Estimating Binomial Proportions

The Confidence Interval

Prof Hans Georg Schaathun

Høgskolen i Ålesund

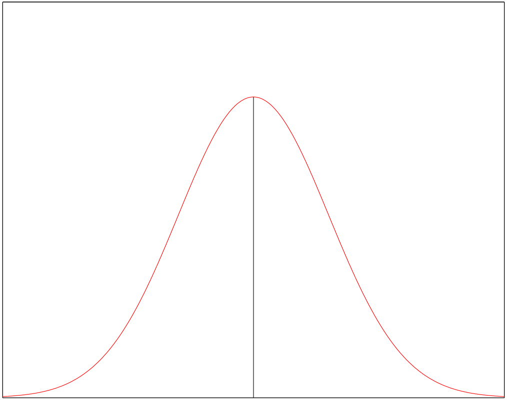
10th January 2014

Confidence Interval

- Point estimator: $\hat{p} = X/n$ where $X \sim B(n,p)$
  - $\hat{p} \sim N(p,\sigma)$
- Interval $(\hat{p}_{\text{low}}, \hat{p}_{\text{high}})$
  - Bounded probability: $P_D(\hat{p}_{\text{low}} \leq p \leq \hat{p}_{\text{high}}) \geq \beta$

HØGSKOLEN I ÅLESUND
Aalesund University College

The Point Estimator as a Start



What is $z_{\alpha/2}$?

- $z_{\alpha/2}$ solves $\alpha/2 = F(z)$
  - where F is the CDF of $N(0,1)$.
- I.e. we need to invert the CDF.
- Consider a 95% confidence interval
  - $\alpha = 0.05$
  - In matlab: `icdf( 'norm', 0.05/2, 0, 1 )`
  - you get: `ans = -1.9600`

What is σ?

Problem *We do not know the standard deviation σ.*

- Binomial distribution $B(n, p)$ has $\sigma^2 = \frac{p(1-p)}{n}$
- ... but we do not know p either ...

What is σ?

Problem *We do not know the standard deviation σ.*

- Binomial distribution $B(n, p)$ has $\sigma^2 = \frac{p(1-p)}{n}$
- ... but we do not know p either ...
- We can estimate p

$$\hat{\sigma} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

HØGSKOLEN I ÅLESUND
Aalesund University College

Confidence Interval for the Binomial Distribution

Summary

$$\hat{p} - z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- We assume
  1. n is sufficiently large
  2. p is not too close to 0 or 1
- Textbook: at least four hits (errors/successes)
- Communications: 100 errors

Exercise

Suppose you want to find out the percentage p of Norwegian students who think they have made a bad choice of degree programme. You poll 1000 students and 177 say they think their choice was bad. Give a 95% confidence interval for p.