

# Kapitel IX - Neyman-Pearson-Tests

## Induktive Statistik

Prof. Dr. W.-D. Heller

Hartwig Senska

Carlo Siebenschuh

Ausgangssituation: *Einfache Testsituation*

$$\Gamma = \{\gamma_0, \gamma_1\}, \Gamma_0 = \{\gamma_0\}, \Gamma_1 = \{\gamma_1\}, \gamma_0 \neq \gamma_1$$

Damit:

$$H_0 : \gamma = \gamma_0, \quad H_1 : \gamma = \gamma_1.$$

Unterscheidung bei zwei verschiedenen radioaktiven Elementen: Anzahl der Zufallsprozesse im Zeitraum  $\Delta t$ :

Poissonverteilt mit Parameter  $\lambda$  hängt zusammen mit der Halbwertszeit des Elements.

Substanz  $A$  habe Parameter  $\lambda_0 = 1$

Substanz  $B$  habe Parameter  $\lambda_1 = 2$

## 5 Versuche:

Anzahl der Zufallsprozesse:  $X_1, X_2, X_3, X_4, X_5$

Zahlenbeispiel:  $x_1 = 0, x_2 = 2, x_3 = 1, x_4 = 0, x_5 = 1$

Wahrscheinlichkeit für dieses Ergebnis beim Parameter  $\lambda$ :

$$P_\lambda(X = x) = \prod_{j=1}^n \frac{\lambda^{x_j}}{x_j!} e^{-\lambda} = \frac{\lambda^{\sum_{j=1}^n x_j} e^{-n\lambda}}{x_1! x_2! x_3! x_4! x_5!}$$

$$\text{Für } \lambda_0 = 1 : \quad = \frac{e^{-5}}{2} = 0.003$$

$$\text{Für } \lambda_1 = 2 : \quad = \frac{2^4 e^{-10}}{2} = 8e^{-10} = 0.0004$$

Also

$$P_{\lambda_0}(X = x) > P_{\lambda_1}(X = x)$$

und damit ist Substanz  $A$  eine sinnvolle Entscheidung.

Bei  $H_0 : \lambda = \lambda_0$  gegen  $H_1 : \lambda = \lambda_1$  entspricht diese Vorgehensweise der Entscheidungsfunktion:

$$\delta(x) = \begin{cases} d_0 & P_{\lambda_0}(X = x) > P_{\lambda_1}(X = x) \\ d_1 & \text{sonst.} \end{cases}$$

Das Entscheidungskriterium für  $d_0$  lässt sich umformen:

$$\begin{aligned} P_{\lambda_0}(X = x) > P_{\lambda_1}(X = x) &\Leftrightarrow \frac{\lambda_0^{\sum_{j=1}^n x_j} e^{-n\lambda_0}}{x_1!x_2!x_3!x_4!x_5!} > \frac{\lambda_1^{\sum_{j=1}^n x_j} e^{-n\lambda_1}}{x_1!x_2!x_3!x_4!x_5!} \\ &\Leftrightarrow \lambda_0^{\sum_{j=1}^n x_j} e^{-n\lambda_0} > \lambda_1^{\sum_{j=1}^n x_j} e^{-n\lambda_1} \\ &\Leftrightarrow \frac{e^{n\lambda_1}}{e^{n\lambda_0}} > \left(\frac{\lambda_1}{\lambda_0}\right)^{\sum_{j=1}^n x_j} \\ &\Leftrightarrow n(\lambda_1 - \lambda_0) > \sum_{j=1}^n x_j (\ln \lambda_1 - \ln \lambda_0) \\ &\Leftrightarrow \sum_{j=1}^n x_j < \frac{n(\lambda_1 - \lambda_0)}{\ln \lambda_1 - \ln \lambda_0} \end{aligned}$$

$$\sum_{j=1}^n x_j < \frac{n(\lambda_1 - \lambda_0)}{\ln(\lambda_1) - \ln(\lambda_0)}$$

Konkret für  $n = 5$ ,  $\lambda_0 = 1$  und  $\lambda_1 = 2$ :

$$P_{\lambda_0}(X = x) > P_{\lambda_1}(X = x) \Leftrightarrow \sum_{j=1}^5 x_j < \frac{5}{\ln(2)} = 7.21$$



Damit lautet die Entscheidungsfunktion in anderer Formulierung:

$$\delta(x) = \begin{cases} d_0 & \sum_{j=1}^5 x_j < 7.21 \text{ (bzw. } \leq 7) \\ d_1 & \text{sonst} \end{cases}$$

Damit gilt für die Fehlerwahrscheinlichkeit 1. Art:

$$\begin{aligned} P_{\lambda_0}(\delta(X) = d_1) &= P_{\lambda_0} \left( \sum_{j=1}^n X_j \geq 7.21 \right) = \sum_{k=8}^{\infty} \frac{(n\lambda_0)^k}{k!} e^{-n\lambda_0} \\ &= \sum_{k=8}^{\infty} \frac{5^k}{k!} e^{-5} = 1 - \sum_{k=0}^7 \frac{5^k}{k!} e^{-5} = 0.133 \end{aligned}$$

Der Fehler 2. Art ( $H_0$  wird fälschlicherweise angenommen, wenn bei  $\lambda_1$  Entscheidung  $d_0$  getroffen wird, also  $\sum_{j=1}^5 x_j \leq 7$  ist.)

