EECS 70 Discrete Mathematics and Probability Theory
Spring 2013 Anant Sahai

HW 13

Due April 29

1. The Rule of Huge Quantities

In all of the following questions assume that $X_1, \ldots, X_n$ are independent and identically distributed random variables. Also $A = (X_1 + \cdots + X_n)/n$ is the average of all X_is.

1. Assuming that each X_i is the outcome of a die roll (i.e. a random value between 1, 2, 3, 4, 5, 6), for $n = 100$, compute the bounds of an interval $[a,b]$ such that $A \in [a,b]$ with probability at least 90% Obviously you should try to find intervals that are as small as you can, not trivial things like $[1,6]$!
2. Again, assuming that each X_i is the outcome of a die roll, but this time with $n = 30$, compute a lower bound on the probability that $3 \leq A \leq 4$.
3. Using the same die rolls example, what is the minimum n for which you can guarantee that with probability at least 99% you have $3 \leq A \leq 4$.
4. Assume that $n = 12$, and each X_i shows the amount of money you spend in month i. Note that we are assuming $X_1, \ldots, X_{12}$ are independent. You know that $E[X_i] = 1500$, i.e. on average you spend 1500 dollars a month and that the standard deviation (the square root of the variance) of X_i is 500. Assuming that you have no income, how much should you have in your bank account at the beginning of the year to make sure that you won't have to borrow money from anyone until the end of the year with probability at least 95%?
5. You have found a notebook of ideas for startups. In order to implement an idea, you would need 50 thousand dollars, and after the implementation your idea works with probability p and generates 150 thousand dollars in revenue for you (so you would net 100 thousand dollars), or doesn't work with probability $1-p$ and your 50 thousand dollars is burnt. The notebook has 10 ideas in it and assume that you have enough money and you decide to implement all of them. You don't like your net return from implementing these ideas to be negative (i.e. you don't want to lose more money than you make), so you want the probability of this happening to be bounded by 5%. For what values of p, can you guarantee that the probability of losing money is at most 5%?
6. Assume that you are using an error-correcting code to protect against erasures (assume that there are only erasures and not errors). Your codes can recover the original message if you have at least 1000 symbols that are not erased. Assuming that each symbol gets erased with probability 0.8 and that you want to be able to recover the message with probability at least 99%, how many symbols should you send?

2. Sample mean and variance

Suppose we have n i.i.d. samples $X_1, X_2, \ldots, X_n$ drawn from some distribution. We wish to estimate the mean and variance of this distribution. The obvious choice for an estimate of the mean is $M = \frac{X_1 + \cdots + X_n}{n}$.

1. Prove that $E[M] = E[X_i]$.

2. Now a natural choice for estimating the variance is $V = \frac{(X_1-m)^2+(X_2-m)^2+\cdots+(X_n-m)^2}{n}$ where m is your estimate of the mean from the previous part. Calculate $E[V]$. Is $E[V] = E[X_i^2] - E[X_i]^2$? If not, how can you modify the estimate V such that $E[V] = E[X_i^2] - E[X_i]^2$?

3. Binary entropy

Let $X_i,\ 1 \le i \le n$ be a sequence of i.i.d. Bernoulli random variables with parameter p, i.e. $Pr(X_i = 1) = p$ and $Pr(X_i = 0) = 1 - p$.

1. Calculate $Pr(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$, where $x_i \in \{0,1\}$ for all $1 \le i \le n$. Note the answer is a function of $(x_1, x_2, \ldots, x_n)$. Let us call it $f(x_1, x_2, \ldots, x_n)$.
2. Define random variable Y_n as $Y_n = -\frac{1}{n}\log f(X_1, X_2, \ldots, X_n)$. What does Y_n tend to as n grows large? [Hint: Try to use the law of large number. Your answer should be a function of p.]
3. Use Stirling approximation to calculate $\lim_{n\to\infty} \frac{1}{n}\log \binom{n}{np}$. How is your answer related to part 2? Comment on what you see.

4. Chernoff

Let $X_1, \ldots, X_n$ be independent Bernoulli random variables that each take value 1 with probability p and 0 with probability $1 - p$. You have learned how to use Chebyshev's inequality to say things about the probability that the sum $S = X_1 + X_2 + \cdots + X_n$ deviates from its mean (pn). In this question you will derive another bound called Chernoff's inequality that is much stronger in most cases.

