Massachusetts Institute of Technology

Department of Electrical Engineering & Computer Science

6.041/6.431: Probabilistic Systems Analysis

(Fall 2009)

Problem Set 10: Solutions
Due December 4, 2009

1. Because of the independence of the per-week profits and the high number of weeks, the desired probability can be well approximated by the Central Limit Theorem. Let X_i be Joe's profit in the ith week. The X_is have a common mean $\mu = 5200$ and variance $\sigma^2 = \frac{1}{12}(6800)^2$.

$$\begin{aligned}
P\left(\sum_{i=1}^{52} X_i \leq 250000\right) &= P\left(\frac{\sum_{i=1}^{52} X_i - 52\mu}{\sigma\sqrt{52}} \leq \frac{250000 - 52\mu}{\sigma\sqrt{52}}\right) \\
&\approx P\left(Z \leq \frac{250000 - 52\mu}{\sigma\sqrt{52}}\right) \text{ where Z is a standard normal r.v. (CLT)} \\
&= P\left(Z \leq \frac{250000 - 52(5200)}{\frac{1}{\sqrt{12}}6800\sqrt{52}}\right) \\
&= P\left(Z \leq \frac{-3\sqrt{3}}{\sqrt{13}}\right) \\
&= 1 - \phi\left(\frac{3\sqrt{3}}{\sqrt{13}}\right) \approx 1 - \phi(1.44) \approx 0.075
\end{aligned}$$

2. The probability that the airline will have to deny passengers from boarding is the probability that more than 300 passengers show up. Let N be the number of passengers that show up, which is a binomial random variable, with parameters n and $p = 0.9$; thus, $\mathbf{E}[N] = 0.9n$ and $\sigma_N = \sqrt{n(0.1)(0.9)} = \sqrt{0.09n}$. Using the de Moivre - Laplace normal approximation to the binomial, we have

$$\begin{aligned}
\mathbf{P}(N > 300) &= \mathbf{P}(N \geq 300 + 0.5) \\
&= \mathbf{P}\left(\frac{N - 0.9n}{\sqrt{0.09n}} \geq \frac{300.5 - 0.9n}{\sqrt{0.09n}}\right) \\
&\approx 1 - \Phi\left(\frac{300.5 - 0.9n}{\sqrt{0.09n}}\right)
\end{aligned}$$

In order for the airline to be approximately 99 percent confident that it will not have to deny boarding to passengers holding tickets,

$$\mathbf{P}(N > 300) \approx 0.01$$

Therefore,

$$\Phi\left(\frac{300.5 - 0.9n}{\sqrt{0.09n}}\right) \approx 0.99$$

and as $\Phi(2.33) \approx 0.99$,

$$\begin{aligned}
\frac{300.5 - 0.9n}{\sqrt{0.09n}} &\approx 2.33 \\
300.5 - 0.9n &\approx 2.33\sqrt{0.09n} \\
(0.9)^2 n^2 - 2(0.9)(300.5)n + (300.5)^2 &\approx (2.33)^2(0.09)n
\end{aligned}$$

Solving the quadratic equation above and choosing the lesser of the two solutions for n tell us that n≈320.

3. (a) Method 1:
No. Using the CLT, the approximate distribution of T_i for large i is that of a gaussian random variable with mean 0 and variance $i\text{var}(X) = i/3$. As $i \to \infty$, $\text{var}(T_i) \to \infty$, and $\mathbf{P}(a \le T_i \le b) \to 0$ for all $a \le b$. So, $\mathbf{P}\left(|T_i - c| \le \epsilon\right) \to 0$ as $i \to \infty$ for all c and $\epsilon > 0$. Therefore T_i cannot converge in probability to any constant.
Method 2:
In order for T_i to converge in probability, $T_i - T_{i-1}$ must converge to zero in probability. Since $T_i - T_{i-1} = X_i$ for all i, $T_i - T_{i-1}$ does not converge to zero, and therefore T_i does not converge in probability.

(b) Yes, to 0. Applying weak law of large numbers, we have

$$\mathbf{P}(|U_i - \mu| > \epsilon) \to 0 \text{ as } i \to \infty, \text{ for all } \epsilon > 0$$

Here $\mu = 0$ since $X_i \sim U(-1.0, 1.0)$.

