

• Bid-ask spread과 수익률 이유

: MM가 호가를 제시할 때 거래 상대방이 informed? uninformed?

=> "adverse selection risk"

$$(\text{Bid-ask spread}) \sum = \frac{\alpha \mu}{\alpha \mu + 2\epsilon} \left( \overline{S} - \underline{S} \right)$$

이런 경우  
보통은  
호가를  
제출한다

정확한 가격  
시장에 도달하는  
가능

거래량  
— " —

이건 = 보통은  
상대방이  
informed

$$\alpha \begin{bmatrix} \text{정확한 가격의 비율} \\ \text{or 거래 상대방이} \\ \text{호가를 제출할 확률} \end{bmatrix} \times \begin{bmatrix} \text{비율} \end{bmatrix}$$



" "

$$PIN = \frac{\alpha \mu}{\alpha \mu + 2\varepsilon}$$

PIN

:  $\alpha, \mu, \varepsilon$  파라미터 추정

MLE  $\Rightarrow$  파라미터 추정  
 일관성, 효율성 갖는다.

$\Rightarrow$  "VPIN" : Volume-Synchronized PIN

- VPIN tries to capture  $\left( \frac{\alpha \mu}{\alpha \mu + 2\varepsilon} \right)$  in the real-time basis with complicated estimation procedure.

Volume size  
 $= \frac{1}{50} \cdot [\text{average daily volume}]$

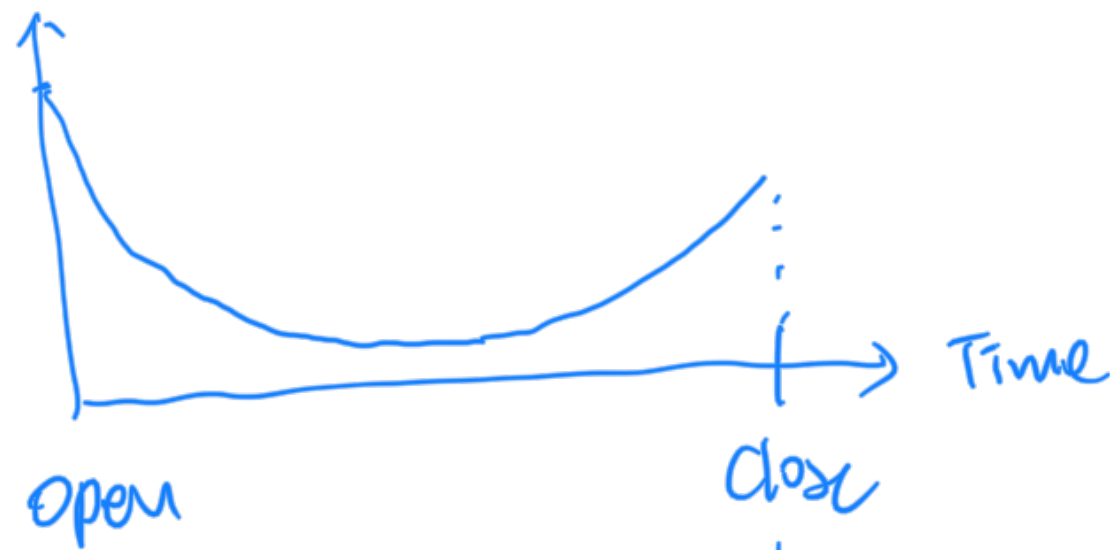
Step 1. Volume bucketing.

: 시간 단위로 나뉘어진 체결 (tick) 데이터 를 "equal volume bucket" 으로 그룹화한다.

