ull6b.tex **Week 6: pm, 1.11.2017**

Chapter VII. INSURANCE MATHEMATICS

I. Insurance Background

The idea of insurance is simple: it is the *spreading*, or *pooling*, of risk. The relevant theory is that of *collective risk*.

History.

Insurance can be traced back to antiquity (Greek and Roman times). Like much else, it disappeared, to be re-developed in Renaissance Italy (Genoa, 14th C.). It received a great impetus in the UK from the Great Fire of London in 1666; fire insurance had started there by 1681. Property insurance had begun in London by 1710, and in Philadelphia (Benjamin Franklin) in 1752.

Shipping insurance grew in London around Edward Lloyd's coffee house in the 1680s. He died in 1710; Lloyd's of London had developed by 1774.

John Graunt (1620-74) published his *Bills of Mortality* in 1662 (breaking down London deaths by cause, age etc.). This was followed by the first life table (Edmund Halley, 1693). Mutual life insurance had begun by 1762. One of the earliest such companies is Scottish Widows (1815) (founded to look after the widows of Presbyterian ministers who died in office, and had to leave the manse – the minister's house).

At a national level, national insurance began in Germany with Bismarck in the 1880s. It developed here with e.g. Lloyd George (pre-WWI), Beveridge and the Beveridge Report (1942), and the founding of the Welfare State post-WWII.

Limited liability.

Lloyd's of London pre-dates limited liability (which developed in the mid-19th C.). The Lloyd's participants, or *names*, had unlimited liability, and were liable for the full extent of losses, irrespective of their investment or their assets. This changed, following the Lloyd's scandal of the 1990s.

Insurance is now done (and most was before the Lloyd's scandal) by limited liability companies. So for these, the possibility of *ruin* is crucial. Not only would this wipe out the company, its assets and expertise, the jobs of its employees etc., but it would leave policy-holders without cover.

Reinsurance.

Because a run of large claims could bankrupt an insurance company, companies seek to lay off large risks – to reinsure – insure themselves – with larger, specialist reinsurance companies.

The question arises as to where reinsurane companies re-reinsure themselves ... This raises the modern form of Juvenal's question (*Satires*, c. 80 AD): *Quis custodiet ipsos custodes* – Who guards the guards? Who polices the police? Reinsurers reinsure insurers, but – who reinsures the reinsurers? – etc.

Regulation.

It is in the interest of some industries to agree to cover each other's liabilities in the event of a bankruptcy. For instance, this happens with *travel firms*. If a travel firm goes bust, leaving large numbers of people stranded abroad, or unable to travel on a foreign holiday booked and paid for, this would destroy public confidence in the whole industry – *unless* other firms, by prior agreement, step in to cover. This is what happens, and works well.

As motor insurance is compulsory by law, motor insurance companies are regulated by the state, and again, this provides a degree of protection in case of bankruptcy.

The actuarial profession.

People involved in the insurance industry have been known as *actuaries* from the early days of insurance. Companies offering insurance employ actuaries, and these need to be qualified. Actuaries become qualified by passing exams set by the *Institute of Actuaries*. London is an important centre for the actuarial/insurance industry, and so is Edinburgh. The mathematics involved is interesting, and useful. Those taking this course would be well advised to consider an actuarial career as one of their career possibilities.

Life v. non-life.

The usual way the modern insurance industry splits is between life and non-life. Life insurance is payable on death, and/or as an annuity ceasing on death. Life insurance is often combined with a mortgage (so that the mortgage is paid if one dies before it expires). Naturally, assessing premiums here depends on a detailed knowledge of mortality rates over ages, etc. The relevant mathematics is largely *Survival Analysis* – hazard rates, etc. Much use is made here nowadays of *martingale methods* (Ch. IV). Non-life splits again into categories: motor; house; (house) contents (these are the only three kinds of insurance ordinary people take out); (personal) accident (the next commonest); travel; commercial property; industrial; ... There are even catastrophe insurance, weather insurance etc. nowadays.

2. The Poisson Process; Compound Poisson Processes

The Poisson distribution.

