Lösungen zu den Aufgaben

1. **Aufgabe**

   In einer Lostrommel befinden sich "sehr viele" Lose, davon ein Anteil p Treffer (und $1 - p$ Nieten), mit zunächst $p = 0.01$.

   Sie kaufen $n = 10$ Lose.

   a. Wie groß ist die Wahrscheinlichkeit für genau $k = 0, 1, ..., 10$ Treffer?
   b. Sagen wir, Sie haben 3 Treffer in den 10 Losen. Yeah! Jetzt sei p unbekannt und Sie sind indifferent zu den einzelnen Werten von p. Visualisieren Sie die Posteriori-Wahrscheinlichkeitsverteilung mit ca. 100 Gridwerten. Was beobachten Sie?
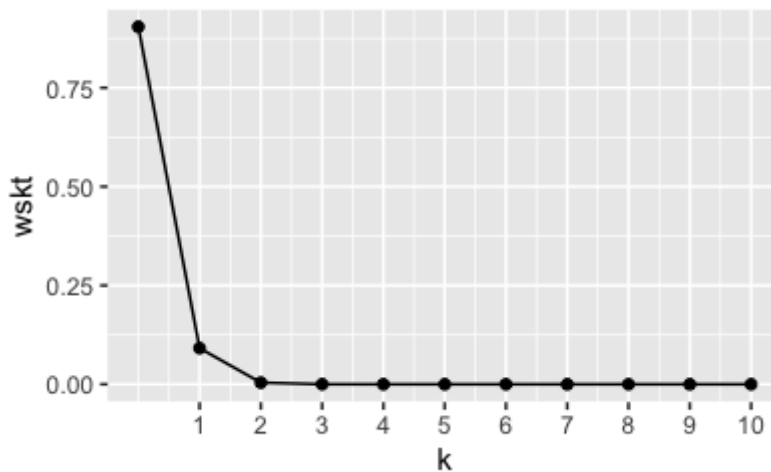   c. Variieren Sie n, aber halten Sie die Trefferquote bei 1/3. Was beobachten Sie?

   Nutzen Sie die Gittermethode. Treffen Sie Annahmen, wo nötig.

   Lösung

   a. Wie groß ist die Wahrscheinlichkeit für genau $k = 0, 1, ..., 10$ Treffer?

```
d_a <-
  tibble(
    k = 0:10,
    wskt = dbinom(k, size = 10, prob = .01))

d_a %>%
  ggplot() +
  aes(x = k, y = wskt) +
  geom_point() +
  geom_line() +
  scale_x_continuous(breaks = 1:10)
```


k	wskt
0	9.04×10^{-1}

k	wskt
1	9.14×10^{-2}
2	4.15×10^{-3}
3	1.12×10^{-4}
4	1.98×10^{-6}
5	2.40×10^{-8}
6	2.02×10^{-10}
7	1.16×10^{-12}
8	4.41×10^{-15}
9	9.90×10^{-18}
10	1.00×10^{-20}

b. Sagen wir, Sie haben 3 Treffer in den 10 Losen. Yeah! Jetzt sei p unbekannt und Sie sind indifferent zu den einzelnen Werten von p. Visualisieren Sie die Posteriori-Wahrscheinlichkeitsverteilung mit ca. 100 Gridwerten. Was beobachten Sie?

```
d2 <-
  tibble(
    p_grid = seq(0, 1, by = 0.01),
    prior = 1,
    Likelihood = dbinom(x = 3, size = 10, prob = p_grid),
    unstand_post = prior * Likelihood,
    std_post = unstand_post / sum(unstand_post)
  )

d2 %>%
  ggplot() +
  aes(x = p_grid, y = std_post) +
  geom_point() +
  geom_line()
```
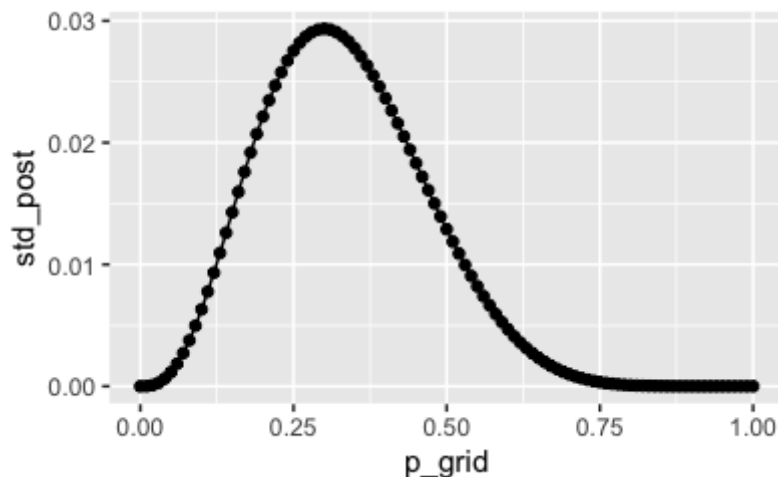
Der Modus liegt bei ca 1/3. Der Bereich plausibler Werte für p liegt ca. zwischen 0.1 und und 0.7, grob visuell geschätzt. Mehr dazu im nächsten Kapitel.