Damit gilt für die Fehlerwahrscheinlichkeit 2. Art:

$$\begin{aligned} P_{\lambda_1}(\delta(X) = d_0) &= P_{\lambda_1}\left(\sum_{j=1}^n X_j < 7.21\right) = \sum_{k=0}^7 \frac{(5\lambda_1)^k}{k!} e^{-5\lambda_1} \\ &= \sum_{k=0}^7 \frac{10^k}{k!} e^{-10} = 0.220 \end{aligned}$$

Für die vier Konstellationen für Parameter und Entscheidung gilt also:

$P.(\cdot)$	$\lambda = \lambda_0$	$\lambda = \lambda_1$
$d = d_0$	0.867	0.220
$d = d_1$	0.133	0.780

Wahrscheinlichkeiten für die Entscheidungen  $d_0$  und  $d_1$

Verkleinerung der Fehlerwahrscheinlichkeit 1. Art durch Lockerung des Kriteriums für  $d_0$ :

$$P_{\lambda_0}(X = x) > P_{\lambda_1}(X = x)$$

Neues Kriterium für  $d_0$  mit  $k > 0$ :

$$k \cdot P_{\lambda_0}(X = x) \geq P_{\lambda_1}(X = x)$$

(Lockerung für  $k > 1$ , Verschärfung für  
 $0 < k < 1$ )

oder

$$\frac{P_{\lambda_1}(X=x)}{P_{\lambda_0}(X=x)} \leq k$$

D.h. Entscheidung  $d_1$  wird nur getroffen, wenn die Wahrscheinlichkeit für  $x$  bei  $\lambda_1$  um den Faktor  $k$  größer ist als die Wahrscheinlichkeit für  $x$  bei  $\lambda_0$ :

$$\frac{P_{\lambda_1}(X = x)}{P_{\lambda_0}(X = x)} > k \Leftrightarrow P_{\lambda_1}(X = x) > kP_{\lambda_0}(X = x)$$

Das Verhältnis

$$\frac{P_{\lambda_1}(X = x)}{P_{\lambda_0}(X = x)}$$

wird als **Wahrscheinlichkeitsquotient** bezeichnet.

Im Beispiel gilt:

$$\frac{P_{\lambda_1}(X = x)}{P_{\lambda_0}(X = x)} = \left(\frac{\lambda_1}{\lambda_0}\right)^{\sum_{j=1}^n x_j} e^{-n(\lambda_1 - \lambda_0)}$$

Damit ist  $(\lambda_0 < \lambda_1)$ :

$$\frac{P_{\lambda_1}(X = x)}{P_{\lambda_0}(X = x)} \leq k \Leftrightarrow \sum_{i=1}^n x_i \leq \frac{\ln k + n(\lambda_1 - \lambda_0)}{\ln \lambda_1 - \ln \lambda_0}$$

Bei  $n = 5$ ,  $\lambda_0 = 1$ ,  $\lambda_1 = 2$ :

$$\dots \Leftrightarrow \sum_{i=1}^n x_i \leq \frac{\ln k + 5}{\ln 2}$$

Damit lautet die Entscheidungsfunktion:

$$\delta_k(x_1, \dots, x_n) = \begin{cases} d_0 & \sum_{i=1}^n x_i \leq \frac{\ln k + 5}{\ln 2} \\ d_1 & \sum_{i=1}^n x_i > \frac{\ln k + 5}{\ln 2} \end{cases}$$

Da  $\sum_{i=1}^n x_i$  ganzzahlig ist, sind nur ganzzahlige Schranken für  $\sum_{i=1}^n x_i$  interessant:

$$\delta_c(x_1, \dots, x_n) = \begin{cases} d_0 & \sum_{i=1}^n x_i \leq c \\ d_1 & \sum_{i=1}^n x_i > c \end{cases}$$

mit  $c = 0, 1, 2, 3, \dots$



Für  $\delta_c$  erhalten wir die Fehlerwahrscheinlichkeit:

$$P_I(\delta_c, \lambda_0) = P_{\lambda_0}\left(\sum_{i=1}^n X_i > c\right) = \sum_{j=c+1}^{\infty} \frac{(n\lambda_0)^j}{j!} e^{-n\lambda_0}$$
$$P_{II}(\delta_c, \lambda_1) = P_{\lambda_1}\left(\sum_{i=1}^n X_i \leq c\right) = \sum_{j=0}^c \frac{(n\lambda_1)^j}{j!} e^{-n\lambda_1}$$

Bei  $n = 5$ ,  $\lambda_0 = 1$ ,  $\lambda_1 = 2$  können wir diese Werte abhängig von  $c$  tabellieren: ...

...

$c$	$P_I(\delta_c, \lambda_0)$	$P_{II}(\delta_c, \lambda_1)$
0	0.9933	0.0000
1	0.9596	0.0005
2	0.8753	0.0028
3	0.7350	0.0103
4	0.5595	0.0293
5	0.3840	0.0671
6	0.2378	0.1301
7	0.1334	0.2202
8	0.0681	0.3328
9	0.0318	0.4579
10	0.0137	0.5830
11	0.0055	0.6968

## Zusammenfassung: Prinzip der Vorgehensweise

- Vergleich der Wahrscheinlichkeiten des eingetretenen Stichprobenergebnisses bei den beiden Parametern
- Entscheidung  $d_1$  für die Gegenhypothese nur, wenn die Plausibilität hoch ist, d.h. die Wahrscheinlichkeit bei  $\gamma_1$  um den Faktor  $k$  größer ist als bei  $\gamma_0$
- Der Faktor  $k$  ist unter Beurteilung der Fehlerwahrscheinlichkeiten 1. und 2. Art festzulegen

Falls

$$P_{\gamma_0}(X = x) \neq 0$$

ist, kann damit die Entscheidung in Abhängigkeit vom **Wahrscheinlichkeitsquotienten**

$$Q(x) = \frac{P_{\gamma_1}(X = x)}{P_{\gamma_0}(X = x)}$$

festgelegt werden (**disketer Fall**). Ergänzt wird diese Definition durch

$$Q(x) = \infty \text{ für } P_{\gamma_0}(X = x) = 0.$$

Im **stetigen Fall** verwenden wir die Dichtefunktion zur Definition des Wahrscheinlichkeitsquotienten

$$Q(x) = \frac{f_{X,\gamma_1}(x)}{f_{X,\gamma_0}(x)},$$

wobei wir ebenfalls

$$Q(x) = \infty \text{ für } f_{X,\gamma_0}(x) = 0$$

ergänzen.

