ECE 313: Problem Set 5
Confidence Interval, Bayes formula, Hypothesis testing

Due: Wednesday, February 20 at 6 p.m.
Reading: *ECE 313 Course Notes,* Sections 2.9-2.11.

1. **[Confidence interval]**
A communication system designer needs to simulate a communication link. The link is being designed for binary transmission, i.e., bits $b_k \in \{0;1\}$ are transmitted, and bits $\hat{b}_k \in \{0,1\}$ are received in the kth bit-period. Noise and other sources of channel impairments result in a bit-errors, i.e., the event $\hat{b}_k \neq b_k$. The bit errors are assumed to occur independently from one bit-period to the next with a probability of error or bit error-rate (BER) of p_e. The communication link is being designed to meet the specifications set by an international standards committee such as ITU, IEEE, or ETSI. The standards document specifies that $p_e \leq 10^{-4}$. The designer estimates p_e by running n bits through her simulation model, counts the number of errors E by comparing the transmitted bit b_k with recovered bit $\hat{b}_k$ and obtains a BER estimate $\hat{p}_e = \frac{E}{n}$ where E is the error count, i.e., the total number of bits in error in a stream of n bits. The designer wants to impress her manager and wishes to report that her design meets the BER specications with a 10% tolerance around the 10^{-4} specified value.
   (a) Reason that E is a binomial random variable with parameters (n, p_e).
   (b) How many bits will she need to simulate the communication link over in order to achieve a condence level of 99%?
   (c) The designer has a desk-top with a dual core CPU running at a clock frequency of 4GHz. The complexity of her simulation program and the associated compiler are such that each bit period takes 40 clock cycles (using both cores) to simulate. How much time in minutes will it take her to simulate all n bits, where n is the answer to part (a).
   (d) Realizing the simulation time in part (c) is too long, she settles upon a strategy to run simulations with fewer bits initially until the design has matured and then run a long one. She decides to give 10 minutes per run. How many bits can she simulate in this time and what is her level of confidence in her design assuming that the tolerance remains fixed at ± 10 %?

2. **[Bayes Formula]**
Dilbert has 3 coins in his pocket, 2 of which are fair coins while the third is a biased coin with $P(H) = p \neq \frac{1}{2}$. The probability that a coin chosen at random from his pocket will land Tails is $\frac{7}{12}$.
   (a) What is the value of p?
   (b) Dilbert picks two coins at random from his pocket, tosses each coin once, and observes a Head and a Tail. What is the conditional probability that both coins are fair?

3. **[Law of Total Probability]**
Tom goes to the casino with \$5. The game consists of Tom tossing a fair coin. If the result of a toss is Heads, the casino pays Tom \$1 while if the result is Tails, Tom pays the casino \$1. Thus, after each coin toss, the amount of money that Tom has — hereinafter referred to as his *wealth* — either increases by \$1 or decreases by \$1. The (independent) coin tosses continue until *one* of the following two events occurs: **(i)** Tom's *wealth* increases to \$8 *or* **(ii)** Tom's *wealth* decreases to \$0. At this point, Tom goes home (with *wealth* \$8 or \$0 as the case may be).
   (a) What is the probability that Tom goes home after three coin tosses?
   (b) What is the probability that Tom goes home after five coin tosses?
   (c) *Conditioned* on the event $A = \{\text{Tom tosses the coin at least 5 times}\}$, what is the conditional pmf of $\mathbb{X}$, Tom's *wealth* after 5 tosses have occurred? Note that the increase or decrease in Tom's *wealth* due to the result of the fifth toss is included in $\mathbb{X}$.

4. **[Bayes Formula]**
Dissatisfied with the candidates proposed by Facespace, Harry turns to Mybook which secretly classifies one-thirds of its members as H-type ("Hot") and two-thirds as N-type ("Not"). Mybook presents a member selected at random from its subscribers to Harry who accepts a H-type candidate with probability $\frac{3}{4}$ and an N-type candidate with probability $\frac{3}{40}$.

   (a) What is the probability that Harry accepts the proposed candidate?

   (b) What is the probability that an accepted candidate is a H-type?

5. **[The (in)famous Monty Hall problem]**
Monty Hall, the host of the yester-year TV game show "Let's Make A Deal" shows you three curtains. One curtain conceals a car, while the other two conceal goats. All three curtains are equally likely to conceal the car. He offers you the following "deal": pick a curtain, and you can have whatever is behind it. When you pick a curtain, instead of giving you your just deserts, Monty (who knows where the car is) opens one of the remaining curtains to show you that there is a goat behind it, and offers the following "new, improved deal": you can either stick with your original choice, or switch to the remaining (unopened) curtain. Amidst the deafening roars of "Stand pat!" and "Switch, you idiot!" from the crowd, Monty points out that previously your chances of winning were 1/3. Now, since you know that the car is behind one of the two unopened curtains, your chances of winning have increased to 1/2, and thus the new improved deal is indeed better.

