# MAST30001 Stochastic Modelling

## Tutorial Sheet 6

1. Yeast microbes from the air outside of a culture float by according to a Poisson process with rate 2 per minute. Each microbe that floats by joins the population of the culture with probability $p$ and with probability $1-p$ the microbe doesn't join the culture, and this choice is made independent from the times of arrival and choice to join of all other microbes.

   (a) What is the chance that exactly four outside microbes float by in the first 3 minutes?

   (b) What is the chance that exactly four outside microbes join the culture in the first 3 minutes?

   (c) Given that 7 outside microbes have floated by the culture in first 3 minutes, what is the chance that at least two of the seven join the culture?

   (d) Given that 7 outside microbes have floated by the culture in first 3 minutes, what is the chance that exactly 3 float by in the first 1 minute?

   (e) What is the chance that in the first 3 minutes, exactly four microbes join the culture and 3 float by that don't join the culture?

   Assume now that a second strain of yeast microbes independently float by the culture according to a Poisson process with rate 1, and each microbe joins the culture with probability $q$, analogous to the previous process.

   (f) What is the chance that exactly four yeast microbes from either strain float by in the first 3 minutes?

   (g) What is the chance that exactly four yeast microbes from either strain join the culture in the first 3 minutes?

   **Ans.** Let $N_t$ be the number of microbes that float by up to time to $t$ and let $M_t$ be the number that join the colony up to to time $t$. Then $N_t$ is Poisson mean $2t$ and $M_t$ is Poisson mean $2pt$, independent of the process $N_t - M_t$ which is Poisson mean $2(1-p)t$. Also conditional on a Poisson process equals $k$ at time $t$, the distribution of the $k$ points in the interval $(0, t)$ are the same as $k$ iid variables that are uniform on $(0, t)$. These facts and the description of the process imply we have the following answers:

   (a) $e^{-6}\frac{6^4}{4!}$ ($N_t$ Poisson process),

   (b) $e^{-6p}\frac{(6p)^4}{4!}$ ($M_t$ Poisson process),

   (c) $1 - 7p(1-p)^6 - (1-p)^7$ (description of the $M_t$ from $N_t$),

   (d) $\binom{7}{3}(1/3)^3(2/3)^4$ (conditional Poisson process description),

   (e) $e^{-6}(6p)^4(6(1-p))^3/(4!3!)$ (Poisson and independence of $M_t$ and $N_t - M_t$),

   Let $N_t$ and $M_t$ be respectively, the number of the first strain that float by, and the number that float by and join the culture up to time $t$. Let $K_t$ and $L_t$ be the analogous processes for the second strain. Then as before, all of these processes are

Poisson processes and with $N_t$ rate 2, $M_t$ rate $2p$, $K_t$ rate 1, $L_t$ rate $q$, and $N_t, M_t$ are independent of $K_t, L_t$. Because of independence, superposition of Poisson processes implies that $N_t + K_t$ and $M_t + L_t$ are Poisson processes with rates 3 and $2p + q$. Using these facts we find:

(f) $P(N_3 + K_3 = 4) = e^{-9}9^4/4!$,

(g) $P(L_3 + M_3 = 4) = e^{-3(2p+q)}(3(2p + q))^4/4!$.

2. In a Poisson process with rate $\lambda$, what is the joint density of the times of the first and second jumps? What is the joint density of the times of the $i$th and $j$th jump for $i < j$? Can you interpret these formulas similar to our discussion in lecture deriving the joint densities of order statistics?

