Solutions to Exam 1

Problem 1 *(12 points)* Let $A_1, A_2, \ldots$ be a sequence of independent random variables, with $P[A_i = 1] = P[A_i = \frac{1}{2}] = \frac{1}{2}$ for all i. Let $B_k = A_1 \cdots A_k$.
(a) Does $\lim_{k\to\infty} B_k$ exist in the m.s. sense? Justify your anwswer.
(b) Does $\lim_{k\to\infty} B_k$ exist in the a.s. sense? Justify your anwswer.
(c) Let $S_n = B_1 + \ldots + B_n$. You can use without proof (time is short!) the fact that $\lim_{m,n\to\infty} E[S_m S_n] = \frac{35}{3}$, which implies that $\lim_{n\to\infty} S_n$ exists in the m.s. sense. Find the mean and variance of the limit random variable.
(d) Does $\lim_{n\to\infty}$ *a.s.* S_n exist? Justify your anwswer.

(a) $E[(B_k - 0)^2] = E[A_1^2]^k = (\frac{5}{8})^k \to 0$ as $k \to \infty$. Thus, $\lim_{k\to\infty}$ *m.s.* $B_k = 0$.
(b) Each sample path of the sequence B_k is monotone nonincreasing and bounded below by zero, and is hence convergent. Thus, $\lim_{k\to\infty}$ *a.s.* B_k exists. (The limit has to be the same as the m.s. limit, so B_k converges to zero almost surely.)
(c) Mean square convergence implies convergence of the mean. Thus, the mean of the limit is $\lim_{n\to\infty} E[S_n] = \lim_{n\to\infty} \sum_{k=1}^n E[B_k] = \sum_{k=1}^\infty (\frac{3}{4})^k = 3$. By the property given in the problem statement, the second moment of the limit is $\frac{35}{3}$, so the variance of the limit is $\frac{35}{3} - 3^2 = \frac{8}{3}$.
(d) Each sample path of the sequence S_n is monotone nondecreasing and is hence convergent. Thus, $\lim_{k\to\infty}$ *a.s.* B_k exists. (The limit has to be the same as the m.s. limit.)

Here is a proof of the claim given in part (c) of the problem statement, although the proof was not asked for on the exam. If $j \le k$, then $E[B_j B_k] = E[A_1^2 \cdots A_j^2 A_{j+1} \cdots A_k] = (\frac{5}{8})^j (\frac{3}{4})^{k-j}$, and a similar expression holds for $j \ge k$. Therefore,

$$
\begin{aligned}
E[S_n S_m] &= E[\sum_{j=1}^n B_j \sum_{k=1}^n B_k] = \sum_{j=1}^n \sum_{k=1}^n E[B_j B_k] \\
&\to \sum_{j=1}^\infty \sum_{k=1}^\infty E[B_j B_k] \\
&= 2\sum_{j=1}^\infty \sum_{k=j+1}^\infty \left(\frac{5}{8}\right)^j \left(\frac{3}{4}\right)^{k-j} + \sum_{j=1}^\infty \left(\frac{5}{8}\right)^j \\
&= 2\sum_{j=1}^\infty \sum_{l=1}^\infty \left(\frac{5}{8}\right)^j \left(\frac{3}{4}\right)^l + \sum_{j=1}^\infty \left(\frac{5}{8}\right)^j \\
&= (\sum_{j=1}^\infty \left(\frac{5}{8}\right)^j)(2\sum_{l=1}^\infty \left(\frac{3}{4}\right)^l + 1) \\
&= \frac{5}{3}(2 \cdot 3 + 1) = \frac{35}{3}
\end{aligned}
$$

Problem 2 *(12 points)* Let X, Y, and Z be random variables with finite second moments and suppose X is to be estimated. For each of the following, if true, give a brief explanation. If false, give a counter example.
(a) TRUE or FALSE: $E[|X - E[X|Y]|^2] \le E[|X - \widehat{E}[X|Y, Y^2]|^2]$.
(b) TRUE or FALSE: $E[|X - E[X|Y]|^2] = E[|X - \widehat{E}[X|Y, Y^2]|^2]$ if X and Y are jointly Gaussian.
(c) TRUE or FALSE? $E[\,|X - \;E[E[X|Z]\;|Y]|^2] \le E[|X - E[X|Y]|^2]$.
(d) TRUE or FALSE? If $E[|X - E[X|Y]|^2] = \mathrm{Var}(X)$ then X and Y are independent.

