A35 Konfidenzintervalle Kurzfragen

Freitag, 19. November 2021 14:30

David Venker
Jan Jäkel
Nico Guth

Aufgabe 35 *Konfidenzintervalle Kurzfragen*

- Ein Konfidenzintervall $[x_1, x_2]$ eines Parameters x zu einem Konfidenzniveau α sei gegeben. Was ist die frequentistische, was die bayesische Interpretation dieses Intervalls?

- frequentistische Interpretation:

  Im relativen Anteil $\alpha = \frac{k}{n}$ von n durchgeführten Experimenten wird das (pro Experiment) konstruierte Konfidenzintervall $[x_1, x_2]$ den wahren Wert x_{wahr} enthalten.

- bayesische Interpretation: (Kredibilitätsintervall)

  Der wahre Wert liegt mit der Wahrscheinlichkeit α im Intervall $[x_1, x_2]$.

- Welche Bedeutung hat der Prior in der bayesischen Statistik?

Der Prior beschreibt die Information über ein Problem, die schon vor der Analyse bekannt ist.

- Welche Freiheiten gibt es bei der Wahl des Intervalls?

Das Konfidenzintervall $[x_1, x_2]$ kann beliebig verschoben und gestreckt/gestaucht werden solange $\int_{x_1}^{x_2} P(x)dx = \alpha$ gilt.

solange $\int_{x_1} P(x)dx = \alpha$ gilt.

drei Beispiele (am häufigsten verwendet):

- symmetrisch um den Erwartungswert
- kürzestes Intervall
- Zentrales Intervall: die (frequentistische) Wahrscheinlichkeit dass $x_{wahr} > x_2$ soll gleich der für $x_{wahr} < x_1$ sein.

$$\Leftrightarrow \int_{-\infty}^{x_1} P(x)\,dx = \int_{x_2}^{\infty} P(x)\,dx \qquad \left(= \frac{1-\alpha}{2}\right)$$

- Was passiert im Spezialfall symmetrischer PDF?

Die genannten möglichen Intervalle sind identisch wenn die PDF symmetrisch ist.

- Was ist der Unterschied von Intervallen und Upper/Lower Limits?

- Intervalle geben die "Wkeit" für einen Bereich: $x_1 \leq x_{wahr} \leq x_2$
- Upper Limits geben die "W'keit" für eine obere Grenze: $-\infty \leq x_{wahr} \leq x_0$
- Lower Limits geben die "W'keit" für eine untere Grenze: $x_0 \leq x_{wahr} \leq \infty$

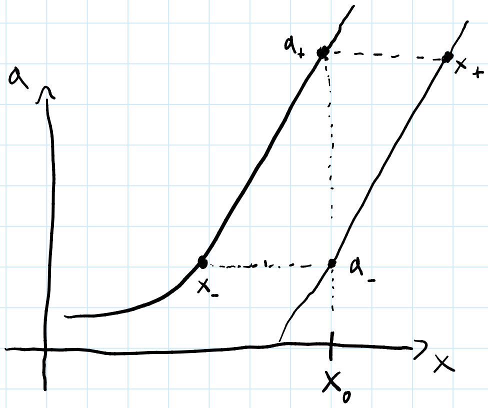
Aufgabe 36 *Konfidenzintervalle*

Die Likelihoodfunktion für einen Messwert x bei einem gegebenen Parameter a sei

$$L(X;a) = \frac{1}{\pi}\,\frac{1}{1+(x-a)^2} \quad \text{mit} \quad a > 0\,. \tag{1}$$

(a) Bestimmen Sie mit Hilfe der Neyman-Konstruktion das zentrale frequentistische 90 % Konfidenzintervall für a, wenn ein Messwert $x = 10$ gemessen wurde.

Beispiel Konfidenzband



$x_0 \mathrel{\hat=}$ gegebener Messwert $\qquad P(x,a) \mathrel{\hat=}$ gegebene "Form" der PDF für x

Gesucht ist das α-Kofidenzintervall $[a_-, a_+]$ bei x_0.

Allerdings kann nur das Konfidenzintervall $[x_-, x_0]$ und $[x_0, x_+]$ berechnet werden.

$\Rightarrow$ Es muss gelten $\boxed{\int_{x_-}^{x_0} P(x,a_-)\,dx = \alpha \quad \text{und} \quad \int_{x_0}^{x_+} P(x,a_+)\,dx = \alpha}$.

Außerdem wird das Zentrale Konfidenzintervall gewählt.

Also muss auch gelten $\boxed{\int_{-\infty}^{x_-} P(x,a_-)\,dx = \int_{x_0}^{\infty} P(x,a_-)\,dx \quad \text{und} \quad \int_{-\infty}^{x_0} P(x,a_+)\,dx = \int_{x_+}^{\infty} P(x,a_+)\,dx}$.

