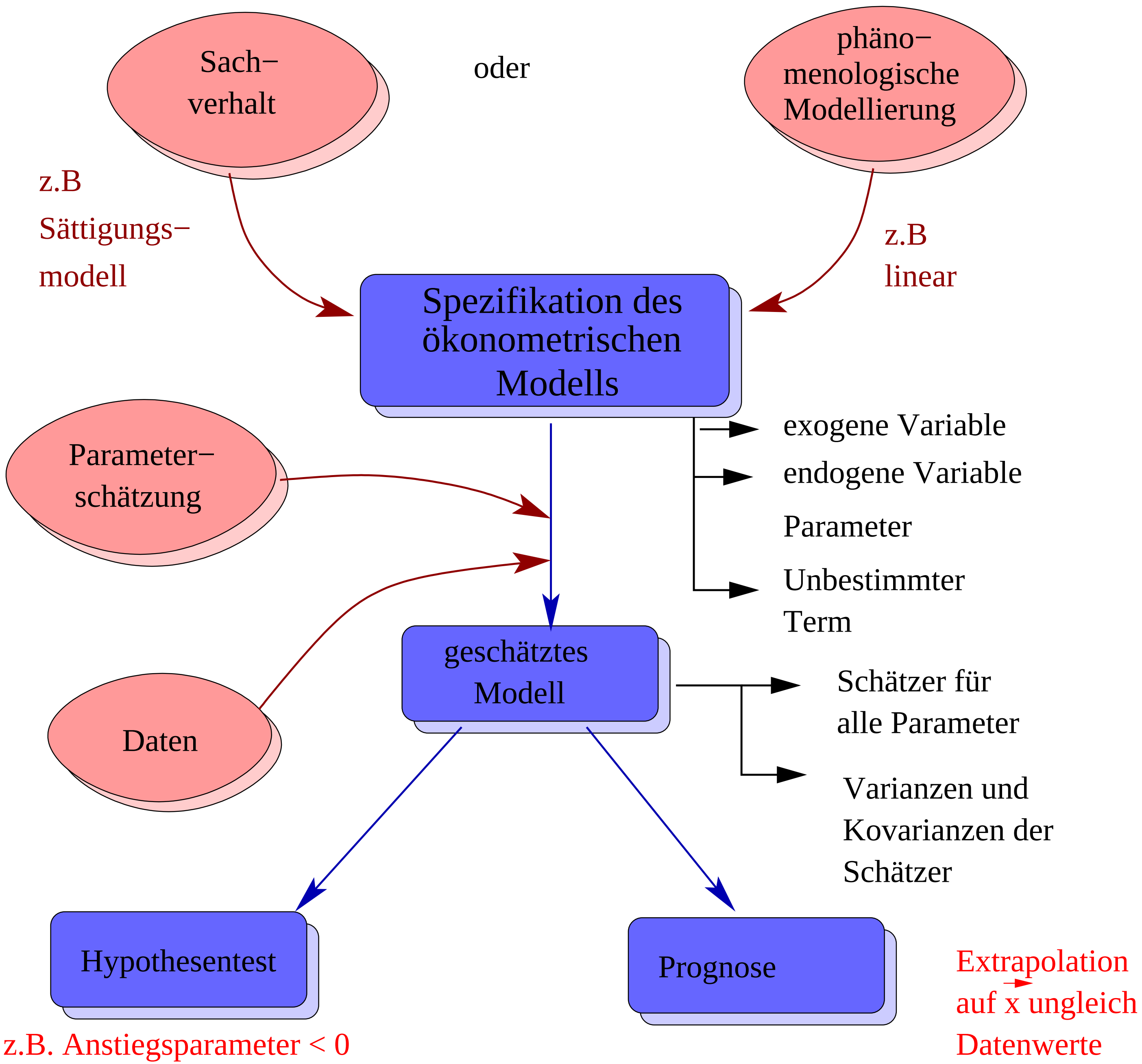
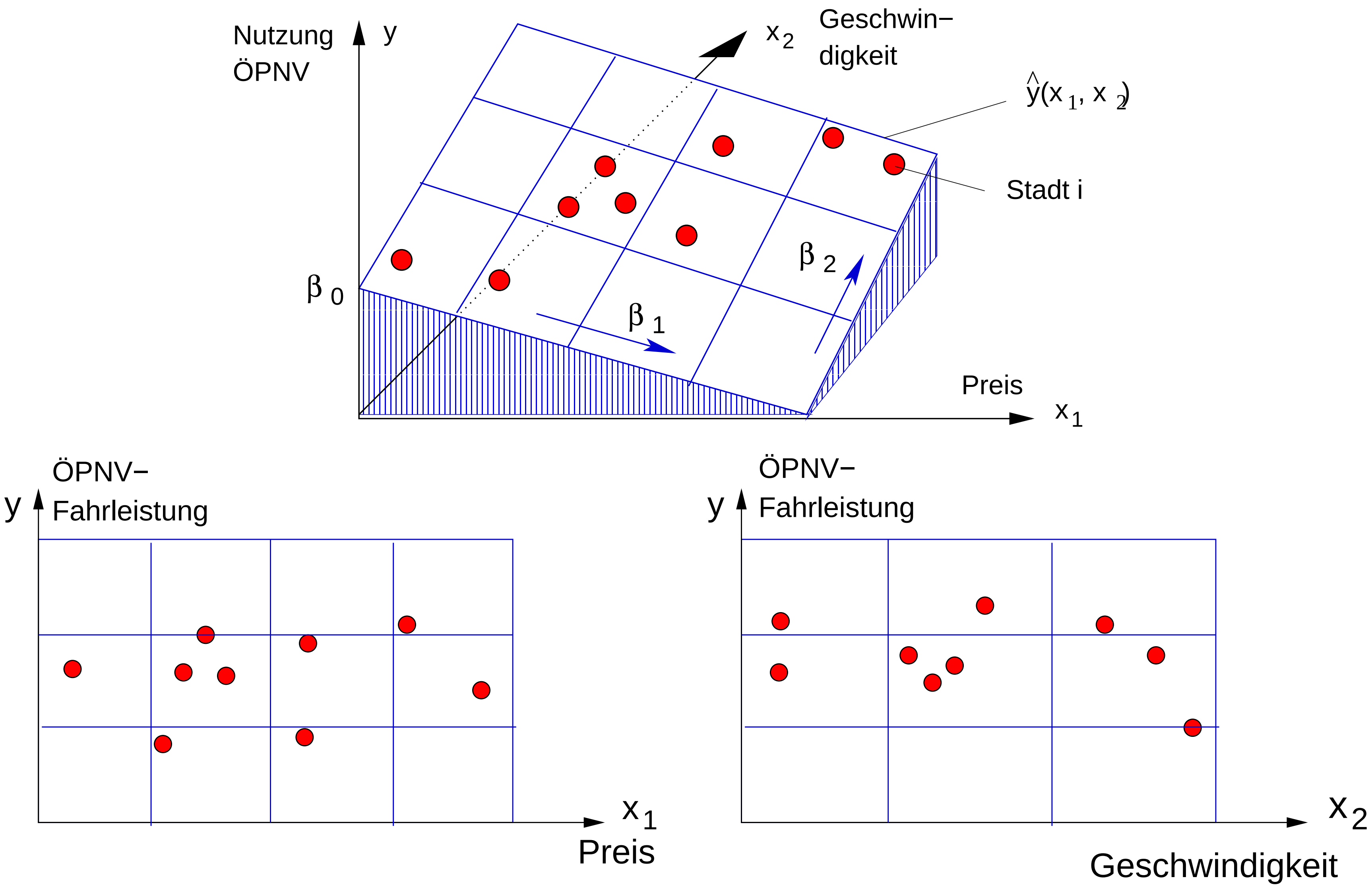
Flussdiagramm der ökonometrischen Methode

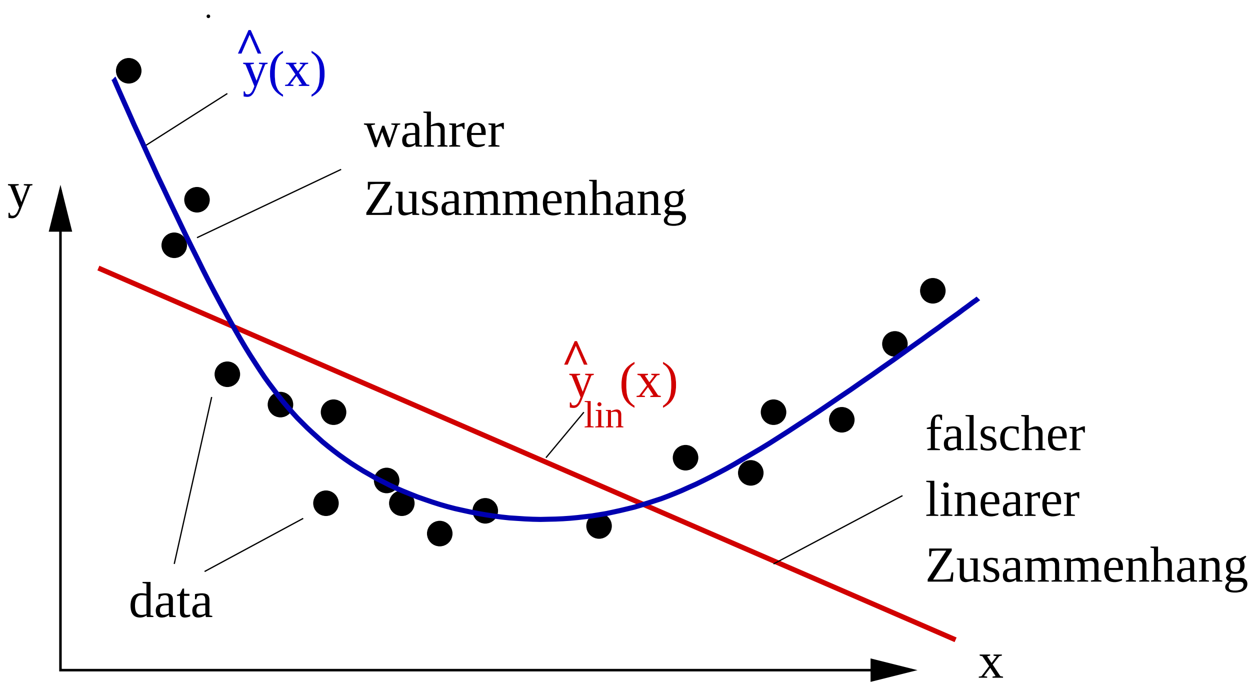


Modellspezifikation I: Funktionale Spezifikation 1



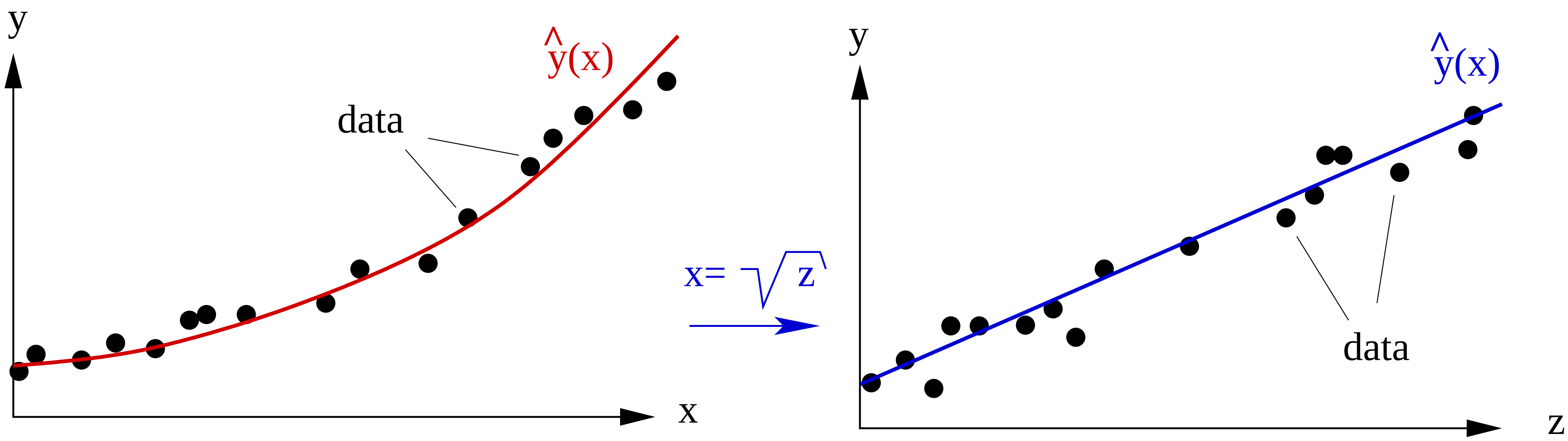
1. Alle relevanten Einflussfaktoren sind berücksichtigt (oben, nicht aber unten)

Modellspezifikation I: Funktionale Spezifikation 2



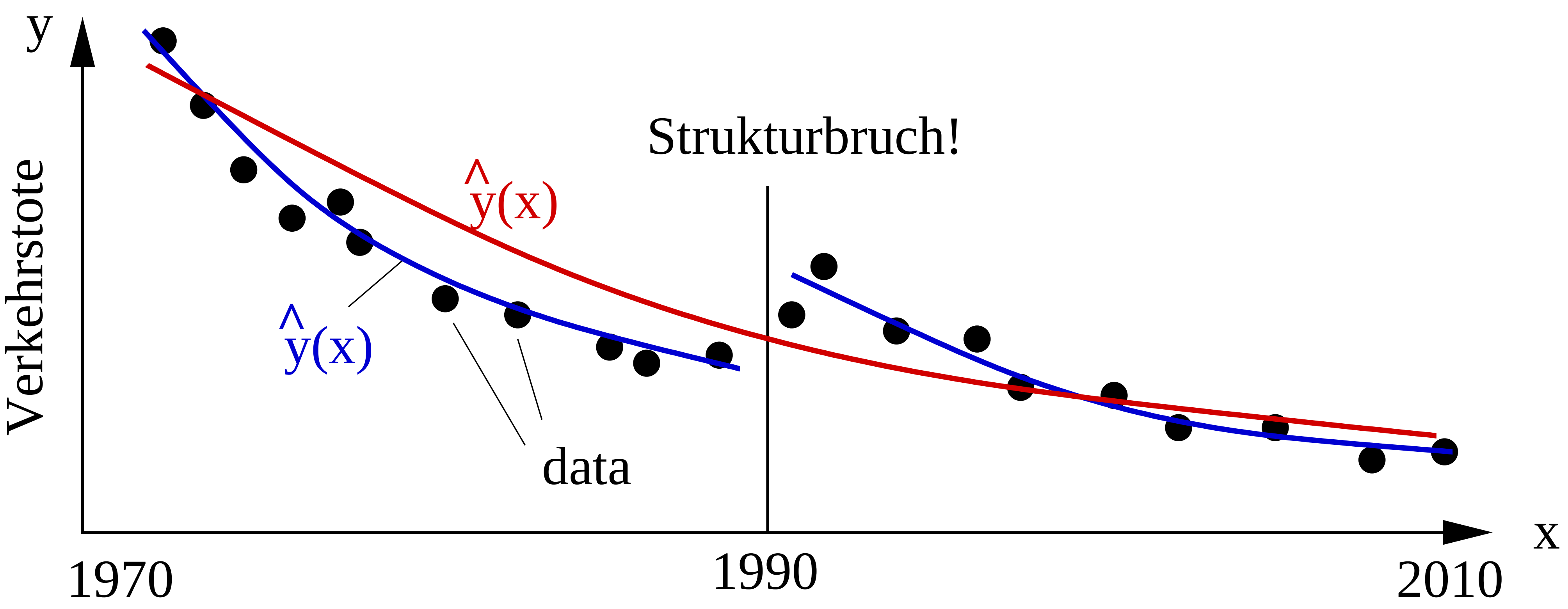
2. Das Modell ist linear, was hier nicht erfüllt ist