   (a) Let A denote the event that your first choice of door has the car behind it. What is $P(A)$?

   Let B denote the event that your second choice of door has the car behind it.

   (b) *Your strategy is to always stay put* and so your second choice is the same as your first choice. What is $P(B \mid A)$ in this case? What is $P(B \mid A^c)$? Use these results to find $P(B)$ for the stay-put strategy.

   (c) *Your strategy is to always switch* and so your second choice is the other unopened door. What is $P(B \mid A)$ in this case? What is $P(B \mid A^c)$? Use these results to find $P(B)$ for the always-switch strategy.

   (d) *Your strategy is to pick randomly* and so your second choice is equally likely to be either unopened door. What is $P(B \mid A)$ in this case? What is $P(B \mid A^c)$? Use these results to find $P(B)$ for the pick-randomly strategy. Is Monty correct in asserting that if you choose randomly between the two unopened curtains, you have a probability of winning of 1/2?

   Note: The rules of the game of parts (a)-(d) are that Monty, who knows which curtain conceals the car, always opens one of the two unchosen curtains and he always offers the "new improved deal," that is, he never opens a curtain to reveal the prize (saying "Oops, you lose; return to your seat.").

6. **[Hypothesis Testing]**
Two different tests A and B are available for detecting the presence of HIV virus in human blood. Test A is *more sensitive* than Test B: when the HIV virus is present (event $H+$), Test A gives a positive result (event $T+$) with probability 0.999 whereas Test B gives a positive result with probability only 0.99. On the other hand, Test B is *more specific* than Test A: when the HIV virus is *not* present (event $H-$), Test B gives positive result $T+$ with probability only 0.001 whereas Test A gives positive result $T+$ with larger probability 0.01. For obvious reasons, such positive results are called *false positives*. False positive results are a huge nuisance in medical testing (and testing in general) because they can lead to the patient receiving unnecessary treatment. In HIV testing, false positive results can also lead to undesirable familial, social, and societal consequences.. *False negative* results (event $T-$ occurring when the virus is actually present) also have undesirable consequences in that a patient may not get treatment for the disease, and may also unknowingly infect others.

   The data above can be summarized as follows:
   Test A: $P(T+|H+) = 0.999$ $\quad P(T+|H-) = 0.01.$
   Test B: $P(T+|H+) = 0.99$ $\quad P(T+|H-) = 0.001.$

(a) Suppose that 2% of the population has HIV virus in the bloodstream, that is, $P(H+) = 0.02$. Find the value of $P(T+)$ for each test.

(b) For each test, compute $P(H + |T+)$.

(c) For each test, compute $P(H + |T-)$.

7. **[Hypothesis testing]**
If H_0 is the true hypothesis, the random variable $\mathbb{X}$ takes on values 0, 1, 2, and 3 with probabilities 0.1, 0.2, 0.3, and 0.4 respectively. If H_1 is the true hypothesis, the random variable $\mathbb{X}$ takes on values 0, 1, 2, and 3 with probabilities 0.4, 0.3, 0.2, and 0.1 respectively.

(a) Find the likelihood matrix L and indicate the *maximum-likelihood decision rule* by shading the appropriate entries in L. What is the false-alarm probability P_{FA} and what is the missed-detection probability P_{MD} for the maximum-likelihood decision rule?

(b) Now suppose that the hypotheses have *a priori* probabilities $\pi_0 = 0.7$ and $\pi_1 = 0.3$. Use the law of total probability to find the average error probability of the maximum-likelihood decision rule that you found in part (a).

(c) Use the *a priori* probabilities given in part (b) to find the joint probability matrix J and indicate on it the Bayesian decision rule, which is also known as the minimum-error-probability (MEP) or maximum *a posteriori* probability (MAP) decision rule. What is the average error probability of the Bayesian decision rule? Is it smaller or larger than the average error probability of the maximum-likelihood decision rule? In the latter case, provide a brief explanation as to why the minimum-error-probability rule has a larger average error probability than another rule.

(d) Show that in each of the two cases $i = 0$ and for $i = 1$, it is true that if $\pi_i > 0.8$, then the Bayesian decision rule always decides that H_i is the true hypothesis, no matter what the value of $\mathbb{X}$ is. Hint: Remember that $\pi_{1-i} = 1 - \pi_i$.