Contents

1 Preliminary **2**

1.1 Poisson & Exponential Distribution . . . 2

1.1.1 Poisson Distribution . . . 2

1.1.2 Exponential Distribution . . . 2

1.1.3 Connection . . . 3

2 Inventory Models **3**

2.1 Garman's Model . . . 3

2.1.1 The Gambler's Ruin Problem . . . 5

2.1.2 The Market Maker's Ruin Problem . . . 5

2.2 The dealer's problem . . . 6

2.3 The intertemporal role of inventory . . . 7

Market Microstructure

Cliff

March 20, 2021

1 Preliminary

1.1 Poisson & Exponential Distribution

1.1.1 Poisson Distribution

A discrete random variable X is said to have a Poisson distribution with parameter $\lambda > 0$ if for $k = 0, 1, 2, \ldots,$ the probability mass function of X is given by:

$$f(k; \lambda) = \Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

where λ is the arrival rate, k is he number of occurrences, within a time interval.

$$Var(X) = E(X) = \lambda$$

1.1.2 Exponential Distribution

The probability density function of an exponential distribution is

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0, \end{cases}$$

and its cdf is

$$F(x; \lambda) = \begin{cases} 1 - e^{-(\lambda x)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

with Mean and Variance

$$\mathrm{E}[X] = \frac{1}{\lambda} \quad Var[X] = \frac{1}{\lambda^2}$$

1.1.3 Connection

- N_t: the number of arrivals during time period t.
- X_t: the time it takes for one additional arrival to arrive assuming that someone arrived at time t.

Be definition, the following conditions are equivalent:

$$(X_t > x) \equiv (N_t = N_{t+x})$$

By the complement rule, we also have

$$P(X_t \leq x) = 1 - P(X_t > x)$$

$$P(X_t \leq x) = 1 - P(N_{t+x} - N_t = 0)$$

But,

$$P(N_{t+x} - N_t = 0) = P(N_x = 0)$$

Using the poisson pmf the above where λ is the average number of arrivals per time unit and x a quantity of time units, simplifies to:

$$P(N_{t+x} - N_t = 0) = \frac{(\lambda x)^0}{0!} e^{-\lambda x}$$

Substituting in our original eqn, we have:

$$P(X_t \leq x) = 1 - e^{-\lambda x}$$

The above is the cdf of a exponential pdf.

2 Inventory Models

2.1 Garman's Model

Assumptions

- Dealer structure & double auction mechanism.
- Market maker only decide on the ask price, p_a and bid price, p_b.
- Orders are represented as independent stochastic process with stationary arrival function $\lambda_a(p)$ and $\lambda_b(p)$.
- Market maker is not permitted by borrowing or lending.
- Poisson order arrival rates essentially require that (1) there are a large number of agents in the market; (2) each agent acts independently in submitting order; (3) no agent can generate an infinite number of orders in a finite period; (4) no subset of agents can dominate order generation.

Order arriving theory

Let $N_a(t)$ (or $N_b(t)$) be the cumulative numbers of shares that have sold (or bought) to traders up to time t. Then inventories are governed by

$$I_c(t) = I_c(0) + p_a N_a(t) - p_b N_b(t)$$

and

$$I_s(t) = I_s(0) + N_b(t) - N_a(t)$$

Then define

$$Q_k(t) := \mathbb{P}[I_c(t) = k]$$
$$R_k(t) := \mathbb{P}[I_s(t) = k]$$

And if market maker hold k units of cash/stocks at time t, the previous state would be

- I = k-1, at t - Δt, with an sell order arrives in the next instance.
- I = k + 1, at t - Δt, with an buy order arrives in the next instance.
- I = k, at t - Δt, nothing happens in the next instance.

By assume the arrive rate as $\lambda_a p_a$, the probability of the market maker holds exactly k unit stocks at time t would be

- $Q_{k-1}(t - \Delta t)[\lambda_a(p_a)p_a\Delta t][1 - \lambda_b(p_b)p_b\Delta t]$
- $Q_{k+1}(t - \Delta t)[1 - \lambda_a(p_a)p_a\Delta t][\lambda_b(p_b)p_b\Delta t]$
- $Q_k(t - \Delta t)[1 - \lambda_a(p_a)p_a\Delta t][1 - \lambda_b(p_b)p_b\Delta t]$

Summing them up,

$$\begin{aligned} Q_k(t) = {} & Q_{k-1}(t - \Delta t)[\lambda_a(p_a)p_a\Delta t][1 - \lambda_b(p_b)p_b\Delta t] \\ & + Q_{k+1}(t - \Delta t)[1 - \lambda_a(p_a)p_a\Delta t][\lambda_b(p_b)p_b\Delta t] \\ & + Q_k(t - \Delta t)[1 - \lambda_a(p_a)p_a\Delta t][1 - \lambda_b(p_b)p_b\Delta t] \end{aligned}$$

To calculate the time derivative of the probability $Q_k(t)$, we take the limit as $\Delta t \to 0$ of $Q_k(t) - Q_k(t - \Delta t)/\Delta t$. This yields

$$\begin{aligned} \frac{\partial Q_k(t)}{\partial t} = {} & Q_{k-1}(t)[\lambda_a(p_a)p_a] + Q_{k+1}(t)[\lambda_{b(p_b)p_b}] \\ & - Q_k(t)[\lambda_a(p_a)p_a + \lambda_b(p_b)p_b] \end{aligned}$$

2.1.1 The Gambler's Ruin Problem

Let's denote $\mathbb{P}[F|S]$ as the failure probability while holding S units of stock.