   c. Variieren Sie n, aber halten Sie die Trefferquote bei 1/3. Was beobachten Sie?

```
# n = 2
d3 <-
  tibble(
    p_grid = seq(0,1, by = 0.01),
    prior = 1,
    Likelihood = dbinom(x = 2, size = 6, prob = p_grid),
    unstand_post = prior * Likelihood,
    std_post = unstand_post / sum(unstand_post)
  )

d3 %>%
  ggplot() +
  aes(x = p_grid, y = std_post) +
  geom_point() +
  geom_line() +
  labs(title = "n=20")


# n = 20
d4 <-
  tibble(
    p_grid = seq(0,1, by = 0.01),
    prior = 1,
    Likelihood = dbinom(x = 20, size = 60, prob = p_grid),
    unstand_post = prior * Likelihood,
    std_post = unstand_post / sum(unstand_post)
  )

d4 %>%
  ggplot() +
  aes(x = p_grid, y = std_post) +
  geom_point() +
  geom_line() +
  labs(title = "n = 20")

# n = 200
d5 <-
  tibble(
    p_grid = seq(0,1, by = 0.01),
    prior = 1,
    Likelihood = dbinom(x = 200, size = 600, prob = p_grid),
    unstand_post = prior * Likelihood,
    std_post = unstand_post / sum(unstand_post)
  )

d5 %>%
  ggplot() +
  aes(x = p_grid, y = std_post) +
  geom_point() +
  geom_line() +
  labs(title = "n = 20")
```
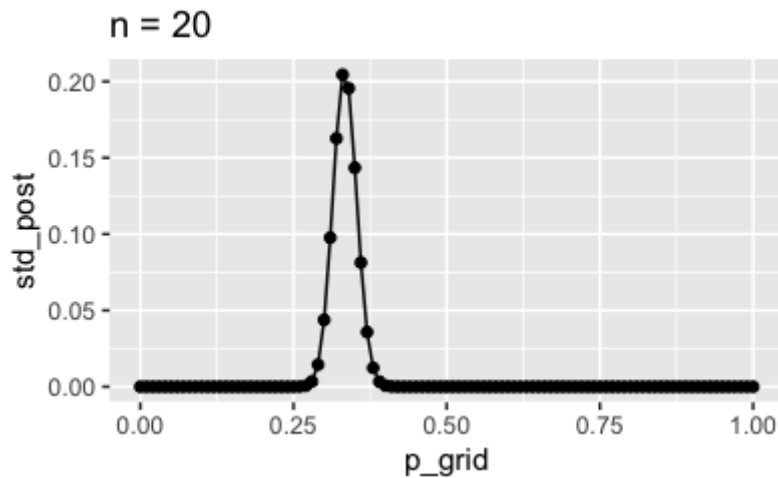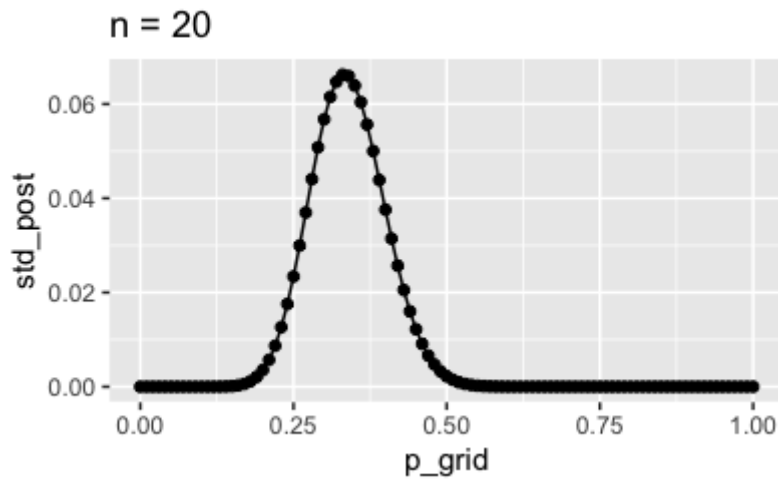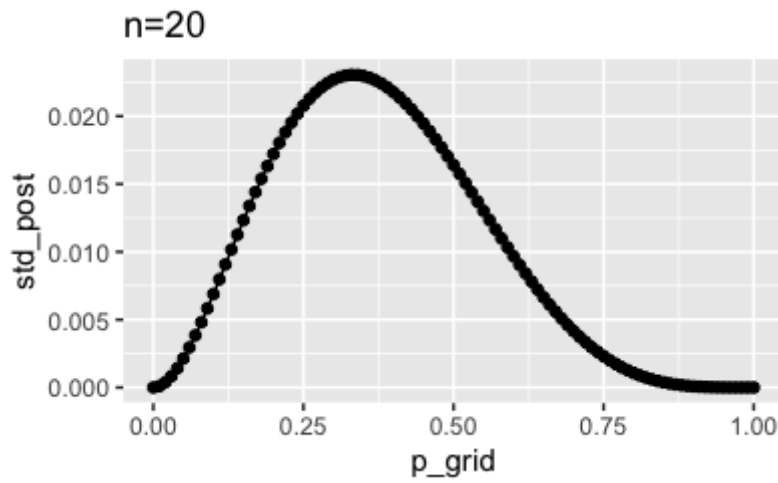
n=20

n = 20

n = 20

Der Modus und andere Maße der zentralen Tendenz bleiben gleich; die Streuung wird geringer.