Ein Test auf der Basis des Wahrscheinlichkeitsquotienten mit Schranke  $k$  lautet damit

$$\delta_k(x_1, \dots, x_n) = \begin{cases} d_0 & Q(x) \leq k \\ d_1 & Q(x) > k \end{cases}$$

Ein solcher Test wird auch als **Neyman-Pearson-Test** bezeichnet.

Zur Beurteilung dieses Tests und der Festlegung der Schranke  $k$  berechnen wir die Fehlerwahrscheinlichkeit 1. Art aus

$$\begin{aligned} P_I(\delta_k, \gamma_0) &= P_{\gamma_0}(Q(X) > k) = 1 - P_{\gamma_0}(Q(X) \leq k) \\ &= 1 - F_{Q(X), \gamma_0}(k). \end{aligned}$$

Dabei ist  $F_{Q(X)}$  die Verteilungsfunktion der Zufallsvariable  $Q(X)$  beim Parameter  $\gamma_0$ .

Soll also die Fehlerwahrscheinlichkeit eine Schranke  $\alpha$  nicht überschreiten (Test zum Niveau  $\alpha$ ), so muss zunächst die Verteilungsfunktion der Zufallsvariable  $Q(X)$  und damit das  $(1 - \alpha)$ -Quantil unter Gültigkeit von Parameter  $\gamma_0$  bestimmt werden.

# Beispiele: Wahrscheinlichkeitsquotient

## 1. Bernoulli-Verteilung

$$\begin{aligned} P_p(X = x) &= p^{\sum_{i=1}^n x_i} (1-p)^{n-\sum_{i=1}^n x_i} \\ \Rightarrow Q(x) = \frac{P_{p_1}(X = x)}{P_{p_0}(X = x)} &= \frac{p_1^{\sum_{i=1}^n x_i} (1-p_1)^{n-\sum_{i=1}^n x_i}}{p_0^{\sum_{i=1}^n x_i} (1-p_0)^{n-\sum_{i=1}^n x_i}} \\ &= \left(\frac{p_1}{p_0}\right)^{\sum_{i=1}^n x_i} \left(\frac{1-p_1}{1-p_0}\right)^{n-\sum_{i=1}^n x_i} \end{aligned}$$

Der Wahrscheinlichkeitsquotient hängt also direkt ab vom Wert der suffizienten und vollständigen Statistik

$$T(x) = \sum_{i=1}^n x_i.$$



Zur Berechnung der Verteilungsfunktion  $F_{Q(x)}$  stellen wir zunächst fest

$$\begin{aligned} Q(x) &= \left(\frac{p_1}{p_0}\right)^{\sum_{i=1}^n x_i} \left(\frac{1-p_1}{1-p_0}\right)^{n-\sum_{i=1}^n x_i} \leq k \\ &\iff \sum_{i=1}^n x_i \ln\left(\frac{p_1}{p_0}\right) + \left(n - \sum_{i=1}^n x_i\right) \ln\left(\frac{1-p_1}{1-p_0}\right) \leq \ln k \\ &\iff \sum_{i=1}^n x_i \left(\ln\left(\frac{p_1}{p_0}\right) - \ln\left(\frac{1-p_1}{1-p_0}\right)\right) \leq \ln k - n \ln\left(\frac{1-p_1}{1-p_0}\right) \end{aligned}$$

Für  $p_0 < p_1$  und damit  $\ln p_1 > \ln p_0$  und  $\ln(1 - p_0) > \ln(1 - p_1)$  :

$$Q(x) \leq k \iff \sum_{i=1}^n x_i \leq \frac{\ln k - n \ln(1 - p_1) + n \ln(1 - p_0)}{\ln p_1 - \ln p_0 + \ln(1 - p_0) - \ln(1 - p_1)}$$

Für  $p_0 > p_1$  und damit  $\ln p_1 < \ln p_0$  und  $\ln(1 - p_0) < \ln(1 - p_1)$  :

$$Q(x) \leq k \iff \sum_{i=1}^n x_i \geq \frac{\ln k - n \ln(1 - p_1) + n \ln(1 - p_0)}{\ln p_1 - \ln p_0 + \ln(1 - p_0) - \ln(1 - p_1)}$$

Damit ist

$$P(Q(X) \leq k) = P\left(\sum_{i=1}^n X_i \leq \frac{\ln k - n \ln(1 - p_1) + n \ln(1 - p_0)}{\ln p_1 - \ln p_0 + \ln(1 - p_0) - \ln(1 - p_1)}\right)$$

bzw.  $P\left(\sum_{i=1}^n X_i \geq \frac{\ln k - n \ln(1 - p_1) + n \ln(1 - p_0)}{\ln p_1 - \ln p_0 + \ln(1 - p_0) - \ln(1 - p_1)}\right)$

$\sum_{i=1}^n X_i$  ist binomialverteilt mit Parameter  $p_0$  bzw.  $p_1$ , also diskret mit ganzzahligen Werten. Der Wert der Verteilungsfunktion ändert sich damit nur bei ganzzahligen Werten der rechten Seite der Ungleichungen. Die Entscheidungsfunktion

$$\delta_k(x_1, \dots, x_n) = \begin{cases} d_0 & Q(x) \leq k \\ d_1 & Q(x) > k \end{cases}$$

kann folglich auch direkt in Abhängigkeit vom Auswertungsergebnis bei der suffizienten und vollständigen Statistik  $\sum_{i=1}^n x_i$  geschrieben werden: ...

