1. Przyszła długość życia x-latka

Rozważmy osobę mającą x lat; oznaczenie: (x). Jej przyszłą długość życia oznaczymy $T(x)$, lub krótko T. Zatem $x+T$ oznacza całkowitą długość życia. T jest zmienną losową, której dystrybuantę oznaczymy $G(t)$:

$$G(t) = \Pr(T \leqslant t),\ t \geqslant 0.$$

Tak więc $G(t)$ oznacza prawdopodobieństwo, że osoba umrze w ciągu t lat. Będziemy zakładać, że dystrybuanta G jest znana oraz że jest ciągła i ma gęstość $g(t) = G'(t)$. Zatem

$$g(t)dt = \Pr(t < T < t + dt)$$

oznacza prawdopodobieństwo, że śmierć nastąpi w nieskończenie małym przedziale $[t, t + dt]$.
Prawdopodobieństwa i wartości oczekiwane można wyrażać używając funkcji g i G. Tradycyjna notacja aktuarialna jest jednak inna. Np. symbol ${}_tq_x$ oznacza prawdopodobieństwo, że x-latek umrze w ciągu t lat. Stąd

$${}_tq_x = G(t).$$

Zatem

$${}_tp_x = 1 - G(t).$$

jest prawdopodobieństwem, że x-latek przeżyje t lat. Następnie

$${}_{s|t}q_x = \Pr(s < T \leqslant s + t) = G(s + t) - G(s) = {}_{s+t}q_x - {}_sq_x = {}_sp_x - {}_{s+t}p_x$$

jest prawdopodobieństwem, że x-latek przeżyje s lat, a następnie umrze w ciągu t lat.

$${}_tp_{x+s} = \Pr(T > s + t|T > s) = \frac{1 - G(s + t)}{1 - G(s)},$$

jest z kolei prawdopodobieństwem warunkowym, że osoba przeżyje dalsze t lat po przeżyciu s lat.
Podobnie

$${}_tq_{x+s} = \Pr(T \leqslant s + t|T > s) = \frac{G(s + t) - G(s)}{1 - G(s)},$$

jest prawdopodobieństwem śmierci w ciągu t lat po przeżyciu s lat. Mamy tożsamości

$${}_{s+t}p_x = 1 - G(s + t) = [1 - G(s)] \cdot \frac{1 - G(s + t)}{1 - G(s)} = {}_sp_x \cdot {}_tp_{x+s},$$

$${}_{s|t}q_x = G(s + t) - G(s) = [1 - G(s)] \cdot \frac{G(s + t) - G(s)}{1 - G(s)} = {}_sp_x \cdot {}_tq_{x+s}.$$

Wartość oczekiwaną pozostałej długości życia x-latka, $\mathrm{E}(T)$ oznaczamy $\mathring{e}_x$.
Zatem

$$\mathring{e}_x = \int_0^\infty tg(t)dt.$$

Ponieważ ze wzoru na całkowanie przez części otrzymujemy:

$$\int_0^\infty [1-G(t)]dt = (1-G(t))t|_0^\infty + \int_0^\infty tg(t)dt = \int_0^\infty tg(t)dt,$$

więc

$$\overset{\circ}{e}_x = \int_0^\infty [1-G(t)]dt = \int_0^\infty {}_tp_x dt.$$

Jeżeli $t=1$, indeks t jest zawsze opuszczany w symbolach ${}_tp_x$, ${}_tq_x$, ${}_{s|t}q_x$. Zatem q_x jest prawdopodobieństwem, że x-latek umrze w ciągu roku, a ${}_{s|}q_x$ jest prawdopodobieństwem, że x-latek przeżyje s lat, a potem umrze w ciągu roku.

1.1. Intensywność wymierania

Intensywność (natężenie) wymierania dla x-latków w wieku $x+t$ lat określamy jako

$$\mu_{x+t} = \frac{g(t)}{1-G(t)} = -\frac{d}{dt}\ln[1-G(t)].$$

Wtedy mamy:

$$\begin{aligned}\Pr(t<T<t+dt) &= g(t)dt = \frac{g(t)}{1-G(t)}[1-G(t)]dt = \\ &= \mu_{x+t}{}_tp_x dt,\end{aligned}$$

a także:

$$\overset{\circ}{e}_x = \int_o^\infty tg(t)dt = \int_o^\infty t{}_tp_x\mu_{x+t}dt.$$

