Alternative Summary Indices: PLC and ASC for the Summary Receiver Operating Characteristic (SROC) Curve

Xuan Zhang

Supervisor: Prof. Stephen Walter

Outline

- Introduction
  - Historical background
  - Lee's findings in the Receiver Operating Characteristic (ROC) curve
  - Walter's findings in the Summary Receiver Operating Characteristic (SROC) curve
- Derivation of equations of the Projected Length of the Curve (PLC), the Area Swept out by the Curve (ASC), and their variances
- The behavior of PLC and its variance
- The behavior of ASC and its variance
- Conclusion

Historical Background

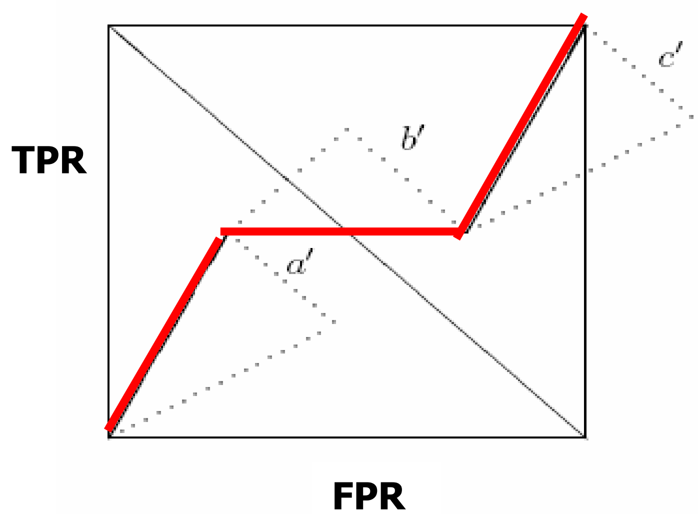
- ROC represents the performance of a diagnostic test.
- Two alternative indices of ROC: PLC and ASC were proposed by Lee (1996)
- SROC curve was proposed to describe a diagnostic test based on data from a meta-analysis.
- The basic properties of the SROC curve were discussed by Walter (2002)
- Little is known about the basic properties of the PLC and ASC in SROC curve
- Based on Lee's and Walter's findings, my project focuses on studying PLC and ASC in SROC curve

Lee’s findings in ROC curve

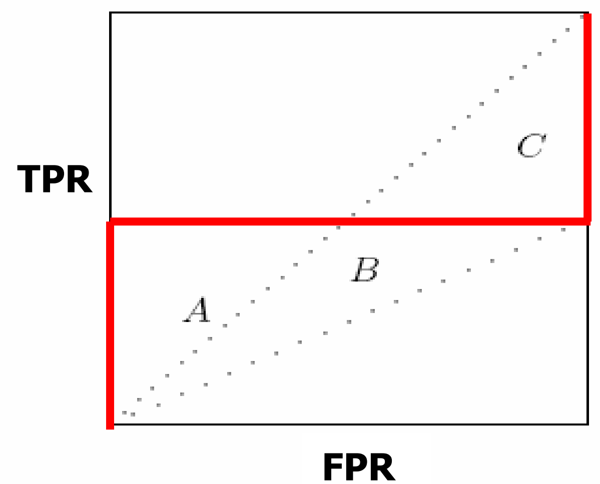
- What’s the ROC curve?
  - ROC is the plot of TPR (y) against FPR (x) in a single study
  - TPR: (true positive rate)
    Pr (test result = positive | disease = present)
  - FPR: (false positive rate)
    Pr (test result = positive | disease = absent)

Lee’s findings in ROC curve

- Two alternative indices of ROC: PLC and ASC
  - PLC: sum of all of the projected lengths of the ROC curve onto the negative diagonal line (a’+b’+c’)
  - ASC: the total area swept out by the ROC curve (A+2B+C)



ROC curves to illustrate PLC



ROC curves to illustrate ASC

Walter’s findings in SROC curve

- What’s the SROC curve?
  - Meta-analysis is the systematic and quantitative review of a set of individual studies all concerning the related question, intended to integrate their findings
  - SROC is proposed as a means of summarizing a test’s TPR and FPR from multiple studies based on data from a meta-analysis.

Walter's findings in SROC curve

- Moses' theory to propose a $SROC$ curve.

$$D = \ln\left(\frac{TPR}{1-TPR}\right) - \ln\left(\frac{FPR}{1-FPR}\right)$$

$$S = \ln\left(\frac{TPR}{1-TPR}\right) + \ln\left(\frac{FPR}{1-FPR}\right)$$

$$D = a + bS$$

Walter's findings in SROC curve

- the relationship between TPR and FPR

$$TPR = \frac{\exp(\frac{a}{1-b})(\frac{FPR}{1-FPR})^{\frac{1+b}{1-b}}}{1+\exp(\frac{a}{1-b})(\frac{FPR}{1-FPR})^{\frac{1+b}{1-b}}}$$

Walter's findings in SROC curve

- What's the 'a'?
  - Odds Ratio (OR)

$$OR = \frac{TPR/(1-TPR)}{FPR/(1-FPR)}$$

  - a = ln(OR) convey the test's accuracy in discriminating cases from non-cases.
  - My project focuses on $|a| \leq 3$, which includes most cases in practice.

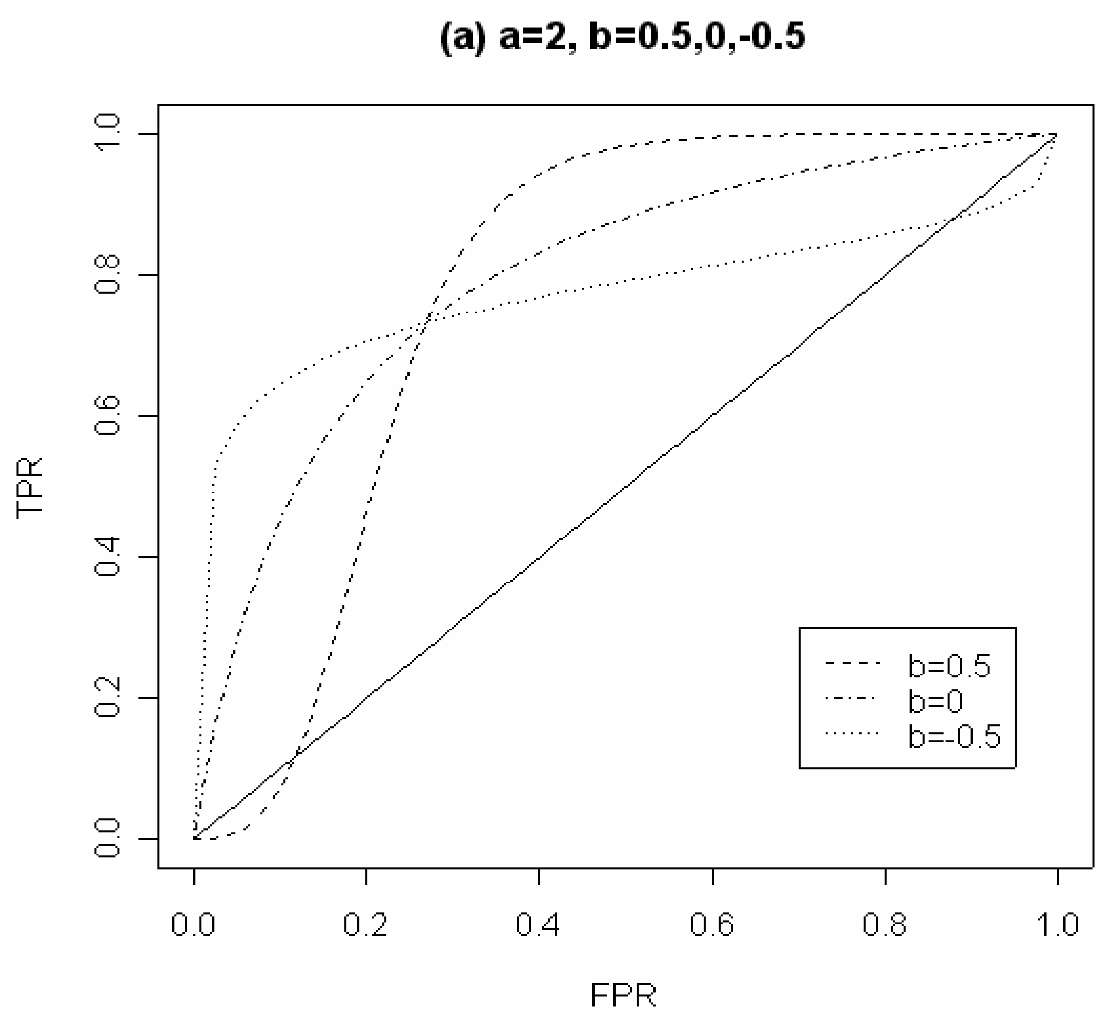
Walter's findings in SROC curve

- What's the 'b' ?