(c) Yes, to 1. Since for $\epsilon > 0$,

$$\begin{aligned}
\lim_{i\to\infty} \mathbf{P}(|W_i - 1| > \epsilon) &\le \lim_{i\to\infty} \mathbf{P}(|\max\{X_1, \cdots, X_i\} - 1| > \epsilon) \\
&= \lim_{i\to\infty} [\mathbf{P}(\max\{X_1, \cdots, X_i\} > 1 + \epsilon) \\
&\quad + \mathbf{P}(\max\{X_1, \cdots, X_i\} < 1 - \epsilon)] \\
&= \lim_{i\to\infty} [0 + (1 - \frac{\epsilon}{2})^i] \\
&= 0.
\end{aligned}$$

(d) Yes, to 0.

Method 1:
$|V_n| \le \min\{|X_1|, |X_2|, \cdots, |X_n|\}$
but $\min\{|X_1|, |X_2|, \cdots, |X_n|\}$ converges to 0 in probability. So, since $|V_n| \ge 0$, $|V_n|$ converges to 0 in probability. To see why $\min\{|X_1|, |X_2|, \cdots, |X_n|\}$ converges to 0 in probability, we calculate:

$$\begin{aligned}
\lim_{i\to\infty} \mathbf{P}\left(|\min\{|X_1|, \cdots, |X_i|\} - 0| > \epsilon\right) &= \lim_{i\to\infty} \mathbf{P}\left(\min\{|X_1|, \cdots, |X_i|\} > \epsilon\right) \\
&= \lim_{i\to\infty} \mathbf{P}(|X_1| > \epsilon) \cdot \mathbf{P}(|X_2| > \epsilon) \cdots \mathbf{P}(|X_i| > \epsilon) \\
&= \lim_{i\to\infty} (1 + \epsilon)^i \text{ since } |X_i| \text{ is uniform between 0 and 1} \\
&= 0.
\end{aligned}$$

Method 2:

$$\begin{aligned}
\mathbf{E}[V_n] &= \mathbf{E}[\prod_{k=1}^{n} X_k] \\
&= \prod_{k=1}^{n} \mathbf{E}[X_k] = 0 \\
\text{var}(V_n) &= \mathbf{E}[\text{var}(V_n|X_n)] + \text{var}(\mathbf{E}[V_n|X_n]) \\
&= \mathbf{E}[X_n^2 \text{var}(V_{n-1})] + \text{var}(X_n \mathbf{E}[V_{n-1}]) \\
&= \mathbf{E}[X_n^2]\text{var}(V_{n-1}) + (\mathbf{E}[V_{n-1}])^2 \text{var}(X_n) \\
&= \frac{1}{3}\text{var}(V_{n-1}) = \left(\frac{1}{3}\right)^{n-1} \text{var}(X_1)
\end{aligned}$$

Notice that as n becomes very large, $\text{var}(V_n)$ approaches 0. By Chebyshev's inequality, we know V_n approaches $\mathbf{E}[V_n] = 0$ in probability.

4. The probability that you will believe the fair coin to be biased is the probability that the fair coin will come up with more than 525 heads out of the 1000 tosses. Let S be the number of times the coin comes up heads, which is a binomial random variable, with parameters $n = 1000$ and $p = 0.5$, so that $\mathbf{E}[S] = (1000)(0.5) = 500$ and $\sigma_S = \sqrt{(1000)(0.5)(0.5)} = 5\sqrt{10}$.

   (a) Using the de Moivre - Laplace normal approximation to the binomial, we have

$$\begin{aligned}
\mathbf{P}(S > 525) &= \mathbf{P}(S \geq 525.5) \\
&= \mathbf{P}\left(\frac{S - 500}{5\sqrt{10}} \geq \frac{525.5 - 500}{5\sqrt{10}}\right) \\
&\approx 1 - \Phi\left(\frac{25.5}{5\sqrt{10}}\right) \\
&= 1 - \Phi(1.6128) \\
&\approx 0.0537.
\end{aligned}$$

   (b) Using the Markov inequality, we have

$$\begin{aligned}
\mathbf{P}(S > 525) &= \mathbf{P}(S \geq 526) \\
&\leq \frac{\mathbf{E}[S]}{526} \\
&= \frac{500}{526} \\
&\approx 0.951.
\end{aligned}$$

   We see that using the Markov inequality gives us a weak upper bound, considering the approximate probability as calculated in part (a).

(c) Using the Chebyshev inequality, we have

$$\begin{aligned}
\mathbf{P}(S > 525) &= \mathbf{P}(S \geq 526) \\
&= \frac{1}{2}(\mathbf{P}(S \geq 526) + \mathbf{P}(S \leq 474)) \text{ (by symmetry, since p = 0.5)} \\
&= \frac{1}{2}P(|S - 500| \geq 26) \\
&\leq \frac{1}{2}\frac{\sigma_S^2}{26^2} \\
&= \frac{1}{2}\frac{25 \cdot 10}{26^2} \\
&\approx 0.185.
\end{aligned}$$

We see that the Chesyshev inequality provides a substantial improvement upon the upper bound calculated by the Markov inequality in part (b).