... abhängig von  $\sum_{i=1}^n x_i$ :

$$\delta_c(x_1, \dots, x_n) = \begin{cases} d_0 & \sum_{i=1}^n x_i \leq c \\ d_1 & \sum_{i=1}^n x_i > c \end{cases}$$

bzw.

$$\delta_c(x_1, \dots, x_n) = \begin{cases} d_0 & \sum_{i=1}^n x_i \geq c \\ d_1 & \sum_{i=1}^n x_i < c \end{cases}$$

Die unterschiedlichen Ungleichheitszeichen sind plausibel, da bei  $p_0 < p_1$  große Werte von  $\sum_{i=1}^n x_i$  für  $p_1$  sprechen, bei  $p_0 > p_1$  aber eher für  $p_0$  naheliegend sind.

Formal folgt dies daraus, dass bei  $p_0 < p_1$  der Wahrscheinlichkeitsquotient monoton in  $\sum_{i=1}^n x_i$  steigt, bei

$p_0 > p_1$  monoton in  $\sum_{i=1}^n x_i$  fällt.

Die Fehlerwahrscheinlichkeit 1. bzw. 2. Art kann dann unmittelbar aus der Binomialverteilung in Abhängigkeit von  $c$  berechnet werden.

Für den Test zum Niveau  $\alpha$  suchen wir ein  $c$  mit

$$P_{p_0} \left( \sum_{i=1}^n X_i > c \right) \leq \alpha \text{ für } p_0 < p_1$$

bzw.

$$P_{p_0} \left( \sum_{i=1}^n X_i < c \right) \leq \alpha \text{ für } p_0 > p_1$$

## Zahlenbeispiel:

Gesucht ist ein Test zum Niveau  $\alpha = 0.05$  für

$$H_0 : p_0 = 0.1 \quad \text{gegen} \quad H_1 : p_1 = 0.2 \quad \text{mit}$$

Stichprobenumfang  $n = 10$ . Der Test lautet damit ( $p_0 < p_1$ )

$$\delta_c(x_1, \dots, x_n) = \begin{cases} d_0 & \sum_{i=1}^n x_i \leq c \\ d_1 & \sum_{i=1}^n x_i > c \end{cases}$$

Also ist die Fehlerwahrscheinlichkeit 1. Art:

$$\begin{aligned} P_{0.1} \left( \sum_{i=1}^n X_i > c \right) &= 1 - P_{0.1} \left( \sum_{i=1}^n X_i \leq c \right) \\ &= 1 - \sum_{k=0}^c \binom{n}{k} p^k (1-p)^{n-k} \\ &= 1 - \sum_{k=0}^c \binom{10}{k} 0.1^k (1-0.1)^{10-k} \end{aligned}$$



Die Werte der Summe entnehmen wir einer Tabelle der kumulierten Binomialverteilung und bekommen:

$c$	$P(\sum X_i \leq c)$	$P(\sum X_i > c)$
0	0.3487	0.6513
1	0.7361	0.2639
2	0.9298	0.0702
3	0.9872	0.0128
4	0.9984	0.0016

Kumulierte Binomialverteilung für  $n = 10$  und  $p = 0.1$   
 $\delta_c$  ist ein Test zum Niveau  $\alpha = 0.05$ , wenn  $c = 3$  oder größer gewählt wird.

Die Fehlerwahrscheinlichkeit 2. Art ist

$$P_{0.2}\left(\sum_{i=1}^n X_i \leq c\right)$$

$c$	$P(\sum X_i \leq c)$
0	0.1074
1	0.3758
2	0.6778
3	0.8791
4	0.9672

Kumulierte Binomialverteilung für  $n = 10$  und  $p = 0.2$   
 $\Rightarrow$  Man wird  $c = 3$  wählen, um diese möglichst klein zu halten.

## Beispiel: Exponentialverteilung

$$f_{X,\lambda}(x_1, \dots, x_n) = \prod_{i=1}^n \lambda e^{-\lambda \cdot x_i} \text{ für } x_1, \dots, x_n > 0$$

$$Q(x) = \frac{f_{X,\lambda_1}(x_1, x_2, \dots, x_n)}{f_{X,\lambda_0}(x_1, x_2, \dots, x_n)} = \frac{\lambda_1^n e^{-\lambda_1 \sum_{i=1}^n x_i}}{\lambda_0^n e^{-\lambda_0 \sum_{i=1}^n x_i}} = \left(\frac{\lambda_1}{\lambda_0}\right)^n e^{-(\lambda_1 - \lambda_0) \sum_{i=1}^n x_i}$$

Folglich gilt

$$Q(x) = \left(\frac{\lambda_1}{\lambda_0}\right)^n e^{-(\lambda_1 - \lambda_0) \sum_{i=1}^n x_i} \leq k$$

$$\Leftrightarrow e^{-(\lambda_1 - \lambda_0) \sum_{i=1}^n x_i} \leq k \left(\frac{\lambda_0}{\lambda_1}\right)^n$$

$$\Leftrightarrow -(\lambda_1 - \lambda_0) \sum_{i=1}^n x_i \leq \ln k + n(\ln \lambda_0 - \ln \lambda_1)$$

$$\text{für } \lambda_1 > \lambda_0 : \Leftrightarrow \sum_{i=1}^n x_i \geq \frac{\ln k + n(\ln \lambda_0 - \ln \lambda_1)}{\lambda_0 - \lambda_1}$$

$$\text{für } \lambda_1 < \lambda_0 : \Leftrightarrow \sum_{i=1}^n x_i \leq \frac{\ln k + n(\ln \lambda_0 - \ln \lambda_1)}{\lambda_0 - \lambda_1}$$



Im Unterschied zu Beispiel 1 sind für  $c$  auch nichtganzzahlige Werte sinnvoll.

