2003 Spring Midterm for Advanced Probability for Communications

1. **(Variant of Bernstein's lemma)** Let $X_1, X_2, \ldots$ be discrete i.i.d. random variables taking values from $\{0, 1, 2\}$ with

$$\begin{cases} \Pr[X_1 = 0] = \dfrac{5 - 2x - \sqrt{9-4x}}{2}, \\ \Pr[X_1 = 1] = x - 3 + \sqrt{9-4x}, \\ \Pr[X_1 = 2] = \dfrac{3 - \sqrt{9-4x}}{2}. \end{cases}$$

Define

$$S_n = X_1 + \cdots + X_n \quad \text{and} \quad g_n(k, x) = \Pr\left[S_n = k\right].$$

Prove that for a real-valued function f that is continuous over a closed-and-bounded set $[0, 2]$, the function

$$B_n(x) = \sum_{k=1}^{2n} g_n(k, x)\, f\left(k/n\right)$$

converges to $f(x)$ uniformly on $[0, 2]$.

(Hint: Derive a bound for $|B_n(x) - f(x)|$ that is independent of x. The mean of X_1 is equal to x. Use Chebyshev's inequality and $\text{Var}[X_1] \leq 4$.)

Proof: Let

$$M = \sup_{x \in [0,2]} |f(x)| \quad \text{and} \quad \delta(\varepsilon) = \sup_{\{(x,y) \in [0,2] \times [0,2] : |x-y| \leq \varepsilon\}} |f(x) - f(y)|.$$

By the continuity of $f(\cdot)$ over a closed and bounded (hence, compact) set $[0, 2]$,

$$\lim_{\varepsilon \downarrow 0} \delta(\varepsilon) = 0.$$

For convenience, denote $\mathbb{N}_n = \{0, 1, 2, \ldots, 2n\}$. Then

$$\begin{aligned}
|B_n(x) - f(x)| &= |E[f(S_n/n)] - f(x)| \\
&= |E[f(S_n/n) - f(x)]| \\
&\leq E[|f(S_n/n) - f(x)|] \\
&= \sum_{k \in \mathbb{N}_n} |f(k/n) - f(x)| \Pr[S_n = k] \\
&= \sum_{\{k \in \mathbb{N}_n : |k/n - x| < \varepsilon\}} |f(k/n) - f(x)| \Pr[S_n = k] \\
&\quad + \sum_{\{k \in \mathbb{N}_n : |k/n - x| \geq \varepsilon\}} |f(k/n) - f(x)| \Pr[S_n = k] \\
&\leq \sum_{\{k \in \mathbb{N}_n : |k/n - x| < \varepsilon\}} \delta(\varepsilon) \Pr[S_n = k] + \sum_{\{k \in \mathbb{N}_n : |k/n - x| \geq \varepsilon\}} (2M) \Pr[S_n = k] \\
&= \delta(\varepsilon) \Pr[|S_n/n - x| < \varepsilon] + (2M) \Pr[|S_n/n - x| \geq \varepsilon] \\
&\leq \delta(\varepsilon) + (2M) \frac{\mathrm{Var}[X_1]}{n\varepsilon^2} \quad \text{(by Chebyshev's ineq.)} \\
&\leq \delta(\varepsilon) + (2M) \frac{E[X_1^2]}{n\varepsilon^2} \\
&\leq \delta(\varepsilon) + \frac{8M}{n\varepsilon^2},
\end{aligned}$$

for which the upper bound is independent of x (and hence, is a uniform bound). □

2. Suppose that we wish to estimate the parameter λ of an exponential random variable N with pdf $\lambda e^{-\lambda x} I_{[x \geq 0]}$ by successive observations $N_1, N_2, \ldots, N_n$ and a function $g(\cdot)$, where $g(\lambda) = E[f(N)]$ for some given $f(\cdot)$. We will always choose $f(\cdot)$ such that $g(\lambda)$ has inverse function at $\lambda \geq 0$. Then by the large deviation technique, we obtain:

$$\begin{aligned}
\Pr[|\hat{\boldsymbol{\lambda}}_n - \lambda| \geq \delta\lambda] &\leq \Pr\left\{ -\varepsilon_l \leq \frac{f(N_1) + f(N_2) + \cdots + f(N_n)}{n} - g(\lambda) \leq \varepsilon_h \right\} \\
&\leq \rho^n + \hat{\rho}^n,
\end{aligned}$$

