Probability Qualifying Review Exam
Friday May 24, 2013
9a.m. to 1p.m.

Instructions:

- You are allowed only pen (or pencil) and paper. Calculators, cell phones, and any other electronic device should not be brought into the exam room.
- You may answer up to seven of the eight questions. Each question carries equal credit. If you submit solutions to all eight questions then only the first seven will be graded.
- Questions 1-4 cover material from Stat 620. Questions 5-8 cover material from Stat 621.

1. This question studies a random variable with a random probability mass function, constructed as follows. Let $Z_1, \ldots, Z_N$ be independent, identically distributed, positive random variables. Let

$$W_n = \frac{Z_n}{\sum_{m=1}^{N} Z_m} \qquad\qquad \text{for } n = 1, \ldots N.$$

Let $x_1, \ldots, x_N$ be a sequence of real numbers. Let X be a random variable defined such that $\mathbb{P}(X = x_n \,|\, W_1, \ldots, W_N) = W_n$.

(a) Find $\mathbb{E}[W_n]$ and $\mathbb{E}[W_n \,|\, W_m]$ for $m \neq n$. Hint: you can appeal to symmetry to argue that the answer should not depend on n or m.

Solution: Note that $\sum_{n=1}^{N} W_n = 1$, so $\sum_{n=1}^{N} \mathbb{E}[W_n] = 1$. By symmetry, we get $\mathbb{E}[W_n] = 1/N$. Similarly, $\mathbb{E}[\sum_{n \neq m} W_n \,|\, W_m] = 1 - W_m$. By symmetry we have $(N-1)\mathbb{E}[W_n \,|\, W_m] = 1 - W_m$, so $\mathbb{E}[W_n \,|\, W_m] = (1 - W_m)/(N-1)$.

(b) Let $\mathrm{Var}(W_n) = \alpha$. Find $\mathrm{Cov}(W_m, W_n)$ in terms of α, for $m \neq n$.

Solution:

$$\begin{aligned}
\mathbb{E}[W_m W_n] &= \mathbb{E}\big[W_m \mathbb{E}[W_n \,|\, W_m]\big] \\
&= \mathbb{E}[W_m(1 - W_m)/(N-1)] \\
&= \frac{1}{N(N-1)} - \frac{\mathrm{Var}(W_m) + \mathbb{E}[W_m]^2}{N-1} \\
&= \frac{1}{N^2} - \frac{\alpha}{N-1}
\end{aligned}$$

Therefore, $\mathrm{Cov}(W_m, W_n) = \mathbb{E}[W_m W_n] = \mathbb{E}[W_m]\mathbb{E}[W_n] = -\alpha/(N-1)$. Another way to get this is to note that $\mathrm{Var}\big(\sum_{n=1}^{N} W_n\big) = 0$ and then, by symmetry, we get $N\mathrm{Var}(W_n) + N(N-1)\mathrm{Cov}(W_m, W_n) = 0$.

(c) Find an expression for $\mathbb{E}[\mathrm{Var}(X \,|\, W_1, \ldots, W_N)]$ in terms of $x_1, \ldots, x_N$ and α. Check whether your expression behaves appropriately in two situations; (i) when $\alpha = 0$, and (ii) when $x_1, x_2, \ldots, x_N$ is replaced by $x_1 + c, x_2 + c, \ldots, x_N + c$ for some constant c.

Solution:

$$\begin{aligned}
\operatorname{Var}(X \mid W_1, \ldots, W_N) &= \sum_{n=1}^N x_n^2 W_n - \Big(\sum_{n=1}^N x_n W_n\Big)^2, \\
\mathbb{E}\big[\operatorname{Var}(X \mid W_1, \ldots, W_N)\big] &= \sum_{n=1}^N x_n^2 \mathbb{E}[W_n] - \sum_{m=1}^N \sum_{n=1}^N x_m x_n \mathbb{E}[W_m W_n] \\
&= \frac{1}{N} \sum_{n=1}^N x_n^2 - \sum_{n=1}^N \left[\alpha + \frac{1}{N^2}\right] - 2 \sum_{m<n} x_m x_n \left(\frac{1}{N^2} - \frac{\alpha}{N-1}\right) \\
&= \frac{1}{N} \sum_{n=1}^N x_n^2 - \Big(\frac{1}{N} \sum_{n=1}^N x_n\Big)^2 - \alpha \left(\sum_{n=1}^N x_n^2 - 2 \sum_{m<n} \frac{x_m x_n}{N-1}\right) \\
&= \frac{1}{N} \sum_{n=1}^N x_n^2 - \Big(\frac{1}{N} \sum_{n=1}^N x_n\Big)^2 - \alpha \sum_{n=1}^N x_n \left[x_n - \frac{1}{N-1} \sum_{m:m \neq n} x_m\right] \\
&= s^2 - \alpha \sum_{n=1}^N \left[\Big(x_n - \frac{1}{N} \sum_{m=1}^N x_m\Big)\Big(x_n - \frac{1}{N-1} \sum_{m:m \neq n} x_m\Big)\right], \qquad (1)
\end{aligned}$$