2. **Aufgabe**

Die Binomialverteilung wird in Lehrbüchern häufig mit Münzwürfen motiviert. In *Statistical Rethinking* muss ein Globus herhalten (also ein Zufallsexperiment mit den Ergebnissen *Wasser* und *Land* unter dem Zeigefinger).

Die Beispiele sind ja gut und schön. Aber was hat das mit der Praxis zu tun? Gute Frage. Nennen Sie Beispiele aus Berufsfeldern der AWM, für die die Binomialverteilung relevant ist.

Sie müssen nichts rechnen, nur Beispiele nennen.

Lösung

Zur Erinnerung: Die Inferenzstatistik macht Aussagen bzgl. einer *Population*, nicht einer Stichprobe. Solche Aussagen sind *ungewiss*, also mit einer Unsicherheit behaftet, da wir nicht die ganze Population kennen. Aber die Daten der Stichprobe werden als Grundlage der Schätzung herangezogen.

- Auswahl geeigneter Kandidatis in einem Assessment-Verfahren. Man hat $n = 40$ Bewerbis, und die Wahrscheinlichkeit geeigneter Kandidatis liege bei $p = 10$. Welche Spannweite an geeigneten Bewerbis kann man erwarten?

- Social Influencing. Sie posten 100 Videoclips; davon werden 9 viral. Welche Spannweite plausibler Werte für eine Erfolgsquote kann man zugrunde legen?

- App-Wartung. Sie prüfen eine Anzahl ($n = 42$) alter Apps, aus einer früheren Kampagne. Sie finden, dass $k = 19$ noch funktionieren. Welche Quote an "technisch veraltet" muss man in der Population erwarten, und in welchem Bereich könnte sich diese Quote bewegen?

- Schulungsprogramm. Sie entwickeln ein Schulungsprogramm, das im großen Stil in einer Firma eingesetzt werden soll; mehrere Tausend Personen sollen das Programm durchlaufen. In einer Pilotstudie mit $n = 90$ Personen erreichen $k = 42$ nicht das Lernziel. Welche Parameterwerte für p (Lernziel erreicht) sind plausibel?

3. **Aufgabe**

This question is taken from McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan* (2. Ed.). Taylor and Francis, CRC Press.

2E4. The Bayesian statistician Bruno de Finetti (1906–1985) began his 1973 book on probability theory with the dedication: "PROBABILITY DOES NOT EXIST." The capitals appeared in the original, so I imagine de Finetti wanted us to shout this statement. What he meant is that probability is a device for describing uncertainty from the perspective of an observer with limited knowledge; it has no objective reality. Discuss the globe tossing example from the chapter, in light of this statement. What does it mean to say "the probability of water is 0.7"?

Lösung

The solution is taken from [this source](#).

The idea is that probability is only a subjective perception of the likelihood that something will happen. In the globe tossing example, the result will always be either "land" or "water" (i.e., 0 or 1). When we toss the globe, we don't know what the result will be, but we know it will always be "land" or "water." To express our uncertainty in the outcome, we use probability. Because we know that water is more likely than land, we may say that the probability of "water" is 0.7;

however, we'll never actually observe a result of 0.7 waters, or observe any probability. We will only ever observe the two results of "land" and "water."

4. **Aufgabe**

This question is taken from McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan* (2. Ed.). Taylor and Francis, CRC Press.

2M1. Recall the globe tossing model from the chapter. Compute and plot the grid approximate posterior distribution for each of the following sets of observations. In each case, assume a uniform prior for *p*.