This is defined on $\mathbb{N}_0 := \{0, 1, 2, \cdots\}$ for a parameter $\lambda > 0$ by

$$p_k := e^{-\lambda}\lambda^k/k! \quad (k = 0, 1, 2, \cdots).$$

From the exponential series, $\sum_k p_k = 1$, so this does indeed give a probability distribution (or law, for short) on $\mathbb{N}_0$. It is called the *Poisson distribution* $P(\lambda)$, with *parameter* λ, after S.-D. Poisson (1781-1840) in 1837.

The Poisson law has *mean* λ. For if N is a random variable with the Poisson law $P(\lambda)$, $N \sim P(\lambda)$, N has mean

$$E[N] = \sum kP(N = k) = \sum kp_k = \sum k.e^{-\lambda}\lambda^k/k! = \lambda \sum e^{-\lambda}\lambda^{k-1}/(k-1)! = \lambda,$$

as the sum is 1 (exponential series – $P(\lambda)$ is a probability law). Similarly,

$$E[N(N-1)] = \sum k(k-1)e^{-\lambda}\lambda^k/k! = \lambda^2 \sum e^{-\lambda}\lambda^{k-2}/(k-2)! = \lambda^2 :$$

$$var(N) = E[N^2] - (E[N])^2) = E[N(N-1)] + E[N] - (E[N])^2) = \lambda^2 + \lambda - (\lambda)^2 = \lambda :$$

the Poisson law $P(\lambda)$ with parameter λ has mean λ and variance λ.
Note. 1. The Poisson law is the commonest one for *count data* on $\mathbb{N}_0$.
2. This property – that the mean and variance are equal (to the parameter, λ) is very important and useful. It can be used as the basis for a test for Poissonianity, the *Poisson dispersion test.* Data with variance greater than the Poisson are called *over-dispersed*; data with variance less than Poisson are *under-dispersed.*
3. The variance calculation above used the (second) *factorial moment*, $E[N(N-1)]$. These are better for count data than ordinary moments.

The Exponential Distribution

A random variable T on $\mathbb{R}_+ := (0, \infty)$ is said to have an *exponential distribution* with *rate* (or parameter) λ, $T \sim E(\lambda)$, if

$$P(T \le t) = 1 - e^{-\lambda t} \qquad \text{for all } \ t \ge 0.$$

So this law has density

$$f(t) := \lambda e^{-\lambda t} \quad (t > 0), \quad 0 \quad (t \le 0)$$

(as $\int_{-\infty}^{t} f(u)du = P(T \leq t)$, as required). So the mean is

$$E[T] = \int tf(t)dt = \int_0^\infty \lambda te^{-\lambda t}dt = 1/\lambda. \int_0^\infty ue^{-u}du = 1/\lambda$$

(putting $u := \lambda t$). Similarly,

$$E[T^2] = \int t^2 f(t)dt = \int_0^\infty \lambda t^2 e^{-\lambda t}dt = 1/\lambda^2 \int_0^\infty u^2 e^{-u}du = 2/\lambda^2,$$

$$var(T) = E[T^2] - (E[T])^2 = 2/\lambda^2 - (1/\lambda)^2 = 1/\lambda^2.$$

The Lack-of-Memory Property.

Imagine components – lightbulbs, say – which last a certain *lifetime*, and are then discarded and replaced. Do we expect to see *aging*? With human lifetimes, of course we do! On average, there is much less lifetime remaining in an old person than in a young one. With some machine components, we also see aging. This is why parts in cars, aeroplanes etc. are replaced after their expected (or 'design') lifetime, at routine servicing. But, some components do *not* show aging. These things change with technology, but in the early-to-mid 20th C. lightbulbs typically didn't show aging. Nor in the early days of television did television tubes (not used in modern televisions!). In Physics, the atoms of radioactive elements show lack of memory. This is the basis for the concept of *half-life*: it takes the same time for half a quantity of radioactive material to decay as it does for half the remaining half, etc.

We can find *which* laws show no aging, as follows. The law F has the *lack-of-memory property* iff the components show no aging – that is, if a component still in use behaves as if new. The condition for this is

$$P(X > s + t | X > s) = P(X > t) \qquad (s, t > 0):$$

$$P(X > s + t) = P(X > s)P(X > t).$$

Writing $\overline{F}(x) := 1 - F(x)$ $(x \geq 0)$ for the *tail* of F, this says that

$$\overline{F}(s + t) = \overline{F}(s)\overline{F}(t) \qquad (s, t \geq 0).$$

Obvious solutions are

$$\overline{F}(t) = e^{-\lambda t}, \qquad F(t) = 1 - e^{-\lambda t}$$

for some $\lambda > 0$ – the exponential law $E(\lambda)$. Now

$$f(s+t) = f(s)f(t) \qquad (s, t \geq 0)$$

is a 'functional equation' – the *Cauchy functional equation* – and we quote that these are the *only* solutions, subject to minimal regularity (such as one-sided boundedness, as here – even on an interval of arbitrarily small length!).