Z równości poprzedniej mamy dla małych s:

$$\Pr(t<T<t+s) \approx {}_tp_x\mu_{x+t}s.$$

Lewa strona tej równości to ${}_{t|s}q_x$, czyli ${}_tp_x{}_sq_{x+t}$. Stąd

$${}_tp_x{}_sq_{x+t} \approx {}_tp_x\mu_{x+t}s,$$

więc

$${}_sq_{x+t} \approx \mu_{x+t}s.$$

Interpretacja: *prawdopodobieństwo śmierci w wieku bliskim $x+t$ jest proporcjonalne do intensywności wymierania w tym wieku.*
Ponieważ z definicji

$$\mu_{x+t} = -\frac{d}{dt}\ln[1-G(t)] = -\frac{d}{dt}\ln({}_tp_x),$$

więc całkując otrzymamy

$$\ln({}_tp_x) = -\int_0^t \mu_{x+s}ds,$$

$${}_tp_x = e^{-\int_0^t \mu_{x+s}ds}.$$

1.2. Analityczne dystrybuanty zmiennej losowej T

Funkcję G nazywamy dystrybuantą analityczną, jeśli można ją określić prostym wzorem. Dawniej próbowano wyprowadzić wzory na $G(t)$ z pewnych podstawowych postulatów, na wzór fizyki. Obecnie te próby wydają się nieco naiwne i mistyczne.
Historycznie pierwszą próbą był postulat de Moivre'a (1724) zakładający, że rozkład zmiennej T jest jednostajny między wiekiem 0 (liczonym od momentu x) a wiekiem $\omega - x$, gdzie ω jest hipotetycznym maksymalnym wiekiem. Stąd otrzymujemy

$$g(t) = \frac{1}{\omega - x} \text{ dla } 0 < t < \omega - x,$$

i dalej

$$G(t) = \int_0^t \frac{1}{\omega - x} ds = \frac{t}{\omega - x},$$

skąd

$$\mu_{x+t} = \frac{g(t)}{1 - G(t)} = \frac{\frac{1}{\omega - x}}{1 - \frac{t}{\omega - x}} = \frac{1}{\omega - x - t}$$

dla $0 < t < \omega - x$. Jak widać intensywność wymierania jest rosnącą funkcją t.
W roku 1824 Gompertz postulował wykładniczy wzrost intensywności wymierania:

$$\mu_{x+t} = Bc^{x+t},\ t > 0,$$

co lepiej oddaje charakter procesu starzenia a ponadto usuwa założenie wieku maksymalnego ω.
Prawo Gompertza zostało uogólnione przez Makehama, który zaproponował wzór

$$\mu_{x+t} = A + Bc^{x+t},\ t > 0.$$

Szczególnym przypadkiem tych praw jest stała intensywność wymierania (przyjąć $c = 1$ w prawie Gompertza lub $B = 0$ w prawie Makehama). Wtedy

$${}_tp_x = e^{-\mu t}.$$

Jest to matematycznie proste lecz niezbyt realistyczne.
Ogólniej, przy założeniu prawa Makehama:

$$\begin{aligned} {}_tp_x &= \exp[-\textstyle\int_0^t (A + Bc^{x+s})ds] = \exp(-[As + \frac{B}{\ln c} c^{x+s}]_0^t) = \\ &= \exp[-At - \frac{B}{\ln c} c^x (c^t - 1)] = \exp(-At - mc^x(c^t - 1)), \end{aligned}$$

gdzie $m = B/\ln c$.
Kolejna propozycja pochodzi od Weibulla (1939):

$$\mu_{x+t} = k(x+t)^n,\ k, n > 0.$$

Wtedy prawdopodobieństwo przeżycia wynosi

$$\begin{aligned} {}_tp_x &= \exp\left(-k\frac{(x+s)^{n+1}}{n+1} \Big|_0^t\right) = \\ &= \exp(-\frac{k}{n+1}[(x+t)^{n+1} - x^{n+1}]). \end{aligned}$$

1.3. Obcięta długość życia x-latka

Określimy pewne zmienne losowe związane ze zmienną T.
$K = [T]$ jest liczbą *pełnych* lat przeżytych później przez x-latka (tzw. *obcięta długość życia*). K jest zmienną losową o wartościach całkowitych i jej rozkład określają wzory:

$$\Pr(K = k) = \Pr(k \leqslant T < k+1) = {}_k p_x q_{x+k}, \ k = 0, 1, \ldots$$

Wartość oczekiwaną zmiennej K oznaczamy e_x. Mamy:

$$e_x = \sum_{k=1}^{\infty} k \Pr(K = k) = \sum_{k=1}^{\infty} k {}_k p_x q_{x+k}.$$