where s^2 is the variance of the random variable with probability mass $1/N$ at each of $x_1, \ldots, x_N$. When $\alpha = 0$, this gives the sample variance of $x_1, \ldots, x_N$. We see, by writing the expression in the form (1), that adding a constant to $x_1, \ldots, x_N$ does not change the answer.

2. Let $\{X(t), 0 \le t \le T\}$ be a Brownian motion with initial distribution $X(0) \sim \text{Normal}(\nu, \tau^2)$. In other words, $\{X(t)\}$ has continuous sample paths and stationary independent increments having distribution $X(t+s) - X(t) \sim \text{Normal}(0, \sigma^2 s)$ for $s > 0$. Let $Y(t) = X(T-t)$ for $0 \le t \le T$.

   (a) Explain why $\{Y(t), 0 \le t \le T\}$ is a diffusion process.

   Solution: $\{Y(t), 0 \le t \le T\}$ inherits continuous sample paths from $\{X(t), 0 \le t \le T\}$. In the symmetric form (that the past and future are conditionally independent given the present) the Markov property for the time-reversed process $\{Y(t), 0 \le t \le T\}$ follows immediately from the corresponding property for $\{X(t), 0 \le t \le T\}$.

   (b) Obtain the infinitesimal parameters of $\{Y(t)\}$.

   Solution:

$$\begin{pmatrix} Y(t) \\ Y(t+\delta) \end{pmatrix} \sim \text{Normal}\left[\begin{pmatrix} \nu \\ \nu \end{pmatrix}, \begin{pmatrix} \sigma^2(T-t) + \tau^2 & \sigma^2(T-t-\delta) + \tau^2 \\ \sigma^2(T-t-\delta) + \tau^2 & \sigma^2(T-t-\delta) + \tau^2 \end{pmatrix}\right], \qquad (2)$$

from which we obtain

$$\begin{aligned}
\mathbb{E}[Y(t+\delta) \mid Y(t)] &= \nu + \frac{\sigma^2(T-t-\delta) + \tau^2}{\sigma^2(T-t) + \tau^2}(Y(t) - \nu) \\
\mathbb{E}[Y(t+\delta) - Y(t) \mid Y(t)] &= (\nu - Y(t)) \left[1 - \frac{\sigma^2(T-t-\delta) + \tau^2}{\sigma^2(T-t) + \tau^2}\right] \\
&= \frac{\delta(\nu - Y(t))}{\sigma^2(T-t) + \tau^2}
\end{aligned}$$

$$
\begin{aligned}
\mathrm{Var}[Y(t+\delta)\,|\,Y(t)] &= \sigma^2(T-t-\delta)+\tau^2-\frac{\{\sigma^2(T-t-\delta)+\tau^2\}^2}{\sigma^2(T-t)+\tau^2} \\
&= \frac{[\sigma^2(T-t-\delta)+\tau^2][\sigma^2(T-t)+\tau^2]-[\sigma^2(T-t-\delta)+\tau^2]^2}{\sigma^2(T-t)+\tau^2} \\
&= \delta\sigma^2+o(\delta)
\end{aligned}
$$

Thus, the infinitesimal mean is $\mu(y,t)=(\nu-y)/[\sigma^2(T-t)+\tau^2]$ and the infinitesimal variance is σ^2.

(c) In the special case with $\nu=0$, $\tau=0$ and $T=1$, comment on how the behavior of $\{Y(t), 0\le t\le T\}$ relates to a well-known diffusion process.

