This homework is entirely optional, due December 15, 2014, at 12:00 noon.

1. Expecting you to integrate by parts!

In Discussion 12W, we derived an alternative form for the expected value of a discrete non-negative random variable Y, $E[Y] = \sum_{i=1}^{n} P(Y \geq i)$, when Y takes on values from 1 to n. In this problem, we will derive the continuous analog of this expression. Throughout this problem, assume X is a *continuous* non-negative random variable with PDF $f_X(x)$.

(a) Write an expression for $P(X \geq x)$ in terms of $f_X(x)$. This is called the *complementary cumulative distribution function* of X, of the CDF of X. For this problem, we denote this as $\bar{F}_X(x)$. What is $\bar{F}_X(0)$? How about $\bar{F}_X(x)$ as $x \to \infty$?

(b) Use integration by parts on $E[X] = \int_0^\infty x\ f_X(x)\ dx$ to derive the expression in question. (*Hint:* What is the antiderivative of $f_X(x)$?)

2. Exponential world

If X is a non-negative and continuous random variable, we call it an exponential random variable with rate λ if and only if for each $t \geq 0$ we have $\Pr[X \geq t] = e^{-\lambda t}$.

(a) Show that an exponential random variable is memoryless. This means that $\Pr[X \geq t_1 + t_2 | X \geq t_2] = \Pr[X \geq t_1]$. Why does this property justify the use of an exponential random variable to approximately model the following: the length of the time it takes from now until the next customer arrives in a shop?

(b) Now consider a shop. Customers are arriving in the shop. We model them this way: the first customer draws a sample from the exponential distribution with rate 1 and arrives at the time specified by that. The second customer independently draws another sample from the exponential distribution and waits that much time after customer 1 before arriving at the shop, and so on.

Now assume that the random variable the first customer draws is X_1 and the one that the second customer draws is X_2. Therefore the first customer arrives at time X_1 and the second customer arrives at time $X_1 + X_2$.

The shop owner is a psychic and knows that in the next hour, there is exactly one customer arriving. So $X_1 < 1$ and $X_1 + X_2 > 1$. Conditioned on this we wish to figure the distribution of X_1.

Fix a tiny interval $[t, t+\varepsilon]$ for some $t \in [0,1]$ and some small $\varepsilon > 0$. What is $\Pr[X_1 \in [t, t+\varepsilon]]$?

(c) Conditioned on $X_1 \in [t, t+\varepsilon]$, the probability of the event $X_1 + X_2 > 1$ is roughly equivalent to the probability of $X_2 > 1 - t$ (this approximation becomes more and more accurate as $\varepsilon \to 0$). Use this to approximate $\Pr[X_1 + X_2 > 1 | X_1 \in [t, t+\varepsilon]]$.

(d) Now use the Bayes formula to compute $\Pr[X_1 \in [t, t+\varepsilon] | X_1 + X_2 > 1]$. Does the result depend on t? Argue that the distribution of X_1 conditioned on the fact that $X_1 + X_2 > 1$ and $X_1 < 1$ is uniform on $[0,1]$.

3. Exponential Distributions: Lightbulbs

A brand new lightbulb has just been installed in our classroom, and you know the life span of a lightbulb is exponentially distributed with a mean of 50 days (see Note 18).

(a) Suppose an electrician is scheduled to check on the lightbulb in 30 days and replace it if it is broken. What is the probability that the electrician will find the bulb broken?

(b) Suppose the electrician finds the bulb broken and replaces it with a new one. What is the probablity that the new bulb will last at least 30 days?

(c) Suppose the electrician finds the bulb in working condition and leaves. What is the probability that the bulb will last at least another 30 days?

4. Maximum Likelihood

Often, when you observe samples from a random distribution, you don't know the parameters which govern the distribution. You would like to estimate these parameters after having seen some number of samples from the distribution. This is similar to what we saw in polling-related questions where we were trying to estimate the true fraction p of some relevant subset of some population.

In this question, you are given the type of the distribution and n samples drawn from it. You are to determine the most likely parameters of the distribution which would give those samples.

For example, consider the case of n tosses of a biased coin which gives heads with probability p. You want to find the most likely p, given the number of heads we see from the coin tosses.

We define the likelihood function $l(p)$ as the likelihood of getting the results we observed, as a function of p (or more generally, all the parameters of the distribution).

In this case, if we observe k heads and $n-k$ tails from our tosses, $l(p) = p^k \cdot (1-p)^{n-k}$. The most likely value of p would be:

$$p' = \texttt{argmax}_p l(p)$$

We can do this by differentiating $l(p)$ with respect to p and setting it to 0. This would give us $p' = \frac{k}{n}$ as the only maximum, which is what we expect.

