

Markov cycle transition probabilities

Relationship with rates and adjustment of time intervals

Andrew J. Sims

2020-05-13

Relationship between rates and probabilities

Briggs *et al* (2006) use the expression $p = 1 - \exp(-rt)$, where r is described as an instantaneous rate and t is a time period of interest. This is derived by assuming that if there is a population of N patients (or atoms of a radioactive element), the rate of events (or radioactive decays) is proportional to the number of patients (or atoms), i.e. that events occur independently:

$$-\frac{dN}{dt} \propto N$$

The number of patients, $N(t)$ who have not experienced an event at time t (or number of undecayed atoms) is therefore given by the solution to the differential equation, i.e.:

$$N(t) = Ne^{-rt}$$

where r is the decay constant or rate, the reciprocal of the mean lifetime. The expected number of events \hat{K} which occur in interval t from a starting population of N is

$$\begin{aligned}\hat{K} &= N - Ne^{-rt} \\ \hat{K} &= Np\end{aligned}$$

where $p = 1 - e^{-rt}$ is the probability with which events arise during interval t .

Estimation of probabilities

Over a given time interval t , assuming the rate is constant, the number of events observed (K) will follow a Poisson point process with mean μ ,

$$P(K = k) = \frac{\mu^k e^{-\mu}}{k!}$$

From the properties of Poisson distributions, the expected number of events (or counts) in time interval t is equal to μ ; i.e. $\mu = \hat{K}$.

The instantaneous rate can be calculated from observations of the number of events (in a trial, say). If the trial population was N , the duration of follow-up was t and K events were observed, $p = K/N$, and $r = -\log(1 - p)/t$. This method allows the conversion of observed probabilities from one time interval to another. If probability p_0 was observed in a trial of duration t_0 , $r = -\log(1 - p_0)/t_0$, and the probability in a trial of duration t_1 is given by

$$\begin{aligned}p_1 &= 1 - e^{-rt_1} \\ &= 1 - e^{(\log(1 - p_0) \frac{t_1}{t_0})}\end{aligned}$$

For example, if the probability of events over 5 years was 5.8%, the probability over 1 year is $1 - \exp(\log(1 - 0.058) \times \frac{1}{5}) = 0.011879$ (1.19%).

Thus, given an instantaneous rate (r) and a time interval (t), the probability of an event during that interval can be estimated (p). If the starting population (N) is known, both the expected number of events in the

time interval (\hat{K}) and the distribution of the observed number of events in the time interval ($P(K = k)$) can be estimated. In practical Markov models, p is the per-cycle transition probability and N is the state population at the start of the cycle, from which the expected number of transitions to other states can be calculated.

Alternative approach using compounding

If the probability of an event in a given time interval is p and there are N patients, Np will have an event in the time interval. If the number of cycles in the time interval is n , it is required to calculate the rate per cycle, p_c , such that in n cycles, Np_c patients will have an event.

For $n = 2$, the number at risk for the first cycle is N , and the number of events in the first cycle is Np_c ; the number at risk at the start of the second cycle is $N - Np_c$ and $(N - Np_c)p_c$ have an event. The total number of patients having an event in both cycles is $Np_c + (N - Np_c)p_c$. Thus

$$\begin{aligned} Np_c + (N - Np_c)p_c &= Np \\ Np_c + Np_c - Np_c^2 &= Np \\ p_c^2 - 2p_c + p &= 0 \\ p_c &= 1 - \sqrt{1 - p} \end{aligned}$$

For example, if 200 patients out of $N = 1000$ followed-up have an event in 1 year, the 1-year probability of the event is $p = 0.2$. The rate for each 6 month cycle is $1 - \sqrt{1 - 0.2} = 0.1056$, so after 6 months, 105.6 patients have an event, and 894.4 patients are at risk at the start of the second cycle. $894.4 \times 0.1056 = 94.4$ have an event in the second cycle, and $105.6 + 94.4 = 200$ patients in total have an event.

In general, the number at risk and number of events at each cycle is as follows

Cycle	Number at risk	Number having event	Cumulative number
1	N	Np_c	Np_c
2	$N(1 - p_c)$	$N(p_c - p_c^2)$	$N(2p_c - p_c^2)$
3	$N(1 - 2p_c + p_c^2)$	$N(p_c - 2p_c^2 + p_c^3)$	$N(3p_c - 3p_c^2 + p_c^3)$
4	$N(1 - 3p_c + 3p_c^2 - p_c^3)$	$N(p_c - 3p_c^2 + 3p_c^3 - p_c^4)$	$N(4p_c - 6p_c^2 + 4p_c^3 - p_c^4)$
n	$N(1 - p_c)^{n-1}$	$Np_c(1 - p_c)^{n-1}$	$N(1 - (1 - p_c)^n)$

Thus p_c is calculated from p as follows:

$$\begin{aligned} N(1 - (1 - p_c)^n) &= Np \\ p_c &= 1 - (1 - p)^{\frac{1}{n}} \end{aligned}$$

which is a generalisation of the example given for $n = 2$. For values n greater than about 4, the two approaches are almost identical. Using the example in the previous section, if $p = 5.8\%$ over 5 years, then the one-year per-cycle probability p_c is equal to $1 - (1 - 0.58)^{\frac{1}{5}} = 0.011879$ (1.19%).

Briggs, Andrew, Karl Claxton, and Mark Sculpher. 2006. *Decision Modelling for Health Economic Evaluation*. Oxford, UK: Oxford University Press.