Parametric Bootstrapping

18.05 Spring 2017

Parametric bootstrapping

Use the estimated parameter to estimate the variation of estimates of the parameter!

- Data: $x_1, \ldots, x_n$ drawn from a parametric distribution $F(\theta)$.
- Estimate θ by a statistic $\hat{\theta}$.
- **Generate many bootstrap samples from $F(\hat{\theta})$.**
- Compute the statistic θ^* for each bootstrap sample.
- Compute the bootstrap difference

$$\delta^* = \theta^* - \hat{\theta}.$$

- Use the quantiles of δ^* to approximate quantiles of

$$\delta = \hat{\theta} - \theta$$

- Set a confidence interval $[\hat{\theta} - \delta^*_{1-\alpha/2},\, \hat{\theta} - \delta^*_{\alpha/2}]$

Parametric sampling in R

```
# Data from binomial(15, θ) for an unknown θ
x = c(3, 5, 7, 9, 11, 13)
binomSize = 15      # known size of binomial
n = length(x)       # sample size
thetahat = mean(x)/binomSize     # MLE for θ
nboot = 5000        # number of bootstrap samples to use

# nboot parametric samples of size n; organize in a matrix
tmpdata = rbinom(n*nboot, binomSize, thetahat)
bootstrapsample = matrix(tmpdata, nrow=n, ncol=nboot)

# Compute bootstrap means thetahat* and differences delta*
thetahatstar = colMeans(bootstrapsample)/binomSize
deltastar = thetahatstar - thetahat

# Find quantiles and make the bootstrap confidence interval
d = quantile(deltastar, c(.1,.9))
ci = thetahat - c(d[2], d[1])
```

Board question

Data: 6 5 5 5 7 4 $\sim$ binomial(8,θ)

1. Estimate θ.

2. Write out the R code to generate data of 100 parametric bootstrap samples and compute an 80% confidence interval for θ.

(Try this without looking at your notes. We'll show the previous slide at the end)

Preview of linear regression

- Fit lines or polynomials to bivariate data
- Model: $y = f(x) + E$
  $f(x)$ function, E random error.
- Example: $y = ax + b + E$
- Example: $y = ax^2 + bx + c + E$
- Example: $y = \mathrm{e}^{ax+b+E}$ (Compute with $\ln(y) = ax + b + E$.)