Statistics 522, Problem Set 6 Solutions

Wellner; 2/28/2008

1. Suppose that $\{N_t : t \ge 0\}$ is a Poisson process with rate $\nu > 0$; recall PfS, Chapter 9, pages 195-196 (or, for example, Billingsley, pages 307-313).

   (a) Fix $0 < s < t$, and let $A = [N_s = 0]$, $B_i = [N_t = i]$, $i = 0, 1, 2, \ldots$, and let $\mathcal{D} \equiv \sigma[B_1, B_2, \ldots]$. Show that $P(N_s = 0|\mathcal{D}) = (1 - s/t)^{N_t}$.

   (b) Fix $0 < s < t$. Show that

$$P(N_s = k|N_t) = \binom{N_t}{k}\left(\frac{s}{t}\right)^k\left(1 - \frac{s}{t}\right)^{N_t - k} 1_{\{0,\ldots,N_t\}}(k).$$

   Thus $(N_s|N_t) \sim \text{Binomial}(N_t, s/t)$.

   (c) Fix $0 \equiv s_0 < s_1 < \cdots < s_{m-1} < s_m \equiv t$. Show that

$$\begin{aligned} &P(\cap_{j=1}^m \left[N_{s_j} - N_{s_{j-1}} = k_j\right] |N_t) \\ &\quad = \frac{N_t!}{k_1! \cdots k_m!} \prod_{j=1}^m \left(\frac{s_j - s_{j-1}}{t}\right)^{k_j} 1_{\{0 \le k_j \le N_t,\ j=1,\ldots,m,\ \sum_{j=1}^m k_j = N_t\}}. \end{aligned}$$

   That is,

$$\begin{aligned} &(N_{s_1}, N_{s_2} - N_{s_1}, \ldots, N_t - N_{s_{m-1}} \big| N_t) \\ &\qquad \sim \text{Mult}_m\left(N_t, \left(\frac{s_1}{t}, \frac{s_2 - s_1}{t}, \ldots, \frac{t - s_{m-1}}{t}\right)\right). \end{aligned}$$

   Remark. The result of (c) implies that

$$(\{N_s :\ 0 \le s \le t\}|N_t = n) \overset{d}{=} \{n\mathbb{F}_n(s) :\ 0 \le s \le t\}$$

   where $\mathbb{F}_n(s) = n^{-1}\sum_{i=1}^n 1_{[0,s]}(X_i)$ is the empirical distribution function of X_i i.i.d. Uniform$(0, t)$.

   Solution: (a) and (b). Note that $\mathcal{D} = \sigma[N_t]$. Thus conditional probabilities given $\mathcal{D}$ are just conditional probabilities given N_t. For

$B_i = [N_t = i]$, by using the independent increments property of the Poisson process,

$$\begin{aligned} P(N_s = k|B_i) &= P(N_s = k|N_t = i) \\ &= \frac{P(N_s = k,\ N_t = i)}{P(N_t = i)} 1_{\{0,\dots,i\}}(k) \\ &= \frac{P(N_s = k,\ N_t - N_s = i - k)}{P(N_t = i)} 1_{\{0,\dots,i\}}(k) \\ &= \frac{P(N_s = k) P(N_t - N_s = i - k)}{P(N_t = i)} 1_{\{0,\dots,i\}}(k) \\ &= \frac{e^{-\nu s}(\nu s)^k/k! \cdot e^{-\nu(t-s)}(\nu(t-s))^{i-k}/(i-k)!}{e^{-\nu t}(\nu t)^i/i!} 1_{\{0,\dots,i\}}(k) \\ &= \frac{i!}{k!(i-k)!} \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{i-k} 1_{\{0,\dots,i\}}(k). \end{aligned}$$

It follows that

$$P(N_s = k|N_t) = \binom{N_t}{k} \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{N_t - k} 1_{\{0,\dots,N_t\}}(k).$$

(c) Similarly to the calculations above,

$$\begin{aligned} & P(\cap_{j=1}^m \left[N_{s_j} - N_{s_{j-1}} = k_j\right] |N_t = i) \\ &= \frac{P(\cap_{j=1}^m \left[N_{s_j} - N_{s_{j-1}} = k_j\right], N_t = i)}{P(N_t = i)} 1_{\{0 \le k_j \le m,\ \sum_1^m k_j = i\}} \\ &= \frac{\prod_{j=1}^m P(N_{s_j} - N_{s_{j-1}} = k_j)}{P(N_t = i)} 1_{\{0 \le k_j \le m,\ \sum_1^m k_j = i\}} \\ &= \frac{\prod_{j=1}^m e^{-\nu(s_j - s_{j-1})} \frac{(\nu(s_j - s_{j-1}))^{k_j}}{k_j!}}{e^{-\nu t} \frac{(\nu t)^i}{i!}} 1_{\{0 \le k_j \le m,\ \sum_1^m k_j = i\}} \\ &= \frac{N_t!}{k_1! \cdots k_m!} \prod_{j=1}^m \left(\frac{s_j - s_{j-1}}{t}\right)^{k_j} 1_{\{0 \le k_j \le N_t,\ j=1,\dots,m,\ \sum_{j=1}^m k_j = N_t\}}. \end{aligned}$$