Modellspezifikation I: Funktionale Spezifikation 2



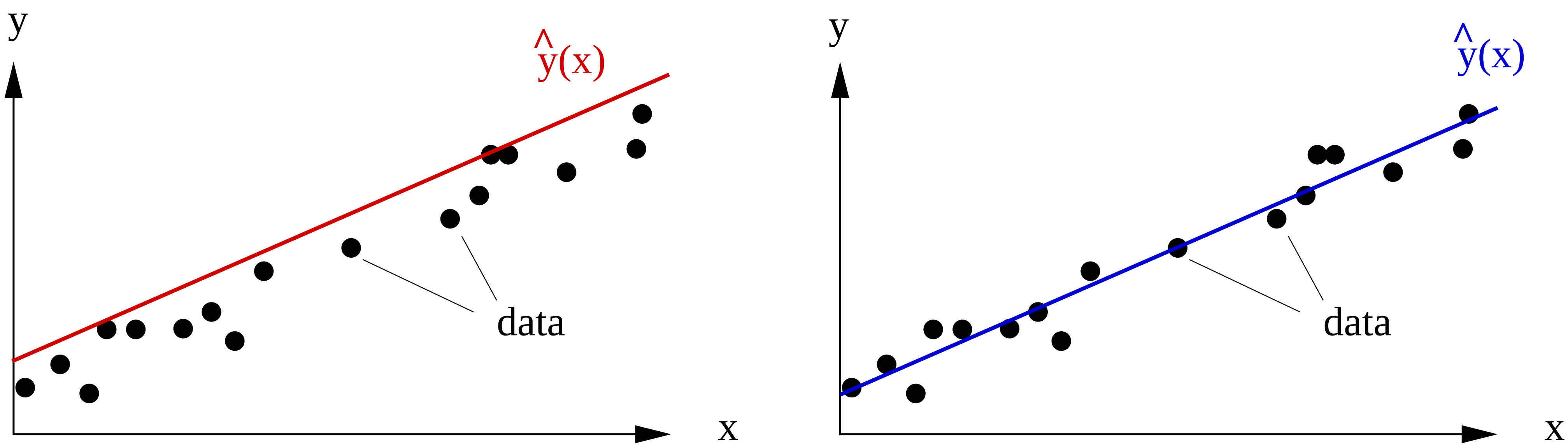
Manchmal kann das Modell durch Transformationen der exogenen und/oder endogenen Variablen linearisiert werden

Modellspezifikation I: Funktionale Spezifikation 3



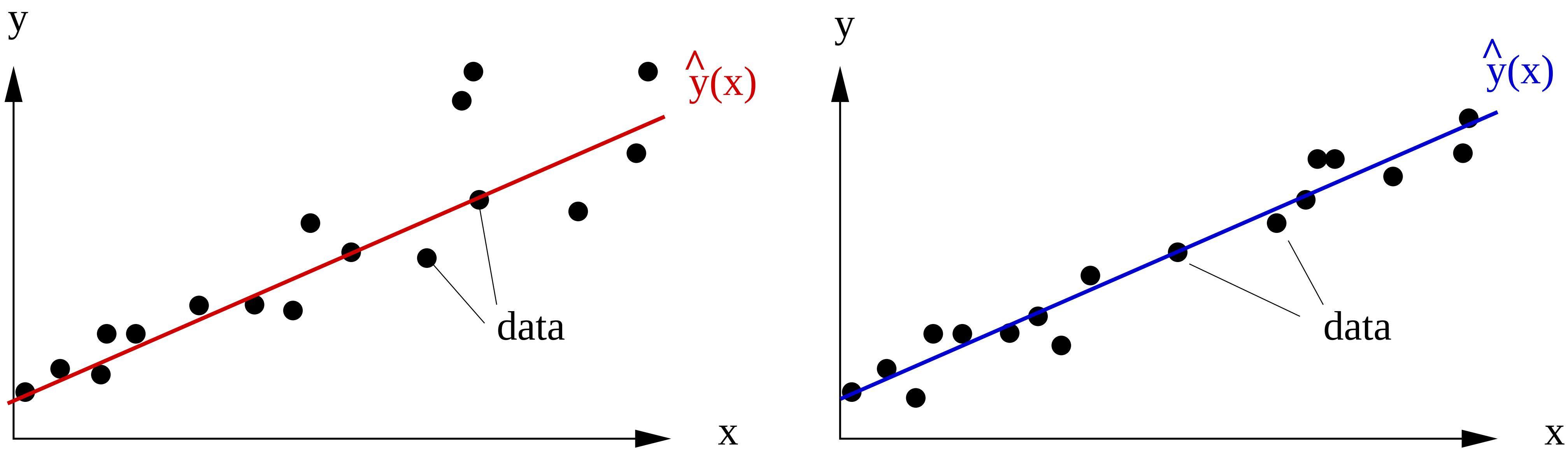
3. Homogenitätskriterium (z.B. kein Strukturbruch im Raum der exogenen Variablen, wie hier gezeigt)

Modellspezifikation II: Statistische Spezifikation 1



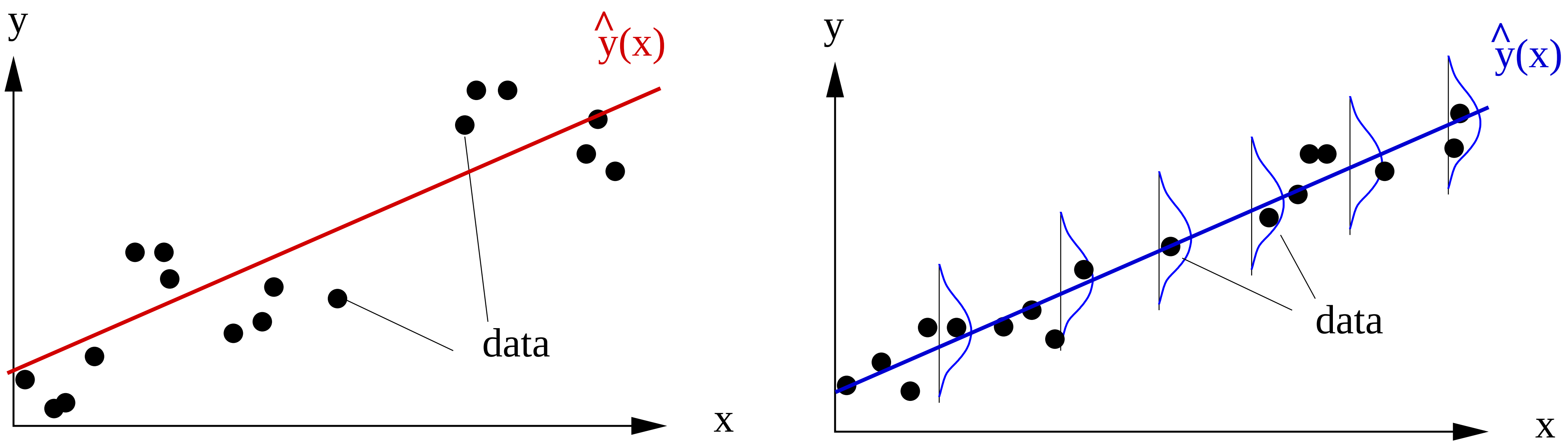
1. Der Erwartungswert der Störgröße muss verschwinden.

Modellspezifikation II: Statistische Spezifikation 2



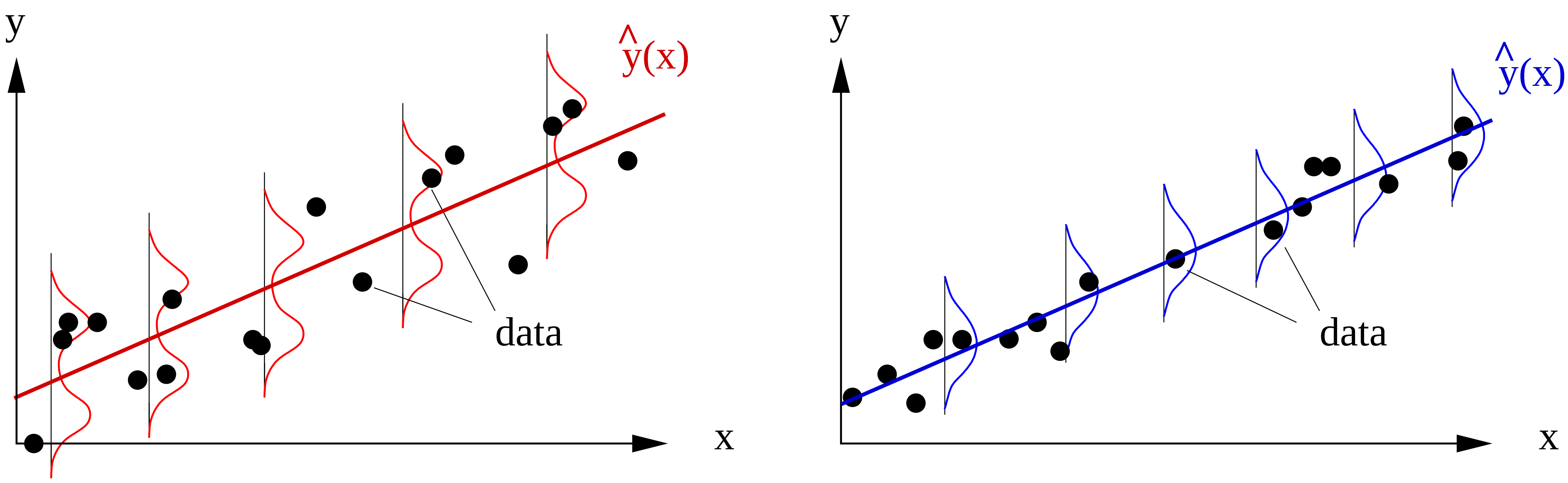
2. Der Residualterm ϵ ist homoskedastisch (rechts),
nicht etwa heteroskedastisch (links)

Modellspezifikation II: Statistische Spezifikation 3



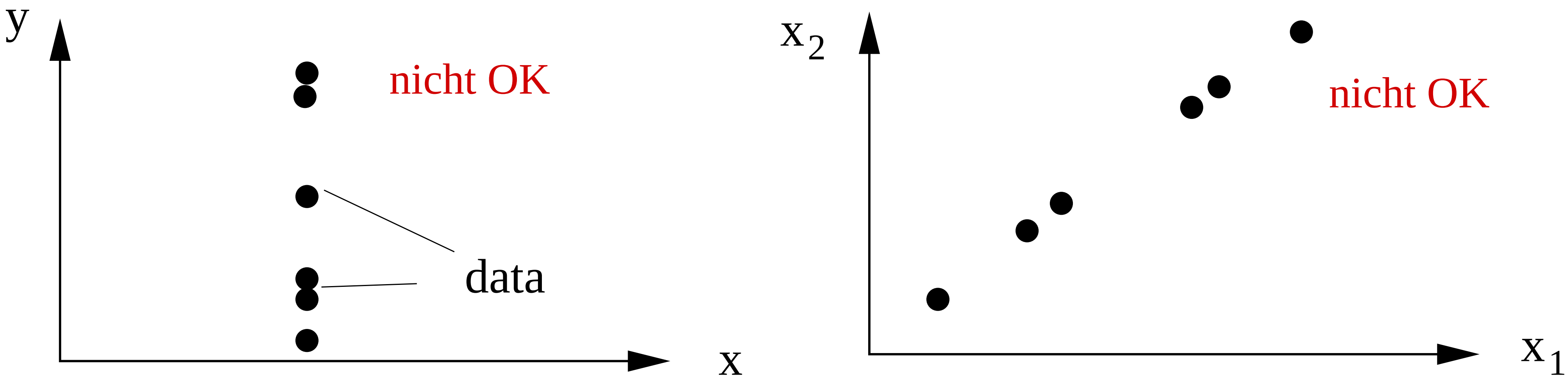
Keine Korrelationen von ϵ bezüglich x_i oder y (rechts),
während das Modell links fehlspezifiziert ist

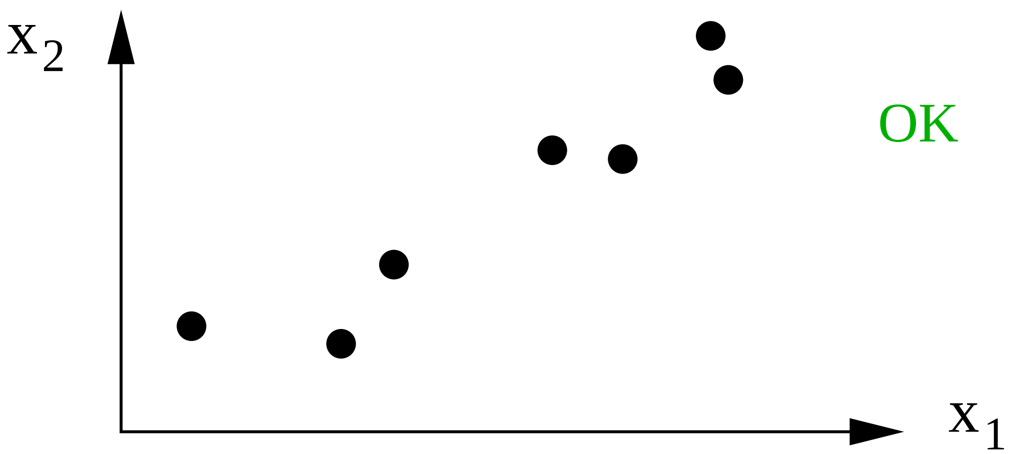
Modellspezifikation II: Statistische Spezifikation 4



Der Residualterm ϵ ist gaußverteilt (rechts),
nicht etwa bimodal verteilt (links)

