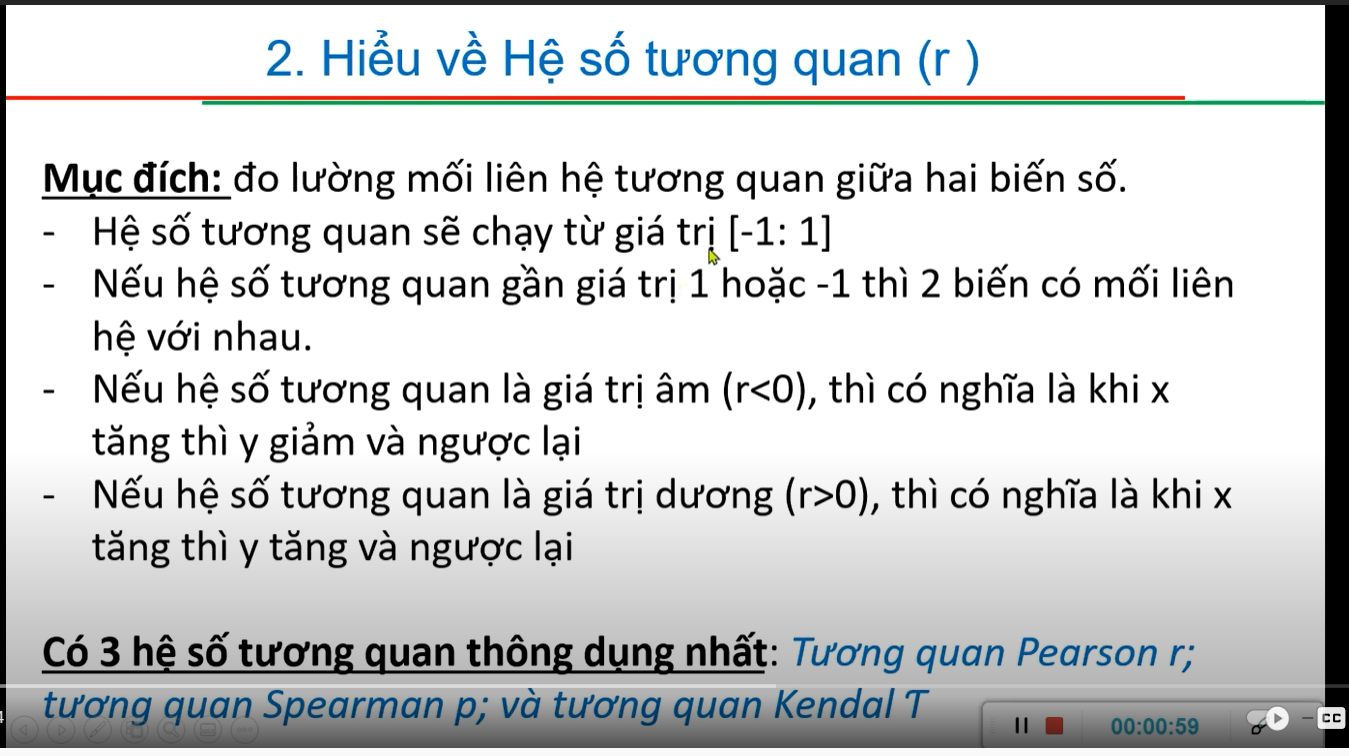
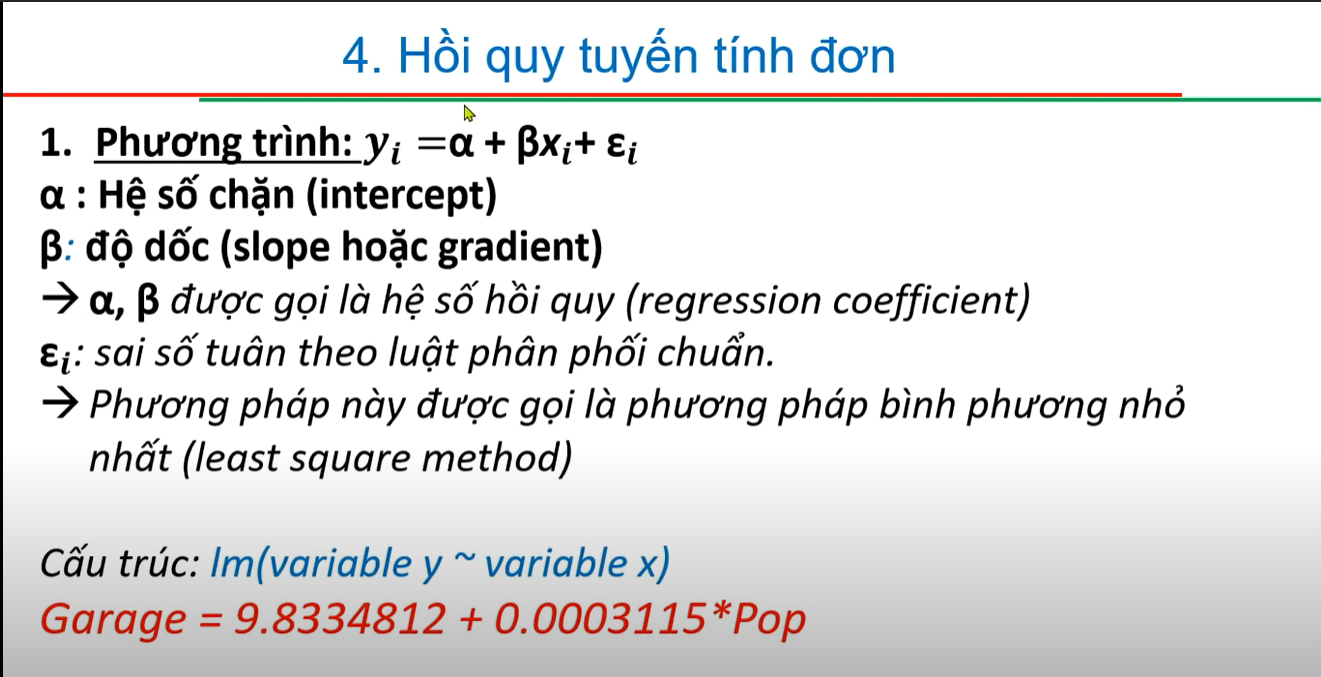
**Hệ số tương quan** là chỉ số thống kê đo lường mức độ mạnh yếu của mối quan hệ giữa hai biến số.

