

Hausaufgaben für LL-08, Statistik zum 18.6.2009

Aufgabe 1

Lesen Sie Abschnitt 4.2 *Punktschätzungen* nochmals. Überspringen Sie den Unterabschnitt *Steigung der Regressionsgeraden*.

Anmerkung zum davorliegenden *Schätzprinzip der schließenden Statistik*: Es sollte besser lauten „... wird durch die entsprechende Formel der Stichprobenwerte geschätzt.“ Da ein Parameter eine konstante Größe der Grundgesamtheit ist, sollte man nicht von einem Parameter der Stichprobe schreiben.

Zu Seite 136 „Nach dem Gesetz der großen Zahlen (3.55) konvergiert dann der Durchschnitt vieler Schätzungen gegen den zu schätzenden Wert.“ Dieses Verhalten sollte das Urnenbeispiel auch verdeutlichen - in der Praxis wird in der Regel aber nur eine Stichprobe genommen.

Zu Satz 4.1: (i) Für die Schätzfunktion „Stichprobenmittel“ wissen wir sogar, dass sie näherungsweise normalverteilt ist. (ii) Die Anteilsschätzfunktion \hat{p} ist nur ein Spezialfall des Stichprobenmittels: Sie verwendet Bernoulli-Zufallsvariablen (X_i zählt, ob das Ereignis in der i -ten Ziehung vorliegt). (iii) Die empirische Varianz ist NICHT normalverteilt. Man beachte z.B., dass der Wertebereich nur positiv ist. Hierzu gibt es die so genannte χ^2 -Verteilung (s. Seite 160).

Der Unterabschnitt *Angabe von Unsicherheiten beim Messen* ist von großer praktischer Bedeutung in der Messtechnik. Die Beispiele 4.3 und 4.4 sollten besonders gut verstanden werden.

Aufgabe 2

Lesen Sie Abschnitt 4.3 *Intervallschätzungen* bis Seite 153.

Der Unterabschnitt zum normalverteilten Merkmal mit bekannter Standardabweichung ist wie in der Vorlesung grundlegend für das weitere Verständnis. Die Annahme, dass die Standardabweichung bekannt ist, ist natürlich trotzdem wenig praxisnah. Deshalb ist auch der folgende Abschnitt mit unbekannter Standardabweichung (Verwendung der t -Verteilung) umso bedeutsamer. Der Entdecker der t -Verteilung entwickelte das Verfahren zum Einsatz beim Bierbrauen (Guinness-Brauerei). Er veröffentlichte unter einem Pseudonym, damit die Konkurrenz den Anwendungsnutzen nicht erkennen sollte. Die t -Verteilung sieht wie eine „fette“ Normalverteilung aus, die mit zunehmender Anzahl Freiheitsgrade sich der Normalverteilung annähert.

Zu unseren Stichproben habe ich die Konfidenzintervalle mit dem Verfahren auf Seite 150 berechnet. Schauen Sie sich dazu die Excel-Datei an. Das dortige Diagramm der Konfidenzintervalle zu den einzelnen Stichproben soll verdeutlichen, was die 95 % bedeuten: Das Verfahren liefert in 95% der Fälle ein Konfidenzintervall, welches den tatsächlichen Wert enthält. Unsere Zahlenbeispiele belegen das sehr schön: 2 von 41 Konfidenzintervalle, also 4,8%, verfehlen den tatsächlichen Wert.

Beachten Sie, dass das Verfahren auf Seite 151 unten bei unbekannter Standardabweichung inkonsistenterweise auf (4.25) verweist, obwohl auf Seite 150 bei derselben Stichprobengröße bereits die Normalverteilung empfohlen wird.

Rechnen Sie die Beispiele 4.6 bis 4.9 nach.