1 Mathematical Background

1. For the general form of the exponential family distribution:

$$p(x;\theta) = \exp\left(\langle\theta, \phi(x)\rangle - T(\theta)\right),$$

the conjugate prior has the form:

$$p(\theta; b_0) = \exp(\langle n\nu, \theta\rangle + nT(\theta) - \psi(\nu, n).$$

- **Random variable:** θ
- **Natural parameter:** $b_0 = (n\nu, n)$

2 Prediction Market Simulation

2.1 Configuring Agents

There are three parameters associated with an agent: $\omega_{j,t}, \theta, c$.

$\omega_{j,t}$ is the private signal for the agent. It is based on the underlying truth at $t - \Delta t$. For instance, if we want to simulate the prediction market for a basketball game, first we need to retrieve the actual score over time (starting from time point 0). When we build the simulation (starting from time point 0), at time point t, the private signal of the agent j would be based on the real score at $t - \Delta t$. If the real score for team A and team B at $t - Deltat$ is $S_{teamA} : S_{teamB}$, then the private signal for the agent at simulation time t might be $S_{teamA}/(S_{teamA} + S_{teamB})$ (if the prediction market is based on event E: team A would win the game and if the agent predicts correctly, the payoff would be \$1).

θ is the weighting parameter affecting the agent's belief. The agent's belief integrates his private signal as well as the current price in the simulation prediction market. The equation is

$$x_{j,t} = (1-\theta)\left(\frac{ASK_t - BID_t}{2}\right) + \theta\omega_{j,t} \tag{1}$$

c is the budget associated with a single agent. The budget limits the amount of shares the agent buys when he enters the prediction market. Thus, when the last agent enters the market, he could not buy shares without any limit and push the price to exactly what he expects (in this situation, the prediction market fails to aggregate information from all buyers).

There are three types of agents in the prediction market: informed agent, less informed agent and noisy agent. Since we simulate the frequency of their arrival by exponential distribution, $\lambda_{informaed}, \lambda_{less\_informed}$ would be the parameters of exponential distribution for informed and less informed agents. The informed agent has a higher frequency of receiving the information (the actual score), that is, $\Delta t_{informed} < \Delta t_{less\_informed}$. Both informed and less informed agents belong to strategic agent, that is, the quantity they purchase is $min\{q_{budget}, q_{best}\}$. For noisy agent, they will buy as much as they could, that is, the quantity they purchase will always be q_{budget}.