Statistics 523, Problem Set 1 Solutions

Wellner; 4/9/10

1. Let $X_1, X_2, \ldots$ be i.i.d. $U(0,1)$, let $S_n = X_1 + \cdots + X_n$ and let $T \equiv \inf\{n : S_n > 1\}$. Show that $P(T > n) = 1/n!$ and hence that $E(T) = e$ and $E(S_T) = e/2$.

   Solution: Note that $[T > n] = [S_n \le 1]$ where $P(S_1 \le 1) = P(X_1 \le 1) = 1$ and $P(S_2 \le 1) = 1/2$ by an easy symmetry argument. Furthermore, for $x \le 1$, with f_n denoting the density of S_n (which clearly exists)

$$\begin{aligned} F_{n+1}(x) &\equiv P(S_{n+1} \le x) = E\left\{E[1_{[S_{n+1} \le x]} | S_n]\right\} \\ &= E\left\{P(X_{n+1} \le x - S_n | S_n) 1\{S_n \le x\}\right\} \\ &= E\left\{(x - S_n) 1\{S_n \le x\}\right\} = \int_{[0,x]} (x - s) f_n(s) ds. \end{aligned}$$

   By differentiating this relation with respect to x we find that

$$f_{n+1}(x) = \int_{[0,x]} f_n(s) ds$$

   Since $f_n(s) = 1_{[0,1]}(s)$, this yields $f_n(x) = x^{n-1}/(n-1)!$ for $x \le 1$ and all $n \ge 1$, so that

$$P(S_n \le 1) = \int_{[0,1]} f_n(s) ds = \int_{[0,1]} \frac{s^{n-1}}{(n-1)!} ds = \frac{1}{n!}.$$

   This is in complete agreement with the more general result of Feller (1968), section I.9.6, page 27, who gives the distribution function of S_n for all x as

$$F_n(x) = P(S_n \le x) = \frac{1}{n!} \sum_{k=0}^{n} (-1)^k \binom{n}{k} (x - k)_+^n.$$

   Thus $P(T > n) = 1/n!$, and

$$\begin{aligned} P(T = n) &= P(T \ge n) - P(T \ge n+1) = P(T > n-1) - P(T > n) \\ &= \frac{1}{(n-1)!} - \frac{1}{n!} = \frac{n-1}{n!}, \quad n \ge 2. \end{aligned}$$

   Thus it follows that

$$E(T) = \sum_{n=2}^{\infty} n P(T = n) = \sum_{n=2}^{\infty} n \frac{n-1}{n!} = \sum_{n=2}^{\infty} \frac{1}{(n-2)!} = \sum_{k=0}^{\infty} \frac{1}{k!} = e,$$

and hence, by Wald's equation, that $E(S_T) = E(T) \cdot E(X_1) = e \cdot (1/2) = e/2$.

Here is a slightly more general problem: for a fixed integer k, let $T_k \equiv \inf\{n \geq 1 : S_n > k\}$. Thus $T_1 = T$ in the problem as posed and treated above. Then $[T_k > n] = [S_n \leq k]$, and by the distribution of S_n as given above (from Feller, vol. II),

$$P(T_k > n) = P(S_n \leq k) = \frac{1}{n!} \sum_{j=0}^{n} (-1)^j \binom{n}{j} (k-j)_+^n.$$

Note that $P(T_k > k) = P(T_k \geq k+1) = 1$ for each $k \geq 1$. Now $P(T_k > n) = P(T_k \geq n+1)$, so $P(T_k = n) = P(T_k > n-1) - P(T_k > n)$ and

$$\begin{aligned} E(T_k) &= \sum_{k=1}^{\infty} P(T \geq n) = k + 1 + \sum_{n=k+2}^{\infty} P(T_k \geq n) \\ &= k + 1 + \sum_{n=k+2}^{\infty} P(T_k > n-1) = k + 1 + \sum_{j=k+1}^{\infty} P(T_k > j) \\ &= k + 1 + \sum_{j=k+1}^{\infty} \frac{1}{j!} \sum_{l=0}^{j} (-1)^l \binom{j}{l} (k-l)_+^j. \end{aligned}$$