1. As an example to help you understand the setting better, assume that X_i is the outcome of a coin flip (that is $X_i = 1$ if the coin flip results in heads and otherwise $X_i = 0$). Then $p = 1/2$ and S is the number of heads you observe. Assume that $n = 100$ is the number of coin flips. The expected number of heads you see is $pn = 50$. Using a computer calculate the probability that $S \ge 80$ (note that since the probability is very small, you can't use simulations). Now using Chebyshev's inequality find an upper bound for this probability. Is your upper bound much larger than the value you computed?
2. Back to the general setting, prove that if $f : \{0,1\} \to \mathbb{R}$ is any function, then $f(X_1), \ldots, f(X_n)$ are independent. Hint: write down the definition of independence. If f takes the same value at 0 and 1 then everything should be obvious. It remains to prove it in the case where $f(0) \ne f(1)$.
3. Now if we fix a number t and let $f(x) = e^{tx}$, then $f(X_i) = e^{tX_i}$. Compute the expected value of $f(X_i) = e^{tX_i}$ and write it in terms of p and t.
4. The following is a famous inequality about real numbers: $1 + x \le e^x$. Plot $1 + x$ and e^x in the same figure and observe that the inequality holds. Another variant of the inequality (which can be derived by replacing x by $x - 1$) is the following: $x \le e^{x-1}$. Apply the latter inequality with x being the expected value you computed in the previous step in order to get an upper bound on $E[f(X_i)]$. (You don't need to prove either of these inequalities.)
5. Remembering that $f(X_1), \ldots, f(X_n)$ are all independent what is $E[f(X_1)f(X_2)\ldots f(X_n)]$ in terms of $E[f(X_1)], \ldots, E[f(X_n)]$? Use the upper bound you got from the previous step to get an upper bound on $E[f(X_1)f(X_2)\ldots f(X_n)]$. You should be able to express your answer in terms of p, n, and t. Now let $\mu = pn$ be the expected value of S. Re-express your upper bound in terms of μ and t (i.e. remove the occurrences of p and n and rewrite them in terms of μ).
6. Observe that $f(X_1)\ldots f(X_n) = e^{t(X_1+\cdots+X_n)} = e^{tS}$. Let us call e^{tS} the random variable Y. Does it always take positive values? Let's say we are interested in bounding the probability that $S \ge (1+\alpha)\mu$ where α is a non-negative number. Prove that $S \ge (1+\alpha)\mu$ is the same event as $Y \ge e^{t\mu(1+\alpha)}$. Use Markov's inequality on the latter event to derive an upper bound for $\Pr[S \ge (1+\alpha)\mu]$ in terms of μ, t, and α.

7. For different values of t you get different upper bounds for the probability that $S \geq (1+\alpha)\mu$. But of course all of them are giving you an upper bound on the same quantity. Therefore it is wiser to pick a t that minimizes the upper bound. This way you get the tightest upper bound you can using this method. Assuming that α is fixed, find the value t that minimizes your upper bound. For this value of t what is the actual upper bound? Your answer should only depend on α and μ. Hint: in order to minimize a positive expression you can instead minimize its ln. Then you can use familiar methods from calculus in order to minimize the expression.

8. Here we want to compare Chernoff's bound and the bound you can get from Chebyshev's inequality. Assume for simplicity that $p = 1/2$, so $\mu = n/2$.

   First compute Chernoff's bound for the probability of seeing at least 80 heads in 100 coin flips (the quantity you bounded in the first part). Compare your answer to that part and see which one is closer to the actual value.

   Now back to the setting with general n and α, write down the Chernoff bound as c^n where c is an expression that only contains α and not n. This shows that for a fixed value of α, Chernoff's bound decays exponentially in n. Now write down Chebyshev's inequality to bound $\Pr[|S-\mu| \geq \alpha\mu]$. Show that this is also a bound on $\Pr[S \geq (1+\alpha)\mu]$. Write down this bound as γn^{β} where γ and β are some numbers that do not depend on n. This shows that Chebyshev's inequality decays like n^{β}. In general an exponential decay (which you get from Chernoff's) is much faster than a polynomial decay (the one you get from Chebyshev's).

5. Binomial Distribution

In this question you will use Stirling's approximation to study the binomial distribution. Let us flip n independent coins (each showing heads with probability $1/2$) and count the number of times we see heads. Let S denote this random variable. We would like to study the distribution of S.

1. Calculate the exact probability $\Pr[S = k]$ for each $0 \leq k \leq n$.
2. Use Stirling's approximation $m! \simeq \sqrt{2\pi m}(m/e)^m$ and replace all of the factorials in the expression you found for $\Pr[S = k]$. Simplify your expression so that you don't have the constant e anymore.
3. Now assume that $k = tn$ where $0 \leq t \leq 1$ is a real number. Simplify the expression you got in the previous step and write it in terms of t and n. You should simplify enough that your expression looks like the following
$$\frac{1}{A\sqrt{n}}B^n$$
where A and B only depend on t and not n.
4. Plot the expression you have for $0 \leq t \leq 1$ for these choices of n: 10, 20, 100. Use a log-scale plot, i.e. instead of plotting $f(t)$, plot $\ln(f(t))$ where $f(t)$ was your approximation for $\Pr[S = tn]$. Do these graphs all look somewhat similar?
5. Plot the expression you obtained as a function of n for three fixed values of t, $t = 1/4$, $t = 1/3$, and $t = 1/2$ (you should have three plots in the same figure). Your plot should go from $n = 1$ to $n = 100$. Again plot using log-scale (use log-scale only for your approximation, i.e. the y-axis, not the x-axis).
6. Note that $\Pr[S \geq tn] \geq \Pr[S = tn]$. For $t > 1/2$ the previous question (Chernoff's inequality) gives you an upper bound on $\Pr[S \geq tn]$. Compute this upper bound in terms of t and n. This upper bound should obviously be greater than $\Pr[S = tn]$. Plot both the approximation for $\Pr[S = tn]$ you got using Stirling's formula, and the upper bound you got using Chernoff's in the log-scale for three different values of n: 10, 20, 100. Let t vary in the interval $[0.5, 1]$ in each plot.

6. Making a triangle

You are given a one meter-long stick. You choose two points $X < 1$ and $Y < 1$ randomly along the stick and cut the stick at those two points. What is the probability that you can make a triangle with the three pieces?

7. Your Own Problem

Write your own problem related to this week's material and solve it. You may still work in groups to brainstorm problems, but each student should submit a unique problem. What is the problem? How to formulate it? How to solve it? What is the solution?