Einfaktorielle Varianzanalyse ohne Messwiederholung

Code

Hide

```
library(car)
library(ggplot2)
library(dplyr)
library(sjstats)
```

Hide

```
# Die Datei "tips.csv" aus Aufgabe 3 kann auch für diese Aufgabe verwendet werden
df_tips <- read.csv("tips.csv")
```

Umsortieren der Wochentage

Hide

```
df_tips$day<-factor(df_tips$day, levels=c("Thur", "Fri", "Sat", "Sun"))
```

1) Hypothese

H1: Es gibt einen Unterschied der Mittelwerte des “Trinkgelds” zwischen den verschiedenen Ausprägungen der Gruppen (Wochentage: Do, Fr, Sa, So).
H0: Es gibt keinen Unterschied der Mittelwerte des “Trinkgelds” zwischen den verschiedenen Ausprägungen der Gruppen (Wochentage: Do, Fr, Sa, So).

2) Voraussetzungen

- Die unabhängige Variable ist nominal- oder ordinalskaliert. => Erfüllt, weil “Wochentage” ist nominalskaliert
- Die abhängige Variable ist intervalskaliert. => Erfüllt, weil “Trinkgeld” ist metrisch
- Die durch die unabhängige Variable gebildeten Gruppen sind unabhängig. => Erfüllt, weil sich die Trinkgelder jeweils nur auf einen Besuch des Restaurants beziehen, der an einem Tag stattgefunden hat
- Die abhängige Variable ist normalverteilt innerhalb jeder der Gruppen (Ab > 25 ProbandInnen pro Gruppe sind Verletzungen in der Regel unproblematisch) => Erfüllt, weil Anzahl der ProbandInnen pro Gruppe ist größer als 25 außer bei der Gruppe “Freitag”. Am Freitag (sowie an den anderen Tagen auch) sind die Trinkgelder annähernd normalverteilt.
- Homogenität der Varianzen: Die Gruppen stammen aus Grundgesamtheiten mit annähernd identischen Varianzen der abhängigen Variablen => Levene-Test (andernfalls Welch-Korrektur)