  - b tests the dependence of the test accuracy on threshold

  - My project focuses on $|b| \le 0.3$, which includes most cases in practice.

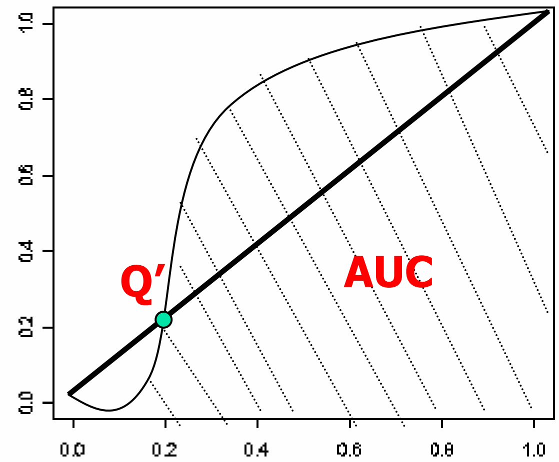
Walter's findings in SROC curve



Walter's findings in SROC curve

- AUC index: area under the curve

$$AUC = \int_0^1 \frac{\exp(\frac{a}{1-b})(\frac{x}{1-x})^{\frac{1+b}{1-b}}}{1+\exp(\frac{a}{1-b})(\frac{x}{1-x})^{\frac{1+b}{1-b}}}dx$$



- Q' index: point where the curve crosses the diagonal

$$TPR = FPR = \frac{\exp(-a/2b)}{1+\exp(-a/2b)}$$

My Research Results

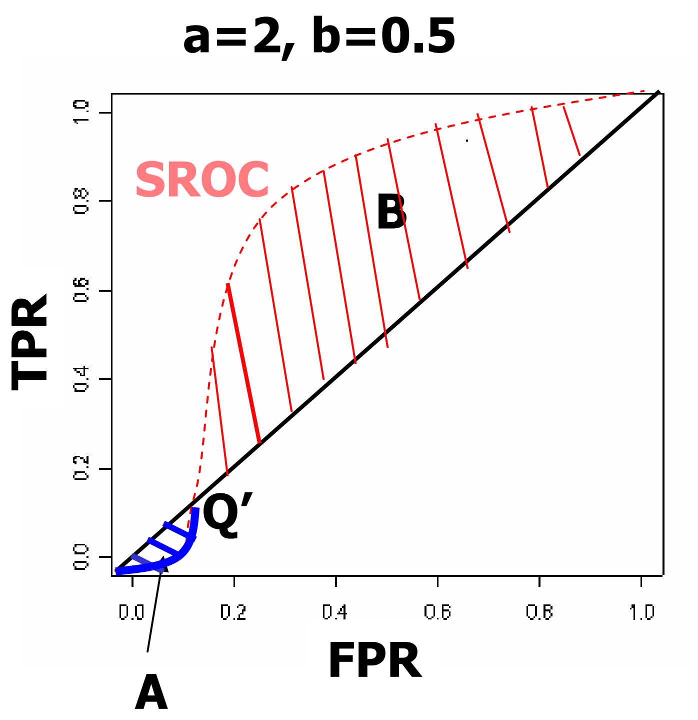
Derivation of Equations of PLC and its Variance

- PLC

- Point Q' separates the whole curve into Region A **(below)** and Region B **(above)**

- $\hat{d}_1$ and $\hat{d}_2$ the maximum distance in Region A and B respectively

- $PLC = 2(\hat{d}_1 + \hat{d}_2)$



Derivation of Equations of PLC and its Variance

- variance of PLC
  - Delta method
  - $\mathrm{var}(PLC) = \left(\frac{\partial PLC}{\partial a}\right)^2 \mathrm{var}(\hat{a}) + \left(\frac{\partial PLC}{\partial b}\right)^2 \mathrm{var}(\hat{b})$

    $+2\left(\frac{\partial PLC}{\partial a}\right)\left(\frac{\partial PLC}{\partial b}\right)\mathrm{cov}(\hat{a},\hat{b})$
  - $\frac{\partial PLC}{\partial a} = 2\frac{\partial \hat{d}_1}{\partial a} + 2\frac{\partial \hat{d}_2}{\partial a}$
  - $\frac{\partial PLC}{\partial b} = 2\frac{\partial \hat{d}_1}{\partial b} + 2\frac{\partial \hat{d}_2}{\partial b}$

Derivation of Equations of ASC and its Variance

- ASC

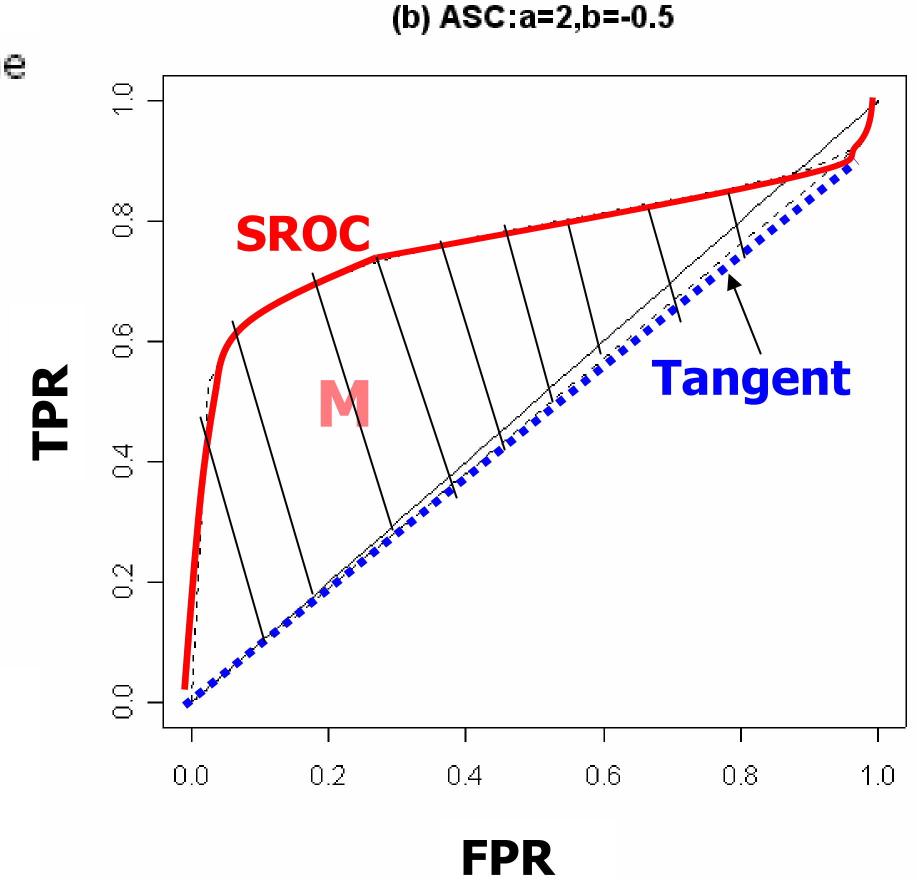
- M: the area between the tangent line and the curve

