

Loesung Aufgabe 1.1

[Download R-Skript](#)

[Download PDF](#)

kommentierte Musterlösungen

```
# Als eine Möglichkeit, die Aufgabe 1.1 zu bearbeiten, nehmen wir hier den
# Datensatz der Gästebefragung NOVANIMAL und gehen der folgenden Frage nach:
# Gibt es einen Zusammenhang zwischen Geschlecht und dem wahrgenommenen
# Milchkonsum (viel vs. wenig Milch/-produkte)

# die Variable wahrgenommener Milchkonsum muss noch in 2 Kategorien zusammengefasst werden: geringer vs.

# Variable milk == wahrgenommener Milchkonsum
# alles kleiner als 4 (3 inklusive) == geringer wahrgenommener Milchkonsum,
# alles grösser als 3 (4 inklusive) == hoher wahrgenommener Milchkonsum
nova2 <- nova_survey %>%
  filter(gender != "x") %>% # x aus der Variable Geschlecht entfernen
  mutate(milkcompt = if_else(milk >= 3, "wenig", "viel")) %>%
  select(gender, milkcompt) %>%
  drop_na() # alle Missings können gestrichen werden

# mal anschauen
table(nova2)

##          milkcompt
## gender viel wenig
##   Frau    23   469
##   Mann    25   623

#Chi-squared Test
chi_sq <- chisq.test(nova2$gender, nova2$milkcompt)
chi_sq

##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  nova2$gender and nova2$milkcompt
## X-squared = 0.28223, df = 1, p-value = 0.5952

#Fisher's Test
fisher.test(nova2$gender, nova2$milkcompt)
```

```
##
## Fisher's Exact Test for Count Data
##
## data: nova2$gender and nova2$milkcompt
## p-value = 0.5523
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
## 0.6536644 2.2746745
## sample estimates:
## odds ratio
## 1.221866
```

Ergebnisse

Der χ^2 -Test sagt uns, dass das Geschlecht und der wahrgenommene Milchkonsum nicht zusammenhängen. Es gibt keine signifikante Unterscheide zwischen dem Geschlecht und dem wahrgenommenen Milchkonsum ($\chi^2(1) = 0.282, p = 0.595$). Es sieht so aus, dass Männer leicht mehr angeben weniger Milch zu konsumieren (Tabelle 1). Die Ergebnisse müssen jedoch mit Vorsicht interpretiert werden, denn der χ^2 -Test gibt uns nur an, dass ein signifikanter Unterschied zwischen Geschlecht und wahrgenommener Milchkonsum vorliegt. Um die Unterschiede innerhalb der Gruppen (z.B. Geschlecht) festzustellen bedarf es weiterer Analysen z. B. Odds Ratios oder einer mehrfaktorieller ANOVA mit anschliessenden Post-hoc Tests (siehe Statistik 3).

```
## 'summarise()' regrouping output by 'gender' (override with '.groups' argument)
```

Table 1: Wahrgenommener Milchkonsum nach Geschlecht

| Geschlecht | wahr. Milchkonsum | absolute Werte | wahr. Milchkonsum (%) |
|------------|-------------------|----------------|-----------------------|
| Frau | viel | 23 | 4.7 |
| Frau | wenig | 469 | 95.3 |
| Mann | viel | 25 | 3.9 |
| Mann | wenig | 623 | 96.1 |

Musterloesung Aufgabe 1.2: t-Test

Meine Empfehlung Kapitel 2 von Manny Gimond

Null- und Alternativhypothese H_0 : Es gibt keine Unterschiede in den Verkaufszahlen zwischen Basis- und Interventionswochen.

