

## 18.1 Abbys Schwangerschaftsberatung

Die folgende Zuschrift wurde in einer Lebensberatungskolumne in einer Zeitung in den USA abgedruckt:

*Dear Abby: Du hast geschrieben, dass eine Frau 266 Tage lang schwanger ist. Ich habe mein Baby nach zehn Monaten und fünf Tagen geboren, und daran gibt es keinen Zweifel, denn ich erinnere mich an das Zugangsdatum sehr genau. Mein Ehegatte ist bei der Flotte, und ich habe ihn damals nur eine Stunde lang gesehen. Während meiner gesamten Schwangerschaft haben wir uns nicht getroffen.*

*Ich trinke nicht, bin auch sonst sehr zuverlässig, und es ist ganz unmöglich, dass das Kind nicht von ihm ist. Bitte drucke eine Berichtigung Deiner Behauptung ab, dass eine Schwangerschaft 266 Tage dauert, denn sonst bekomme ich viel Ärger.*

- a) Wie kann Abby helfen?
- b) Es stellt sich die Frage, ab wann der Ärger berechtigt ist. Vielleicht hilft der Hinweis, dass an dem vorbestimmten Geburtsdatum nur ca. 5 % der Geburten eintreten, 30 % der verbleibenden vor dem Datum geboren werden und außerdem von den anderen 70 % ca. 90 % innerhalb der ersten 10 Tage nach dem berechneten Datum auf die Welt kommen.

**Lösung** Der Aufgabenteil a) zielt auf einen Vorgehensvorschlag ab. Ärger für die Betroffene lässt sich vermeiden, wenn man begründen kann, dass solche Schwangerschaftsdauern wie die der Briefschreiberin durchaus normal sind. Zu dieser Frage kann man Stellung nehmen, wenn man eine begründete Wahrscheinlichkeit für ein so extremes Ereignis wie das vorgestellte angeben kann. Damit bietet sich folgende Strategie an:

- i) Wähle Modell für Schwangerschaftszeiten.
- ii) Bestimme Parameter des Modells.
- iii) Ermittle die gesuchte Wahrscheinlichkeit.

Der Aufgabenteil b) fordert die Umsetzung der Strategie ein. Zunächst sammeln wir Fakten und formulieren, was gesucht ist und entdecken folgende Fakten:

- 266 Tage Schwangerschaft ist Durchschnitt.
- Die besagte Schwangerschaft soll 10 Monate und 5 Tage (= 305 Tage) gedauert haben.
- 5% der Geburten treffen den prognostizierten Tag.
- 30% der verbleibenden 95% der Geburten tragen sich von dem anvisierten Tag zu, also  $.3 \cdot .95 \times 100\% = .3350 \times 100\%$  aller Geburten.
- 70% der verbleibenden 95% der Geburten erfolgen nach dem anvisierten Tag, also  $.7 \cdot .95 \times 100\% = .665 \times 100\%$  aller Geburten.

- die letztgenannten 66.5% zerteilen sich in 10% und 90% in der Form, dass sich 10% · 70% · 95% nach über 10 Tagen nach dem prognostizierten Geburtstermin ereignen.

Mit diesen Fakten soll die folgende Frage geklärt werden:

- Kann sich der Fall so zugetragen haben, oder wie groß ist die Wahrscheinlichkeit, eine Schwangerschaftsdauer von 305 oder mehr Tagen zu beobachten?

Kurz:

geg.: Durchschnittsschwangerschaftszeit: 266, spezielle Beobachtung: 305 sowie die weiteren genannten Fakten

ges.:  $P(\text{Schwangerschaftszeit} \geq 305) = ?$

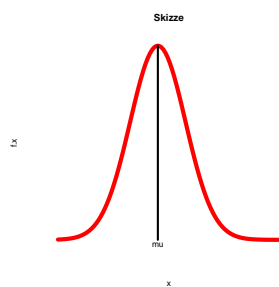
**ad i)** In dem Wissen, dass sehr kurze und sehr lange Schwangerschaften sehr selten sein dürften, wollen wir an die Gegebenheiten ein Normalverteilungsmodell anpassen. Denn in NV-Modellen sind ebenfalls die meisten Fälle in einem mittleren Bereich zu erwarten. Auch ist eine unimodale Verteilung plausibel. Bezüglich der Schiefe drängen sich keine Argumente auf. Im Anschluss an die Bearbeitung der Aufgabe werden wir übrigens noch einen Blick auf eine entsprechende empirische Verteilung werfen.

