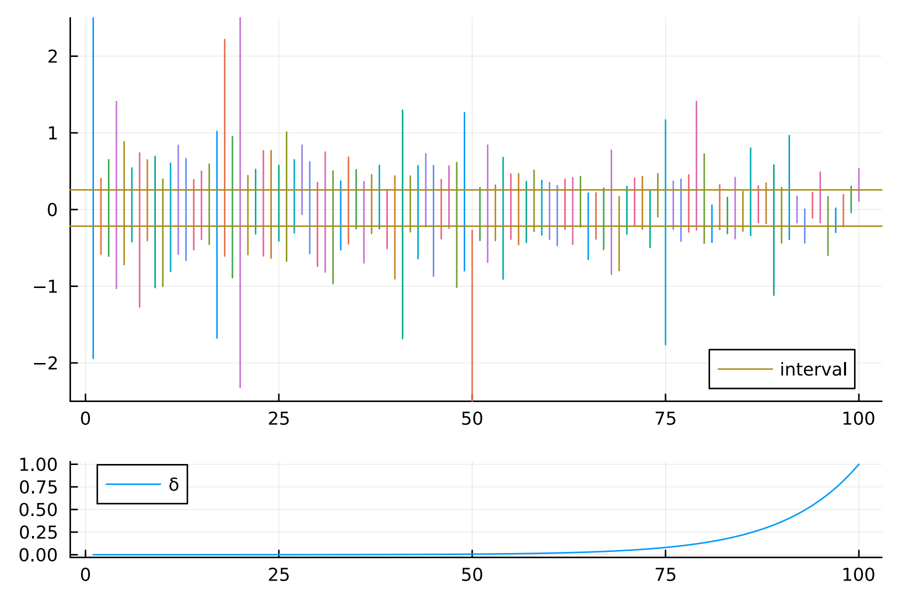
Combining Interval Data

Wouter M. Koolen
CWI and University of Twente

E-Day, CWI, November 8, 2024

Birds Eye



Confidence Interval Data

We are getting a sequence of **confidence regions** R_t (e.g. intervals) with associated confidences δ_t for a parameter μ of interest.

The guarantee is that these are valid, i.e. whenever $\mathbb{P}$ has parameter μ,

$$\mathbb{P}\{\mu \in R_t\} \;\geq\; 1 - \delta_t$$

for each t, conditioned on everything that happened before.

What are we supposed to infer from these regions? Our goal is to make a confidence region (fixed n) or confidence sequence (anytime) combining all data at **fixed confidence** β.

Approach

Use testing – confidence interval duality, i.e. we test every candidate μ separately.

For the *true* parameter μ, the binary outcome

$$Z_t \ := \ \mathbf{1}_{\mu \notin R_t}$$

is $\mathrm{Ber}(\theta)$ for some $\theta \leq \delta_t$.

Approach

Use testing – confidence interval duality, i.e. we test every candidate μ separately.

For the *true* parameter μ, the binary outcome

$$Z_t \;:=\; \mathbf{1}_{\mu \notin R_t}$$

is Ber(θ) for some $\theta \leq \delta_t$.

O.t.o.h. if μ is false, $Z_t = 1$ **may** happen too often.

Let's get rich in that scenario and disqualify μ.

One-round GRO

Conditionally, the null is $\{\mathrm{Ber}(\theta) | \theta \leq \delta_t\}$.

When we believe Z_t is $\mathrm{Ber}(q)$ for alternative $q \geq \delta_t$, the GRO e-variable is the likelihood ratio

$$S(Z_t) \;:=\; \frac{P_q(Z_t)}{P_{\delta_t}(Z_t)}$$

and it has expected log-return (implied target)

$$\mathrm{KL}(q, \delta_t)$$

Case closed?

One-round GRO

Conditionally, the null is $\{\mathrm{Ber}(\theta) | \theta \leq \delta_t\}$.

When we believe Z_t is $\mathrm{Ber}(q)$ for alternative $q \geq \delta_t$, the GRO e-variable is the likelihood ratio

$$S(Z_t) \;:=\; \frac{P_q(Z_t)}{P_{\delta_t}(Z_t)}$$

and it has expected log-return (implied target)

$$\mathrm{KL}(q, \delta_t)$$

Case closed?

No! A fixed q may not stay feasible, as it needs to be $\geq \delta_t$ for all t.

Next idea: pick a fixed function $q : [0,1] \to [0,1]$ with $q(\delta) \geq \delta$ (inflationary) and use alternative $q_t = q(\delta_t)$ in round t.

Which function?

Which Function

We can fix, or try to learn, inflationary $q : [0,1] \to [0,1]$ from some given function class:

- Arbitrary
- Increasing
- Parametric, e.g.

Function class	alternative $q(\delta)$	$S(0)$	$S(1)$
Alternative q	q	$\frac{1-q}{1-\delta}$	$\frac{q}{\delta}$
Payoff $x \in [0,1]$ at 0	$1-(1-\delta)x$	x	$x + \frac{1-x}{\delta}$
Factor $\rho \geq 1$	$\rho\delta$	$\frac{1-\rho\delta}{1-\delta}$	ρ
Ratio $c \geq 1$	$\frac{c\delta}{1-\delta+c\delta}$	$\frac{1}{1-\delta+\delta c}$	$\frac{c}{1-\delta+\delta c}$
Offset ϵ	$\delta + \epsilon$	$1 - \frac{\epsilon}{1-\delta}$	$1 + \frac{\epsilon}{\delta}$
$\text{KL}(q,\delta) \approx \frac{1}{2} D_{\chi^2}(q,\delta) = \tau$	$\delta + \sqrt{2\tau\delta(1-\delta)}$	$1 - \sqrt{2\tau\frac{\delta}{1-\delta}}$	$1 + \sqrt{2\tau\frac{1-\delta}{\delta}}$

Some desiderata

We want to take special notice when $Z_t = 1$ but $\delta_t \approx 0$.

A standard log-betting score scaling with

$$\sum_{t=1}^{n} (Z_t - \delta_t)$$

does not do that.

We propose to look at

$$\sum_{t=1}^{n} (Z_t - \delta_t) \ln \frac{1}{\delta_t}$$

Result

Lemma (Main Lemma)

$(Z_t - \delta_t) \ln \frac{1}{\delta_t}$ is sub-Gamma. I.e. for every $\lambda \in [0, 1)$, the following is an e-variable

$$S_t \;=\; e^{\lambda (Z_t - \delta_t) \ln \frac{1}{\delta_t} - \frac{-\ln(1-\lambda)-\lambda}{2}}.$$

We can multiply over rounds, mix/tune λ, e.g. with knowledge of β and n to find

Corollary

For every $\beta \in [0, 1]$

$$\mathbb{P}\left(\sum_{t=1}^{n} (Z_t - \delta_t) \ln \frac{1}{\delta_t} \;\geq\; \sqrt{n \ln \frac{1}{\beta}} + \frac{2}{3} \ln \frac{1}{\beta}\right) \leq \beta.$$

Conclusion

We found some way to test a Forecaster making one-sided Bernoulli claims.

Why is the above evalue a good/natural idea?

Let's talk!

References i