Now, do the same maximum likelihood estimate for the following. Note, this information might be useful: $\texttt{argmax}_p l(p) = \texttt{argmax}_p \log(l(p))$

(a) Consider the binomial distribution with parameters m, p given by the PMF:

$$P(i) = \binom{m}{i} p^i (1-p)^{m-i}$$

Here, assume you already know m. Given n samples $X_1 = x_1, \ldots, X_n = x_n$, what is the maximum likelihood estimate of p?

(b) Consider the geometric distribution with parameter p given by the PMF:

$$P(i) = (1-p)^{i-1} p \qquad i > 0$$

Given n samples $X_1 = x_1, \ldots, X_n = x_n$, what is the maximum likelihood estimate of p?

(c) Consider the poisson distribution with parameter λ given by the PMF:

$$P(i) = \frac{\lambda^i \cdot e^{-\lambda}}{i!} \qquad i \geq 0$$

Given n samples $X_1 = x_1, \ldots, X_n = x_n$, what is the maximum likelihood estimate of λ?

5. Maximum Likelihood revisited

(a) Consider the uniform distribution with parameters a and b ($a < b$) given by the PDF:

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

Given n samples $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$ drawn from a uniform distribution, what are the maximum likelihood estimates of a and b?

(b) Consider the exponential distribution with parameter λ given by the PDF:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Given n samples $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$ drawn from an exponential distribution, what is the maximum likelihood estimate of λ?

(c) Consider the Gaussian distribution with parameters μ and σ^2 given by the PDF:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Given n samples $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$ drawn from a Gaussian distribution, what are the maximum likelihood estimate of μ and σ^2? Find μ first by differentiating the log-likelihood with respect to μ and then find σ^2.

6. Uniform Distribution

You have two spinners, each having a circumference of 10, with values in the range $[0,10)$. If you spin both (independently) and let X be the position of the first spinner and Y be the position of the second spinner, what is the probability that $X \geq 5$, given that $Y \geq X$?

7. Normal Distribution

The average jump of a certain frog is 3 inches. However, because of the wind, the frog does not always go exactly 3 inches. A zoologist tells you that the distance the frog travels is normally distributed with mean 3 and variance $\frac{1}{4}$. Recall that given a random variable X distributed normally with mean μ and variance σ^2, the random variable $Z = \frac{X-\mu}{\sigma}$ is distributed normally with mean 0 and variance 1. Observe that this implies that $\Pr[Z \leq z] = \Pr[X \leq \sigma z + \mu]$. Let us define the function $\Phi(z)$ as the CDF of the standard normal, i.e., $\Phi(z) = \Pr[Z \leq z]$ where $Z = N(0,1)$. For the following problems, write your answer in terms of $\Phi(z)$. Then use the table on the back of this sheet to get a numerical answer.

(a) What is the probability that the frog jumps more than 4 inches?

(b) What is the probability that the distance the frog jumps is between 2 and 4 inches?

8. Count it!

For each of the following collections, determine and briefly explain whether it is finite, countably infinite (like the natural numbers), or uncountably infinite (like the reals):

(a) The integers which divide 8.

(b) The integers which 8 divides.

(c) The functions from $\mathbb{N}$ to $\mathbb{N}$.

(d) Computer programs that halt.

(e) Computer programs that always correctly tell if a program halts or not.

(f) Numbers that are the roots of nonzero polynomials with integer coefficients.

(g) The number of points in the unit square $[0,1] \times [0,1]$

(h) Computer programs that correctly return the product of their two integer arguments

9. Compute this

(a) Can you write a program that gets n (a natural number) as input and finds the shortest formula that computes n? A formula is a valid sequence consisting of decimal digits, the operators $+$, $\times$, ^(raising to the power), and parentheses. The length of a formula is simply the number of characters you need to use to type it (i.e. each operator, decimal digit, or paranthesis counts as one character).

(b) Now assume that you want to write a computer program that given the input n (a natural number) finds another computer program (in a specific language, e.g. C or Python) that prints n. The program that is found has to have the minimum length plus execution time amongst all programs that print n, where length is measured by the number of characters in the source code and execution time is measured by a concrete number such as the number of CPU instructions executed. Can this be done?

(c) Consider the set of programs (or functions) that take a single natural number n as input and output a natural number in at most $10^6 + 2^n$ steps (i.e. they always terminate after $10^6 + 2^n$ steps). Let this set be L. A member of L is called **thorough** if every natural number m can be produced as its output (by an appropriate input). Can you write a program that takes a member of L as input and determines whether that member is thorough? The given member of L is guaranteed to be in L, there is no need for your program to verify the membership.

(HINT: If you had such a program, could you somehow use it to solve the halting problem? If so, what would that mean?)

10. Write 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?

Here is a table of approximations of the CDF of the standard normal distribution. To look up the $\Pr\left[N(0,1) \leq x\right]$, you look up the row of the first two digits of x and then the column of the third digit. For example, to look up $x = 1.34$, first find the row labeled 1.3 and then the column 0.04. The result is the entry in that cell, 0.9099.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990