Conformal Prediction as Bayesian Quadrature

Jake C. Snell Thomas L. Griffiths
Princeton University

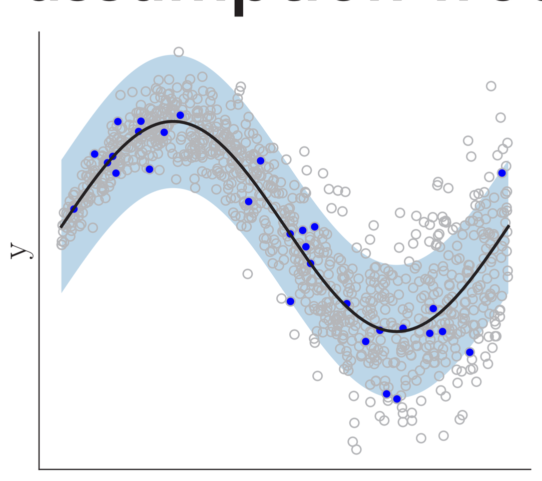
Contributions

1. We **reformulate** conformal prediction (CP) as prior-agnostic **Bayesian quadrature**.
2. We show how to **recover conformal prediction** from the **posterior mean**.
3. Our **distributional view** controls risk better with just a **small modification** to vanilla CP.

We revisit the foundations of conformal prediction and discover exciting new connections to Bayesian quadrature.

Conformal Prediction

Conformal prediction (Vovk et al., 2005) offers **assumption-free coverage** guarantees.



$$\Pr(Y_{n+1} \notin \mathcal{C}_{\lambda_{\mathrm{cp}}}(X_{n+1})) \leq \alpha$$

$$\mathcal{C}_\lambda(X_{n+1}) \triangleq \{y : s(X_{n+1}, y) \leq \lambda\}$$

$$\lambda_{\mathrm{cp}} \triangleq s_{(\lceil (n+1)(1-\alpha) \rceil)}$$

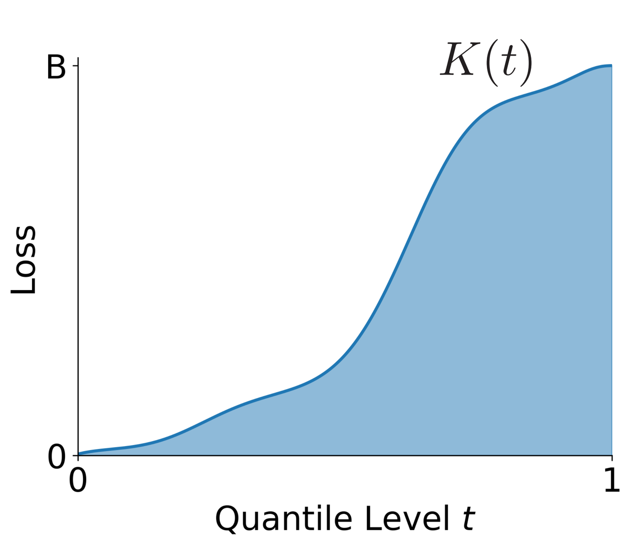
Conformal risk control (Angelopoulos et al., 2024) extends this to **more general losses**.

$$E(\ell(\mathcal{C}_{\lambda_{\mathrm{crc}}}(X_{n+1}), Y_{n+1})) \leq \alpha$$

$$\lambda_{\mathrm{crc}} \triangleq \inf\left\{\lambda : \frac{1}{n+1}\sum_{i=1}^{n} \ell_i(\lambda) + \frac{B}{n+1} \leq \alpha\right\}$$

Quantile Functions

We consider **probabilistic inference** over the **quantile function** of the loss distribution.



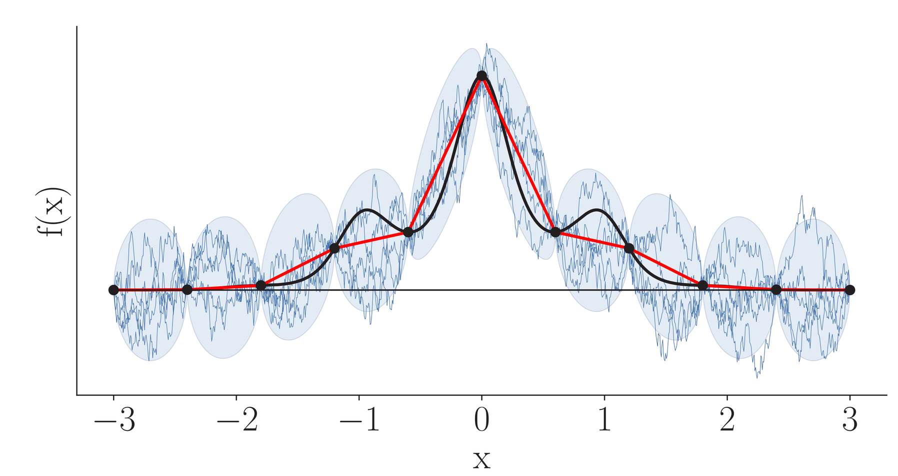
Risk (expected loss) is simply the **integral** of the quantile function:

$$L = \int_0^1 K(t)\, dt$$

Bayesian Quadrature

Bayesian quadrature (Diaconis, 1988) **estimates integrals** that have **uncertainty**.

1. Prior
2. Observe
3. Posterior
4. Integrate



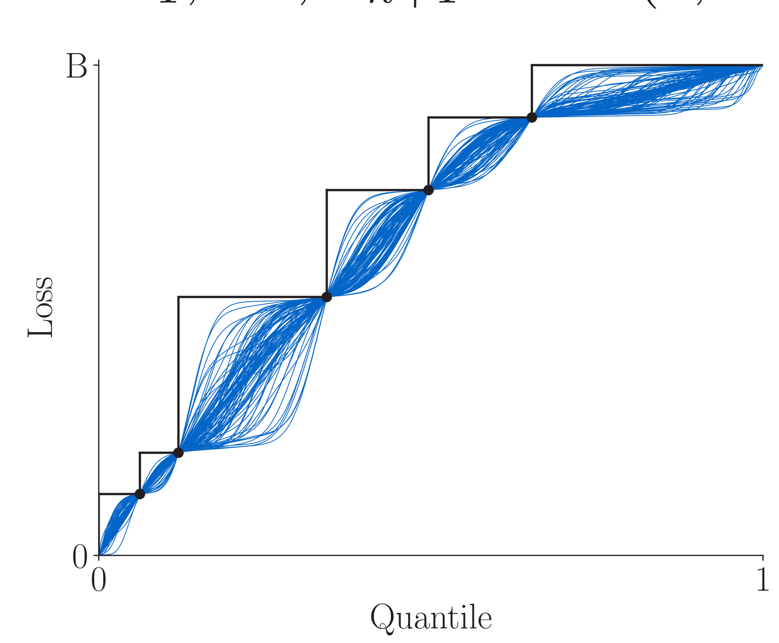
GP **posterior mean** recovers **trapezoid rule**.

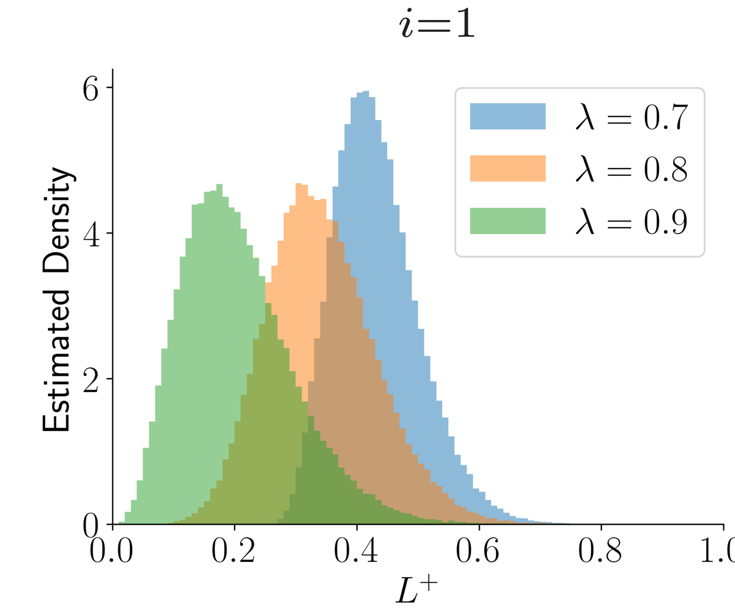
Our Approach

Our approach is based on two key insights:

1. Rectangular rule **upper bounds any prior**.
2. Widths distributed as **uniform Dirichlet**.

$$U_1, \ldots, U_{n+1} \sim \mathrm{Dir}(1, \ldots, 1) \qquad L^+ \triangleq \sum_{i=1}^{n+1} U_i \ell_{(i)}$$





Posterior mean recovers **CP/CRC selection**!