Solution: In this case, the infinitesimal parameters match those of the Brownian bridge. $\{Y(t), 0\le t\le T\}$ is not quite the usual Brownian bridge, since $Y(0)\sim \mathrm{Normal}(\nu,\sigma^2T+\tau^2)$. However, if you condition $\{Y(t), 0\le t\le T\}$ on $\{Y(0)=0\}$, you get exactly a Brownian bridge.

3. We wish to show that not every discrete time Markov chain can be embedded in a continuous time Markov chain. To formalize this, let $\{X_n, n=0,1,2,\dots\}$ be a discrete time Markov chain with states $\{1,2,\dots,K\}$ having transition probabilites P_{ij}, and let $\{X(t), t\ge 0\}$ be a continuous time Markov chain with states $\{1,2,\dots,K\}$ having transition rates q_{ij}. We say that $\{X_n, n=0,1,2,\dots\}$ is embedded in $\{X(t), t\ge 0\}$ if $\{X(n), n=0,1,2,\dots\}$ has the same distribution as $\{X_n, n=0,1,2,\dots\}$.

   Specifically, you are asked to consider the case $K=2$ with $\{X_n\}$ having a symmetric transition probability matrix,

$$
P=[P_{ij}]=\begin{pmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{pmatrix} \qquad \text{for } 0<\alpha<1.
$$

   For what values of α does an embedding exist?

   Solution: The initial distribution can be assigned to be the same for X_0 and $X(0)$, so we need to match the transition probabilities. A general Q-matrix for $\{X(t)\}$ is

$$
Q=\begin{pmatrix} -\beta & \beta \\ \gamma & -\gamma \end{pmatrix}.
$$

   Intuitively, it may seem necessary that $\beta=\gamma$ since the embedded process is symmetric. One way to check this is to match the limiting distributions. Solving detailed balance gives $\lim_{n\to\infty}\mathbb{P}(X_n=1)=1/2$ and $\lim_{t\to\infty}\mathbb{P}(X(t)=1)=\gamma/(\beta+\gamma)$, implying $\gamma=\beta$. Then, the eigenvalues of Q are 0 and -2β, and the transition probabilities are therefore a linear combination of 1 and $e^{-2\beta t}$, which can be computed as

$$
P(t)=e^{Qt}=\begin{pmatrix} (1+e^{-2\beta t})/2 & (1-e^{-2\beta t})/2 \\ (1-e^{-2\beta t})/2 & (1+e^{-2\beta t})/2 \end{pmatrix}.
$$

   The requirement that $P(1)=P$ gives $\beta=-\frac{1}{2}\log(2\alpha-1)$, which has a solution only when $\alpha>1/2$.

4. Consider a partially observed discrete stochastic process $\{(X_n, Y_n), n = 1, 2, \dots\}$, where $\{Y_n, n = 1, 2, \dots\}$ is observed and $\{X_n, n = 1, 2, \dots\}$ is unobserved. Suppose that, marginally, $\{X_n\}$ is a Markov chain taking values in a finite set $\mathbb{X}$. Suppose $\{Y_n\}$ takes values in a finite set $\mathbb{Y}$. To model Y_n as a noisy observation of X_n, we suppose $\{Y_n\}$ has a conditional independence property that

$$\mathbb{P}(Y_n = j_n \mid X_1 = i_1, \dots, X_n = i_n, Y_1 = j_1, \dots, Y_{n-1} = j_{n-1}) = \mathbb{P}(Y_n = j_n \mid X_n = i_n).$$

We are interested in computing the following quantities, concerning the conditional distribution of $\{X_n\}$ given observed values $Y_1 = j_1, Y_2 = j_2, \dots$

$$\begin{aligned} C(i,n) &= \mathbb{P}(X_n = i \mid Y_1 = j_1, Y_2 = j_2, \dots, Y_n = j_n) \\ P(i,n) &= \mathbb{P}(X_n = i \mid Y_1 = j_1, Y_2 = j_2, \dots, Y_{n-1} = j_{n-1}) \\ S(i,n) &= \mathbb{P}(X_n = i \mid Y_1 = j_1, Y_2 = j_2, \dots, Y_N = j_N) \quad \text{for } N > n \end{aligned}$$

Here, $C(i,n)$ is current state estimation given observations up to time n; $P(i,n)$ is prediction of the state at time n using observations up to time $n-1$; $S(i,n)$ is so-called smoothed estimation of the state at time n using observations up to time N.