So the exponential laws $E(\lambda)$ are *characterized* by the lack-of-memory property. Also, the lack-of-memory property corresponds in the renewal context to the *Markov property*. The renewal process generated by $E(\lambda)$ is called the *Poisson (point) process* with *rate* λ, $Ppp(\lambda)$. So: among renewal processes, the only Markov processes are the Poisson processes. We meet Lévy processes below: among renewal processes, the only Lévy processes are the Poisson processes.

It is the lack of memory property of the exponential distribution that (since the inter-arrival times of the Poisson process are exponentially distributed) makes the Poisson process the basic model for events occurring 'out of the blue'. Typical examples are accidents, insurance claims, hospital admissions, earthquakes, volcanic eruptions etc. So it is not surprising that Poisson processes and their extensions (compound Poisson processes) dominate in the actuarial and insurance professions, as well as geophysics, etc.

Gamma distributions.

Recall the *Gamma function*,

$$\Gamma(x) := \int_0^\infty t^{x-1} e^{-t} dt, \qquad (x > 0)$$

($x > 0$ is needed for convergence at the origin). One can check (integration by parts, and induction) that

$$\Gamma(x+1) = x\Gamma(x) \quad (x > 0), \qquad \Gamma(n+1) = n! \quad (n = 0, 1, 2, \cdots);$$

thus Gamma provides a *continuous extension to the factorial.* One can show

$$\Gamma(\frac{1}{2}) = \sqrt{\pi}$$

(the proof is essentially that $\int_{\mathbb{R}} e^{-\frac{1}{2}x^2} dx = \sqrt{2\pi}$, i.e. that the standard normal density integrates to 1). The Gamma function is needed for Statistics, as it commonly occurs in the normalisation constants of the standard densities.

The *Gamma distribution* $\Gamma(\nu, \lambda)$ with parameters $\nu, \lambda > 0$ is defined to have density

$$f(x) = \frac{\lambda^\nu}{\Gamma(\nu)}.x^{\nu-1}e^{-\lambda x} \qquad (x > 0).$$

This has MGF

$$\begin{aligned} M(t) := \int e^{tx} f(x) dx &= \frac{\lambda^\nu}{\Gamma(\nu)}.\int_0^\infty e^{tx}.x^{\nu-1}e^{-\lambda x}dx \\ &= \frac{\lambda^\nu}{\Gamma(\nu)}.\int_0^\infty x^{\nu-1}e^{-(\lambda-t)x}dx \\ &= \frac{\lambda^\nu}{\Gamma(\nu)}.\frac{1}{(\lambda-t)^\nu}\int_0^\infty u^{\nu-1}e^{-u}du \\ &= \Big(\frac{\lambda}{\lambda-t}\Big)^\nu \qquad (t < \lambda). \end{aligned}$$

Sums of exponential random variables.

Let $W_1, W_2, \dots W_n$ be independent exponentially distributed random variables with parameter λ ('W for waiting time' – see below): $W_i \sim E(\lambda)$. Then

$$S_n := W_1 + \cdots + W_n \sim \Gamma(n, \lambda).$$

For, each W_i has *moment-generating function* (MGF)

$$M(t) := E[e^{tW_i}] = \int_0^\infty e^{tx} f(x) dx = \int_0^\infty e^t x.\lambda e^{-\lambda x}dx$$

$$= \lambda.\int_0^\infty e^{-(\lambda-t)}dx = \lambda/(\lambda - t) \qquad (t < \lambda).$$

The MGF of the sum of independent random variables is the product of the MGFs (same for characteristic functions, CFs, and for probability generating functions, PGFs – check). So $W_1+\cdots+W_n$ has MGF $(\lambda/(\lambda-t))^n$, the MGT of $\Gamma 9, n, \lambda)$ as above:

$$S_n := W_1 + \cdots W_n \sim \Gamma(n, \lambda).$$

The Poisson Process

Definition. Let $W_1, W_2, \dots W_n$ be independent exponential $E(\lambda)$ random variables, $T_n := W_1, + \dots + W_n$ for $n \geq 1$, $T_0 = 0$, $N(s) := \max\{n : T_n \leq s\}$.