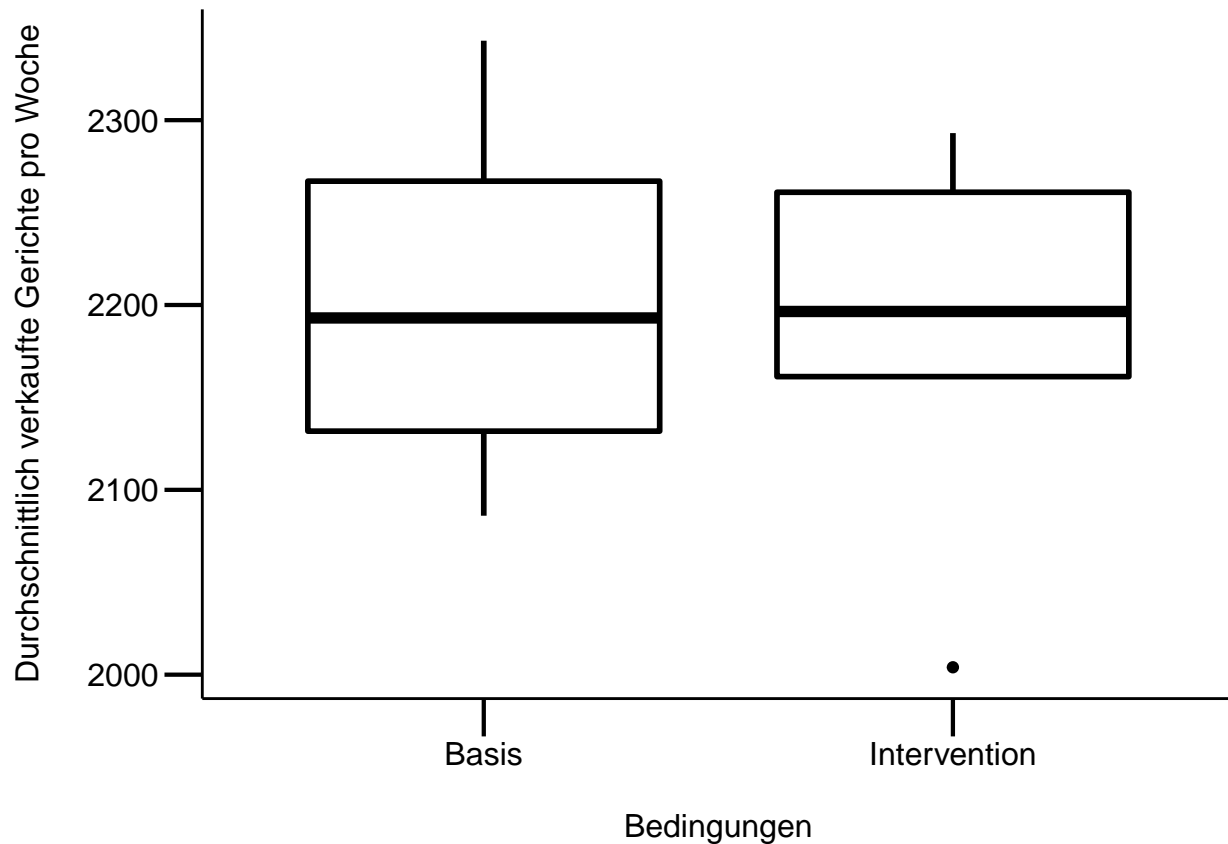
H_1 : Es gibt Unterschiede in den Verkaufszahlen zwischen Basis- und Interventionswochen.

*# Gemäss Aufgabenstellung müsset die Daten zuerst nach Kalenderwochen "week"
und Bedingungen "condition" zusammengefasst werden*

```
df <- nova %>%
  group_by(week, condit) %>%
  summarise(tot_sold = n())
```

```
## 'summarise()' regrouping output by 'week' (override with '.groups' argument)
```

```
# überprüft die Voraussetzungen für einen t-Test
ggplot(df, aes(x = condit, y= tot_sold)) + # achtung 0 Punkt fehlt
  geom_boxplot(fill = "white", color = "black", size = 1) +
  labs(x="\nBedingungen", y="Durchschnittlich verkaufte Gerichte pro Woche\n") +
  mytheme
```



```
# Auf den ersten Blick scheint es keine starken Abweichungen zu einer
# Normalverteilung zu geben resp. es sind keine extremen schiefen Verteilungen
# ersichtlich (vgl. Skript Statistik 2)
```

```
# führt einen t-Tests durch;
# es wird angenommen, dass die Verkaufszahlen zwischen den Bedingungen
# unabhängig sind
```

```
t_test <- t.test(tot_sold~condit, data=df)
```

```
#alternative Formulierung
```

```
t.test(df[df$condit == "Basis", ]$tot_sold,
       df[df$condit == "Intervention", ]$tot_sold)
```

```
##
```

```
## Welch Two Sample t-test
```

```
##
```

```
## data: df[df$condit == "Basis", ]$tot_sold and df[df$condit == "Intervention", ]$tot_sold
```

```
## t = 0.27168, df = 9.9707, p-value = 0.7914
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -115.2743  147.2743
## sample estimates:
## mean of x mean of y
##      2203      2187
```