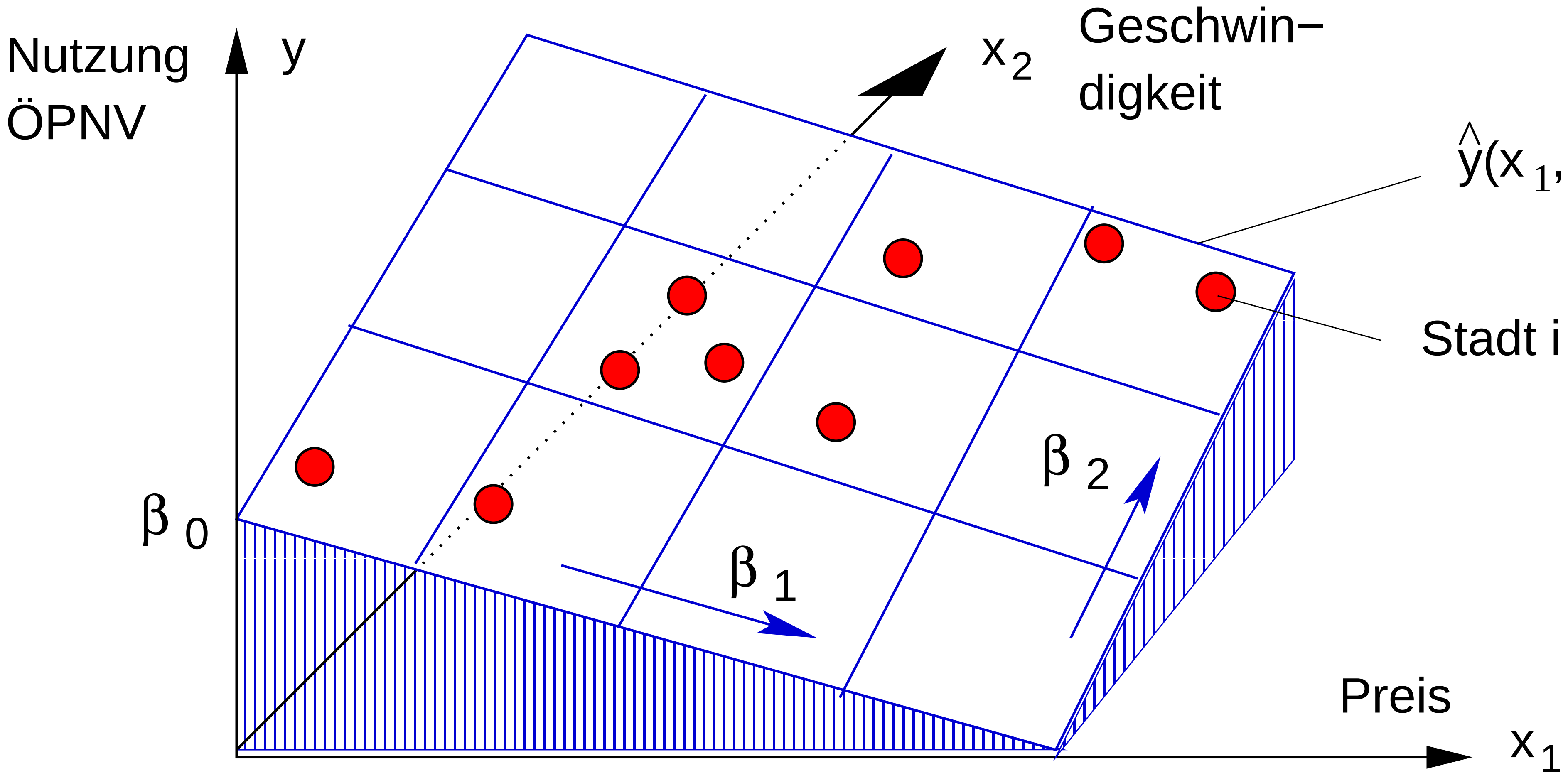
Modellspezifikation III: Datenspezifikation





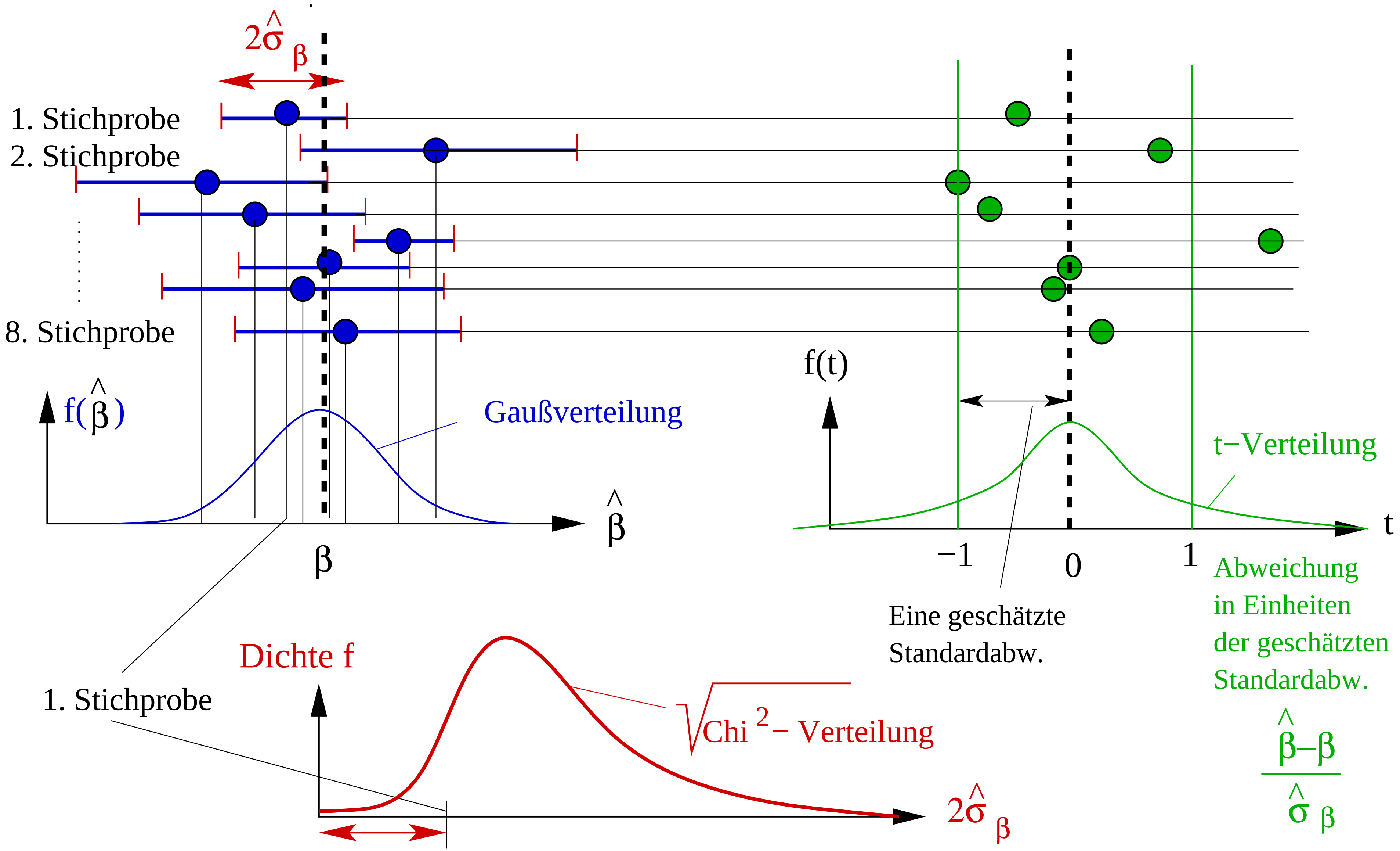
Keine der exogenen Variablen darf sich als Linearkombination aus Konstanten und anderen exogenen Variablen darstellen lassen (oben); nichtperfekte Korrelationen sind aber erlaubt (unten)

Lineares Modell mit zwei exogenen Variablen (schematisch)

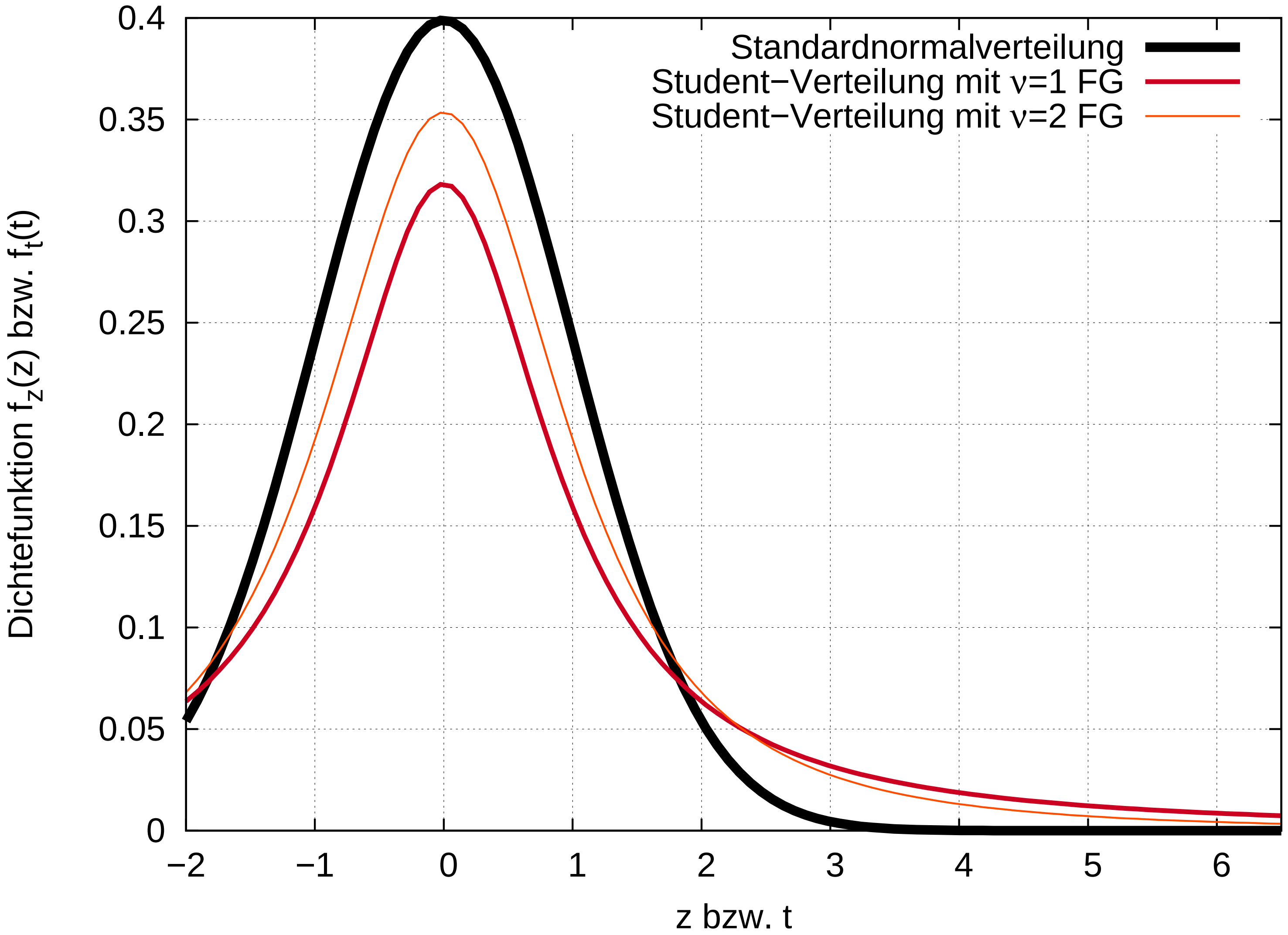


Achsabschnitt β_0, Anstiegsparameter β_1 und β_2. Wenn die Datenpunkte genau auf der Ebene liegen, gehorchen sie dem linearen Modell exakt

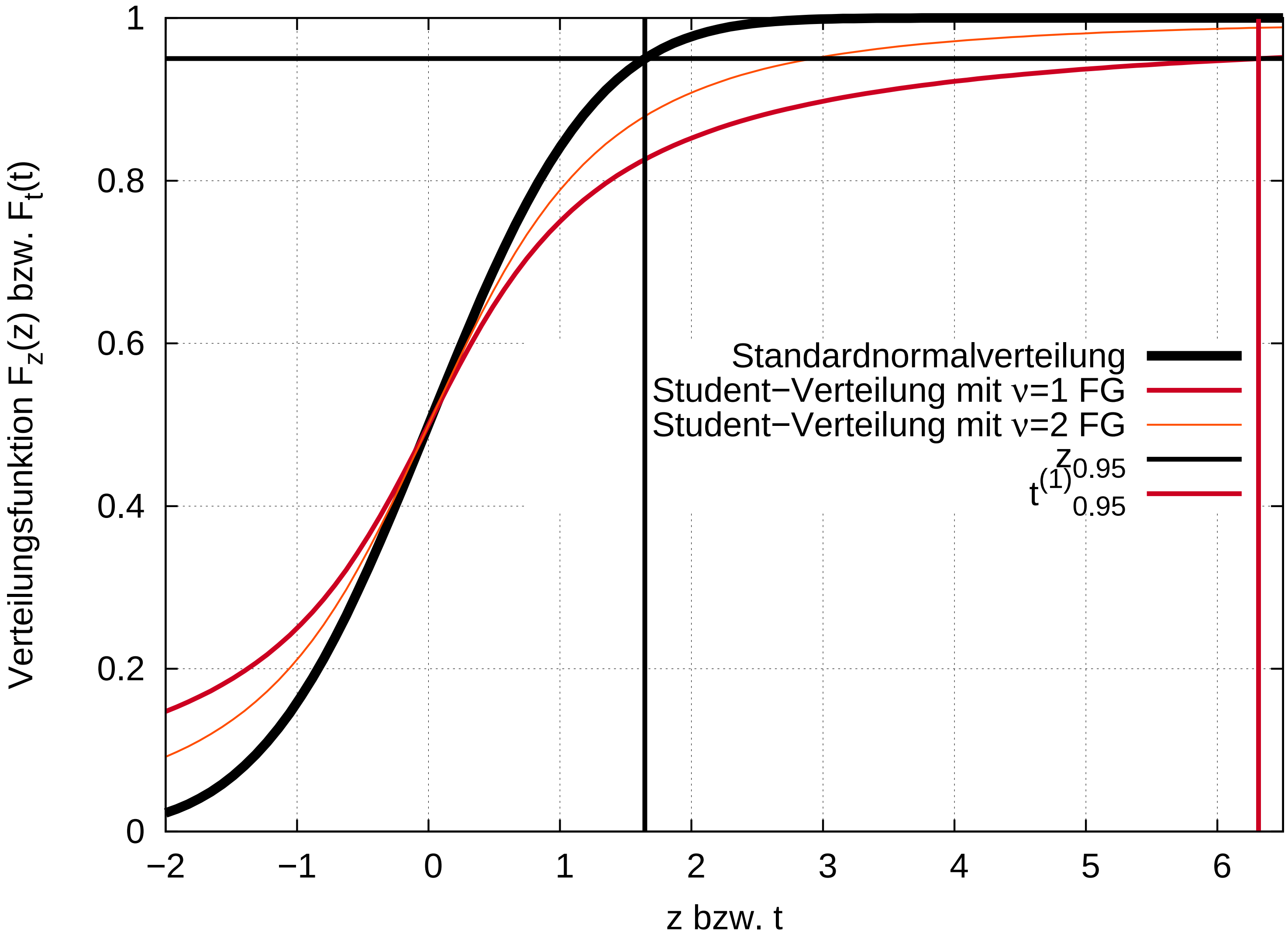
Konfidenzintervalle und die Entstehung der Student-Verteilung



