

## Wie finde ich den optimalen Cutpoint in SAS?

Andrea Gonnermann  
 Med. Hochschule Hannover -  
 Institut für Biometrie  
 Carl-Neuberg-Str. 1  
 30625 Hannover  
 Gonnermann.Andrea@mh-  
 hannover.de

Eva Budde  
 Med. Hochschule Hannover -  
 Institut für Biometrie  
 Carl-Neuberg-Str. 1  
 30625 Hannover  
 Budde.Eva@mh-hannover.de

Paraskevi Fytili  
 Med. Hochschule Hannover -  
 Klinik für Gastroenterologie, Hepatologie und Endokrinologie  
 Carl-Neuberg-Str. 1  
 30625 Hannover  
 Fytili.Paraskevi@mh-hannover.de

### Zusammenfassung

In Diagnosestudien mit stetigen Parametern werden für die Bewertung häufig Receiver Operator Characteristics (ROC) Analysen durchgeführt. Das Ziel solcher Untersuchungen ist die Berechnung der Area Under the Curve (AUC) und das Auffinden eines optimalen Schwellenwerts (Cutpoint). Der Cutpoint dient dazu, mit Hilfe des stetigen Merkmals einen Test zur Einteilung in z.B. gesund/krank zu ermöglichen. Für die Bestimmung des Cutpoints können verschiedene Methoden verwendet werden. Während die AUC in SAS über Proc Logistic berechnet werden kann, ist das Auffinden des optimalen Cutpoints nicht trivial. Verschiedene Macros werden über das Internet zur Verfügung gestellt. Die Beschreibung dieser Macros ist jedoch häufig unzureichend, so dass an vielen Stellen nicht klar ist, welche Funktionen vorhanden sind und welche Parameter eingegeben werden sollen. Daher wurden die verschiedenen Macros anhand eines realen Beispiels evaluiert und zusätzlich mit einer grafischen Methode zum Auffinden des optimalen Cutpoints (ohne Macro) und zu einem Paket in der R-Software verglichen.

Das finale Ergebnis der verschiedenen Möglichkeiten den optimalen Cutpoint zu bestimmen ist in unserem Beispiel nicht eindeutig. Zu beachten ist, dass weder die SAS-Macros noch das R-Paket validiert sind. Am Ende muss jeder Anwender über eine genaue Prüfung und ggf. durch eine unabhängige zweite Person sicherstellen, dass die Ergebnisse korrekt sind.

**Schlüsselwörter:** Diagnosestudie, Cutpoint, AUC, ROC-Kurve, Biomarker

## 1 Einleitung

Zunehmend werden in klinischen Studien Biomarker gemessen, die als diagnostischer Test herangezogen werden sollen. Da Biomarker als stetige Parameter vorliegen, werden für die Bewertung eines Biomarker als diagnostischer Test häufig Receiver Operator Characteristics (ROC) Analysen durchgeführt. Das Ziel solcher Untersuchungen ist

die Berechnung der Area Under the Curve (AUC) und ggf. das Auffinden eines optimalen Schwellenwerts (Cutpoint). Der Cutpoint dient dazu, mit Hilfe des stetigen Merkmals einen Test zur Einteilung in z.B. gesund/krank zu ermöglichen. In der herangezogenen Beispielstudie wurden Biomarker bei lebertransplantierten Patienten gemessen, um herauszufinden, ob eine Komplikation in Form einer Abstoßungsreaktion oder einer Re-Infektion vorliegen könnte. Da es sich bei Transplantation um eine seltene Erkrankung handelt, wurden für die Darstellung der statistischen Methoden im Rahmen dieses Vortrags zusätzliche Daten per Simulation generiert.

