

# Interval Estimation

## What is the Confidence Interval?

Prof Hans Georg Schaathun

Høgskolen i Ålesund

10th February 2014

# Point Estimation

- Experiment generates data  $D \in \mathcal{D}$
- Estimator of  $\theta$ :  $\hat{\theta} : \mathcal{D} \rightarrow \mathbb{R}$ 
  - $D$  is stochastic
  - hence  $\hat{\theta}(D)$  is a stochastic variable
- Estimation error  $|\theta - \hat{\theta}(D)|$ 
  - the estimation error is **probably** small
  - but it is stochastic and usually unbounded

# Interval estimation

- Experiment  $\rightarrow$  data  $D \in \mathcal{D}$
- Two 'estimators'
  - $\hat{\theta}_{\text{low}} : \mathcal{D} \rightarrow \mathbb{R}$  *Lower bound*
  - $\hat{\theta}_{\text{high}} : \mathcal{D} \rightarrow \mathbb{R}$  *Upper bound*
- Level of confidence  $\beta$  *Large, e.g. 95%, 98%, 99%*

## Definition

Given a pair of functions  $\hat{\theta}_{\text{low}}$  and  $\hat{\theta}_{\text{high}}$  from the observed data into the real numbers, and

$$P_D(\hat{\theta}_{\text{low}}(D) \leq \theta \leq \hat{\theta}_{\text{high}}(D)) \geq \beta.$$

Then  $(\hat{\theta}_{\text{low}}(D), \hat{\theta}_{\text{high}}(D))$  is a  $(100\beta)\%$  confidence interval.

# Motivation of confidence interval

- Two key quantities combined
  - approximate estimated value (interval)
  - **level of confidence**
- Opinion polls sometimes use only error margins
  - $\hat{\theta} \pm e$
  - We do **not know** the level of confidence

# Warning! Pitfall

## Confidence level versus probability

$$P_D(\hat{\theta}_{\text{low}}(D) \leq \theta \leq \hat{\theta}_{\text{high}}(D)) \geq \beta$$

- The confidence level is *a priori* probability
  - that the confidence interval will enclose the parameter  $\theta$
- **It is not**
  - the probability that  $\theta$  is within the interval
  - because  $\theta$  is not a stochastic variable

$$P_{\theta}(\hat{\theta}_{\text{low}}(D) \leq \theta \leq \hat{\theta}_{\text{high}}(D)) \geq \beta$$

# Summary

$$P_D(\hat{\theta}_{\text{low}}(D) \leq \theta \leq \hat{\theta}_{\text{high}}(D)) \geq \beta$$

*In the next couple of videos, **Confidence Intervals for particular quantities** (mean, binomial proportion, etc.)*