Damit ist die Fehlerwahrscheinlichkeit 1. Art

$$\text{für } \lambda_0 < \lambda_1 : \quad P_I(\delta_c, \lambda_0) = P_{\lambda_0} \left( \sum_{i=1}^n X_i < c \right)$$

$$\text{für } \lambda_0 > \lambda_1 : \quad P_I(\delta_c, \lambda_0) = P_{\lambda_0} \left( \sum_{i=1}^n X_i > c \right)$$

## Zahlenbeispiel:

Gesucht ist ein Test zum Niveau  $\alpha = 0.1$  für

$$H_0 : \lambda_0 = 0.10$$

gegen

$$H_1 : \lambda_1 = 0.15$$

beim Stichprobenumfang  $n = 8$ .

$\sum X_i$  ist Erlang-verteilt mit Stufenzahl 8 und Parameter  $\lambda$ .

Die Forderung lautet

$$P_I(\delta_c, \lambda_0) = P_{\lambda_0} \left( \sum_{i=1}^n X_i < c \right) \leq 0.1 \text{ für } \lambda_0 = 0.1 \text{ und } n = 8.$$

Gesucht ist also das 0.1-Quantil der Erlang-Verteilung mit Parameter 0.1

und Stufenzahl 8, bzw. der Gamma-Verteilung mit  $\alpha = 8$  und  $\beta = 0.1^{-1} = 10$ .

Diesen Wert erhalten wir mit Excel 7.0 und der Funktion GAMMAINV

zu 46.56... .



Test zum Niveau  $\alpha = 0.1$  für

$$H_0 : \lambda = 0.1 \text{ gegen } H_1 : \lambda = 0.15$$

ist

$$\delta_c(x_1, \dots, x_n) = \begin{cases} d_0 & \sum_{i=1}^n x_i \geq 46.56 \\ d_1 & \sum_{i=1}^n x_i < 46.56 \end{cases}$$

**Bemerkung:**

Den Wert für  $\lambda_1$  benötigen wir nicht zur Bestimmung des Tests zum Niveau  $\alpha = 0.1$  auf der Basis des Wahrscheinlichkeitsquotienten. Er wird nur zur Berechnung der Fehlerwahrscheinlichkeit 2. Art benötigt.

Auch diese Fehlerwahrscheinlichkeit 2. Art

$$P_{II}(\delta_c, \lambda_1) = P_{\lambda_1} \left( \sum_{i=1}^n X_i \geq c \right) = 1 - P_{\lambda_1} \left( \sum_{i=1}^n X_i < c \right)$$

erhalten wir mit Excel (jetzt mit der Funktion GAMMAVERT) mit:

$$1 - 0.3989 = 0.601.$$

Dieser Wert ist hoch. Bessere Werte bei diesem Niveau ( $\alpha = 0.1$ ) sind durch eine Erhöhung des Stichprobenumfangs möglich.

Bei  $n = 16$  erhalten wir die Testschranke  $c = 111.35$  bei Beibehaltung des Niveaus  $\alpha = 0.1$ . Die

Fehlerwahrscheinlichkeit 2. Art ist 0.40. Bei  $n = 100$  ist die Testschranke 874.18 und die Fehlerwahrscheinlichkeit 2. Art ist 0.002.

Liegen die beiden Parameterwerte  $\gamma_0$  und  $\gamma_1$  nicht so dicht zusammen, ist die Testsituation günstiger:

Für  $\lambda_1 = 0.3$  erhalten wir bei  $n = 8$  den identisch aussehenden Test zum Niveau  $\alpha = 0.1$  für

$$H_0 : \lambda = 0.1 \quad \text{gegen} \quad H_1 : \lambda = 0.3$$

$$\delta_c(x_1, \dots, x_n) = \begin{cases} d_0 & \sum_{i=1}^n x_i \geq 46.56 \\ d_1 & \sum_{i=1}^n x_i < 46.56 \end{cases}$$

Die Fehlerwahrscheinlichkeit 2. Art ist jetzt:

$$\begin{aligned} P_{II}(\delta_c, 0.3) &= P_{0.3}\left(\sum_{i=1}^n X_i \geq 46.56\right) \\ &= 1 - P_{0.3}\left(\sum_{i=1}^n X_i < 46.56\right) \\ &= 0.0379. \end{aligned}$$

Zusammenfassung der Erkenntnisse aus den Beispielen:

1. Die Tests auf der Basis des Wahrscheinlichkeitsquotienten liefern plausible Ergebnisse, da die Entscheidung jeweils von einer suffizienten Statistik abhängt.
2. Die Ungleichheitszeichen bei den Tests mit der vollständigen und suffizienten Statistik als Testgröße hängen davon ab, ob der Wahrscheinlichkeitsquotient monoton steigend oder fallend in der Statistik ist. In den Beispielen hängt dies davon ab, ob  $\gamma_0 < \gamma_1$  oder  $\gamma_0 > \gamma_1$  ist.
3. Zur Berechnung der Testschranke bei der vollständigen und suffizienten Statistik ist die Kenntnis des Parameterwerts  $\gamma_1$  nicht erforderlich.  $\gamma_1$  wird nur zur Berechnung der Fehlerwahrscheinlichkeit 2. Art benötigt.

## Frage:

Ist ein Test zum Niveau  $\alpha$  auf der Basis des  
Wahrscheinlichkeitsquotienten wirklich das beste  
Entscheidungsverfahren, also der beste Test zum Niveau  $\alpha$ ?  
Zur Konkurrenz zugelassen sind also nur die Tests zum  
Niveau  $\alpha$ .