(a) TRUE. The estimator $E[X|Y]$ yields a smaller MSE than any other function of Y, including $\widehat{E}[X|Y, Y^2]$.
(b) TRUE. Equality holds because the unconstrained estimator with the smallest mean squre error, $E[X|Y]$, is linear, and the MSE for $\widehat{E}[X|Y, Y^2]$ is less than or equal to the MSE of any linear estimator.
(c) FALSE. For example if X is a nonconstant random variable, $Y = X$, and $Z \equiv 0$, then, on one hand, $E[X|Z] = E[X]$, so $E[E[X|Z]|Y] = E[X]$, and thus $E[|X - E[E[X|Z]|Y]|^2] = \mathrm{Var}(X)$. On the other hand, $E[X|Y] = X$ so that $E[|X - E[X|Y]|^2] = 0$.
(d) FALSE. For example, suppose $X = YW$, where W is a random variable independent of Y with mean zero and variance one. Then given Y, the conditional distribution of X has mean zero and variance Y^2. In particular, $E[X|Y] = 0$, so that $E[|X - E[X|Y]|^2] = \mathrm{Var}(X)$, but X and Y are not independent.

Problem 3 *(6 points)* Recall from a homework problem that if $0 < f < 1$ and if S_n is the sum of n independent random variables, such that a fraction f of the random variables have a CDF F_Y and a fraction $1-f$ have a CDF F_Z, then the large deviations exponent for $\frac{S_n}{n}$ is given by:

$$l(a) = \max_{\theta} \{\theta a - f M_Y(\theta) - (1-f) M_Z(\theta)\}$$

where $M_Y(\theta)$ and $M_Z(\theta)$ are the log moment generating functions for F_Y and F_Z respectively.

Consider the following variation. Let $X_1, X_2, \ldots, X_n$ be independent, and identically distributed, each with CDF given by $F_X(c) = f F_Y(c) + (1-f) F_Z(c)$. Equivalently, each X_i can be generated by flipping a biased coin with probability of heads equal to f, and generating X_i using CDF F_Y if heads shows and generating X_i with CDF F_Z if tails shows. Let $\widetilde{S}_n = X_1 + \cdots + X_n$, and let $\widetilde{l}$ denote the large deviations exponent for $\frac{\widetilde{S}_n}{n}$.
(a) Express the function $\widetilde{l}$ in terms of f, M_Y, and M_Z.
(b) Determine which is true and give a proof: $\widetilde{l}(a) \le l(a)$ for all a, or $\widetilde{l}(a) \ge l(a)$ for all a.

(a) $\widetilde{l}(a) = \max_\theta \{\theta a - M_X(\theta)\}$ where

$$\begin{aligned} M_X(\theta) &= \log E[\exp(\theta X)] \\ &= \log\{f E[\exp(\theta Y)] + (1-f) E[\exp(\theta Z)]\} \\ &= \log\{f \exp(M_Y(\theta)) + (1-f)\exp(M_Z(\theta))\} \end{aligned}$$

(b) View $f\exp(M_Y(\theta)) + (1-f)\exp(M_Z(\theta))$ as an average of $\exp(M_Y(\theta))$ and $\exp(M_Z(\theta))$. The definition of concavity (or Jensen's inequality) applied to the concave function $\log u$ implies that $\log(\textit{average}) \ge \textit{average}(\log)$, so that $\log\{f\exp(M_Y(\theta)) + (1-f)\exp(M_Z(\theta)) \ge f M_Y(\theta) + (1-f) M_Z(\theta)$, where we also used the fact that $\log \exp M_Y(\theta) = M_Y(\theta)$. Therefore, $\widetilde{l}(a) \le l(a)$ for all a.

Remark: This means that $\frac{\widetilde{S}_n}{n}$ is more likely to have large deviations than $\frac{S_n}{n}$. That is reasonable, because $\frac{\widetilde{S}_n}{n}$ has randomness due not only to F_Y and F_Z, but also due to the random coin flips. This point is particularly clear in case the Y's and Z's are constant, or nearly constant, random variables.