Stochastik **Lösungen zu Serie 5** **HSLU**

Vorbesprechung:
2013

Aufgabe 1

(**a**) P(genau 200 Unfälle)=

```
> dpois(x=200, lambda=200)
[1] 0.02819773
```

(**b**) P(höchstens 210 Unfälle) =

```
> ppois(q=210, lambda=200)
[1] 0.772708
```

(**c**) P(zwischen 190 und 210 Unfälle) =

```
> ppois(q=210, lambda=200) - ppois(q=189, lambda=200)
[1] 0.5422097
```

Aufgabe 2

Da $X \sim \text{Poisson}(\lambda)$ mit $\lambda = 2$ gilt: $P(X = x) = \exp(-2)\frac{2^x}{x!}$

(**a**) $P(X = 0) = \exp(-2)\frac{2^0}{0!} = \exp(-2)\frac{1}{1} \approx 0.135$

(**b**)

$$\begin{aligned} P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= 0.135 + 0.271 + 0.271 + 0.180 \\ &\approx 0.857 \end{aligned}$$

(**c**) $P(X > 3) = 1 - P(X \leq 3) = 1 - 0.857 \approx 0.143$

(**d**) Nach Kapitel 3.7.2 folgt: $Y \sim \text{Poisson}(6 \cdot \lambda) = \text{Poisson}(12)$

Aufgabe 3

Es gilt: $X_1 \sim \text{Bin}(n_1, \pi)$ und $X_2 \sim \text{Bin}(n_2, \pi)$; X_1 und X_2 sind unabhängig.

(a) Da X_1 und X_2 unabhängig sind, gilt:

$$P(X_1 = x_1 \cap X_2 = x_2) = P(X_1 = x_1) \cdot P(X_2 = x_2) \,,$$

wobei $P(X_1 = x_1) = \binom{n_1}{x_1}\pi^{x_1}(1-\pi)^{n_1-x_1}$ und $P(X_2 = x_2) = \binom{n_2}{x_2}\pi^{x_2}(1-\pi)^{n_2-x_2}$.

(b)

$$\begin{aligned}
\log(P(X_1 = x_1 \cap X_2 = x_2)) &= \log(P(X_1 = x_1) \cdot P(X_2 = x_2)) \\
&= \log(P(X_1 = x_1)) + \log(P(X_2 = x_2)) \\
&= \log\left(\binom{n_1}{x_1}(1-\pi)^{n_1-x_1}\right) + \log\left(\binom{n_2}{x_2}\pi^{x_2}(1-\pi)^{n_2-x_2}\right) \\
&= \log\left(\binom{n_1}{x_1}\right) + x_1 \cdot \log(\pi) + (n_1 - x_1) \cdot \log(1-\pi) \\
&+ \log\left(\binom{n_2}{x_2}\right) + x_2 \cdot \log(\pi) + (n_2 - x_2) \cdot \log(1-\pi) \,.
\end{aligned}$$

(c)

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\pi}\left\{\log\left(\binom{n_1}{x_1}\right) + x_1 \cdot \log(\pi) + (n_1 - x_1) \cdot \log(1-\pi)\right. \\
&\left.+\log\left(\binom{n_2}{x_2}\right) + x_2 \cdot \log(\pi) + (n_2 - x_2) \cdot \log(1-\pi)\right\} \\
&= \frac{x_1}{\pi} - (n_1 - x_1) \cdot \frac{1}{1-\pi} + \frac{x_2}{\pi} - (n_2 - x_2) \cdot \frac{1}{1-\pi} \\
&= \frac{x_1 + x_2}{\pi} - \frac{((n_1 + n_2) - (x_1 + x_2))}{1-\pi} \,.
\end{aligned}$$

Wenn wir diesen Ausdruck gleich Null setzen und nach π auflösen, erhalten wir:

$$\pi = \frac{x_1 + x_2}{n_1 + n_2} \,.$$

Das Ergebnis ist also identisch mit dem Ergebnis, das wir erhalten hätten, wenn eine Person 30 + 50 = 80 Lose gezogen hätte und dabei 2 + 4 = 6 Gewinne gezogen hätte (da $X_1 + X_2 \sim \text{Bin}(n_1 + n_2, \pi)$).
Das hier gesehene Prinzip, einen Parameter zu schätzen, indem man mehrere unabhängige Beobachtungen kombiniert, ist die mit Abstand häufigste Schätzmethode in der Statistik.