## 1.1 Medizinischer Hintergrund

Lebertransplantierte Patienten mit Hepatitis C laufen Gefahr nach der Transplantation entweder eine erneute Infektion der Leber mit Hepatitis C (Re-Infektion) oder eine Abstoßung der Leber (ACR=acute cellular reaction) zu erleiden. Problematisch ist hierbei, dass bei beiden Komplikationen die Symptomatik nahezu gleich ist, sich die Therapie aber grundsätzlich unterscheidet. Während bei einer Re-Infektion immunsuppressive Medikamente (Immunsuppressiva) erniedrigt werden, werden bei einer Abstoßungsreaktion Immunsuppressiva erhöht. In einer retrospektiven Studie wurden an lebertransplantierten Patienten mit Re-Infektion bzw. Abstoßungsreaktion verschiedene Biomarker untersucht. Ziel der Diagnosestudie war es Biomarker aufzufinden, die zwischen Re-Infektion und Abstoßungsreaktion differenzieren können. Dafür wurde die AUC der ROC-Kurve für jeden Biomarker berechnet. Solche Biomarker, die eine  $AUC > 0.8$  aufweisen, werden im Allgemeinen als geeignet angesehen. Für diese Biomarker wurde der optimale Cutpoint bestimmt, um zwischen „gesund“ und „krank“ zu unterscheiden.

## 1.2 Maßzahlen in Diagnosestudien

Wichtige Maßzahlen in Diagnosestudien sind die Sensitivität und Spezifität. Diese berechnen sich jeweils als richtig positiv Rate und richtig negativ Rate im Vergleich zum jeweiligen Goldstandard.

**Tabelle 1:** 4-Felder-Tafel einer Diagnosestudie

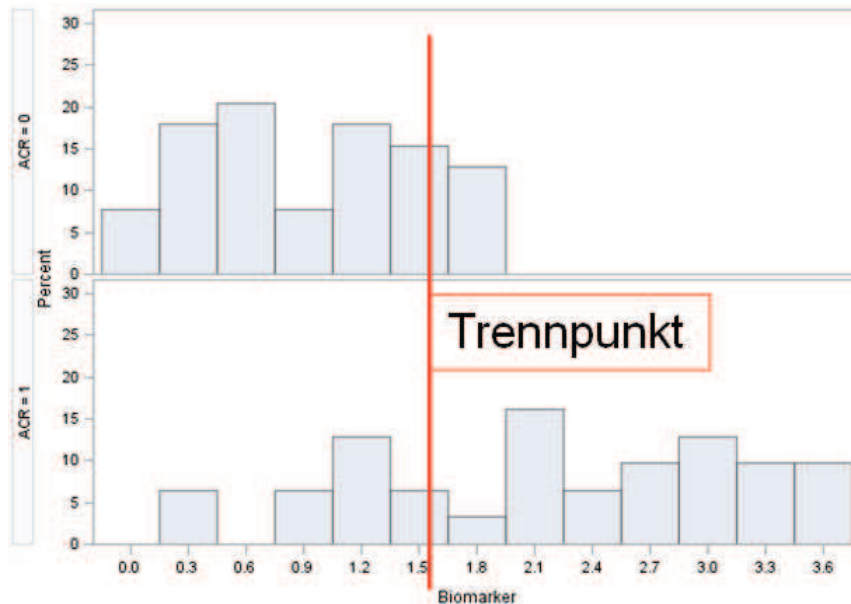
		Goldstandard		
		Krank	Gesund	
Test	Positiv	a	b	a+b
	Negativ	c	d	c+d
		a+c	b+d	N=a+b+c+d

$$\text{Sensitivität} = \frac{a}{(a+c)}$$

$$\text{Spezifität} = \frac{d}{(b+d)}$$

Stetige Werte müssen zunächst dichotomisiert werden. Dazu ist es notwendig einen Trennpunkt (Cutpoint) zu wählen. Folglich sind Sensitivität und Spezifität von der

Wahl des jeweiligen Cutpoints abhängig. Die folgende Grafik veranschaulicht die Verteilung eines Biomarker für den jeweiligen Status (ACR=0 bzw. ACR=1).



**Abbildung 1:** Verteilung des Biomarker nach ACR-Status und fiktiver Cutpoint

Anhand der Grafik lässt sich erkennen, dass manche Patienten richtig, andere wiederum falsch zugeordnet werden: Patienten mit einem Biomarkerwert kleiner als der Cutpoint würden bei ACR=0 richtig negativ (Spezifität) und bei ACR=1 als falsch negativ eingestuft werden. Patienten mit einem Wert größer als der gewählte Cutpoint würden bei ACR=1 als richtig positiv eingestuft (Sensitivität); bei ACR=0 wären sie hingegen falsch positiv.