Then $N = (N(t) : t \geq 0)$ (or $(N_t : t \geq 0)$) is called the *Poisson process* (or *Poisson point process*) with *rate* λ, $Pp(\lambda)$ (or $Ppp(\lambda)$).

Interpretation: Think of the W_i as the waiting times between arrivals of events, then T_n is the arrival time of the nth event and $N(s)$ the number of arrivals by time s. Then $N(s)$ has a Poisson distribution with mean λs:

Theorem. If $\{N(s), s \geq 0\}$ is a Poisson process, then
(i) $N(0) = 0$;
(ii) $N(t+s) - N(s)$ is Poisson $P(\lambda t)$. In particular, $N(t) \sim P(\lambda t)$;
(iii) $N(t)$ has independent increments.
Conversely, if (i),(ii) and (iii) hold, then $\{N(s), s \geq 0\}$ is a Poisson process.

Proof. Part (i) is clear: the first lifetime is positive (they all are).

The link between the Poisson *process*, defined as above in terms of the exponential distribution, and the Poisson *distribution*, is as follows. First,

$$P(N_t = 0) = P(t < X_1) = e^{-\lambda t}.$$

This starts an induction, which continues (using integration by parts):

$$\begin{aligned}
P(N_t = k) &= P(S_k \leq t < S_{k+1}) = P(S_k \leq t) - P(S_{k+1} \leq t) \\
&= \int_0^t \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x} dx - \int_0^t \frac{\lambda^{k+1}}{\Gamma(k+1)} x^k e^{-\lambda x} dx \\
&= \int_0^t \Big[\frac{\lambda^k}{\Gamma(k+1)} . x^k - \frac{\lambda^{k-1}}{\Gamma(k)} . x^{k-1}\Big] d(e^{-\lambda x}) \\
&= \Big[\frac{\lambda^k}{\Gamma(k+1)} . t^k - \frac{\lambda^{k-1}}{\Gamma(k-1)} . t^{k-1}\Big] e^{-\lambda t} - \int_0^t e^{-\lambda x} \Big[\frac{\lambda^k}{\Gamma(k)} . x^{k-1} - \frac{\lambda^{k-1}}{\Gamma(k-1)} . x^{k-2}\Big] dx \\
&= \Big[\frac{\lambda^k}{\Gamma(k+1)} . t^k - \frac{\lambda^{k-1}}{\Gamma(k-1)} . t^{k-1}\Big] e^{-\lambda t} + \int_0^t e^{-\lambda x} \Big[\frac{\lambda^{k-1}}{\Gamma(k-1)} . x^{k-2} - \frac{\lambda^k}{\Gamma(k)} . x^{k-1}\Big] dx.
\end{aligned}$$

But the integral here is $P(N_t = k-1)$. So (passing from Gammas to factorials)

$$P(N_t = k) - e^{-\lambda t} \frac{(\lambda t)^k}{k!} = P(N_t = k-1) - e^{-\lambda t} \frac{(\lambda t)^{k-1}}{(k-1)!},$$

completing the induction. This shows that

$$N(t) \sim P(\lambda t).$$

This gives (ii) also: re-start the process at time t, which becomes the new time-origin. The re-started process is a new Poisson process, by the lack-of-memory property applied to the current item (lightbulb above); this gives (ii) and (iii). Conversely, independent increments of N corresponds to the lack-of-memory property of the lifetime law, and we know that this characterises the exponential law, and so the Poisson process. //

Time-dependent rates.

The parameter λ is called the *rate* or *intensity* of the Poisson process. Think of it as the rate at which accidents happen (or telephone calls arrive at an exchange), or the intensity of a bombardment, etc. The above extends to include time-dependent intensities. We say that $\{N(s), s \geq 0\}$ is a *Poisson process* with *rate* $\lambda(r)$ if
(i) $N(0) = 0$,
(ii) $N(t+s) - N(s)$ is Poisson with mean $\int_s^t \lambda(r)dr$, and
(iii) $N(t)$ has independent increments.