**ad ii)** Zunächst fertigen wir eine Skizze an, die die gegebene Situation reflektiert. Für die Skizze wählen wir  $\sigma = 10$  als unfachmännische Schätzung, dass sich wohl die Geburten innerhalb eines Intervalls der Länge 40 Tage um die mittlere Schwangerschaftsdauer zutragen werden. Die Mitte heben wir durch einen senkrechten Strich hervor.

140

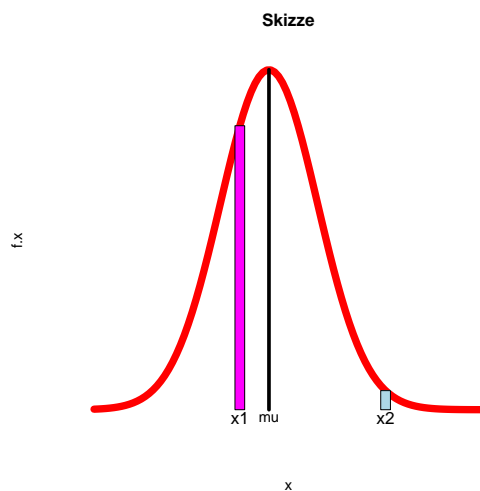
$\langle * 2 \rangle + \equiv$

*(zeichne Normalverteilung mit  $\mu = 266$  147)*



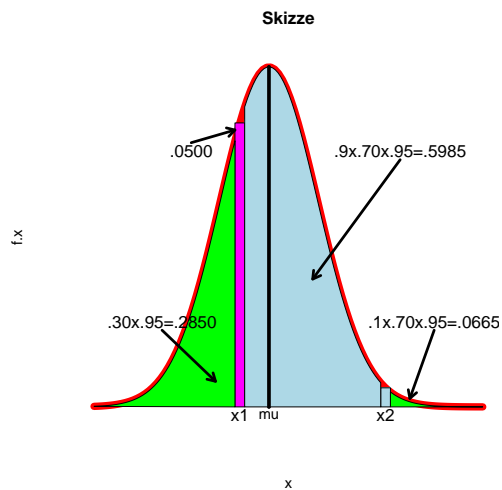
Nur für die Skizze platzieren wir den Tag für den prognostizierten Geburtstag links vom Mittelwert und bezeichnen ihn mit  $x_1$  oder  $x_1$ . Um zu verdeutlichen, dass zu einem Tag auf der Zeitachse ein Intervall gehört, haben wir die Stelle  $x_1$  durch einen kleinen Balken dargestellt. Der zweite interessante Punkt ist 10 Tage von  $x_1$  entfernt. Wir nennen ihn  $x_2$  (oder  $x_2$ ), wählen für die Anschauung einen Punkt rechts von der Mitte und erweitern auch damit unsere Skizze.

141  $\langle * 2 \rangle + \equiv$   
*(zeichne  $x_1$  und  $x_2$  ein 148)*



Zu den Bereichen kennen wir aufgrund der Fakten die Wahrscheinlichkeiten bzw. Flächengrößen. Mit diesen ergänzen wir unsere Skizze.

142  $\langle * 2 \rangle + \equiv$   
*(zeichne Skizze mit Flächenstreifen 149)*



Wenn wir das Bild betrachten, sehen wir die relevanten Bereiche sowie die Wahrscheinlichkeiten, mit denen die Bereiche getroffen werden. Unterstellen wir 10000 Geburten, dann können wir durch Weglassen der Dezimalpunkte auch die erwarteten Anzahlen der Geburten in den Bereichen ablesen.

**Wie lässt sich jetzt der korrekte Wert für  $x_1$  bzw.  $x_1$  berechnen?**  
Ausgehend von Wahrscheinlichkeiten können wir Prozentpunkte der Standardnormalverteilung  $z_p$  ermitteln. Durch die Transformationsbeziehung

$$x_p = z_p \sigma + \mu$$

sind diese mit den Prozentpunkten  $x_p$  anderer Normalverteilungen verbunden. In unserem Fall betrachten wir  $\mu = 266$  als bekannt. Wenn wir die empirisch gewonnenen Prozentangaben als Wahrscheinlichkeiten interpretieren, liegt links von  $x_2 = x_{p=.9335}$  gerade 93.35% der Wahrscheinlichkeitsmasse. Weiterhin ist  $x_2$  um 10 (Tage) größer als  $x_1$ . Dann folgt:

$$x_{p=.9335} = x_2 = x_1 + 10$$

Um in den linken Bereich den Tag den prognostizierten Geburtstag (PGT) mit einzuschließen, addieren wir die Wahrscheinlichkeit für eine Geburt vor dem PGT und die Wahrscheinlichkeit einer Niederkunft genau am Prognosetag. Dann erhalten wir:

$$x_1 = x_{p=.2850+0.050} = x_{p=.3350}$$

Mit diesen Vorüberlegungen können wir uns auf die Suche nach den nun nur noch zwei Unbekannten machen:  $x_1$  und  $\sigma$ . Mit Hilfe der Wahrscheinlichkeiten erhalten wir für die beiden Stellen  $x_1$  und  $x_1 + 10$  zwei Gleichungen:

$$x_{p=.3350} = x_1 = \sigma z_{p=.3350} + 266 \quad (1)$$

$$x_{p=.9335} = x_1 + 10 = \sigma z_{p=.9335} + 266 \quad (2)$$

Jetzt haben wir zwei Gleichungen mit zwei Unbekannten und formen (1) um:

$$x_1 - 266 = \sigma z_{p=.3350} \Rightarrow \sigma = \frac{x_1 - 266}{z_{p=.3350}}$$

Diese Größe können wir nach Auflösung von (2) nach  $x_1$  einsetzen und erhalten mit  $Q := z_{p=.9335}/z_{p=.3350}$ :

$$x_1 = -10 + \left( \frac{x_1 - 266}{z_{p=.3350}} \right) \cdot z_{p=.9335} + 266 =: 256 + Q \cdot (x_1 - 266) = x_1$$

Mit wenig Mühe folgt für  $x_1$  eine Bestimmungsgleichung:

$$x_1 - Qx_1 = 256 - Q \cdot 266 \Rightarrow x_1 = \frac{256 - 266 \cdot Q}{1 - Q}$$

Auch können wir nun schnell das unbekannte  $\sigma$  bestimmen, denn aus (1) folgt:

$$\sigma = \frac{x_1 - 266}{z_{p=.3350}}$$

Das werden wir sofort einmal ausrechnen:

143

```

(*)+≡
z1 <- z_3350 <- qnorm(.3350)
z2 <- z_9335 <- qnorm(.9335)
Q <- z2/z1
x1 <- (256 - 266*Q)/(1 - Q)
sigma <- (x1 - 266)/z1
cat("Wir erhalten x1 =",x1,"und sigma =", sigma)

```

Wir erhalten  $x_1 = 263.7903$  und  $\sigma = 5.185312$

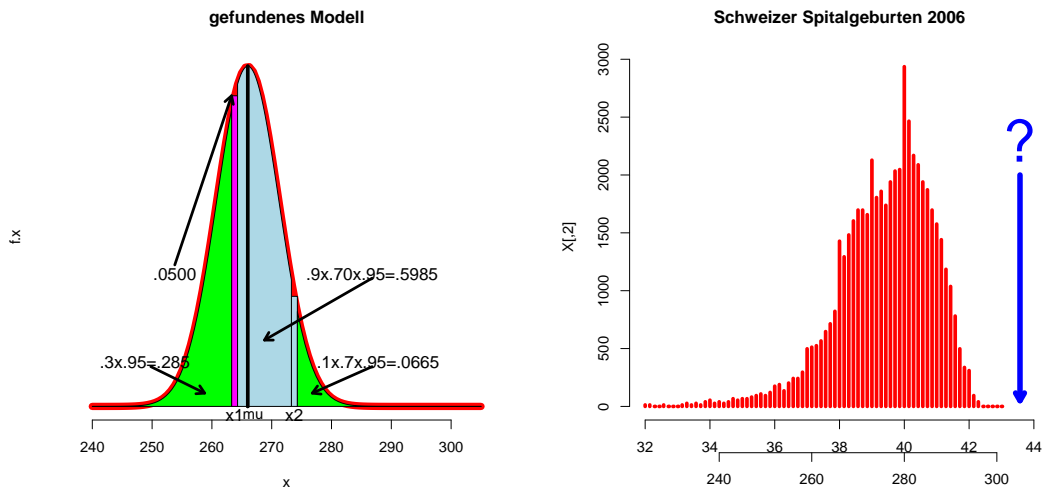
**Welche Wahrscheinlichkeit folgt** für das fragliche Ereignis? Mit diesen Berechnungen können wir den letzten Schritt vollziehen und die gesuchte Wahrscheinlichkeit nach dem ausgewählten Modell berechnen:

144 `< * 2>+ ≡`  
`cat("WS fuer Geburt nach ueber 304 Tagen betraegt", 1-pnorm(304.5, 266, sigma))`  
 Als Ergebnis bekommen wir:

WS fuer Geburt nach ueber 304 Tagen betraegt 5.651035e-14

**Reflexion** Mit den erarbeiteten Ergebnissen passen wir die Skizze von oben an unser Modellierungsergebnis an, ...

145 `< * 2>+ ≡`  
`<zeichne an das gefundene Modell angepasste Skizze 150>`  
 ... siehe Darstellung links. Auch in dieser sehen wir, dass um 305 keine Wahrscheinlichkeitsmasse zu erkennen ist.



