1.1 **▶ Geeignete Berechnung der fehlenden Wahrscheinlichkeitswerte** (11BE)

Damit eine Wahrscheinlichkeitsverteilung entsteht, muss die Summe aller Einzelwahrscheinlichkeiten $P(k)$ gerade 1 sein. Wegen

$P(1) + P(2) + P(3) + P(4) + P(5) + P(6) = 0{,}30 + 0{,}18 + 0{,}12 + 0{,}10 + 0{,}08 + 0{,}07 = 0{,}85$

muss die Summe der fehlenden Wahrscheinlichkeitswerte 0,15 betragen. Daher wird wie folgt gerundet:

$$P(7) = \log_{10}\frac{8}{7} = 0{,}057991... \approx 0{,}06;$$

$$P(8) = \log_{10}\frac{9}{8} = 0{,}051152... \approx 0{,}05;$$

$$P(9) = \log_{10}\frac{10}{9} = 0{,}045757... \approx 0{,}05.$$

1.2 **▶ Berechnung der Wahrscheinlichkeit $P(A)$**

Unter der 100-elementigen Zufallsprobe hat die Ziffer 1 mit $P(1) = 0{,}30$ den Erwartungswert $E(1) = n \cdot p = 100 \cdot 0{,}30 = 30$.

Wenn X die Anzahl der auftretenden Zahlen mit der Ziffer 1 ist, gilt für die gesuchte Wahrscheinlichkeit

$$P(A) = P(X = 30) = \binom{100}{30} \cdot 0{,}30^{30} \cdot 0{,}70^{70} \approx 0{,}087 = 8{,}7\%.$$

▶ Berechnung der Wahrscheinlichkeit $P(B)$

Eine σ-Umgebung ist die Umgebung, die vom Erwartungswert μ nach oben und unten gerade um die Standardabweichung σ abweicht, also das Intervall $[\mu - \sigma;\ \mu + \sigma]$.

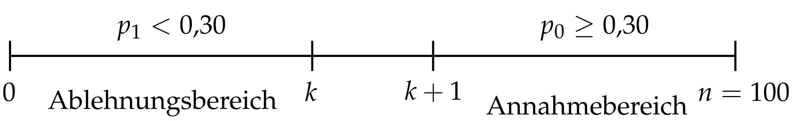
Mit $\mu = E(1) = 30$ und $\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{100 \cdot 0{,}30 \cdot 0{,}70} = \sqrt{21}$ folgt:

$$\begin{aligned} P(\mu - \sigma \leq X \leq \mu + \sigma) &= P\left(30 - \sqrt{21} \leq X \leq 30 + \sqrt{21}\right) \\ &= P\left(25{,}4174 \leq X \leq 34{,}5826\right) = P\left(26 \leq X \leq 34\right) \\ &= \sum_{i=26}^{34} \binom{100}{i} \cdot \left(0{,}3^i \cdot 0{,}7^{100-i}\right) \approx 0{,}674 \end{aligned}$$

Also liegt die Anzahl der Zahlen mit Ziffer 1 mit einer Wahrscheinlichkeit von etwa 67,4% in der 1-σ-Umgebung um den Erwartungswert.

2.1 **▶ Entwicklung eines Signifikanztests mit Signifikanzniveau 5%** (14BE)

Es sei X die Anzahl der auftretenden Zahlen mit Anfangsziffer 1 im Test. Getestet wird vorsichtshalber die Nullhypothese H_0: $p_0 \geq 0{,}30$, also dass die Bilanz nicht gefälscht ist, damit sich die Prüfer sicher sein können, nicht falsch zu liegen. Sie wird abgelehnt, falls sich in dem Test **sehr wenige** Zahlen mit Anfangsziffer 1 feststellen lassen. Als Ablehnungsbereich wird daher $\overline{A} = \{0; 1; ...; k\}$ festgelegt.