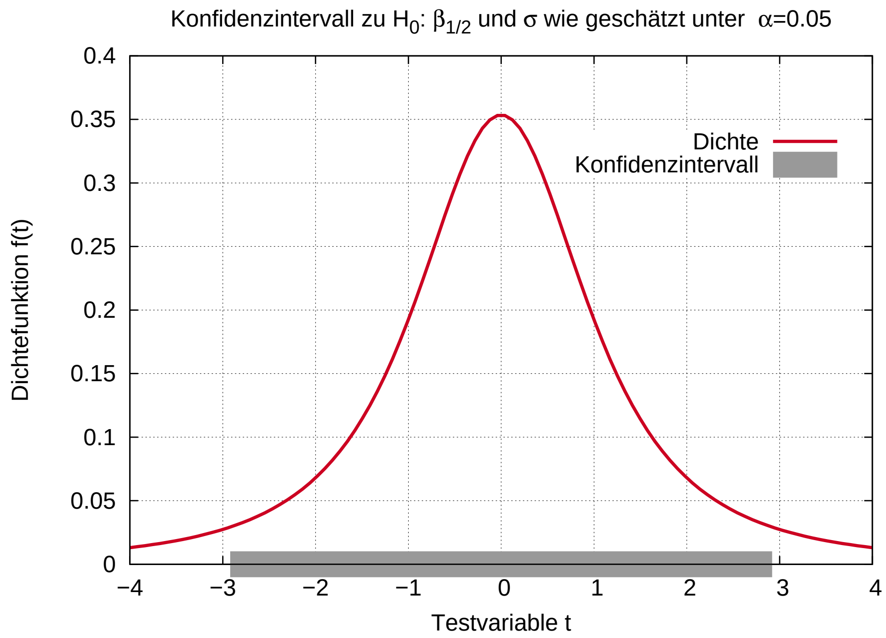
Dichten der Standardnormal *vs.* Student-t-verteilung

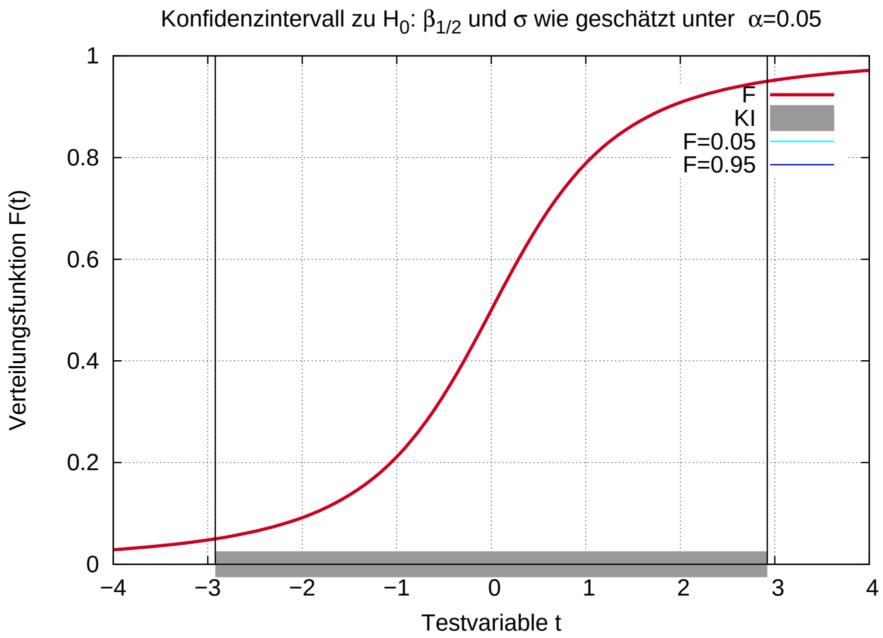


Standardnormal *vs.* Student-t-verteilung



Konfidenzintervalle

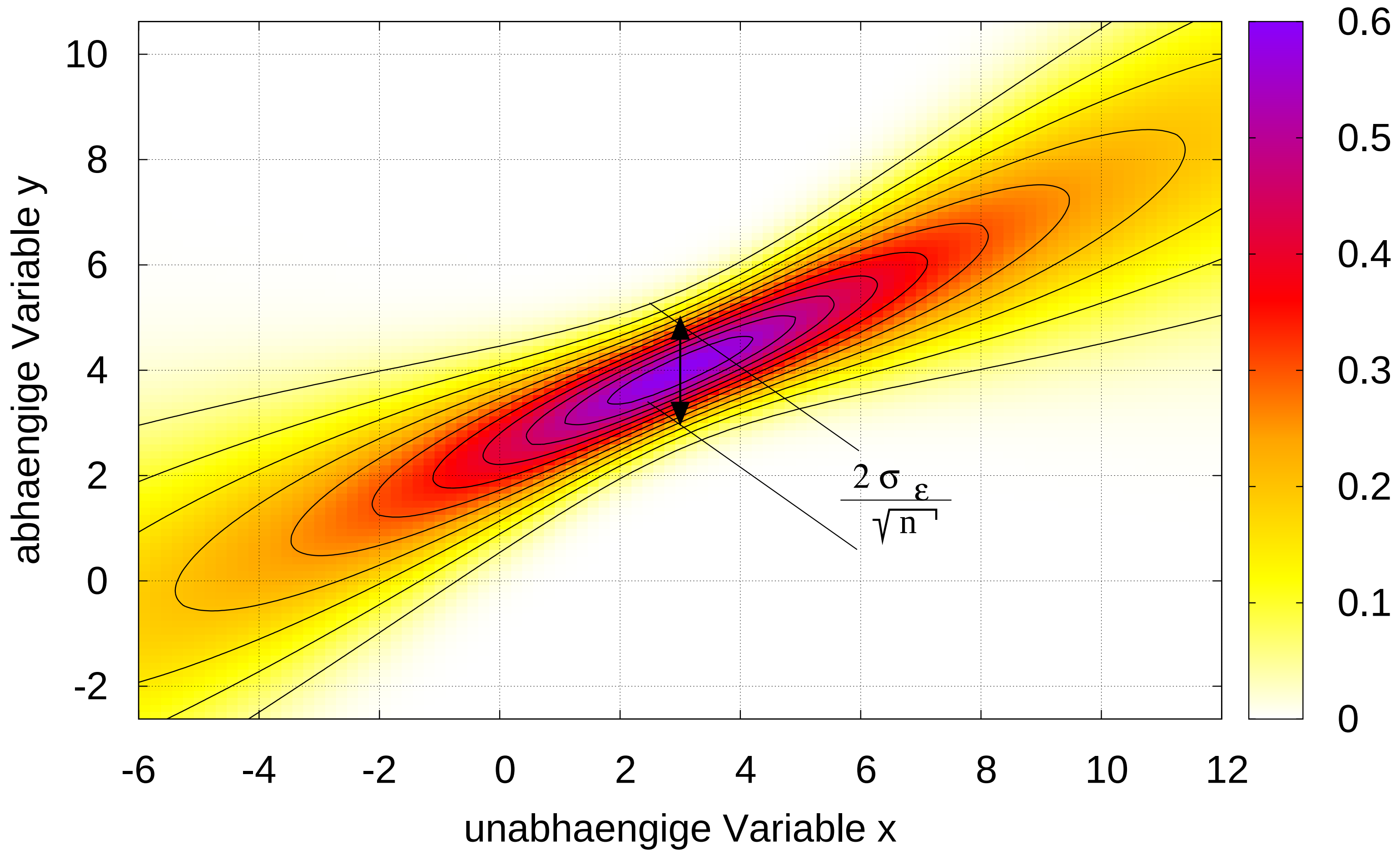




KI zu einer Fehlerwahrscheinlichkeit $\alpha = 5\,\%$ für $n - J - 1 = 2$ Freiheitsgrade.

Parameter-Schätzfehler (bedingte W-Dichte) bei linearer Einfachregression

$$\hat{y}(x) = \bar{y} + \beta_1(x - \bar{x}), n = 20, \bar{x} = 3, \bar{y} = 4, \beta_1 = 0.5$$

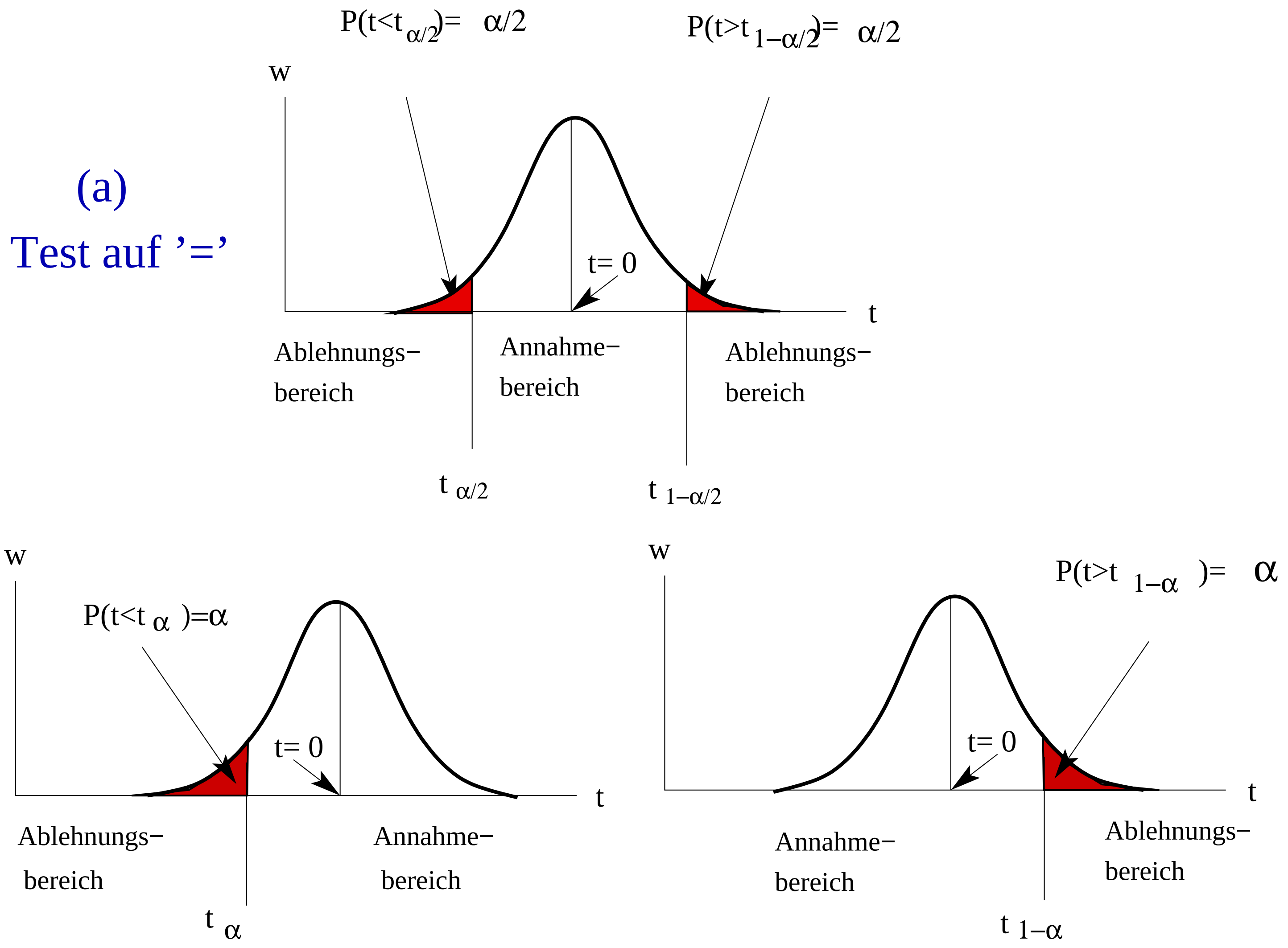


Die Residualfehler sind i.i.d. $N(0, 3^2)$ verteilt

Definition der Fehler erster und 2. Art bei Signifikanztests

	H_0 nicht abgelehnt	H_0 abgelehnt
H_0 trifft zu	✓	Fehler erster Art
H_0 trifft nicht zu	Fehler zweiter Art	✓

Annahme und Ablehnungsbereiche

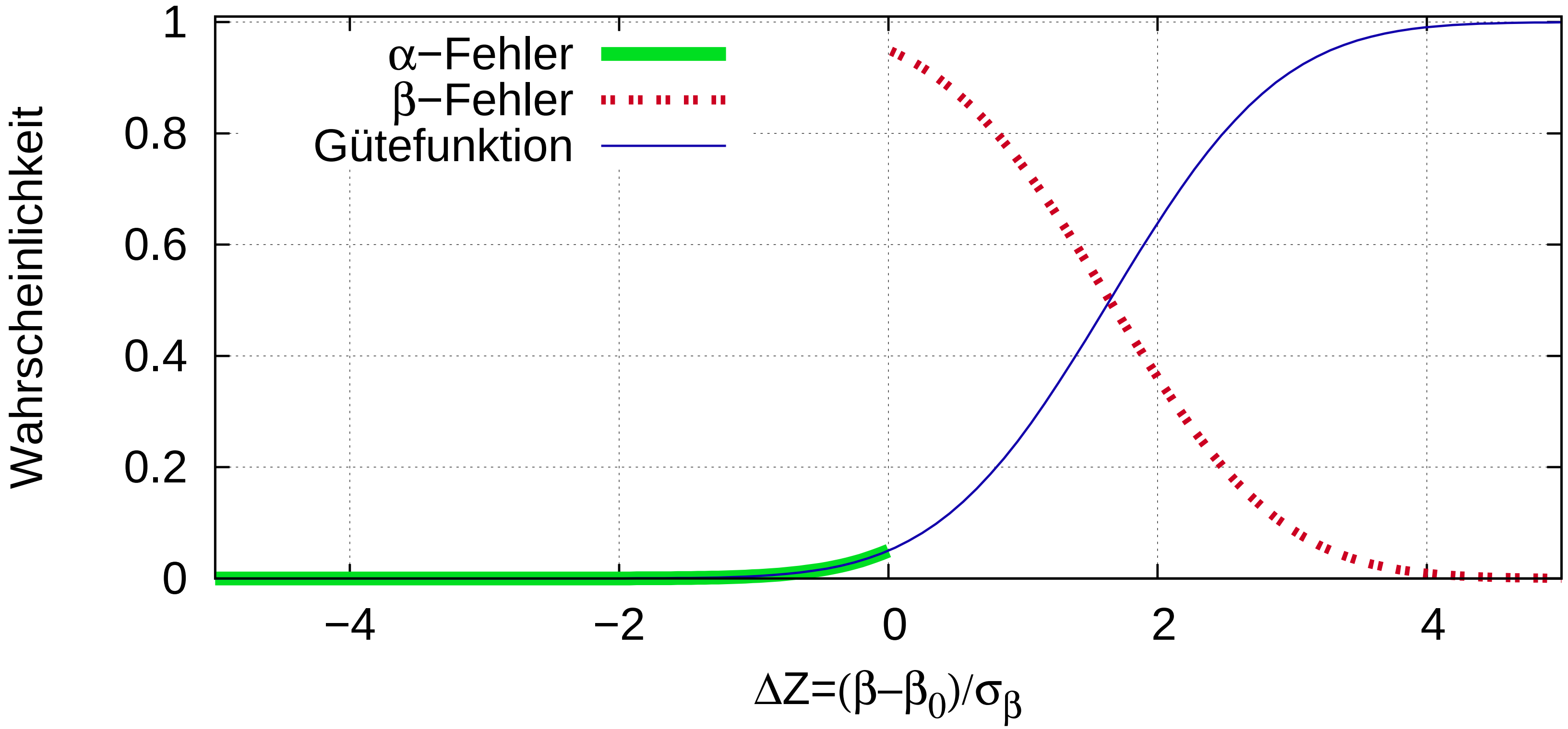


(b) Test auf ’>’ oder ’>=’