Verschiebt man diesen Cutpoint, so erhält man jeweils unterschiedliche Werte für Sensitivität und Spezifität. Der jeweilige Cutpoint sollte jedoch nicht willkürlich gewählt werden. Der optimale Cutpoint kann beispielsweise über den Youden-Index bestimmt werden. Dabei ist der jeweilige Cutpoint zu wählen, bei dem der Youden-Index am größten ist.

$$\text{Youden-Index} = \text{Sensitivität} + \text{Spezifität} - 1$$

Zur Bestimmung des Cutpoints sind auch andere Verfahren möglich. Beispielsweise können Sensitivität und Spezifität unterschiedlich gewichtet werden.

### 1.3 Berechnung der AUC einer ROC-Kurve in SAS

Bevor der optimale Cutpoint für einen stetigen Biomarker bestimmt wird, sollte zunächst geprüft werden, ob sich der Biomarker überhaupt dazu eignet, Vorhersagen für den Gesundheitsstatus machen zu können. Dazu wird die AUC der ROC-Kurve bestimmt. In unserem Beispiel wurden nur Biomarker größer 0.8 weiter untersucht. Die Bestimmung der AUC kann über die Logistic-Prozedur in SAS erfolgen:

```
ods graphics on;  
proc logistic data = KSFE.sim2 plots(only)=(roc(id=casenum));  
model ACR= biomarker;  
run;
```

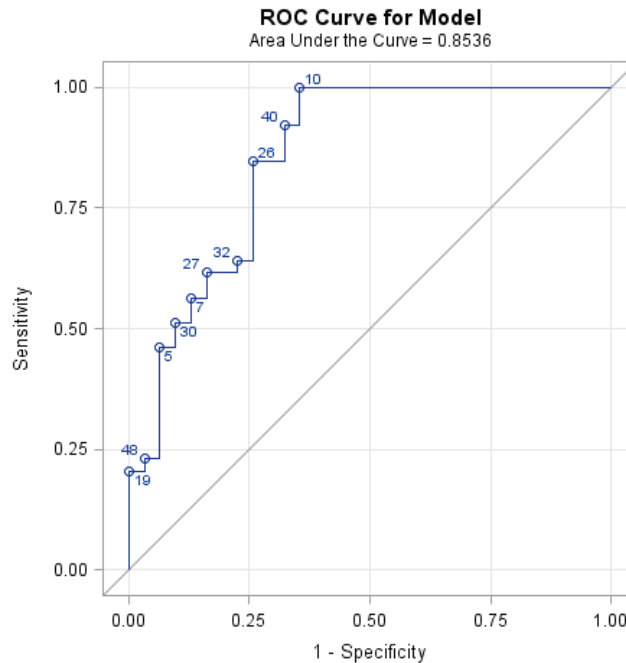


Abbildung 2: ROC-Kurve mit Kennzeichnung der Patienten-IDs

Um die ROC-Kurve betrachten können, muss zunächst das `ods graphics` – Statement aktiviert werden: `ods graphics on`. Das `plots(only)` – Statement dient dazu, andere voreingestellte Grafiken außer der ROC-Kurve zu unterdrücken. Markierungen an der ROC-Kurve können mit Hilfe des `id=` – Statements vorgenommen werden. Die Ausgabe kann um weitere Einstellungen, z.B. das Anzeigen des 95%-Konfidenzintervalls der AUC erweitert werden.

## 2 Auffinden des optimalen Cutpoints in SAS

Obwohl häufig benötigt, ist das Auffinden des optimalen Cutpoints keine Standardprozedur in SAS. Mittlerweile gibt es jedoch verschiedene Macros in SAS, die dies ermöglichen sollen. Allerdings sind die bestehenden Macros, deren Inhalte und Funktionen mitunter nur unzureichend beschrieben. Auch ist die Benennung der Macros nicht immer eindeutig, da es gleiche Macros mit verschiedenen Namen und verschiedenen Macros mit gleichem Namen gibt. Im Rahmen dieser Ausarbeitung wurden folgende Macros genauer betrachtet:

- %ROC (Gönen, o.J.)
- %PLOTROC (Gönen, o.J.)
- %ROCPLLOT (Gönen, 2007)
- %ROCPLLOT (Friendly, 2012)
- %ROCCUTOFF (Chen, 2010)

Das %ROC-Macro [1] bietet keine Möglichkeit zur Berechnung des optimalen Cutpoints, liefert aber die Möglichkeit, weitere Informationen für die AUC berechnen zu können (z.B. Standardfehler sowie Konfidenzintervall für die AUC). Außerdem kann es verwendet werden, wenn verschiedene AUCs miteinander verglichen werden sollen. Bei der Verwendung sollte die Aktualisierungsprüfung im Macro Code auskommentiert bzw. ausgeschaltet werden, da sich das Macro ansonsten mit einer nicht mehr existierenden Website zu verbinden versucht.

