

## Musterlösung zu Serie 7

1. a) Das Intervall sollte in 95% der Fälle den wahren Wert enthalten. Da wir 20 Realisationen betrachten, erwarten wir, dass der Wert im Schnitt 1 mal nicht im Intervall enthalten ist.

Wir verwenden folgenden R-Code:

```
set.seed(79) ## Macht Resultate reproduzierbar

p <- 0.3
x <- rbinom(20, 50, p) ## 20 Werte simulieren

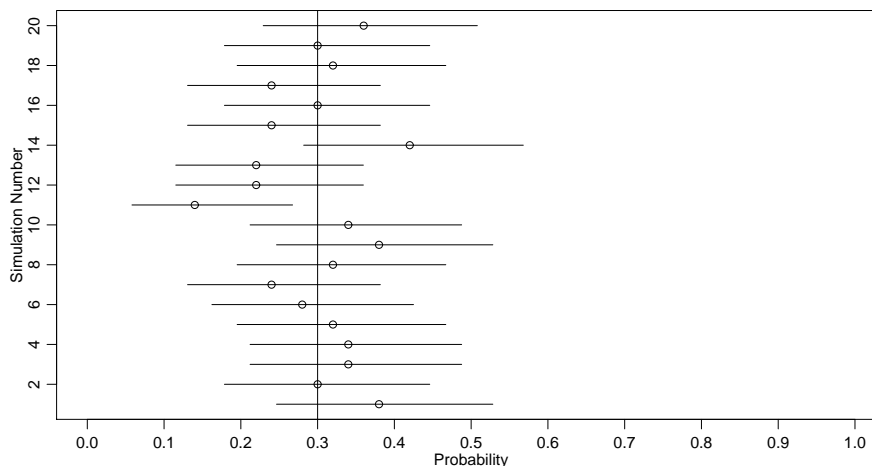
## Grenzen der Intervalle in Matrix speichern
## 1. Spalte ist untere Grenze, 2. Spalte obere
confint.bound <- matrix(0, nrow = 20, ncol = 2)
contains.truth <- logical(20)
## Alle 20 Faelle untersuchen und Grenzen speichern
for(i in 1:20){
  test <- binom.test(x[i], 50, p)
  confint.bound[i,] <- test$conf.int
  contains.truth[i] <-
    (p >= confint.bound[i,1]) & (p <= confint.bound[i,2])
}
sum(contains.truth) ## 19
```

Für unsere Simulationen ist der wahre Wert wie erwartet in 19 der Vertrauensintervalle enthalten. Je nach Simulation kann es natürlich sein, dass der Wert immer enthalten oder in weniger als 19 Fälle enthalten ist (die Anzahl Intervalle, die den wahren Wert enthalten, ist binomialverteilt mit  $n = 20$  und Erfolgswahrscheinlichkeit 0.95)

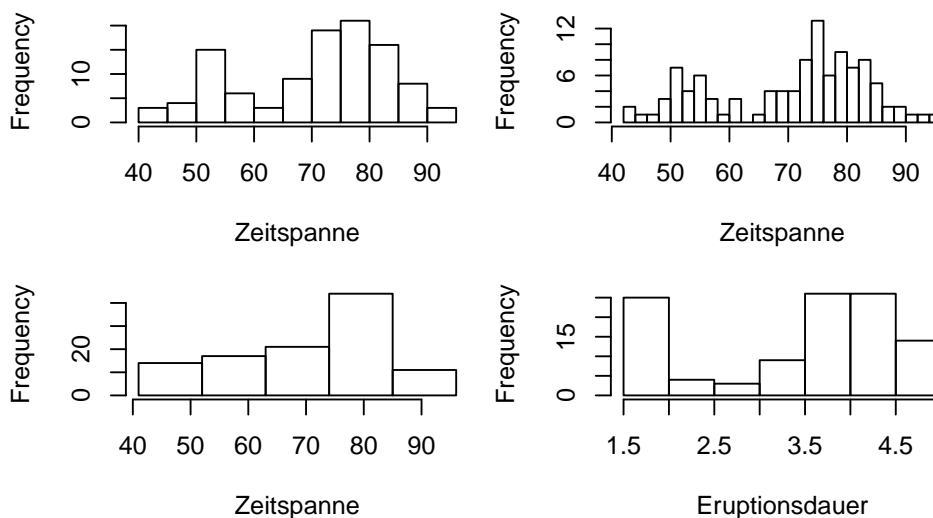
- b) R-Code:

```
## Relative Haeufigkeiten plotten
plot(x / 50, 1:20, xlim = c(0, 1), xlab = "Probability",
      ylab = "Simulation Number")
## Vertrauensintervalle als Liniensegmente plotten
for(i in 1:20){
  segments(confint.bound[i,1], i, confint.bound[i,2], i)
}
## Wahrer Wert als vertikale Linie einzeichnen
abline(v = p)
```