(a) Prove the following recursions:

$$\begin{aligned} C(i,n) &= \frac{P(i,n)\,\mathbb{P}(Y_n = j_n \mid X_n = i)}{\sum_{k\in\mathbb{X}} P(k,n)\,\mathbb{P}(Y_n = j_n \mid X_n = k)} \\ P(i,n+1) &= \textstyle\sum_{k\in\mathbb{X}} C(k,n)\,\mathbb{P}(X_n = i \mid X_{n-1} = k) \end{aligned}$$

Solution:

$$\begin{aligned} C(i,n) &= \mathbb{P}(X_n = i \mid Y_1 = j_1, Y_2 = j_2, \dots, Y_n = j_n) \\ &= \frac{\mathbb{P}(Y_n = j_n \mid X_n = i, Y_1 = j_1, Y_2 = j_2, \dots, Y_{n-1} = j_{n-1})\mathbb{P}(X_n = i \mid Y_1 = j_1, \dots, Y_{n-1} = j_{n-1})}{\mathbb{P}(Y_n = j_n \mid Y_1 = j_1, Y_2 = j_2, \dots, Y_{n-1} = j_{n-1})} \\ &= \frac{\mathbb{P}(Y_n = j_n \mid X_n = i)P(i,n)}{\sum_k \mathbb{P}(Y_n = j, X_n = k \mid Y_1 = j_1, Y_2 = j_2, \dots, Y_{n-1} = j_{n-1})} \\ &= \frac{\mathbb{P}(Y_n = j_n \mid X_n = i)P(i,n)}{\sum_k \mathbb{P}(Y_n = j_n \mid X_n = k, Y_1 = j_1, \dots, Y_{n-1} = j_{n-1})\mathbb{P}(X_n = k \mid Y_1 = j_1, Y_2 = j_2, \dots, Y_{n-1} = j_{n-1})} \\ &= \frac{\mathbb{P}(Y_n = j_n \mid X_n = i)P(i,n)}{\sum_k \mathbb{P}(Y_n = j_n \mid X_n = k)P(k,n)} \end{aligned}$$

$$\begin{aligned} P(i,n+1) &= \mathbb{P}(X_{n+1} = i \mid Y_1 = j_1, Y_2 = j_2, \dots, Y_n = j_n) \\ &= \sum_k \mathbb{P}(X_{n+1} = i \mid X_n = k, Y_1 = j_1, Y_2 = j_2, \dots, Y_n = j_n)\mathbb{P}(X_n = k \mid Y_1 = j_1, \dots, Y_n = j_n) \qquad (3) \\ &= \sum_k \mathbb{P}(X_{n+1} = i \mid X_n = k)\mathbb{P}(X_n = k \mid Y_1 = j_1, Y_2 = j_2, \dots, Y_n = j_n) \qquad (4) \\ &= \sum_k C(n,k)\mathbb{P}(X_{n+1} = i \mid X_n = k) \end{aligned}$$

where (4) follows from (3) by checking that

$$\mathbb{P}(X_{n+1} = i \mid X_n = k, Y_1 = j_1, Y_2 = j_2, \dots, Y_n = j_n) = \mathbb{P}(X_{n+1} = i \mid X_n = k).$$

(b) Prove the relationship

$$S(i,n) = \frac{P(i,n)R(i,n)}{\sum_{k\in\mathbb{X}} P(k,n)R(k,n)},$$

where

$$R(i,n) = \mathbb{P}(Y_n = j_n, Y_{n+1} = j_{n+1}, \dots Y_N = j_N \,|\, X_n = i).$$

Solution:

$$\begin{aligned}
&S(i,n) = \mathbb{P}(X_n = i \,|\, Y_1 = j_1, Y_2 = j_n, \dots, Y_N = j_N) \\
&= \frac{\mathbb{P}(Y_n = j_n, \dots, Y_N = j_N \,|\, X_n = i, Y_1 = j_1, \dots, Y_{n-1} = j_{n-1})\mathbb{P}(X_n = i \,|\, Y_1 = j_1, \dots, Y_{n-1} = j_{n-1})}{\sum_k \mathbb{P}(Y_n = j_n, \dots, Y_N = j_N \,|\, X_n = k, Y_1 = j_1, \dots, Y_{n-1} = j_{n-1})\mathbb{P}(X_n = k \,|\, Y_1 = j_1, \dots, Y_{n-1} = j_{n-1})} \\
&= \frac{\mathbb{P}(Y_n = j_n, \dots, Y_N = j_N \,|\, X_n = i)\mathbb{P}(X_n = i \,|\, Y_1 = j_1, \dots, Y_{n-1} = j_{n-1})}{\sum_k \mathbb{P}(Y_n = j_n, \dots, Y_N = j_N \,|\, X_n = k)\mathbb{P}(X_n = k \,|\, Y_1 = j_1, \dots, Y_{n-1} = j_{n-1})} \\
&= \frac{R(n,i)P(n,i)}{\sum_k R(n,k)P(n,k)}
\end{aligned}$$