Sei  $\delta_k$  ein Test auf der Basis des  
Wahrscheinlichkeitsquotienten zum Niveau  $\alpha$  im diskreten  
Fall:

$$\delta_c(x_1, \dots, x_n) = \begin{cases} d_0 & k P_{\gamma_0}(X = x) \geq P_{\gamma_1}(X = x) \\ d_1 & \text{sonst} \end{cases}$$

Das Niveau wird eingehalten, falls

$$P_I(\delta_k, \gamma_0) \leq \alpha$$

ist. Angenommen es gilt  $P_I(\delta_k, \gamma_0) = \alpha$ , d.h.  $k$  und  $\alpha$  sind so  
gewählt, dass die Fehlerwahrscheinlichkeit 1. Art genau  $\alpha$  ist.

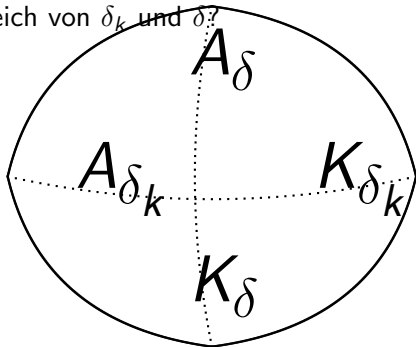
## Bemerkung:

Da bei diskreten Zufallsvariablen nicht jedes Niveau exakt eingestellt werden kann, wird man zu jedem Faktor  $k$  (Schranke für den Wahrscheinlichkeitsquotienten) die Fehlerwahrscheinlichkeit 1. Art bestimmen und kennt so die Kombinationen von  $k$  und zugehörigem exakt eingehaltenem Niveau.

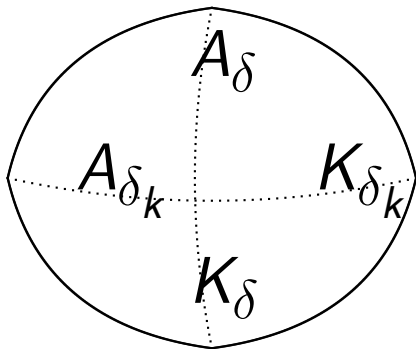
Wenn der Wahrscheinlichkeitsquotient  $Q(X)$  eine stetige Zufallsvariable ist, gibt es zu jedem Niveau eine Schranke  $k$  für den Wahrscheinlichkeitsquotienten, so dass der Test genau dieses Niveau exakt einhält.



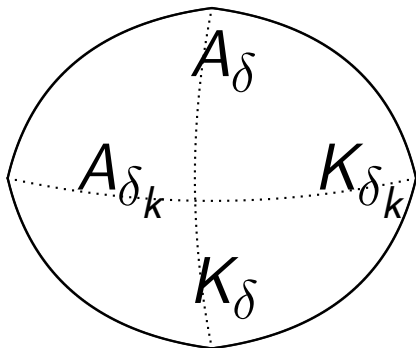
Sei  $\delta$  ein anderer Test zum Niveau  $\alpha$ . Was ergibt ein Vergleich von  $\delta_k$  und  $\delta$ ?



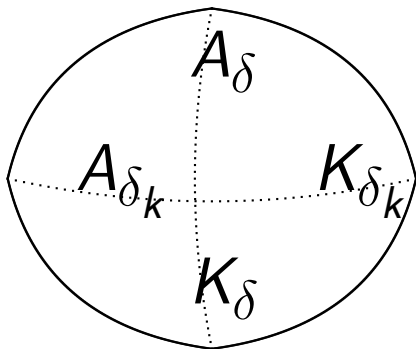
Der grüne Bereich in der Graphik, der zum kritischen Bereich von  $\delta$  gehört, ist bei  $\delta_k$  dem Annahmebereich zugeteilt, der blaue Bereich des Annahmebereichs von  $\delta$  wird bei  $\delta_k$  dem kritischen Bereich zugeordnet.



Für einen Punkt  $x_1$  des blauen Bereichs ist die Wahrscheinlichkeit bei  $\gamma_1$  um den Faktor  $k$  höher als bei  $\gamma_0$ . Bei  $\delta$  gehört  $x_1$  zum Annahmehbereich, trägt also viel zur Fehlerwahrscheinlichkeit 2. Art bei  $\delta$  bei.



Ändern wir  $\delta$  im Punkt  $x_1$  von der Entscheidung  $d_0$  zur Entscheidung  $d_1$ , so reduziert sich die Fehlerwahrscheinlichkeit 2. Art. Im Gegenzug dazu wird die Wahrscheinlichkeit für  $d_0$  auch bei  $\gamma_0$  kleiner  $\Rightarrow$  die Fehlerwahrscheinlichkeit 1. Art steigt um  $P_{\gamma_0}(X = x_1)$ .



Dies kompensieren wir durch einen Punkt  $x_2$  aus dem grünen Bereich mit  $P_{\gamma_0}(X = x_2) = P_{\gamma_0}(X = x_1)$ , indem wir die Entscheidung von  $d_1$  auf  $d_0$  ändern.  $x_2$  gehört zum Annahmebereich von  $\delta_k \Rightarrow k P_{\gamma_0}(X = x_2) \geq P_{\gamma_1}(X = x_2)$ .

Durch die Änderung der Entscheidung bei  $\delta$  in  $x_1$  und  $x_2$  wird bewirkt:

- bei der Fehlerwahrscheinlichkeit 1. Art:

$$P_I(\delta, \gamma_0) + P_{\gamma_0}(X = x_1) - P_{\gamma_0}(X = x_2) = P_I(\delta, \gamma_0)$$

- bei der Fehlerwahrscheinlichkeit 2. Art:

$$P_{II}(\delta, \gamma_1) - P_{\gamma_1}(X = x_1) + P_{\gamma_1}(X = x_2)$$

$$< P_{II}(\delta, \gamma_1) - k P_{\gamma_0}(X = x_1) + k P_{\gamma_0}(X = x_2) = P_{II}(\delta, \gamma_1)$$

Also:

Der Wechsel der Entscheidung in  $x_1$  und  $x_2$  beim Test  $\delta$  wirkt sich bei der Fehlerwahrscheinlichkeit 1. Art nicht aus, bringt aber eine Verbesserung bei der Fehlerwahrscheinlichkeit 2.