5. Consider a random variable X with PMF

$$p_X(x) = \begin{cases} p, & \text{if } x = \mu - c; \\ p, & \text{if } x = \mu + c; \\ 1 - 2p, & \text{if } x = \mu. \end{cases}$$

The mean of X is μ, and the variance of X is $2pc^2$. To make the variance equal σ^2, set $p = \frac{\sigma^2}{2c^2}$. For this random variable, we have

$$\mathbf{P}(|X - \mu| \geq c) = 2p = \frac{\sigma^2}{c^2},$$

and therefore the Chebyshev inequality is tight.

6. (a) From the central limit theorem we know that as $n \to \infty$, S_n behaves as a normal random variable with mean $n\lambda$ and variance $n\lambda(1-\lambda)$. Equivalently, $\frac{S_n - n\lambda}{\sqrt{n\lambda(1-\lambda)}}$ behaves as standard normal random variable. Therefore,

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{n\lambda - m \leq i \leq n\lambda + m} \mathbf{P}(S_n = i) \\
&= \lim_{n\to\infty} \mathbf{P}\left[n\lambda - m \leq S_n \leq n\lambda + m\right] \\
&= \lim_{n\to\infty} \mathbf{P}\left[\frac{-m}{\sqrt{n\lambda(1-\lambda)}} \leq \frac{S_n - n\lambda}{\sqrt{n\lambda(1-\lambda)}} \leq \frac{m}{\sqrt{n\lambda(1-\lambda)}}\right] \\
&= \lim_{n\to\infty} \Phi\left(\frac{m}{\sqrt{(n\lambda(1-\lambda))}}\right) - \Phi\left(\frac{-m}{\sqrt{(n\lambda(1-\lambda))}}\right) \\
&= \Phi(0) - \Phi(0) = 1/2 - 1/2 = 0
\end{aligned}$$

where Φ is the standard normal distribution.

(b) We are still using the fact that S_n behaves as a normal random variable as $n \to \infty$.

$$\begin{aligned}
\lim_{n\to\infty} \sum_{0 \le i \le n\lambda+m} \mathbf{P}(S_n = i) &= \lim_{n\to\infty} \mathbf{P}\left[0 \le S_n \le n\lambda + m\right] \\
&= \lim_{n\to\infty} \mathbf{P}\left[\frac{-n\lambda}{\sqrt{n\lambda(1-\lambda)}} \le \frac{S_n - n\lambda}{\sqrt{n\lambda(1-\lambda)}} \le \frac{m}{\sqrt{n\lambda(1-\lambda)}}\right] \\
&= \lim_{n\to\infty} \Phi\left(\frac{m}{\sqrt{(n\lambda(1-\lambda))}}\right) - \Phi\left(\frac{-n\lambda}{\sqrt{(n\lambda(1-\lambda))}}\right) \\
&= \Phi(0) - \Phi(-\infty) = 1/2 - 0 = 1/2
\end{aligned}$$

(c)

$$\begin{aligned}
\lim_{n\to\infty} \sum_{i:n(\lambda-1/m)\le i \le n(\lambda+1/m} \mathbf{P}(S_n = i) &= \lim_{n\to\infty} \mathbf{P}\left[\frac{-n}{m\sqrt{n\lambda(1-\lambda)}} \le \frac{S_n - n\lambda}{\sqrt{n\lambda(1-\lambda)}} \le \frac{n}{m\sqrt{n\lambda(1-\lambda)}}\right] \\
&= \lim_{n\to\infty} \Phi\left(\frac{n}{m\sqrt{(n\lambda(1-\lambda))}}\right) - \Phi\left(\frac{-n}{m\sqrt{(n\lambda(1-\lambda))}}\right) \\
&= \Phi(\infty) - \Phi(-\infty) = 1 - 0 = 1
\end{aligned}$$

7. (a) To recover the loss he sustains from a single tails, the player requires the profit from two heads. Thus to break even or win, the player requires at least twice as many heads as tails. For 100 tosses, he needs at least 67 heads.

