1. Let $N \sim Pois(\lambda t)$ be the number of passengers which arrive between buses and let $T_1, \ldots, T_N$ be the arrival times of these passengers in the interval $[0, t]$ where 0 is right after the previous bus departs and t is right before the next bus arrives, so that $W_i = t - T_i$ is the amount of time passenger T_i waits. Given N, each $T_i|N \sim Unif(0, t)$ so $W_i|N \sim Unif(0, t)$ and $E(W_i|N) = t/2$. Thus to find the total waiting time, apply Adam's law: $E(\sum_i W_i) = E(E(\sum_i W_i|N)) = E(Nt/2) = \lambda t^2/2$.

2. Let $N \sim Pois(\lambda t)$ be the number of earthquakes in $[0, t]$. Then the cumulative intensity is $Z_1 + \cdots + Z_N$. Using Adam's law, $E(E(\sum_i Z_i|N)) = E(\mu N) = \mu\lambda t$. Using Eve's law, $Var(\sum_i Z_i) = E(Var(\sum_i Z_i|N)) + Var(E(\sum_i Z_i|N)) = E(N\sigma^2) + Var(N\mu) = \sigma^2\lambda t + \mu^2\lambda t$.

3. $N(T)|T \sim Pois(\lambda T)$. By Adam's law, $E(N(T)) = E(E(N(T)|T)) = E(\lambda T) = \lambda\mu$. By Eve's law, $Var(N(T)) = E(Var(N(T)|T)) + Var(E(N(T)|T)) = E(\lambda T) + Var(\lambda T) = \lambda\mu + \lambda^2\sigma^2$. Finally, $Cov(T, N(T)) = E(TN(T)) - E(T)E(N(T)) = E(E(TN(T)|T)) - \mu^2\lambda = E(\lambda T^2) - \mu^2\lambda = \lambda E(T^2) - \mu^2\lambda = \lambda\sigma^2$.

4. By the chicken-egg story (or thinning), the distribution of work emails is $W \sim Pois(\lambda tp)$ while the distribution of personal emails is $P \sim Pois(\lambda t(1-p))$. Then the amount of time to respond to work emails is $T_1 + \cdots + T_W$ while the amount of time to respond to personal emails is $S_1, \cdots, S_P$.

   Thus the expected amount of time to answer all emails is $E(\sum_i T_i + \sum_j S_j) = E(\sum_i T_i) + E(\sum_j S_j) = E(E(\sum_i T_i|W)) + E(E(\sum_j S_j|P)) = E(W\mu_W) + E(P\mu_P) = \lambda tp\mu_W + \lambda t(1-p)\mu_P$. The variance is $Var(\sum_i T_i + \sum_j S_j) = Var(\sum_i T_i) + Var(\sum_j S_j) = E(Var(\sum_i T_i|W)) + Var(E(\sum_i T_i|W)) + E(Var(\sum_j S_j)) + Var(E(\sum_j S_j)) = E(W\sigma_W^2) + Var(W\mu_W) + E(P\sigma_P^2) + Var(P\mu_P) = (\lambda tp)\sigma_W^2 + \mu_W^2(\lambda tp) + (\lambda t(1-p))\sigma_P^2 + \mu_P^2(\lambda t(1-p))$.

5. (a) Given that there are n goals, each goal is made by team A with probability p and by team B with probability $1-p$. Let I_j be the indicator of whether goal j is a turnaround. Then $E(I_j) = P(I_j = 1) = P(G_{j-1} = A, G_j = B) + P(G_{j-1} = B, G_j = A) = 2p(1-p)$. Thus the expected number of turnarounds is $2p(1-p)(n-1)$ since there are n possible turnarounds.