- when $a > 0,\ b > 0$ and $a < 0,\ b > 0$,

$$ASC = AUC + 2M - \frac{1}{2}$$

- when $a > 0,\ b < 0$ and $a < 0,\ b < 0$,

$$ASC = 2M + \frac{1}{2} - AUC$$



Derivation of Equations of ASC and its Variance

- variance of ASC

- Delta method

$$- \ \text{var}(A\hat{S}C) = \left(\frac{\partial ASC}{\partial a}\right)^2 \text{var}(\hat{a}) + \left(\frac{\partial ASC}{\partial b}\right)^2 \text{var}(\hat{b})$$

$$+2\left(\frac{\partial ASC}{\partial a}\right)\left(\frac{\partial ASC}{\partial b}\right)\text{cov}(\hat{a}, \hat{b})$$

Derivation of Equations of ASC and its Variance

- when $a > 0$, $b > 0$ and $a < 0$, $b > 0$,

$$\frac{\partial ASC}{\partial a} = \frac{\partial AUC}{\partial a} + 2\frac{\partial M}{\partial a}$$

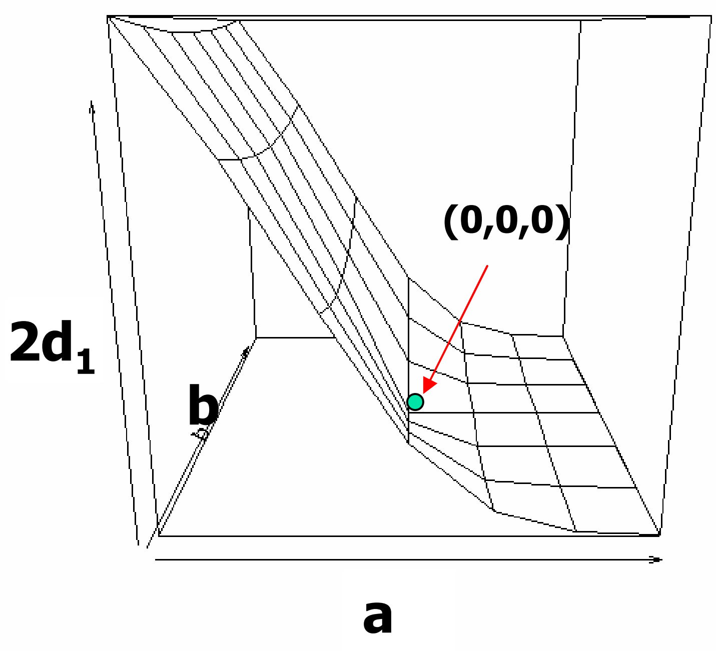
$$\frac{\partial ASC}{\partial b} = \frac{\partial AUC}{\partial b} + 2\frac{\partial M}{\partial b}$$

- when $a > 0$, $b < 0$ and $a < 0$, $b < 0$,

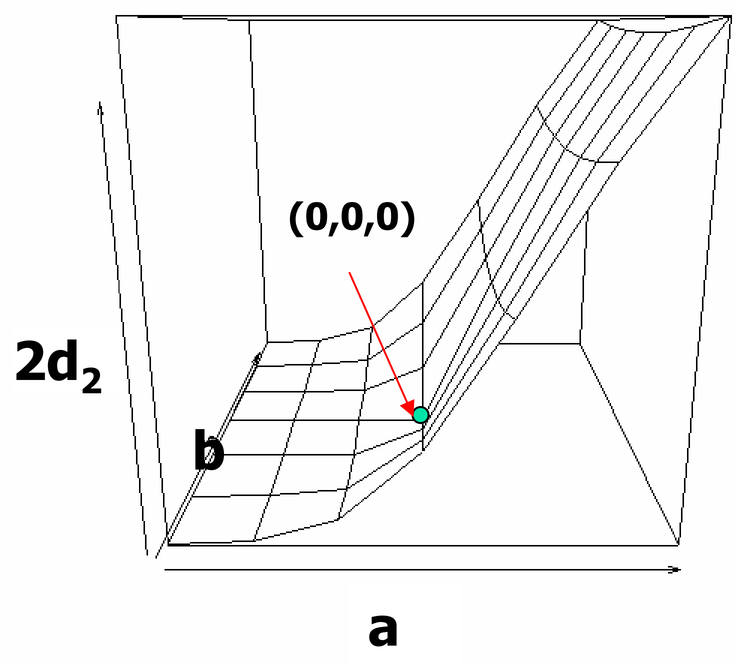
$$\frac{\partial ASC}{\partial a} = 2\frac{\partial M}{\partial a} - \frac{\partial AUC}{\partial a}$$

$$\frac{\partial ASC}{\partial b} = 2\frac{\partial M}{\partial b} - \frac{\partial AUC}{\partial b}$$

The Behavior of PLC



maximum values of $2d_1$



Maximum values of $2d_2$

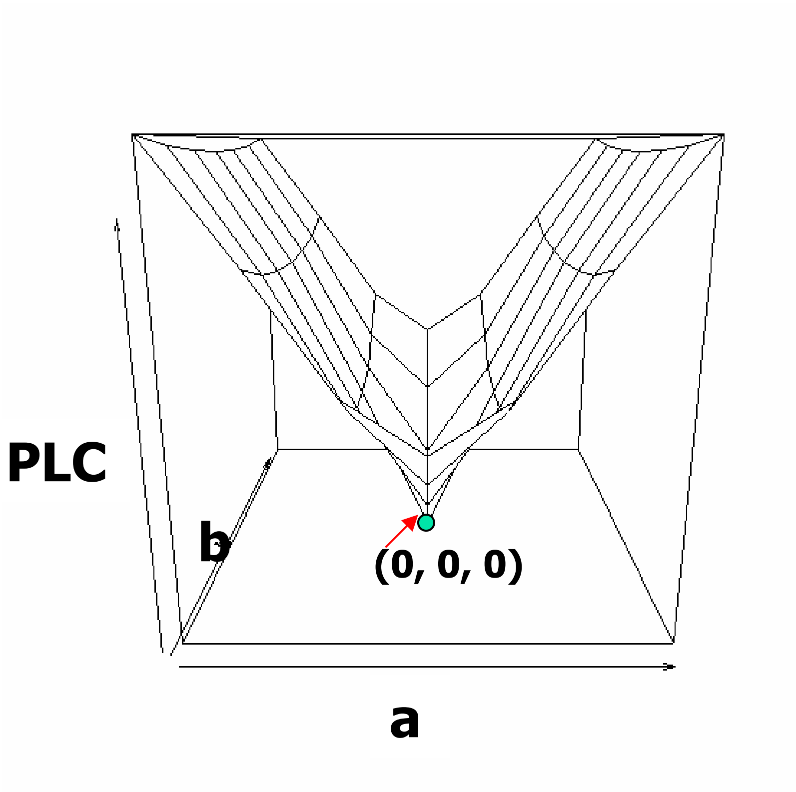
The Behavior of PLC

- $\tilde{d}_1$ and $\tilde{d}_2$ are only symmetric for the fixed value of a

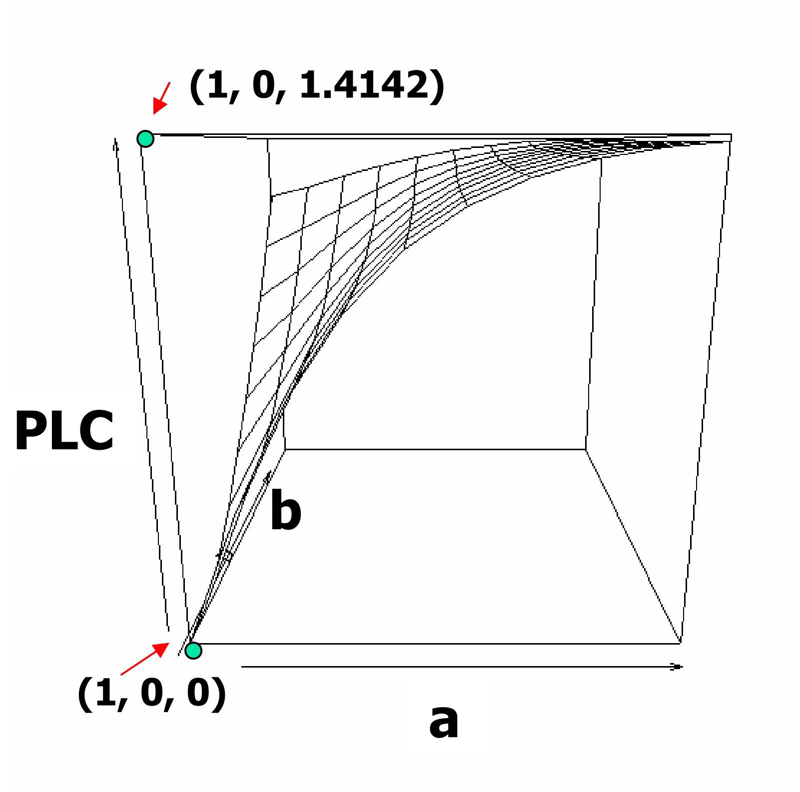
- $\tilde{d}_1$ and $\tilde{d}_2$ perform same for fixed a; opposite for fixed b