1. WWW
2. WWWL
3. LWWLWWW

Lösung

The solution is taken from [this source](#).

```
library(tidyverse)

dist <-
  tibble(
    # Gridwerte bestimmen:
    p_grid = seq(from = 0, to = 1, length.out = 20),
    # Priori-Wskt bestimmen:
    prior = rep(1, times = 20)) %>%
  mutate(
    # Likelihood berechnen:
    likelihood_1 = dbinom(3, size = 3, prob = p_grid),  # WWW
    likelihood_2 = dbinom(3, size = 4, prob = p_grid),  # WWWL
    likelihood_3 = dbinom(5, size = 7, prob = p_grid),  # LWWLWWW
    # unstand. Posterior-Wskt:
    unstand_post_1 = likelihood_1 * prior,
    unstand_post_2 = likelihood_2 * prior,
    unstand_post_3 = likelihood_3 * prior,
    # stand. Post-Wskt:
    std_post_1 = unstand_post_1 / sum(unstand_post_1),
    std_post_2 = unstand_post_2 / sum(unstand_post_2),
    std_post_3 = unstand_post_3 / sum(unstand_post_3)
    )
```
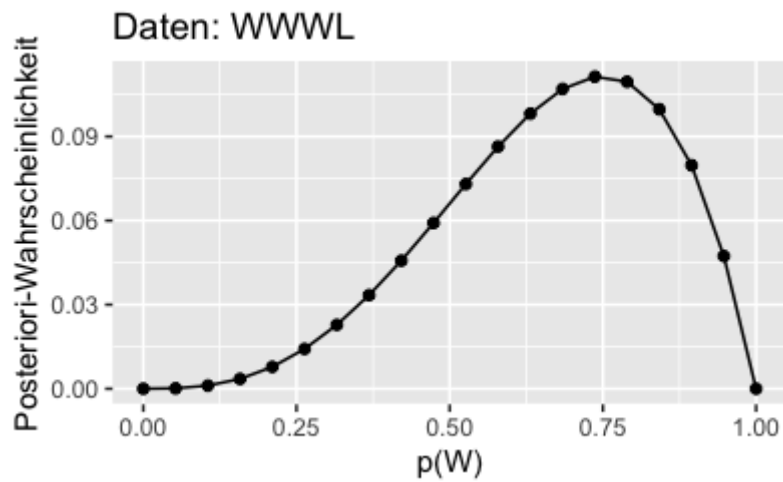
Jetzt können wir das Diagramm zeichnen:

```
ggplot(dist) +
  aes(x = p_grid, y= std_post_1) +
  geom_line()+
  geom_point() +
  labs(x = "p(W)",
       y = "Posteriori-Wahrscheinlichkeit",
       title = "Daten: WWW")
```

Daten: WWW


```
ggplot(dist) +
  aes(x = p_grid, y= std_post_2) +
  geom_line()+
  geom_point() +
  labs(x = "p(W)",
       y = "Posteriori-Wahrscheinlichkeit",
       title = "Daten: WWWL")
```

Daten: WWWL


```
ggplot(dist) +
  aes(x = p_grid, y= std_post_3) +
  geom_line()+
  geom_point() +
  labs(x = "p(W)",
       y = "Posteriori-Wahrscheinlichkeit",
       title = "Daten: LWWLWWW")
```
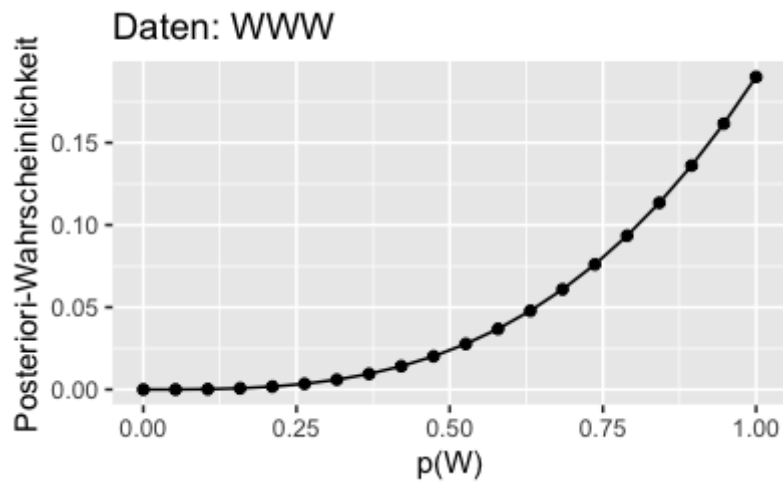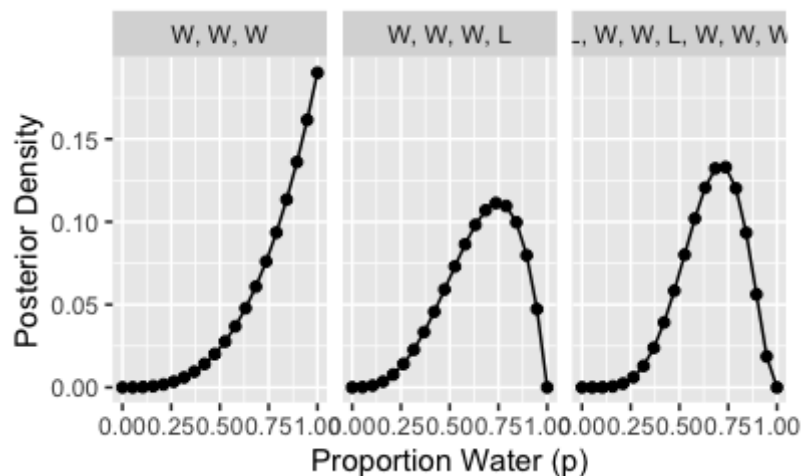
Daten: LWWLWWW
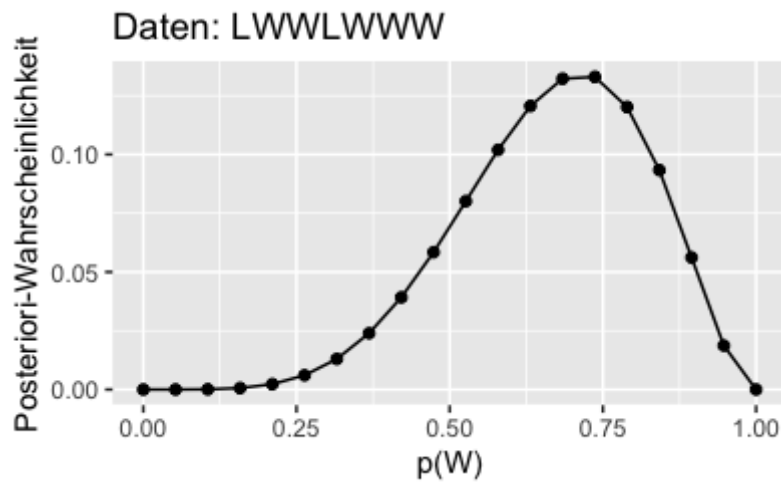
Etwas eleganter (und komplizierter) kann man es auch so in R schreiben (Quelle):

```
library(tidyverse)

dist <- tibble(p_grid = seq(from = 0, to = 1, length.out = 20),
               prior = rep(1, times = 20)) %>%
  mutate(likelihood_1 = dbinom(3, size = 3, prob = p_grid),
         likelihood_2 = dbinom(3, size = 4, prob = p_grid),
         likelihood_3 = dbinom(5, size = 7, prob = p_grid),
         across(starts_with("likelihood"), ~ .x * prior),
         across(starts_with("likelihood"), ~ .x / sum(.x))) %>%
  pivot_longer(cols = starts_with("likelihood"), names_to = "pattern",
               values_to = "posterior") %>%
  separate(pattern, c(NA, "pattern"), sep = "_", convert = TRUE) %>%
  mutate(obs = case_when(pattern == 1L ~ "W, W, W",
                         pattern == 2L ~ "W, W, W, L",
                         pattern == 3L ~ "L, W, W, L, W, W, W"))

ggplot(dist, aes(x = p_grid, y = posterior)) +
  facet_wrap(vars(fct_inorder(obs)), nrow = 1) +
  geom_line() +
  geom_point() +
  labs(x = "Proportion Water (p)", y = "Posterior Density")
```