**Empirische Untermauerung** Eine Graphik einer Wikipedia-Seite stützt unsere Ergebnisse. Siehe dazu:

<http://de.wikipedia.org/wiki/Schwangerschaftsdauer>

und betrachte dort die Darstellung der Verteilung des Geburtszeitpunktes. Die Verteilung der Schweizer Spitalgeburten von 2006 ist zwar etwas linksschief, doch wir erkennen, dass auch hier nach der Woche 42 bzw. nach 294 Tagen keine Geburten mehr zu erwarten sind. Die referenzierte Graphik ist grob rekonstruiert in der rechten Hälfte der letzten Abbildung zu finden.

146 `< * 2>+ ≡`  
`<zeige Schweizer Verteilung 151>`

Mit diesen Ausführungen, die leider für die Anfragerin insgesamt negativ ausfallen, kann zumindest die Aufgabe als hinreichend beantwortet angesehen werden.

Für Formalisten sei noch erwähnt, dass wir zwecks Lesbarkeit empirische Häufigkeiten als Wahrscheinlichkeiten interpretiert haben. Damit konnten wir Schätzfragen ausklammern und auf die  $\hat{\cdot}$  über Parametern wie  $\mu$  und  $\sigma$  verzichten.

## 18.2 Anhang mit Code-Sequenzen

- 147 *(zeichne Normalverteilung mit  $\mu = 266$  147)*  $\equiv$  C 140, 148
- ```
xmin <- 230; xmax <- 310
x <- seq(xmin, xmax, length=100)
f.x <- dnorm(x, m=266, sd=10)
plot(x, f.x, type="l", lwd=8, col="red", axes=FALSE)
x <- 266; segments(x,.000, x, dnorm(x, m=266, sd=10), lwd=4)
text(x, -0.001, "mu"); title("Skizze")
```
- 148 *(zeichne x1 und x2 ein 148)*  $\equiv$  C 141, 149
- (zeichne Normalverteilung mit  $\mu = 266$  147)*
- ```
x1 <- x <- 260; text(x, -0.001, "x1", cex=1.2)
fx <- dnorm(x, m=266, sd=10)
polygon(c(x-1,x-1,x+1,x+1),c(0, fx, fx, 0),col="magenta")
x2 <- x <- x + 30; text(x, -0.001, "x2", cex=1.2); fx <- dnorm(x, m=266, sd=10)
polygon(c(x-1,x-1,x+1,x+1),c(0, fx, fx, 0),col="lightblue")
```
- 149 *(zeichne Skizze mit Flächenstreifen 149)*  $\equiv$  C 142
- (zeichne x1 und x2 ein 148)*
- ```
# prognosetag
text(x1-10, .03, ".0500", cex=1.2)
arrows(x1-10, .031, x1-1, .96*dnorm(x1, m=266, sd=10)+.0005, lwd=3, length=.1)
# links
xx <- seq(xmin, x1-1, length=50)
polygon(c(xx, x1-1), c(dnorm(xx, m=266, sd=10),0), col="green")
text(x1-16, .01, ".30x.95=.2850", cex=1.2)
arrows(x1-16, .009, x1-5, .003, lwd=3, length=.1)
# rechts
xx <- seq(xmax, x2+1, length=50)
polygon(c(xx, x2+1), c(dnorm(xx, m=266, sd=10),0), col="green")
text(x2+10, .01, ".1x.70x.95=.0665", cex=1.2)
arrows(x2+10, .009, x2+5, .001, lwd=3, length=.1)
# Mitte
xx <- seq(x1+1, x2-1, length=50)
polygon(c(x1+1, xx, x2-1), c(0, dnorm(xx, m=266, sd=10),0), col="lightblue")
x <- 266
segments(x,.000, x, dnorm(x, m=266, sd=10), lwd=4)
text(x2+3, .03, ".9x.70x.95=.5985", cex=1.2)
arrows(x2+3, .029, x2-15, .015, lwd=3, length=.1)
```
- 150 *(zeichne an das gefundene Modell angepasste Skizze 150)*  $\equiv$  C 145
- ```
xmin <- 240; xmax <- 305; sigma <- 5.185312; x1 <- 263.7903
x <- seq(xmin, xmax, length=100)
f.x <- dnorm(x, m=266, sd=sigma)
plot(x, f.x, type="l", lwd=8, col="red", axes=FALSE); axis(1)
x <- 266
segments(x,.000, x, dnorm(x, m=266, sd=sigma), lwd=4)
```

```

text(266.9, -0.002, "mu")
title("gefundenes Modell")
x1 <- x <- x1; x2 <- x1+10
text(x, -0.002, "x1", cex=1.2)
#segments(x,.000, x, dnorm(x, m=266, sd=sigma), lwd=10)
fx <- dnorm(x, m=266, sd=sigma)
polygon(c(x-.5,x-.5,x+.5,x+.5),c(0, fx, fx, 0),col="magenta")
x <- x2
fx <- dnorm(x, m=266, sd=sigma)
polygon(c(x-.5,x-.5,x+.5,x+.5),c(0, fx, fx, 0),col="lightblue")
text(x, -0.002, "x2", cex=1.2)
# prognosetag
text(x1-10, .03, ".0500", cex=1.2)
arrows(x1-10, .032, x1-.5, dnorm(x1, m=266, sd=sigma)+.000, lwd=3, length=.1)
# links
xx <- seq(xmin, x1-.5, length=50)
polygon(c(xx, x1-.5), c(dnorm(xx, m=266, sd=sigma),0), col="green")
text(x1-15, .01, ".3x.95=.285", cex=1.2)
arrows(x1-14, .009, x1-5, .003, lwd=3, length=.1)
# rechts
xx <- seq(xmax, x2+.5, length=50)
polygon(c(xx, x2+.5), c(dnorm(xx, m=266, sd=sigma),0), col="green")
text(x2+14, .01, ".1x.7x.95=.0665", cex=1.2)
arrows(x2+13, .009, x2+3, .003, lwd=3, length=.1)
# Mitte
xx <- seq(x1+.5, x2-.5, length=50)
polygon(c(x1+.5, xx, x2-.5), c(0,dnorm(xx, m=266, sd=sigma),0), col="lightblue")
x <- 266
segments(x,.000, x, dnorm(x, m=266, sd=sigma), lwd=4)
text(x2+13, .03, ".9x.70x.95=.5985", cex=1.2)
arrows(x2+13, .029, x2-5, .015, lwd=3, length=.1)

