LMU München **Prof. F. Merkl**
Mathematisches Institut
Sommersemester 2018

Übungen zur Mathematischen Statistik
Blatt 10

1. Gegeben sei eine einfache Hypothese $H_0 = \{P_0\}$ und eine zugehörige Familie $(v_c = 1_{\{T \le c\}})_{c \in \mathbb{R}}$ von statistischen Tests zu variablen kritischen Werten c mit einer Teststatistik T. Die Verteilung von T unter P_0 sei atomlos: $P_0(T = c) = 0$ für alle $c \in \mathbb{R}$. Zeigen Sie, dass der zugehörige p-Wert unter P_0 uniform auf dem Einheitsintervall verteilt ist.

2. Die χ^2-Verteilung mit n Freiheitsgraden, kurz χ^2_n, ist die Verteilung von $\sum_{j=1}^n X_j^2$, wenn $X_1, \ldots, X_n$ i.i.d. standardnormalverteilt sind. Die Gammaverteilung $\gamma_{s,a}$ mit dem Formparameter $s > 0$ und dem Skalenparameter $a > 0$ ist das folgende Wahrscheinlichkeitsmaß auf $\mathbb{R}$:

$$\gamma_{s,a}(dx) = \frac{a^s}{\Gamma(s)} x^{s-1} e^{-ax} 1_{\{x>0\}}\, dx$$

Zeigen Sie $\chi^2_n = \gamma_{\frac{n}{2},\frac{1}{2}}$ für alle $n \in \mathbb{N}$.

3. Gegeben seien $n \ge 2$ unabhängige normalverteilte Datenpunkte $X_1, \ldots, X_n$ mit Erwartungswert 0 und Varianz σ^2 mit Mittel $\bar{X} = \frac{1}{n}\sum_{j=1}^n X_j$, empirischer Varianz $s^2 = \frac{1}{n-1}\sum_{j=1}^n (X_j - \bar{X})^2$ sowie die Z- und t-Statistiken

$$Z = \sqrt{n}\frac{\bar{X}}{\sqrt{\sigma^2}},$$
$$t = \sqrt{n}\frac{\bar{X}}{\sqrt{s^2}}.$$

Das zugrundeliegende Wahrscheinlichkeitsmaß heiße P. Die Z-Statistik ist standardnormalverteilt, aber auch die t-Statistik ist für $n \to \infty$ asymptotisch standardnormalverteilt, weil s^2/σ^2 für $n \to \infty$ in Wahrscheinlichkeit gegen 1 geht.

   (a) Zeigen Sie bei fixierter Stichprobengröße n:

$$\frac{P[t > c]}{P[Z > c]} \overset{c\to\infty}{\longrightarrow} \infty.$$

   Trotz gleicher Asymptotik für $n \to \infty$ untscheiden sich also die t-Statistik und die Z-Statistik bei großen kritischen Werten c stark.

   (b) Zeigen Sie, dass s^2 unter $P[\cdot \mid t > c]$ für $c \to \infty$ in Wahrscheinlichkeit gegen 0 geht, d.h.

$$\forall \epsilon > 0 : \lim_{c\to\infty} P[|s^2| > \epsilon \mid t > c] = 0.$$

   Große Abweichungen von t werden also typischerweise von kleinen Nennern $\sqrt{s^2}$ verursacht.

Keine Abgabe.