2. Suppose that $X_1, \dots, X_n, \dots$ are i.i.d. with distribution function F, and let N be a Poisson random variable with parameter λ which is independent of the X_i's. Define the stochastic process $\{M_t : t \in \mathbb{R}\}$ by

$$M_t = \sum_{i=1}^N 1_{(-\infty,t]}(X_i) = N\mathbb{F}_N(t)$$

where $\mathbb{F}_n(t) = n^{-1}\sum_{i=1}^n 1_{(-\infty,t]}(X_i)$ is the empirical distribution function of $X_1, \ldots, X_n$.
(a) Show that $M_t \sim Poisson(\lambda F(t))$ for each fixed $t \in \mathbb{R}$.
(b) Show that for fixed $-\infty \equiv s_0 < s_1 < s_2 < \cdots < s_{m-1} < s_m \equiv \infty$, the random variables $M_{s_1}, M_{s_2} - M_{s_1}, \ldots, M_{s_m} - M_{s_{m-1}}$ are independent Poisson random variables with parameters $\lambda F(s_1), \lambda(F(s_2) - F(s_1)), \ldots, \lambda(1 - F(s_{m-1}))$ respectively.
Remark: (a) and (b) imply that the process M_t is a Poisson process with intensity (and mean) function λF; this fact generalizes to the empirical measure of observations X_i i.i.d. P on $(\mathcal{X}, \mathcal{A})$; see e.g. van der Vaart and Wellner (1996), section 3.5.2.

Solution: (a) Since $(M_t|N) \sim \text{Binomial}(N, F(t))$, for $k \in \{0, 1, 2, \ldots\}$,

$$\begin{aligned}
P(M_t = k) &= EP(M_t = k|N) = E\left\{\binom{N}{k} F(t)^k (1 - F(t))^{N-k} 1_{\{0,\ldots,N\}}(k)\right\} \\
&= \frac{1}{k!}\left(\frac{F(t)}{1 - F(t)}\right)^k E\left\{\frac{N!}{(N-k)!}(1 - F(t))^N 1_{\{0,\ldots,N\}}(k)\right\} \\
&= \frac{1}{k!}\left(\frac{F(t)}{1 - F(t)}\right)^k \sum_{j=k}^{\infty} \frac{j!}{(j-k)!}(1 - F(t))^j e^{-\lambda}\frac{\lambda^j}{j!} \\
&= \frac{e^{-\lambda}}{k!}\left(\frac{F(t)}{1 - F(t)}\right)^k \sum_{j=k}^{\infty} \frac{\lambda^j (1 - F(t))^j}{(j-k)!} \\
&= \frac{e^{-\lambda}(\lambda F(t))^k}{k!} \sum_{r=0}^{\infty} \frac{[\lambda(1 - F(t))]^r}{r!} \\
&= \frac{e^{-\lambda}(\lambda F(t))^k}{k!} e^{\lambda(1-F(t))} \\
&= e^{-\lambda F(t)} \frac{(\lambda F(t))^k}{k!}.
\end{aligned}$$

Thus $M_t \sim \text{Poisson}(\lambda F(t))$ as claimed.
(b) Since

$$\begin{aligned}
&(M_{s_1}, M_{s_2} - M_{s_1}, \ldots, M_{s_m} - M_{s_{m-1}}|N) \\
&\quad \sim \text{Mult}_m(N, (F(s_1), F(s_2) - F(s_1), \ldots, 1 - F(s_{m-1}))),
\end{aligned}$$