Korzystając ze wzoru

$$\sum_{k=1}^{\infty} \sum_{i=1}^{k} a_{ik} = \sum_{i=1}^{\infty} \sum_{k=i}^{\infty} a_{ik},$$

możemy napisać

$$\begin{aligned} e_x &= \sum_{k=1}^{\infty} k \Pr(K = k) = \sum_{k=1}^{\infty} \sum_{i=1}^{k} \Pr(K = k) = \\ &= \sum_{i=1}^{\infty} \sum_{k=i}^{\infty} \Pr(K = k) = \sum_{i=1}^{\infty} \Pr(K \geqslant i), \end{aligned}$$

czyli

$$e_x = \sum_{k=1}^{\infty} \Pr(K \geqslant k) = \sum_{k=1}^{\infty} {}_k p_x.$$

Niech S oznacza tę część roku śmierci, którą przeżywa (x), tj.

$$T = K + S.$$

Zmienna losowa S ma ciągły rozkład między 0 i 1. Przybliżając jej wartość oczekiwaną przez $\frac{1}{2}$ mamy:

$$\overset{\circ}{e}_x \approx e_x + \frac{1}{2}.$$

Załóżmy, że K i S są niezależnymi zmiennymi losowymi, tj. że dystrybuanta zmiennej S pod warunkiem K nie zależy od K, czyli:

$$\Pr(S \leqslant u | K = k) = \frac{{}_u q_{x+k}}{q_{x+k}}$$

nie zależy od k, więc można napisać

$${}_u q_{x+k} = H(u) q_{x+k},$$

dla $k = 0, 1, \ldots$, $0 \leqslant u \leqslant 1$ i pewnej funkcji $H(u)$. Jeżeli założymy, że $H(u) = u$ (rozkład jednostajny), to mamy równość dokładną:

$$\overset{\circ}{e}_x = e_x + \frac{1}{2}$$

oraz

$$\mathrm{Var}(T) = \mathrm{Var}(K) + \frac{1}{12}$$

(bo $\mathrm{Var}(S) = \int_0^1 (x - \frac{1}{2})^2 dx = \frac{1}{12}$).

1.4. Tablice trwania życia

Rozkład prawdopodobieństwa przyszłej długości życia x-latka można zbudować w oparciu o odpowiednią *tablicę trwania życia*, zbudowaną na podstawie danych statystycznych. Tablica taka jest w zasadzie tablicą prawdopodobieństwa śmierci dotyczących kolejnych lat. Tworzy się tablice dla różnych grup populacji (np. wg płci, rasy, pokolenia). Wiek początkowy x może znacząco wpływać na taką tablicę. Np. niech x oznacza wiek w którym dana osoba kupiła ubezpieczenie na życie. Ponieważ ubezpieczenie oferuje się tylko ludziom zdrowym (czasem po badaniach lekarskich), więc należy oczekiwać, że osoba, która właśnie nabyła ubezpieczenie będzie lepszego zdrowia niż ta, która je kupiła kilkanaście lat temu (chociaż teraz wiek jest ten sam). To zjawisko uwzględniają tzw. *ścięte tablice trwania życia*. W takiej tablicy prawdopodobieństwa śmierci są stopniowane zależnie od wieku wejścia do systemu ubezpieczeń. Zatem $q_{[x]+t}$ jest prawdopodobieństwem śmierci (w ciągu roku) dla $x+t$-latka, gdzie x jest wiekiem wejścia. Mamy nierówności

$$q_{[x]} < q_{[x-1]+1} < q_{[x-2]+2} < \dots$$

Efekt selekcji zanika po kilku, powiedzmy r, latach po wejściu. Zakładamy więc, że

$$q_{[x-r]+r} = q_{[x-r-1]+r+1} = \dots = q_x.$$

Okres r nazywa się okresem selekcji.

1.5. Prawdopodobieństwo śmierci dla części roku

Rozkład zmiennej K i jej charakterystyki można wyznaczyć na podstawie tablic trwania życia. Np.

$${}_kp_x = p_x p_{x+1} p_{x+2} \cdots p_{x+k-1}.$$

Aby otrzymać rozkład zmiennej T czyni się założenia odnośnie prawdopodobieństwa śmierci ${}_uq_x$ lub intensywności wymierania μ_{x+u} w wieku $x+u$ (x całkowite, $0 < u < 1$). Podamy trzy takie założenia.
Założenie a: Liniowość ${}_uq_x$.
To założenie nazywa się **hipotezą jednostajności** (*uniform distribution of deaths*).
Zakładamy, że ${}_uq_x$ jest liniową funkcją u, tj. ${}_uq_x = uq_x$. Jak już wiadomo jest to przypadek gdy K i S są niezależne i S ma rozkład jednostajny. Wtedy

$${}_up_x = 1 - uq_x$$

oraz

$$\mu_{x+u} = -\frac{d}{du}\ln({}_up_x) = -\frac{d}{du}\ln(1-uq_x) = \frac{q_x}{1-uq_x}.$$