- $\tilde{d}_1, \tilde{d}_2$ attain the minimum value in the homogeneous case

The Behavior of PLC



Values of PLC

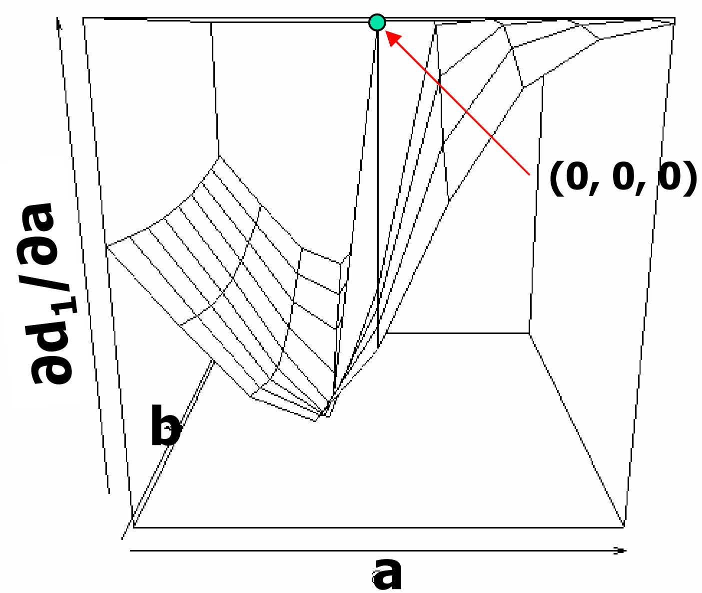


Values of PLC as $1 \leq a \leq 10$

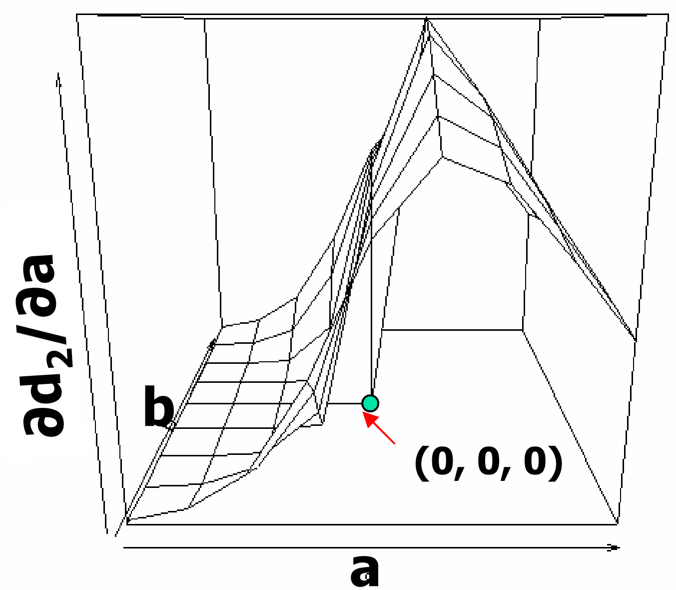
The Behavior of PLC

- PLC index is symmetric for the fixed values of |a| and |b|
- PLC (homo) provide a good approximation of PLC (hetero) for big values of |a|
- As a → ∞, PLC → 1.4142 the maximum
- Zero value PLC corresponds to the SROC curve running along the diagonal line

