

Übungsblatt 11

Übung 11.1 Vergleichen Sie die verschiedenen Konfidenz-Intervalle für $\alpha = 0.05$ für die verschiedenen in der folgenden Tabelle angegebenen Werte.

m	\hat{p}	<i>klassisch</i>	<i>Wilson</i>	<i>HPD</i>
100	0.01			
1000	0.01			
50	0.4			
500	0.4			

Hinweis: Das HPD-Intervall lässt sich nur numerisch berechnen. Dazu gibt es in R die Funktion *hdi* aus dem Paket *HDInterval* ($m = \text{Anzahl}$, $k = \text{Treffer}$):

```
library(HDInterval)
hdi(qbeta, shape1=(k+1), shape2=(m-k+1))
```

Übung 11.2 Vergleichen Sie für die in Übung 8.2 verwendeten Classifier auf dem IRIS-Datensatz die *Holdout Error Rate* e_T , die *Apparent Error Rate* e_A und die *Leave-One-Out Error Rate* e_{CV} .

- Wählen Sie dazu zunächst die Hälfte der Zeichen (zufällige Auswahl) als Trainingsdaten, so dass Sie alle drei Fehlerraten bestimmen können.
- Nehmen Sie anschließend alle Daten als Trainingsdaten und bestimmen Sie danach e_A und e_{CV} .

Wie verhalten sich die Fehlerraten zueinander?

Hinweis: Die Methode *knn.cv* im paket *class* bestimmt bereits die Leave-One-Out Rate eines knn-Classifiers