3) Grundlegende Konzepte

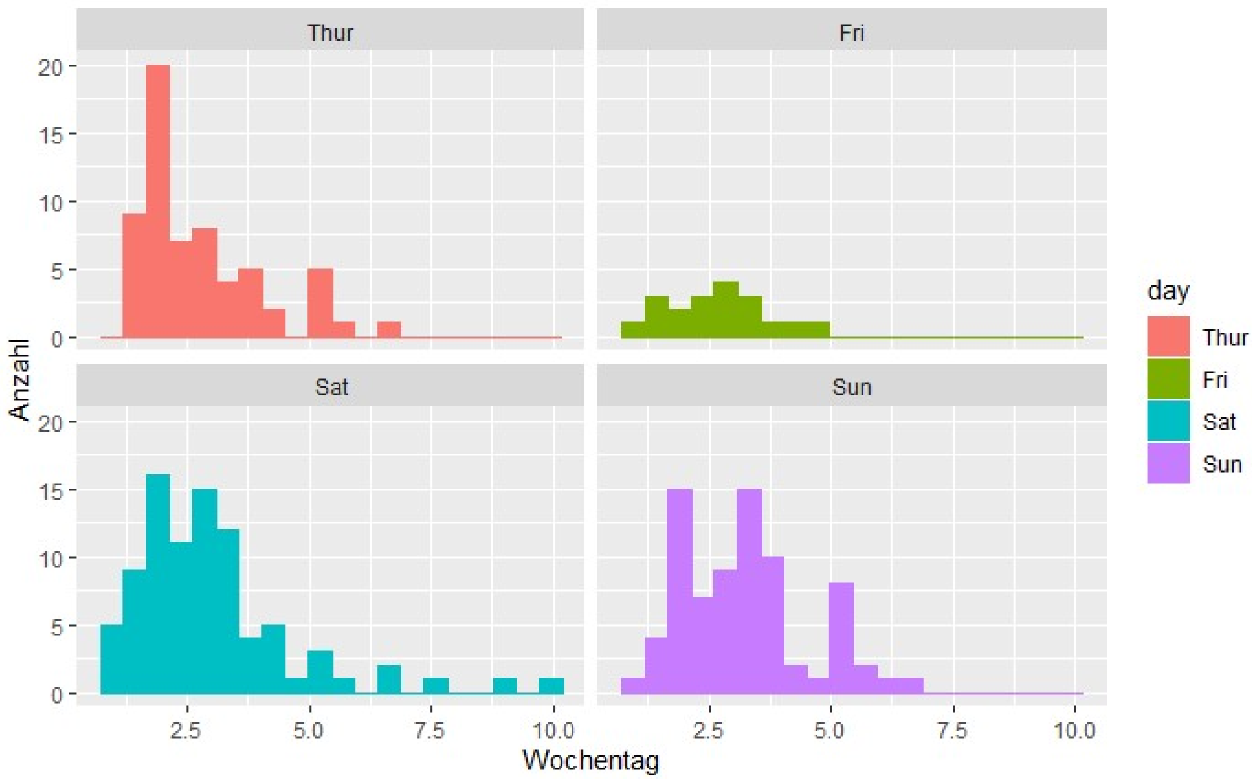
Die Varianzanalyse ist daran interessiert, die Unterschiede zwischen den beobachteten Werten und dem Gesamtmittelwert der Stichprobe zu erklären. Die Grundidee ist, dass jeder Messwert in drei Anteile zerlegt werden kann:
- "Grundniveau": Dies ist der Mittelwert der gesamten Stichprobe (ungeachtet der Gruppenzugehörigkeit), der sogenannte "Gesamtmittelwert". Im Beispiel ist dieses Ausgangsniveau der Gesamtmittelwert des Trinkgeldes. Dieser ist für die Berechnung der Varianzanalyse zwar wichtig, doch die Varianzanalyse versucht nicht, diesen Gesamtmittelwert zu erklären.
- Effekt der Gruppenzugehörigkeit: Dieser Anteil repräsentiert die Gruppenzugehörigkeit und entspricht der Differenz zwischen dem Gesamtmittelwert und dem jeweiligen Gruppenmittelwert.
- Fehlerwert: Dieser Anteil umfasst Messfehler, doch daneben fliessen auch alle individuellen Abweichungen vom Gesamtmittelwert ein, die nicht durch die Gruppenzugehörigkeit erklärt werden können. Je mehr Variation durch die Gruppenzugehörigkeit erklärt wird, desto höher fällt der F-Wert aus (da $MS_{zwischen}$ ein Mass für die erklärte Varianz darstellt, während $MS_{innerhalb}$ ein Mass für die Residualvarianz des Modells darstellt). Dieser F-Wert wird mit dem kritischen Wert auf einer durch die Freiheitsgrade $df_{zwischen}$ und $df_{innerhalb}$ charakterisierten F-Verteilung verglichen. Ist der F-Wert höher als der kritische Wert, so ist der Test signifikant.

4) Deskriptive Statistik

Normalverteilung => Prüfung durch Histogramm

Hide

```
df_tips %>%
  group_by(day) %>%
  ggplot(aes(tip, color=day)) +
  geom_histogram(aes(fill=day), bins=20) +
  facet_wrap(~day) +
  labs(x="Wochentag", y="Anzahl")
```