The behavior of the variance of PLC

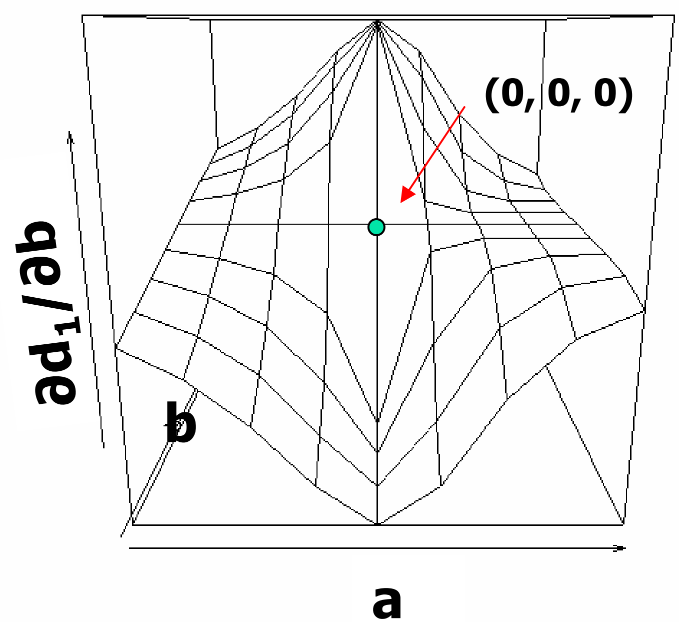


Values of $\partial d_1/\partial a$

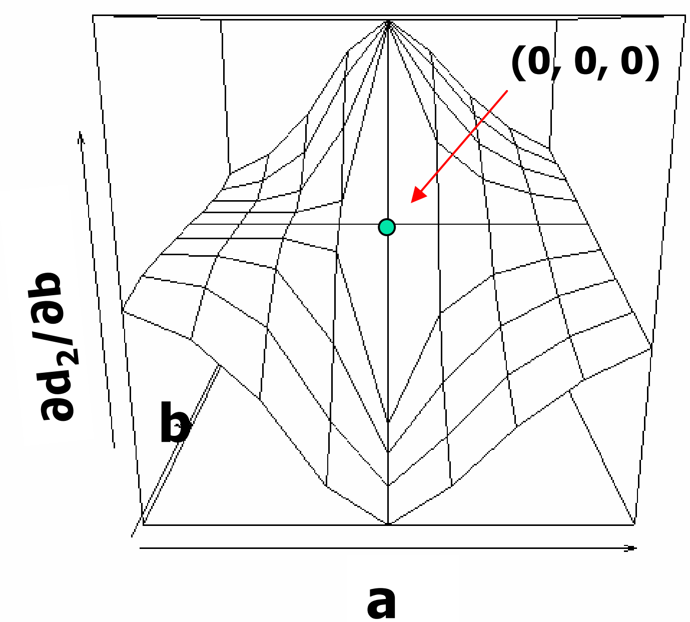


Values of $\partial d_2/\partial a$

The behavior of the variance of PLC



Values of $\partial d_1/\partial b$



Values of $\partial d_2/\partial b$

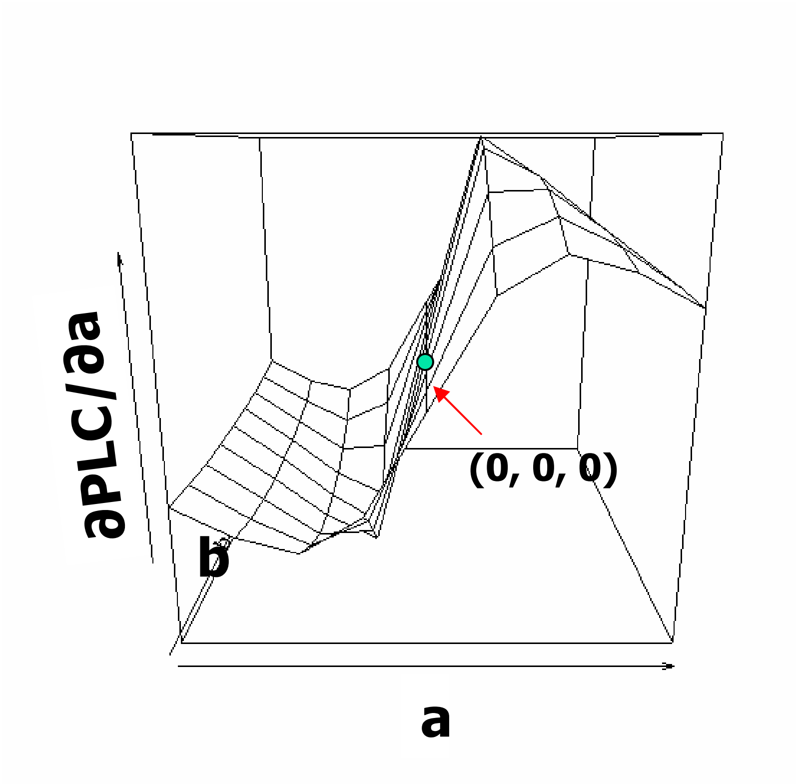
The behavior of the variance of PLC

- as $-3 \leq a \leq 3$, $\hat{d}_1$ decreasing function of a for fixed b; $\hat{d}_2$ on the opposite

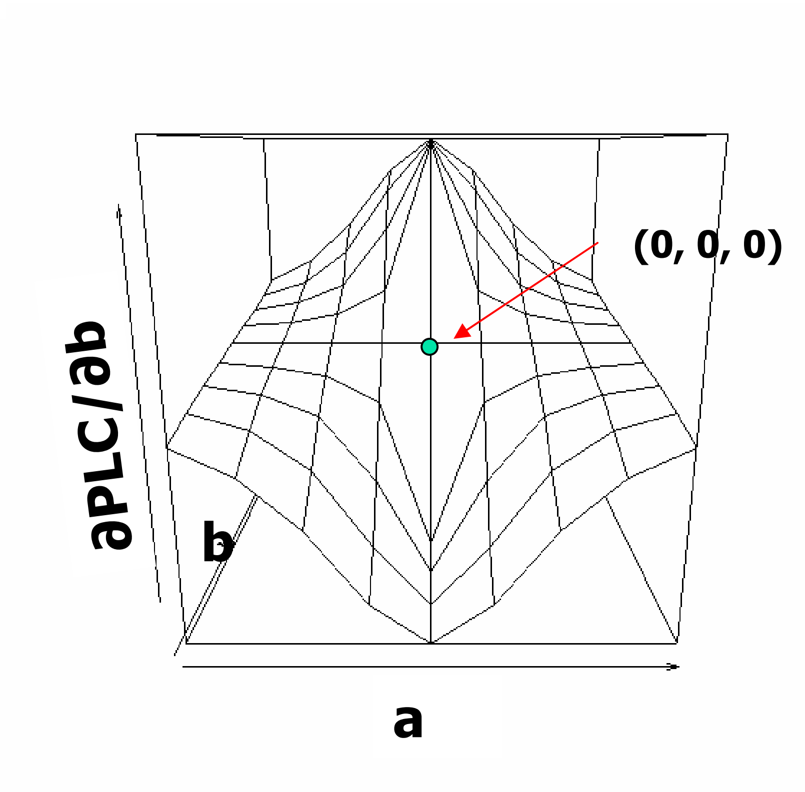
- both $\hat{d}_1$ and $\hat{d}_2$ changes progressively steeper for the smaller values of a

- $\partial d_1/\partial b$ and $\partial d_2/\partial b$ change sign from negative to positive at $b = 0 \Rightarrow$ both attain the minimum in homogeneous case

The behavior of the variance of PLC



Values of ∂PLC/∂a



Values of ∂PLC/∂b

The behavior of the variance of PLC

- ∂PLC/∂a changes sign from negative to positive at a = 0, PLC attains its minimum value at a = 0

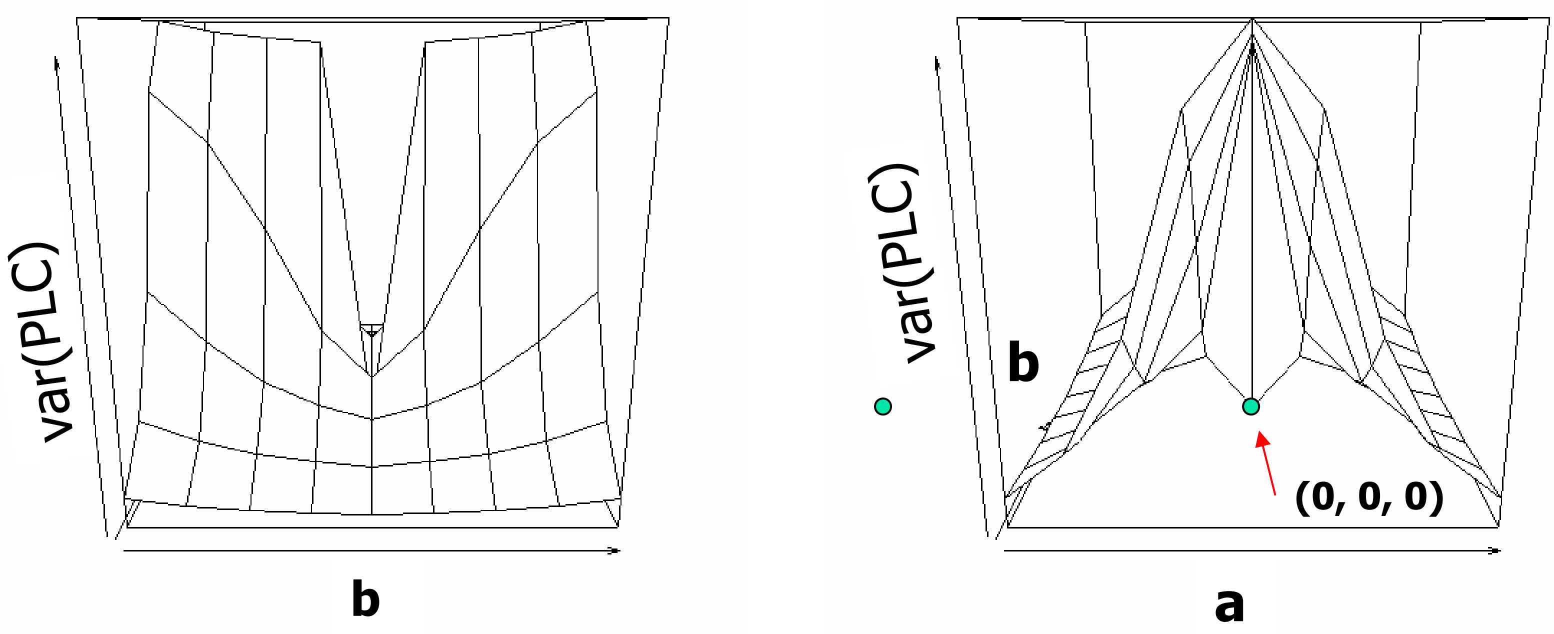
- ∂PLC/∂b changes sign from negative to positive at b = 0, PLC attains its minimum value at b = 0, the homogeneous case

The behavior of the variance of PLC

- To explore the effects of var(PLC) data for the lymphangiography test for cervical cancer metastatatases are used for example

- The values of var(a), var(b) and cov(a,b) are calculated by using the standard regression software

