# Musterlösung zu Serie 5

**1.**

1. $\mathrm{P}[X = 10] = F(10) - F(9) = 0.176$
2. $\mathrm{P}[X > 12] = 1 - P[X \leq 12] = 1 - F(12) = 0.132$
3. $\mathrm{P}[5 \leq X \leq 15] = \mathrm{P}[X \leq 15] - \mathrm{P}[X < 5] = F(15) - F(4) = 0.988$
4. Es soll gelten $\mathrm{P}[X < k] = F(k-1) \leq 0.025 < F(k)$. Aus der Tabelle folgt $k = 6$.
5. Es muss $\mathrm{P}[X > k] \leq 0.05 < \mathrm{P}[X > k-1]$, also $F(k) \geq 0.95 > F(k-1)$ gelten. Aus der Tabelle folgt $k = 14$.

**2.**

1. **Modell:** $X$: Anzahl defekter Reagenzgläser in einer Stichprobe aus 50 Reagenzgläsern. $X \sim Bin(50, \pi)$.
2. **Nullhypothese:** $H_0 : \pi = 0.1$
   **Alternative:** $H_A : \pi < 0.1$
3. **Teststatistik:** $T$: Anzahl defekter Reagenzgläser in einer Stichprobe aus 50 Reagenzgläsern.
   **Verteilung der Teststatistik unter $H_0$:** $T \sim Bin(50, 0.1)$
4. **Signifikanzniveau:** $\alpha = 0.05$
5. **Verwerfungsbereich:** Falls $H_0$ stimmt, gilt:

| | | | | | |
|---|---|---|---|---|---|
| $P(X = 0)$ | = | 0.0052 | $P(X \leq 0)$ | = | 0.0052 |
| $P(X = 1)$ | = | 0.0286 | $P(X \leq 1)$ | = | 0.0338 |
| $P(X = 2)$ | = | 0.0779 | $P(X \leq 2)$ | = | 0.1117 |

   Der Verwerfungsbereich $K$ für ein Signifikanzniveau von 5% ist also gegeben durch $K = \{0, 1\}$.
6. **Testentscheid:** Der beobachtete Wert der Teststatistik ist $t = 3$. Der beobachtete Wert der Teststatistik ($t = 3$) liegt nicht im Verwerfungsbereich der Teststatistik ($K = \{0, 1\}$). Die Nullhypothese kann daher auf dem 5% Signifikanzniveau nicht verworfen werden. Es kann also durchaus sein, dass der Anteil minderwertiger Gläser in der ganzen Lieferung 10% ist. Der Hersteller sollte also seine Lieferung nicht losschicken sondern genauer untersuchen.

**3.**

**a)** Weil wir im zweiten Schritt 200 Fische ziehen, ist $n = 200$.
Die Wahrscheinlichkeit, im zweiten Schritt einen markierten Fisch zu fangen ist

$$\pi = \frac{\text{Anzahl markierte Fische}}{\text{Totale Anzahl Fische im See}} = \frac{500}{N}. \tag{1}$$

Für $N = 2000$ ergibt dies $\pi = 1/4$, für $N = 5000$ erhält man $\pi = 1/10$.

**b)** Wir schätzen $\pi$ durch $\widehat{\pi} = x/n = 40/200 = 1/5$. Wenn wir (1) nach $N$ auflösen, ergibt sich die folgende Schätzung für die Gesamtzahl Forellen im See:

$$\widehat{N} = \frac{500}{\widehat{\pi}} = 2500.$$

**c)**

1. **Modell:** $X$: Anzahl markierter Forellen.
   $X \sim \mathrm{Bin}(200, \pi)$.
2. Wegen (1) gilt $N > 2000$ genau dann wenn $\frac{500}{\pi} > 2000$, was äquivalent zu $\frac{1}{4} > \pi$ ist. Deshalb:
   **Nullhypothese:** $H_0 : \pi = \pi_0 = \frac{1}{4}$.
   **Alternative:** $H_A : \pi < \pi_0 = \frac{1}{4}$.
3. **Teststatistik:** $T$: Anzahl markierter Forellen.
   **Verteilung der Teststatistik unter $H_0$:** $T \sim \mathrm{Bin}(200, 1/4)$
4. **Signifikanzniveau:** $\alpha = 0.05$.

5. **Verwerfungsbereich:** Wir müssen den Verwerfungsbereich $K = [0, c]$ so bestimmen, dass $P_{H_0}(T \leq c) \leq \alpha$ und $P_{H_0}(T \leq c+1) > \alpha$ ist. Man berechnet mit
`pbinom(0:50, size=200,prob=1/4)`
die Werte folgender Tabelle:

| $k$ | 0 | 1 | ... | 38 | 39 | 40 | 41 | ... | 50 |
|---|---|---|---|---|---|---|---|---|---|
| $F(k)$ | 0.000 | 0.000 | ... | 0.0276 | 0.0405 | 0.0578 | 0.0804 | ... | 0.5379 |

Daraus kann man $c = 39$ ablesen. Oder schneller:
`max(which(pbinom(0:200, size=200, prob=1/4)<=.05))-1 ## 39 .`
Also $K = [0, 39]$.

6. **Testentscheid:** Der beobachtete Wert der Teststatistik ist $t = 40$. Der beobachtete Wert der Teststatistik liegt nicht im Verwerfungsbereich der Teststatistik. Die Nullhypothese kann daher auf dem 5% Signifikanzniveau nicht verworfen werden.

**d)**
- Da es sich um Ziehungen ohne zurücklegen handelt (alle Fische werden auf einmal gefangen), ist streng genommen die Unabhängigkeitsannahme beim Ziehen verletzt.
- Falls die 20 Forellen an derselben Stelle im See gefischt werden, kann die Annahme der Unabhängigkeit der einzelnen Fangereignisse verletzt sein (die miteinander markierten Fische könnten beispielsweise bevorzugt beeinander bleiben).
- Denkbar sind auch Lerneffekte der einmal gefangenen Fische, d.h. die Erfolgswahrscheinlichkeit $\pi$, einen markierten Fisch zu fangen, wäre dann kleiner als oben berechnet, da ein einmal gefangener (und markierter) Fisch nicht so schnell ein zweites Mal ins Netz oder an den Haken geht.
- Einen Effekt auf die Erfolgswahrscheinlichkeit $\pi$, einen markierten Fisch zu fangen, kann auch eine Zu- oder Abwanderung von Fischen haben.

Bemerkung: Die beiden letztgenannten Punkte sprechen eigentlich nicht gegen die Annahme, dass $X$ binomialverteilt ist. Da sie aber $\pi$ verändern können, führen sie zu systematischen Fehlern, wenn man aufgrund des in der Aufgabe beschriebenen Vorgehens über eine Schätzung des Parameters $\pi$ die Anzahl $N$ der Forellen im See schätzen will.