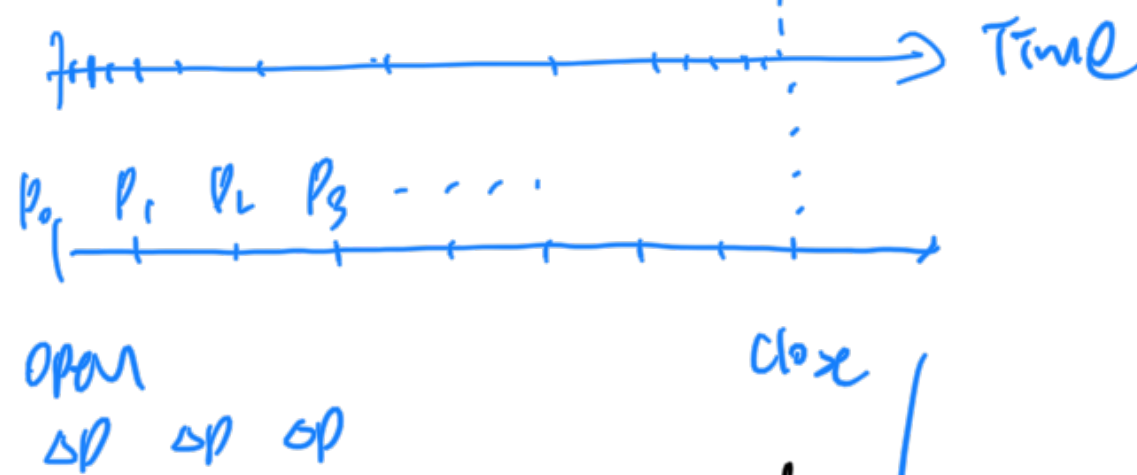
Panel-time intervals.

Note : Equal - Volume vs. Equal - Time

① Volume



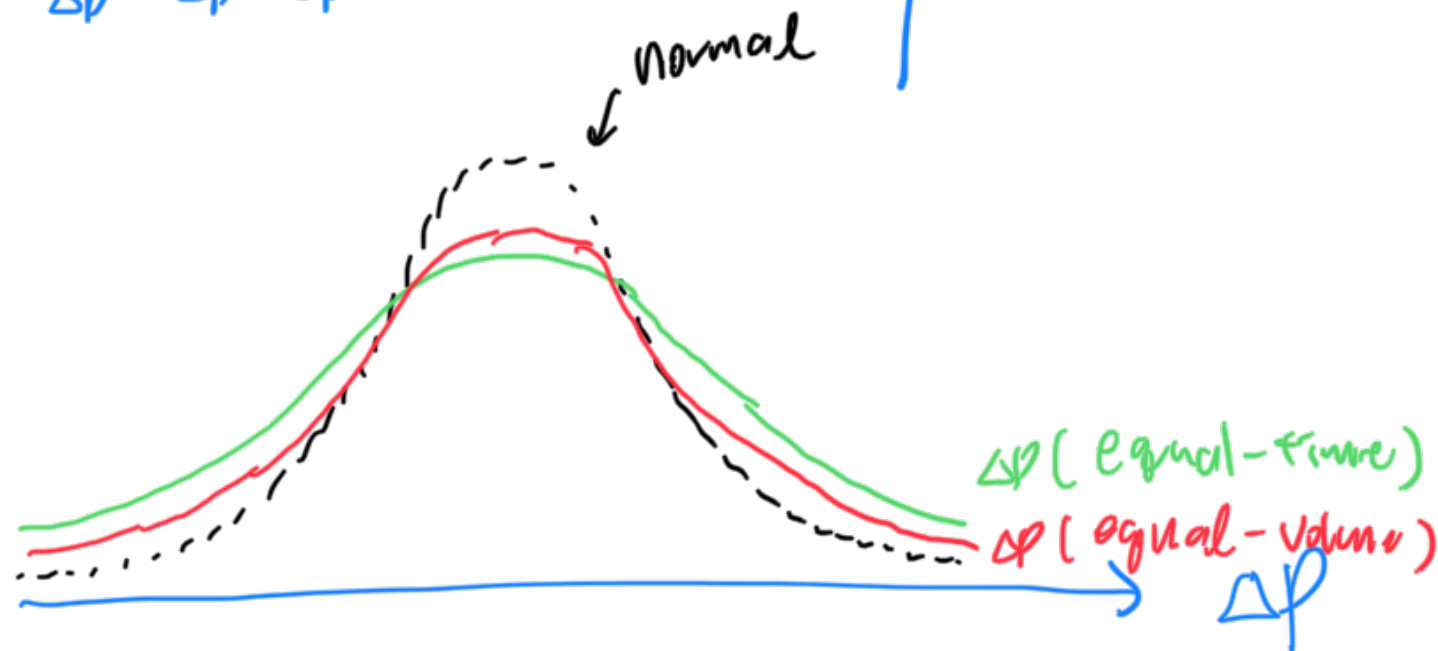
: U-shaped  
volume pattern



: Volume Interval

: Time Interval

②



③ equal-time interval에 비해 사람이 인식하기 어렵고  
각 interval이 언제 끝났는지 불확실하지만,  
동일한 정보량으로 인식을 수행할 수 있다.

(4) 거래가 많이 일어날 때 처리할 수 많은  
인식작업을 해야한다.