5. **Aufgabe**

This question is taken from McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan* (2. Ed.). Taylor and Francis, CRC Press.

Recall the globe tossing model from the chapter. Compute and plot the grid approximate posterior distribution for each of the following sets of observations. In each case, assume a uniform prior for *p*.

Data:

1. WWW
2. WWWL
3. LWWLWWW

Now assume a prior for p that is equal to zero when p < 0.5 and is a positive constant when p ≥ 0.5. Again compute and plot the grid approximate posterior distribution for each of the sets of observations in the problem just above.

Lösung

The solution is taken from [this source](#).

```
library(tidyverse)

dist <-
  tibble(
    # Gridwerte bestimmen:
    p_grid = seq(from = 0, to = 1, length.out = 20),
    # Priori-Wskt bestimmen:
    prior  = case_when(
      p_grid < 0.5 ~ 0,
      p_grid >= 0.5 ~ 1)) %>%
  mutate(
    # Likelihood berechnen:
    likelihood_1 = dbinom(3, size = 3, prob = p_grid),
    likelihood_2 = dbinom(3, size = 4, prob = p_grid),
    likelihood_3 = dbinom(5, size = 7, prob = p_grid),
    # unstand. Posterior-Wskt:
    unstand_post_1 = likelihood_1 * prior,
    unstand_post_2 = likelihood_2 * prior,
    unstand_post_3 = likelihood_3 * prior,
    # stand. Post-Wskt:
    std_post_1 = unstand_post_1 / sum(unstand_post_1),
    std_post_2 = unstand_post_2 / sum(unstand_post_1),
    std_post_3 = unstand_post_3 / sum(unstand_post_1)
    )
```
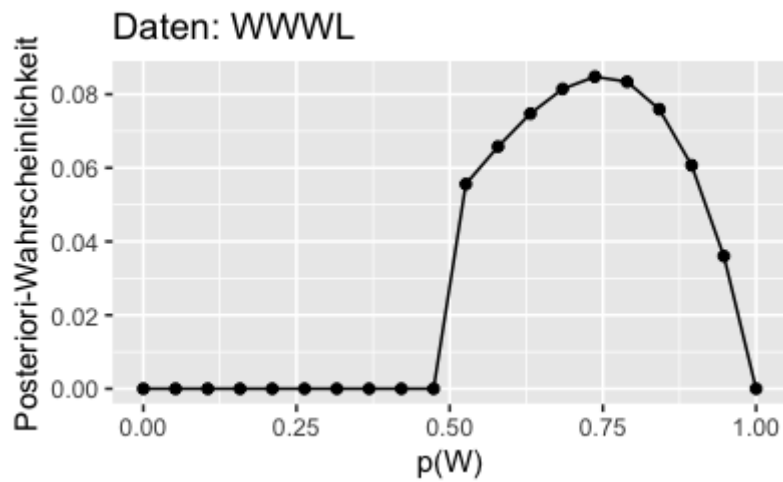
Jetzt können wir das Diagramm zeichnen:

```
ggplot(dist) +
  aes(x = p_grid, y= std_post_1) +
  geom_line()+
  geom_point() +
  labs(x = "p(W)",
       y = "Posteriori-Wahrscheinlichkeit",
       title = "Daten: WWW")
```