where

$$\text{estimate } \hat{\boldsymbol{\lambda}}_n = g^{-1}\left(\frac{f(N_1) + f(N_2) + \cdots + f(N_n)}{n} \right),$$

$$\varepsilon_h = \max_{\{x\in\Re\,:\,|x-\lambda|\le\delta\lambda\}} g(x) - g(\lambda), \quad \varepsilon_l = g(\lambda) - \min_{\{x\in\Re\,:\,|x-\lambda|\le\delta\lambda\}} g(x),$$

$$\rho = \exp\left\{-I_{f(N)}\left(g(\lambda)+\varepsilon_h\right)\right\}, \quad \hat{\rho} = \exp\left\{-I_{f(N)}\left(g(\lambda)-\varepsilon_l\right)\right\},$$

and $I_{f(N)}(x) = \sup_{t\in\Re}[tx - C_{f(N)}(t)]$ and $C_{f(N)}(t)$ are respectively the large deviation rate function and the cumulant generating function of $f(N)$.

(a) Determine $-\log\rho$ and $-\log\hat{\rho}$ for $f(x) = x$.

(b) Determine $-\log\rho$ and $-\log\hat{\rho}$ for $f(x) = I_{[x\le a]}$ for a pre-specified constant a, where $I_{[\cdot]}$ is the indicator function. (Hint: Denote $e^{-a\lambda}$ by u, and use the binary divergence formula $D(p\|q) = p\log(p/q) + (1-p)\log[(1-p)/(1-q)]$ in your final answer.)

Answer:

(a) The moment generating function of N is:

$$\begin{aligned} M_N(t) &= E[e^{tN}] \\ &= \int_0^\infty e^{tu}\lambda e^{-\lambda u}du \\ &= \lambda\int_0^\infty e^{-(\lambda-t)u}du \\ &= \begin{cases} \dfrac{\lambda}{\lambda-t}, & \text{if } t<\lambda; \\ \infty, & \text{otherwise.} \end{cases} \\ \Rightarrow C_{N^2}(t) &= \begin{cases} \log(\lambda)-\log(\lambda-t), & \text{if } t<\lambda; \\ \infty, & \text{otherwise.} \end{cases} \end{aligned}$$

Then

$$\begin{aligned} I_N(x) &= \sup_{t\in\Re}[tx - C_N(t)] \\ &= \sup_{t<\lambda}\left[tx - \log(\lambda) + \log(\lambda-t)\right] \\ &= \begin{cases} \infty, & \text{if } x\le 0; \\ \lambda x - \log(\lambda x) - 1, & \text{if } x>0. \end{cases} \end{aligned}$$

In addition, $g(\lambda) = E[N] = 1/\lambda$. Hence, $\varepsilon_h = \frac{\delta}{1-\delta}(1/\lambda)$ and $\varepsilon_l = \frac{\delta}{1+\delta}(1/\lambda)$. Accordingly,

$$-\log\rho = I_N(1/\lambda+\varepsilon_h) = \log(1-\delta) + \frac{\delta}{1-\delta}$$

and

$$-\log\hat{\rho} \;=\; I_N(\lambda-\varepsilon_l) = \log(1+\delta) - \frac{\delta}{1+\delta}.$$

(b) The moment generating function of $I_{[N\le a]}$ is:

$$\begin{aligned} M(t) &= E[e^{tI_{[N\le a]}}] \\ &= \int_0^a e^t \lambda e^{-\lambda u} du + \int_a^\infty \lambda e^{-\lambda u} du \\ &= e^t(1-e^{-a\lambda}) + e^{-a\lambda}. \end{aligned}$$

Then

$$\begin{aligned} & I(x) \\ = \;& \sup_{t\in\Re}[tx - \log(M(t))] \\ = \;& \sup_{t\in\Re}\left[tx - \log\left(e^t(1-e^{-a\lambda}) + e^{-a\lambda}\right)\right] \\ = \;& \begin{cases} \infty, & \text{if } x\le 0 \text{ or } x\ge 1; \\ x\log\dfrac{x}{1-e^{-a\lambda}} + (1-x)\log\dfrac{1-x}{e^{-a\lambda}}, & \text{if } 0<x<1 \end{cases} \end{aligned}$$