```

151

(zeige Schweizer Verteilung 151)  $\equiv$   $\subset$  146

```

X <- structure(c(32, 32.143, 32.286, 32.429, 32.571, 32.714, 32.857,
33, 33.143, 33.286, 33.429, 33.571, 33.714, 33.857, 34, 34.143,
34.286, 34.429, 34.571, 34.714, 34.857, 35, 35.143, 35.286, 35.429,
35.571, 35.714, 35.857, 36, 36.143, 36.286, 36.429, 36.571, 36.714,
36.857, 37, 37.143, 37.286, 37.429, 37.571, 37.714, 37.857, 38,
38.143, 38.286, 38.429, 38.571, 38.714, 38.857, 39, 39.143, 39.286,
39.429, 39.571, 39.714, 39.857, 40, 40.143, 40.286, 40.429, 40.571,
40.714, 40.857, 41, 41.143, 41.286, 41.429, 41.571, 41.714, 41.857,
42, 42.143, 42.286, 42.429, 42.571, 42.714, 42.857, 43, 13, 13,
0, 0, 13, 0, 0, 0, 13, 27, 13, 27, 13, 40, 54, 27, 40, 27, 40,
67, 54, 67, 67, 81, 94, 108, 94, 121, 175, 189, 135, 202, 242,
242, 296, 498, 512, 525, 566, 647, 714, 822, 1428, 1293, 1482,
1603, 1697, 1697, 1657, 2128, 1805, 1859, 1738, 1940, 2034, 2047,
2936, 2465, 2169, 2088, 1940, 1872, 1697, 1576, 1441, 1185, 1037,
781, 498, 337, 310, 94, 40, 0, 0, 0, 0, 0), .Dim = c(78L, 2L))
plot(X,type="h",lwd=4,col="red", xlim=c(32,44), xlab="",axes=FALSE)
arrows(305/7, 2000, 305/7, 50, lwd=6, col="blue", length=.1)
text(305/7, 2300, "?", cex=4, col="blue")
title("Schweizer Spitalgeburten 2006"); axis(2); axis(1)
usr <- par()$usr; usr[1:2] <- usr[1:2]*7; par(usr=usr); axis(1,pos=-400)

```