Daten: WWW


```
ggplot(dist) +
  aes(x = p_grid, y= std_post_2) +
  geom_line()+
  geom_point() +
  labs(x = "p(W)",
       y = "Posteriori-Wahrscheinlichkeit",
       title = "Daten: WWWL")
```

Daten: WWWL

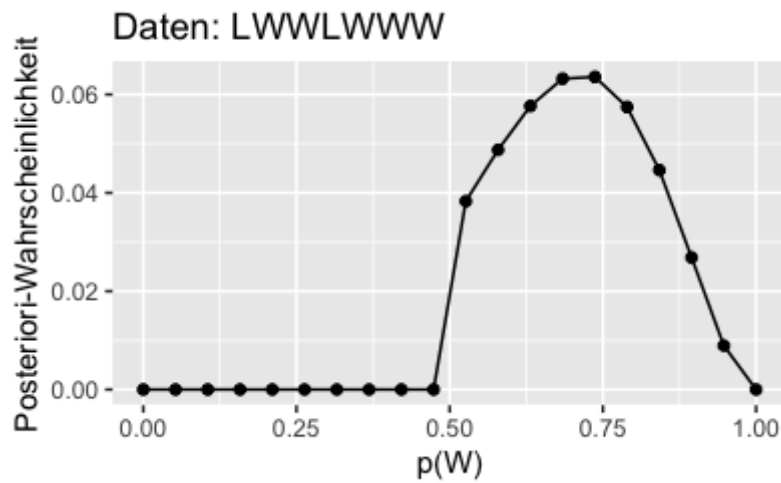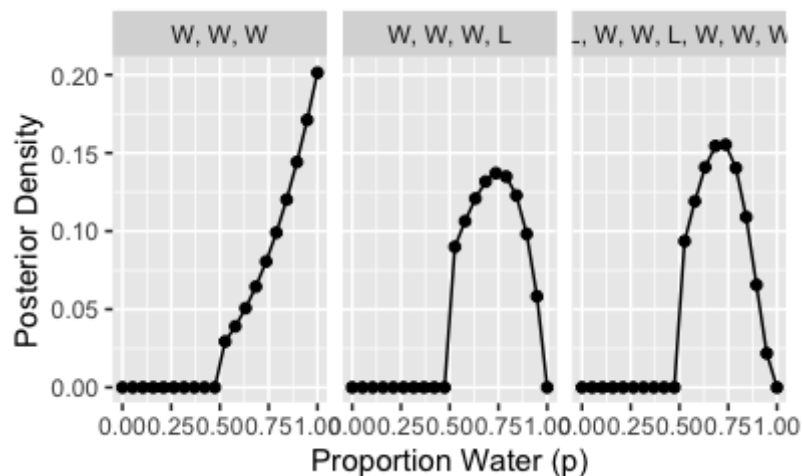

```
ggplot(dist) +
  aes(x = p_grid, y= std_post_3) +
  geom_line()+
  geom_point() +
  labs(x = "p(W)",
       y = "Posteriori-Wahrscheinlichkeit",
       title = "Daten: LWWLWWW")
```
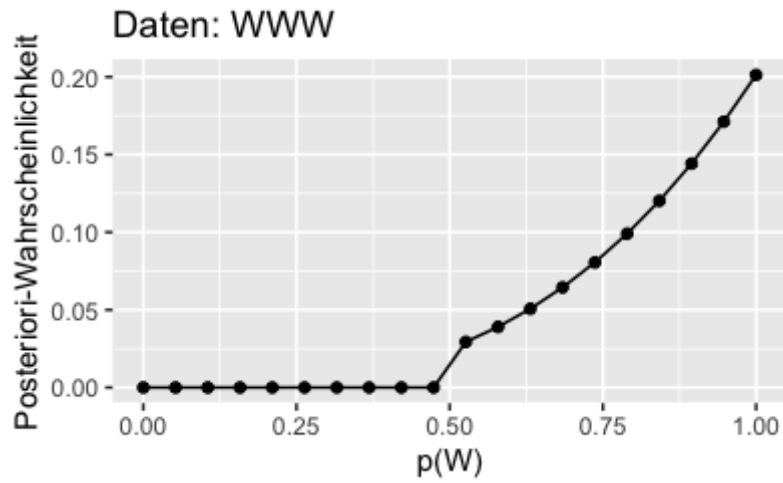
Daten: LWWLWWW

Etwas eleganter (und komplizierter) kann man es auch so in R schreiben (Quelle):

```
dist <- tibble(p_grid = seq(from = 0, to = 1, length.out = 20)) %>%
  mutate(prior = case_when(
    p_grid < 0.5 ~ 0L,
    TRUE ~ 1L),
    likelihood_1 = dbinom(3, size = 3, prob = p_grid),
    likelihood_2 = dbinom(3, size = 4, prob = p_grid),
    likelihood_3 = dbinom(5, size = 7, prob = p_grid),
    across(starts_with("likelihood"), ~ .x * prior),
    across(starts_with("likelihood"), ~ .x / sum(.x))) %>%
  pivot_longer(cols = starts_with("likelihood"), names_to = "pattern",
               values_to = "posterior") %>%
  separate(pattern, c(NA, "pattern"), sep = "_", convert = TRUE) %>%
  mutate(obs = case_when(pattern == 1L ~ "W, W, W",
                         pattern == 2L ~ "W, W, W, L",
                         pattern == 3L ~ "L, W, W, L, W, W, W"))

ggplot(dist, aes(x = p_grid, y = posterior)) +
  facet_wrap(vars(fct_inorder(obs)), nrow = 1) +
  geom_line() +
  geom_point() +
  labs(x = "Proportion Water (p)", y = "Posterior Density")
```


6. **Aufgabe**

This question is taken from McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan* (2. Ed.). Taylor and Francis, CRC Press.

