CSE 312A: Foundations of Computing II
Assignment #8
March 1
due: March 7, at noon.

Instructions:

Answers: When asked for a short answer (such as a single number), also *show and explain your work* briefly. Simplify your final formula algebraically as much as possible, without using your calculator. Then, if the answer is a number rather than a function of some variables, use a calculator to evaluate it and provide the number. For example, for counting problems, your answer might look like this:

Answer: $\binom{5}{2} - \binom{4}{2} = 4$.

Explanation: There are $\binom{5}{2}$ ways to select 2 fingers out of the 5, and $\binom{4}{2}$ of them do not involve the thumb.

Solutions that do not show enough work may not get full credit.

Turn-in: Do not write your name on your pages (your Gradescope account will identify you to us) and do not include a copy of the exercise's question in what you turn in. You must use Gradescope to upload your homework solutions. You will submit a single PDF file containing your solutions to all the exercises in the homework. Each numbered homework question must be answered on its own page (or pages). You must follow the Gradescope prompts that have you link exercise numbers to your pages. You may typeset your solutions on a computer (see here for tutorials and templates) or you can handwrite them, take a picture of (or scan) each handwritten page, and convert the pictures into a single PDF file. You are responsible for making sure that your solution is easily readable, and submitted on time.

CDF of normal Some of the problems require looking up the CDF of the normal distribution. You can use this table to find the values you need.

1. In class we discussed the distributed load balancing problem. We showed that if n jobs are randomly distributed to k servers, then the probability that there is a server that gets more than $n/k + 3\sqrt{n\ln(k)/k}$ jobs is at most $1/k^2$. Use the Chernoff bound to show that the probability that there is a server that gets fewer than $n/k - 2\sqrt{n\ln(k)/k}$ jobs is at most $1/k$.

2. A certain city is experiencing a big crime wave. The city decides that it needs to put its police officers out into the streets to bring back security. The city is conveniently arranged into a 100×100 grid of streets. Each street intersection can be identified by two integers (a, b) where $1 \leq a \leq 100$ and $1 \leq b \leq 100$. The city only has 1000 police officers, so it decides to send each police officer to a uniformly random grid location, independent of each other. The city wants to make sure that every 20×20 subgrid (corresponding to grid points (a, b) with $A \leq a \leq A + 19$ and $B \leq b \leq B + 19$ for valid A, B) gets more than 10 police officers.

   (a) Use the Chernoff bound to compute the probability that a single subgrid gets at most 10 police officers.

   (b) Use the union bound together with the result from above to calculate the probability that the city fails to meet its goal.

3. Professor Lazy decides to assign final grades in CSE 312 by ignoring all the work the students have done and instead using the following probabilistic method: each student independently will be assigned an A with probability θ, a B with probability 3θ, a C with probability $\frac{1}{2}$, and an F with probability $\frac{1}{2} - 4\theta$. When the quarter is over, you discover that only 2 students got an A, 10 got a B, 60 got a C, and 40 got an F. Find the maximum likelihood estimate for the parameter θ that Professor Lazy used. Give an exact answer as a simplified fraction.

4. (a) Let $x_1, x_2, \ldots, x_n$ be independent samples from a geometric distribution with unknown parameter p. What is the maximum likelihood estimator for p?

   (b) If the samples from the geometric distribution are 5, 4, 10, 2, 9, 5, 6, 13, 9, what is the maximum likelihood estimator for p? Give an exact answer as a simplified fraction.

5. Let $x_1, x_2, \ldots, x_n$ be independent samples from an exponential distribution with unknown parameter λ. What is the maximum likelihood estimator for λ? (If you do the previous exercise correctly, you should see a very natural relationship between the estimators.)

6. (a) Let $x_1, x_2, \ldots, x_n$ be independent samples from the continuous uniform distribution on $[0, \theta]$, where θ is the unknown parameter. What is the maximum likelihood estimator $\hat{\theta}$ for θ?

   Hint: It might help to roughly sketch out the shape of the likelihood function as a function of θ, that is, θ on the horizontal axis and likelihood L on the vertical axis. Start your sketch by asking what happens to L as θ goes to infinity. Then figure out the shape as you get closer to $\theta = 0$. From your sketch, at what value of θ do you maximize L?

   (b) In the remaining parts of this exercise you will compute the bias of the estimator $\hat{\theta}$. Begin by computing the cumulative distribution function for the random variable $\hat{\theta}$. Recall that this is simply the function $F(x) = P(\hat{\theta} < x)$. Focus first on the interval $0 \leq x \leq \theta$, but when you're done with that, don't forget to also define $F(x)$ on the rest of the real numbers.

   (c) From your answer to part (c), compute the probability density function $f(x)$ of $\hat{\theta}$.

   (d) From your answer to part (d), compute $E[\hat{\theta}]$. Is $\hat{\theta}$ an unbiased estimator of θ?