Các sai lầm trong suy luận

Lần trươc chúng ta đã phát triển nhiều loại kiểm định giả thiết khác nhau.

Khi chúng ta bác bỏ hay không thể bác bỏ một giả thiết, chúng ta có thể phạm sai lầm. Có hai loại sai lầm:

Sai lầm loại 1: chúng ta bác bỏ giả thiết "không" (null = còn gọi là giả thiết ban đầu) khi thực ra nó đúng.

Sai lầm loại 2 : chúng ta không thể bác bỏ giả thiết "không" khi thực sự nó sai.

Rất cần thiết khi chúng ta dành một ít thời gian cho những khái niệm này vì chúng sẽ hương dẫn chúng ta khi lựa chọn các giá trị cho mức ý nghĩa (α) đối vơi những kiểm định nào đó trong môn học kinh tế lượng.

Sai lầm loại 1

Xét giả thiết sau :

$$H_0 : \mu \le \mu_0$$
$$H_1 : \mu > \mu_0$$

Điều kiện khách quan : giả thiết "không" là đúng .

Loại sai lầm có thể : Loại 1

P(Sai lầm loại 1) = $P\left(\overline{X}_{(1-\alpha)} < \overline{X} \mid \mu = \mu_0\right)$

Xác suất này là xác suất bác bỏ giả thiết "không" khi thực ra là nó đúng.

Nên nhơ là khi giá trị trung bình tổng thể bằng vơi số tiêu chuẩn được xác định trong giả thiết "không", thì xác suất của sai lầm loại 1 là α. Khi thảo luận sai lầm loại 1, chúng ta gọi xác suất phạm sai lầm này là rủi ro α .

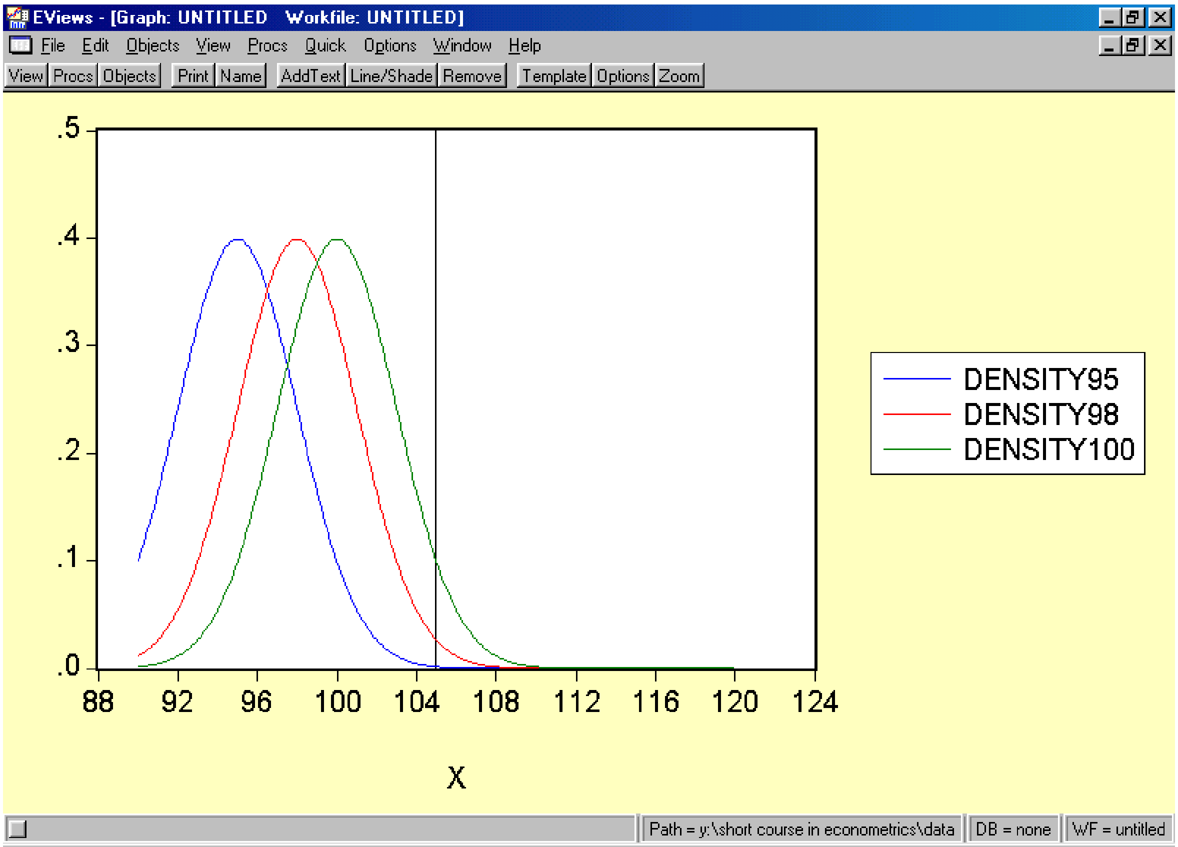
Bằng cách đánh giá xác suất của sai lầm loại 1 cho nhiều giá trị có khả năng là thực của trung bình tổng thể phù hợp vơi giả thiết "không", chúng ta có thể xây dựng đường sai số là một hàm của giá trị trung bình thực.