5. Let $X_n,\ n = 1, 2, \cdots$ be independent, identically distributed random variables.

(a) If $X_1 \in L^p,\ p \geq 1$, then show that

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow{L^p} \mathbb{E}(X_1), \qquad \text{as } n \to \infty.$$

(b) If X_n/n converges almost surely to zero, as $n \to \infty$, then show that $\mathbb{E}|X_1| < \infty$.

Solution: **(a):** By the SLLN, we have that $S_n/n \to \mu$, almost surely, where $\mu = \mathbb{E}(X_1)$. The L^p convergence then holds if and only if $\{|S_n/n|^p,\ n \geq 1\}$ is u.i. Observe that since $x \mapsto |x|^p$, is convex for $p \geq 1$, then the Jensen's inequality implies:

$$\mathbb{E}|S_n/n|^p \leq \sum_{k=1}^n \mathbb{E}|X_k|^p/n = \mathbb{E}|X_1|^p.$$

This shows that $\sup_{n\geq 1} \mathbb{E}|S_n/n|^p < \infty$. To prove desired the u.i. it is enough to show that for all $\epsilon > 0$, there exists $\delta > 0$, such that

$$\mathbb{E}\Big(|S_n/n|^p 1_A\Big) \leq \epsilon, \quad \text{for all events } A \text{ with } P(A) < \delta.$$

Note, as above, that by Jensen's

$$\mathbb{E}\Big(|S_n/n|^p 1_A\Big) \leq \frac{1}{n}\sum_{k=1}^n \mathbb{E}(|X_k|^p 1_A) \leq \sup_{k\geq 1} \mathbb{E}|X_k|^p 1_A.$$

Since the X_k's are iid and belong to L^p, it follows that $\{|X_k|^p,\ k \geq 1\}$ is u.i. and the latter supremum (by definition) can be made to be less than $\epsilon > 0$, provided $\delta > 0$ is sufficiently small.

(b): For any fixed $\epsilon > 0$, we have that $|X_n/n| < \epsilon$, eventually. That is, $\mathbb{P}(|X_n/n| > \epsilon,\ i.o.) = 0$. But since the X_n's are independent, the Borel zero-one law implies that

$$\sum_{n=1}^{\infty} \mathbb{P}(|X_n/n| > \epsilon) < \infty.$$

Since the X_n's are identically distributed and by taking $\epsilon := 1$, we then have that

$$\mathbb{E}|X_1| \le \mathbb{E}\sum_{n=0}^{\infty} 1_{|X_1|>n} = \sum_{n=0}^{\infty} \mathbb{P}(|X_n/n| > 1) < \infty.$$

$\square$

6. Let $X_n,\ n = 1, 2, \cdots$ be independent random variables such that $X_n \sim \text{Gamma}(\alpha_n, 1),\ \alpha_n > 0$. That is, the probability density of X_n is

$$f_{X_n}(x) = \begin{cases} x^{\alpha_n - 1}e^{-x}/\Gamma(\alpha_n) & ,\ \text{if } x > 0 \\ 0 & ,\ \text{otherwise,} \end{cases}$$

where $\Gamma(\alpha) := \int_0^\infty x^{\alpha-1}e^{-x}dx,\ \alpha > 0$ is the Euler gamma function.

(a) If $\sum_{n=1}^\infty \alpha_n < \infty$, then show that the series $\sum_{n=1}^\infty X_n$ converges almost surely and in the L^2 sense.

(b) If $\sum_{n=1}^\infty \alpha_n = \infty$, then show that

$$\frac{S_n - \mathbb{E}(S_n)}{\sqrt{\text{Var}(S_n)}} \Longrightarrow \text{Normal}(0, 1), \quad \text{as } n \to \infty,$$

where $S_n = \sum_{k=1}^n X_k$.