(c) Test auf ’<’ oder ’<=’

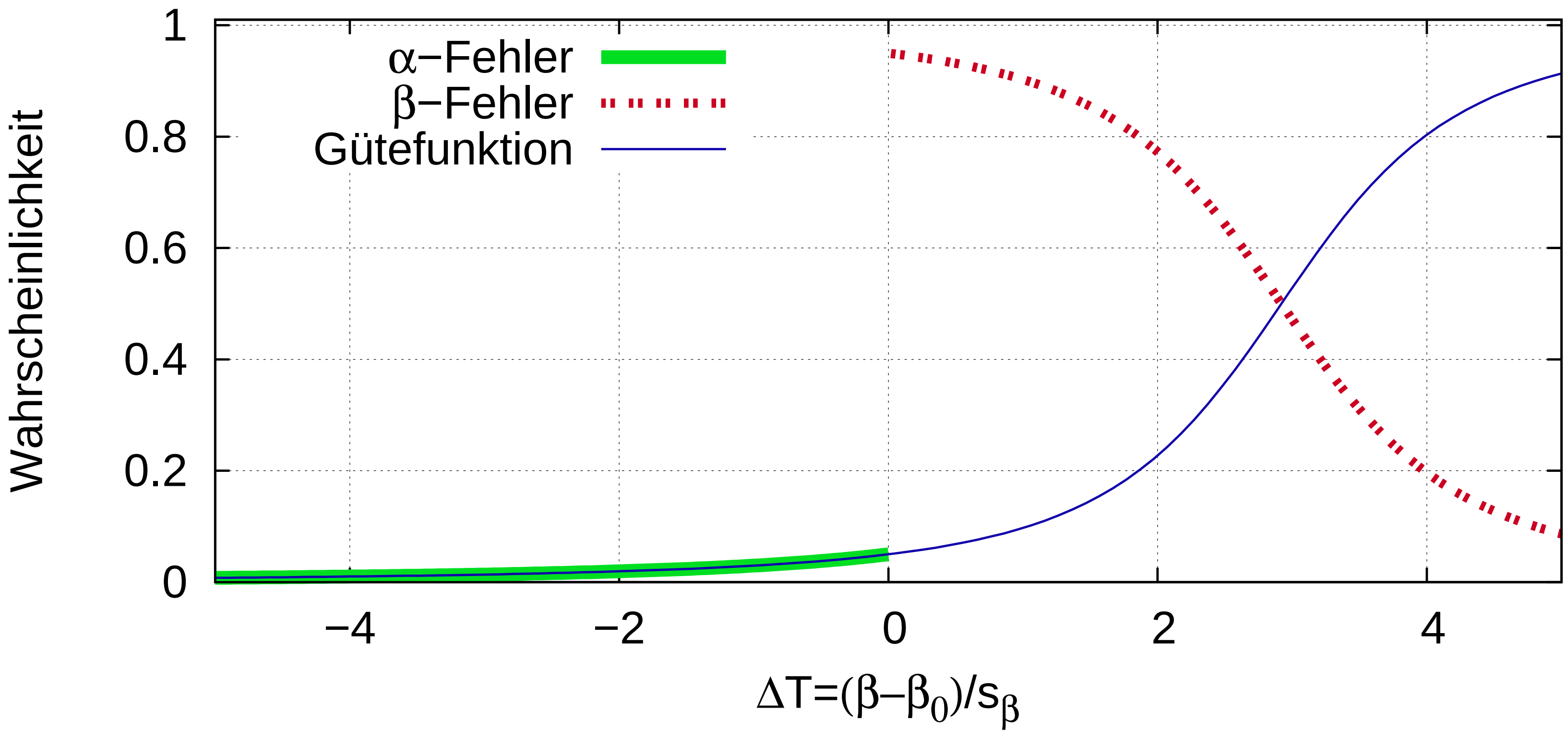
Die schwarze Kurve ist nur dann die Dichte der Testfunktion, wenn H_0 <u>grenzwertig</u> zutrifft!

Fehler erster und 2. Art bei “$<$” und “$\leq$”-Tests, bekannte Varianz



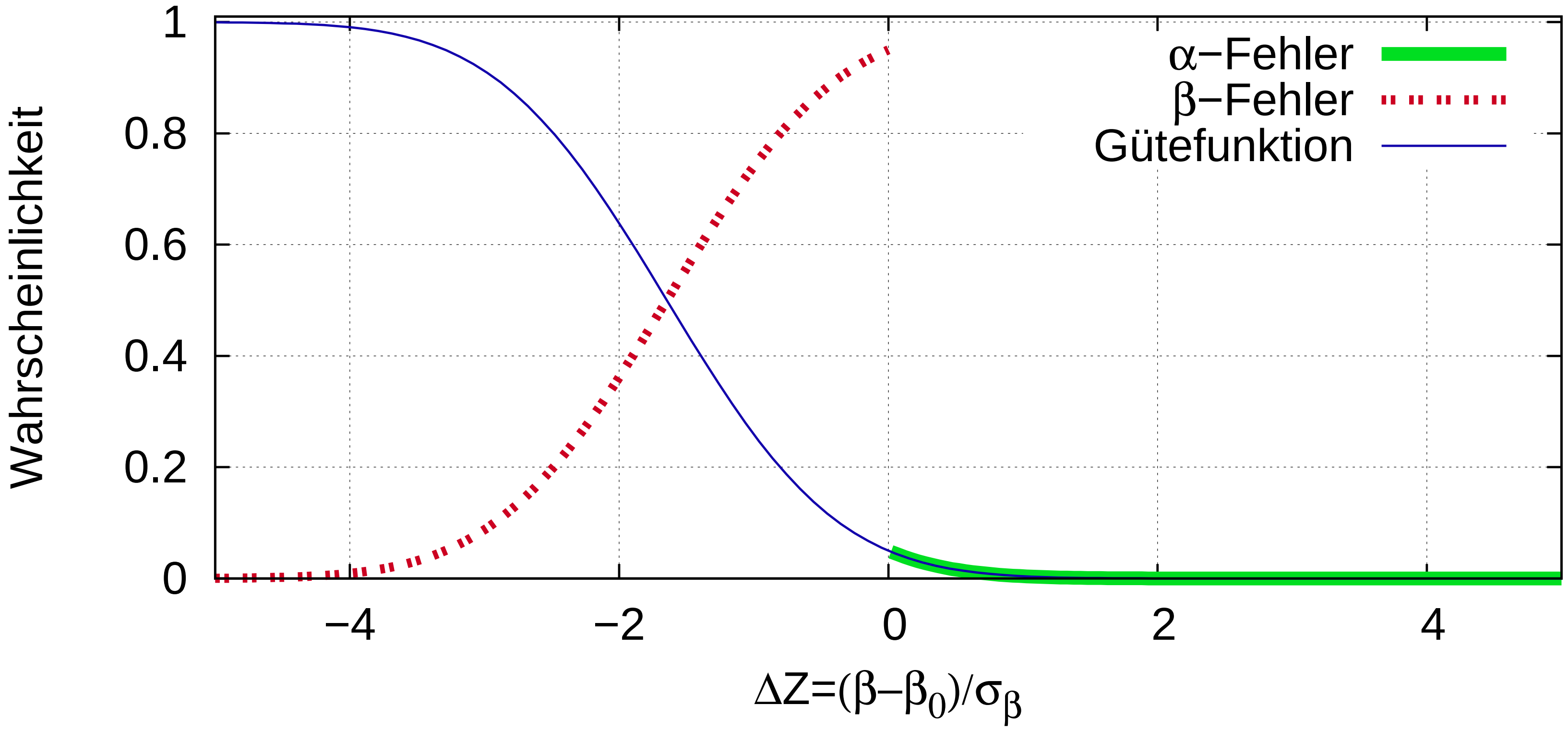
Fehlerwahrscheinlichkeit für Fehler 1. und 2. Art (α- bzw. β-Fehler) bei einem Signifikanzniveau $\alpha = 5\,\%$ in Abhängigkeit des wahren Abstandes $(\beta - \beta_0)/\sigma$ des Parameters vom H_0-Grenzwert in Einheiten der wahren Standardabweichung des Schätzers

Fehler erster und 2. Art bei "$<$" und "$\leq$"-Tests, unbekannte Varianz, $n - J - 1 = 2$ Freiheitsgrade



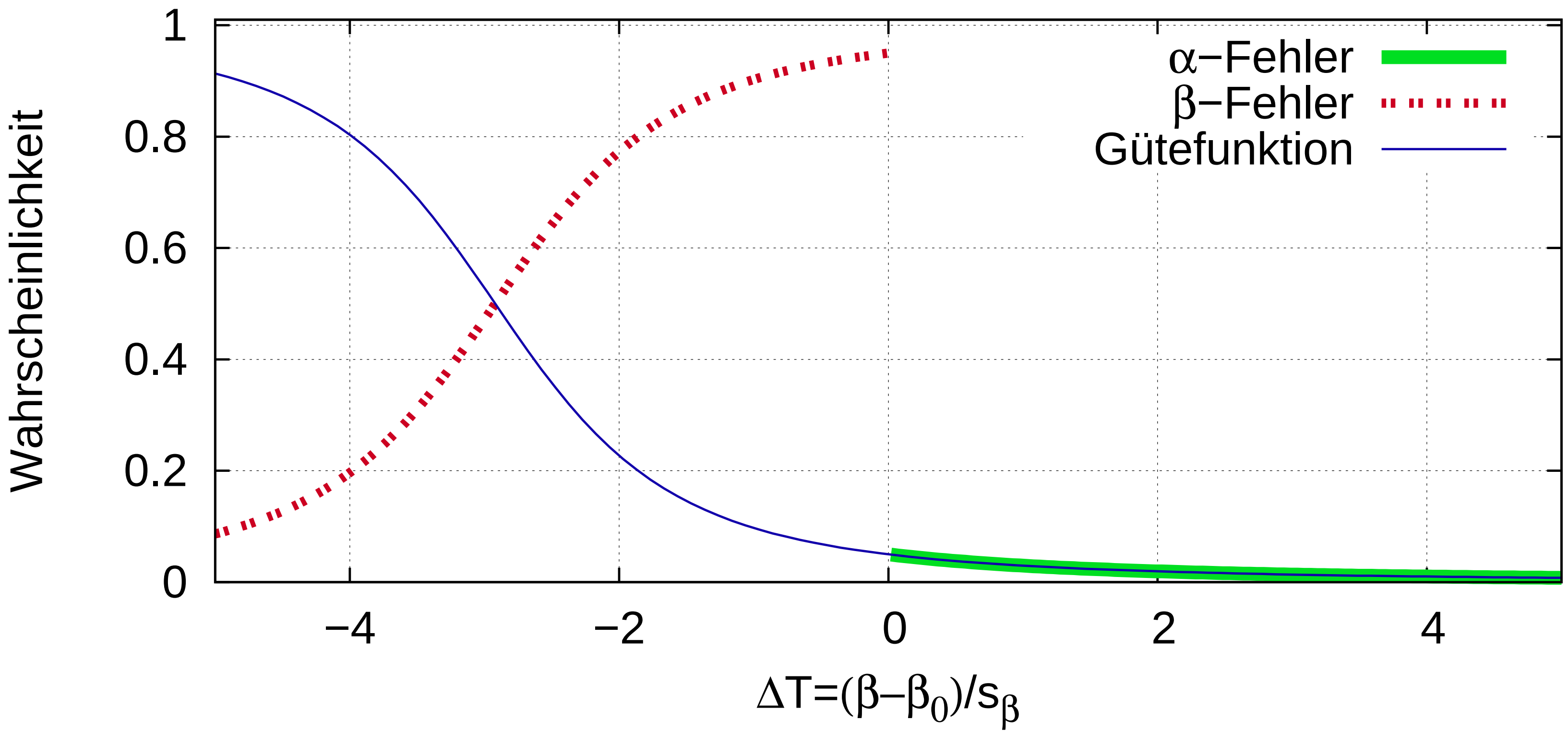
Der skalierte Abstand des wahren β-Wertes von der Grenze der Nullhypothese ist immer noch in Einheiten der wahren Standardabweichung gegeben

Fehler erster und 2. Art bei “$>$” und “$\geq$”-Tests, bekannte Varianz

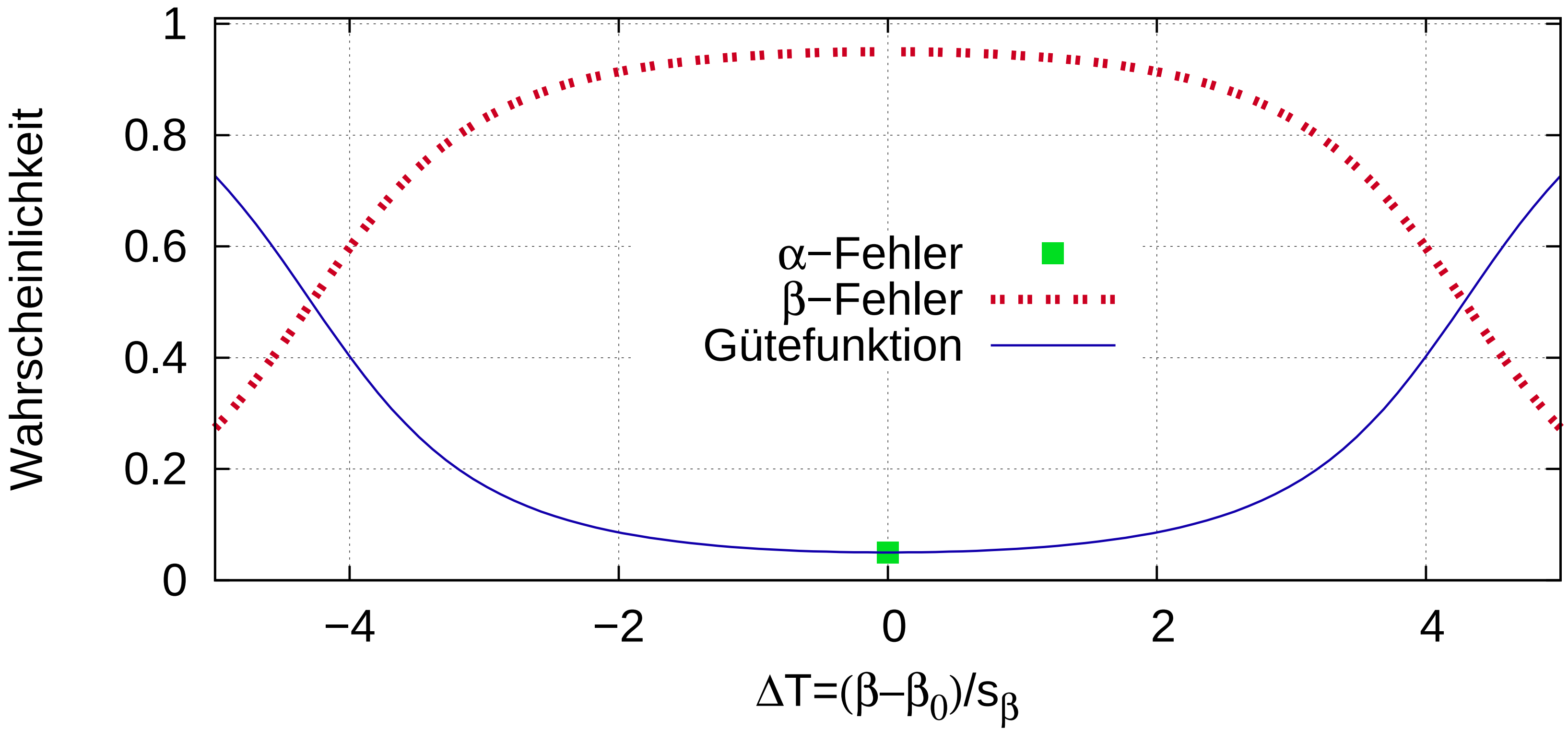


Wieder ist das Signifikanzniveau=maximale Wahrscheinlichkeit eines Fehlers 1. Art durch $\alpha = 5\,\%$ gegeben