$\Rightarrow$ 4 Gleichungen, 4 Unbekannte (x_-, x_+, a_-, a_+)

---

gegeben: $P(x,a) = \frac{1}{\pi}\left(1+(x-a)^2\right)^{-1} \qquad \alpha = 90\%\,, \quad x_0 = 10$

unbestimmtes Integral:

$\int P(x,a)\,dx = \int \frac{1}{\pi}\left(1+(x-a)^2\right)^{-1} dx \qquad$ mit $\int (1+u)\,du = \arctan(u)$

subst. $u = x-a \Leftrightarrow dx = du$

$= \frac{1}{\pi}\arctan(x-a)$

subst. $u = x - a \iff dx = du$

$$= \frac{1}{\pi} \arctan(x-a)$$

nun die 4 Gleichungen

$$\text{I}: \int_{x_-}^{x_0} P(x,a_-)\,dx = \left[\frac{1}{\pi}\arctan(x-a_-)\right]_{x_-}^{x_0} \overset{!}{=} \alpha$$

$$\text{II}: \int_{x_0}^{x_+} P(x,a_+)\,dx = \left[\frac{1}{\pi}\arctan(x-a_+)\right]_{x_0}^{x_+} \overset{!}{=} \alpha$$

$$\text{III}: \int_{-\infty}^{x_-} P(x,a_-)\,dx = \left[\frac{1}{\pi}\arctan(x-a_-)\right]_{-\infty}^{x_-} \overset{!}{=} \int_{x_0}^{\infty} P(x,a_-)\,dx = \left[\frac{1}{\pi}\arctan(x-a_-)\right]_{x_0}^{\infty}$$

mit $\arctan(-\infty) = -\frac{\pi}{2}$, $\arctan(\infty) = \frac{\pi}{2}$

$$\text{IV}: \int_{-\infty}^{x_0} P(x,a_+)\,dx = \left[\frac{1}{\pi}\arctan(x-a_+)\right]_{-\infty}^{x_0} \overset{!}{=} \int_{x_+}^{\infty} P(x,a_+)\,dx = \left[\frac{1}{\pi}\arctan(x-a_+)\right]_{x_+}^{\infty}$$

sind 2 unabhängige Gleichungssysteme

1) I und III für x_- und a_-

2) II und IV für x_+ und a_+

1)

$$\text{I}: \arctan(x_0 - a_-) - \arctan(x_- - a_-) = \alpha \cdot \pi$$

$$\text{III}: \arctan(x_- - a_-) - \left(-\frac{\pi}{2}\right) = \left(\frac{\pi}{2}\right) - \arctan(x_0 - a_-)$$

$$\text{III} \iff \arctan(x_- - a_-) = -\arctan(x_0 - a_-)$$ mit $\tan(-\arctan(x)) = -x$ wegen Punktsymmetrie um $x=0$

$$\iff x_- - a_- = -(x_0 - a_-)$$

$$\iff x_- = 2a_- - x_0$$

x_- in I einsetzen:

$$\arctan(x_0 - a_-) - \arctan(2a_- - x_0 - a_-) = \alpha\pi$$

$$\iff \arctan(x_0 - a_-) - \arctan(-(x_0 - a_-)) = \alpha\pi$$

$$\iff 2\arctan(x_0 - a_-) = \alpha\pi$$

$$\iff \boxed{a_- = x_0 - \tan\left(\alpha\frac{\pi}{2}\right)}$$ Werte Einsetzen: $a_- \approx 3{,}6862$

2)

$\text{II}: \arctan(x_+ - a_+) - \arctan(x_0 - a_+) = \alpha\pi$

$\text{IV}: \arctan(x_0 - a_+) - (-\frac{\pi}{2}) = \frac{\pi}{2} - \arctan(x_+ - a_+)$

$\text{IV} \Leftrightarrow \arctan(x_0 - a_+) = -\arctan(x_+ - a_+)$

$\Leftrightarrow x_+ = a_+ - (x_0 - a_+)$

$\Leftrightarrow x_+ = 2a_+ - x_0$

x_+ in II einsetzen:

$\arctan((2a_+ - x_0) - a_+) - \arctan(x_0 - a_+) = \alpha\pi$

$\Leftrightarrow \quad 2\arctan(a_+ - x_0) = \alpha\pi$

$\Leftrightarrow \quad \boxed{a_+ = x_0 + \tan(\alpha\frac{\pi}{2})} \quad \Rightarrow \quad a_+ \approx 16,3138$

$\Rightarrow [a_-, a_+] = [3.69, 16.31]$ ist das 90% Konfidenzintervall für $x = 10$

(b) Bestimmen Sie unter Annahme einer gleichverteilten Priorverteilung in a das zentrale bayesianische Konfidenzintervall. (Beide Seiten außerhalb des zentralen Konfidenzintervalls haben den gleichen Wahrscheinlichkeitsinhalt.)