   (b) In this Poisson process, the probability that the next goal is scored by team A is p. Thus the number of goals until the next turnaround is distributed as $G \sim FS(p)$. Then the time until the next turnaround is $T_1 + \cdots + T_G$ where T_i is the increment of time between the $(i-1)$ and i goals, which is distributed as $Expo(\lambda)$. Thus the expected time until the next $B \to A$ turnaround is $E(\sum_i T_i) = E(E(\sum_i T_i|G)) = E(G/\lambda) = 1/(\lambda p)$.

6. $N_t < n$ means there are fewer than n arrivals by time t so this is the same event as $T_n > t$. $N_t > n$ means there are more than n arrivals by time t so this is the same event as $T_n < t$. The analogous statements hold for the weak inequalities by arguments of the form $P(N_t \leq n) = 1 - P(N_t > n) = 1 - P(T_n < t) = P(T_n \geq t)$.

7. (a) Given that N claims were received, the arrival time of each claim is distributed as $Unif(0, t)$ over the period $[0, t]$, so the distribution of N_1 is $Bin(N, t_1/t)$.

(b) $E(W_1|N) = E(E(W_1|N_1, N)|N) = E(N_1\mu|N) = N\mu t_1/t$. $Var(W_1|N) = E(Var(W_1|N_1, N)|N)+$ $Var(E(W_1|N_1, N)|N) = E(N_1\sigma^2|N)+Var(N_1\mu|N) = \sigma^2 N t_1/t+\mu^2 N(t_1/t)(t_2/t)$.

8. (a) Given the posting time $T \sim Unif(0, 1)$, then the probability that it is unanswered is $P(Expo(\lambda_2) > 1-T|T) = 1-P(Expo(\lambda_2) \leq 1-T|T) = 1-(1-e^{-\lambda_2(1-T)}) = e^{-\lambda_2(1-T)}$. Thus, the probabilty that it is unanswered is $p = \int_0^1 e^{-\lambda_2(1-t)}dt = 1/\lambda_2(1-e^{-\lambda_2})$.

   (b) By the chicken-egg story, the number of unanswered and number of answered questions are independent with distributions $Pois(\lambda_1 p)$ and $Pois(\lambda_1(1-p))$ respectively.

9. (a) Let $t' < t'', t''' < t'''' \in [t_1, t_2)$ with (t', t'') and (t''', t'''') disjoint. Then given N, the distribution of the number of points whose x coordinate is in $[t', t'']$ is $Bin(N, (t''-t')/(t_2-t_1))$ and the number of points whose x coordinate is in $[t''', t'''']$ is $Bin(N, (t''''-t'''))$. Now by the chicken-egg story, the number of points between t' and t'' is distributed as $Pois(\lambda_{max}(t''-t'))$, the number of points between t''' and t'''' is distributed as $Pois(\lambda_{max}(t''''-t'''))$ and importantly these are independent, so disjoint intervals are independent. We may now assume WLOG that $t' = t_1, t'' = t_2$.

   Thus it only remains to show that the number of accepted points is distributed as $Pois(\int_{t_1}^{t_2} \lambda(t)dt)$. $\int_{t_1}^{t_2} \lambda(t)dt$ is the area under the curve $\lambda(t)$ from t_1 to t_2. Thus, since the points are distributed uniformly, then given N, the distribution of the number of points under $\lambda(t)$ is $Bin(N, \int_{t_1}^{t_2} \lambda(t)dt/\lambda_{max}(t_2-t_1))$. Thus by the chicken-egg story again, the distribution of the number of accepted points is $Pois(\lambda_{max}(t_2-t_1)\int_{t_1}^{t_2} \lambda(t)dt/\lambda_{max}(t_2-t_1)) = Pois(\int_{t_1}^{t_2} \lambda(t)dt)$ as desired.

   (b) $f(n, t_1, \ldots, t_n) = f(t_1, \ldots, t_n|n)P(N(t) = n)$ by the definition of conditional probability. Since $N(t) \sim Pois(\int_0^t \lambda(s)ds) = Pois(\lambda_{total})$, then $P(N(t) = n) = e^{-\lambda_{total}}\lambda_{total}^n/n!$. Given the number of arrivals (and the generation process), the probability that the first arrival occurs before time t_1 is $\int_0^{t_1} \lambda(t)/\lambda_{total}dt$ so the corresponding density is $\lambda(t_1)/\lambda_{total}$, noting that there are n, so the density from the first arrival is $n\lambda(t_1)/\lambda_{total}$.