Das Signifikanzniveau von $\alpha = 5\%$ gibt an, mit welcher Wahrscheinlichkeit die Nullhypothese irrtümlich abgelehnt wird, also dass sich höchstens k Zahlen mit Anfangsziffer 1 in der Stichprobe finden lassen, obwohl die Wahrscheinlichkeit dafür tatsächlich $p_0 \geq 0{,}30$ beträgt:

$\alpha = P_{\{H_0 \text{ wahr}\}}(X \in \overline{A}) = P(X \leq k) \leq 0{,}05.$

$$P\,(X \leq k) = \sum_{i=0}^{k} \binom{100}{i} \cdot 0{,}3^i \cdot 0{,}7^{100-i} \approx 0{,}05$$

Eine Tabelle zur kumulierten Binomialverteilung liefert $P(X \leq 22) = 0{,}0479$, also $k \leq 22$.

Daher wird die Nullhypothese angenommen, wenn mindestens 23 Zahlen die Anfangsziffer 1 aufweisen.

▶ Beschreibung der Fehler 1. und 2. Art im Sachzusammenhang

Der **Fehler 1. Art** gibt die Wahrscheinlichkeit dafür an, die Nullhypothese irrtümlich **abzulehnen**. Im Sachzusammenhang ist dies die Möglichkeit, dass eine Bilanz als **gefälscht** erachtet wird, obwohl sie eigentlich in Ordnung ist.

Der **Fehler 2. Art** gibt die Wahrscheinlichkeit dafür an, die Nullhypothese irrtümlich **anzunehmen**. Im Sachzusammenhang ist dies die Möglichkeit, dass die Bilanz in Wirklichkeit **gefälscht** ist, aber im Test nicht als solche erkannt wird.

2.2 **▶ Wahrscheinlichkeit, dass die Fälschung unentdeckt bleibt**

Ist die Bilanz gefälscht und hat nur jede 5. Zahl die Anfangsziffer 1, so ist die Wahrscheinlichkeit für das Finden einer solchen Zahl gerade $p = 1/5 = 0{,}2$. Die Nullhypothese wird trotzdem angenommen (und die Fälschung damit nicht aufgedeckt), wenn sich in dem Test trotz allem mindestens 22 Zahlen mit Anfangsziffer 1 finden lassen. Die Wahrscheinlichkeit hierfür ist:

$P(X \geq 23) = 1 - P(X \leq 22) = 1 - 0{,}7389 = 0{,}2611 = 26{,}11\%.$

Für $n = 100$ und $p = 0{,}2$ sind die summierten Binomialverteilungen in den Tabellen aufgelistet.

3. **▶ Alternatives Testverfahren vorschlagen** (5BE)

Eine Möglichkeit, das Testverfahren zu ändern, wäre, den **Ablehnungsbereich** der Nullhypothese zu vergrößern. Der aktuelle Test ist ein Signifikanztest auf dem Signifikanzniveau von 5 %, bei dem die Nullhypothese $H_0 : p_0 \geq 0{,}3$ geprüft wird. Die Vermutung, dass die Hypothese Auf diese Weise kann der Fehler 1. Art kontrolliert und festgelegt werden und es wird vermieden, dass eine ungefälschte Bilanz versehentlich als gefälscht eingestuft wird.

Eine mögliche Änderung wäre, statt der Nullhypothese $H_0 : p_0 \geq 0{,}3$ die Hypothese $H : p \leq 0{,}3$ zu prüfen. Mit einem Signifikanzniveau von 5 % würde somit gewährleistet, dass eine gefälschte Bilanz nur mit einer Wahrscheinlichkeit von 5 % als ungefälscht eingestuft wird.

Das Problem hierbei ist, dass nun **viele** ungefälschte Bilanzen als gefälscht eingestuft werden.

Eine andere Möglichkeit wäre, statt des Signifikanztests einen **Alternativtest** durchzuführen. Hierzu sind allerdings **zwei** Hypothesen notwendig, die gegeneinander getestet werden. Die Bilanzprüfer müssen somit eine Vermutung haben, die sie testen können.

Wollen die Bilanzprüfer beispielsweise testen, ob die Hypothese $H_0 : p_0 = 0{,}3$ oder die Hypothese $H_1 : p_1 = 0{,}2$ (wie in 2.2) wahrscheinlicher ist, so bietet sich ein solcher Alternativtest an.

Auf diese Weise lassen sich sowohl der Fehler 1. als auch 2. Art „steuern".