Statistics 521, Problem Set 10, Solutions

Wellner; 12/6/2007

1. Let $K \geq 3$ be a prime and let X and Y be independent random variables that are uniformly distributed on $\{0, 1, \ldots, K-1\}$. For $0 \leq n < K$ let $Z_n = (X + nY) mod(K)$. Show that $Z_0, \ldots, Z_{K-1}$ are *pairwise independent*; that is, each pair is independent, but if we know the values of two of the variables, then we know the values of all the variables.

   Solution: Suppose that $Z_m = a$ and $Z_n = b$ with $m, n \in \{0, \ldots, K-1\}$ and $m \neq n$. Then

   $$a-b = Z_m - Z_n = (x+my)\text{mod}(K) - (x+ny)\text{mod}(K) = [(m-n)y]\text{mod}(K)$$

   and hence $y = [(a-b) + Ks]/(m-n)] = y(a,b)$ for some integer s. Also,

   $$\begin{aligned} mb - na &= mZ_n - nZ_m \\ &= (mx + mny)\text{mod}(K) - (nx + nmy)\text{mod}(K) \\ &= [(m-n)x]\text{mod}(K)\,, \end{aligned}$$

   so that $x = [(mb - na) + Kr]/(m-n) \equiv x(a,b)$ for some r. Thus the values of Z_m and Z_n determine the values of X and Y, and hence also the values of Z_j for $j \notin \{m, n\}$.

   To show pairwise independence, we begin by showing that Z_m is uniformly distributed over $\{0, \ldots, K-1\}$ for each $m \in \{0, \ldots, m\}$: (The following solution is that of Dick Hwang.) Let $c \in \{0, \ldots, K-1\}$, and consider the mapping $\sigma : \{0, \ldots, K-1\} \to \{0, \ldots, K-1\}$ defined by $\sigma(i) \equiv (ni+c)\text{mod}(K)$, $i \in \{0, \ldots, K-1\}$. We claim that σ is a permutation of $\{0, \ldots, K-1\}$. To see this, suppose that $\sigma(i) = \sigma(j)$ for some $i \neq j$, $0 \leq i, j < K$. Then there exists m such that $ni+c = nj+c+Km$ with $m \neq 0$ since $i \neq j$. Thus $i = j + Km/n$. Since K is prime, n divides m, and hence $|m| \geq n$. Then $|i-j| \geq K$, which contradicts $0 \leq i, j < K$. Hence $\sigma(i) \neq \sigma(j)$ whenever $i \neq j$; and thus σ is one-to-one. Thus (by the pigeon-hole principle), σ is onto, and is therefore, a

permutation on $\{0, \ldots, K-1\}$. Thus if Y is uniformly distributed on $\{0, \ldots, K-1\}$, then $(nY + c)\mathrm{mod}(K)$ is also uniformly distributed on $\{0, \ldots, K-1\}$.

Now, for $0 < n < K$, define $\sigma(i, j) \equiv (ni + j)\mathrm{mod}(K)$ for $0 \leq i, j < K$. We claim that σ maps exactly K pairs of $(i, j) : 0 \leq i, j < K$ into each element of $\{0, \ldots, K-1\}$. To see this, fix $j = c$, and then $\sigma(i, j) = \sigma(i, c) = \sigma(i)$ (of the preceding paragraph) which maps one (i, j), $j = c$ into each element of $\{0, \ldots, K-1\}$. Now j can take on K different vlaues, and the claim follows. Hence if X and Y are uniformly and independently distributed on $\{0, \ldots, K-1\}$, then $(X + nY)\mathrm{mod}(K)$ is uniformly distributed on $\{0, \ldots, K-1\}$. for $n = 1, \ldots, K-1$. Note that for $n = 0$ this is also true since then $(X)\mathrm{mod}(K) = X$, which is uniform on $\{0, \ldots, K-1\}$ by assumption.