Results

Provides **greater control** over risk: **full distribution** instead of point estimate only.

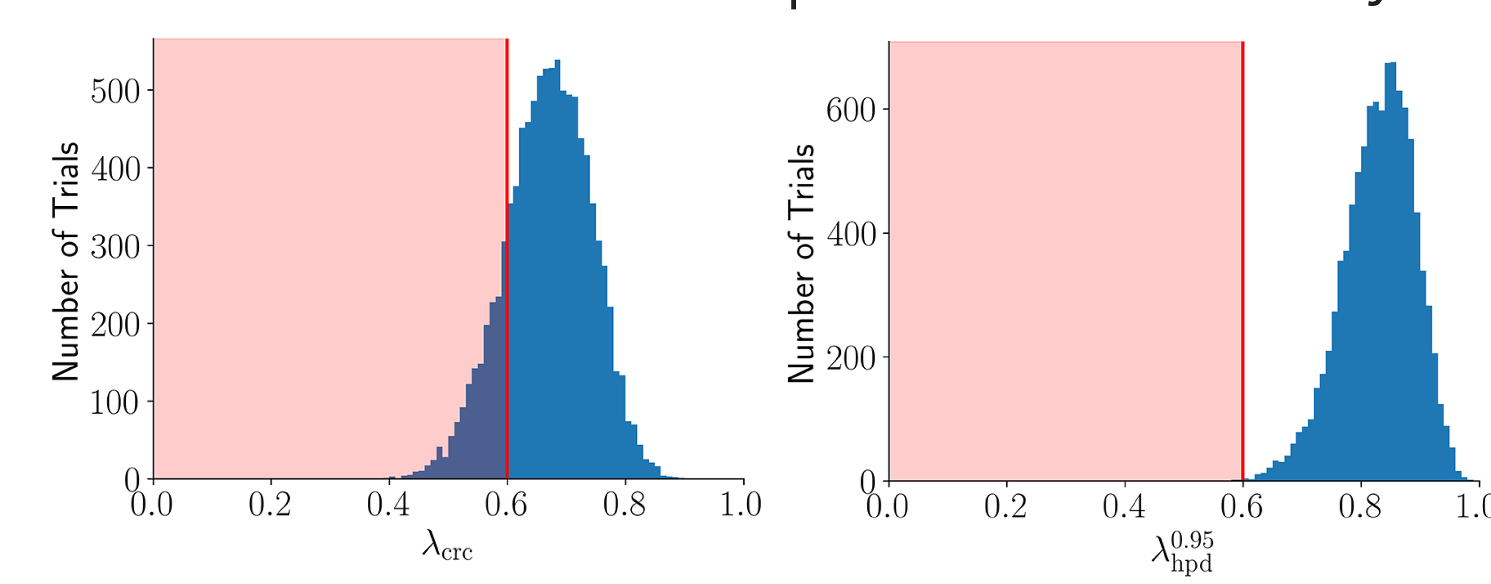


Figure. Comparison of actual risk incurred on synthetic binomial data.

Table 1. Risk control for synthetic heteroskedastic data.

Method	Failure Rate	95% C.I.	Pred. Int. Len.
Conformal Prediction	46.2%	[45.2%, 47.2%]	8.0
RCPS	0.0%	[0.00%, 0.04%]	14.3
Ours ($\beta = 0.95$)	3.4%	[3.1%, 3.8%]	9.5

Table 2. MS-COCO false negative rate control.

Method	Failure Rate	95% C.I.	Pred. Set Size
Conformal Risk Control	45.0%	[44.1%, 46.0%]	2.9
RCPS	0.0%	[0.00%, 0.04%]	3.6
Ours ($\beta = 0.95$)	5.4%	[5.0%, 5.9%]	3.0

References

Angelopoulos, A. N., Bates, S., Fisch, A., Lei, L., & Schuster, T. (2024). Conformal risk control. *The Twelfth International Conference on Learning Representations*.

Diaconis, P. (1988). Bayesian numerical analysis. *Statistical Decision Theory and Related Topics IV*, volume 1, pp. 163-175.

Vovk, V., Gammerman, A., and Shafer, G. (2005). *Algorithmic Learning in a Random World*.

Acknowledgments

This work was supported by the Office of Naval Research grant N00014-23-1-2510.