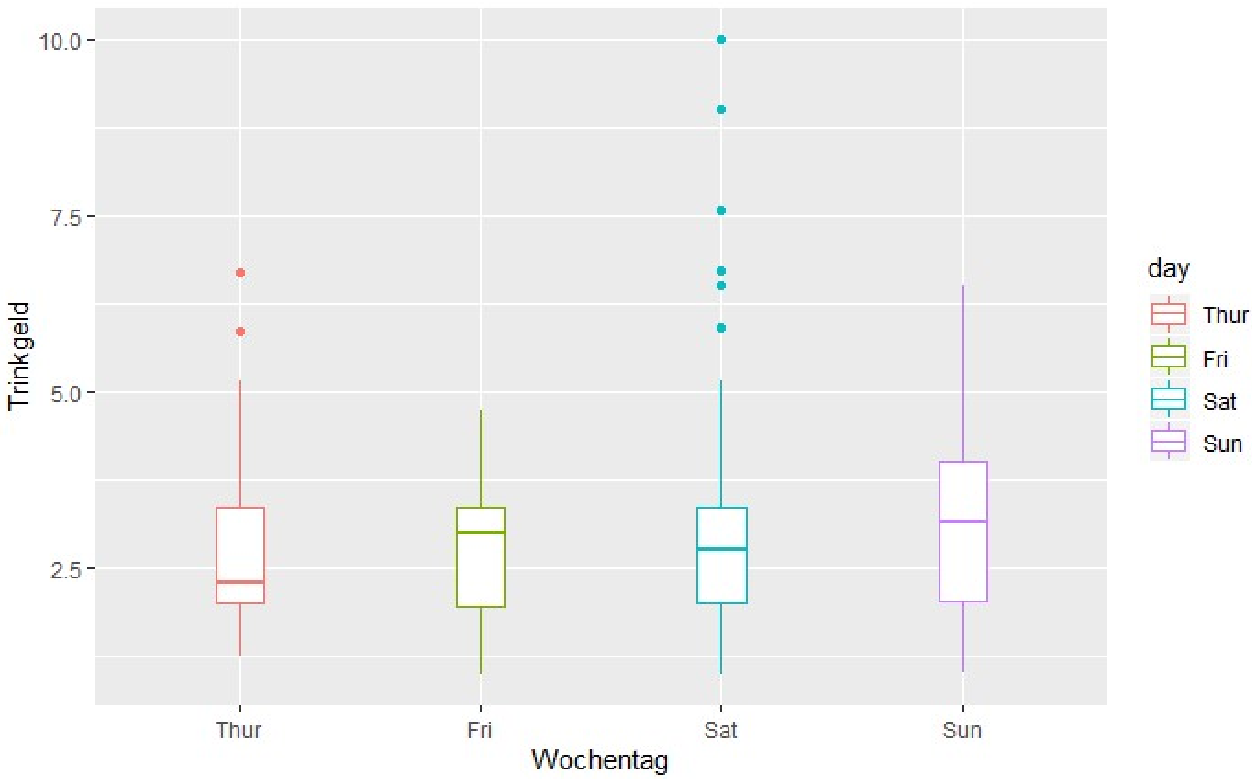


Die Daten sind normalverteilt, wenn auch nicht perfekt.

Boxplots

Hide

```
ggplot(df_tips, aes(x=day, y=tip, color=day)) +
  geom_boxplot(width=0.2) +
  labs(x="Wochentag", y="Trinkgeld")
```



Die Boxplots der verschiedenen Gruppen ("Wochentag") zeigen am Donnerstag ("Thur") sowie am Samstag ("Sat") wenige Ausreißer. Die Mediane der verschiedenen Gruppen zeigen geringe Unterschiede, ebenso wie der Bereich des 2. und 3. Quartils. Am niedrigsten liegt der Median der Gruppe "Donnerstag" ("Thur"), am höchsten der der Gruppe "Sonntag". Der IQR der Gruppen "Donnerstag", "Freitag" und "Samstag" ist augenscheinlich gleich groß (Überlappung der verteilung). Lediglich am Sonntag zeigt sich ein größerer IQR.

Tabelle verschiedener statistischer Parameter der einzelnen Gruppen

Hide

```
df_tips %>%
  group_by(day) %>%
  summarize(Anzahl=n(), Mittelwert=mean(tip), Median=median(tip), Standardabweichung=s
d(tip), IQR=IQR(tip)) %>%
  mutate_if(is.numeric, round, 2)
```

day <fctr>	Anzahl <dbl>	Mittelwert <dbl>	Median <dbl>	Standardabweichung <dbl>	IQR <dbl>
Thur	62	2.77	2.30	1.24	1.36
Fri	19	2.73	3.00	1.02	1.41
Sat	87	2.99	2.75	1.63	1.37

day	Anzahl	Mittelwert	Median	Standardabweichung	IQR
<fctr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
Sun	76	3.26	3.15	1.23	1.96

4 rows

Auf den ersten Blick fällt auf, dass die Gruppe “Freitag” nur sehr wenige Elemente (n=19 < 25!) aufweist. Weil die Elemente aber normalverteilt sind kann mit der Untersuchung fortgefahren werden. Außerdem fällt auf, dass sich Mittelwerte und Mediane teilweise deutlich unterscheiden. Wie bereits im Boxplot zu erkennen war, hat der IQR der Gruppen “Donnerstag”, “Freitag” und “Samstag” ähnliche Werte. Den höchsten Mittelwert und Median hat die Gruppe “Sonntag” (M=3.26, SD=1,23, n=76), den geringsten Median hat die Gruppe “Donnerstag” (Median=2.3, SD=1.24, n=62).