Satz von Bayes: $P(a|x) = \dfrac{P(x|a) \cdot P(a)}{\int_{-\infty}^{\infty} P(x|a)\, P(a)\, da}$

gegebene Likelihood: $P(x|a) = \frac{1}{\pi}\left(1 + (x-a)^2\right)^{-1} \quad a > 0$

$x_0 = 10 \qquad \alpha = 90\%$

Prior: $P(a) = \begin{cases} \frac{1}{a_{max}} & \text{für } 0 < a \leq a_{max} \\ 0 & \text{sonst} \end{cases}$

Normierung:

$$\int_{-\infty}^{\infty} P(x|a)\, P(a)\, da = \int_0^{a_{max}} \frac{1}{\pi}(1+(x-a)^2)^{-1} \frac{1}{a_{max}}\, da$$

$$= \frac{1}{a_{max}\pi}\left[-\arctan(x-a)\right]_0^{a_{max}}$$

$$= \frac{1}{a_{max}\pi}\left(\arctan(x) - \arctan(x - a_{max})\right)$$

$$= \frac{1}{a_{max}\pi}\left(\arctan(x) - \arctan(x - a_{max})\right)$$

$$\Rightarrow P(a|x) = \frac{P(x|a)\cdot P(a)}{\int_{-\infty}^{\infty} P(x|a)\, P(a)\, da}$$

$$= \left(1 + (x-a)^2\right)^{-1}\left(\arctan(x) - \arctan(x - a_{max})\right)^{-1}$$

für das α- Kredibilitätsinterval $[a_-, a_+]$ muss gelten:

$$\mathrm{I}: \int_{a_-}^{a_+} P(a|x)\, da \overset{!}{=} \alpha$$

für das Zentrale Intervall muss gelten:

$$\mathrm{II}: \int_{0}^{a_-} P(a|x)\, da \overset{!}{=} \int_{a_+}^{a_{max}} P(a|x)\, da$$

$$\mathrm{I}: \left[\arctan(x-a)\right]_{a_-}^{a_+} \cdot \left(\arctan(x) - \arctan(x - a_{max})\right)^{-1} = \alpha$$

$$\Leftrightarrow \; -\arctan(x - a_+) + \arctan(x - a_-) = \alpha\left(\arctan(x) - \arctan(x - a_{max})\right)$$

$$\mathrm{II}: \left[-\arctan(x-a)\right]_{0}^{a_-} = \left[-\arctan(x-a)\right]_{a_+}^{a_{max}}$$

$$\Leftrightarrow \; -\arctan(x - a_-) + \arctan(x) = -\arctan(x - a_{max}) + \arctan(x - a_+)$$

$$\Leftrightarrow \; \arctan(x - a_-) = \arctan(x) + \arctan(x - a_{max}) - \arctan(x - a_+)$$

in I einsetzen:

$$-\arctan(x - a_+) + \arctan(x) + \arctan(x - a_{max}) - \arctan(x - a_+) = \alpha\left(\arctan(x) - \arctan(x - a_{max})\right)$$

$$\Leftrightarrow \; -2\arctan(x - a_+) = (\alpha - 1)\arctan(x) - (\alpha + 1)\arctan(x - a_{max})$$

$$\Leftrightarrow \; \boxed{\begin{aligned} a_+ &= x + \tan\left(\frac{\alpha - 1}{2}\arctan(x) - \frac{\alpha + 1}{2}\arctan(x - a_{max})\right) \\ a_- &= x - \tan\left(\arctan(x) + \arctan(x - a_{max}) - \arctan(x - a_+)\right) \end{aligned}}$$

entscheide $a_{max} \to \infty$

$$\boxed{\begin{aligned} a_+ &= x + \tan\left(\frac{\alpha - 1}{2}\arctan(x) - \frac{\alpha + 1}{2}\left(-\frac{\pi}{2}\right)\right) \\ a_- &= x - \tan\left(\arctan(x) + \left(-\frac{\pi}{2}\right) - \arctan(x - a_+)\right) \end{aligned}}$$

Werte einsetzen:

$a_+ = 16{,}5240$

$a_- = 6{,}1123$

(c) Betrachten Sie den Unterschied zwischen beiden Methoden für $x \to \infty$.
Hinweis:

$$\int \frac{1}{1+x^2}\,\mathrm{d}x = \arctan x$$

a)

$a_+ = x + \tan\left(\alpha \frac{\pi}{2}\right)$

$a_- = x - \tan\left(\alpha \frac{\pi}{2}\right)$

$x \to \infty$

$a_+ \to \infty$

$a_- \to \infty$

b)

$a_+ = x + \tan\Big(\overbrace{\frac{\alpha-1}{2} \overbrace{\arctan(x)}^{\to \frac{\pi}{2}} - \frac{\alpha+1}{2}\left(-\frac{\pi}{2}\right)}^{\to \frac{\alpha\pi}{2}}\Big)$

$a_- = x - \tan\Big(\underbrace{\arctan(x) + \left(-\frac{\pi}{2}\right)}_{\to 0} - \underbrace{\arctan(x - a_+)}_{\to \frac{\alpha\pi}{2}}\Big)$

$x \to \infty$

$a_+ \to \infty$

$a_- \to \infty$

Im Unendlichen zeigen beide Methoden das gleiche Verhalten.