<u>Hint:</u> Although it would in principle be possible to use a Lindeberg-Feller argument, you are not advised to follow that approach.

<u>Solution</u>: **(a):** Recall that $\mathbb{E}(X_k) = \alpha_k$ and $\text{Var}(X_k) = \alpha_k$. Therefore, $\sum_{n=1}^\infty \text{Var}(X_n) < \infty$ and the Kolmogorov's convergence criterion implies that

$$\sum_{n=1}^{\infty}(X_n - \mathbb{E}(X_n))$$

converges almost surely. Since however $\sum_{n=1}^\infty \mathbb{E}(X_n) = \sum_{n=1}^\infty \alpha_n < \infty$, it also follows that $\sum_{n=1}^\infty X_n$ converges almost surely.

(b): Note that $S_n \sim \text{Gamma}(v_n, 1)$, where $v_n = \sum_{k=1}^n \alpha_k \to \infty$, as $n \to \infty$. Let $\xi_k,\ k \ge 1$ be iid $\text{Gamma}(1, 1)$ and observe that

$$S_n \stackrel{d}{=} \widetilde{S}_n + R_n,$$

where $\widetilde{S}_n = \xi_1 + \cdots + \xi_{[v_n]}$ and $R_n \sim \text{Gamma}(\{v_n\}, 1)$ is independent of the ξ_k's.

By the classic CLT, we have that $(\widetilde{S}_n - \mathbb{E}(\widetilde{S}_n))/\sqrt{\mathrm{Var}(\widetilde{S}_n)} \Rightarrow \mathcal{N}(0,1)$. Note also that $\mathbb{E}\widetilde{S}_n = [v_n]$ and $\mathrm{Var}(\widetilde{S}_n) = [v_n]$. Therefore, it is easy to see that by Slutsky's theorem, it is enough to show that $R_n/\sqrt{[v_n]} \to 0$, in probability. We have that

$$\mathbb{E}|R_n| = \mathbb{E}(R_n) = \int_0^\infty x^{\{v_n\}} e^{-x} dx/\Gamma(\{v_n\}) = \Gamma(1+\{v_n\})/\Gamma(\{v_n\}) = \{v_n\} \le 1.$$

This implies $R_n/\sqrt{[v_n]} \to 0,\ n \to \infty$, in L^1-sense and hence in probability. This completes the proof.
□

7. Let $X_0 = 1$ and define the random variables $X_n,\ n \ge 1$ recursively as follows. Introduce the σ-algebra $\mathcal{F}_n := \sigma(X_0, \cdots, X_n),\ n \ge 0$. Conditionally on $\mathcal{F}_n$, the random variable X_{n+1} is uniformly distributed in the interval $(-S_n, S_n)$, where $S_n = X_0 + \cdots + X_n,\ n \ge 0$.

   (a) Show that $\{S_n,\ n \ge 1\}$ is a martingale and compute $\mathbb{E}(S_n)$ and $\mathrm{Var}(S_n)$.

   (b) Show that $S_n > 0$ almost surely, and that for some positive sequences a_n and b_n, the random variables $S_n^{a_n}/b_n$ converge in distribution to a non-trivial limit, as $n \to \infty$. Identify explicitly the sequences $\{a_n\}$ and $\{b_n\}$, and the limit distribution.

   Hint: Express S_n in terms of independent random variables.

   Solution: **(a):** Observe that $X_{n+1} = S_n U_{n+1}$, where $U_{n+1} \sim U(-1,1)$ is independent of $X_0, X_1, \cdots, X_n$. Therefore,

$$S_{n+1} = S_n(1+U_{n+1}) = \prod_{k=1}^{n+1}(1+U_k),$$

where $U_k,\ k = 1, 2, \cdots$ are iid $U(-1,1)$. We then have that

$$\mathbb{E}(S_{n+1}|\mathcal{F}_n) = S_n \mathbb{E}(1+U_{n+1}) = S_n,$$

showing that $S_n,\ n \ge 0$ is a martingale.

