Stochastic Calculus Assignment 2 Solution

Chen-Hung Wu

December 10, 2012

1. *(Urn process)* The urn process is a simple but not trivial one dimensional random walk. In later classes we will come back to it to see how it goes over to the Ornstein Uhlenbeck process in the limit $m \to \infty$, $T \to \infty$, but m and T related by a *scaling* that we will figure out.

(a) Calculate the transition probabilities $c_i = P(i \to i+1)$ and $a_i = P(i \to i-1)$. Here i is the number of red balls. The formulas depend on m (the total number of balls), and p (the probability to put back a red ball).

Sol: The transition probability c_i means that in this process you choose a red ball and replace it with a blue ball, thus $c_i = \frac{m-i}{m}p$. If one choose a blue ball and replace it with a red ball, then the transition probability $a_i = \frac{i}{m}(1-p)$. Also

$$\begin{aligned} b_i &= P(i \to i) \\ &= \frac{m-i}{m}(1-p) + \frac{i}{m}p \\ &= 1 - a_i - c_i \end{aligned}$$

(b) Figure out the forward equation for $u_{n+1,i}$ in terms of $u_{n,i-1}$, $u_{n,i}$, and $u_{n,i+1}$, and the numbers a_i, b_i, and c_i from part a.

Sol: Let $u_{n,i} := P(X_n = i)$, then the forward equation is a formula for the number

$$\begin{aligned} u_{n+1,i} &= P\left(X_{n+1} = i\right) \\ &= \sum_{y \in S} P\left(X_{n+1} = i \middle| X_n = y\right) P\left(X_n = y\right) \\ &= u_{n,i-1}P(i-1 \to i) + u_{n,i}P(i \to i) + u_{n,i+1}P(i+1 \to i) \\ &= u_{n,i-1}c_{i-1} + u_{n,i}b_i + u_{n,i+1}a_{i+1}. \end{aligned}$$

(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)

$$\pi_i = p^i(1-p)^{m-i} \left(\begin{array}{c} m \\ i \end{array} \right). \tag{1}$$

The *binomial coefficient* is

$$\left(\begin{array}{c} m \\ i \end{array} \right) = \frac{m!}{i!(m-i)!}.$$

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

$$\left(\begin{array}{c} m \\ i+1 \end{array} \right) = \frac{m!}{(i+1)!(m-i-1)!} = \frac{m-i}{i+1} \left(\begin{array}{c} m \\ i \end{array} \right).$$

Sol: A probability distribution, π, is *stationary*, or *steady state*, or *statistical steady state*, if $u_n = \pi \Rightarrow u_{n+1} = \pi$. That is the same as saying that $X_n \sim \pi \Rightarrow X_{n+1} \sim \pi$. The forward equation implies that a stationary probability distribution must satisfy the equation $\pi = \pi P$, which determines π. Let us consider the equations

$$\begin{aligned}
\pi_0 &= \pi_0(1-p) + \pi_1 \frac{1}{m}(1-p) \\
\pi_1 &= \pi_0 p + \pi_1 \left(\frac{1}{m} p + \frac{m-1}{m}(1-p) \right) + \pi_2 \frac{2}{m}(1-p) \\
\pi_2 &= \pi_1 \frac{m-1}{m} p + \pi_2 \left(\frac{2}{m} p + \frac{m-2}{m}(1-p) \right) + \pi_3 \frac{3}{m}(1-p) \\
&\vdots \\
&\vdots \\
\pi_m &= \pi_{m-1} \frac{1}{m} p + \pi_m p,
\end{aligned}$$

combined with the constrain $\sum_i \pi_i = 1$. Solving above system by the tridiagonal matrix algorithm we would have

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

as claimed.

(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

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

Maximize ϕ over i (use calculus, differentiate with respect to i, ...). Show that you get $i_* \approx pm$, 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 i_* into the pre-factor. Do you get the same result as the CLT? Note (*not an action item*) that you find from this a *scaling* that $i - i_*$ is on the order of $\sqrt{m}$.

