.7125**	**0.95**
H_1: good student	0.005	0.0075	0.03	0.0075	0.05
H_2: average student	0	0	0	0	0
Column sums	0.024	0.0835	0.1725	0.72	1.00

7.(a) You don't *have* to buy the salesman's machine. In fact, your attitude is essentially "You claim your machine is better. Prove it!" Since the alternative hypothesis H_1 is usually taken to be the one that is to be proved, the two hypotheses should be taken to be

H_0: The new machine has P(defect) = 0.1 = p_0

and H_1: The new machine has P(defect) = 0.05 = p_1.

(b) Let **X** denote the number of defects observed on the test run of 100 chips. For i = 0, 1, if hypothesis H_i is true, then **X** is a binomial random variable with parameters (100, p_i). If **X** = 7, the likelihood ratio has value

$$\Lambda(7) = \frac{P\{\mathbf{X} = 7\} \text{ assuming } H_1 \text{ is true}}{P\{\mathbf{X} = 7\} \text{ assuming } H_0 \text{ is true}} = \frac{\binom{100}{7}(p_1)^7(1-p_1)^{93}}{\binom{100}{7}(p_0)^7(1-p_0)^{93}} = \frac{(0.05)^7(0.95)^{93}}{(0.1)^7(0.90)^{93}} = 1.193.$$ Hence,

the maximum-likelihood decision is that H_1 is the true hypothesis, and the salesman's claim should be believed.

(c) If **X** = N is observed, the likelihood ratio has value $\Lambda(N) = \frac{(p_1)^N(1-p_1)^{100-N}}{(p_0)^N(1-p_0)^{100-N}}$. Thus, the *log* likelihood ratio is $\ln \Lambda(N) = N\bullet\ln\left(\frac{p_1}{p_0}\right) + (100-N)\bullet\ln\left(\frac{1-p_1}{1-p_0}\right) = N\bullet\ln\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right) + 100\bullet\ln\left(\frac{1-p_1}{1-p_0}\right) = -0.7472N + 5.4067$ which is negative for $N > 7.23\ldots$ The maximum-likelihood rule is thus: "If the number of defectives in a run of 100 *exceeds* 7, do not buy the new machine." Check: $\Lambda(8) = 0.5649\ldots$ so H_1 is rejected.

(d) H_1 is falsely accepted if 7 or fewer defectives are observed when H_0 is the true hypothesis. But, when H_0 is the true hypothesis, **X** is a binomial random variable with parameters (100,0.1). Thus, $P_{FA} = P\{\mathbf{X} \le 7\} = 0.2060\ldots$ is readily computed. On the other hand, H_1 is correctly accepted if 7 or fewer defectives are observed when H_1 is the true hypothesis. But, when H_1 is the true hypothesis, **X** is a binomial random variable with parameters (100,0.05). Thus, $P_{MD} = 1 - P\{\mathbf{X} \le 7\} = 1 - 0.872\ldots = 0.12796\ldots$ is readily computed. Note that P_{FA} and P_{MD} are different.

(e) The general analysis of part (c) applies to the case of a test run of 1000 chips if we replace 100 by 1000 to get that $\Lambda(N) = -0.7472N + 54.067$ which says that the new machine should be bought only if the number of defectives is 72 or fewer. Note that *roughly* 50 and 100 defective chips will be produced under the two hypotheses. We can compute $P\{\mathbf{X} \le 72\}$ under the two hypotheses to get that $P_{FA} = 0.00127\ldots$ and $P_{MD} = 0.0010\ldots$ which are *much* smaller than with a test run of 100 chips. This is pretty much as can be expected: the more extensive the testing, the smaller the chances of making a mistake. The problem here quantifies how *much* smaller the chances are.