Since $S_0 = 1$, and since $\{S_n,\ n \ge 0\}$ is a martingale, we have that $\mathbb{E}(S_n) = 1$. By the smoothing property and independence, we also have that

$$\mathbb{E}(S_{n+1}^2) = \mathbb{E}\mathbb{E}(S_{n+1}^2|\mathcal{F}_n) = \mathbb{E}(S_n^2)\mathbb{E}(1+U_{n+1})^2.$$

Note that $1+U_{n+1} \sim U(0,2)$, and therefore $\mathbb{E}(1+U_{n+1})^2 = \int_0^2 x^2 dx/2 = 4/3$. Applying the last relation recursively, we obtain

$$\mathbb{E}(S_n^2) = (4/3)^n$$

and hence $\mathrm{Var}(S_n) = \mathbb{E}(S_n^2) - (\mathbb{E}S_n)^2 = (4/3)^n - 1,\ n \ge 0$.

(b): The above representation shows that $\mathbb{P}(S_n > 0) = 1$ since $\mathbb{P}(U_n > -1) = 1,\ n \geq 1$. By taking logarithms, we have

$$\log(S_n) = \sum_{k=1}^{n} \log(1 + U_k).$$

Observe that $\log(2) - \log(1 + U_k) =: X_k \sim \text{Exp}(1)$ are standard iid exponentially distributed random variables. Therefore, the classic CLT for $\sum_{k=1}^n X_k$ implies that

$$(\log(S_n) - \mathbb{E}(\log(S_n)))/\sqrt{\text{Var}(\log(S_n))} \Rightarrow \mathcal{N}(0,1), \quad \text{as } n \to \infty.$$

Note that $\mathbb{E}(\log(S_n)) = n(\log(2) - 1)$ and $\text{Var}(\log(S_n)) = n$. The last relation along with the CMT applied to the mapping $x \mapsto e^x$ implies

$$S_n^{1/\sqrt{n}} e^{\sqrt{n}(1-\log(2))} \Longrightarrow e^Z, \quad \text{as } n \to \infty,$$

where $Z \sim \mathcal{N}(0,1)$. That is, for the sequences a_n and b_n we have

$$a_n = 1/\sqrt{n} \quad \text{and} \quad b_n = e^{\sqrt{n}(\log(2)-1)},$$

and the limit distribution is log-Normal. $\square$

8. Suppose that $Y_k,\ k = 1, 2, \cdots$ are independent and identically distributed random variables such that $\mathbb{E}(Y_1) = 0$ and $\text{Var}(Y_1) = 1$. Let $\{X_k,\ k = 1, 2, \cdots\}$ be a sequence of positive random variables, independent of $\{Y_k, k = 1, 2, \cdots\}$, such that

$$\frac{\max_{k=1,\cdots,n} X_k^2}{\sum_{j=1}^n X_j^2} \xrightarrow{a.s.} 0, \quad \text{as } n \to \infty.$$

(a) Let

$$Z_n := \frac{\sum_{k=1}^n X_k Y_k}{\sqrt{\sum_{j=1}^n X_j^2}}$$

and calculate $\mathbb{E}(Z_n)$ and $\text{Var}(Z_n)$.

(b) Let $\mathcal{F}_X := \sigma(X_k,\ k \geq 1)$ and show that for all bounded and continuous functions f, we have

$$\mathbb{E}(f(Z_n)|\mathcal{F}_X) \longrightarrow \mathbb{E}f(Z), \quad \text{almost surely},$$

as $n \to \infty$, where Z is a standard Normal random variable.

(c) Using part (b), argue that

$$Z_n \Longrightarrow \text{Normal}(0,1), \quad \text{as } n \to \infty.$$

Solution: **(a):** By the smoothing property of conditional expectation, we have that

$$\mathbb{E}Z_n = \mathbb{E}\mathbb{E}(Z_n|\mathcal{F}_X) = \mathbb{E}(0) = 0.$$

Similarly

$$\mathrm{Var}(Z_n) = \mathbb{E}(Z_n^2) = \mathbb{E}\mathbb{E}(Z_n^2|\mathcal{F}_X) = \mathbb{E}\Big(\frac{1}{\sum_{k=1}^n X_k^2}\mathbb{E}(\sum_{j=1}^n Y_j X_j)^2|\mathcal{F}_X\Big) = \mathbb{E}(1) = 1.$$