Das %PLOTROC-Macro stammt wie das %ROC PLOT-Macro von Gönen [1,2]. Ein Anwendungsbeispiel für das %PLOTROC-Macro wird in [1] gegeben. Das Macro ist auf Anfrage beim Autor erhältlich. Das %ROC PLOT-Macro kann hingegen direkt von der SAS-Webseite heruntergeladen werden. Auf Anfrage, welche Unterschiede zwischen den beiden Macros bestehen hat sich jedoch gezeigt, dass es sich um die gleichen Macros handelt.

Das %ROC PLOT-Macro von Friendly [3] zeichnet die ROC-Kurve mit kleineren optischen Veränderungen. Wenig praktikabel ist jedoch, dass ein ID-statement angegeben werden muss, so dass die ROC-Kurve immer mit den entsprechenden IDs (z.B. Patientennummer) gekennzeichnet wird. Beachtet werden sollte auch, dass im Macro das Symbol-Statement mit `interpolation = join` verwendet wird, das temporär die Grafikeinstellungen auch für z.B. nachfolgende Histogramm-Plots verändert. Ein Cutpoint kann mit dem Macro nicht berechnet werden.

Das %ROCCUTOFF-Macro [4] sowie das %ROC PLOT-Macro von Gönen bieten beide eine Möglichkeit zur Berechnung des optimalen Cutpoints. Beide Macros werden in den Abschnitten 2.2 und 2.3 ausführlicher dargestellt.

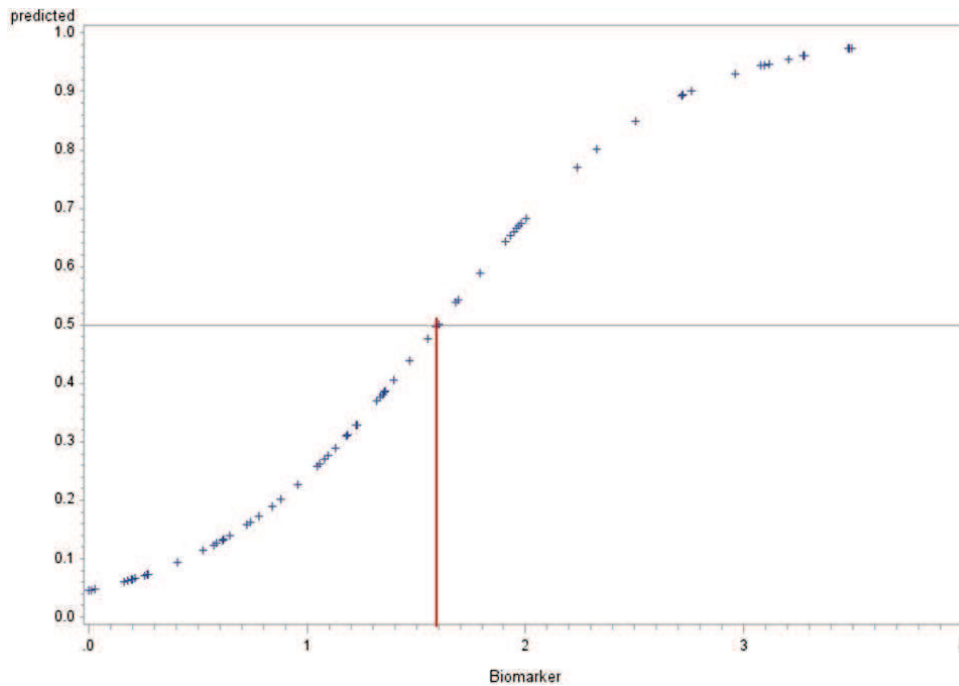
## 2.1 Grafische Darstellung der angepassten logistischen Funktion

Eine grafische Möglichkeit zum Auffinden des optimalen Cutpoints bietet die Darstellung der angepassten logistischen Funktion, wie bereits in einem früheren KSFE-Beitrag von Altenburg (2009) zu Diagnosestudien mit stetigen Vorhersagewerten nachgelesen werden kann [5]. Über diese Methode lässt sich der optimale Cutpoint grafisch abschätzen.