2M3. Suppose there are two globes, one for Earth and one for Mars. The Earth globe is 70% covered in water. The Mars globe is 100% land. Further suppose that one of these globes—you don't know which—was tossed in the air and produced a "land" observatiion. Assume that each globe was equally likely to be tossed. Show that the posterior probability that the globe was the Earth, conditional on seeing "land" (Pr(Earth|land)), is 0.23.

Lösung

Zur Erinnerung:

$$\begin{aligned}
Pr(A) &= Pr(A \cap B) + Pr\left(A \cap A^C\right) \qquad \text{| bei disjunkten Ereignissen} \\
Pr(A \cap B) &= Pr(A|B) \cdot Pr(B) \\
Pr\left(A \cap B^C\right) &= Pr\left(A|B^C\right) \cdot Pr\left(B^C\right)
\end{aligned}$$

The solution is taken from this source.

```
# probability of land, given Earth:
p_le <- 0.3

# probability of land, given Mars:
p_lm <- 1.0

# probability of Earth:
p_e <- 0.5

# prob. of Mars:
p_m <- 0.5

# probability of land:
# das ist die Summe zweier gemeinsamer Wahrscheinlichkeiten (Randwahrscheinlichkeit)
# also die Summe zweier Zellen aus der Kontingenztabelle
p_l <- (p_e * p_le) + (p_m * p_lm)


# probability of Earth, given land (using Bayes' Theorem):
p_el <- (p_le * p_e) / p_l
p_el

## [1] 0.23

#> [1] 0.231
```

7. **Aufgabe**

This question is taken from McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan* (2. Ed.). Taylor and Francis, CRC Press.

2M4. Suppose you have a deck with only three cards. Each card has only two sides, and each side is either black or white. One card has two black sides. The second card has one black and one white side. The third card has two white sides. Now suppose all three cards are placed in a

bag and shuffled. Someone reaches into the bag and pulls out a card and places it flat on a table. A black side is shown facing up, but you don't know the color of the side facing down. Show that the probability that the other side is also black is 2/3. Use the counting method (Section 2 of the chapter) to approach this problem. This means counting up the ways that each card could produce the observed data (a black side faceing up on the table).

Lösung

Let's label the cards *bb* (black on both sides), *bw* (black on one, white on the other), and *ww* (both sides are white), respectively.

Wanted is the probability that both sides are black (*bb*), given one side is black (*1b*): $Pr(bb|1b)$.

Let's count the ways how the data - one black side - can come up in each conjecture (hypothesis), *bb*, *bw*, *ww*. Let's denote "first side white" as *1b*" and "first side black" as *2b*.

bb: 2 valid paths
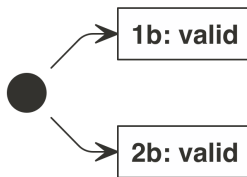
bw: 1 valid path
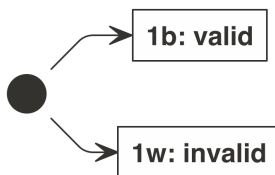
ww: 0 valid path
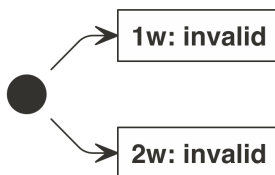

```
d <-
  tibble::tribble(
  ~Hyp,  ~Prior,
  "bb",     1,
  "bw",     1,
  "ww",     1,
  ) %>%
  mutate(Likelihood = c(2,1,0),
```

```
        unstand_post = Prior*Likelihood,
        std_post = unstand_post / sum(unstand_post))
```

Hyp	Prior	Likelihood	unstand_post	std_post
bb	1	2	2	0.67
bw	1	1	1	0.33
ww	1	0	0	0.00

The following solution is taken from [this source](#).

```
card_bb_likelihood <- 2
card_bw_likelihood <- 1
card_ww_likelihood <- 0

likelihood <- c(card_bb_likelihood, card_bw_likelihood, card_ww_likelihood)
prior <- c(1, 1, 1)
posterior <- likelihood * prior
posterior <- posterior / sum(posterior)

posterior[1]

## [1] 0.67
```

8. Aufgabe

This question is taken from McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan* (2. Ed.). Taylor and Francis, CRC Press.

2M5. Now suppose there are four cards: B/B, B/W, W/W, and another B/B. Again suppose a card is drawn from the bag and a black side appears face up. Again calculate the probability that the other side is black.