1. Hệ số tương quan dùng để xác định mức độ hiệu quả của một quĩ tương hỗ so với chỉ số chuẩn của nó (benchmark index), hoặc các tài sản hoặc quĩ khác. Bằng cách thêm một quĩ tương hỗ thấp có mức độ tương quan thấp hoặc tương quan nghịch vào danh mục đầu tư hiện có, nhà đầu tư đạt được lợi ích đa dạng hóa.

Rsquare= RSS/TSS





## Ý nghĩa của độ lệch chuẩn

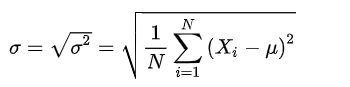
Độ lệch chuẩn đo tính biến động của giá trị mang tính thống kê. Nó cho thấy sự chênh lệch về giá trị của từng thời điểm đánh giá so với giá trị trung bình.

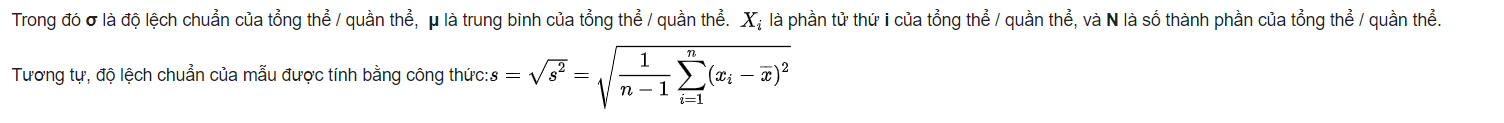
Tính biến động cũng như độ lệch chuẩn sẽ cao hơn nếu giá đóng cửa và giá đóng cửa trên trung bình khác nhau đáng kể.

Nếu sự chênh lệch không đáng kể thì độ lệch chuẩn và tính biến động ở mức thấp. Sự đảo chiều xu thế tạo các vùng đáy hoặc đỉnh của thị trường được xác định thời cơ bằng các mức độ biến động cao.