Über die Logistik Prozedur passt man zunächst ein logistisches Modell an und speichert die geschätzten Werte des logistischen Modells (Vorhersagewerte) in einer Datei:

```
proc logistic data=KSFE.sim2 descending;
  model ACR = biomarker;
  output out=estoutboth p=predicted;
run;
```

Die so berechneten Vorhersagewerte können nun über eine simple Grafikprozedur (z.B. `proc gplot`) gegen die Biomarker-Variable dargestellt werden.



**Abbildung 3:** Grafik der angepassten logistischen Regression

Diese angepasste logistische Funktion verläuft in Form einer S-Kurve und steigt bei Populationen die sich gut trennen lassen (krank/gesund) umso steiler an [5]. Eine Hilfslinie an der der Stelle 0.5 zeigt die Trennung zwischen "eher krank" und "eher gesund" an. Den optimalen Trennpunkt kann man an der x-Achse ablesen, an der Stelle wo sich die angepasste Kurve und die Hilfslinie schneiden (predicted=0.5). Die x-Achse zeigt dabei die Skala des Biomarkers an. In unserem Beispiel kann man den optimalen Cutpoint für den entsprechenden Biomarker bei etwa 1.6 ablesen.

## 2.2 %ROC PLOT

Eine weitere Möglichkeit zum Bestimmen des optimalen Cutpoints bietet das %ROC PLOT-Macro von Gönen [2]:

```
%ROCplot(dsn=KSFE.sim2,  
          marker=biomarker,  
          gold=ACR,  
          anno=0,  
          tlist=1.9,  
          round=0.1,  
          optimal=);
```

Folgende Eingabeparameter sind zu definieren:

- dsn = Name des Datensatzes
- marker = Name der Variable, für die der Cutpoint gesucht wird
- gold = dichotomer Zielparameter
- anno = spezifiziert den gewünschten Output der ROC-Kurve:



- 0: kein grafischer Output
  - 1: es wird nur die ROC-Kurve ausgegeben
  - 2: zeigt eine diagonale Linie (45°) an
  - 3: die ROC Kurve wird an den in `tlist=` spezifizierten Punkten markiert
  - 4: fügt horizontale und vertikale Referenzlinien ein
- `tlist` = spezifiziert die Punkte an denen die ROC-Kurve spezifiziert werden soll (optional)
  - `round` = spezifiziert die Zwischenstufen für die der Cutpoint ausgegeben werden soll (z.B. 0.5 → 0 / 0.5 / 1 ...); es ist nicht die Anzahl der Nachkommastellen gemeint
  - `optimal` = Ausgabe einer Liste mit den jeweiligen Abständen der Cutpoints.

Das Macro gibt mit der Parameterspezifikation `anno=0` für jeden möglichen Cutpoint die jeweiligen Abstände zur linken oberen Ecke und der 45° Linie in Form einer Liste aus. Die Abstände zur linken oberen Ecke sollten möglichst klein sein, da hier die Sensitivität und die Spezifität am höchsten ist (=perfekter Marker). Dies entspricht der Bestimmung des optimalen Cutpoints nach dem Euklidischen Abstand. Die Abstände zur 45° Linie sollten möglichst hoch sein (=nicht informativer Marker). Bei dieser Methode werden die Cutpoints entsprechend des Youden-Index bewertet. Ohne Spezifikation des `optimal=` Parameters wird die Liste nach den Werten der Cutpoints sortiert, mit `optimal=1` wird die Liste nach dem Abstand zur linken oberen Ecke sortiert. Es ist zu beachten, dass die Liste nur ausgegeben wird, wenn in der Parameterspezifikation mit `anno=0` kein Output spezifiziert ist. Es gibt keine Möglichkeit die Liste und eine Grafik gemeinsam ausgeben zu lassen. Die Parameterspezifikation `anno=4` fügt für die in `tlist=` spezifizierten Parameter eine horizontale und eine vertikale Referenzlinie zur x- und y-Achse ein.

Der optimale Cutpoint kann nun anhand der jeweiligen Abstände herausgesucht werden. In unserem Beispiel findet sich der optimale Cutpoint für den entsprechenden Biomarker bei 1.5 oder 1.9, je nachdem ob mehr Wert auf einen kleinen Abstand zur linken oberen Ecke oder einen großen Abstand zur 45° Linie gelegt wird (Euklidischer Abstand vs. Youden-Index).