Aufgabe 4

(a)

$$P(X=2) = \binom{10}{2} 0.3^2 0.7^8 = 0.23$$

$$P(X \le 2) = P(X=0) + P(x=1) + P(X=2) = 0.7^{10} + \binom{10}{1} 0.3^1 0.7^9 + \binom{10}{2} 0.3^2 0.7^8 = 0.38$$

(b) **1.** Modell: X ist die Anzahl erfolgreich behandelter Patienten, $X \sim \text{Bin}(10, \pi)$.

2. Die Nullhypothese ist $H_0 : \pi = 0.3$, die Alternative ist $H_A : \pi > 0.3$.

3. Die Teststatistik ist $T : P(T = t|H_0) = \binom{10}{t} 0.3^t 0.7^{10-t}$

4. Das Signifikanzniveau ist $\alpha = 0.05$.

5. Verwerfungsbereich:

	$t=5$	$t=6$	$t=7$	$t=8$	$t=9$	$t=10$
$P(T \ge t)$	0.1503	0.0473	0.0106	0.0016	0.0001	5.9×10^{-6}

Daher ist der Verwerfungsbereich $K = \{6, 7, 8, 9, 10\}$.

6. Testentscheid: Da $4 \notin K$ wird H_0 nicht verworfen. Eine erhöhte Wirksamkeit des neuen Medikaments kann nicht nachgewiesen werden.

(c) Die Macht eines Tests ist die Wahrscheinlichkeit, dass die Nullhypothese verworfen wird, wenn die Alternative stimmt: $P(T \in K|H_A)$. (Alternativ: Macht $= 1 - P(\text{Fehler 2. Art}) = 1 - P(T \notin K|H_A)$)

Im konkreten Fall:

$$\text{Macht} = \binom{10}{6} 0.6^6 0.4^4 + \binom{10}{7} 0.6^7 0.4^3 + \binom{10}{8} 0.6^8 0.4^2 + \binom{10}{9} 0.6^9 0.4 + 0.6^1 0 = 0.6331\,.$$

Aufgabe 5

1. **Modell**: X: Anzahl defekter Reagenzgläser in einer Stichprobe aus 50 Reagenzgläsern. $X \sim \text{Bin}(50, \pi)$.

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 $H_0 : T \sim \text{Bin}(50, 0.1)$

4. **Signifikanzniveau**: $\alpha = 0.05$

5. Verwerfungsbereich: Falls H_0 stimmt, gilt:

$$P(T = 0) = 0.0052 \qquad P(T \leq 0) = 0.0052$$
$$P(T = 1) = 0.0286 \qquad P(T \leq 1) = 0.0338$$
$$P(T = 2) = 0.0779 \qquad P(T \leq 2) = 0.1117$$

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.

Aufgabe 6

(**a**) Wir müssen $P[X \leq 1]$ berechnen im Falle von $\pi = 0.075$.

```
> pbinom(1,50,0.075)

[1] 0.1025006
```

Wenn die Lieferung also nur 7.5% defekte Gläser enthält, so können wir dies mit unserem Test (mit 50 Proben) nur in ca. 10% der Fälle nachweisen!

(**b**) Mit **pbinom(0:50, 150, 0.1)** sehen wir, dass der Verwerfungsbereich $K = \{T \leq 8\}$ ist. Wir erhalten

```
> pbinom(8,150,0.075)

[1] 0.2000952
```

Dank der grösseren Stichprobe ist auch die Macht grösser geworden.