- $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$

  family of probability distributions

  θ - index or parameter

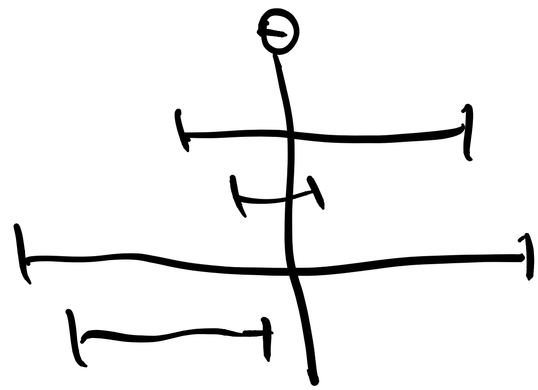
- data $X_i \sim P_\theta$ (θ unknown)

  i.i.d. $i = 1, \dots, n$

- Confidence intervals (frequentist)

  $P_\theta \to X \quad X = (X_1, \dots, X_n)$

  $\to$ Conf interval for θ



"covers θ"

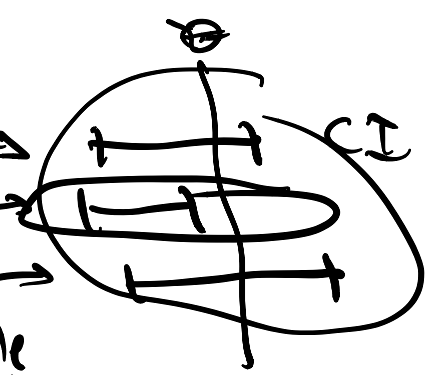
thought experiment

$P_\theta \to X \to$

$P_\theta \to X \to$

$\vdots$

$P_\theta \to X \to$



multiple datasets

- "ask" - 95% of the CI's

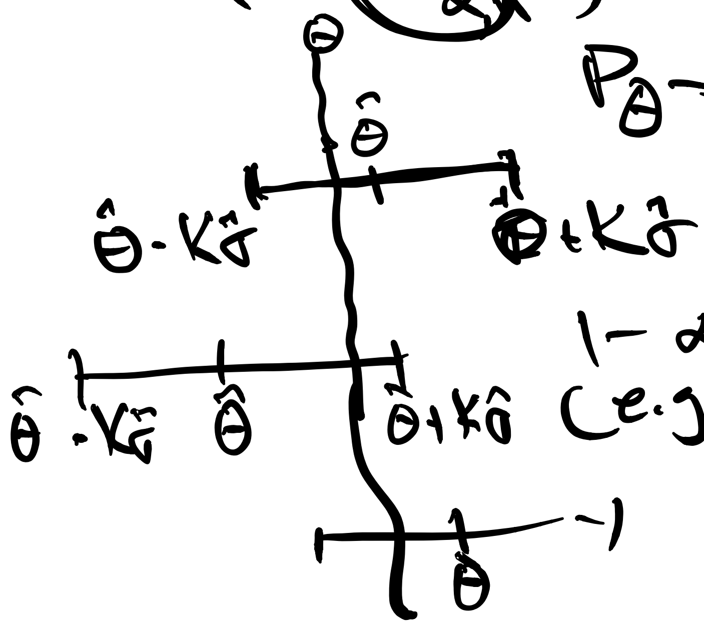
cover the unknown θ

- meaning of "confidence interval"
- true for any θ

- how do we compute confidence intervals?
  - typical recipe
    a) define an estimator: $\hat{\theta}$
    b) estimate a standard deviation of X_i: $\hat{\sigma}$
    c) CI:

$$(\hat{\theta} - K_{\alpha,n}\hat{\sigma},\ \hat{\theta} + K_{\alpha,n}\hat{\sigma})$$

$P_\theta \rightarrow X \rightarrow \hat{\theta}$
$\qquad\ \ \searrow \hat{\sigma}$



$1-\alpha$ coverage (e.g. $\alpha = .05$)

- bootstrap estimates

$$P_\theta \to X \to \hat{P}$$

thought experiment becomes real

histogram

$$\hat{P} \to X \to CI$$
$$\hat{P} \to X \to CI$$
$$\vdots$$
$$\hat{P} \to X \to CI$$

- bootstrap can be slow, and is not informative a priori

- alternative I

$$\left(\hat{\theta} - \frac{1.96}{\sqrt{n}}\hat{\sigma},\ \hat{\theta} + \frac{1.96}{\sqrt{n}}\hat{\sigma}\right)$$

$$K_{\alpha,n} = \frac{1.96}{\sqrt{n}}$$

- comes from the CLT for $\hat{\theta}$
- "asymptotic" "asymptopia"
  $\hat{\theta}$ becomes Gaussian eventually

- alternative II

Concentration inequalities

non-asymptotic

- re-express CI's:

$$P\Big(\underbrace{\theta}_{\substack{\text{fixed}\\\text{constant}\\\text{(not random)}}} \in \underbrace{(\hat\theta - K_{\alpha,n}\hat\sigma,\ \hat\theta + K_{\alpha,n}\hat\sigma)}_{\text{random interval}}\Big) \geq 1-\alpha$$