Threshold	False Positive Rate	True Positive Rate	Distance from Perfect Marker	Distance from Non-Informative Marker
1.5	0.17949	0.74194	0.31435	0.56245
1.4	0.20513	0.74194	0.32966	0.53681
1.7	0.07692	0.67742	0.33163	0.60050
1.6	0.12821	0.67742	0.34712	0.54921
1.9	0.00000	0.64516	0.35484	0.64516

Abbildung 4: Ausgabe der optimalen Cutpoints über `%ROCPLLOT`

Vorteil dieser Liste ist, dass die möglichen Cutpoints der entsprechenden Biomarker, anders als bei der grafischen Methode (siehe 2.1), direkt abgelesen werden können. Allerdings können sich beide Methoden (Euklidischer Abstand vs. Youden-Index) in den Ergebnissen unterscheiden. Auch kann man sich vorstellen, dass die Liste bei einer großen Anzahl möglicher Cutpoints (Spannweite des Biomarkers) schnell sehr unübersichtlich werden kann. Außerdem bietet das Macro leider keine Möglichkeit, die Liste als Datensatz zu speichern, um sie später oder in anderer Form weiter verwenden zu können.

## 2.3 %ROCCUTOFF

Beim %ROCCUTOFF-Macro [4] von Chen (2010) sieht die Eingabe wie folgt aus:

```
%ROCcutoff(DSName= KSFE.sim2,  
            Dir= Q:\Gonnermann\KSFE\  
            ID= PatID,  
            Outcome= ACR,  
            PredictNum= Biomarker,  
            Format= fyn.,  
            WSpec= 0.5,  
            OutForm= HTML);
```

Folgende Eingabeparameter sind zu definieren:

- DSName = Name des SAS Datensatzes
- Dir = Speicherort für die Ergebnisse
- ID = Name der eindeutigen Patientent-ID für jede Beobachtung
- Outcome = Dichotomer Zielparameter
- PredictNum = Name der Variable, für die der Cutpoint gesucht wird
- Format = Format des Zielparameters (definiert über **proc format**)
- WSpec = Zahlen zwischen 0 und 1 sind möglich, z.B. wird bei 1 der Cutpoint so gewählt, dass die Spezifität maximal, bei 0 die Sensitivität maximal ist. Bei 0.5 werden beide gleich gewichtet.
- OutForm = Für die Ausgabe kann HTML oder RTF gewählt werden.

Im Macro werden zunächst mit `Proc logistic` Parameter berechnet, z.B. Sensitivität und 1-Spezifität, in eine Datei herausgeschrieben und die ROC-Kurve ausgegeben.

Diese Werte werden verwendet, um zu berechnen, bei welchem Biomarker der Wert „cut“ am größten ist. Wird beispielsweise Sensitivität und Spezifität gleich gewichtet, so wird in diesem Macro der Cutpoint gewählt, bei dem der Youden-Index am größten ist. Bei der Ausgabe wird der Cutpoint und die entsprechenden Patienten-ID angegeben:



**logistic regression with the continuous predictor**

Obs	PatID	biomarker_cutoff
1	10	1.94531

**Abbildung 5:** Ausgabe des optimalen Cutpoints über %ROCCUTOFF

Anschließend wird im nächsten Schritt der Biomarker nach dem optimalen Cutpoint dichotomisiert und erneut die logistische Regression durchgeführt und somit das Odds Ratio für den dichotomen Biomarker ausgegeben.

Zu beachten ist, dass nachdem das Macro verwendet wurde, ein `ods html`-Statement erforderlich ist, da das Macro die Output-Fenster deaktiviert.

Bei diesem Macro erhält man einen eindeutigen Wert, während anhand des %ROCPLLOT-Macros zwei Werte in Frage kommen würden (1.5 oder 1.9). Nach der grafischen Methode erhält man einen Wert, der nicht diesem entspricht (ca. 1.6). Somit stellt sich die Frage, welcher Cutpoint nun gewählt werden sollte? Im Folgenden werden die Berechnungen mit Hilfe eines R-Pakets dargestellt und mit den SAS-Berechnungen verglichen.