Trong đồ thị sau chúng ta chỉ ra nhiều mật độ có thể so sánh vơi giả thiết sau:

H_0 : 100 $\leq$ μ
H_1 : 100 $>$ μ

Giá trị tới hạ n đối với kiểm định của giả thiết nà y có mức ý nghĩa 5% đượ c thể hiện bởi đườ ng thẳng tạ i

$$\overline{X}_{(0.95)} = 103{,}29$$



Các xác suất của Sai lầm loạ i 1 đượ c mô tả bằng các diện tích vù ng nằm dưới nhữ ng đườ ng cong mật độ cho tới phiá bên phải của đườ ng tới hạ n. Trong Bà i tập 7, Anh/Chị sẽ xây dự ng nhữ ng đồ thị của các xác suất phạ msai lầm.

Sai lầm loạ i 2

Một lần nữ a lạ i xét giả thiết y như vậy:

H_0 :μ $\leq$ μ_0
H_1 :μ $>$ μ_0

Điều kiện khách quan : giả thiết "không" là sai .

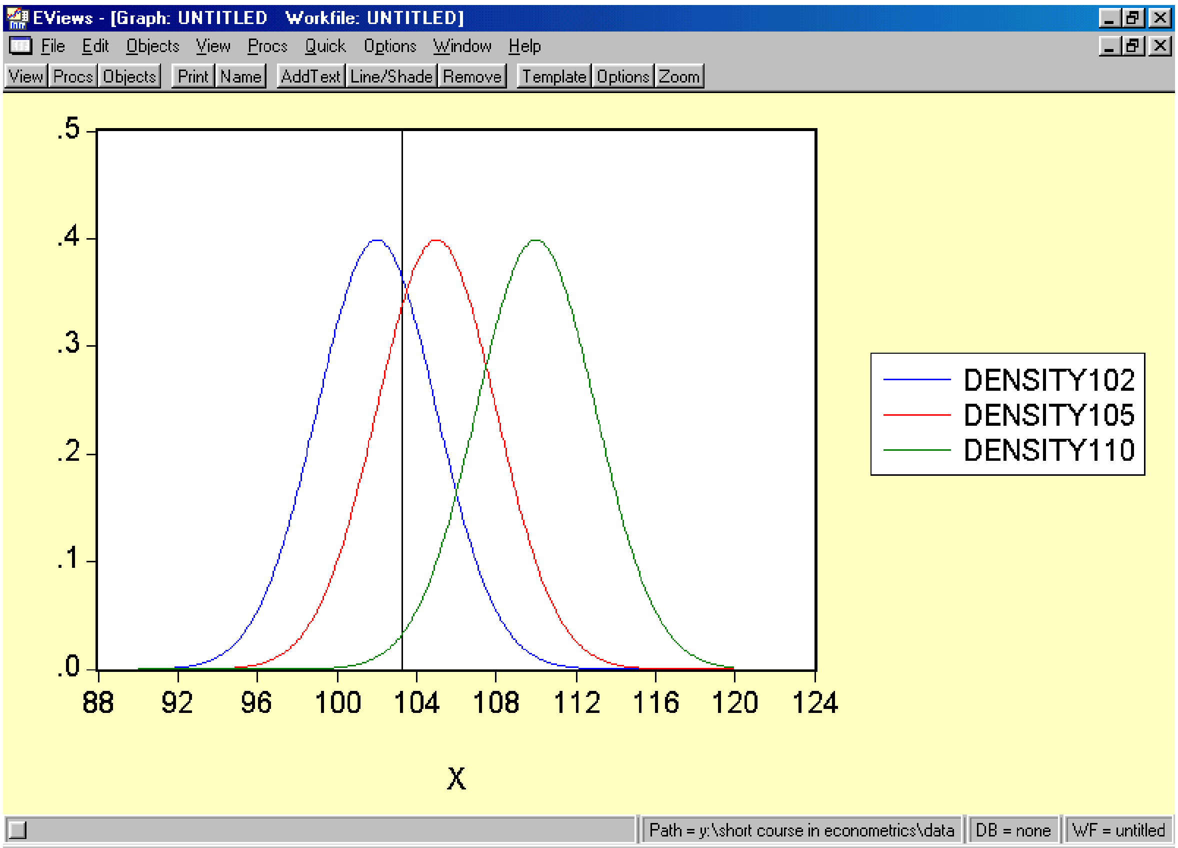
Loạ i sai lầm có thể : Loạ i 2

P(Sai lầm loạ i 2) = $P\left(\overline{X} < \overline{X}_{(1-\alpha)} \mid \mu = \mu_1\right)$ trong đó $\mu_0 < \mu_1$

Xác suất nà y là xác suất không thể bác bỏ đượ c giả thiết "không" khi thự c ra nó là sai. Xác suất nà y đượ c gọ i là rủi ro β .