Now to show the pairwise independence it suffices, since each Z_m is marginally uniformly distributed by the above, that for $m \neq n$, $0 \leq m, n < K-1$, we have

$$P(Z_m = a, Z_n = b) = \frac{1}{K^2} = P(Z_n = a)P(Z_n = b)$$

for $a, b \in \{0, \ldots, K-1\}$. But by the first paragraph of the solution, with $x(a, b), y(a, b) \in \{0, \ldots, K-1\}$,

$$\begin{aligned} & P(Z_m = a, Z_n = b) \\ &= P\left(X = x(a, b), Y = y(a, b)\right) \\ &= P\left(X = x(a, b)\right) P\left(Y = y(a, b)\right) \\ &= \frac{1}{K} \cdot \frac{1}{K} \end{aligned}$$

by the independence of X and Y. Thus Z_m and Z_n are independent.

2. Show that if X_n is any sequence of random variables, there are constants $c_n \to \infty$ so that $X_n / c_n \to_{a.s.} 0$.

   Solution: Define $F_n(x) \equiv P(|X_n| \leq x)$, the distribution function of $|X_n|$. Set $b_n = F_n^{-1}(1 - n^{-2})$ for $n = 1, 2, \ldots$, and let $\{a_n\}$ be any sequence with $a_n \to \infty$. Let $c_n = a_n b_n$. Then, for any $\epsilon > 0$ we have

$\epsilon a_n \geq 1$ for $n \geq N_\epsilon$ and hence, using $F_n \circ F_n^{-1}(t) \geq t$ for all $0 < t < 1$,

$$\begin{aligned} P(|X_n| > \epsilon c_n) &= P(|X_n| > \epsilon a_n b_n) \\ &\leq P(|X_n| > b_n) \\ &= 1 - F_n(b_n) = 1 - F_n(F_n^{-1}(1 - n^{-2})) \\ &\leq n^{-2}, \quad n \geq N_\epsilon. \end{aligned}$$

Hence by the first Borel-Cantelli lemma, $P(|X_n| > \epsilon c_n \text{ i.o.}) = 0$ for every $\epsilon > 0$; that is, $X_n/c_n \to_{a.s.} 0$.

3. Show that if $P(A_n) \to 0$ and $\sum_{n=1}^\infty P(A_n \cap A_{n+1}^c) < \infty$, then $P(A_n \text{ i.o.}) = 0$.

Solution: If the A_n's are decreasing, this is fairly easy: then we can write

$$\begin{aligned} \cup_{k=n}^\infty A_k &= \cup_{k=n}^\infty A_k \cap A_{k+1}^c + \cup_{k=n}^\infty A_k \cap A_{k+1} \\ &= \cup_{k=n}^\infty A_k \cap A_{k+1}^c + A_{n+1}, \end{aligned}$$

and this yields

$$\begin{aligned} P(A_n \text{ i.o.}) &= \lim_{n\to\infty} P(\cup_{k=n}^\infty A_k) \\ &\leq \lim_{n\to\infty} \left\{ P(\cup_{k=n}^\infty A_k \cap A_{k+1}^c) + P(A_{n+1}) \right\} \\ &= P(A_n \cap A_{n+1}^c \text{ i.o.}) + 0 = 0\,. \end{aligned}$$

On the other hand, if the A_n's are not decreasing, this is a bit trickier. Now use the usual disjointfication procedure: let $B_1 = A_1$, $B_2 = A_n^c A_{n+1}$, ... , $B_k = A_n^c A_{n+1}^c \cdots A_{n+k-2}^c A_{n+k-1}, \cdots$. Then

$$\begin{aligned} \cup_{k=n}^\infty A_k &= \sum_{k=1}^\infty B_k \\ &= A_n + \cup_{k=n} A_{k+1} \cap_{j=n}^\infty A_j^c \\ &\subset A_n + \cup_{k=n} A_{k+1} A_k^c\,. \end{aligned}$$