In addition, $g(\lambda) = E[I_{[N\le a]}] = 1-e^{-a\lambda}$. Accordingly, for $u = e^{-a\lambda}$,

$$\begin{aligned} -\log\rho &= I(1-e^{-a\lambda(1+\delta)}) \\ &= (1-u^{1+\delta})\log\frac{1-u^{1+\delta}}{1-u} + u^{1+\delta}\log\frac{u^{1+\delta}}{u} \\ &= D(u^{1+\delta}\|u), \end{aligned}$$

and

$$\begin{aligned} -\log\hat{\rho} &= I(1-e^{-a\lambda(1-\delta)}) \\ &= (1-u^{1-\delta})\log\frac{1-u^{1-\delta}}{1-u} + u^{1-\delta}\log\frac{u^{1-\delta}}{u} \\ &= D(u^{1-\delta}\|u). \end{aligned}$$

3. Construct an example that X_n has density $f_{X_n}(x)$ for every n, X has density $f(x)$, and $X_n \Rightarrow X$; however, $f_{X_n}(x)$ does not converge to

$f_X(x)$ for those x in the support of X.
(Hint: You may focus your construction on a uniform X over $(0,1)$, whose density is 1. Let $f_{X_n}(x) = 1 + \delta_n(x)$ for some can-be-negative $\delta_n(x)$, where $\delta_n(x) \geq -1$, but

$$\int_0^x \delta_n(t)dt = \int_0^x \max\{\delta_n(t), 0\}dt + \int_0^x \min\{\delta_n(t), 0\}dt$$

vanishes for every x.)

Answer: $\delta_n(x) = \cos(2\pi n x)$. □

4. (a) Prove that for any sequence of non-negative random variable $\{Y_n\}_{n=1}^\infty$ with finite positive mean $E[Y_n]$ satisfying:

$$\lim_{n\to\infty} \frac{E[|Y_n - E[Y_n]|^k]}{|E[Y_n]|^k} = 0 \text{ for some } k > 0, \tag{1}$$

the mean-normalized sequence $\{Y_n/E[Y_n]\}_{n=1}^\infty$ converges in distribution to X, where $\Pr[X = 1] = 1$.
(Hint: Markov's inquality)

(b) Show that if Y_n is Poisson distributed with mean n, then $\{Y_n\}_{n=1}^\infty$ satisfies (1) with $k = 2$.
(Hint: To save your time, you may use the fact directly that the mean and variance of a Poisson random variable are equal to each other.)

(c) Create a sequence of random variables $Y_1, Y_2, \ldots$ that satisfy (1) for every $k > 0$. (Hint: $Y_n/E[Y_n]$ must converge in distribution to X.)

Proof:

(a) By Markov's inequality,

$$\begin{aligned} \Pr\left[\left|\frac{Y_n}{E[Y_n]} - 1\right| \geq \varepsilon\right] &= \Pr\left[|Y_n - E[Y_n]| \geq \varepsilon\,|E[Y_n]|\right] \\ &\leq \frac{E[|Y_n - E[Y_n]|^k]}{\varepsilon^k |E[Y_n]|^k}. \end{aligned}$$

Hence, $\{Y_n/E[Y_n]\}_{n=1}^\infty$ converges in distribution to X, where $\Pr[X = 1] = 1$.

(b) It can be checked that

$$\frac{E[|Y_n - E[Y_n]|^2]}{|E[Y_n]|^2} = \frac{n}{n^2} = \frac{1}{n} \to 0.$$

(c) Any example with $\Pr[Y_n = E[Y_n]] = 1$ for every n and $E[Y_n] \to \infty$ as $n \to \infty$. □

5. (a) Prove that $E[X] = \int_0^\infty \Pr[X > t]dt$ for non-negative bounded random variable X with density $f(x)$. (Hint: Fubini's Theorem.)