Limit Theory.

For independent, identically distributed (iid for short) random variables $X_1, X_2, \cdots$, the *sample mean* (a *statistic*: a function of the data – random, as the data is, but known, after sampling, when you have the data) is

$$\overline{X} := \frac{1}{n}\sum_1^n X_k.$$

The *mean*, or *population mean*, $E[X]$ is defined as in Measure Theory, though we can restrict here to the discrete and density cases – a weighted average $\sum x_k f(x_k)$ in the discrete case where X takes values x_k with probability $f(x_k)$, and in the density case by the continuous analogue $\int xf(x)dx$ when X has density f. Always, the sum or integral is *absolutely convergent*:

$$E[|X|] < \infty; \qquad \sum |x_k| f(x_k) < \infty; \qquad \int |x| f(x)dx < \infty.$$

One would expect that $\overline{X}$ would tend to $E[X]$ as the sample size n increases. This is exactly right. By Kolmogorov's Strong Law of Large Numbers of 1933 (SLLN, or just LLN for short), convergence takes place *with probability one* (*almost surely*, or a.s. for short):

$$\overline{X} \to E[X] \qquad (n \to \infty) \qquad a.s.$$

The Conditional Mean Formula

Theorem (Conditional Mean Formula. For $\mathcal{B}$ any σ-field,

$$E[E[X|\mathcal{B}]] = E[X].$$

Proof. In the tower property, take $\mathcal{C}$ the trivial σ-field $\{\emptyset, \Omega\}$. This contains no information, so an expectation conditioning on it is the same as an unconditional expectation. The first form of the tower property now gives

$$E[E[X|\mathcal{B}]\ |\{\emptyset, \Omega\}] = E[X|\{\emptyset, \Omega\}] = E[X]. \qquad //$$

The Conditional Variance Formula

Theorem (Conditional Variance Formula).

$$var(Y) = E[var(Y|X)] + var(E[Y|X]).$$

Proof. Recall $varX := E[(X - EX)^2]$. Expanding the square,

$$varX = E[X^2 - 2X.(EX) + (EX)^2] = E(X^2) - 2(EX)(EX) + (EX)^2 = E(X^2) - (EX)^2.$$

Conditional variances can be defined in the same way. Recall that $E(Y|X)$ is constant when X is known ($= x$, say), so can be taken outside an expectation over X, E_X say. Then

$$var(Y|X) := E(Y^2|X) - [E(Y|X)]^2.$$

Take expectations of both sides over X:

$$E_X var(Y|X) = E_X[E(Y^2|X)] - E_X[E(Y|X)]^2.$$

Now $E_X[E(Y^2|X)] = E(Y^2)$, by the Conditional Mean Formula, so the right is, adding and subtracting $(EY)^2$,

$$\{E(Y^2) - (EY)^2\} - \{E_X[E(Y|X)]^2 - (EY)^2\}.$$

The first term is $varY$, by above. Since $E(Y|X)$ has E_X-mean EY, the second term is $var_X E(Y|X)$, the variance (over X) of the random variable $E[Y|X]$ (random because X is). Combining, the result follows. //

Interpretation.
$varY$ = total variability in Y,
$E_X var(Y|X)$ = variability in Y not accounted for by knowledge of X,
$var_X E(Y|X)$ = variability in Y accounted for by knowledge of X.
In words:
variance = mean of conditional variance + variance of conditional mean, with these interpretations. This is extremely useful in Statistics, in breaking down uncertainty, or variability, into its contributing components. There is a whole area of Statistics devoted to such Components of Variance.

Compound Poisson Processes
We now associate i.i.d. random variables X_i with each arrival and consider

$$S(t) = X_1 + \ldots + X_{N(t)}, \qquad S(t) = 0 \text{ if } N(t) = 0.$$

Thus $S(t)$ is a *random sum* – a sum of a random number of random variables.

A typical application in the insurance context is a Poisson model of claim arrivals with random claim sizes The claims arrive at the epochs of a Poisson process with rate λ. The claims are independent (different motor accidents are independent; so are different house-insurance claims for fire damage, burglary etc.). Then the claim-total mean is the claim-number mean times the claim-amount mean. This is a special case of *Wald's identity* (below).

