Stat 516, Homework 1

Due date: Thursday, October 5, at the beginning of class.

1. *Gamma-Poisson mixture*

   Recall that if $Z \sim \text{Poisson}(\lambda)$, then $\text{E}(Z) = \text{var}(Z) = \lambda$, where $\lambda > 0$ is the intensity parameter. This property of the Poisson distribution is restrictive since often in practice *excess-Poisson variation* is encountered, i.e. the variance of observed count data exceeds the mean. One possible solution to this discrepancy is to use a Gamma-Poisson mixture, which is generated by first randomly drawing $X \sim \text{Gamma}(\alpha, \beta)$, where $\alpha, \beta > 0$ (the parameterization is such that that $\text{E}(X) = \alpha/\beta$ and $\text{var}(X) = \alpha/\beta^2$) and then generating $Y \mid X \sim \text{Poisson}(X)$.

   (a) Derive $\text{E}(Y)$ and $\text{var}(Y)$ and show that $\text{E}(Y) < \text{var}(Y)$.

$$
\begin{aligned}
E[Y] &= E[E(Y|X)] \\
&= E[X] \\
&= \frac{\alpha}{\beta}
\end{aligned}
$$

$$
\begin{aligned}
Var[Y] &= E[Var(Y|X)] + Var[E(N|X)] \\
&= \frac{\alpha}{\beta} + \frac{\alpha}{\beta^2} \\
&= \frac{\beta\alpha + \alpha}{\beta^2} \\
&= \frac{\alpha(\beta + 1)}{\beta^2}
\end{aligned}
$$

$$
\begin{aligned}
E[Y] &< Var[Y] \\
\frac{\alpha}{\beta} &< \frac{\alpha(\beta+1)}{\beta^2} \\
\alpha\beta &< \alpha(\beta + 1) \\
\beta &< \beta + 1 \\
0 &< 1
\end{aligned}
$$

   (b) Derive the marginal distribution of Y.

$$P(Y=k|\alpha,\beta) = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}\frac{\lambda^k e^{-\lambda}}{k!}d\lambda$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)k!}\int_0^\infty \lambda^{\alpha+k-1}e^{-\lambda(\beta+1)}d\lambda$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)k!}\frac{\Gamma(\alpha+k)}{(\beta+1)^{a+k}}\int_0^\infty \frac{(\beta+1)^{a+k}}{\Gamma(\alpha+k)}\lambda^{\alpha+k-1}e^{-\lambda(\beta+1)}d\lambda$$
$$= \frac{\beta^\alpha}{\Gamma(\alpha)k!}\frac{\Gamma(\alpha+k)}{(\beta+1)^{a+k}}$$

$$= \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\Gamma(k+1)}\left(\frac{\beta}{\beta+1}\right)^\alpha\left(1-\frac{\beta}{\beta+1}\right)^k$$

2. *Brémaud, 1.7.1.*

   Let X_1 and X_2 be two independent random variables taking their values in $\{1, 2, \ldots, N\}$, and uniformly distributed, that is, $\Pr(X_i = k) = \Pr(X_2 = k) = \frac{1}{N}$, $1 \le k \le N$. Compute $\mathrm{E}[X_1 \mid \max\{X_1, X_2\}]$.

$$
\begin{aligned}
Z = max\{X_1, X_2\} \\
P(X_1 > Z|Z) &= 0 \\
P(X_1 = Z|Z) &= \frac{Z}{2Z-1}
\end{aligned}
$$

$$
\begin{aligned}
V = \{1, 2, ...Z-1\} \\
v \in V \\
P(X_1 = v|Z) &= \frac{1 - P(X_1 = Z|Z)}{Z-1} \\
P(X_1 = v|Z) &= \frac{1 - \frac{Z}{2Z-1}}{Z-1} \\
&= \frac{\frac{2Z-1}{2Z-1} - \frac{Z}{2Z-1}}{Z-1} \\
&= \frac{\frac{Z-1}{2Z-1}}{Z-1} \\
&= \frac{Z-1}{(Z-1)(2Z-1)} \\
&= \frac{1}{2Z-1}
\end{aligned}
$$

$$
\begin{aligned}
E[X_1|Z] &= \sum_{x=1}^{Z} p(x)x \\
&= \frac{Z^2}{2Z-1} \sum_{x=1}^{Z-1} \frac{x}{2Z-1} \\
&= \frac{Z^2 + \sum_{x=1}^{Z-1} x}{2Z-1}
\end{aligned}
$$

3. *Brémaud, 1.4.1: Exponential races*

   Let X_i, $i = 1, \ldots, n$, be independent exponential random variables with intensities/rates $\lambda_i > 0$, $i = 1, \ldots, n$ respectively. Let $Z = \min(X_1, \ldots, X_n)$ and $J = \arg\min_j X_j$. In other words, J is a random index corresponding to $Z = X_j$; J is well defined, because the probability that one or more X_i's attain the same value is 0. Show that Z and J are independent, and give their respective distributions.

