1/14/2023:

I did some scratchwork to calculate the hazard function, but I've stalled on it for a while. Right now, I just want to continue learning survival analysis.

X. Math models in survival analysis

Describe relationship between exposure variable E and outcome variable D after controlling for the possible confounding and interaction effects of additional variables. $C_1, \ldots, C_p$

E: social network index SNI
D: survival time variable
C_1: AGE
C_2: systolic blood pressure SBP
C_3: chronic disease CHR
C_4: Quetelet's body size index QUET
C_5: social class SOCL

Model	Outcome
Survival analysis	Time to event w/ censoring
Linear regression	Continuous output
Logistic regression	Dichotomous output

No follow up time info.

One goal of survival analysis is to obtain some measure of effect describing the exposure-outcome relationship adjusted for relevant extraneous variables.

Model	Measure of effect
Linear regression	β regression coefficient
Logistic regression	e^{β} odds ratio
Survival analysis	e^{β} hazard ratio

Hazard ratio is the hazard given that you're exposed over the hazard given that you're not exposed.

XI: Censoring Assumptions

There are three assumptions about censoring:
- Independent censoring
- Random censoring
- Non-informative censoring

- Independent censoring is the most useful and affects validity.
- Random censoring is a stronger assumption. random ⇒ independent

Random censoring: subjects who are censored at time t should be representative of all the study subjects who remained at risk at time t with respect to their survival experience.

$$h_{\text{Censored}}(t) = h_{\text{Not Censored}}(t)$$

Independent censoring: Within any subgroup of interest, the
What? subjects who are censored at time t should be representative of all the subjects in that subgroup who remained at risk at time t w.r.t. their survival experience.

I need to be careful about how I design my experiment. ~~The way I have it now~~
I believe I should only track one message at a time. The reason is that if I track two or more messages, then they are competing for nodes, so the survival of one may make the extinction of another more likely.

Another thing to consider is what the ergodicity adjustment actually means. For an ε-BrAn(A, B) system, the smaller ε is, the slower the rate of convergence to the stationary distribution. So I may want to set ε to something non-negligible so that convergence is faster, maybe 10^{-3}.

Oh yes, also: tracking one message at a time means that the simulation runs longer, hence the initial state for a given message is more and more likely to be distributed according to the stationary distribution as the message number grows. This means that if someone manages to find an efficient way of calculating the stationary distribution, then these predictions can actually be tested.