5) Prüfung der Varianzhomogenität (Levene-Test)

Hide

```
leveneTest(df_tips$tip ~ df_tips$day, center="mean")
```

```
Levene's Test for Homogeneity of Variance (center = "mean")
       Df F value Pr(>F)
group   3  0.5161 0.6716
      240
```

Der Levene-Test ist nicht signifikant (F(3, 240)=0.52, p-Value=0.67 > 0.05). Das bedeutet, dass Varianzhomogenität in den Daten vorliegt. => *Eine Welch-Korrektur ist nicht notwendig!*

6) Ergebnisse der einfaktoriellen Varianzanalyse

ANOVA ohne Welch-Korrektur:

Hide

```
result_anova_ohne <- aov(df_tips$tip ~ df_tips$day)
summary(result_anova_ohne)
```

```
             Df Sum Sq Mean Sq F value Pr(>F)
df_tips$day   3    9.5   3.175   1.672  0.174
Residuals   240  455.7   1.899
```

Das Gesamtmodell ist nicht signifikant (F(3, 240)=1.67, p=0.174 > 0.05).

7) Post-hoc-Tests

Wird nur ein t-Test mit einem Signifikanzlevel von .05 durchgefuehrt, so betraegt die Wahrscheinlichkeit des Nicht-Eintreffens des Alpha-Fehlers 95%. Werden jedoch sechs solcher Paarvergleiche vorgenommen, so betraegt die Nicht-Eintreffens-Wahrscheinlichkeit des Alpha-Fehlers (.95)6 = .735.
Um die Wahrscheinlichkeit des Eintreffens des Alpha-Fehlers zu bestimmen, wird 1 - .735 = .2649 gerechnet. Die Wahrscheinlichkeit des Eintreffens des Alpha-Fehlers liegt somit bei 26.49%. Diese Fehlerwahrscheinlichkeit wird als "Familywise Error Rate" bezeichnet.
Um dieses Problem zu beheben kann zum Beispiel der Tukey-Test angewendet werden. Hierbei wird durch die Anzahl der Paarvergleiche dividiert. Im hier aufgefuehrten Fall ist dies .05/6 = .0083. Das heisst, jeder Test wird gegen ein Niveau von .0083 geprueft.

Anzahl der Kombinationen ist: 6 Jeder Test wird gegen ein Niveau von 0.05/6 = 0.0083 geprüft.

Hide

```
TukeyHSD(result_anova_ohne)
```

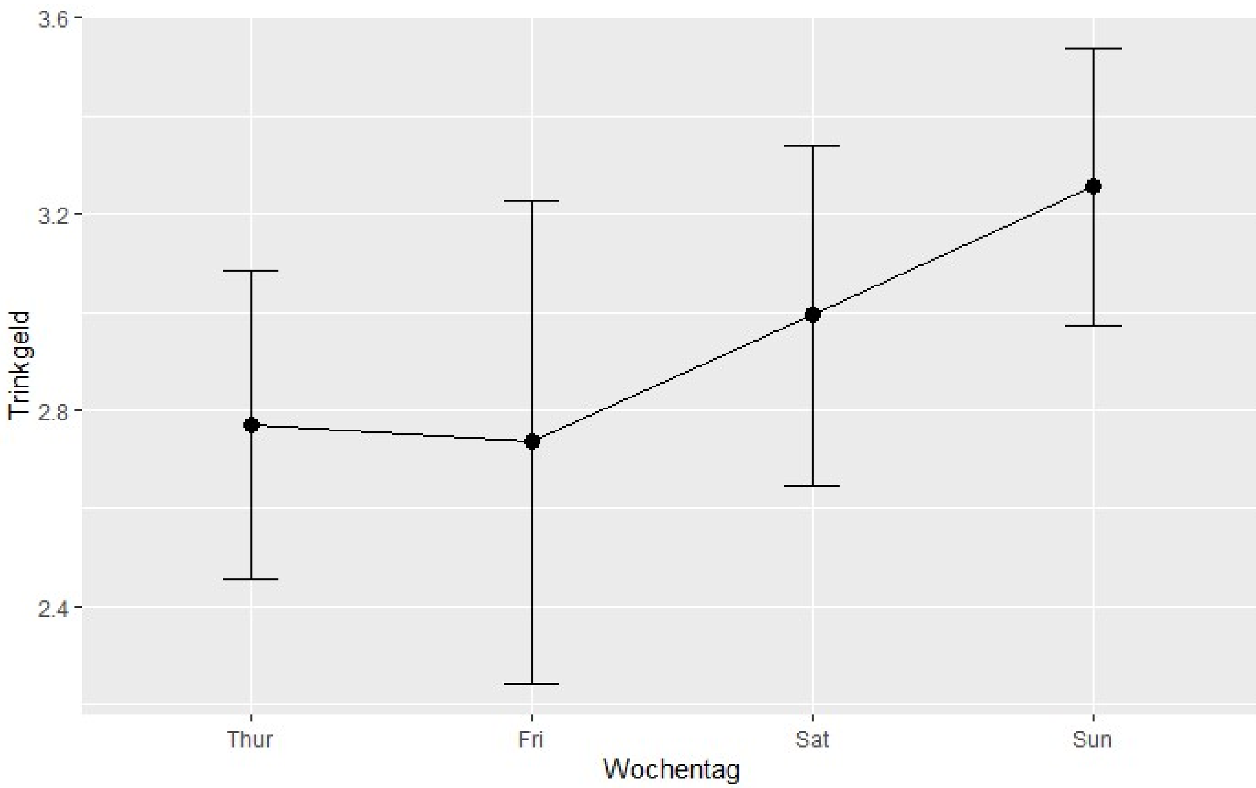
```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = df_tips$tip ~ df_tips$day)

$`df_tips$day`
                diff        lwr       upr     p adj
Fri-Thur -0.03671477 -0.9715049 0.8980753 0.9996235
Sat-Thur  0.22165184 -0.3708394 0.8141430 0.7678581
Sun-Thur  0.48367997 -0.1263921 1.0937520 0.1724212
Sat-Fri   0.25836661 -0.6443694 1.1611026 0.8806455
Sun-Fri   0.52039474 -0.3939763 1.4347658 0.4558054
Sun-Sat   0.26202813 -0.2976929 0.8217492 0.6203822
```

