Stochastic Calculus, Courant Institute, Fall 2014
http://www.math.nyu.edu/faculty/goodman/teaching/StochCalc2014/index.html
Always check the class message board before doing any work on the assignment.

Assignment 3, due October 20

Corrections (check the class message board): (none yet.)

1. (*Conditioning, etc.*) Suppose $X_1, \ldots, X_4$ are the results of 4 coin tosses. The possible results of a single toss are $X_k = \texttt{H}$ or $X_k = \texttt{T}$. (An old American coin had the head, for H, of a buffalo on one side and the tail, for T, on the other.) The probability space Ω is the space of all 16 possible sequences. The $\sigma-$ algebra $\mathcal{F}$ reflects the state where you know only the number of X_k with $X_k = \texttt{H}$, but not the values of k. For example, in $\mathcal{F}$, the outcomes $\omega = \texttt{HHTT}$ and $\omega' = \texttt{THTH}$ are indistinguishable. Consider the function $f(\omega)$ which records the first k with $X_k = \texttt{H}$. Set $f(\texttt{TTTT}) = 5$ and $f(\texttt{HTHT}) = 1$. Let $\mathcal{P}$ be the partition corresponding to $\mathcal{F}$.

   (a) Show that there are 5 elements of $\mathcal{P}$.

   (b) List the elements of the equivalence class $[\texttt{HHHT}] \in \mathcal{P}$.

   (c) Calculate the numbers $P(B_j)$.

   (d) Suppose that $P(\omega) = 1/16$ for each $\omega \in \Omega$. Calculate conditional expectation $g = E[f \mid \mathcal{F}]$. You may express this as $g(j)$, where j is the number of H tosses in $B_j \in \mathcal{P}$.

2. (*Urn process*) The urn process is a simple but not trivial one dimensional random walk. The urn has m balls in all. At each time, t, X_t of them are red and the rest are blue. To go from t to $t+1$, you select one of the m balls, each being equally likely to be chosen. You replace the selected ball with a new independent one, making the new one red with probability p and blue with probability $1-p$. We will use it in later classes to see how the Ornstein Uhlenbeck process arises as a limit.

   (a) Calculate the transition probabilities $c_x = \mathrm{P}(\, x \to x+1)$, and $a_x = \mathrm{P}(\, x \to x-1)$, and $b_x = \mathrm{P}(\, x \to x) = \mathrm{P}(\,\text{new ball same color as old ball})$. Here x is the number of red balls before a ball is replaced. The formulas depend on m (the total number of balls), and p (the probability to put back a red ball).

   (b) Figure out the forward equation for $u_{n+1,x}$ in terms of $u_{n,x-1}$, $u_{n,x}$, and $u_{n,x+1}$, and the numbers a_x, b_x, and c_x from part a.

   (c) Write the equations satisfied by the steady state probabilities π_i. Show using algebra that these equations are satisfied by (possibly a small variation on)

$$\pi_x = p^x (1-p)^{m-x} \binom{m}{x} \; . \tag{1}$$

The *binomial coefficient* is

$$\binom{m}{x} = \frac{m!}{x!(m-x)!} \ .$$

Hint: you can relate neighboring binomial coefficients using reasoning such as (approximately)

$$\binom{m}{x+1} = \frac{m!}{(x+1)!(m-x-1)!} = \frac{m-x}{x+1}\binom{m}{x} \ .$$

(d) Give a more conceptual derivation of the solution formula (1) as follows. Imagine that when you start, all the balls in the urn are "stale". Each time you put a new ball in, that ball is "fresh". The colors on the fresh balls are independent of each other, and each fresh ball has probability p of being red. Eventually, all the balls will be fresh. When that happens, the probability distribution of the number of red balls is binomial.

(e) *Stirling's formula* is the approximation

$$n! \approx \sqrt{2\pi n}\, n^n e^{-n} = \sqrt{2\pi n}\, e^{n\log(n)-n} \ .$$

Use Stirling's formula (treating it as exact) to write an approximate formula for π_i when m i, and $m-i$ are all large. Write this in the form

$$\pi_x \approx \sqrt{\frac{m}{2\pi x(m-x)}} e^{-\phi(x,m)} \ .$$

Maximize ϕ over x (use calculus, differentiate with respect to x, ...). Show that you get $x_* \approx pm$, and argue that this is the right answer, using part c if necessary. Make a quadratic approximation to ϕ about i_* and use that to make a Gaussian approximation to π. Just substitute x_* into the prefactor. Do you get the same result as the CLT? Note (*not an action item*) that you find from this a *scaling* that $x - x_*$ is on the order of $\sqrt{m}$.

3. The *ansatz* method for solving equations is to guess the form of the solution, then find the precise solution by plugging your guess into the equation. It is not always satisfying, but it is great when it works. Consider a simple random walk on $\mathbb{Z}$ with transition probabilities a, b, and c independent of x.

   (a) Write the backward equation for this process.

   (b) Show that the backward equation has solutions of the form $f_{n,x} = \alpha_n + (x-\beta_n)^2$. Find the *recurrence relations* for α_n and β_n in terms of α_{n+1} and β_{n+1}.

   (c) Directly from the process, derive equations for $\mu_n = \mathrm{E}[\,X_n]$, and $\sigma_n^2 = \mathrm{var}(X_n)$. You may assume $\mu_0 = 0$ and $\sigma_0 = 0$.

(d) Show that parts (b) and (c) are consistent, using the definition of the quantity $f_{n,i}$ in the backward equation.

4. (*computing*) This assignment will get you started doing stochastic simulation. You will simulate the simple urn process described above and check that some of its properties agree with theory.

   (a) Download the files `UrnProcess.R` and `UrnProcessCheck.pdf`. Put them in a convenient directory where you will put your `R` files. Open your `R` application in that directory, or, if you will use the command line, just `cd` to that directory. If you use the `R` application, type: `source("UrnProcess.R")` It should create a file `UrnProcess.pdf` that is identical to `UrnProcessCheck.pdf`.

   (b) Experiment with different values of p and m. Make another plot or two showing what things can happen.

   (c) Run the code with smaller values of T so see that the probabilities $u(x,t)$ converge to π_x as $t \to \infty$. Find T where this convergence has not happened and larger T where it has. Show that you need larger T if you have larger m. You may also need more paths, as the histograms might start looking ragged.

   (d) Add another curve to the graph that represents the CLT predictions of π_x. Make a couple of plots showing that this is not accurate for small m, but gets better as m increases.