We know that if each event of a Poisson process is independently classified as a type I event with probability p and as a type II event with probability $1-p$ then the counting processes of type I and type II events are independent Poisson processes with respective rates λp and $\lambda(1-p)$. Suppose now, however, that there are k possible types of events and that the probability that an event is classified as a type i event, $i = 1, \dots, k$, depends on the time the event occurs. Specifically, suppose that if an event occurs at time y then it will be classified as a type i event, independently of anything that has previously occurred, with probability $P_i(y)$, $i = 1, \dots, k$ where $\sum_{i=1}^{k} P_i(y) = 1$. Now we have

Proposition 1. *If $X_i(t)$, $i = 1, \dots, k$, represents the number of type i events occurring by time t then $X_i(t)$, $i = 1, \dots, k$, are independent Poisson random variables having means*

$$E[X_i(t)] = \lambda \int_0^t P_i(s) ds.$$

Problem 1. [Tracking the Number of HIV Infections]There is a relatively long incubation period from the time when an individual becomes infected with the HIV virus, which causes AIDS, until the symptoms of the disease appear. As a result, it is difficult for public health officials to be certain of the number of members of the population that are infected at any given time. Find an approximation model for this phenomenon, which can be used to obtain a rough estimate of the number of infected individuals. Suppose that

- individuals contract the HIV virus in accordance with a Poisson process whose rate λ is unknown,
- the time from when an individual becomes infected until symptoms of the disease appear is a random variable having a known distribution G,
- the incubation times of different infected individuals are independent.

Let $N_1(t)$ denote the number of individuals who have shown symptoms of the disease by time t . Also, let $N_2(t)$ denote the number who are HIV positive but have not yet shown any symptoms by time t.

Solution 1. Since an individual who contracts the virus at time s will have symptoms by time t with probability $G(t-s)$ and will not with probability $1 - G(t-s) = \bar{G}(t-s)$, it follows that $N_1(t)$ and $N_2(t)$ are independent Poisson random variables with respective means

$$E(N_1(t)) = \lambda \int_0^t G(t-s)\, ds = \lambda \int_0^t G(y)\, dy,$$

$$E(N_2(t)) = \lambda \int_0^t \bar{G}(t-s)\, ds = \lambda \int_0^t \bar{G}(y)\, dy.$$

Now, if we knew λ, then we could use it to estimate $N_2(t)$, the number of individuals infected but without any outward symptoms at time t, by its mean value $E[N_2(t)]$. However, since λ is unknown, we must first estimate it. Now, we will presumably know the value of $N_1(t)$, and so we can use its known value as an estimate of its mean $E[N_1(t)]$. That is, if the number of individuals who have exhibited symptoms by time t is n_1, then we can estimate that

$$n_1 \approx E(N_1(t)) = \lambda \int_0^t \bar{G}(y)\,dy.$$

Therefore, we can estimate λ by the quantity $\hat{\lambda}$ given by

$$\hat{\lambda} = \frac{n_1}{\int_0^t G(y)\,dy}.$$

Using this estimate of λ, we can estimate the number of infected but symptomless individuals at time t by

$$\text{estimate of } N_2(t) = \hat{\lambda} \int_0^t \bar{G}(y)\,dy = n_1 \frac{\int_0^t \bar{G}(y)\,dy}{\int_0^t G(y)\,dy}.$$

For example, suppose that G is exponential with mean μ. Then $\bar{G}(y) = e^{-\frac{y}{\mu}}$, and a simple integration gives that

$$\text{estimate of } N_2(t) = \frac{n_1 \mu (1 - e^{-\frac{t}{\mu}})}{t - \mu(1 - e^{-\frac{t}{\mu}})}.$$

If we suppose that $t = 16$ years, $\mu = 10$ years, and $n_1 = 220$ thousand, then the estimate of the number of infected but symptomless individuals at time 16 is

$$\text{estimate of } N_2(16) = \frac{220 \cdot 10(1 - e^{-1.6})}{16 - 10(1 - e^{-1.6})} = 218.96.$$

That is, if we suppose that the foregoing model is approximately correct (and we should be aware that the assumption of a constant infection rate λ that is unchanging over time is almost certainly a weak point of the model), then if the incubation period is exponential with mean 10 years and if the total number of individuals who have exhibited AIDS symptoms during the first 16 years of the epidemic is 220 thousand, then we can expect that approximately 219 thousand individuals are HIV positive though symptomless at time 16.