Fehler erster und 2. Art bei “$>$” und “$\geq$”-Tests, unbekannte Varianz, $n - J - 1 = 2$ Freiheitsgrade

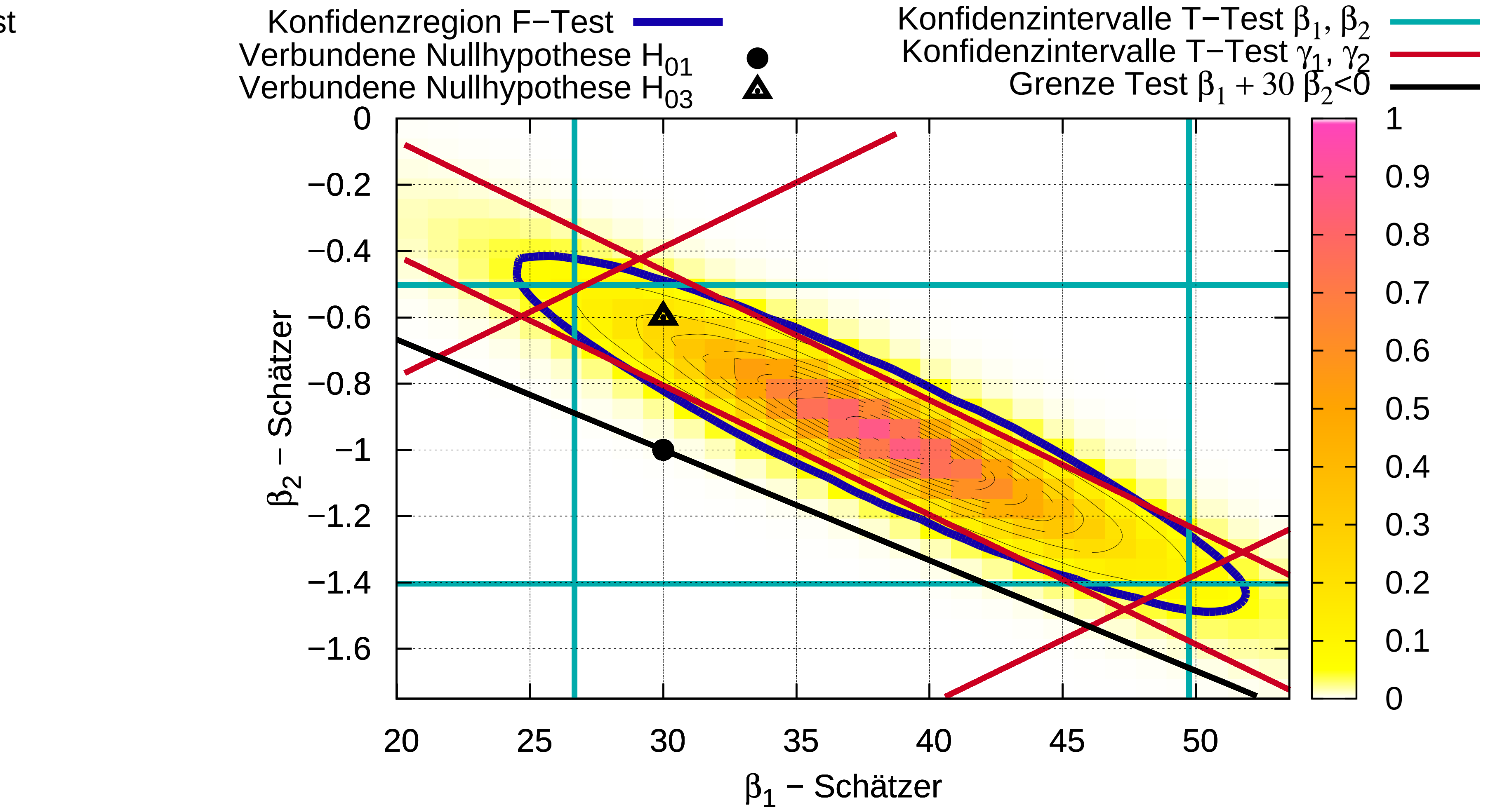


Fehler erster und 2. Art bei zweiseitigen (Punkt-)Tests, unbekannte Varianz, $n - J - 1 = 2$ Freiheitsgrade



Signifikanzniveau=maximale Wahrscheinlichkeit $\alpha = 5\,\%$

Hotelbeispiel: Verbundene Tests und 2d Konfidenzregionen



$\Rightarrow$ Nichtverschwindende Korrelation $V_{12}(\hat{\boldsymbol{\beta}}) = -0.926$ der Schätzfehler!

Einfache Nullhypothesen (t-Test):

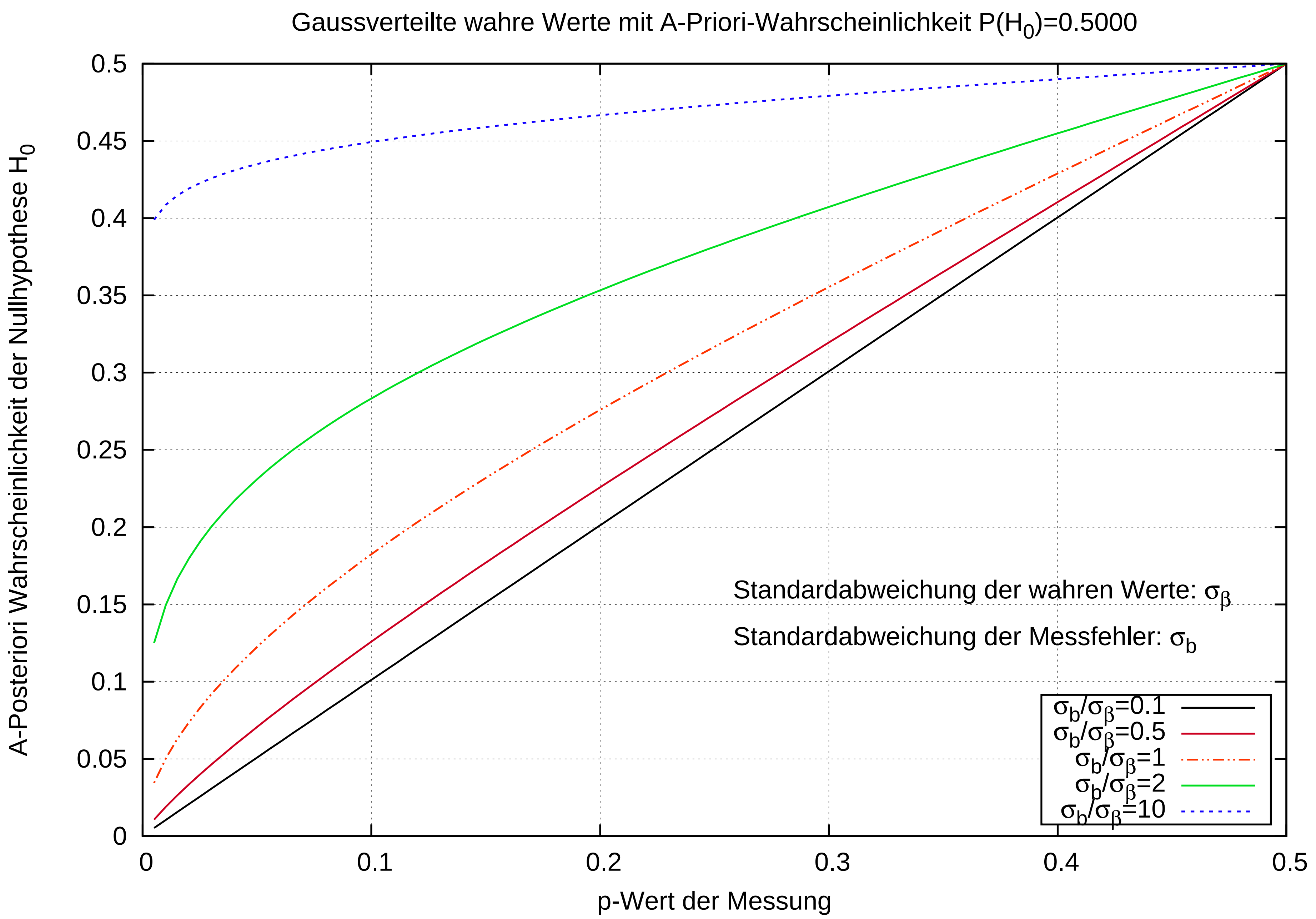
$H_{02} : \beta_1 = \beta_{10} = 27$
$H_{04} : \beta_2 = \beta_{20} = -1.4$ usw.

Verbundene Nullhypothesen (F-Test):

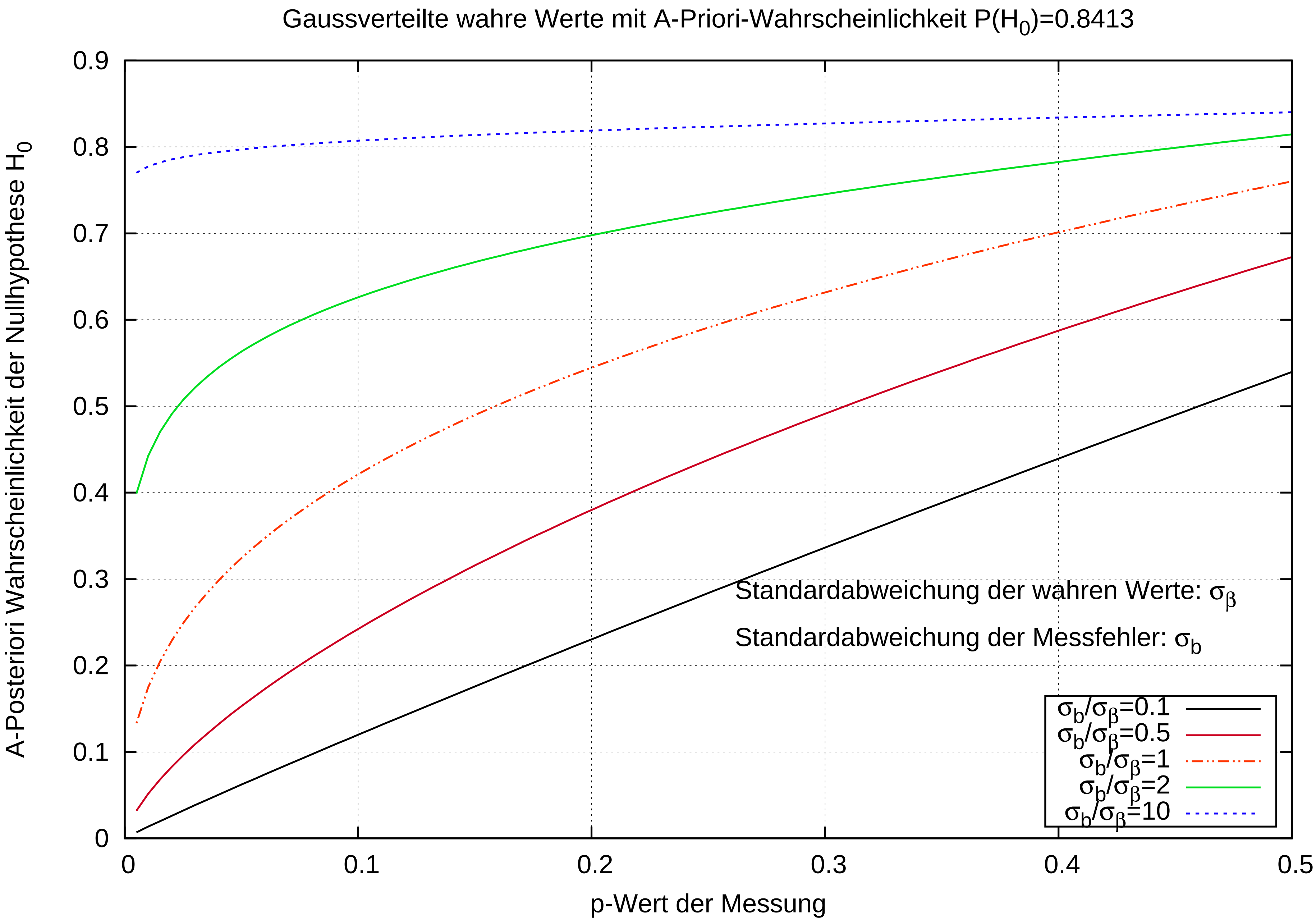
$\bullet = H_{01} : \beta_{10} = 30$ und $\beta_{20} = -1$
$\triangle = H_{03} : \beta_{10} = 34$ und $\beta_{20} = -1$

$\Rightarrow$ extra Folien zum Hotelbeispiel!