Step 2. Buy-volume / Sell-volume classification

: Step 1에서 자른 volume bucket 별로 각 bucket 내  
buy volume or sell volume 을 계산한다.

① Tick-based classification ( Lee-Ready algorithm )

↓  
tick + 최근의 order book 정보

$\perp A(t)$

$\wedge M(t)$

거래가격 :  $|X(t)$

$T B(t)$

if  $p(t) > M(t)$  :  $\sum_{i=1}^N t_{ij}^b$  buy volume trade  
(buyer-initiated trade)

else if  $p(t) < M(t)$  :  $\sum_{i=1}^N t_{ij}^s$  sell  $\sum_{i=1}^N t_{ij}^b$

else  $p(t) = M(t)$

if  $p(t) > p(t-1)$  : buy volume

else if  $p(t) < p(t-1)$  : sell volume

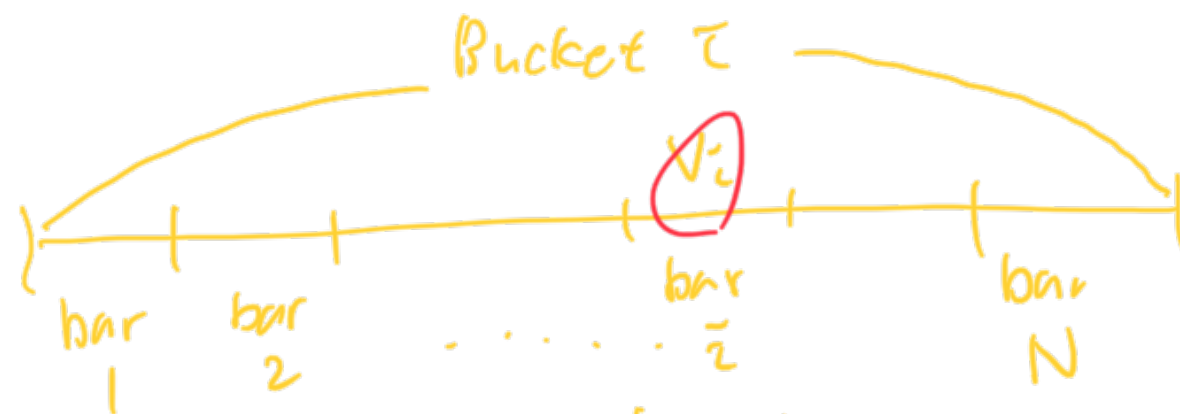
else  $p(t) = p(t-1)$  : undefined.

## ② Bulk classification



if  $p(t) > p(t-1)$  : buy volume  $\uparrow$   
if  $p(t) < p(t-1)$  : sell volume  $\downarrow$

$\rightarrow$  long min volume  $\rightarrow$  buy volume  $\downarrow$   
 sell volume  $\uparrow$



COF of standard normal dist.

$\sim N(0, 1^2)$

$\rightarrow$  buy volume  $V_i^B = V_i \times z \left( \frac{P_i - P_{i-1}}{\sigma_{op}} \right)$

$\rightarrow$  sell volume  $V_i^S = V_i \times \left( 1 - z \left( \frac{P_i - P_{i-1}}{\sigma_{op}} \right) \right)$

$= V_i - V_i^B$

$V_{\tau}^B = \sum_i V_i \times z \left( \frac{P_i - P_{i-1}}{\sigma_{op}} \right)$

$V_{\tau}^S = \sum_i V_i \times \left( 1 - z \left( \frac{P_i - P_{i-1}}{\sigma_{op}} \right) \right) = V - V_{\tau}^B$



Step 3 : Calculate VPIN

;  $n$  개 bucket on order imbalance  $\frac{1}{2}$   
이동 평균을  $2$  사용한다.

$$VPIN = \frac{1}{n} \sum_{i=1}^n \frac{|V_i^B - V_i^S|}{V}$$

$\nwarrow$  bucket size

$$E[|V_i^B - V_i^S|] \approx \alpha \mu$$

$$E[V_i^B + V_i^S] \approx \alpha \mu + 2\varepsilon.$$

$$VPIN \approx \frac{\alpha \mu}{\alpha \mu + 2\varepsilon}$$

• (1) VPIN  $\propto$  Intraday order volume bucket size  $\frac{1}{2}$

가MLE가, 이 분포상에서  $\theta_1$ 이 엠펙시미트 됨.

② MLE와 같은 분포상에서 추정 방식  $X$ .