### 3 R-Paket "OptimalCutpoints"

Das R-Paket „OptimalCutpoints“ von Lopez-Raton und Rodriguez-Alvarez [6,7] enthält numerische und grafische Ausgaben, um den optimalen Cutpoint zu finden. Dabei sind verschiedene Methoden und Einstellungen möglich. In diesem Fall beschränken wir uns auf den Youden-Index, um den Vergleich zu den vorgestellten Macros herzustellen.

Nachdem das Paket erfolgreich geladen wurde, kann mit folgendem Aufruf die Berechnung analog zu dem SAS-Macro %ROCCUTOFF erfolgen:

```
optimal.cutpoints(X="Biomarker",
                 status="ACR",
                 tag.healthy = 1,
                 method="Youden",
                 data=daten,
                 direction=c(">"),
                 conf.level = 0.95);
```

Über `x` wird die stetige Variable spezifiziert, über `status=` das dichotome Ergebnis des Goldstandards. Über `tag.healthy=` wird angegeben, in welcher Form der Status codiert wurde. Üblicherweise ist die Diagnose bei 0 negativ und bei 1 positiv. In unserem Beispiel ist dies auch der Fall, jedoch weicht dann das Ergebnis vom SAS-Macro %ROCCUTOFF ab. Man kann jedoch über `direction=` ändern, ob hohe oder niedrige Werte eine Person als krank oder gesund klassifizieren. Automatisch ist eingestellt, dass kleinere Werte für gesund stehen. In unserem Beispiel haben wir die Richtung geändert, so dass größere Werte den Status „gesund“ kennzeichnen. Aufgrund der doppelten

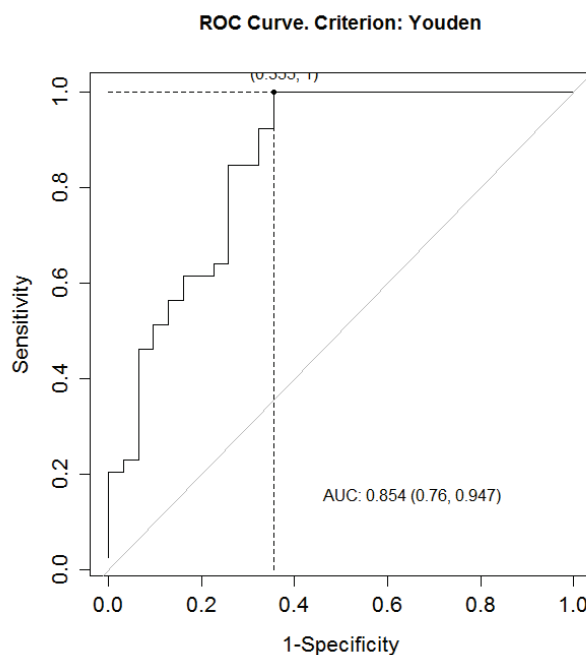
Richtungsänderung (`tag.healthy= 1` und `direction=c(">")`) hat man nun wieder die gleiche Einstellung gewählt: Hohe Werte stehen für den Status 1 (hier: krank). Über `method=` wurde hier der Youden-Index gewählt, zu dem auch ein 95%-Konfidenzintervall (`conf.level =`) ausgegeben werden soll. Die Daten werden über `data=` eingegeben.

Das Ergebnis ist mit dem vom `%ROCCUTOFF`-Macro vergleichbar:

```
Optimal cutoffs:  
  Youden  
1 1.9453  
  
Area under the ROC curve (AUC): 0.854 (0.76, 0.947)
```

**Abbildung 6:** Ausgabe des optimalen Cutpoints über das R-Paket

Auch Grafiken (z.B. die ROC-Kurve) können direkt über dieses Paket ausgegeben werden. In unserem Beispiel wäre die Grafik allerdings noch zu optimieren, da das Konfidenzintervall nicht erkennbar ist:



**Abbildung 7:** ROC-Kurve mit Hilfe des R-Pakets

## 4 Fazit

Obwohl in Diagnosestudien das Auffinden des Cutpoints oft gebraucht wird, gibt es keine Funktion in SAS, mit der diese Auswertung direkt durchgeführt werden kann. Mittlerweile stehen jedoch im Internet verschiedene Macros in SAS bzw. Pakete in R zur Verfügung, mit denen diese Auswertung prinzipiell möglich ist. Die Beschreibung dieser Macros ist allerdings oft unzureichend, so dass an vielen Stellen nicht klar ist, welche Funktionen vorhanden sind und welche Parameter eingegeben werden sollen. Bei den betrachteten Macros stellte sich außerdem heraus, dass manche Aufrufe z.B. Grafikeinstellungen in SAS temporär oder dauerhaft verstellen können. Zusätzlich gab es Funktionen innerhalb eines Macros (`%ROC-Macro`), die veraltet sind und dadurch nicht mehr funktionieren. Bei den R-Paketen kann es hingegen vorkommen, dass sie zu einem späteren Zeitpunkt nicht mehr zur Verfügung stehen.

Eine Validierung der Ergebnisse, indem man die Ergebnisse der Macros miteinander vergleicht bzw. mit Hilfe des R-Pakets die SAS-Ergebnisse bestätigt, ist nicht gelungen. Das finale Ergebnis der verschiedenen Möglichkeiten den optimalen Cutpoint zu bestimmen ist in unserem Beispiel nicht eindeutig. Anhand der grafischen Methode lässt sich der optimale Cutpoint auf etwa 1.6 abschätzen. Verwendet man das `%ROCPLOT` -Macro von Gönen ergibt sich der optimale Cutpoint, je nachdem ob mehr Wert auf einen großen Abstand zum nicht-informativen Marker oder einen geringen Abstand zum perfekten Marker gelegt wird, bei 1.5 oder 1.9. Hier sollte man sich vorher überlegen, ob man den Cutpoint entsprechend dem Euklidischen Abstand oder nach dem Youden-Index bestimmen möchte. Mit dem `%ROCCUTOFF` -Macro wird der optimale Cutpoint bei 1.9453 bestimmt. Unter Verwendung der R Software findet sich der gleiche Wert für den optimalen Cutpoint: 1.9453. Allerdings sind die Ergebnisse des R-Pakets und die des `%ROCCUTOFF`-Macro nur vergleichbar, wenn die Eingabe im R-Paket wie oben beschrieben angepasst wird.

Die Daten, auf deren Basis der Cutpoint im `%ROCCUTOFF` -Macro bestimmt wurde, können in SAS problemlos eingesehen und nachvollzogen werden, so dass wir zu dem Schluss kommen, dass die Einstellungen des SAS-Macros `%ROCCUTOFF` korrekt sind. Insgesamt muss jeder Anwender über die genaue Prüfung und ggf. durch eine unabhängige zweite Person sicherstellen, dass die Ergebnisse korrekt sind.

## Literatur

- [1] M. Gönen: Receiver Operating Characteristics (ROC) Curves, SUGI-31. Verfügbar im Internet:  
<http://www2.sas.com/proceedings/sugi31/210-31.pdf> [23.03.2015]
- [2] M. Gönen: Analyzing Receiver Operating Characteristics Curves with SAS, SAS Press Series, North Carolina 2007.
- [3] M. Friendly: %Rocplot, GitHub 2012. Verfügbar im Internet:  
<https://github.com/friendly/SAS-macros/blob/master/rocplot.sas> [23.03.2015]
- [4] R. Chen: %ROCcutoff, 2010. Verfügbar im Internet:  
<https://www.ctspedia.org/do/view/CTSpedia/ROCcutoff> [13.03.2015]
- [5] H.P. Altenburg: Analyse von ROC-Kurven mit Hilfe von SAS, 2009. Verfügbar im Internet:  
[http://de.saswiki.org/images/9/9d/13.KSFE-2009-Altenburg-Analyse\\_von\\_ROC-Kurven.pdf](http://de.saswiki.org/images/9/9d/13.KSFE-2009-Altenburg-Analyse_von_ROC-Kurven.pdf) [03.02.2015]
- [6] M. Lopez-Raton, M.X. Rodriguez-Alvarez: Computing optimal cutpoints in diagnostic tests, 2014. Verfügbar im Internet:  
<http://cran.r-project.org/web/packages/OptimalCutpoints/OptimalCutpoints.pdf> [05.02.2015]
- [7] R Development Core Team. R: A language and environment for statistical computing, reference index version 3.1.0. R Foundation for Statistical Computing, Vienna, Austria, 2014. Verfügbar im Internet:  
<http://www.R-project.org> [05.02.2015].