Utility Metric

$$U = \sum_{t=0}^{\infty} \beta^t u_t$$ ← agents maximize this (total lifetime utility)

Measuring the total; running sim infinitely

$$\sum_{t=0}^{\infty} \beta^t u_t = u_0 + \beta u_1 + \beta^2 u_2 + \beta^3 u_3 + \dots$$

Measuring only last day T, with stopping probability $(1-\beta)$

$$\begin{aligned} E[U_T] &= p(T=0)\,u_0 + p(T=1)\cdot u_1 + \dots \\ &= (1-\beta)u_0 + \beta(1-\beta)u_1 + \beta^2(1-\beta)u_2 + \dots \\ &= (1-\beta)\sum_{t=0}^{\infty}\beta^t u_t = (1-\beta)\,U \end{aligned}$$

⇒ equivalent up to a constant factor $(1-\beta)$

⇒ maximizing total lifetime utility is the same as maximizing the utility of your last day (with constant death propability)

What about maximizing average expected utility?

$$\begin{aligned} E\Big[\sum_{t=0}^{T} u_t / T\Big] &= p(T=0)\,u_0 + p(T=1)\Big(\frac{u_0+u_1}{2}\Big) + p(T=2)\,\frac{u_0+u_1+u_2}{3} + \dots \\ &= (1-\beta)u_0 + (1-\beta)\beta\,\frac{u_0+u_1}{2} + (1-\beta)\beta^2\,\frac{\dots}{3} + \dots \\ &= (1-\beta)\Big(u_0 + \beta\frac{u_0}{2} + \beta^2\frac{u_0}{3} + \beta^3\frac{u_0}{4} + \dots\Big) \\ &\quad + \beta(1-\beta)\Big(\frac{u_1}{2} + \beta\frac{u_1}{3} + \beta^2\frac{u_1}{4} + \dots\Big) + \dots \end{aligned}$$

⇒ different factor for every $\beta^t u_t$ ⇒ not the same

Maximizing the expected average utility leads to a different strategy than maximizing expected total utility!

(Intuition: otherwise a one-day life with daily utility 1 would be as good as a 1000-day life with the same daily utility.)
(Note that there is no difference in rankings when running the simulation only once.)

What about using exponentially moving averages?

$$\begin{aligned} E\Big[\sum_{t=0}^{T} \alpha^{T-t} u_t\Big] &= (1-\beta)u_0 + (1-\beta)\beta(\alpha u_0 + u_1) + (1-\beta)\beta^2(\alpha^2 u_0 + \alpha u_1 + u_2) + \dots \\ &= (1-\beta)\big(u_0 + \alpha\beta u_0 + (\alpha\beta)^2 u_0 + \dots\big) + (1-\beta)\beta\big(u_1 + \alpha\beta u_1 + \dots\big) + \dots \\ &= (1-\beta)\frac{u_0}{1-\alpha\beta} + (1-\beta)\beta\frac{u_1}{1-\alpha\beta} + \dots \\ &= \frac{(1-\beta)}{(1-\alpha\beta)}\Big(\sum_{t=0}^{\infty}\beta^t u_t\Big) = \frac{(1-\beta)}{(1-\alpha\beta)}\,U \quad ! \end{aligned}$$

⇒ nice, exp. average works!