   (b) Using (a) to prove that $E[X] = \int_0^\infty \Pr[X > t]dt$ for non-negative (possibly unbounded) random variable X with density $f(x)$. (Hint: Define $Y_n = XI_{[x \le n]}$, and Fatou's lemma tells that

$$E[X] \le \liminf_{n\to\infty} E[Y_n] \quad \text{and} \quad \int_0^\infty \Pr[X > t]dt \le \liminf_{n\to\infty} \int_0^\infty \Pr[Y_n > t]dt.)$$

Proof:

(a) Denote by M the bound for random variable X, i.e., $\Pr[X \le M] = 1$. By Fubini's Theorem (the condition for Fubini's Theorem is always valid for bounded integration region),

$$\begin{aligned}
\int_0^\infty \Pr[X > t]dt &= \int_0^M \Pr[X > t]dt \\
&= \int_0^M \left(\int_t^M f(x)dx \right) dt \\
&= \int_0^M \left(\int_0^x f(x)dt \right) dx \\
&= \int_0^M xf(x)dx \\
&= E[X].
\end{aligned}$$

(b) Define $Y_n = XI_{[X \le n]}$. Then $\Pr[Y_n > t] \to \Pr[X > t]$ as $n \to \infty$, or equivalently, $Y_n \Rightarrow X$.

By Fatou's Lemma,

$$E[Y_n] \le E[X] \le \liminf_{n\to} E[Y_n].$$

Therefore, $E[X] = \lim_{n\to\infty} E[Y_n]$.
On the other hand,

$$\int_0^\infty \Pr[Y_n > t]dt \le \int_0^\infty \Pr[X > t]dt \le \liminf_{n\to\infty} \int_0^\infty \Pr[Y_n > t]dt,$$

which implies that

$$\int_0^\infty \Pr[X > t]dt = \lim_{n\to\infty} \int_0^\infty \Pr[Y_n > t]dt.$$

The proof is completed by noting that $E[Y_n] = \int_0^\infty \Pr[Y_n > t]dt$ from (a). □

6. A random variable has a *lattice distribution* if for some a and b, where $b \neq 0$, the set $\{a + nb \ : \ n = 0, \pm 1, \pm 2, \ldots\}$ supports the distribution of X.

   (a) Prove that X has a lattice distribution, if its characteristic function $|\varphi_X(t)| = 1$ for some $t \neq 0$.
   (Hint: A complex value $\varphi_X(t_0)$ satisfying $|\varphi_X(t_0)| = 1$ can be represented as $\varphi_X(t_0) = e^{it_0a}$ for some a. That a non-negative function like $[1 - \cos(it_0(x - a))]$ has integral zero implies that the function is exactly zero in the integration domain.)

   (b) Prove that if X has a lattice distribution, then its characteristic function $|\varphi_X(t)| = 1$ for some $t \neq 0$.
   (Hint: Let $b = 2\pi/t_0$.)

Proof:

(a) If $|\varphi_X(t_0)| = 1$ for some $t_0 \neq 0$, then $\varphi_X(t_0) = e^{it_0a}$ for some a.

$$\begin{aligned} 0 &= 1 - e^{-it_0a}\varphi_X(t_0) \\ &= \int_{-\infty}^\infty dF_X(x) - e^{-it_0a}\int_{-\infty}^\infty e^{it_0x}dF_X(x) \\ &= \int_{-\infty}^\infty (1 - e^{it_0(x-a)})dF_X(x) \\ &= \int_{-\infty}^\infty [1 - \cos(it_0(x-a))]dF_X(x) + i\int_{-\infty}^\infty \sin(it_0(x-a))dF_X(x). \end{aligned}$$

As $[1-\cos(it_0(x-a))] \geq 0$ and $\int_{-\infty}^{\infty}[1-\cos(it_0(x-a))]dF_X(x) = 0$, we conclude that

$$\Pr[\cos(it_0(X-a)) = 1] = 1,$$

which implies that

$$\Pr[t_0(X-a) = 2n\pi \text{ for } n = 0, \pm 1, \pm 2, \ldots] = 1.$$

(b) Let $b = 2\pi/t_0$.

$$\begin{aligned} \varphi_X(t_0) &= \sum_{n=0,\pm 1,\pm 2,\ldots} e^{it_0(a+nb)} p_n \\ &= \sum_{n=0,\pm 1,\pm 2,\ldots} e^{it_0(a+2\pi n/t_0)} p_n \\ &= e^{it_0 a} \sum_{n=0,\pm 1,\pm 2,\ldots} e^{i2\pi n} p_n \\ &= e^{it_0 a}. \end{aligned}$$

So $|\varphi_X(t_0)| = 1$. □