Since the left side is decreasing in n, this implies that for each fixed n we have

$$[A_n \text{ i.o.}] \subset A_n + [A_{n+1} \cap A_n^c \text{ i.o.}]\,.$$

But $[A_{n+1} \cap A_n^c \text{ i.o.}] = [A_n \cap A_{n+1}^c \text{ i.o.}]$, so

$$[A_n \text{ i.o.}] \subset A_n + [A_n \cap A_{n+1}^c \text{ i.o.}]$$

for every fixed n. Thus

$$P(A_n \text{ i.o.}) \leq P(A_n) + P(A_n \cap A_{n+1}^c \text{ i.o.}) = P(A_n) + 0$$

by the first Borel-Cantelli lemma. Since $P(A_n) \to 0$, it follows that $P(A_n \text{ i.o.}) = 0$.

4. Let $X_1, X_2, \ldots$ be independent. Show that $\sup X_n < \infty$ almost surely if and only if $\sum_n P(X_n > M) < \infty$ for some $M < \infty$.

   Solution: Suppose that $\sum_n P(X_n > M) < \infty$ for some $M < \infty$. Then by the first Borel-Cantelli lemma, $P(X_n > M \text{ i.o.}) = 0$; i.e. for $n \geq N_\omega$ we have $X_n(\omega) \leq M$. Thus

   $$\sup_n X_n(\omega) \leq \left(\max_{1 \leq k < N_\omega} X_k \right) \vee M < \infty .$$

   Now suppose that $\sup X_n < \infty$ almost surely. If $\sum_n P(X_n > M) = \infty$ for every $M < \infty$, then, by the second Borel-Cantelli lemma, $P(X_n > M \text{ i.o.}) = 1$ for every M; i.e. $\limsup_{n\to\infty} X_n \geq M$ a.s. for every $M > 0$, and this implies, by taking a sequence $M_k \nearrow \infty$, that $\limsup_{n\to\infty} X_n = \infty$ a.s., which contradicts $\sup X_n < \infty$ almost surely. We therefore conclude that $\sum_n P(X_n > M) < \infty$ for some $M < \infty$.

5. Let $X_1, X_2, \ldots$ be independent with $P(X_n = 1) = p_n$ and $P(X_n = 0) = 1 - p_n$. Show that: (i) $X_n \to_p 0$ if and only if $p_n \to 0$, and $X_n \to_{a.s.} 0$ if and only if $\sum_n p_n < \infty$.

   Solution: Here, for any $0 < \epsilon < 1$, $P(|X_n| > \epsilon) = P(X_n = 1) = p_n$. Thus $X_n \to_p 0$ if and only if $p_n \to 0$.
   By the first and second Borel-Cantelli lemmas and the first part, for $0 < \epsilon < 1$ we have

   $$P(|X_n| > \epsilon \text{ i.o.}) = \left\{ \begin{array}{c} 0 \\ 1 \end{array} \right\} \text{ according as } \sum_{n=1}^{\infty} p_n \left\{ \begin{array}{c} < \infty \\ = \infty \end{array} \right\} .$$

   It follows that $X_n \to_{a.s.} 0$ if and only if $\sum_{n=1}^{\infty} p_n < \infty$.

6. Suppose that $X_1, X_2, \ldots$ are independent with $P(X_n > x) = x^{-5}$ for all $x \geq 1$ and $n = 1, 2, \ldots$. Show that $\limsup_{n\to\infty}(\log X_n)/\log n = c$ almost surely for some number c, and find c.

Solution: Let $c > 0$. Then

$$P(\log X_n > c \log n) = P(X_n > n^c) = n^{-5c}.$$

Hence by the first and second Borel-Cantelli lemmas

$$P(\log X_n > c \log n \text{ i.o.}) = \begin{cases} 0 & \text{if } c > 1/5 \\ 1 & \text{if } c \leq 1/5 \end{cases}.$$

Hence

$$\limsup_{n\to\infty}(\log X_n)/(\log n) = 1/5 \qquad \text{almost surely}.$$