   **Ans.** If $T_1, T_2$ are the times of the first and second jumps, then $T_1$ is exponential and $T_2 - T_1$ is exponential and independent of $T_1$ so we find the following formula for the conditional density: for $t_2 > t_1 > 0$,

$$f_{T_2|T_1}(t_2|t_1) = \lambda e^{-\lambda(t_2-t_1)},$$

   and so the joint density is, over the same range of $t_1, t_2$,

$$f_{T_2|T_1}(t_2|t_1)f_{T_1}(t_1) = \lambda^2 e^{-\lambda t_2}.$$

   A similar calculation, but replacing exponentials with gammas yields for $t_j > t_i > 0$,

$$f_{T_j,T_i}(t_j, t_i) = \lambda^2 \frac{(\lambda t_i)^{i-1}e^{-\lambda t_i}}{(i-1)!} \frac{(\lambda(t_j - t_i))^{j-i-1}e^{-\lambda(t_j-t_i)}}{(j-i-1)!}.$$

   To interpret this similar to our discussion of order statistics, the $\lambda^2$ arises as the "density" of having points at $t_i$ and $t_j$ [One way to understand this more rigorously is that

$$\lim_{h\to 0} \frac{\mathbb{P}(N_{t_i+h} - N_{t_i-h} > 0)}{2h} = \lambda.]$$

   and the two fractions are the (Poisson) chances of having exactly $i-1$ jumps in the interval $(0, t_i)$ and $j - i - 1$ jumps in the interval $(t_i, t_j)$.

3. Let $U_{(1)}, \ldots, U_{(n)}$ be order statistics of independent variables, uniform on the interval $(0, 1)$. For $0 < x < y < 1$ what is

   (a) $\mathbb{P}(U_{(1)} > x, U_{(n)} < y)$,

   (b) $\mathbb{P}(U_{(1)} < x, U_{(n)} < y)$,

   (c) $\mathbb{P}(U_{(k)} < x, U_{(k+1)} > y)$?

   **Ans.**

   (a) The event $\{U_{(1)} > x, U_{(n)} < y\}$ is the same as all the $U_i$'s are between $x$ and $y$ which occurs with chance $(y - x)^n$.

   (b) $\mathbb{P}(U_{(1)} < x, U_{(n)} < y) + \mathbb{P}(U_{(1)} > x, U_{(n)} < y) = \mathbb{P}(U_{(n)} < y) = y^n$ and then use the previous answer.

   (c) The event $\{U_{(k)} < x, U_{(k+1)} > y\}$ is the same as $k$ of the $U_i$'s are smaller than $x$ and the rest are larger than $y$, which occurs with probability

$$\binom{n}{k} x^k(1 - y)^{n-k}.$$

4. From Tutorial 1: If $N$ is geometric with parameter $p$ ($\mathbb{P}(N = j) = p(1-p)^j$, $j = 0, 1, 2, \ldots$) and given $N = n$, $X$ has density

$$f_{X|N}(x|n) = \frac{x^n e^{-x}}{n!},$$

what is the density of $X$? Another question: If $S$ is exponential with rate $\lambda$ and given $S = s$, $M$ is Poisson with mean $s$, then what is the distribution of $M$? A third question: If $K$ is Poisson with mean $\mu$ and given $K = k$, $J$ is binomial with parameters $k$ and $p$, then what is the distribution of $J$? Can you explain (or even derive) the answers to these three questions through superposition and thinning of Poisson processes?

**Ans.** Straightforward computing shows $X$ is exponential with rate $p$, $M$ is geometric (started from 0) with parameter $\lambda/(1+\lambda)$, and $J$ is Poisson with mean $p\mu$. To understand these identities in the Poisson process picture, we can think of $X$ as the first arrival in a thinned (with probability $p$) rate one Poisson process. $M$ as the number of arrivals in a rate one Poisson process before the first arrival of an independent rate $\lambda$ Poisson process, which is the same as the number of marked arrivals of the super-positioned process of rate $1+\lambda$, where the marking occurs with probability $\lambda/(1+\lambda)$. $K$ is the number of arrivals of a rate one Poisson process up to time $\mu$ and $J$ is the number of arrivals on this interval of the thinned (with probability $p$) process.