(b): Without loss of generality, we will suppose that the random variables are defined on a product probability space, so that $\omega = (\omega_x, \omega_y) \in \Omega_X \times \Omega_Y$, where $\mathbb{P}(d\omega) = \mathbb{P}_X(d\omega_x)\mathbb{P}_Y(d\omega_y)$, and where the events in the σ-algebras $\mathcal{F}_X$ and $\mathcal{F}_Y = \sigma(Y_k,\ k \geq 1)$ are of the type $A \times \Omega_Y$ and $\Omega_X \times B$, respectively. Observe that by using Fubini's theorem, and the definition of conditional expectation, one can show that

$$\mathbb{E}(\xi|\mathcal{F}_X)(\omega_x, \omega_y) = \int_{\Omega_Y} \xi(\omega_x, \omega_y)\mathbb{P}_Y(d\omega_y), \tag{5}$$

for all $\mathbb{P}$-integrable ξ.

Fix an $\omega_x \in \Omega_X$, $t > 0$, and consider the Lindeberg-Feller condition

$$LC_n(\omega_x) = \frac{1}{S_n^2}\sum_{k=1}^n X_k^2 \mathbb{E}\Big(Y_k^2 1_{\{Y_k^2 > t^2 S_n^2/X_k^2\}}|\mathcal{F}_X\Big),$$

where $S_n^2 = \sum_{k=1}^n X_k^2$.

Observe that the Lindeberg-Feller theorem is valid if $LC_n \to 0$ for a countable set of t's that have the zero as a limit point. We will show that for all fixed $t > 0$ (and therefore for a countable collection of t's), we have $LC_n(\omega_x) \to 0$, as $n \to \infty$, for $\mathbb{P}_X$-almost all $\omega_x \in \Omega_X$. This would imply that over a $\mathbb{P}_X$-probability one set, we have that $Z_n(\cdot, \omega_x) \Rightarrow Z$, as $n \to \infty$, which in view of (5) would yield $\mathbb{E}(f(Z_n)|\mathcal{F}_X) \to \mathbb{E}f(Z)$, with probability one, for any bounded and continuous function f.

We have that $\{Y_k^2 > t^2 S_n^2/X_k^2\}$ implies $\{Y_k^2 > t^2 S_n^2/\max_{k=1,\cdots,n} X_k^2\}$ and hence

$$\begin{aligned}\mathbb{E}\Big(Y_k^2 1_{\{Y_k^2 > t^2 S_n^2/X_k^2\}}|\mathcal{F}_X\Big) &\leq \mathbb{E}_Y(Y_k^2 1_{\{Y_k^2 > t^2 S_n^2/\max_{j=1,\cdots,n} X_j^2\}}) \\ &= \mathbb{E}_Y(Y_1^2 1_{\{Y_1^2 > t^2 S_n^2/\max_{j=1,\cdots,n} X_j^2\}}).\end{aligned}$$

The latter expectation vanishes $\mathbb{P}_X$-a.s. since $\mathbb{E}_Y(Y_1^2) < \infty$ and since by assumption, we have $S_n^2/\max_{j=1,\cdots,n} X_j^2 \to \infty$, $\mathbb{P}_X-$a.s. This completes the solution of part **(b)** since in view of the last inequality, we have

$$LC_n(\omega_x) \leq \frac{1}{S_n^2}\sum_{k=1}^n X_k^2 \mathbb{E}_Y(Y_1^2 1_{\{Y_1^2 > t^2 S_n^2/\max_{j=1,\cdots,n} X_j^2\}}) = \mathbb{E}_Y(Y_1^2 1_{\{Y_1^2 > t^2 S_n^2/\max_{j=1,\cdots,n} X_j^2\}}) \to 0,$$

for $\mathbb{P}_X$-almost all ω_x.

(c): By part **(b)**, for any bounded and continuous function f, we have that $\mathbb{E}(f(Z_n)|\mathcal{F}_X) \to \mathbb{E}(f(Z))$, almost surely. Since $|\mathbb{E}(f(Z_n)|\mathcal{F}_X)| \leq \sup_{x\in\mathbb{R}} |f(x)|$, by the Lebesgue DCT applied to the random variables $\mathbb{E}(f(Z_n)|\mathcal{F}_X)$, we obtain

$$\mathbb{E}(f(Z_n)) = \mathbb{E}\Big(\mathbb{E}(f(Z_n)|\mathcal{F}_X)\Big) \to \mathbb{E}f(Z).$$

Since the latter convergence is valid for all bounded and continuous functions f, it follows that $Z_n \Rightarrow Z,\ n \to \infty$. $\square$