Massachusetts Institute of Technology
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Spring 2011)

Problem Set 10: Solutions

1. (a) We want to find the probability that there are at least 45 successes out of 50 total trials, where the probability of success is given to be .95. Using the Normal approximation to the binomial (where $\mu = 47.5$ and $\sigma \approx 1.54$), we find:

$$\begin{aligned}
\mathbf{P}(\text{45 to 50 successes}) &\approx 1 - \Phi\left(\frac{44.5-\mu}{\sigma}\right) \\
&\approx 1 - \Phi(-1.95) \\
&= \Phi(1.95) \\
&= 0.9744
\end{aligned}$$

(b) To be able to use the Poisson approximation p has to be small and n has to be relatively large. Therefore, using $p = 0.95$ will not give a good approximation. Instead, we define a new random variable, I, to be the number of incorrect predictions out of 50.

$$\begin{aligned}
\mathbf{P}(\text{45 to 50 successes}) &= \mathbf{P}(I=0) + \mathbf{P}(I=1) + \ldots + \mathbf{P}(I=5) \\
&\approx \sum_{k=0}^{5} \frac{2.5^k e^{-2.5}}{k!} \approx 0.9582
\end{aligned}$$

The second method, although more tedious, is perhaps more appropriate. The Normal approximation works well with sums of symmetric distributions, which for the binomial is satisfied when p is close to .5. Here p is quite far from that. Of course, the Normal distribution makes it quite convenient to calculate, especially when the number of terms in the sum grows.

2. First, let's calculate the expectation and the variance for Y_n, T_n, and A_n.

$$\begin{aligned}
Y_n &= (0.5)^n X_n \\
T_n &= Y_1 + Y_2 + \cdots + Y_n \\
A_n &= \frac{1}{n} T_n
\end{aligned}$$

$$
\begin{aligned}
\mathbf{E}[Y_n] &= \mathbf{E}\left[(\tfrac{1}{2})^n X_n\right] = (\frac{1}{2})^n \mathbf{E}[X_n] = \mathbf{E}[X]\,(\tfrac{1}{2})^n = 2(\tfrac{1}{2})^n \\
\mathrm{var}(Y_n) &= \mathrm{var}\left((\tfrac{1}{2})^n X_n\right) = (\tfrac{1}{2})^{2n}\mathrm{var}(X_n) = \mathrm{var}(X)\,(\tfrac{1}{2})^{2n} = 9(\tfrac{1}{4})^n \\
\mathbf{E}[T_n] &= \mathbf{E}[Y_1 + Y_2 + \cdots + Y_n] = \mathbf{E}[Y_1] + \mathbf{E}[Y_2] + \cdots + \mathbf{E}[Y_n] \\
&= 2\sum(\tfrac{1}{2})^i = 2\frac{0.5(1-0.5^n)}{1-0.5} = 2\left(1-(\tfrac{1}{2})^n\right) \\
\mathrm{var}(T_n) &= \mathrm{var}(Y_1 + Y_2 + \cdots + Y_n) = \sum_{i=1}^{n}(\tfrac{1}{4})^i \mathrm{var}(X_i) \\
&= 9\left(\frac{\frac{1}{4}\left(1-(\frac{1}{4})^n\right)}{1-\frac{1}{4}}\right) = 3\left(1-(\tfrac{1}{4})^n\right) \\
\mathbf{E}[A_n] &= \mathbf{E}\left[\frac{1}{n}T_n\right] = \frac{1}{n}\mathbf{E}[T_n] = \frac{2}{n}\left(1-(\tfrac{1}{2})^n\right) \\
\mathrm{var}(A_n) &= \mathrm{var}\left(\frac{1}{n}T_n\right) = \left(\frac{1}{n}\right)^2 \mathrm{var}(T_n) = \frac{3}{n^2}\left(1-\left(\frac{1}{4}\right)^n\right)
\end{aligned}
$$

(a) Yes. Y_n converges to 0 in probability. As n becomes very large, the expected value of Y_n approaches 0 and the variance of Y_n approaches 0. So, by the Chebychev Inequality, Y_n converges to 0 in probability.

(b) No. Assume that T_n converges in probability to some value a. We also know that:

$$
\begin{aligned}
T_n &= Y_1 + (Y_2 + Y_3 + \ldots.. Y_n) \\
&= Y_1 + ((0.5)^2 X_2 + (0.5)^3 X_3 + \cdots + (0.5)^n X_n) \\
&= Y_1 + \frac{1}{2}(0.5X_2 + (0.5)^2 X_3 + \cdots + (0.5)^{n-1} X_n).
\end{aligned}
$$

Notice that $0.5X_2 + (0.5)^2 X_3 + \cdots + (0.5)^{n-1}X_n$ converges to the same limit as T_n when n goes to infinity. If T_n is to converge to a, Y_1 must converge to $a/2$. But this is clearly false, which presents a contradiction in our original assumption.

(c) Yes. A_n converges to 0 in probability. As n becomes very large, the expected value of A_n approaches 0, and the variance of A_n approaches 0. So, by the Chebychev Inequality, A_n converges to 0 in probability. You could also show this by noting that the A_ns are i.i.d. with finite mean and variance and using the WLLN.

