Confidence Interval

Interval Estimation

Prof Hans Georg Schaathun

Høgskolen i Ålesund

9th January 2014

Point Estimation

- Experiment $\rightarrow$ data D
- Estimator of θ: $\hat{\theta} : D \rightarrow \mathbb{R}$
  - $\hat{\theta}(D)$ is a stochastic variable (because D is)
- Unbiased estimator: $E_D(\hat{\theta}(D)) = \theta$
- but $|\theta - \hat{\theta}(D)|$ may be large

Interval estimation

- Experiment $\rightarrow$ data D
- Two 'estimators'
  - $\hat{\theta}_{\text{low}} : D \rightarrow \mathbb{R}$
  - $\hat{\theta}_{\text{high}} : D \rightarrow \mathbb{R}$
- Bounded probability: $P_D(\hat{\theta}_{\text{low}} \leq \theta \leq \hat{\theta}_{\text{high}}) \geq \beta$
- Level of confidence β
  - should be large (95%, 98%, 99%)
- $(\hat{\theta}_{\text{low}}, \hat{\theta}_{\text{high}})$ 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
  - $\hat{\theta} \pm e$
  - interval without confidence level
  - unbounded stochastic variables $\Rightarrow$ the error may be larger

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 paramter θ
- It is not
  - the probability that θ is within the interval
  - because θ is not a stochastic variable

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

HØGSKOLEN I ÅLESUND
Ålesund University College

Summary

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

Next video: confidence interval for the binomial proportion

HØGSKOLEN I ÅLESUND
Ålesund University College