Art.

Möglicherweise wird man keine Punkte  $x_1$  und  $x_2$  mit übereinstimmender Wahrscheinlichkeit bei  $\gamma_0$  finden. Dies ist aber auch nicht nötig, da es genügt, eine Kombination von Punkten zu finden, bei denen dies richtig ist.

# Satz von Neyman-Pearson:

1. Für ein einfaches Testproblem

$$H_0 : \gamma = \gamma_0 \quad \text{gegen} \quad H_1 : \gamma = \gamma_1$$

ist der **Neyman-Pearson-Test**

$$\delta_c(x_1, \dots, x_n) = \begin{cases} d_0 & k P_{\gamma_0}(X = x) \geq P_{\gamma_1}(X = x) \\ d_1 & \text{sonst} \end{cases}$$

für  $k \in \mathbb{R}, k > 0$  bester Test zum Niveau

$$P_I(\delta_k, \gamma_0) = P_{\gamma_0}(\delta_k(X) = d_1).$$

# Satz von Neyman-Pearson:

2. Jeder andere Test  $\delta$  mit übereinstimmenden Fehlerwahrscheinlichkeiten

1. und 2. Art stimmt fast sicher mit  $\delta_k$  überein:

$$P(\delta(X) = \delta_k(X)) = 1$$



## Achtung:

Bei diskreten Zufallsvariablen nimmt  $P_I(\delta_k, \gamma_0)$  für Neyman-Pearson-Tests  $\delta_k$  nur diskrete Werte bei der Variation von  $k$  an. Es kann also **nicht zu jedem vorgegebenen Niveau**  $\alpha$  ein Neyman-Pearson-Test als bester Test zu diesem Niveau gefunden werden.

Wahrscheinlichkeitsquotient bei Existenz einer suffizienten Statistik:

Nach dem Faktorisierungstheorem gilt im **diskretem** Fall:

$$P_{\gamma}(X = x) = g(T(x), \gamma) h(x)$$

und im *stetigen* Fall:

$$f_{X,\gamma}(x) = g(T(x), \gamma) h(x)$$

Damit gilt für den Wahrscheinlichkeitsquotienten:

$$Q(x) = \frac{P_{\gamma_1}(X = x)}{P_{\gamma_0}(X = x)} = \frac{g(T(x), \gamma_1) h(x)}{g(T(x), \gamma_0) h(x)} = \frac{g(T(x), \gamma_1)}{g(T(x), \gamma_0)} \quad (\text{diskret})$$

bzw.

$$Q(x) = \frac{f_{X, \lambda_1}(x)}{f_{X, \lambda_0}(x)} = \frac{g(T(x), \lambda_1) h(x)}{g(T(x), \lambda_0) h(x)} = \frac{g(T(x), \lambda_1)}{g(T(x), \lambda_0)} \quad \text{stetig}$$

## Folgerung:

Der Wahrscheinlichkeitsquotient hängt nur vom Auswertungsergebnis  $T(x)$  der suffizienten Statistik ab, d.h.

Seien  $x, \tilde{x}$  Stichprobenergebnisse:

$$T(x) = T(\tilde{x}) \Rightarrow Q(x) = Q(\tilde{x})$$

Der Neyman-Pearson-Test ist also eine Entscheidungsfunktion auf der Basis jeder suffizienten Statistik.

Der Wahrscheinlichkeitsquotient für einparametrische Exponentialfamilien:

$Y$  sei eine stetige Zufallsvariable mit einparametrischer Exponentialfamilie, d.h.

$$f_{Y,\gamma}(y) = a(\gamma)h(y) e^{b(\gamma)\tau(y)}$$

Dann gilt für die Stichprobe mit Zurücklegen:

$$\begin{aligned} f_{X,\gamma}(x_1, \dots, x_n) &= \prod_{i=1}^n a(\gamma)h(x_i) e^{b(\gamma)\tau(x_i)} \\ &= a(\gamma)^n \left( \prod_{i=1}^n h(x_i) \right) e^{\sum_{i=1}^n b(\gamma)\tau(x_i)} \end{aligned}$$

und für den Wahrscheinlichkeitsquotienten

$$\begin{aligned} Q(x) &= \frac{f_{X,\gamma_1}(x_1, \dots, x_n)}{f_{X,\gamma_0}(x_1, \dots, x_n)} = \frac{a(\gamma_1)^n \left( \prod_{i=1}^n h(x_i) \right) e^{\sum_{i=1}^n b(\gamma_1)\tau(x_i)}}{a(\gamma_0)^n \left( \prod_{i=1}^n h(x_i) \right) e^{\sum_{i=1}^n b(\gamma_0)\tau(x_i)}} \\ &= \left( \frac{a(\gamma_1)}{a(\gamma_0)} \right)^n \frac{e^{\sum_{i=1}^n b(\gamma_1)\tau(x_i)}}{e^{\sum_{i=1}^n b(\gamma_0)\tau(x_i)}} \\ &= \left( \frac{a(\gamma_1)}{a(\gamma_0)} \right)^n e^{(b(\gamma_1) - b(\gamma_0)) \sum_{i=1}^n \tau(x_i)} \end{aligned}$$

## Folgerungen:

1.  $Q(x)$  hängt nur von der Statistik  $T(x) = \sum_{i=1}^n \tau(x_i)$  ab.
2. a)  $b(\gamma_1) - b(\gamma_0) > 0$ :  
 $Q(x)$  steigt monoton in  $T(x)$ , also

$$Q(x) \leq k \iff T(x) = \sum_{i=1}^n \tau(x_i) \leq c$$

für zusammengehörende Werte von  $k$  und  $c$ .