5. Customers enter a bank according to a Poisson process $(N_t)_{t\geq 0}$ with rate $\lambda = 10$ per hour and each customer makes a deposit or withdrawal. If $X_j$ is the amount brought in by the $j$th customer, assume that the $X_j$ are i.i.d. and independent of the arrivals of customers with distribution uniform on $\{-4, -3, \ldots, 4, 5\}$ (negative amounts correspond to withdrawals). Then the balance of the bank over $t$ hours is given by a compound Poisson process

$$Y_t = \sum_{j=1}^{N_t} X_j.$$

   (a) Draw a typical trajectory of the process $Y_t$.

   (b) Calculate the mean and variance of the money brought into the bank over an eight hour business day.

   (c) Use the central limit theorem to approximate the probability that the bank has a total balance greater than \$4500 over 100 business days.

**Ans.**

(a) The process is piecewise constant with jumps of size distributed according to $X_j$ at times of the jumps of a Poisson process.

(b) If $E[X_j] = \mu$ and $Var(X_j) = \sigma^2$, then formulas from lecture imply

$$E[Y_t] = \lambda t \mu, \quad Var(Y_t) = \lambda t(\sigma^2 + \mu^2).$$

Simple calculations show $\mu = 0.5$ and $\sigma^2 + \mu^2 = 8.5$. Thus setting $t = 8$ and $\lambda = 10$ in the formulas above, we have

$$E[Y_8] = 40, \quad Var(Y_8) = 680.$$

(c) If $W$ is the balance over 100 business days, then we can represent

$$W = \sum_{j=1}^{100} Y_8^{(j)},$$

where the $Y_8^{(j)}$ are i.i.d. having distribution $Y_8$. Thus the CLT says that $W$ is approximately normal with mean and variance read from (b) and so if $Z$ is standard normal, then

$$P(W > 4500) = P\left(\frac{W - 4000}{\sqrt{68000}} > \frac{500}{\sqrt{68000}}\right) \approx P(Z > 1.917) \approx 0.0276.$$

6. For $r > 0$ and $0 < p < 1$, let $N_t$ be a Poisson process with rate $\lambda = r\log(1/p)$ and $X_1, X_2, \ldots$ be i.i.d. with distribution

$$P(X_1 = k) = \frac{(1-p)^k}{k\log(1/p)}, \quad k = 1, 2, \ldots$$

Use moment generating functions to show that the compound Poisson variable

$$Y_t = \sum_{j=1}^{N_t} X_j$$

has the negative binomial distribution (started from zero) with parameters $rt$ and $p$; that is, that

$$P(Y_t = k) = \binom{k + rt - 1}{k}(1-p)^k p^{rt}, \quad k = 0, 1, 2, \ldots$$

**Ans.** By conditioning on $N_t$ and taking expectations, a computation shows that if $\phi_X$ is the moment generating function of $X_1$, then the moment generating function $\varphi_t$ of $Y_t$ is

$$\varphi_t(\theta) = \exp\{\lambda t(\varphi_X(\theta) - 1)\}.$$

We compute

$$\varphi_X(\theta) = \sum_{k\geq 1} \frac{e^{\theta k}(1-p)^k}{k\log(1/p)} = \frac{1}{\log(1/p)}\sum_{k\geq 1}\frac{(e^\theta(1-p))^k}{k} = \frac{-\log(1 - e^\theta(1-p))}{\log(1/p)};$$

the last equality is by Taylor series or integrating the geometric series and $\varphi_X$ is defined for $|e^\theta(1-p)| < 1$. Thus we find that over the same range of $\theta$,

$$\varphi_t(\theta) = \exp\left\{rt\log(1/p)\left(\frac{-\log(1 - e^\theta(1-p))}{\log(1/p)} - 1\right)\right\} = \left(\frac{p}{1 - e^\theta(1-p)}\right)^{rt}.$$

On the other hand, this is the same as the moment generating function of the negative binomial distribution in the problem, shown by computing Taylor series.