The behavior of the variance of PLC



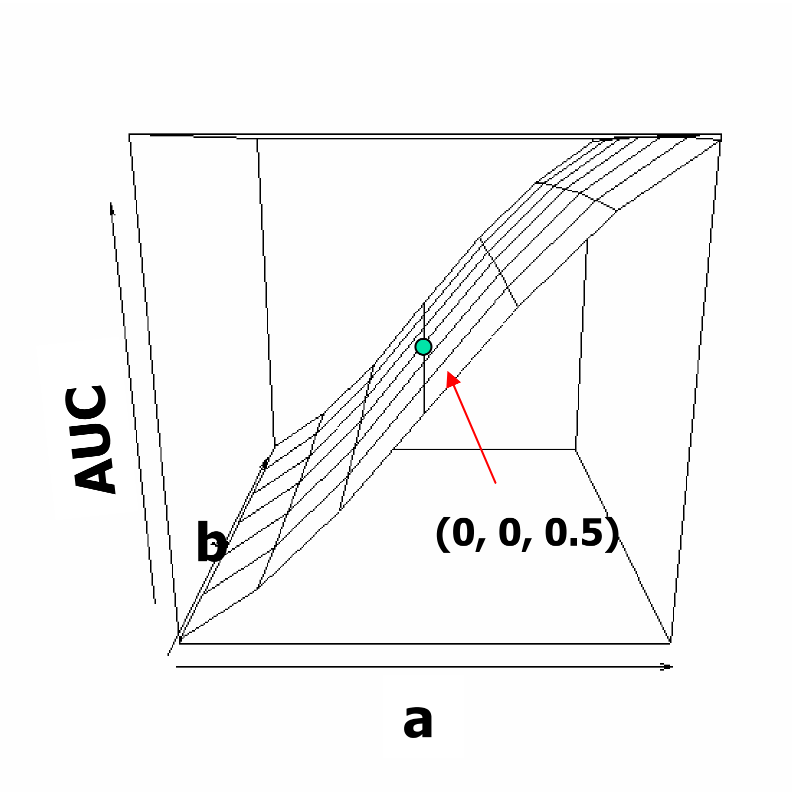
Values of var(PLC)

The behavior of the variance of PLC

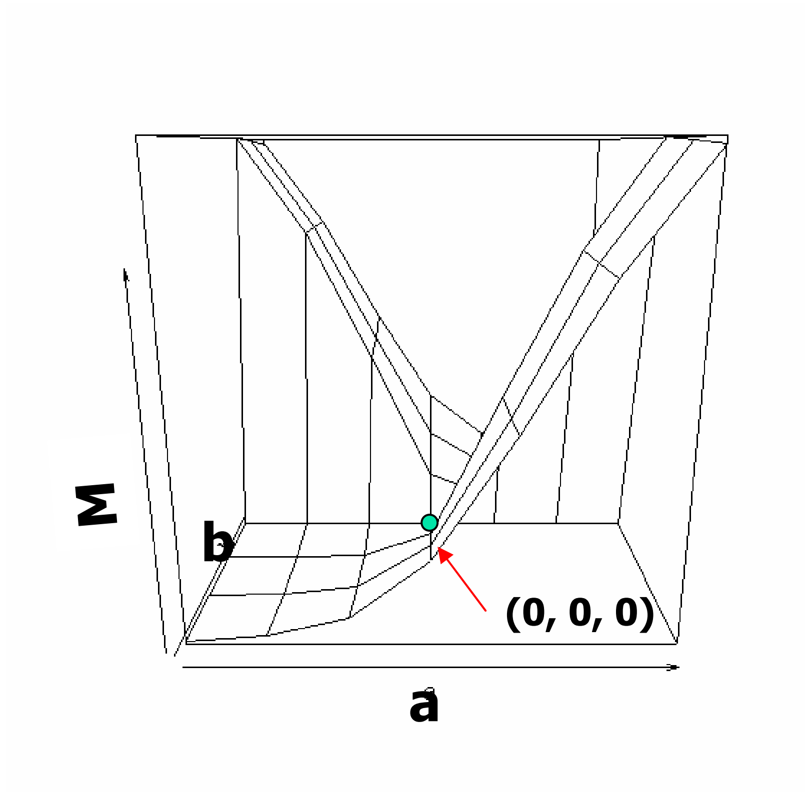
- The example illustrates that the heterogeneous variances are larger than the homogeneous estimates

- The big values of var(PLC) indicate the worst situation for estimating PLC index in the diagnostic test

The behavior of ASC



Values of AUC

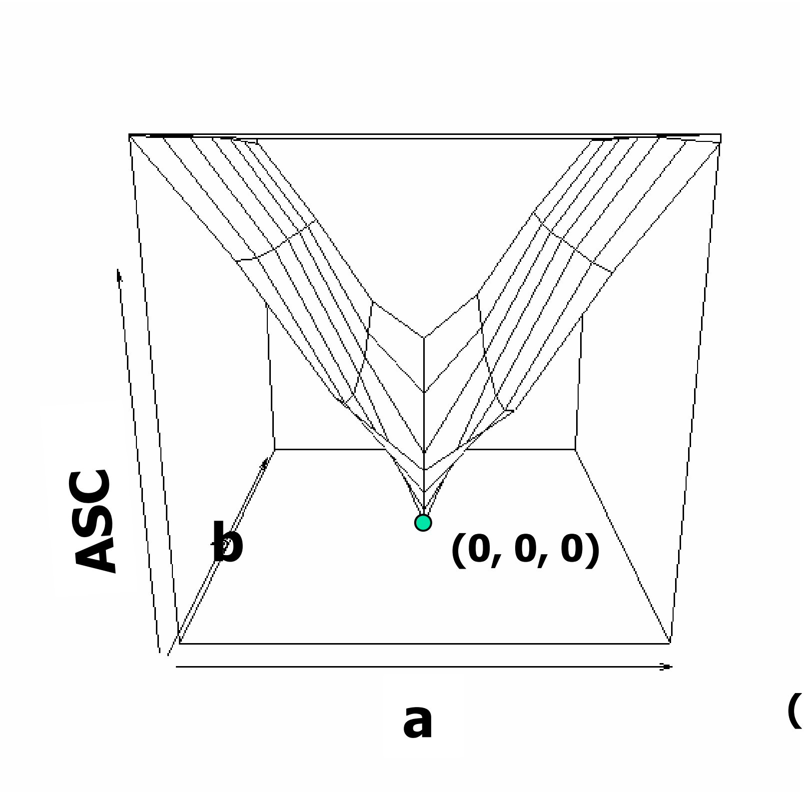


Values of M

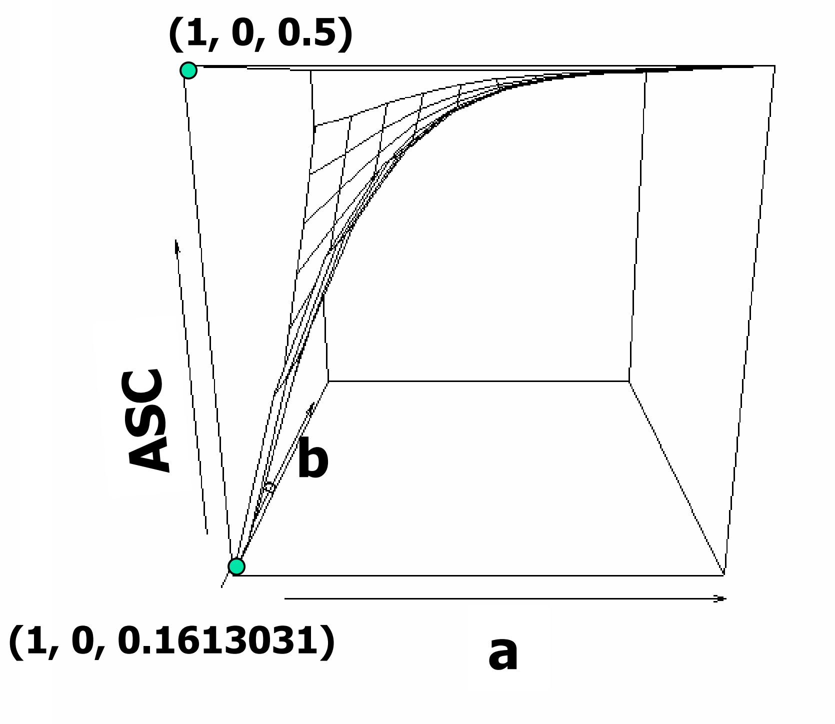
The behavior of ASC

- AUC and M index are all symmetric
- When b = 0, M → 0
- For fixed a M do not change smoothly
- AUC (homo) provide a good approaximation of ASC (hetero)