Lösung

The only difference to the question 2M4 is that we now have two *bb* cards, rendering the *prior* plausibility twice as high.

Let's label the cards *bb* (black on both sides), *bw* (black on one, white on the other), and *ww* (both sides are white), respectively.

Wanted is the probability that the *second* side of the card is black (*2b*), given one side is black (*1b*): $Pr(2b|1b)$.

```
d <-
  tibble::tribble(
  ~Hyp, ~Prior,
  "bb",    2L,
  "bw",    1L,
  "ww",    1L,
  ) %>%
  mutate(Likelihood = c(2,1,0),
```

```
        unstand_post = Prior*Likelihood,
        std_post = unstand_post / sum(unstand_post))
```

Hyp	Prior	Likelihood	unstand_post	std_post
bb	2	2	4	0.80
bw	1	1	1	0.20
ww	1	0	0	0.00

The following solution is taken from [this source](#).

```
card_bb_likelihood <- 2
card_bw_likelihood <- 1
card_ww_likelihood <- 0

likelihood <- c(card_bb_likelihood, card_bw_likelihood, card_ww_likelihood,
                card_bb_likelihood)
prior <- c(1, 1, 1, 1)
posterior <- likelihood * prior
posterior <- posterior / sum(posterior)

posterior[1] + posterior[4]

## [1] 0.8
```

9. Aufgabe

This question is taken from McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan* (2. Ed.). Taylor and Francis, CRC Press.

2M6. Imagine that black ink is heavy, and so cards with black sides are heavier than cards with white sides. As a result, it's less likely that a card with black sides is pulled from the bag. So again assume there are three cards: B/B, B/W, and W/W. After experimenting a number of times, you conclude that for every way to pull the B/B card from the bag, there are 2 ways to pull the B/W card and 3 ways to pull the W/W card. Again suppose that a card is pulled and a black side appears face up. Show that the probability the other side is black is now 0.5. Use the counting method, as before.

Lösung

Let's label the cards *bb* (black on both sides), *bw* (black on one, white on the other), and *ww* (both sides are white), respectively.

Wanted is the probability that the *second* side of the card is black (*2b*), given one side is black (*1b*): $Pr(2b|1b)$.

```
d <-
  tibble::tribble(
  ~Hyp, ~Prior,
  "bb",    1L,
  "bw",    2L,
```

```
    "ww",      3L,
    ) %>%
    mutate(Likelihood = c(2,1,0),
           unstand_post = Prior*Likelihood,
           std_post = unstand_post / sum(unstand_post))

d %>%
  gt() %>%
  fmt_number(columns = 5)
```

Hyp	Prior	Likelihood	unstand_post	std_post
bb	1	2	2	0.50
bw	2	1	2	0.50
ww	3	0	0	0.00

10. **Aufgabe**

This question is taken from McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan* (2. Ed.). Taylor and Francis, CRC Press.

2M7. Assume again the original card problem, with a single card showing a black side face up. Before looking at the other side, we draw another card from the bag and lay it face up on the table. The face that is shown on the new card is white. Show that the probability that the first card, the one showing a black side, has black on its other side is now 0.75. Use the counting method, if you can. Hint: Treat this like the sequence of globe tosses, counting all the ways to see each observation, for each possiible first card.

Lösung

Let's label the cards *bb* (black on both sides), *bw* (black on one, white on the other), and *ww* (both sides are white), respectively.

Wanted is the probability $Pr(c1 = bb|1b)$, the probability of drawing (as card 1) a *bb* card, given that we observerd b in the first draw, denoted as *1b*.

Here, we have to consider two cards. Let's use this notation *ww-bb* for the sequence "first card is white on both sides, second card is black on both sides".

The data observed is: first card has one black side, the second card has one white side, i.,e *b-w*.

Consider the following simple "tree diagram" showing the paths according to the four hypothesis, first level denotes the first card, the second (more indented) denotes the second card:

Hyp. A: bb-ww:

- b
    - w
    - w
- b

- - w
  - w

There are 4 paths consistend with the data, *b-w*.

Hyp. B: bw-ww:

- ○ b
  - - w
    - w
- ○ w
  - - w
    - w

There are 2 paths consistent with the data,* b-w*.

Hyp. A: bb-bw:

- ○ b
  - - b
    - w
- ○ b
  - - b
    - w

There are 2 paths consistent with the data.+, *b-w*.

Looking at the tree, we realize that out of all 8 allowed paths, 6 feature the *bb* card as first card:

$$Pr\left(c1 = bb | b - w\right) = 6/8 = 3/4 = 0.75$$

where `c1` means "card 1", and `b-w` means "first draw showed a *b* side, and second card showed a *w* face".

As for an other approach, consider the solution taken from [this source](#).

```
# 2 choices for first card (either bw or bb), with 3 options for second card: 2 ww + 1 bw
card_bb_likelihood <- 2 * 3
card_wb_likelihood <- 1 * 2
card_ww_likelihood <- 0

likelihood <- c(card_bb_likelihood, card_wb_likelihood, card_ww_likelihood)
prior <- c(1,1,1)
posterior <- prior * likelihood
posterior <- posterior / sum(posterior)

posterior[1]

## [1] 0.75
```