Theorem. (i) For N Poisson distributed with parameter λ and $X_1, X_2, \ldots$ independent of each other and of N, each with distribution F with mean μ, variance σ^2 and characteristic function $\phi(t)$, the compound Poisson distribution of

$$Y := X_1 + \ldots + X_N$$

has characteristic function $\psi(u) = \exp\{-\lambda(1-\phi(u))\}$, mean $\lambda\mu$ and variance $\lambda E[X^2]$.
(ii) For $N = (N_t)$ a compound Poisson process with rate λ and jump-distribution F with mean μ and variance σ^2, N_t has CF $\psi(u) = \exp\{-\lambda t(1-\phi(u))\}$, mean $\lambda t\mu$ and variance $\lambda t E[X^2]$.

Proof. (i) The characteristic function (CF) follows from

$$\psi(t) = E[e^{itY}] \;=\; E[\exp\{it(X_1 + \ldots + X_N)\}]$$

$$
\begin{aligned}
&= \sum_n E[\exp\{it(X_1 + \ldots + X_N)\}|N = n].P(N = n) \\
&= \sum_n e^{-\lambda}\lambda^n/n!.E[\exp\{it(X_1 + \ldots + X_n)\}] \\
&= \sum_n e^{-\lambda}\lambda^n/n!.(E[\exp\{itX_1\}])^n \\
&= \sum_n e^{-\lambda}\lambda^n/n!.\phi(t)^n \\
&= \exp\{-\lambda(1 - \phi(t))\}.
\end{aligned}
$$

We give two proofs for the mean and variance, (a) by differentiating the CF, (b) from the Conditional Mean and Conditional Variance Formulae. Recall that if X has CF ϕ,

$$\phi(t) = E[e^{iXt}].$$

Differentiating formally (this is justified here – we quote this),

$$\phi'(t) = E[iXe^{iXt}] : \qquad \phi'(0) = iE[X]; \qquad E[X] = -i\phi'(0);$$

$$\phi''(t) = E[-X^2 e^{iXt}] : \qquad \phi''(0) = -E[X^2]; \qquad E[X^2] = -\phi''(0).$$

(a) Differentiate the CF:

$$\psi'(t) = \psi(t).\lambda\phi'(t),$$

$$\psi''(t) = \psi'(t).\lambda\phi'(t) + \psi(t).\lambda\phi''(t).$$

By above, ($\phi(0) = 1$ and) $\phi'(0) = i\mu$, $\phi''(0) = -E[X^2]$,

$$\psi'(0) = \lambda\phi'(0) = \lambda.i\mu,$$

and as also $\psi'(0) = iEY$, this gives

$$E[Y] = \lambda\mu.$$

Thus the mean of the random sum $Y := X_1 + \cdots + X_N$ is the product of the means of X (short for a typical X_i) and N:

$$E[Y] := E[X_1 + \cdots + X_N] = E[X].E[N].$$

This is (a special case of) *Wald's identity* (Abraham Wald (1902-1950) in 1944). Similarly,

$$\psi''(0) = i\lambda\mu.i\lambda\mu + \lambda\phi''(0) = -\lambda^2\mu^2 - \lambda E[X^2],$$

and also ($\psi(0) = 1$, $\psi'(0) = i\lambda\mu$ and) $\psi''(0) = -E[Y^2]$. So

$$var\ Y = E[Y^2] - [EY]^2 = \lambda^2\mu^2 + \lambda E[X^2] - \lambda^2\mu^2 = \lambda E[X^2].$$

(b) Given N, $Y = X_1 + \ldots + X_N$ has mean $NEX = N\mu$ and variance $N\ var\ X = N\sigma^2$. As N is Poisson with parameter λ, N has mean λ and variance λ. So by the Conditional Mean Formula,

$$EY = E[E(Y|N)] = E[N\mu] = \lambda\mu.$$

By the Conditional Variance Formula,

$$\begin{aligned} var\ Y &= E[var(Y|N)] + var\ E[Y|N] \\ &= E[Nvar\ X] + var([N\ E[X]) \\ &= E[N].var\ X + var\ N.(EX)^2 \\ &= \lambda[E[X^2] - (E[X])^2] + \lambda.(E[X])^2 \\ &= \lambda E[X^2] = \lambda(\sigma^2 + \mu^2). \end{aligned}$$

(ii) Apply (i): N_t has mean λt and variance λt. //