2. Let $\mathcal{A}_k$ be an increasing family of $\sigma-$fields, $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \mathcal{A}_3 \subset \cdots$. Then N is a stopping time relative to $\{\mathcal{A}_k\}$ if $[N = k] \in \mathcal{A}_k$ for all $k \geq 1$. Let N_1, N_2 be stopping times relative to $\{\mathcal{A}_k\}$. Show that $N_1 \wedge N_2$, $N_1 \vee N_2$, and $N_1 + N_2$ are stopping times relative to $\{\mathcal{A}_k\}$.

   Solution: (i) Let $N = N_1 \wedge N_2$. Then

$$[N = n] = [N_1 \wedge N_2 = n] = [N_1 = n, N_2 \geq n] \cup [N_2 = n, N_1 \geq n].$$

   Now $[N_1 = n] \in \mathcal{A}_n$ since N_1 is a stopping time, while $[N_2 \geq n] = [N_2 < n]^c = [N_2 \leq n-1]^c \in \mathcal{A}_{n-1} \subset \mathcal{A}_n$ since N_2 is a stopping time and the $\mathcal{A}_k$'s are increasing. Thus $[N_1 = n, N_2 \geq n] \in \mathcal{A}_n$. By symmetry, $[N_2 = n, N_1 \geq n] \in \mathcal{A}_n$ and we conclude that $[N = n] \in \mathcal{A}_n$. Thus $N = N_1 \wedge N_2$ is a stopping time.
   (ii) Let $N = N_1 \vee N_2$. Then

$$[N = n] = [N_1 \vee N_2 = n] = [N_1 = n, N_2 \leq n] \cup [N_2 = n, N_1 \leq n].$$

   Now $[N_1 = n] \in \mathcal{A}_n$ since N_1 is a stopping time, while $[N_2 \leq n] = \cup_{k=1}^n [N_2 = k] \in \mathcal{A}_n$ since N_2 is a stopping time and the $\mathcal{A}_k$'s are increasing. Thus $[N_1 = n, N_2 \leq n] \in \mathcal{A}_n$. By symmetry, $[N_2 = n, N_1 \leq n] \in \mathcal{A}_n$ and we conclude that

$[N = n] \in \mathcal{A}_n$. Thus $N = N_1 \vee N_2$ is a stopping time.
(iii) Let $N = N_1 + N_2$. Then

$$[N = n] = [N_1 + N_2 = n] = \cup_{k=1}^{n-1}[N_1 = k, N_2 \le n - k].$$

Now $[N_1 = k] \in \mathcal{A}_k \subset \mathcal{A}_n$ since N_1 is a stopping time, while $[N_2 \le n - k] \in \mathcal{A}_{n-k} \subset \mathcal{A}_n$ since N_2 is a stopping time and the $\mathcal{A}_k$'s are increasing. Thus $[N_1 = k, N_2 = n - k] \in \mathcal{A}_n$ for each $k \in \{1, \ldots, n-1\}$, and we conclude that $[N = n] \in \mathcal{A}_n$. Thus $N = N_1 + N_2$ is a stopping time.

3. In the proof of the existence of Brownian motion as a continuous process on 4/2, we used a 4th moment version of Markov's inequality and reached the conclusion the Brownian motion is Hölder continuous with exponent $\gamma < 1/4$. Can this be strengthened by using a higher-order version of Markov's inequality? What can be concluded if the 4th moment bound is replaced by an 8th moment bound, or by a $2m$th moment bound, or by an exponential bound? (The moment formula given in PfS (12.3.6), page 308 is relevant in this connection.)