it follows that

$$\begin{aligned}
&P(\cap_{j=1}^{m}\left[M_{s_j} - M_{s_{j-1}} = k_j\right]) \\
&\quad = \quad E\left\{P(\cap_{j=1}^{m}\left[M_{s_j} - M_{s_{j-1}} = k_j\right] | N)\right\} \\
&\quad = \quad E\left\{\frac{N!}{\prod_{j=1}^{m} k_j}\prod_{j=1}^{m}(F(s_j) - F(s_{j-1}))^{k_j} 1_{\{0\le k_j \le N,\ \sum_1^m k_j = N\}}\right\} \\
&\quad = \quad \frac{n!}{\prod_{j=1}^{m} k_j}\prod_{j=1}^{m}(F(s_j) - F(s_{j-1}))^{k_j}\exp(-\lambda)\frac{\lambda^n}{n!}1_{[\sum_1^m k_j = n]} \\
&\quad = \quad \frac{1}{\prod_{j=1}^{m} k_j}\prod_{j=1}^{m}\left[\lambda(F(s_j) - F(s_{j-1}))\right]^{k_j}\exp(-\prod_{j=1}^{m}\lambda(F(s_j) - F(s_{j-1})) \\
&\quad = \quad \prod_{j=1}^{m}\exp(-\lambda(F(s_j) - F(s_{j-1})))\frac{\left[\lambda(F(s_j) - F(s_{j-1}))\right]^{k_j}}{k_j!}.
\end{aligned}$$

Thus $(M_{s_1}, M_{s_2} - M_{s_1}, \ldots, M_{s_m} - M_{s_{m-1}})$ are independent Poisson random variables with parameters $(\lambda F(s_1), \lambda(F(s_2) - F(s_1)), \ldots, \lambda(1 - F(s_{m-1})))$ as claimed.

3. Suppose that $X_1, \ldots, X_n, \ldots$ are independent and identically distributed with $E|X_1| < \infty$, and let $S_n = X_1 + \cdots + X_n$. Let $\mu = E(X_1)$. Suppose that N is an integer-valued random variable independent of the X_i's.
(a) If $EN < \infty$, show that $E(S_N) = \mu E(N)$.
(b) If $Var(N) < \infty$, and $Var(X_1) = \sigma^2 < \infty$, compute $Var(S_N)$.
Remark: Consider Example 10.5.3, page 245, and Theorem 11.6.1, page 309, in this connection, noting that the random number of terms in these results need *not* be independent of the X_i's being summed.

   Solution: (a) $E(S_N) = E\{E(S_N|N)\} = E\{N\mu\} = \mu E(N)$.
(b) Now

$$\begin{aligned}
Var(S_N) &= EVar(S_N|N) + Var\left[E(S_N|N)\right] \\
&= E\{N\sigma^2\} + Var\left[N\mu\right] \\
&= \sigma^2 E(N) + \mu^2 Var[N].
\end{aligned}$$

4. Exercise 11.8.16, page 56, Wellner, Chapter 11, notes. If $H : \ell^\infty(T) \mapsto \mathbb{R}$ is bounded and continuous, and $K \subset \ell^\infty$ is compact, then for every

$\epsilon > 0$ there is a $\delta > 0$ such that: if $x \in K$ and $y \in \ell^\infty(T)$ with $\|y - x\|_T < \delta$, then $|H(y) - H(x)| < \epsilon$.

Solution: Suppose $K \subset \ell^\infty(T)$ is compact and H is continuous. Fix $\epsilon > 0$. For each $x \in K$, continuity of H implies existence of a $\delta = \delta_x \equiv \delta_{x,\epsilon} > 0$ such that $y \in B(x, \delta_{x,\epsilon}) \equiv \{y \in \ell^\infty(T) : \ \|y - x\|_T < \delta_{x,\epsilon}\}$ implies that $|H(y) - H(x)| < \epsilon/2$. Note that the family of open sets $\{B(x, \delta_{x,\epsilon}/2) : \ x \in K\}$ covers K. By compactness of K there is a finite sub-cover

$$B(x_1, \delta_{x_1}/2), \ldots, B(x_m, \delta_{x_m}/2).$$

Let $\delta = (1/2) \min_{1 \le i \le m} \delta_{x_i}$.

Now let $x \in K$ and suppose $y \in \ell^\infty(T)$ satisfies $\|x - y\|_T < \delta$. Thus $y \in B(x, \delta)$. Furthermore, $x \in B(x_i, \delta_{x_i}/2)$ for some $i \le m$. Then by the triangle inequality

$$\|y - x_i\|_T \ \le \ \|y - x\|_T + \|x - x_i\|_T < \delta + \delta_i/2 < \delta_{x_i}/2 + \delta_{x_i}/2 = \delta_{x_i}.$$

By choice of the ball $B(x_i, \delta_{x_i})$ it follows that

$$\begin{aligned} |H(y) - H(x)| &\le |H(y) - H(x_i)| + |H(x_i) - H(x)| \\ &\le \epsilon/2 + \epsilon/2 = \epsilon. \end{aligned}$$

Thus the claim holds for this choice of δ.