Những xu thế mới của giá sau thời kỳ thoái trào của thị trường (tức là giai đoạn điều chỉnh) thường được xác định thời cơ bằng những mức độ biến động thấp. Sự thay đổi đáng kể về dữ liệu giá đem lại giá trị độ lệch chuẩn cao và dữ liệu giá ổn định hình thành độ lệch chuẩn ở mức thấp.

Độ lệch chuẩn là căn bậc hai của phương sai. Do đó, công thức của độ lệch chuẩn của tổng thể / quần thể là:

{\displaystyle \sigma ={\sqrt {\sigma ^{2}}}={\sqrt {{\frac {1}{N}}\sum \_{i=1}^{N}\left(X\_{i}-\mu \right)^{2}}}} 



Ta cần phân biệt rõ 2 ký hiệu:

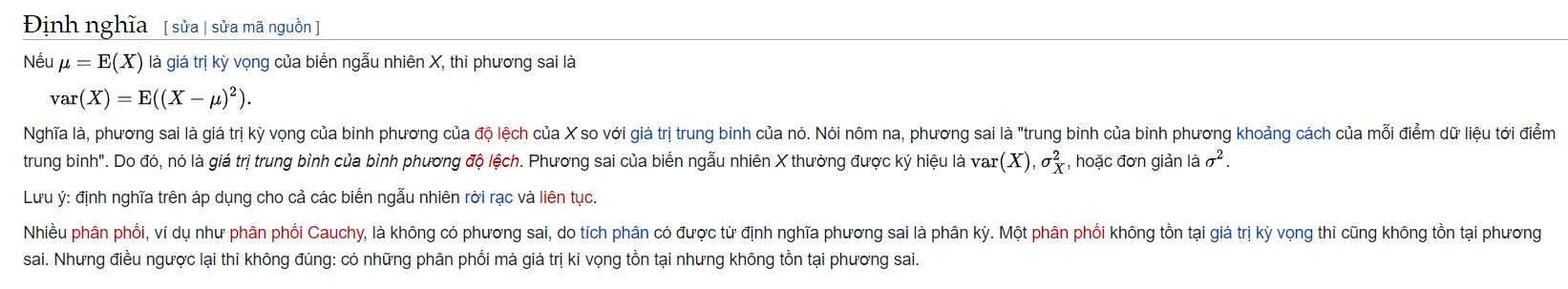
* **σ**: Dùng khi nói về quần thể
* **s**: Dùng khi nói về mẫu

❤️

1

1. Trong [lý thuyết xác suất](https://vi.wikipedia.org/wiki/L%C3%BD_thuy%E1%BA%BFt_x%C3%A1c_su%E1%BA%A5t" \o "Lý thuyết xác suất) và [thống kê](https://vi.wikipedia.org/wiki/Khoa_h%E1%BB%8Dc_Th%E1%BB%91ng_k%C3%AA" \o "Khoa học Thống kê), **phương sai** của một [biến ngẫu nhiên](https://vi.wikipedia.org/wiki/Bi%E1%BA%BFn_ng%E1%BA%ABu_nhi%C3%AAn" \o "Biến ngẫu nhiên) là một độ đo sự [phân tán thống kê](https://vi.wikipedia.org/w/index.php?title=Ph%C3%A2n_t%C3%A1n_th%E1%BB%91ng_k%C3%AA&action=edit&redlink=1" \o "Phân tán thống kê (trang chưa được viết)) của biến đó, nó hàm ý các giá trị của biến đó thường ở cách [giá trị kỳ vọng](https://vi.wikipedia.org/wiki/Gi%C3%A1_tr%E1%BB%8B_k%E1%BB%B3_v%E1%BB%8Dng" \o "Giá trị kỳ vọng) bao xa.