3. (a) To use the Markov inequality, let $X = \sum_{i=1}^{10} X_i$. Then,

$$\mathbf{E}[X] = 10\mathbf{E}[X_i] = 5,$$

and the Markov inequality yields

$$\mathbf{P}(X \geq 7) \leq \frac{5}{7} = 0.7142.$$

(b) Using the Chebyshev inequality, we find that

$$\mathbf{P}(X - 5 \geq 2) \leq \frac{5}{48} = 0.104$$

(c) Finally, using the Central Limit Theorem, we find that

$$\begin{aligned}
\mathbf{P}\left(\sum_{i=1}^{10} X_i \geq 7\right) &= 1 - \mathbf{P}\left(\sum_{i=1}^{10} X_i \leq 7\right) \\
&= 1 - \mathbf{P}\left(\frac{\sum_{i=1}^{10} X_i - 5}{\sqrt{10/12}} \leq \frac{7-5}{\sqrt{10/12}}\right) \\
&= 1 - \Phi(2.19) \\
&= 0.0143
\end{aligned}$$

4. (a) The prior PDF of Θ is

$$f_\Theta(\theta) = \begin{cases} 1 & \text{if } 0 \leq \theta \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and the conditional PDF of the observation is

$$f_{X|\Theta}(x \mid \theta) = \begin{cases} \frac{1}{\theta} & \text{if } 0 \leq x \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

We can easily compute

$$f_{\Theta,X}(\theta, x) = f_\Theta(\theta) f_{X|\Theta}(x \mid \theta) = \begin{cases} \frac{1}{\theta} & \text{if } 0 < x \leq \theta \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and

$$f_X(x) = \int_0^1 f_\Theta(\theta) f_{X|\Theta}(x \mid \theta)\, d\theta = \int_x^1 \frac{1}{\theta}\, d\theta = \ln\frac{1}{x}.$$

Using Bayes' rule, we obtain the posterior PDF of Θ

$$f_{\Theta|X}(\theta \mid x) = \frac{f_\Theta(\theta) f_{X|\Theta}(x \mid \theta)}{f_X(x)} = \boxed{\begin{cases} \frac{1}{\theta \ln \frac{1}{x}} & \text{if } 0 < x \leq \theta \leq 1 \\ 0 & \text{otherwise} \end{cases}}$$

(b) Similar to the case when $n = 1$, we have

$$f_{X_1,\ldots,X_n|\Theta}(x_1, \ldots, x_n \mid \theta) = \begin{cases} \frac{1}{\theta^n} & \text{if } 0 < x_{\max} \leq \theta \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

where

$$x_{\max} = \max\{x_1, \cdots, x_n\}.$$

We can now obtain the joint distribution of $X_1, \ldots, X_n$ for $n > 1$:

$$\begin{aligned}
f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) &= \int_0^1 f_\Theta(\theta) f_{X_1,\ldots,X_n|\Theta}(x_1, \ldots, x_n \mid \theta)\, d\theta = \int_{x_{\max}}^1 \frac{1}{\theta^n}\, d\theta \\
&= \begin{cases} \frac{\frac{1}{x_{\max}^{n-1}} - 1}{n-1} & \text{if } 0 < x_{\max} \leq 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

The posterior PDF of Θ is therefore

$$f_{\Theta|X_1,\ldots,X_n}(\theta \mid x_1, \ldots, x_n) = \boxed{\begin{cases} \frac{n-1}{(x_{\max}^{1-n} - 1)\,\theta^n} & \text{if } 0 < x_{\max} \leq \theta \leq 1 \\ 0 & \text{otherwise} \end{cases}}$$

G1[†]. (a) Let $X_k, k \geq 1$ be a set of i.i.d. Poisson random variables with mean 1. Then,

$$N^{(m)} = \sum_{k=1}^{m} X_k,\ m \geq 1.$$

By the central limit theorem,

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{\sum_{k=1}^{m}(X_k - \mathbf{E}[X_k])}{\sqrt{m}\sigma_{X_k}} \leq a \right) = \Phi(a) = \int_{-\infty}^{a} \frac{e^{-(x^2/2)}}{\sqrt{2\pi}} dx$$

where, since each X_k has $\mathbf{E}[X_k] = \sigma^2{}_{X_k} = 1$,

$$\mathbf{P}\left(\frac{\sum_{k=1}^{m}(X_k - \mathbf{E}[X_k]}{\sqrt{m}\sigma_{X_k}} \leq a \right) = \Phi(a) = \mathbf{P}\left(\frac{N^{(m)} - m}{\sqrt{m}} \leq a \right) = \mathbf{P}\left(N^{(m)} \leq m + a\sqrt{m} \right).$$

(b) Since $\frac{N^{(m)}-m}{\sqrt{m}}$ has a CDF that is approximately $N(\mu = 0, \sigma^2 = 1)$ for large m by the central limit theorem, $N^{(n)}$ must have a CDF that is approximately $N(\mu = n, \sigma^2 = n)$. therefore,

$$\mathbf{P}\left(N^{(n)} = n\right) \approx \int_{n-1/2}^{n+1/2} \frac{1}{\sqrt{2\pi n}} e^{-\frac{(x-n)^2}{2n}} dx \approx \int_{n-1/2}^{n+1/2} \frac{1}{\sqrt{2\pi n}} dx = \frac{1}{\sqrt{2\pi n}},$$

and, since $N^{(n)}$ is Poisson with mean n:

$$\mathbf{P}\left(N^{(n)} = n\right) = \frac{e^{-n} n^n}{n!}.$$

(c) Therefore,

$$\frac{e^{-n} n^n}{n!} \approx \frac{1}{\sqrt{2\pi n}},$$

which implies that

$$n! \approx (n/e)^n \sqrt{2\pi n}.$$

[†]Required for 6.431; optional for 6.041