und wir erhalten so folgende Grafik



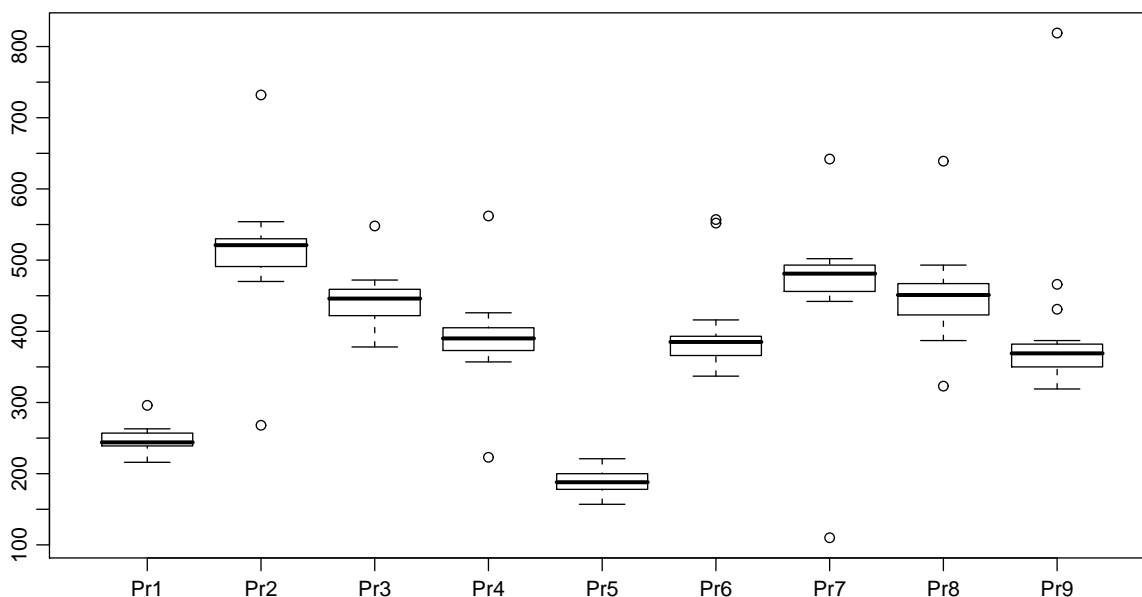
2. a) b) Die ersten drei Histogramme in der Abbildung unten zeigen die Intervalle zwischen zwei Ausbrüchen von Old Faithful. Auffallend ist, dass Zeitspannen um 55 Minuten aber auch zwischen 70 bis 85 Minuten häufiger vorkommen als andere Intervalle. So eine Verteilung mit zwei Gipfeln heisst auch *bimodal*.



Werden die Klassenbreiten ungeschickt gewählt, entdeckt man diese Besonderheit der Geysirdaten nicht. Das ist im dritten Histogramm passiert. Das Beispiel illustriert, dass die richtige Wahl der Klassenbreiten- bzw. grenzen wohlüberlegt sein muss.

Das vierte Histogramm schliesslich zeigt die Häufigkeiten verschiedener Eruptionsdauern. Hier sind die beiden Gipfel sehr deutlich erkennbar: “Entweder ist der Ausbruch sofort wieder vorbei, oder er dauert mindestens dreieinhalb Minuten”. Ob die Dauer eines Ausbruchs aber etwas zu tun hat mit der Dauer des vorangegangenen Ruheintervalls (mit anderen Worten: ob die Gipfel des Histogramms aus Teilaufgabe b) den Gipfeln der Histogramme aus Teilaufgabe a) entsprechen), kann man aufgrund dieser Darstellungen *nicht* sagen.

3. a) Aus den Boxplots erkennen wir, dass es vor allem bei den Proben 2, 4, 6, 7, 8 und 9 Ausreisser gibt. Das arithmetische Mittel und der Median unterscheiden wesentlich bei den Proben 2, 6, 7 und 9.



	PR1	PR2	PR3	PR4	PR5	PR6
Median	244.000	521.000	446.000	390.000	188.000	385.000
Mean	246.143	511.429	443.429	389.238	188.238	394.857

---

	PR7	PR8	PR9
Median	481.000	451.000	369.000
Mean	465.524	449.952	388.905

Bei den Proben 1 und 5 ist es plausibel, dass die Konzentration unter 400 mg/kg liegt, während wir bei Probe 2, 3, 7 und 8 dazu tendieren, den Grenzwert 400 mg/kg als überschritten zu betrachten. Die übrigen Proben, Probe 4, 6 und 9 sind eher Grenzfälle. Die Konzentrationen scheinen zwar unter 400 mg/kg zu liegen, die drei Proben weisen jedoch jeweils extreme Ausreisser über dem Grenzwert auf.

- b) Als erstes stechen die Messungen der Labors 15 und 21 ins Auge. Beide haben sowohl eine grosse Standardabweichung als auch systematische Fehler. Die Labors 6 und 12 haben beide Ausreisser zu verzeichnen. Die Labors 1, 7, 12, 13, 14, 17, 18, 20 und 21 geben systematisch zu kleine Werte an, während die Labors 6, 8, 10 und 15 zu grosse Werte erhalten. Die Labors 2, 3, 4, 5 und 19 scheinen zuverlässige Untersuchungen durchzuführen. Sowohl systematische wie auch Zufallsfehler scheinen sich hier in Grenzen zu halten.