Phương sai của biến ngẫu nhiên giá trị [thực](https://vi.wikipedia.org/wiki/S%E1%BB%91_th%E1%BB%B1c" \o "Số thực) là [moment trung tâm](https://vi.wikipedia.org/w/index.php?title=M%C3%B4men_quanh_gi%C3%A1_tr%E1%BB%8B_trung_b%C3%ACnh&action=edit&redlink=1), nó còn là [nửa bất biến](https://vi.wikipedia.org/w/index.php?title=N%E1%BB%ADa_b%E1%BA%A5t_bi%E1%BA%BFn&action=edit&redlink=1" \o "Nửa bất biến (trang chưa được viết)) (*cumulant*) thứ hai của nó. Phương sai của một biến ngẫu nhiên là [bình phương](https://vi.wikipedia.org/wiki/B%C3%ACnh_ph%C6%B0%C6%A1ng" \o "Bình phương) của [độ lệch chuẩn](https://vi.wikipedia.org/wiki/%C4%90%E1%BB%99_l%E1%BB%87ch_chu%E1%BA%A9n" \o "Độ lệch chuẩn).



* plot q-q kiểm tra xem liệu mọi thứ có ở nơi mà chúng nên được gọi là biểu đồ lượng tử, trên trục x là các lượng tử lý thuyết, y là phân phối chuẩn. nếu dữ liệu thực sự có phần dư tuân theo phân phối chuẩn thì mọi thứ sẽ rơi dọc theo đường thẳng đỏ

hầu hết lý thuyết thống kê được xây dựng trên nền tảng của phân phối chuẩn.

Như vậy từ “normal” được dùng theo thói quen nhưng thực ra không đúng, vì vậy trong tiếng Việt ta không thể dịch là phân phối “bình thường” mà gọi là phân phối chuẩn.

Hai thông số quan trọng trong một phân phối là giá trị trung tâm hay gọi là trung bình µ và phương sai σ2 (hoặc độ lệch chuẩn σ) và thường biểu thị bằng X ~ N (µ,σ2) (N viết tắt từ normal).

Nếu phân phối chuẩn được chuẩn hóa với trung bình μ = 0 và độ lệch chuẩn σ = 1, được viết tắt là: Z ~ N (μ=0, σ=1), được gọi là phân phối chuẩn chuẩn hóa (standardized normal distribution) nghe có vẽ không được xuôi tai như tiếng Anh vì chữ normal được dịch là chuẩn rồi, do vậy dùng từ phân phối chuẩn tắc có vẽ ổn hơn !

Đồ thị này cho t thấy được các giá trị phần dư tập trung quanh đường y đỏ, cho nên giả định ei là chấp nhận được

* Residuals vs Fitted (phần dư so với trang bị) lỗi được vẽ trên trục y, trên trục x là giá trị trung bình được dự đoán cho mỗi lần quan sát, red line đường cho giá trị dự đoán nếu giả định của chúng tôi là đúng, các điểm sẽ được phân tán đều ở trên và bên dưới dòng này, không quan trọng bạn có đang ở gần hay không, bạn đang ở dưới này hay ở trên kia, bạn chỉ nên có một loại phân tán đẹp và đồng đều và để giúp bạn với điều đó, đường này nên bằng phẳng đường này và gần bằng 0.
* Scale-location (quy mô và vị trí) được gọi là biểu đồ vị trí tỷ lệ như Resi vs Fit sự khác biệt là trục x giống nhau nhưng trục y đã được thay đổi tỷ lệ theo hai cách trước tiên chúng lấy giá trị tuyệt đối của phần dư và sau đó cũng chia tỷ lệ chúng bằng sai số ước tính trong phương thức của chúng tôi và sau đó lấy căn bậc hai của số đó, tất cả điều đó kết hợp với nhau có nghĩa là nếu các giả định của chúng em nếu phần dư trung bình của chúng em sẽ giảm xung quanh giá trị 1 thì độ lệch chuẩn của nó là 1
* đồ thị cuối cùng ở đây, chúng em nhận được phần dư chuẩn hóa của chúng tôi trên trục y nhưng bây giờ không phải là giá trị tuyệt đối, vì vậy có một số trong số chúng là âm và một số trong số chúng là cố định và sau đó những gì được vẽ trên trục x bây giờ được gọi là đòn bẩy. (đòn bẩy là thước đo mức độ mà một quan sát cá nhân ảnh hưởng đến sự phù hợp của mô hình của bạn) vì vậy nếu bạn có một đòn bẩy rất thấp khi xóa đi thì nó cũng sẽ không ảnh hưởng tới model của bạn

đòn bẩy (leverage) tức là số lượng biến độc lập trong một mô hình hồi quy tuyến tính

số lượng các biến độc lập trong một chế độ hồi quy