The behavior of ASC



Values of ASC

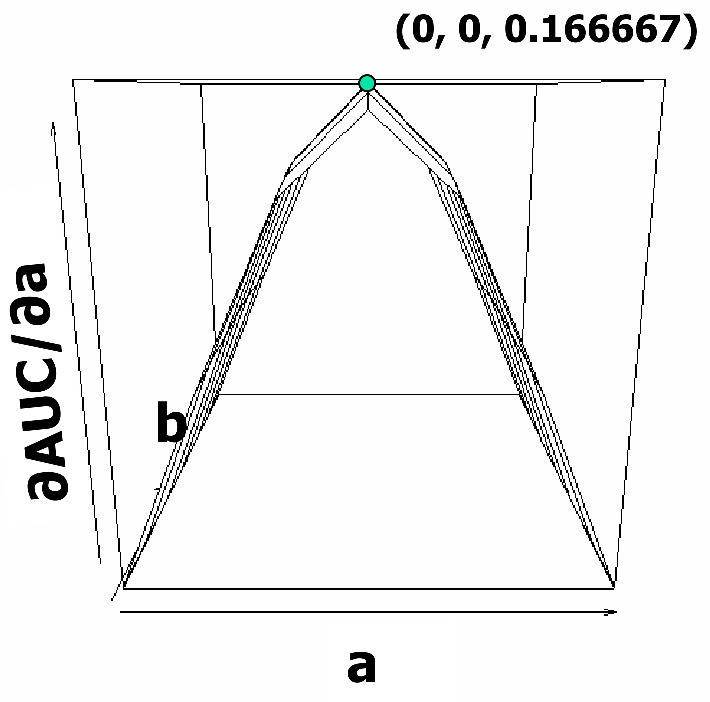


Values of ASC as 1≤a≤10

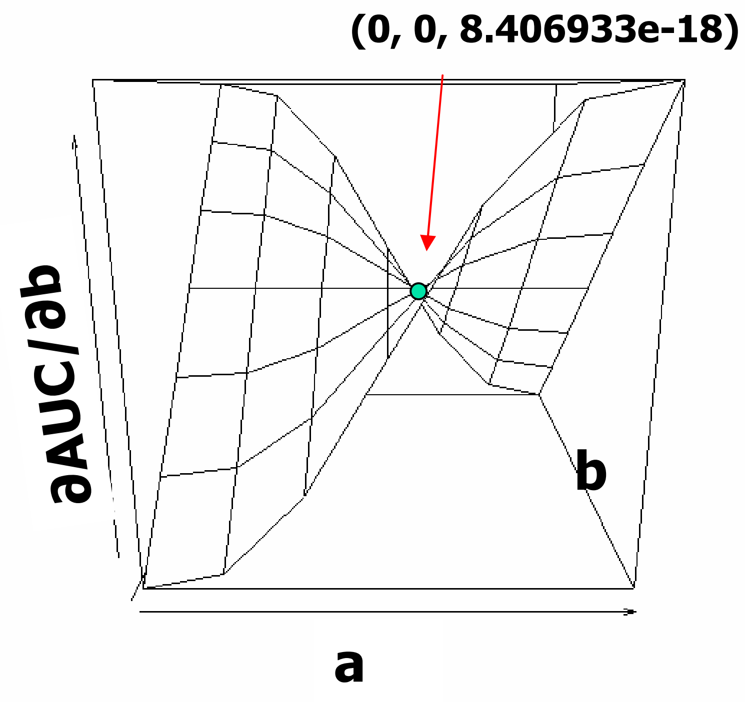
The behavior of ASC

- ASC as a and b are positive symmetric to that as a and b are negative
- As a → ∞, ASC → 0.5, the maximum
- Zero value ASC corresponds to the SROC curve running along the diagonal line

