

4.3.10 Verteilungshypothese und Likelihood-Wert

Seien x_1, x_2, \dots, x_n unabhängige Realisierungen einer Zufallsgröße X . Je nach Verteilungstyp-Annahme kann man *Formeln* für die Größe der „Likelihood“ angeben.

a) Die Exponentialverteilung (EV)

$$\max_{\lambda > 0} L(x_1, \dots, x_n; \lambda) = \max_{\lambda > 0} \prod_{k=1}^n \lambda e^{-\lambda x_k} = (\lambda e^{-\lambda \bar{x}})^n \Big|_{\lambda=1/\bar{x}} = \left(\frac{1}{e \bar{x}}\right)^n;$$

wegen $\frac{\partial}{\partial \lambda} (\lambda e^{-\lambda a}) = e^{-\lambda a} (1 - \lambda a)$ tritt das Maximum für $\lambda = 1/\bar{x}$ ein.

b) Die Normalverteilung (NV)

$$\max_{\mu \in \mathbb{R}, \sigma > 0} \prod_{k=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-(x_k - \mu)^2 / 2\sigma^2} = \left(\frac{1}{\sqrt{2\pi} s_*} e^{-\frac{1}{2s_*^2} s_*^2}\right)^n = \left(\frac{1}{\sqrt{2\pi e} s_*}\right)^n,$$

da das Maximum für $\mu = \bar{x}$, $\sigma^2 = s_*^2 := \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2$ angenommen wird. Mit $\tau := \sigma^{-2}$, $\varphi(\mu) := \frac{1}{n} \sum (x_k - \mu)^2$ und $\psi(\mu, \tau) := \sqrt{\tau} e^{-\frac{\tau}{2} \varphi(\mu)}$ gilt $\psi_\mu = \psi \cdot \tau \sum (x_k - \mu)$, $\psi_\tau = \frac{\psi}{2\tau} (1 - \tau \varphi(\mu))$; also tritt das Maximum für $\mu = \bar{x}$, $\tau = 1/\varphi(\bar{x})$ ein. Die übliche *empirische* Varianz s^2 hat den Vorfaktor $\frac{1}{n-1}$ statt $\frac{1}{n}$; daher s_*^2 statt s^2 .

c) Die Gleichverteilung (GV)

Sei $x_* := \min x_k$, $x^* := \max x_k$. Dann gilt

$$\max_{a \leq x_*, b \geq x^*} L(x_1, \dots, x_n; a, b) = \left(\frac{1}{x^* - x_*}\right)^n.$$

Für eine *sehr große* gleichmäßig auf das Intervall $[x_*, x^*]$ verteilte Stichprobe x_1, \dots, x_n gilt $\bar{x} \approx \frac{x_* + x^*}{2} = x_* + \frac{x^* - x_*}{2}$ und $x^* - x_* \approx 2\sqrt{3} s_*$. Wegen $e/2 > 1$ und $\sqrt{2\pi e} > 2\sqrt{3}$ ist in diesem Fall die GV-Likelihood immer die bei weitem *größte* unter den dreien. Aber ändern wir bei festen x_* und x^* die Anordnung der x_k , bleibt die GV-Likelihood immer *dieselbe*, während die beiden anderen *variieren*. (Bei EV ist $x_* \geq 0$ vorauszusetzen.)

Etwa für $x_1 = \dots = x_{n-1} = x_*$, $x_n = x^*$ gilt $\bar{x} = x_* + \frac{x^* - x_*}{n}$ und $x^* - x_* = \sqrt{n+1 + \frac{1}{n-1}} s_*$. Die EV-Likelihood ist für $n \geq 3$ je nach $x_* \geq 0$ größer oder kleiner als die GV-Likelihood. Die NV-Likelihood ist für $n < 17$ kleiner als die GV-Likelihood, ansonsten größer.

Für $x_1 = x_*$, $x_2 = \dots = x_{n-1} = \frac{x_* + x^*}{2}$, $x_n = x^*$ gilt $\bar{x} = x_* + \frac{x^* - x_*}{2}$ und $x^* - x_* = \sqrt{2n} s_*$. In diesem Fall ist die EV-Likelihood stets kleiner als die GV-Likelihood, während die NV-Likelihood für $n \leq 8$ kleiner ist als die GV-Likelihood und ansonsten größer.

Tendenziell zeigen diese Beispiele, dass die GV-Likelihood zu oft auch bei ganz ungleichmäßigen Stichproben die größte ist. Zufallszahlen-Testbeispiele:

Ein mit **R** erzeugter Vektor $[x_1, \dots, x_{25}]$ unabhängiger $N(0;1)$ -Zufallszahlen hatte die Kenngrößen $-2.162363 \leq x_i \leq 2.268689$, $s_* = 1.097877$, also die GV- und NV-Likelihood 0.22568^{25} bzw. 0.2203987^{25} . Ausführliche Tests zeigten: Bei 50 $N(0;1)$ -Zahlen ist meistens, aber *nicht immer*, die NV-Likelihood größer; bei 25 Zahlen wechselt es ab, und bei 10 Zahlen ist fast immer die GV-Likelihood größer. Die **R**-Befehle (für 25 Zahlen):
`test_i<-rnorm(25) s1_i<-sqrt((24/25)*var(test_i)) nvbas_i<-1/(sqrt(2*pi*exp(1))*s1_i)`
`gvbas_i<-1/(max(test_i)-min(test_i))`

Fazit: Likelihood-Werte eignen sich grundsätzlich *nicht* als Kriterium dafür, ob das eine oder das andere Verteilungs-Modell besser zu einer Stichprobe passt. Im Falle *stetiger* Verteilungen ist der Kolmogorov/Smirnov-Test zu empfehlen, und allgemein kann man den Chiquadrat-Anpassungstest anwenden, der allerdings nur approximativ und bei kleinem n etwas ungenau ist.