(b) $P_w = \text{P(at least 67 heads in 100 tosses)} = \sum_{k=67}^{100} \binom{100}{k} \left(\frac{1}{2}\right)^{100}$

(c) Let S_{100} be the number of heads in 100 tosses of a fair coin. Then $S_{100} \ge 0$ and $E[S_{100}] = 50$, so, we can use the Markov bound:

$$P(S_{100} \ge 67) \le \frac{\mathbf{E}[S_{100}]}{67} \approx \frac{\frac{1}{2}\cdot 100}{\frac{2}{3}\cdot 100} \approx 0.75$$

(d) Let $X_n = 1$ if the nth toss is heads and $X_n = 0$ if the nth toss is tails. Then $\text{var}(X_n) = \frac{1}{4}$, the X_n's are independent, and

$$S_{100} = \sum_{k=1}^{100} X_k$$

Therefore $(\sigma_{S_{100}})^2 = 25$, $\sigma_{S_{100}} = 5$, so from the Chebyshev bound,

$$\mathbf{P}(S_{100} \ge 67) \le \mathbf{P}(|S_{100} - E[S_{100}]| \ge 17) \le \frac{(\sigma_{S_{100}})^2}{17^2} = \left(\frac{5}{17}\right)^2 \approx 0.087$$

From the symmetry of the problem we see that $\mathbf{P}(S_{100} - \mathbf{E}[S_{100}] \geq 17) = \mathbf{P}(S_{100} - \mathbf{E}[S_{100}] \leq -17)$. So we can halve the above bound:

$$P(S_{100} \geq 67) \leq \frac{{\sigma_{S_{100}}}^2}{2 \cdot (17)^2} = 0.043$$

(e) Using $S_{100} \geq 67 \approx \mathbf{E}[S_{100}] + 3.4\sigma_{S_{100}}$ in the central limit theorem gives

$$\mathbf{P}(S_{100} \geq 67) \approx 1 - \Phi(3.4) \approx 0.0003.$$

So the Double or Quarter game (at least when played this way) looks like a poor way to attempt to make one's fortune. The game is deceptive in that your expected wealth (i.e, $(9.8)^n$) becomes very large due to an ever smaller probability of ever larger wins as $n \to \infty$.

G1†. Using the Chernoff bound, we have for $r > 0$

$$\mathbf{P}(S_{100} \geq 67) = \mathbf{P}(e^{rS_{100}} \geq e^{67r}) \leq g_{S_{100}}(r)e^{-67r},$$

We seek the minimum value of $[g_{S_{100}}(r)e^{-67r}] = \left(\frac{1+e^r}{2}\right)^{100} e^{-67r}$, since S_{100} is binomial(100,$\frac{1}{2}$). Differentiating to find the minimum,

$$\begin{aligned}
\frac{d}{dr}\left[\left(\frac{1+e^r}{2}\right)^{100} e^{-67r}\right] &= 0 \\
&= 100\left(\frac{1+e^r}{2}\right)^{99} \frac{e^r}{2} e^{-67r} - (67)\left(\frac{1+e^r}{2}\right)^{100} e^{-67r} \\
&= \left(\frac{1+e^r}{2}\right)^{99} e^{-67r}\left(100 \cdot \frac{e^r}{2} - 67 \cdot \left(\frac{1+e^r}{2}\right)\right) \\
&\implies 33e^{r_{opt}} = 67 \\
&\implies r_{opt} = \ln(\frac{67}{33}) = 0.7082
\end{aligned}$$

This gives

$$\mathbf{P}(S_{100} \geq 67) \leq \left(\frac{1+e^{r_{opt}}}{2}\right)^{100} e^{-67r_{opt}} = 0.00275$$

This bound is roughly one order of magnitude tighter than the Chebyshev bound for this example, but one order of magnitude looser than the CLT approximation.

G2†. (a) Using the lemma, let $Z_n = \frac{1}{n}\sum_{k=1}^{n} \log_b(X_k)$ and $R_n = b^{Z_n}$. We know from the weak law of large numbers that $Z_n \xrightarrow{prob.} \mathbf{E}[\log_b(X)]$. Since exponentiation is continuous, it follows from the lemma that $R_n \xrightarrow{prob.} b^{\mathbf{E}[\log_b(X)]}$. For the double or quarter game, using b=2, we have:

$$\mathbf{E}[\log_2(X)] = (1/2)(-2) + (1/2)(1) = -1/2.$$

$$R_n \xrightarrow{prob.} \mathbf{E}[\log_2(X)] = 2^{-1/2} = \frac{1}{\sqrt{2}} = 0.7071$$