Methoden

Ziel war es die aggregierten Verkaufszahlen zwischen den Interventions- und Basiswochen zu vergleichen. Die Annahme war, dass die wöchentlichen Verkaufszahlen unabhängig sind. Daher können die mittleren Verkaufszahlen pro Woche zwischen den beiden Bedingungen mittels t-Test geprüft werden. Obwohl die visuelle Inspektion keine schwerwiegenden Verletzungen der Modelvoraussetzung zeigte, wurde einen Welch t-Test gerechnet.

Ergebnisse

In den Basiswochen werden mehr Gerichte pro Woche verkauft als in den Interventionswochen (siehe Abbildung 1). Die wöchentlichen Verkaufszahlen zwischen den Bedingungen (Basis oder Intervention) unterscheiden sich gemäss Welch t-Test jedoch nicht signifikant ($t(10) = 0.272$, $p = 0.791$).

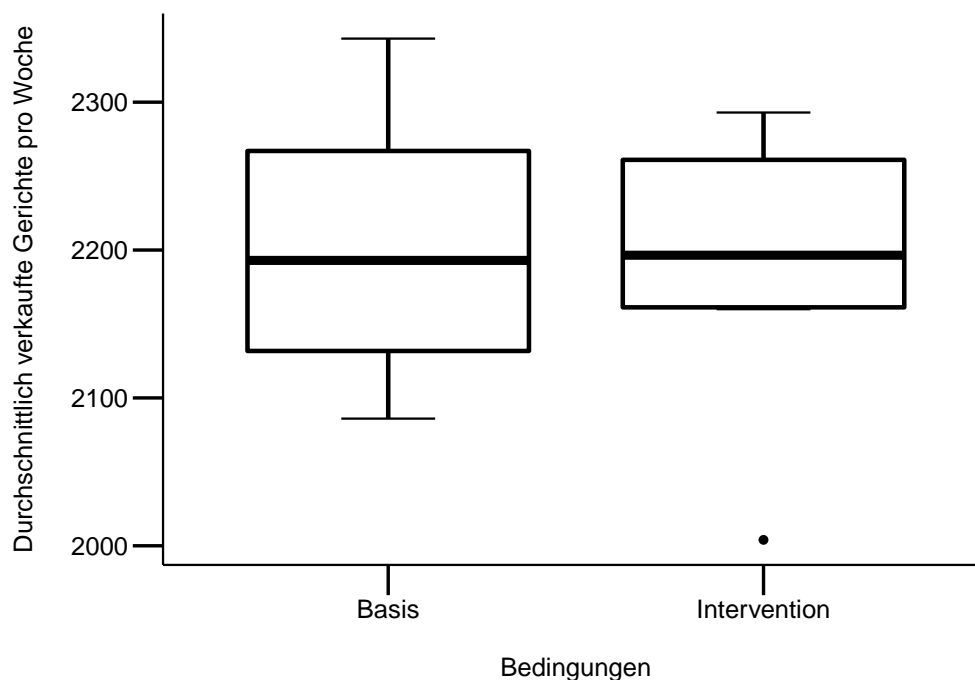


Figure 1: Die wöchentlichen Verkaufszahlen für die Interventions- und Basiswochen unterscheiden sich nicht signifikant.