10. (a) Knowing how many arrivals there are in an interval gives information about λ. However, given λ then disjoint intervals would be independent.

    (b) $Cov(N_1, N_2) = E(N_1N_2)-E(N_1)E(N_2) = E(E(N_1N_2|\lambda))-E(E(N_1|\lambda))E(E(N_2|\lambda)) = E(E(N_1|\lambda)E(N_2|\lambda)) - E(\lambda t)E(\lambda s) = E(\lambda^2 st) - E(\lambda)^2 st = stVar(\lambda) = st(a/b^2)$.

11. (a) In the superposition, we can consider the following story. First, we wait for an arrival from either process. Since the first arrival has waiting time $Expo(\lambda)$ in both processes, the waiting time for any arrival is $\min(Expo(\lambda), Expo(\lambda)) \sim Expo(2\lambda)$. Once this occurs, then we are again waiting for an arrival from either process. Now, one process has waiting time $Expo(2\lambda)$ while the other has waiting time $Expo(\lambda)$, so the waiting time for any arrival is $\min(Expo(\lambda), Expo(2\lambda)) \sim Expo(3\lambda)$. The remaining can be argued analogously by induction (e.g. if the waiting time for the n^{th} arrival is distributed as $Expo((n+1)\lambda)$, then $a+b = n+1$

and in one of the processes (whichever one had the arrival), the next arrival time increments by 1 so the waiting time for the next arrival is $\min(Expo((a+1)\lambda), Expo(b\lambda)) \sim Expo((a+b+1)\lambda)$ or $\min(Expo(a\lambda), Expo((b+1)\lambda)) \sim Expo((a+b+1)\lambda))$.

(b) When we are waiting for the first arrival from either process, it is equally likely that the first arrival is type 1 or type 2, agreeing with step 1. Once we have an arrival from a process, then the rate of that process increases by λ, corresponding to adding a ball with the same number as the one just removed, agreeing with steps $2, 3$.

12. (a) Since the probability that a toy has type j is p_j, the number of toys until collecting the j^{th} toy type is distributed as $N_j \sim FS(p_j)$. Thus $Y_j = W_1 + \cdots + W_{N_j}$ where each $W_i \sim Expo(1)$, so $Y_j|N_j \sim Gamma(N_j, 1)$. The summing out the N_j gives $f(y) = \sum_n f(y|n)P(N_j = n) = \sum_n 1/\Gamma(n)y^n/ye^{-y}q_j^{n-1}p_j = p_je^{-y}\sum_n(q_jy)^{n-1}/(n-1)! = p_je^{-y}e^{q_jy} = p_je^{-p_jy}$ which is the $Expo(p_j)$ PDF. We could also have come to this conclusion by realizing that since the arrivals are distributed as $Expo(1)$, we can thin the Poisson into n separate Poisson processes each with rate p_j. By the thinning story, we can view the $Expo(1)$ waiting time process as the superposition of n independent $Expo(p_j)$ processes.

(b) T is the time until every toy type has been collected at least once, or equivalently the time until collecting the last type of toy, requiring N toys to be collected total. The time to collect N toys is the sum of the times to collect each individual toy $X_1 + \cdots + X_N$. $E(T) = E(\sum_i X_i) = E(E(\sum_i X_i|N)) = E(N(1/1)) = E(N)$.

(c) $E(T) = \int_0^\infty P(T > t)dt = \int_0^\infty 1 - P(T \leq t)dt = \int_0^\infty 1 - \prod_i P(Y_i \leq t)dt = \int_0^\infty 1 - \prod_i(1 - e^{-p_jt})dt$.