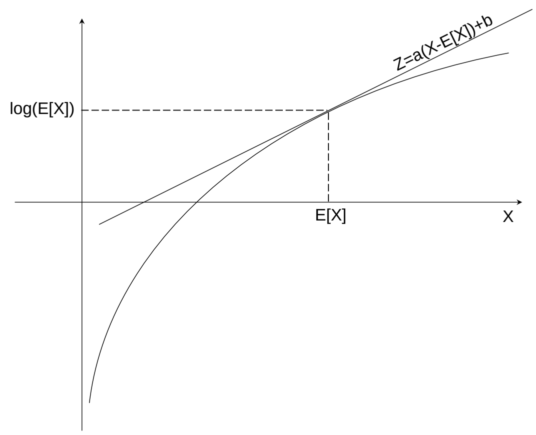

Figure 1: The log function

(b) For each ω and each $n \geq 1$, $W_n(\omega) = [R_n(\omega)]^n$. Since $R_n \xrightarrow{prob.} r = \sqrt{2}/2$, for each $\epsilon > 0$, $\lim_{n\to\infty} \mathbf{P}(r_n \geq r + \epsilon) = 0$. Choose, say, $\epsilon = 0.8 - \sqrt{2}/2$. Then $\lim_{n\to\infty} \mathbf{P}(W_n = (R_n)^n \geq (0.8)^n) = 0$. Since $(0.8)^n \xrightarrow{n\to\infty} 0$, it follows that for any $\delta > 0$, $\lim_{n\to\infty} \mathbf{P}(W_n \geq \delta) = 0$, i.e., (since $W_n > 0$), $W_n \xrightarrow{prob.} 0$. This doesn't look like a promising method to make one's fortune!

(c) Since the logarithm function is concave (i.e., $\frac{d\log(x)}{dx} < 0$ everywhere), the tangent line to y = log(x) at any point lies strictly above the plot of $y = \log(x)$. (Refer to the figure above). In particular, the tangent line at $x = \mathbf{E}[X]$ lies above the plot and is given by the equation:

$$y = f_L(x) = \log(\mathbf{E}[X]) + \frac{d\log(x)}{dx}\Big|_{x=\mathbf{E}[X]}(x - \mathbf{E}[X]).$$

Since $f_L(x) \geq \log(x)$ for all x, it follows that

$$\mathbf{E}[f_L(X)] = \log(\mathbf{E}[X]) \geq \mathbf{E}[\log(X)].$$

.

Therefore, As $n \to \infty$, $R_n = b^{\frac{1}{n}\sum_{k=1}^{n}\log_b(X_k)} \xrightarrow{prob.} b^{\mathbf{E}[\log_b(X)]} \leq b^{\log_b(\mathbf{E}[X])} = \mathbf{E}[X]$

(d) We are given

$$\begin{aligned} W_{n+1} &= fW_nX_{n+1} + (1-f)W_n \\ &= [fX_{n+1} + (1-f)]W_n \end{aligned}$$

Let $Z_n = fX_n + (1-f)$. Because $W_1 = 1$, we have

$$\begin{aligned} W_n &= Z_1 \cdot Z_2 \ldots Z_n \\ \Longrightarrow R_n &= W_n{}^{\frac{1}{n}} = (Z_1 \cdot Z_2 \ldots Z_n)^{\frac{1}{n}},\ Z_i \text{ are i.i.d.} \end{aligned}$$

As in part (a), we have $R_n \xrightarrow{prob} r = 2^{\mathbf{E}[\log_2(Z_i)]}$. For $r > 1$ we require

$$\begin{aligned}
&\mathbf{E}[\log_2(Z_i)] > 0 \\
\Longrightarrow\ &\frac{1}{2}\log(1+f) + \frac{1}{2}\log\left(1 - \frac{3f}{4}\right) > 0 \\
\Longrightarrow\ &log\left[(1+f)\left(1 - \frac{3f}{4}\right)\right] > 0 \\
\Longrightarrow\ &1 + \frac{f}{4} - \frac{3}{4}f^2 > 1 \\
\Longrightarrow\ &\frac{f}{4}(1-3f) > 0 \\
\Longrightarrow\ &0 < f < \frac{1}{3}
\end{aligned}$$

To find the value of f for which the gains will be fastest, we need to maximize the value of $r = 2^{\mathbf{E}[\log_2 Z_i]}$. To find the optimizing f, we set the derivative of $\mathbf{E}[\log(Z_i)]$ to 0;

$$\begin{aligned}
&\frac{d}{df}\log\left(1 + \frac{f}{4} - \frac{3}{4}f^2\right) = 0 \\
\Longrightarrow\ &f = \frac{1}{6}
\end{aligned}$$

$r_{opt} = \frac{1}{2}\log_2(7/6) + \frac{1}{2}\log_2(7/8) = 0.0149$