Many epidemiologic textbooks give the mathematical expression that links the cumulative incidence (CI) or "risk" function, or its complement the "survival" function, with the integral of the incidence density (ID) function. Of the 15 modern texts I have examined, only one derives the relationship. Unfortunately, the formal geometric and calculus-based derivation used does not provide any insight into 'why' or 'how' the exp function comes into it, so epidemiologists are forced to accept it as a mere mathematical 'fact. Here we derive the formula heuristically. By working through a simple example, we try to make clear the difference beween rate and risk, and the units involved, and when one is numerically close to the other.

1 Simplest case

We begin with an exercise which, unless explicitly given in the context of this formula, tends to perplex many first year epidemiology trainees. We base it on data from Ayas et al (2006). In a large study, the observed rate of reported percutaneous injuries (PIs) among residents/interns in obstetrics/gynecology (ob/gyn) programs was 94 injuries in 964 intern-months, or (to the first 2 significant digits) 0.10 injuries per intern-month. We ask students to assume uniform 250-work-hours each month, with injury rates of 0.1 per intern-month that are constant, both within and across the hours and months in question. We then ask them to "calculate the probability that an average-risk ob/gyn resident would suffer at least one PI by the end of 1, 6 and 12 months of experience." We do not explicitly describe each of the probabilities as a 'cumulative incidence' or 'risk, but we do tell them that if they prefer, they may calculate the (complementary) probability of '*surviving*' these lengths of work-time *without* a PI.

Many students readily volunteer answers of $0.1 \times 1 = 0.1 = 10\%$ and $0.1 \times 6 = 0.6 = 60\%$ for the 1- and 6-month risks, before realizing when they try to calculate the 12-month risk that it cannot be $0.1 \times 12 = 1.2 = 120\%$. And while they are unable to now give an exact 12-month risk, many are confident that the 1-month risk is indeed 0.1 or 10%.

They have all been taught very early on how 'person-time' rates are calculated, and that a rate, which has dimension events/person-time, is entirely different conceptually from a risk, which is a (dimensionless) proportion. It is interesting to try to understand why there is such difficulty going back and forth between the two, in appreciating whether the one-month risk is less than or more than 10%, and in estimating how much less than 120% the 12-month risk is!

2 More generally

One heuristic way to begin might be to imagine a physical or human system consisting of say 100 workstations, each one in *continuous operation*. Figure 1 shows the 12-month log for a system in which the physical devices (humans) failed (were injured), independently of each other and of the duration they had been operating, and where, if such events occurred, they were immediately replaced. The expected failure rate (incidence or incidence density) is the expected number of events (120) per 1200 device-months or person-months, 0.1 per device-month or person-month, or 1.2 per device-year or person-year of operation. As we will show below, one would expect approximately 70 of the 100 *initial* devices or operators to fail before the end of the year, so that the one-year risk is in fact considerably less than 100%. The 120 failures or injuries in that first year of the system occur in an average of 70 of the 100 *first generation* members, and in 34 of their 70 *replacements*, and in 12 of *their* 34 *replacements*, and so on. In all, it takes an average of 220 different (100 *initial*, plus 120 *replacement*) devices or humans to keep the 100 workstations in continuous operation for 1 year.

Some of the reasons for the disconnect is our propensity to think in terms of *individual devices* rather than the continuous device-time or person-time needed to maintain the service. In effect, the devicemoments or person-moments are entirely *interchangeable*. We tend to draw person time as separete parallel lines, as if a station belonged to a device or person, but the 'up-time' can be generated by having some replacement devices or persons use the same stations as others.

1-month risk

If one understands the Poisson distribution, and how exactly it is derived, it is easy to move from a failure rate (*ID*) to a 1-month or x-month *risk*: the number of device failures in a period of 1 device-month of operation (up-time) is a Poisson random variable, with possible values 0, 1, 2, .. , and the expected (mean) number of failures is $\mu = 0.1$. Thus the probability of no (zero) PI injuries or failures is $P[0] = \exp(-0.1) = 0.90484$, so the 1-month risk or cumulative incidence is $1 - 0.90484 = 0.09516$ or 9.516%; the x-month risk is obtained similarly, using $\mu = 0.1 \times 6 = 0.6$, to arrive at a risk of $1 - P[0] = \exp(-0.6) = 1 - 0.54881 = 0.45118 = 45.115\%$.

However, just as with the relationship between incidence density (failure rates) and risk, the Poisson distribution is seldom well explained in introductory or epidemiology biostatistics texts, and so many would not be further enlightened by this 'explanation.'

The key to understanding how the exp function is involved in the transition from PI *rate* to PI *risk* is to express the injury rate not as 0.1 per intern-month, but as 0.0004 injuries/intern-hour, or an average of 1 injury per 2500 intern-hours. (we could equally use the rate of failures of the physical devices). The number of events in such a small time unit is again a random variable with possible values $0, 1, 2, \ldots$ but because one intern-hour is so small, the chance of 1 event in that amount of experience is already very small, and the chance of 2 or more is less than 1 in 10 million. Thus, one can very accurately regard the 1-hour risk as $0.1 \times 0.004 = 0.1/250 = 0.0004$, and its complement, the 1-hour 'survival' probability, as $10.1/250 = 0.9996$. Thus, the 1-month survival probability can be approximated by $(10.1/250)^{250} = 0.90482 = 90.482\%$ and its complement, the risk or cumulative incidence, by 9.518%.