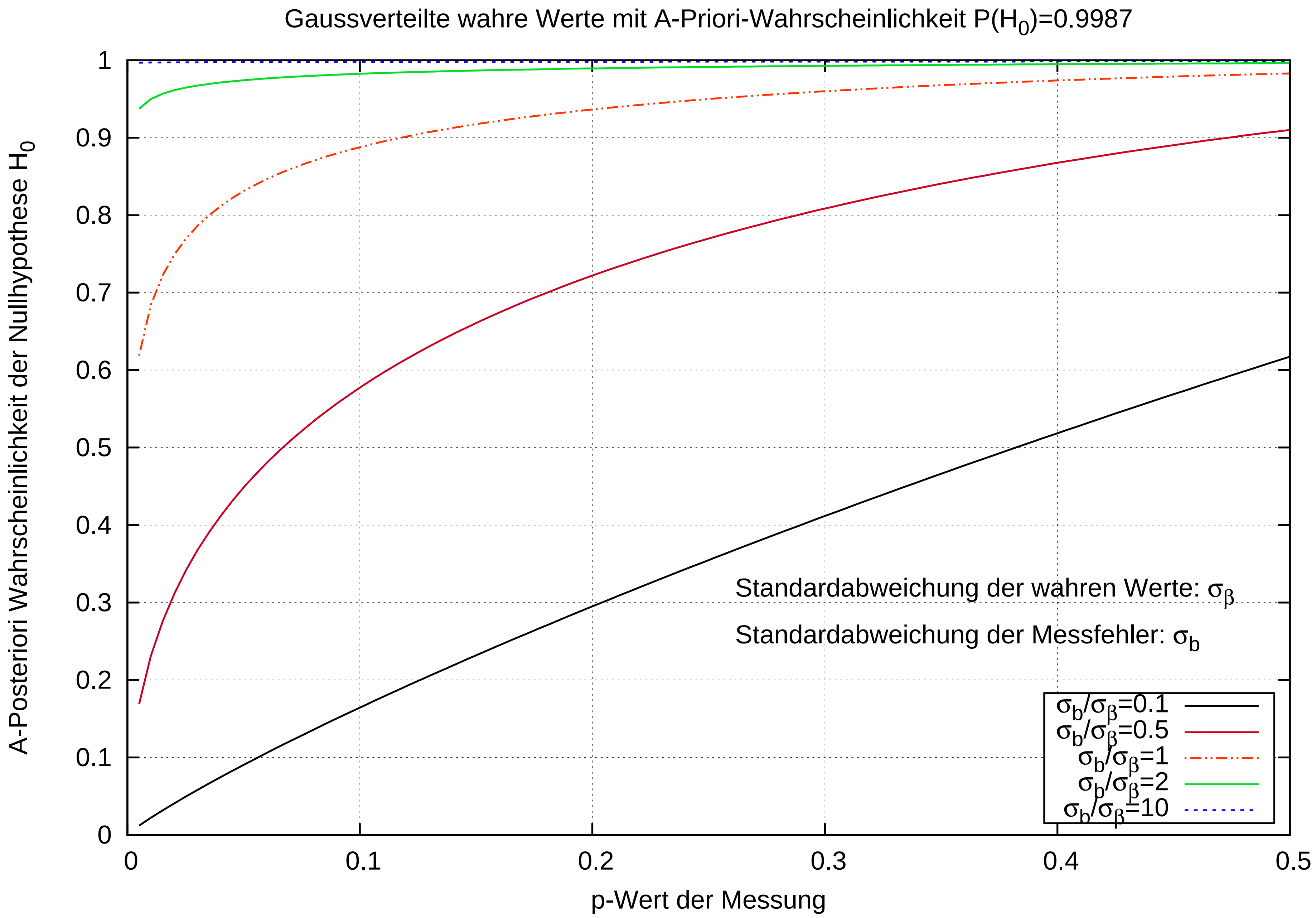
Bayes'sche Statistik bei normalverteilter A-Priori-Verteilung



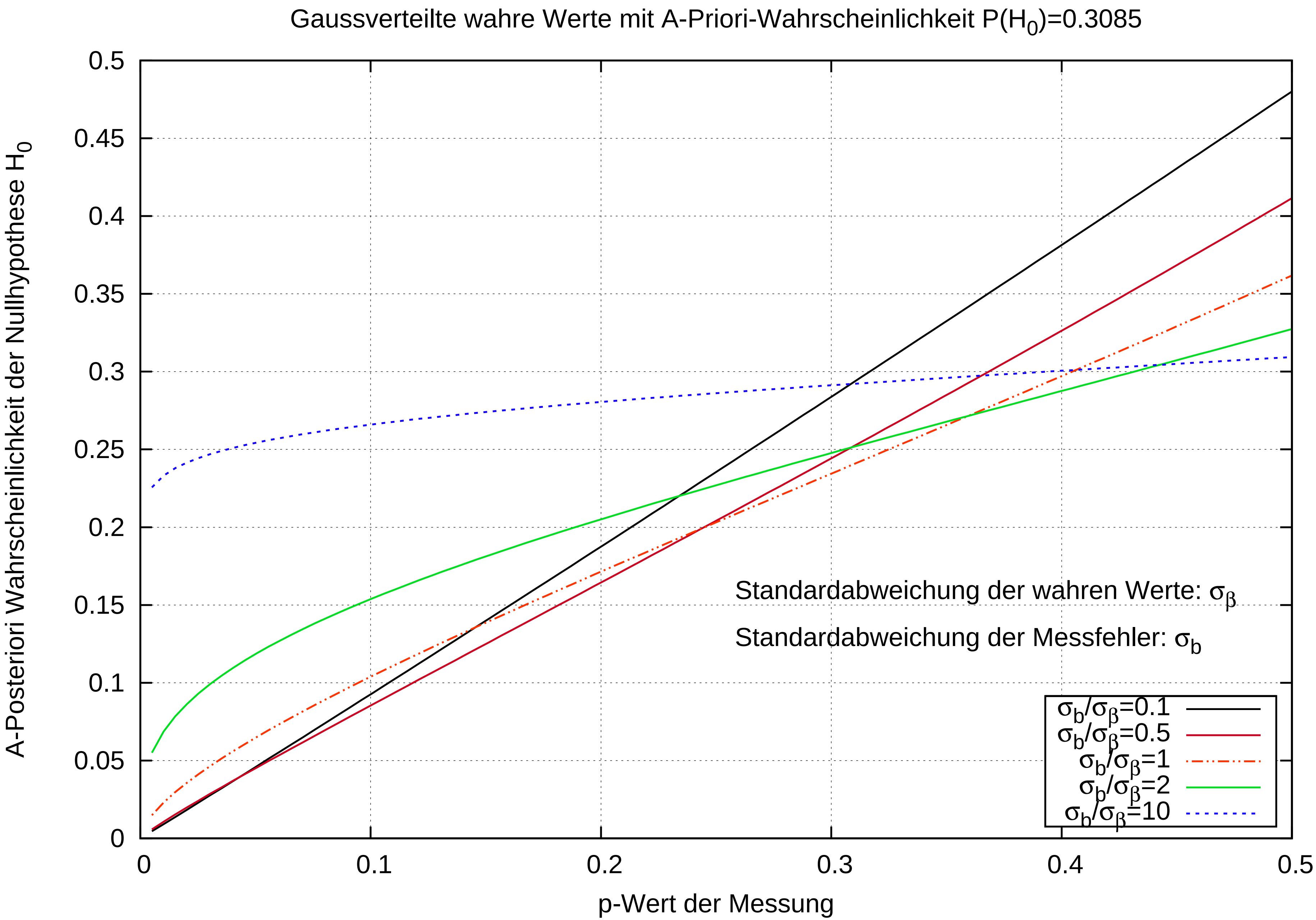
Bayes'sche Statistik bei normalverteilter A-Priori-Verteilung

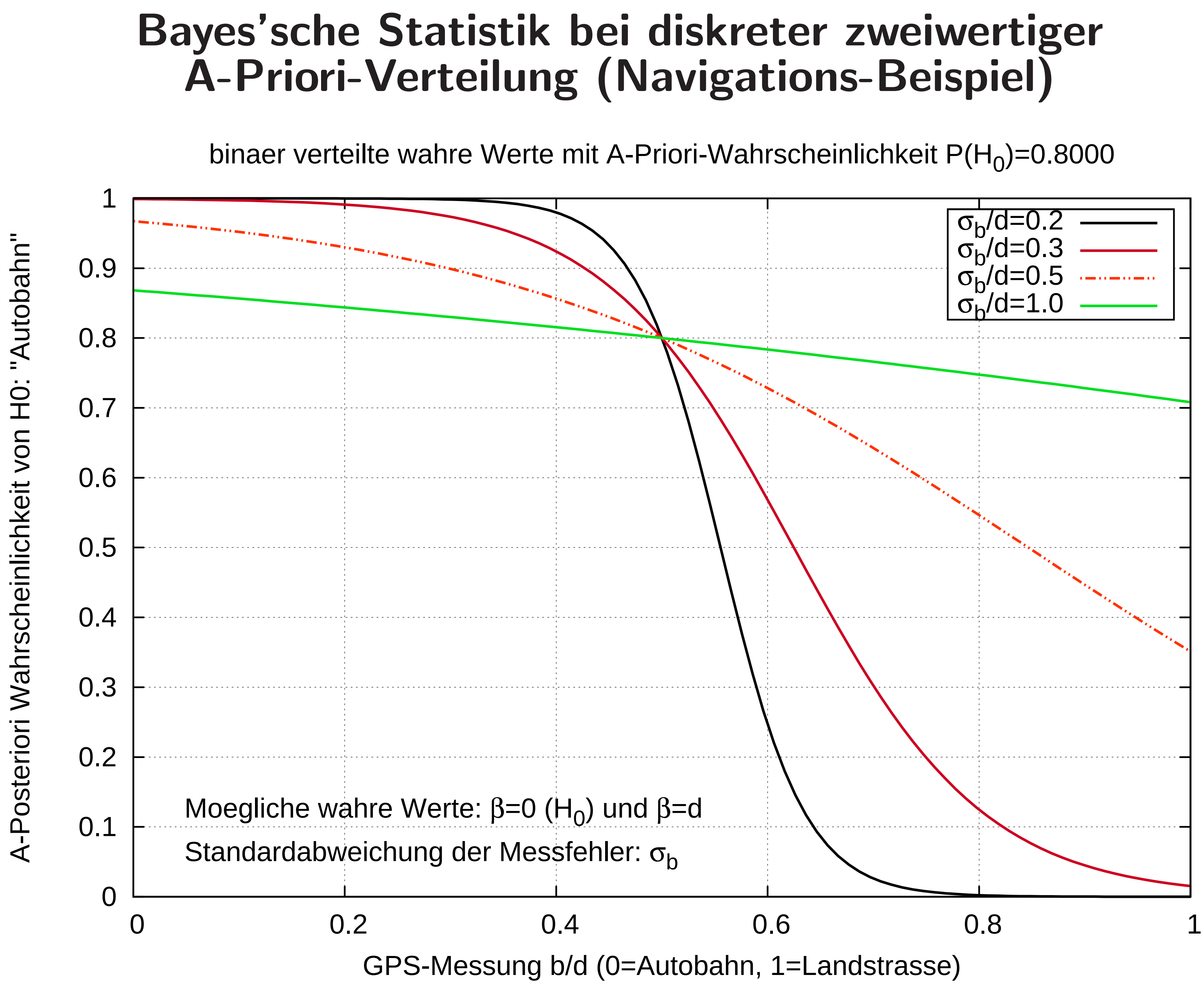


Bayes'sche Statistik bei normalverteilter A-Priori-Verteilung



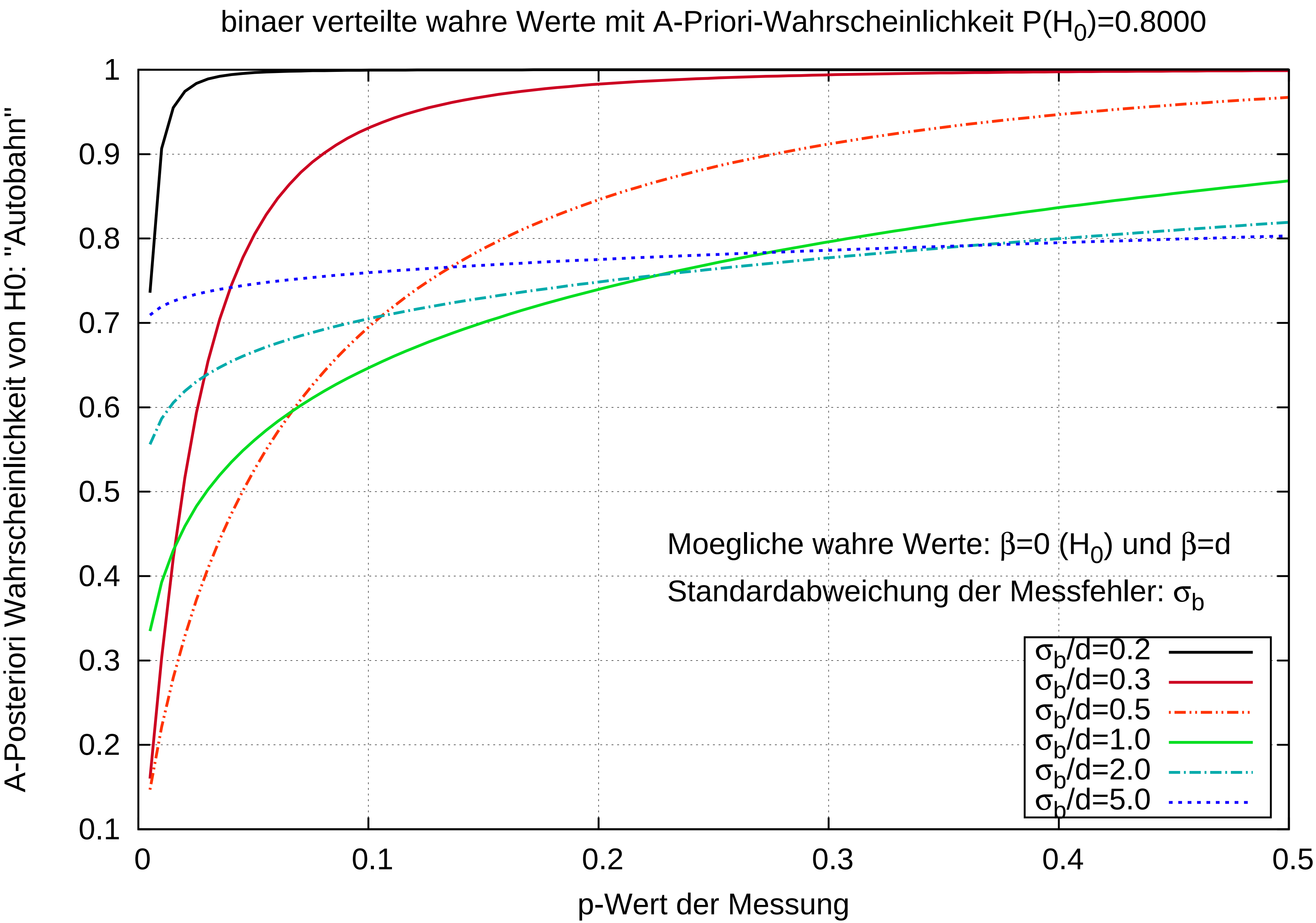
Bayes'sche Statistik bei normalverteilter A-Priori-Verteilung





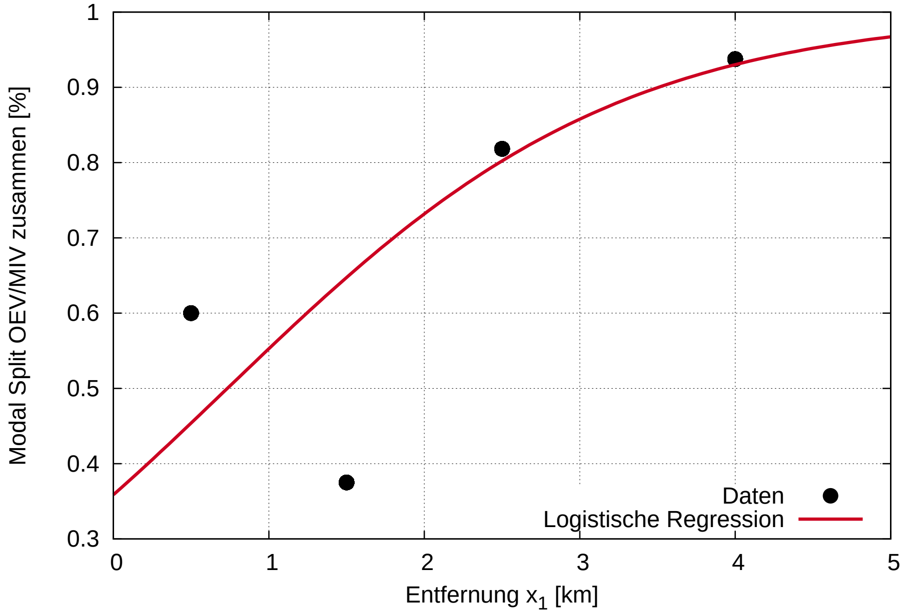
Bayes'sche Statistik bei diskreter zweiwertiger
A-Priori-Verteilung (Navigations-Beispiel)
binaer verteilte wahre Werte mit A-Priori-Wahrscheinlichkeit $P(H_0)$=0.8000
A-Posteriori Wahrscheinlichkeit von H0: "Autobahn"
GPS-Messung b/d (0=Autobahn, 1=Landstrasse)
σ_b/d=0.2
σ_b/d=0.3
σ_b/d=0.5
σ_b/d=1.0
Moegliche wahre Werte: β=0 (H_0) und β=d
Standardabweichung der Messfehler: σ_b
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1