Prawdziwa jest także równość:

$${}_{n+u}p_x = (1-u)\cdot {}_np_x + u\cdot {}_{n+1}p_x,$$

gdyż

$${}_{n+u}p_x = {}_np_x\cdot {}_up_{x+n} = {}_np_x(1-uq_{x+n}) = {}_np_x(1-u+up_{x+n}) = {}_np_x(1-u)+u\cdot {}_{n+1}p_x.$$

Założenie b: μ_{x+u} jest stała.
To założenie nazywa się **hipotezą stałego natężenia zgonów** (*constant force of mortality*).
Zakładamy, że μ_{x+u} jest stała dla $0 < u < 1$, i oznaczamy tę wartość $\mu_{x+1/2}$.
Mamy

$$\mu_{x+1/2} = -\frac{d}{dt}\ln {}_tp_x,$$

czyli

$$\int_0^u \mu_{x+1/2} ds = -\ln {}_up_x, \ 0 < u < 1,$$

więc

$$u\mu_{x+1/2} = -\ln({}_up_x) \tag{1}$$

skąd dla $u = 1$ otrzymujemy:

$$\mu_{x+1/2} = -\ln p_x.$$

Ponadto z równości (1) mamy

$${}_up_x = e^{-u\mu_{x+\frac{1}{2}}} = \left(e^{-\mu_{x+\frac{1}{2}}}\right)^u,$$

więc

$${}_up_x = (p_x)^u.$$

Stąd uzyskujemy

$$\begin{aligned}\Pr(S \leqslant u|K = k) &= \frac{{}_uq_{x+k}}{q_{x+k}} = \frac{1 - {}_up_{x+k}}{1 - p_{x+k}} = \\ &= \frac{1 - (p_{x+k})^u}{1 - p_{x+k}}.\end{aligned}$$

Rozkład warunkowy zależy od K, więc zmienne S i K nie są teraz niezależne.
Przykład. Na podstawie tablic trwania życia wiemy, że q_{26} jest równe: dla mężczyzny 0,00116, a dla kobiety 0,00029. Obliczyć prawdopodobieństwo, że osoba 26-letnia przeżyje najbliższe pół roku.
Przyjmijmy *Założenie a.* Wtedy

$${}_{0,5}p_{26} = 1 - 0,5 \cdot q_{26},$$

więc dla mężczyzny ${}_{0,5}p_{26} = 0,99942$, a dla kobiety 0,999855.
Przy *Założeniu b* mamy

$${}_{0,5}p_{26} = \sqrt{p_{26}} = \sqrt{1 - q_{26}},$$

czyli tak samo: dla mężczyzny ${}_{0,5}p_{26} = 0,99942$, a dla kobiety 0,999855.
Założenie c: Liniowość ${}_{1-u}q_{x+u}$.
To założenie zaproponował Balducci. Stwierdza ono, że

$${}_{1-u}q_{x+u} = (1 - u)q_x.$$

Aby wyznaczyć μ_{x+u} obliczymy ${}_u p_x$ z równości

$$p_x = {}_u p_x \cdot {}_{1-u} p_{x+u},$$

$$\begin{aligned} {}_u p_x &= \frac{p_x}{{}_{1-u} p_{x+u}} = \frac{1-q_x}{1-(1-u)q_x}, \\ -\ln_u p_x &= \ln[1-(1-u)q_x] - \ln(1-q_x), \end{aligned}$$

więc

$$\mu_{x+u} = -\frac{d}{du}(\ln {}_u p_x) = \frac{q_x}{1-(1-u)q_x}.$$

Ponadto

$$\begin{aligned} \Pr(S \leqslant u | K = k) &= \frac{{}_u q_{x+k}}{q_{x+k}} = \frac{1 - {}_u p_{x+k}}{q_{x+k}} = \\ &= \frac{1 - \frac{1-q_{x+k}}{1-(1-u)q_{x+k}}}{q_{x+k}} = \\ &= \frac{-(1-u)q_{x+k} + q_{x+k}}{q_{x+k}[1-(1-u)q_{x+k}]} = \\ &= \frac{u}{1-(1-u)q_{x+k}}. \end{aligned}$$

Zatem zmienne losowe K i S nie są niezależne.