An *even more accurate approximation* to the survival probability can be obtained by further dividing the 250 hours into 15000 minutes, so that the injury rate is 0.1 per 15000 intern-minutes, and calculating $(1 - (0.1/15000)^{15000} = 9.484\%$ so that the 1-month risk is 9.516%. Subdiving the time units further does not change these decimal places; the function $(1 - 0.1/LargeNumber)^{LargeNumber}$ converges to a constant which is solely a function of the 0.1. The function is the exp function. Indeed, one formal definition of exp x is that it is the limit,

$$\lim_{N \to \infty} (1 + x/N)^N.$$

In our example, $x = -0.1$, and the exact survival probability, to 6 decimal places, is $\exp(-0.1) = 0.904837$.

6-month and 12-month risk

As in standard survival calculations, the 6-month survival probability $S_{0\to6}$ is the product of 6 conditional probabilities:

$$S_{0\to6} = S_{0\to1} \times S_{1\to2} \times S_{2\to3} \times S_{3\to4} \times S_{4\to5} S_5 \to 6.$$

In our example the constant PI rate implies that each $S_{t\to(t+1)}$ equals $\exp[-0.1 \times 1]$ and so the 6-month survival probability is

$$S_{0\to6} = \exp[-0.1 \times 1] \times \cdots \times \exp[-0.1 \times 1] = \exp[-0.6] = 0.548 = 54.8\%,$$

so that the 6-month risk is $100 - 54.8 = 45.2\%$.

Had the PI rate *varied over the period at risk*, say as an (equal-) step-function, starting at 0.05 PI/intern-month in month 1 and rising to 0.10 PI/intern-month in month 6, then the 6-month survival probability is again obtained by summing the area under the ID curve to obtain $\int_{t=0}^{6} ID[t]\,dt = 0.45$, and by then calculating

$$S_{0\to6} = \exp\left[-\int_{t=0}^{6} ID[t]\,dt\right] = \exp[-0.45] = 0.64$$

If, as appears to be the case, the injury rate is closer to 0.16 PI/intern-month when working an extended shift, and 0.08 PI/intern-month when working regular shifts, then the risk for a resident over 3 months of extended shift is $1 - \exp[-(0.16 \times 3)] = 38\%$. The corresponding PI risk for the 9 months on regular shifts is $1 - \exp[-(0.08 \times 9)] = 51\%$. The chance of escaping injury-free for the entire 12 months is $\exp[-\{(0.16 \times 3) + (0.08 \times 9)] = \exp[-1.2]$.

The above calculations further illustrate the '*interchangeability*' of the contributions to the integral involved in the cumulative incidence (CI), and the fact that the CI only depends on the integral itself: the overall 6- or 12-month risk is the same whether the higher- and lower-risk blocks of time are interspersed or contiguous: the overall risk is determined by the integral, also called the cumulative hazard $H[T] = \int_{t=0}^{t=T} h[t]\,dt$.

This exponential formula for $S[.]$ is the same as the one for the depreciation/appreciation of a financial fund, where $A_{t=0}$ is the amount at $t = 0$, and $\delta(t)$ / $\alpha(t)$ is the rate of depreciation/appreciation, expressed as a smooth function.

Depreciation: $A_{t=T} = A_0 \times exp[-\int_{t=0}^{t=T} \delta[t]\,dt]$. Appreciation: $A_{t=T} = A_0 \times exp[\ \int_{t=0}^{t=T} \alpha[t]\,dt]$.

3 Approximation to CI

The fact that the risk function is a 1:1 function of the integral of the incidence-density function has implications for when one can obtain acceptably accurate approximations to the risk. The $1 - \exp[-H]$ function can be closely approximated by H over the range $H = 0$ to $H = 0.1$, but this approximation becomes less accurate thereafter. As is shown by the following table

H:	0.05	0.10	0.20	0.30	0.50	1.00	1.50
$(1 - \exp[-H])$:	0.049	0.095	0.181	0.259	0.393	0.632	0.777
% over	3	5	10	16	27	58	93

The percentage over-estimation by using $\mathrm{CI}_{approx} = H$, rather than $\mathrm{CI}_{exact} = 1 - \exp[-H]$, is close to $50 \times H$.

Large values of H can arise from a low rate operating over a longer time-interval, or higher ones over a shorter one.

12-month log for a computer system (workplace) consisting of 100 workstations, represented by 100 horizontal white lines. The dots – if applicable – for a station represent the times at which the devices at that station failed (workers at that station were injured). Failed devices (injured workers) were immediately replaced, so that each station remained in continuous operation. Devices failed (workers were injured) independently of each other and of the duration they had been operating. On average, some 120 failures (injuries) occurred in 1200 operator-months of operation. Thus, the failure(injury) rate was 0.1/operator-month, or 1.2/operator-year.