Sol: Straightforward calculation,

$$\begin{aligned}
\pi_i &= p^i(1-p)^{m-i}\frac{m!}{i!(m-i)!} \\
&\approx \sqrt{\frac{m}{2\pi i(m-i)}}e^{m\log m - m - i\log i + i - (m-i)\log(m-i) + (m-i) + i\log p + (m-i)\log(1-p)} \\
&= \sqrt{\frac{m}{2\pi i(m-i)}}e^{m\log m + i\log\frac{p}{i} + (m-i)\log\frac{1-p}{m-i}} \\
&= \sqrt{\frac{m}{2\pi i(m-i)}}\exp\left[-\underbrace{\left(-m\log m - i\log\frac{p}{i} - (m-i)\log\frac{1-p}{m-i}\right)}_{:=\phi(i,m)}\right]
\end{aligned}$$

To maximize ϕ, we differentiate ϕ w.r.t. i,

$$\begin{aligned}
\partial_i\phi &= \log p - \log i - 1 - \log(1-p) + \log(m-i) + 1 \\
&= 0,
\end{aligned}$$

which means

$$p(m-i_*) = i_*(1-p),$$

and thus $i_* = pm$.

Notice that

$$\begin{aligned}
\phi(i_*, m) &= -m\log m - mp\log\frac{p}{mp} - (m-mp)\log\frac{1-p}{m-mp} \\
&= -2pm\log m,
\end{aligned}$$

and

$$\phi''(i_*) = \frac{1}{\sigma^2}.$$

Consider the Taylor expansion of $\phi(i)$ around i_*,

$$\phi(i) \approx \phi(i_*) + \phi'(i_*)(i-i_*) + \frac{1}{2}\phi''(i_*)(i-i_*)^2 + h.o.t.$$

Then we can approximate

$$\pi_i \approx C\exp\left(-\phi(i_*) + \phi'(i_*)(i-i_*) + \frac{1}{2}\phi''(i_*)(i-i_*)^2 + h.o.t.\right).$$

2. 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 i.

(a) Write the backward equation for this process.

Sol:

$$\begin{aligned} f_{n,i} &= E\left[\, f_{n+1} \middle| \, \mathcal{F}_n\right](x_n = i) \\ &= \sum_{x_{n+1}\in\mathcal{S}} P\left(X_{n+1} = x_{n+1} \middle| \, X_n = x_n\right) f_{n+1}(x_{n+1}) \\ &= af_{n+1}(x_{n+1} = i-1) + bf_{n+1}(x_{n+1} = i) + cf_{n+1}(x_{n+1} = i+1) \\ &= af_{n+1,i-1} + bf_{n+1,i} + cf_{n+1,i+1}. \end{aligned} \tag{2}$$

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

Sol: Plugging $f_{n,i}$ into (2),

$$\alpha_n + (i-\beta_n)^2 = a\alpha_{n+1} + a(i-1-\beta_{n+1})^2 + b\alpha_{n+1} + b(i-\beta_{n+1})^2 + c\alpha_{n+1} + c(i+1-\beta_{n+1})^2.$$

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

Sol:

$$\begin{aligned} \mu_1 &= c - a \\ \sigma_1^2 &= c + a - (c-a)^2 \\ \mu_2 &= 2a^2 + 2ab - 2bc - 2c^2 \\ &= 2\left[(a+c)(a-c) + b(a-c)\right] \\ &= 2(a-c)(a+b+c) \\ &= 2(a-c) \\ \sigma_2^2 &= 4a^2 + 4c^2 + 2bc + 2ab - 4(a-c)^2 \\ &= 2bc + 2ab + 8ac \\ &= 2(1-a-c)(a+c) + 8ac \\ &= 2(a+c-a^2-2ac-c^2) + 8ac \\ &= 2(c+a) - 2(a^2-4ac+c^2) \\ &= 2\left[c+a-(c-a)^2\right] \\ &\vdots \\ &\vdots \end{aligned}$$

We can conclude that,

$$\begin{aligned} \mu_n &= \sum_{i=1}^{n} \mu_i \\ &= n(c-a), \end{aligned}$$

and

$$\sigma_n^2 = n\left[c + a - (c-a)^2\right].$$

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