Es wird deutlich, dass es keine signifikanten Unterschiede zwischen den Gruppen "Donnerstag", "Freitag", "Samstag" und "Sonntag" gibt; alle p-Werte des Post-Hoc-Test liegen über dem Niveau von 0.0083.
Der deutlich geringste p-Wert ist der p-Wert zwischen den Gruppen "Sonntag" und "Donnerstag". Ein Blick auf das Boxplot-Diagramm zeigt, dass sich die Mediane der beiden Gruppen am deutlichsten unterscheiden. Es können keine Gruppen gebildet werden.

8) Plot der Mittelwerte / Profildiagramm

Hide

```
ggplot(df_tips, aes(x=df_tips$day, y=df_tips$tip, group=1)) +
  stat_summary(fun.y=mean, geom="point", size=3) +
  stat_summary(fun.y=mean, geom="line") +
  stat_summary(fun.data=mean_cl_normal, geom="errorbar", width=.2, size=.25) +
  labs(x="Wochentag", y="Trinkgeld")
```



9) Berechnung der Effektstärke

Das partielle Eta-Quadrat

Hide

```
eta <- eta_sq(result_anova_ohne, partial=TRUE)
eta
```

```
         term partial.etasq
1 df_tips$day          0.02
```

Das partielle Eta-Quadrat beträgt 0.02. Das heisst, es wird 2% der Variation des Trinkgeldes durch die verschiedenen Gruppen (“Donnerstag”, “Freitag”, “Samstag” und “Sonntag”) erklärt.

Die Effektstärke

Hide

```
eff <- sqrt(eta$partial.etasq / (1-eta$partial.etasq))
eff
```

```
[1] 0.1445837
```

Um zu beurteilen, wie gross dieser Effekt ist, kann man sich an der Einteilung von Cohen (1988) orientieren:
f = .10 entspricht einem schwachen Effekt
f = .25 entspricht einem mittleren Effekt
f = .40 entspricht einem starken Effekt
Damit entspricht die Effektstärke von 0.14 einem schwachen bis mittleren Effekt.

10) Aussage

Die unterschiedlichen Gruppen haben keinen signifikanten Einfluss auf das Trinkgeld ($F(3, 240)=1.67$, $p=0.174 > 0.05$, $n=244$). 2% der Streuung des Trinkgeldes um den Gesamtmittelwert können durch die verschiedenen Gruppen ("Donnerstag", "Freitag", "Samstag" und "Sonntag") erklärt werden.
Die Effektstärke nach Cohen (1988) von 0.14 entspricht einem schwachen Effekt bis mittleren Effekt.
Post-hoc Tests mit Tukey zeigen, dass es keine signifikanten Unterschiede zwischen den Gruppen gibt.
Den höchsten Mittelwert und Median hat die Gruppe "Sonntag" ($M=3.26$, $SD=1,23$, $n=76$), den geringsten Median hat die Gruppe "Donnerstag" ($Median=2.3$, $SD=1.24$, $n=62$).

H0 wird angenommen. (Es gibt keinen signifikanten Unterschied der Mittelwerte)