   Solution: Since $E|B_t - B_s|^8 = C|t - s|^4$ with $C = E|B_1|^8 = 1 \cdot 3 \cdot 5 \cdot 7 = 105$, the proof yields

$$\begin{aligned}
&P(|B(j2^{-n}) - B(i2^{-n})| > ((j-i)2^{-n})^{\gamma} \text{ for some } 0 \le i,j \le 2^n,\ 0 < j - i < 2^{n\delta}) \\
&\le \sum_{(i,j):j-i \le 2^{n\delta}} ((j-i)2^{-n})^{-8\gamma} E|B(j2^{-n}) - B(i2^{-n})|^8 \\
&\le C \sum_{(i,j):j-i \le 2^{n\delta}} ((j-i)2^{-n})^{-8\gamma}((j-i)2^{-n})^4 \\
&\le C 2^n 2^{n\delta}(2^{n\delta}2^{-n})^{-8\gamma+4} \\
&= C 2^{-n\epsilon}
\end{aligned}$$

   where $\epsilon \equiv (1-\delta)(4 - 8\gamma) - (1+\delta) > 0$ if $\gamma < 3/8$ and δ is chosen so that $4 - 8\gamma > (1+\delta)/(1-\delta)$. Repeating the rest of the argument shows that B is Hölder of order $\gamma < 3/8$. Repeating this argument with 8 replaced by $2k$ for large $k \ge 2$ yields

$$\begin{aligned}
&P(|B(j2^{-n}) - B(i2^{-n})| > ((j-i)2^{-n})^{\gamma} \text{ for some } 0 \le i,j \le 2^n,\ 0 < j - i < 2^{n\delta}) \\
&\le \sum_{(i,j):j-i \le 2^{n\delta}} ((j-i)2^{-n})^{-2k\gamma} E|B(j2^{-n}) - B(i2^{-n})|^{2k} \\
&\le C \sum_{(i,j):j-i \le 2^{n\delta}} ((j-i)2^{-n})^{-2k\gamma}((j-i)2^{-n})^k \\
&\le C 2^n 2^{n\delta}(2^{n\delta}2^{-n})^{-2k\gamma+k} \\
&= C 2^{-n\epsilon}
\end{aligned}$$

where $\epsilon \equiv (1-\delta)(k-2k\gamma)-(1+\delta) > 0$ if $\gamma < (k-1)/(2k)$ and δ is chosen so that $k-2k\gamma > (1+\delta)/(1-\delta)$. Thus B is Hölder of order $\gamma < (k-1)/(2k)$ for each k, and by choosing k large, B is Hölder of order $\gamma < 1/2$. By using the following exponential bound for the tail of a standard normal random variable $Z \stackrel{d}{=} B_1$ we can improve this still further: for all $z > 0$

$$P(|Z| > z) \le 2\exp(-z^2/2).$$

[Proof: $P(Z > z) = P(rZ > rz)$ for $r > 0$, so

$$P(Z > z) = P(\exp(rZ) > \exp(rz)) \le E\{e^{rZ}/e^{rz}\} = e^{-rz}e^{r^2/2};$$

Choosing $r = z$ gives the one-sided bound.] Now for $\epsilon > 0$,

$$\begin{aligned}
&P\Big(|B(j2^{-n}) - B(i2^{-n})| > \big(2(1+\epsilon)(j-i)2^{-n}\log(2^n/(j-i))\big)^{1/2} \\
&\qquad\qquad \text{for some } 0 \le i,j \le 2^n,\ 0 < j-i < 2^{n\delta}\Big) \\
\le &\sum_{(i,j):j-i\le 2^{n\delta}} P\Big(|B(j2^{-n}) - B(i2^{-n})| > \big(2(1+\epsilon)(j-i)2^{-n}\log(2^n/(j-i))\big)^{1/2}\Big) \\
= &\sum_{(i,j):j-i\le 2^{n\delta}} P\Big(|B(1)| > (2(1+\epsilon)\log(2^n/(j-i)))^{1/2}\Big) \\
\le &\ 2^n 2^{n\delta} P\Big(|B(1)| > (2(1+\epsilon)\log(2^n))^{1/2}\Big) \\
\le &\ 2^{(1+\delta)n} 2\exp\left(-\left((1+\epsilon)\log(2^n)\right)\right) \\
= &\ 2\cdot 2^{\{(1+\delta)-(1+\epsilon)\}n} = 2\cdot 2^{-(\epsilon-\delta)n},
\end{aligned}$$

and this converges if $\delta < \epsilon$. This is (a substantial part of) the proof of Lévy's Theorem 12.9.2, PfS, page 331.