Bayes'sche Statistik des Navigations-Beispiels, aber nun als Funktion des p-Wertes



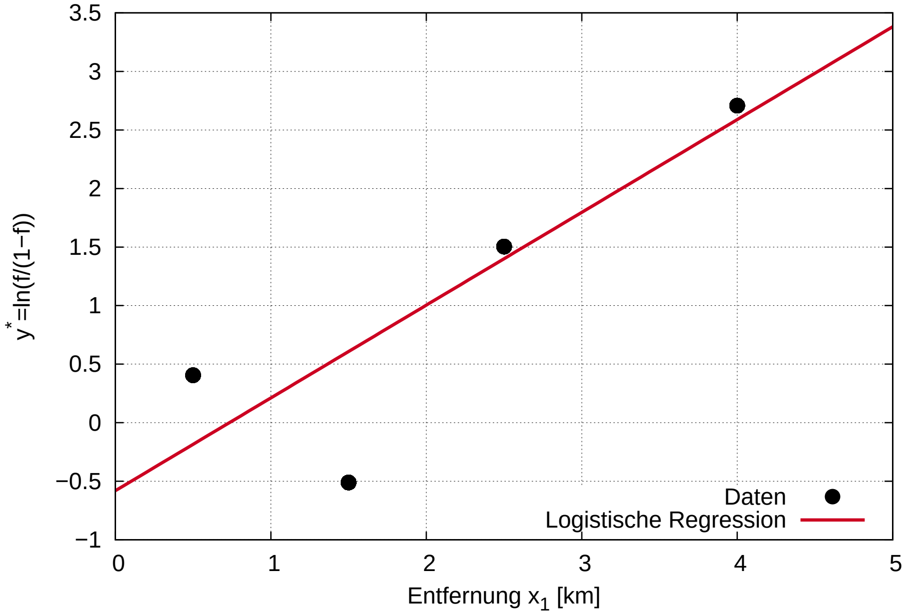
Logistische Regression
mit naiver LSE-Schätzung der log-Odd-Ratios:
RC-Umfrage WS14/15 und WS15/16 kumuliert

Daten und Ergebnis
mit 4 Entfernungsklassen



Unbeobachtete Variable

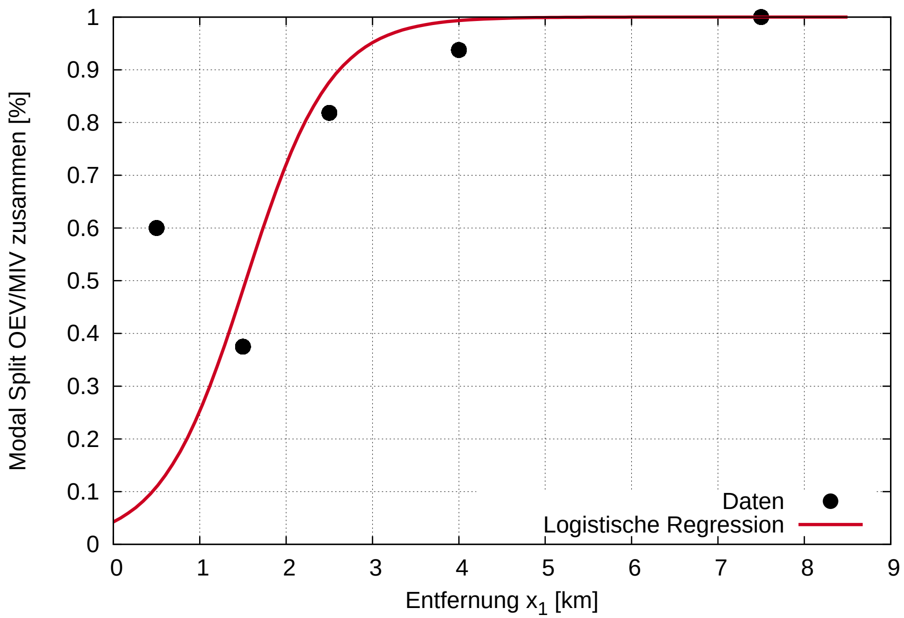
$$y^* = \ln(f_1/(1 - f_1))$$



$$\beta_0 = -0.58, \quad \beta_1 = 0.79$$

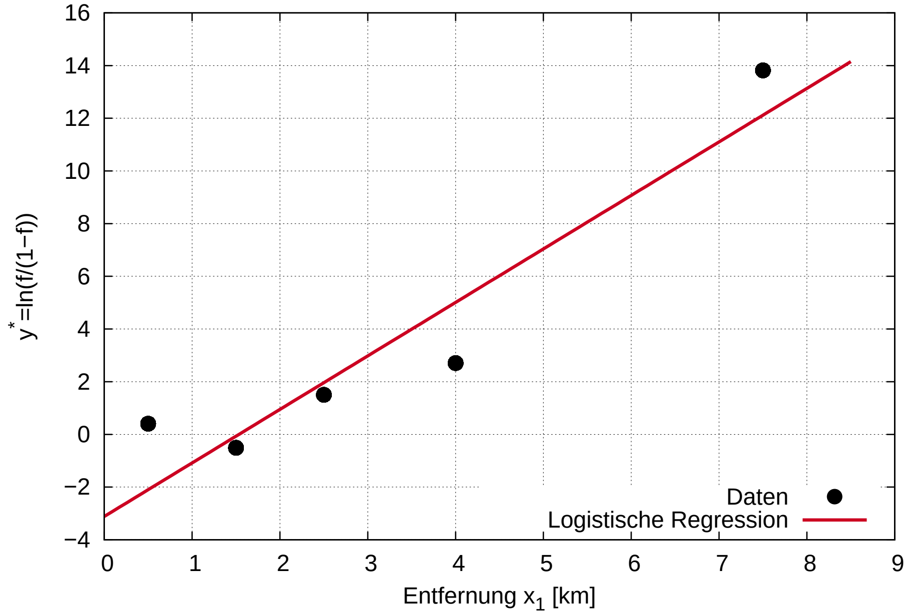
Logistische Regression mit naiver LSE-Schätzung der log-Odd-Ratios: 5. Datenpunkt addiert mit f=0.9999

Daten und Ergebnis
mit 4 Entfernungsklassen



Unbeobachtete Variable

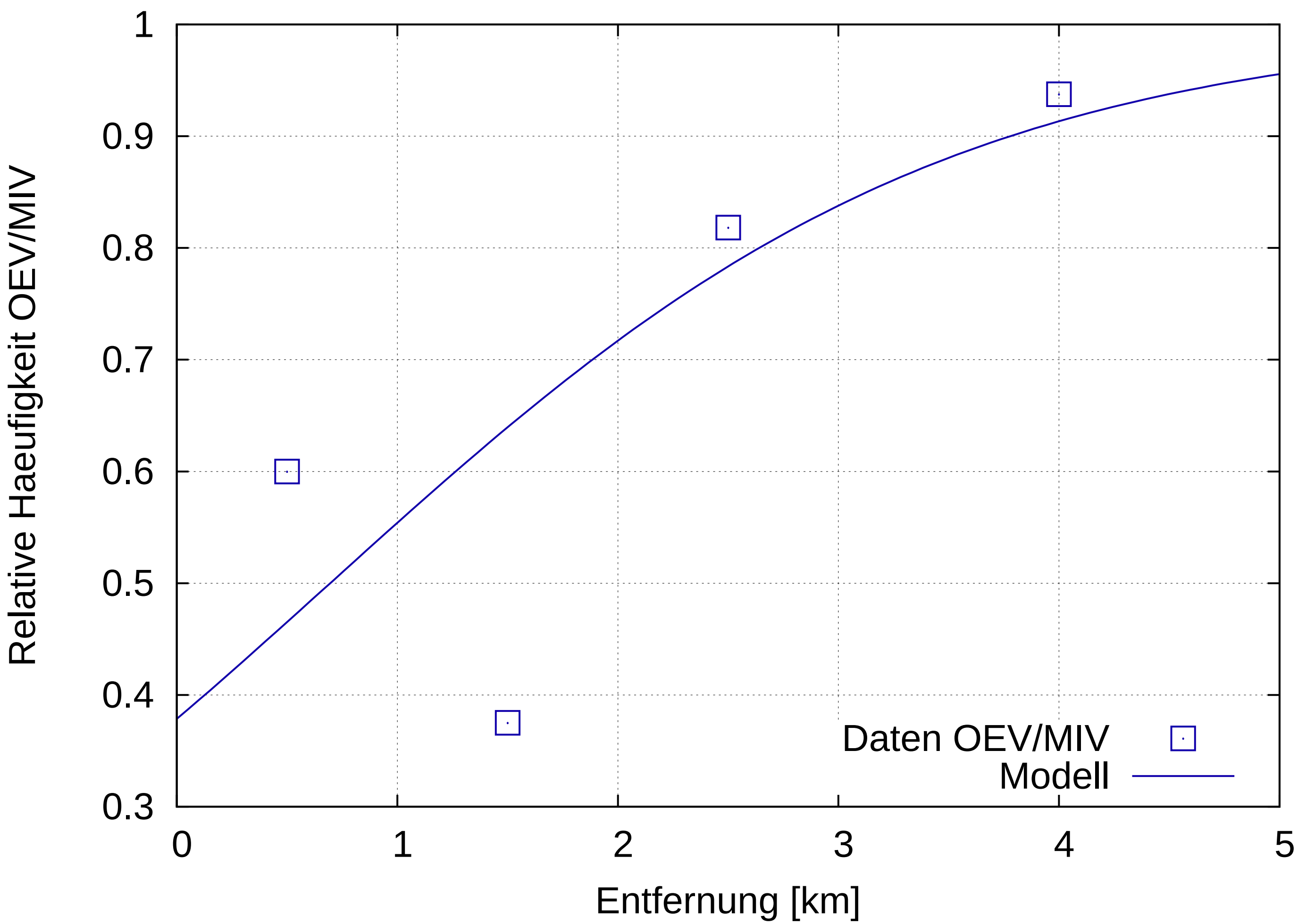
$$y^* = \ln(f_1/(1 - f_1))$$



$$\beta_0 = -3.12, \quad \beta_1 = 2.03$$

Vergleich: “echte” Maximum-Likelihood-Schätzung

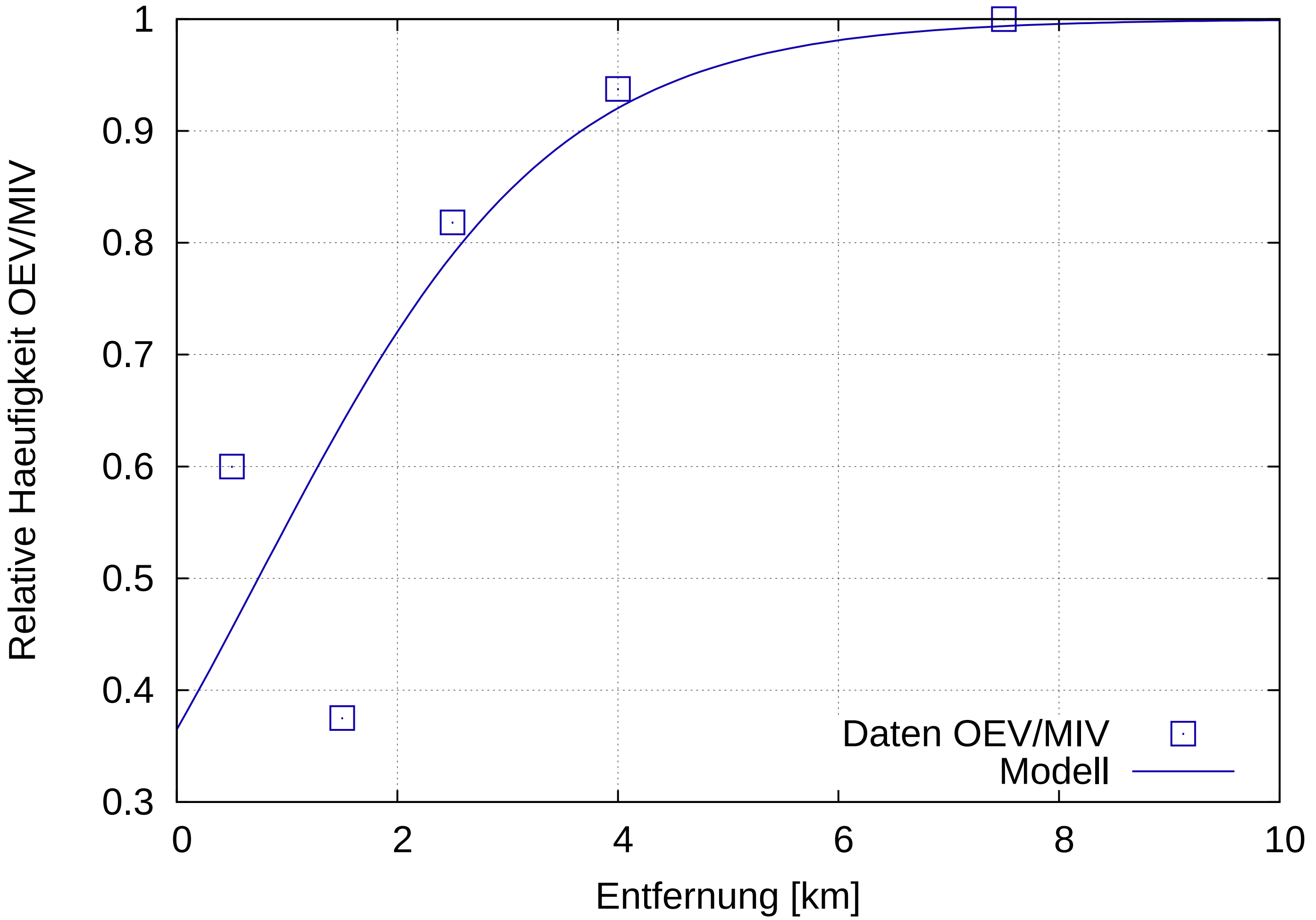
Alternativen 0 (kein ÖV) und 1 (ÖV) $\quad V_i(r) = \beta_0 \delta_{i1} + \beta_1 r \delta_{i1}$



$$\beta_0 = -0.50 \pm 0.65,$$
$$\beta_1 = +0.71 \pm 0.30$$

Vergleich: "echte" Maximum-Likelihood-Schätzung mit 5. Datenpunkt

Alternativen 0 (kein ÖV) und 1 (ÖV) $V_i(r) = \beta_0 \delta_{i1} + \beta_1 r \delta_{i1}$



$$\beta_0 = -0.55 \pm 0.63,$$
$$\beta_1 = +0.75 \pm 0.27$$