**Analog im diskreten Fall.**

b)  $b(\gamma_1) - b(\gamma_0) < 0$ :

$Q(x)$  fällt monoton in  $T(x)$ ,  
also

$$Q(x) \leq k \iff T(x) = \sum_{i=1}^n \tau(x_i) \geq c$$

für zusammengehörende Werte von  $k$  und  $c$ .

**Analog im diskreten Fall.**



# Anwendung:

## 1. Bernoulli-Verteilung

$$\begin{aligned}P_p(Y = y) &= p^y (1 - p)^{1-y} = e^{y \ln p} e^{(1-y) \ln(1-p)} \\ &= e^{y(\ln p - \ln(1-p))} e^{\ln(1-p)}\end{aligned}$$

$$b(p) = \ln p - \ln(1 - p)$$

$$p_0 < p_1 \Rightarrow b(p_1) - b(p_0) > 0 \Rightarrow \delta(x) = \begin{cases} d_0, & \sum_{i=1}^n x_i \leq c \\ d_1, & \sum_{i=1}^n x_i > c \end{cases}$$

$$p_0 > p_1 \Rightarrow b(p_1) - b(p_0) < 0 \Rightarrow \delta(x) = \begin{cases} d_0, & \sum_{i=1}^n x_i \geq c \\ d_1, & \sum_{i=1}^n x_i < c \end{cases}$$

## 2. Binomialverteilung

$$\begin{aligned}P_p(Y = k) &= \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} e^{k \ln p} e^{(n-k) \ln(1-p)} \\ &= \binom{n}{k} e^{k(\ln p - \ln(1-p))} e^{n \ln(1-p)} \\ b(p) &= \ln p - \ln(1-p)\end{aligned}$$

### 3. Poisson-Verteilung

$$\begin{aligned}P_{\lambda}(Y = k) &= \frac{\lambda^k}{k!} e^{-\lambda} = \frac{1}{k!} e^{k \ln \lambda} e^{-\lambda} \\b(\lambda) &= \ln \lambda\end{aligned}$$

## 4. Exponentialverteilung

$$\begin{aligned}f_{Y,\lambda}(y) &= \mathbf{1}_{[0,\infty)}(y)\lambda e^{-\lambda y} \\ b(\lambda) &= -\lambda\end{aligned}$$

## 5. Normalverteilung

a)  $\sigma^2$  bekannt (Test bezüglich des Mittelwerts)

$$b(\mu) = \frac{\mu}{\sigma^2}, \tau(y) = y, \text{ d.h. } T(x) = \sum_{i=1}^n x_i$$

$$\mu_0 < \mu_1 \Rightarrow b(\mu_1) - b(\mu_0) > 0 \Rightarrow \delta(x) = \begin{cases} d_0 & \sum_{i=1}^n x_i \leq c \\ d_1 & \sum_{i=1}^n x_i > c \end{cases}$$
$$\mu_0 > \mu_1 \Rightarrow b(\mu_1) - b(\mu_0) < 0 \Rightarrow \delta(x) = \begin{cases} d_0 & \sum_{i=1}^n x_i \geq c \\ d_1 & \sum_{i=1}^n x_i < c \end{cases}$$

## 5. Normalverteilung

b)  $\mu$  bekannt (Test bezüglich Varianz)

$$b(\sigma^2) = -\frac{1}{2\sigma^2}, \tau(y) = (y - \mu)^2, \text{ d.h. } T(x) = \sum_{i=1}^n (x_i - \mu)^2$$

$$\sigma_0 < \sigma_1 \Rightarrow b(\sigma_1) - b(\sigma_0) > 0 \Rightarrow \delta(x) = \begin{cases} d_0 & \sum_{i=1}^n (x_i - \mu)^2 \leq c \\ d_1 & \sum_{i=1}^n (x_i - \mu)^2 > c \end{cases}$$
$$\sigma_0 > \sigma_1 \Rightarrow b(\sigma_1) - b(\sigma_0) < 0 \Rightarrow \delta(x) = \begin{cases} d_0 & \sum_{i=1}^n (x_i - \mu)^2 \geq c \\ d_1 & \sum_{i=1}^n (x_i - \mu)^2 < c \end{cases}$$

# Zusammenfassung: Neyman-Pearson-Tests

Verteilung	$T(x)$	Parameter- beziehung	Monotonie des Wahrscheinlichkeits- quotienten in $T(x)$	Neyman- Pearson- Test
Bernoulli- und Binomialverteilung	$\sum x_i$	$p_0 < p_1$	steigend	(1)
		$p_0 > p_1$	fallend	(2)
Poisson-Verteilung	$\sum x_i$	$\lambda_0 < \lambda_1$	steigend	(1)
		$\lambda_0 > \lambda_1$	fallend	(2)
Exponential- verteilung	$\sum x_i$	$\lambda_0 < \lambda_1$	fallend	(2)
		$\lambda_0 > \lambda_1$	steigend	(1)
Normalverteilung $\sigma^2$ bekannt	$\sum x_i$	$\mu_0 < \mu_1$	steigend	(1)
		$\mu_0 > \mu_1$	fallend	(2)
Normalverteilung $\mu$ bekannt	$\sum (x_i - \mu)^2$	$\sigma_0^2 < \sigma_1^2$	steigend	(1)
		$\sigma_0^2 > \sigma_1^2$	fallend	(2)

# Zusammenfassung

$$\delta_c(x) = \begin{cases} d_0 & T(x) \leq c \\ d_1 & T(x) > c \end{cases} \quad (1)$$

$$\delta_c(x) = \begin{cases} d_0 & T(x) \geq c \\ d_1 & T(x) < c \end{cases} \quad (2)$$

Testschranke  $c$  steuert Fehlerwahrscheinlichkeit 1. Art, d.h. das Niveau.

Bei vorgegebenem Niveau ist  $c$  geeignet zu bestimmen.