

Übung 6

1. Unterhalb einer Kläranlage wurden 16 unabhängige Wasserproben aus einem Fluss entnommen und jeweils deren Ammoniumkonzentration X_i (in $\mu\text{gNH}_4\text{-N/l}$) mit einem Messgerät bestimmt. Der Mittelwert der Proben ergab $\bar{x} = 204.2$.

Wir wollen nun wissen, ob mit diesem Experiment eine Überschreitung des Grenzwerts von $200 \mu\text{gNH}_4\text{-N/l}$ nachgewiesen werden kann (auf dem 5% Niveau).

- a) Nimm an, die Standardabweichung der Messungen sei im voraus aufgrund früherer Studien bekannt. Sie betrage $10 \mu\text{gNH}_4\text{-N/l}$.
Führe unter dieser Annahme einen z -Test durch, um zu überprüfen, ob eine Grenzwertüberschreitung nachgewiesen werden kann.
Gib die Modellannahmen, H_0 , H_A , den Verwerfungsbereich, den Wert der Teststatistik und das Testergebnis explizit an.
- b) Wie wahrscheinlich ist es, dass man mit 16 unabhängigen Wasserproben eine Grenzwertüberschreitung nachweisen kann, wenn die wahre Ammoniumkonzentration tatsächlich über dem Grenzwert und zwar bei $205 \mu\text{gNH}_4\text{-N/l}$ liegt?
- c) Wie wahrscheinlich ist es, dass man mit 16 unabhängigen Wasserproben fälschlicherweise eine Grenzwertüberschreitung nachweist, obwohl die wahre Ammoniumkonzentration bei $200 \mu\text{gNH}_4\text{-N/l}$ liegt und den Grenzwert somit genau einhält?

2. In dieser Aufgabe untersuchen wir die Wirkung des Zentralen Grenzwertsatzes mittels Simulation. Wir gehen von einer Zufallsvariablen X aus, die folgendermassen verteilt ist: die Werte 0, 10 und 11 werden je mit einer Wahrscheinlichkeit $\frac{1}{3}$ angenommen. Nun simulieren wir die Verteilung von X sowie die Verteilung des Mittelwerts \bar{X} von mehreren X .

- a) Simuliere X . Stelle die Verteilung von X mittels eines Histogramms von 1000 Realisierungen von X dar, und vergleiche sie mittels des Normalplots mit der Normalverteilung.


```
> par(mfrow=c(4,2)) # Mehrere Grafiken neben- und untereinander
> werte <- c(0,10,11) # moegliche Werte von X
> sim <- sample(werte,1000, replace = T) # X simulieren
> hist(sim, main = "Original") # Histogramm erstellen
> qqnorm(sim) # Normalplot erstellen
```
- b) Simuliere nun $\bar{X} = \frac{X_1+X_2+X_3+X_4+X_5}{5}$, wobei die X_i die gleiche Verteilung haben wie X und unabhängig sind. Stelle die Verteilung von \bar{X} anhand von 1000 Realisierungen von \bar{X} dar, und vergleiche sie mit der Normalverteilung.


```
> n <- 5
> sim <- matrix(sample(werte, n*1000, replace = T), ncol=n)
# X_1, ..., X_n simulieren und in einer n-spaltigen Matrix (mit 1000 Zeilen) anordnen
> sim.mean <- apply(sim, 1, "mean") # In jeder Matrixzeile X-bar berechnen
> hist(sim.mean, main = paste("Mittelwerte von", n, "Beobachtungen"))
> qqnorm(sim.mean)
```
- c) Simuliere nun die Verteilung von \bar{X} auch für die Fälle, wo \bar{X} das Mittel von 10 resp. 200 X_i ist.