.95

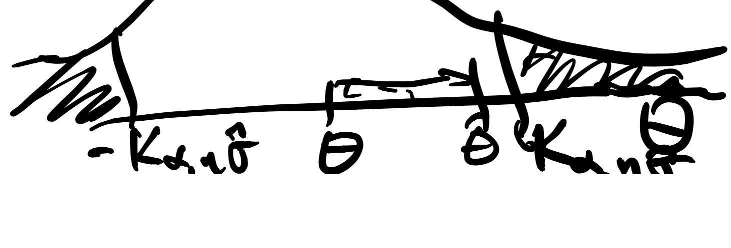
$$= P(\hat\theta - K_{\alpha,n}\hat\sigma \leq \theta \leq \hat\theta + K_{\alpha,n}\hat\sigma) \geq 1-\alpha$$

$$= P(|\hat\theta - \theta| \leq K_{\alpha,n}\hat\sigma) \geq 1-\alpha$$

$$\boxed{= P(|\hat\theta - \theta| \geq K_{\alpha,n}\hat\sigma) \leq \alpha}$$

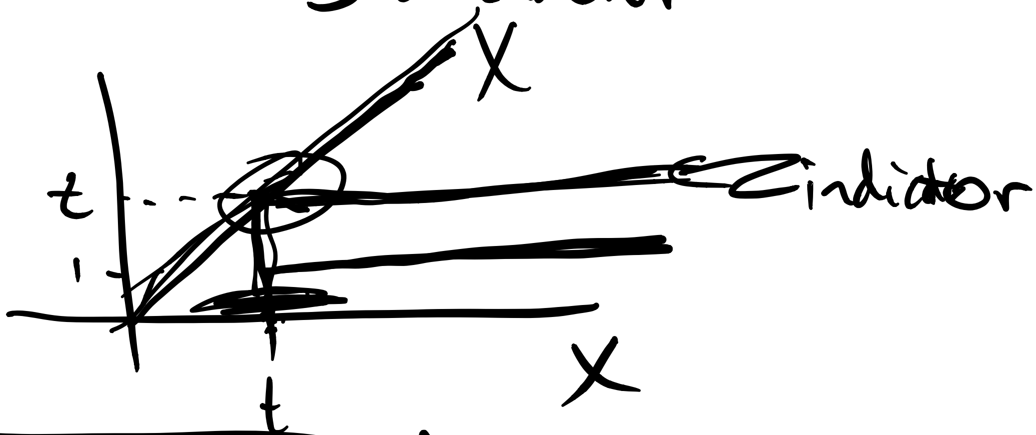
$$\Rightarrow \quad P(\hat\theta - \theta > K_{\alpha,n}\hat\sigma) \leq \frac{\alpha}{2}$$

$$P(\theta - \hat\theta < K_{\alpha,n}\hat\sigma) \leq \frac{\alpha}{2}$$



- assume $\hat{\sigma}$ is known
- $P(|\hat{\theta} - \theta| > t) \le \alpha$

study how to obtain this kind of probability statement



$$\boxed{t \cdot \mathbb{1}(X \ge t)}$$

$$X \ge t\,\mathbb{1}(X \ge t)$$

take Expectation on both sides

$$E(X) \ge t\,P(X \ge t)$$

$$P(X \ge t) \le \frac{E(X)}{t}$$

$$t$$

$$P(Z-\mu \geq t)$$

$$= P(\lambda(Z-\mu) \geq \lambda t) \quad \lambda > 0$$

$$= P(e^{\lambda(Z-\mu)} \geq e^{\lambda t})$$

Markov

$$\leq \frac{E\, e^{\lambda(Z-\mu)}}{e^{\lambda t}}$$

$$= \frac{mgf(Z-\mu)}{e^{\lambda t}}$$

$$P(Z-\mu \geq t) \leq e^{\frac{\sigma^2\lambda^2}{2} - \lambda t}$$

true for any λ

$$\lambda^* = \arg\min_\lambda \exp\left(\frac{\sigma^2\lambda^2}{2} - \lambda t\right)$$

$$\frac{d}{d\lambda} \exp\left(\frac{\sigma^2\lambda^2}{2} - \lambda t\right)$$

$$= \exp\left(\frac{\sigma^2\lambda^2}{2} - \lambda t\right) \cdot \left(\sigma^2\lambda - t\right)$$

Set $= 0$

$$\Rightarrow \quad \sigma^2\lambda - t = 0$$

$$\lambda = \frac{t}{\sigma^2} \qquad \lambda^2 = \frac{t^2}{\sigma^4}$$

$$P(Z - \mu \geq t) \leq \exp\left(\frac{\lambda^2\sigma^2}{2} - \lambda t\right)$$

$$= \exp\left(\frac{t^2 \cancel{\sigma^2}}{2\sigma^{\cancel{4}\,2}} - \frac{t^2}{\sigma^2}\right)$$

$$= \exp\left(\frac{t^2}{2\sigma^2} - \frac{t^2}{\sigma^2}\right)$$

$$= \exp\left(-\frac{1}{2}\frac{t^2}{\sigma^2}\right.$$

$\phi(\lambda)$ $E(X^2)$ $\quad$ $E(X^3)$

mgf $E\left(e^{\lambda X}\right) = E\left(1 + \lambda X + \frac{\lambda^2}{2}X^2 + \dots\right.$

$$= 1 + \lambda EX + \frac{\lambda^2}{2}EX^2 + \dots$$

[illegible]