4. Exercise 12.3.1, PfS, page 309. In the proof of (17), that $t\mathbb{S}(1/t)$ is a Brownian motion, can you use a weaker result than the LIL of (10) to conclude continuity of the sample paths at 0?

   Solution: (11) $\mathbb{S}(t) = \mathbb{V}(t) + tZ$ is a Brownian motion on $[0,1]$ since it is Gaussian with continuous and has

$$\begin{aligned}
&E\mathbb{S}(t) = 0, \qquad \text{and} \\
&E\{\mathbb{S}(s)\mathbb{S}(t)\} = E\mathbb{V}(s)\mathbb{V}(t)\} + stE(Z^2) = s\wedge t - st + st = s\wedge t.
\end{aligned}$$

   (12) $\mathbb{S}(at)/\sqrt{a}$, $0 \le t < \infty$ is a Brownian motion process on $[0,\infty)$ since it is Gaussian with continuous sample paths and since $E\mathbb{S}(at)/\sqrt{a} = 0$ while

$$E\left\{\frac{\mathbb{S}(as)}{\sqrt{a}}\cdot\frac{\mathbb{S}(at)}{\sqrt{a}}\right\} = \frac{1}{a}\{(as)\wedge(at)\} = s\wedge t.$$

(13) $\mathbb{Y}(t) \equiv \mathbb{S}(a+t) - \mathbb{S}(a)$ is a Brownian motion on $[0, \infty)$ by the independent increments property, or, since it is clearly Gaussian with continuous sample paths and has $E\mathbb{Y}(t) = 0 - 0 = 0$ and

$$\begin{aligned} E\{\mathbb{Y}(s)\mathbb{Y}(t)\} &= E(\mathbb{S}(a+s) - \mathbb{S}(a))(\mathbb{S}(a+t) - \mathbb{S}(a)) \\ &= (a+s) \wedge (a+t) - (a+s) \wedge a - a \wedge (a+t) + a \\ &= a + s \wedge t - a - a + a = s \wedge t. \end{aligned}$$

(14)
(15)
(16)
(17) $\mathbb{Z}(t) \equiv t\mathbb{S}(1/t)$ is Gaussian with $E\mathbb{Z}(t) = tE\mathbb{S}(1/t) = 0$, and

$$\begin{aligned} E\{\mathbb{Y}(s)\mathbb{Y}(t)\} &= stE\mathbb{S}(1/s)\mathbb{S}(1/t) = st\left\{\frac{1}{s} \wedge \frac{1}{t}\right\} \\ &= st/\max\{s, t\} = s \wedge t. \end{aligned}$$

Also note that $\mathbb{Z}(0) = 0$ almost surely by the strong law of large numbers applied to $\mathbb{S}$:

$$\begin{aligned} \lim_{t \downarrow 0} \mathbb{Z}(t) &= \lim_{t \downarrow 0} t\frac{\mathbb{S}(1/t)}{1/t} = \lim_{n \to \infty} \frac{\mathbb{S}(n)}{n} \\ &= \lim_{n \to \infty} \frac{\sum_{k=1}^{n}\{\mathbb{S}(k) - \mathbb{S}(k-1)\}}{n} \\ &= 0 \quad \text{almost surely} \end{aligned}$$

since the increments $\mathbb{S}(k) - \mathbb{S}(k-1)$ are i.i.d. $N(0, 1)$ rv's.