Bằng cách đánh giá xác suất của sai lầm loạ i 2 đối vớ i nhiều giá trị có khả năng là có thự c của trung bình tổng thể phù hợ p vớ i giả thuyết "không", chúng ta có thể xây dự ng đượ c đườ ng sai lầm (error curve), và đườ ng nà y là một hà m của giá trị trung bình thự c.

Đồ thị sau đây chỉ ra các mật độ đối vớ i các giá trị trung bình 102, 105 và 110. Các diện tích nằm dưới nhữ ng đườ ng cong mật độ về phiá trái của giá trị tới hạ n $\overline{X}_{(0,95)}$ = 103,29 là nhữ ng xác suất sai lầm loạ i 2.



Đặc trưng hoạ t động và sức mạ nh (Operating Characteristic and Power)

Nhữ ng kết quả mong muốn dĩ nhiên là nhữ ng kết quả không sai lầm.

Khi giả thiết "không" là đúng và chúng ta không thể bác bỏ đượ c nó thì xác suất nà y là $(1 - \alpha)$, là đạ i lượ ng trong trườ ng nà y đượ c gọ i là đặc trưng hoạ t động (operating characteristic) của kiểm định nà y .

Khi giả thuyết "không" là sai , và chúng ta bác bỏ nó, thì xác suất nà y là $(1 - \beta)$, đượ c gọ i là sức mạ nh của kiểm định.

Đánh đổi giữ a các rủi ro

Giả sử Anh/Chị muốn giảm rủi ro α đối vớ i việc kiểm định giả thiết "không" đã cho trên đây. Anh/Chị có thể là m việc nà y đơn giản bằng cách chọ n mức ý nghĩa α nhỏ hơn. Tuy nhiên, hã y lưu ý rằng điều nà y là m di chuyển giá trị tớ i hạ n sang phải. Hậu quả của việc nà y là nó là m tăng rủi ro β và giảm sức mạ nh kiểm định của Anh/Chị.

Có cách nà o giảm cả hai rủi ro cù ng một lúc không? Câu trả lờ i là có: Anh/chị có thể giảm cả hai xác suất phạ m sai lầm bằng cách bảo đảm các ướ c lượ ng chính xác hơn, anh / chị là m điều nà y bằng cách tăng qui mô mẫu.

Qui mô mẫu nà o đảm bảo mụ c tiêu của chúng ta nêu trên ?

$$P\left(\overline{X}_{(1-\alpha)} < \overline{X} \mid \mu = \mu_0\right) = \alpha$$

Dạ ng chuẩn hoá của biểu thức nà y là : $P\left(\dfrac{\overline{X}_{(1-\alpha)} - \mu_0}{\sigma/\sqrt{n}} < Z \mid \mu = \mu_0\right) = \alpha$

Giá trị tớ i hạ n là : $\overline{X}_{(1-\alpha)} = \mu_0 + Z_{(1-\alpha)}\dfrac{\sigma}{\sqrt{n}}$

Đồng thờ i chúng ta muốn rằng

$$P\left(\overline{X} < \overline{X}_{(\beta)} \mid \mu = \mu_1\right) = \beta$$

Dạ ng chuẩn hoá của biểu thức nà y là : $P\left(Z < \dfrac{\overline{X}_{(\beta)} - \mu_0}{\sigma/\sqrt{n}} \mid \mu = \mu_1\right) = \beta$

Giá trị tớ i hạ n là : $\overline{X}_{(\beta)} = \mu_1 + Z_{(\beta)}\dfrac{\sigma}{\sqrt{n}}$

Hai giá trị tới hạn này phải giống nhau, nên chúng ta cân bằng hai biểu thức về giá trị tới hạn và điều này cho ta một biểu thức từ đó chúng ta có thể giải để tìm n là qui mô mẫu cần thiết.

$$\mu_1 + Z_{(\beta)}\frac{\sigma}{\sqrt{n}} = \mu_0 + Z_{(1-\alpha)}\frac{\sigma}{\sqrt{n}}$$

$$n = \sigma^2\left[\frac{Z_{(1-\alpha)} - Z_{(\beta)}}{\mu_1 - \mu_0}\right]^2$$