$$\mathbb{P}[F|S] = q\mathbb{P}[F|S-1] + p\mathbb{P}[F|S+1]$$

The solution to this difference equation yields the general expected failure probability

$$\mathbb{P}[F|S_0] = \left(\frac{q}{p}\right)^{S_0}$$

2.1.2 The Market Maker's Ruin Problem

In the continuous time context

$$\begin{aligned} \lim_{t\to\infty} Q_0(t) \approx \left(\frac{\lambda_b(p_b)p_b}{\lambda_a(p_a)p_a}\right)^{I_c(0)/\bar{p}} &\quad \text{if } \lambda_a(p_a)p_a > \lambda_b(p_b)p_b, \\ = 1 &\quad \text{otherwise} \end{aligned}$$

where $\bar{p}$ is defined to be the average price.

$$\begin{aligned} \lim_{t\to\infty} R_0(t) \approx \left(\frac{\lambda_a(p_a)}{\lambda_b(p_b)}\right)^{I_s(0)} &\quad \text{if } \lambda_b(p_b) > \lambda_a(p_a), \\ = 1 &\quad \text{otherwise} \end{aligned}$$

Notes

We can substitute the above equations into equation at Sect 2.1.1

$$\left(\frac{qp_b}{\pi p_a}\right)^{\frac{w}{p}} = q\left(\frac{qp_b}{\pi p_a}\right)^{\frac{w-p_b}{\bar{p}}} + \pi\left(\frac{qp_b}{\pi p_a}\right)^{\frac{w+p_a}{\bar{p}}}$$

In this framework, the dealer's failure probability is always positive, and in some circumstances, the dealer fails with probability one.

So the requirements of bid& ask which dealer should set are:

$$\lambda_a(p_a)p_a > \lambda_b(p_b)p_b$$
$$\lambda_a(p_a) < \lambda_b(p_b)$$

It implies that the market maker would buy low (bid price) and sell high (ask price).

2.2 The dealer's problem

One way to analyze the dealer's decision making is through risk, utility and dealer's optimal behavior. And it can be regarded as the compensation of specialist service of providing immediacy. Instead of assuming market maker as a risk neutral monopolist in Garman's model, Stoll assumed dealer as a risk averse trader who need to be paid by taking risk which was the cost derived by the bid-ask spread.

The cost can be divided into three points: (1) holding costs from suboptimal portfolio position; (2) transaction cost from exchange fees, taxes, etc; (3) asymmetric information, some counter-parties might have more knowledge on stocks.

Theorem 1 (Stoll's 2-date model)**.** ***Assumption***

- *Unlimited lending/borrowing at risk free rate, R_f.*
- *Dealer holds a belief on "true" price.*
- *Dealer buys/sells the asset at time 1, and liquids the asset at time 2.*

Let's denote as

- W_0*: initial position of efficient optimal portfolio.*
- R_f*: risk-free rate.*
- Q_p*: true value of dealer's trading account with any remaining fund.*
- C_i*: dealer's exposure cost of holding sub optimal portfolio.*
- Q_i*: true value of stock position (true stock price $\times$ shares).*

Then the dealer's terminal wealth is presented as

$$\tilde{W} = W_0(1 + \tilde{R}^*) + (1 + \tilde{R}_i)Q_i - (1 + R_f)(Q_i - C_i) \tag{1}$$

The dealer is assumed to be willing to undertake any transactions that leaves his expected utility unchanged.

$$E[U(W_0(1 + \tilde{R}^*))] = E[U(\tilde{W})] \tag{2}$$

and simplify by taylor expansion,

$$\frac{C_i}{Q_i} = c_i = \frac{z}{W_0}\sigma_{ip}Q_p + \frac{1}{2}\frac{z}{W_0}\sigma_i^2 Q_i \tag{3}$$

- *z is the dealer's coefficient of relative risk aversion.*
- *Q_p dealer's total inventory.*
- *σ_{ip} is the correlation between the rate of return on stock i and the rate of return on the optimal efficient portfolio.*
- *σ_i^2 is the variance of stock i's return .*

In the end, the bid & ask prices can be expressed as

$$(P_i^* - P_b)/P_i^* = c_i(Q_i^b) \tag{4}$$

and

$$(P_a - P_b)/P_i^* = c_i(Q_i^b) - c_i(Q_i^a) = [z/W_0]\sigma_i^2|Q| \quad for \quad |Q_i^a| = |Q_i^b| = |Q| \tag{5}$$

where P_i^* is the "true" price.

This theorem illustrates several interest features of the bid-ask spread,

(1) The spread increases linearly with the trading size;

(2) this spread does not change in response to the inventory as an argument, but inventory decides the position of spread;

The larger inventory causes higher cost for absorbing more inventory, and verse vice.

2.3 The intertemporal role of inventory

In Stoll[1978], the model assumes that the true value of the stock is fixed at some value p, and so the dealer's prices can be written as $p_a = p + a$ and $p_b = p - b$. Without any transactions, the value of dealer's portfolio is

$$dX = r_X X dt + X dZ_X \tag{6}$$

where r_X is the mean return per unit of time, and Z_X is a Wiener process with mean zero and instantaneous variance rate σ_X^2.

And the value of cash account, F

$$dF = rFdt - (p - b)dq_b + (p + a)dq_a \tag{7}$$

where q_b and q_a are dealer buys and sales of securities, respectively. Similarly, the value of the dealer's inventory position, I, follows

$$dI = r_I Idt + pdq_b - pdp_a + IdZ_I \tag{8}$$

There're two interesting features,

(1) The inventory position doesn't change with the bid&ask price, but in its intrinsic price.

(2) The value of inventory doesn't change due to both changes in its size and changes in its value resulting from the diffusion term IdZ_I and the drift term $r_I Idt$.