   Hint: Start with $\Pr(J = k, Z \geq t)$, then condition on the random variable Z. You may find Brémaud's Exercise 4.2 (*Freezing a Random Variable*) useful, if you keep in mind that for any event A, $\Pr(A) = \mathrm{E}(1_A)$. This exercise states that:

   Let $X_1, ..., X_n$ be independent random variables with respective p.d.f.'s f_1,...,f_n. Show that

$$
\mathrm{E}[g(X_1, ..., X_n)] = \int_{-\infty}^{\infty} \mathrm{E}[g(y, X_2, ..., X_n)] f_1(y)\, dy
$$

   and that

$$
\Pr(X_1 \leq X_2, ..., X_1 \leq X_n, X_1 \leq x) = \int_{-\infty}^{x} \Pr(X_2 \geq y) \times ... \times \Pr(X_n \geq y) f_1(y)\, dy.
$$

$$
\begin{aligned}
F_Z(t) = P(Z \leq t) &= P(min(X_1, ...X_n) \leq t) \\
&= 1 - P(min(X_1, ...X_n) > t) \\
&= 1 - P(X_1 > t, ...X_n > t) \\
&= 1 - \prod_{n=1}^{N} 1 - CDF_{X_i}(t) \\
&= 1 - \prod_{n=1}^{N} e^{-t\lambda_i} \\
&= 1 - e^{-t\sum_{n=1}^{N} \lambda_i} \\
f_Z(t) = F_Z'(t) &= pdf_Z \\
&= \sum_{n=1}^{N} \lambda_i e^{-t\sum_{n=1}^{N} \lambda_i} \\
&\therefore\ Z \sim exp(rate = \sum_{n=1}^{N} \lambda_i)
\end{aligned}
$$

$$
\begin{aligned}
P(X_j \leq X_{i \neq j}, X_j \leq x) &= \int_{\infty}^{x} f_1(y) \prod_{i \neq j}^{N} P(X_i \geq y) dy \\
&= \int_0^x PDF_{X_j}(y) \prod_{i \neq j}^{N} (1 - CDF_{X_i}(y)) dy \\
&= \lambda_j \int_0^x \prod_{i=1}^{N} e^{-\lambda_i x} \\
&= \lambda_j \int_0^x e^{-x\sum_{i=1}^{N} \lambda_i} \\
&= \frac{\lambda_j}{\sum_{i=1}^{N} \lambda_i} (1 - e^{-x\sum_{i=1}^{N} \lambda_i})
\end{aligned}
$$

$$
\begin{aligned}
P(J = k) &= \int_0^{\infty} P(X_j = x, J = k) dx \\
&= \frac{\lambda_j}{\sum_{i=1}^{N} \lambda_i} (1 - e^{-\infty}) \\
&= \frac{\lambda_j}{\sum_{i=1}^{N} \lambda_i} \\
\therefore P(Z, J) &= P(Z)P(J)
\end{aligned}
$$

4. *Brémaud, 1.8.1: Applying the SLLN*

   (a) State Kolmogorov's Strong Law of Large Numbers (SLLN).

   The sample average converges almost surely to the expected value.

   $\overline{X}_n \overset{a.s.}{\to} \mu$ where $n \to \infty$

   or

   $P(\lim_{n \to \infty} \overline{X}_n = \mu) = 1$

(b) Let $S_1, S_2, \dots$ be a sequence of independent and identically distributed (i.i.d.) random variables, with $\Pr(0 < S_1 < \infty) = 1$ and $\mathbb{E}[S_1] < \infty$. For $t \geq 0$, let $N_t = \sum_{n=1}^{\infty} 1_{(0,t]}(T_n)$, where $T_n = S_1 + \cdots + S_n$. Prove that, almost surely,

$$\lim_{t \to \infty} \frac{N_t}{t} = \frac{1}{\mathbb{E}[S_1]}.$$