The behavior of the variance of ASC

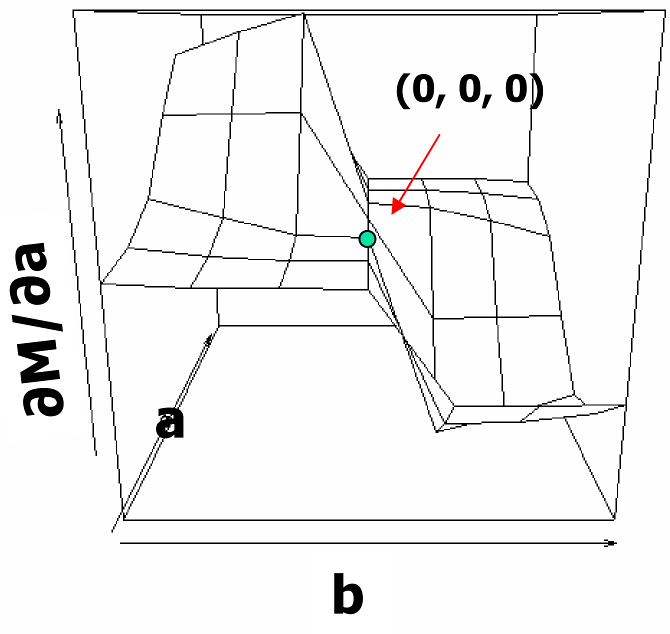


Values of ∂AUC/∂a

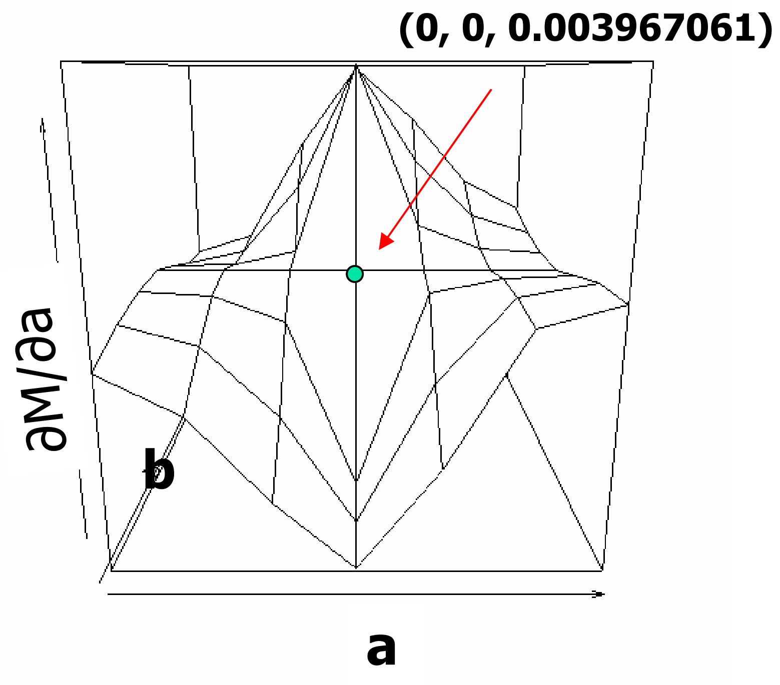


Values of ∂AUC/∂b

The behavior of the variance of ASC



Values of **∂M/∂a**

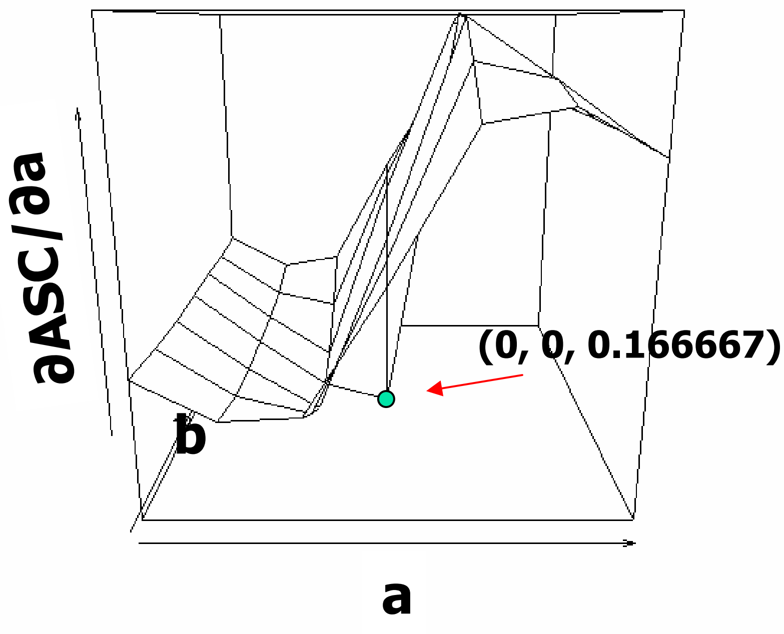


Values of **∂M/∂b**

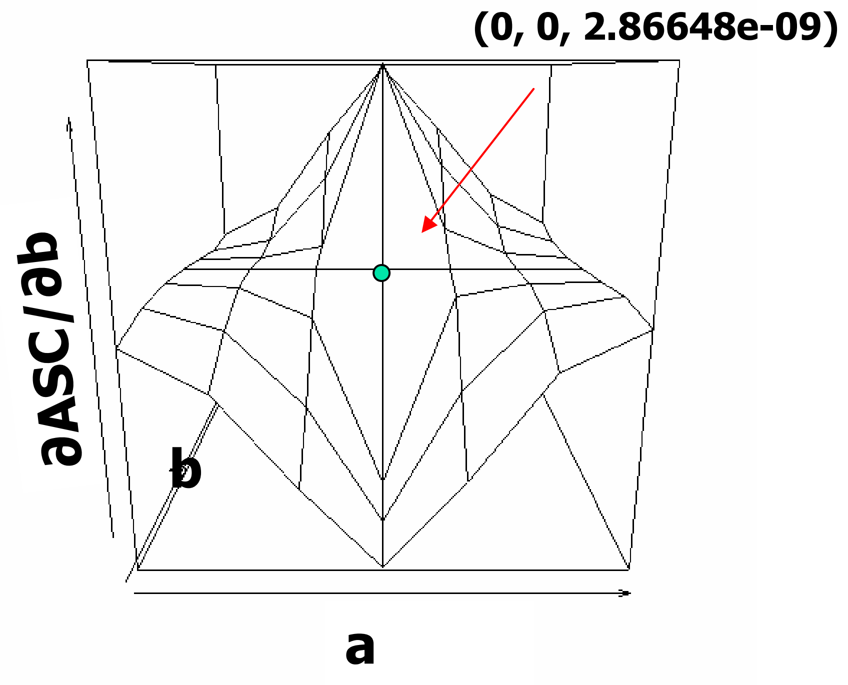
The behavior of the variance of ASC

- $\partial AUC/\partial a$ is an increasing function of a
- $\partial AUC/\partial b$ is a decreasing function of b in the interval of $a<0$, $b<0$ or $a>0$, $b>0$
- $\partial AUC/\partial b$ minimized at $b = 0$ ($a<0$) and maximized at $b = 0$ ($a>0$)
- $\partial M/\partial a$ is an increasing function of a ($b<0$); a decreasing function of a ($b>0$)
- $\partial M/\partial b$ is an increasing function of b ($a>0$); a decreasing function of b ($a<0$)

The behavior of the variance of ASC

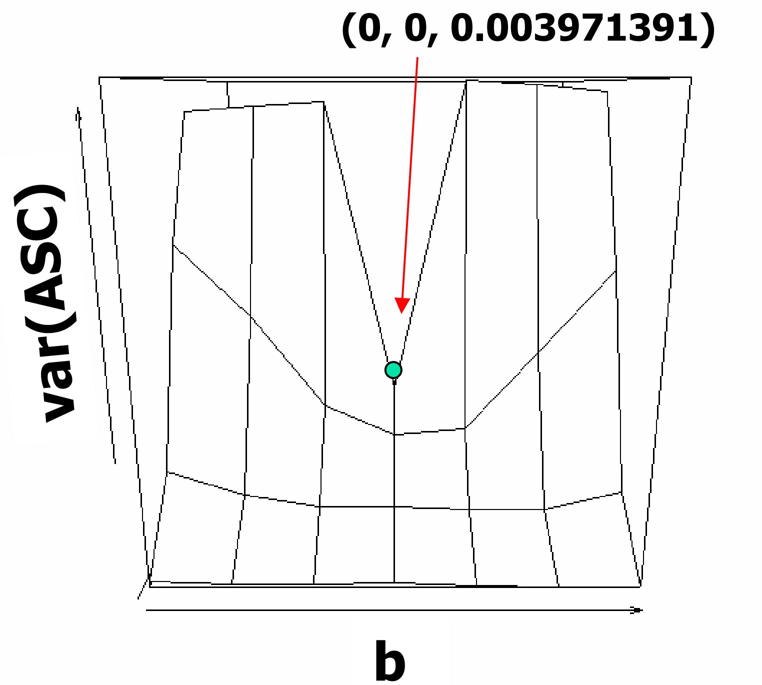


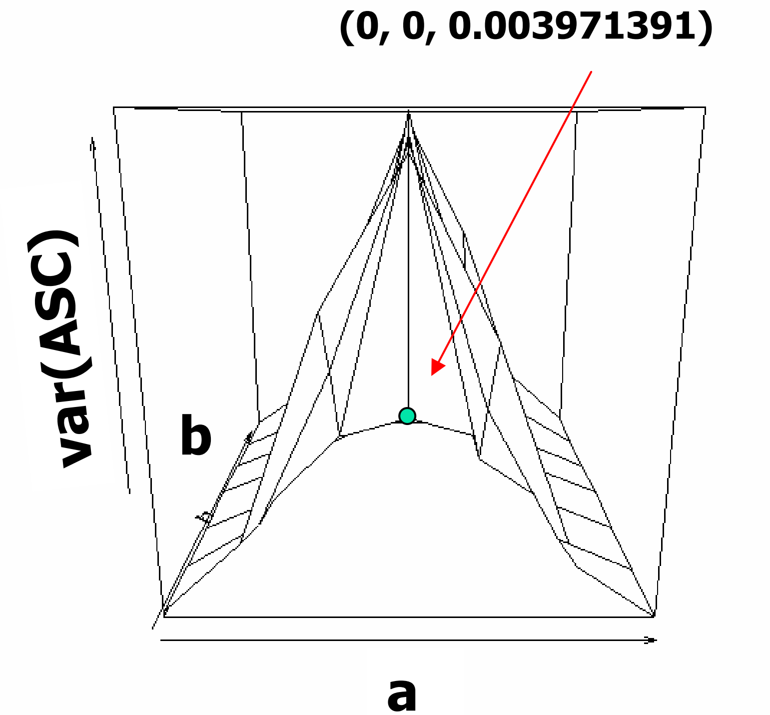
Values of ∂ASC/∂a



Values of ∂ASC/∂b

The behavior of the variance of ASC





Values of var(ASC)

The behavior of the variance of ASC

- Data for the lymphangiography test for cervical cancer metastatases are used for example again
- The example illustrates that the heterogeneous variances are larger than the homogeneous estimates
- The big values of var(ASC) indicate the worst situation for estimating ASC index in the diagnostic test

Conclusion

- The expressions of PLC index and ASC index and their basic properties have been established in this project
- All the results are predicted on the validity of the regression model proposed by Moses
- Numerical integration are used to deal with all the deduced expressions of PLC and ASC when $-3 \leq a \leq 3$ and $-0.3 \leq b \leq 0.3$
- As $a \rightarrow \infty$, PLC $\rightarrow$ 1.4142; ASC $\rightarrow$ 0.5
- Both the heterogeneous variance of PLC and ASC are larger than the homogeneous estimates
- The big values of var(PLC) and var(ASC) indicate the worst situation for estimating PLC and ASC index in the diagnostic test